the plant journal

SEB Society for Experimental Biology

The Plant Journal (2016) **87,** 103–117 doi: 10.1111/tpj.13176

SI SYNTHETIC BIOLOGY

Biotechnology and synthetic biology approaches for metabolic engineering of bioenergy crops

Patrick M. Shih[1,2,†], Yan Liang[1,2,†] and Dominique Loqué[1,2,3,*]

[1]*Joint BioEnergy Institute, Emery Station East, 5885 Hollis St, 4th Floor, Emeryville, CA 94608, USA,*

[2]*Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA, and*

[3]*Université Lyon 1, INSA de Lyon, CNRS, UMR5240, Microbiologie, Adaptation et Pathogénie, 10 rue Raphaël Dubois, F-69622 Villeurbanne, France*

Received 19 December 2015; revised 18 March 2016; accepted 22 March 2016; published online 30 March 2016.

*For correspondence (e-mail dloque@lbl.gov)

†These authors contributed equally.

SUMMARY

The Green Revolution has fuelled an exponential growth in human population since the mid-20th century. Due to population growth, food and energy demands will soon surpass supply capabilities. To overcome these impending problems, significant improvements in genetic engineering will be needed to complement breeding efforts in order to accelerate the improvement of agronomical traits. The new field of plant synthetic biology has emerged in recent years and is expected to support rapid, precise, and robust engineering of plants. In this review, we present recent advances made in the field of plant synthetic biology, specifically in genome editing, transgene expression regulation, and bioenergy crop engineering, with a focus on traits related to lignocellulose, oil, and soluble sugars. Ultimately, progress and innovation in these fields may facilitate the development of beneficial traits in crop plants to meet society's bioenergy needs.

Keywords: synthetic biology, metabolic engineering, bioenergy, energy crops, genome editing, transgene expression control.

INTRODUCTION

Crop domestication is a prime example of how humans have modified the many species surrounding them for thousands of years to meet their needs. Humans have progressed from accidentally selecting for the most 'useful' flora, to actively growing designated crops in the form of agriculture, to the calculated breeding that is performed today. As these improvements evolved into modern agriculture, humans' ability to modify the traits of plants for societal needs has become more targeted. Although many of these traits have revolved around agricultural yield for the purpose of food production, modern society also requires the use of crops as a clean energy source in the form of biofuels.

One goal of synthetic biology is to support rapid, precise, and robust engineering of organisms for useful societal purposes. Civilization was built on the invention of agriculture, so it is of course natural to ask: how can plant biologists further improve upon crops that humans have already spent millennia selecting and breeding? Furthermore, how can scientists coax plants bred to meet traditional agricultural needs into serving new purposes in addressing modern-day issues (i.e., climate change, green technology, and bioenergy)? In general, decreasing the amount of inputs and maximizing product yield outputs are overly simplified but accurate goals of crop engineering.

Because the bulk of synthetic biology has focused on microbes, plant synthetic biology has inherent and specific challenges. For starters, a diversity of tools to robustly engineer plants is almost nonexistent, in contrast to the microbial engineering field. *E. coli* has been the predominant chassis of choice for the synthetic biology community, and for many obvious reasons, the tools and methodologies in plants have not progressed as far. Furthermore, each plant has different physiologies, metabolisms, lifecycles, and environmental niches, making it

difficult to expect that all synthetic biology tools can be broadly transferable between plants. Thus, one inherent goal of the plant synthetic biology community is to continue building robust tools that can be used to engineer a diversity of crop plants. Although synthetic biology has greatly increased the throughput of DNA assembly, the low transformation efficiency of plants is a major bottleneck that will be needed to advance the field, especially to develop high-throughput assays in non-model systems.

Great progress has been made in plant engineering and plant synthetic biology in the recent years, but there are still many gaps in these fields that need to be addressed before both fields become comparable to those used in microbial systems. With traditional food and agricultural traits (e.g., nutritional value, yield, biofortification, and biotic and abiotic stress) already improved upon by breeding and further being advanced by molecular and synthetic biology, bioenergy crops provide a relatively new concept to the old agricultural system. Crop plants that have customarily been grown for food may need to be drastically modified to address growing energy needs and the impending crisis of climate change. In this review, we highlight some recent progress in the fields of plant synthetic biology and bioenergy crop engineering, with an emphasis on genome editing, gene expression regulation, and key bioenergy traits, and discuss some of the benefits of merging both fields (Figures 1 and 2).

GENOME MANIPULATION WITH RNA-GUIDED NUCLEASES AS EDITING TOOLS

In contrast to random mutagenesis and the search for natural allelic variants in traditional breeding, genome editing is aimed at the targeted modification of genomes in a precise manner. There are two basic steps in plant genome editing. Firstly, various genome editing tools are used to induce DNA double-stranded breaks at the target site. Secondly, double-stranded breaks are repaired by endogenous DNA repair machinery. One type of DNA repair mechanism is non-homologous end joining (NHEJ). NHEJ is error prone, leading to random insertion or deletion at the repaired site and subsequent loss of function of the target gene (Gorbunova and Levy, 1997). Earlier genome editing tools represented by meganucleases (D'Halluin *et al.*, 2013), zinc-finger nucleases (ZFNs, Bibikova *et al.*, 2002; Weinthal *et al.*, 2010), and transcription activator-like effector nucleases (TALENs, Li *et al.*, 2011; Zhang *et al.*, 2013) utilize protein motifs to recognize target DNA. The DNA binding motif is fused to the nuclease domain of

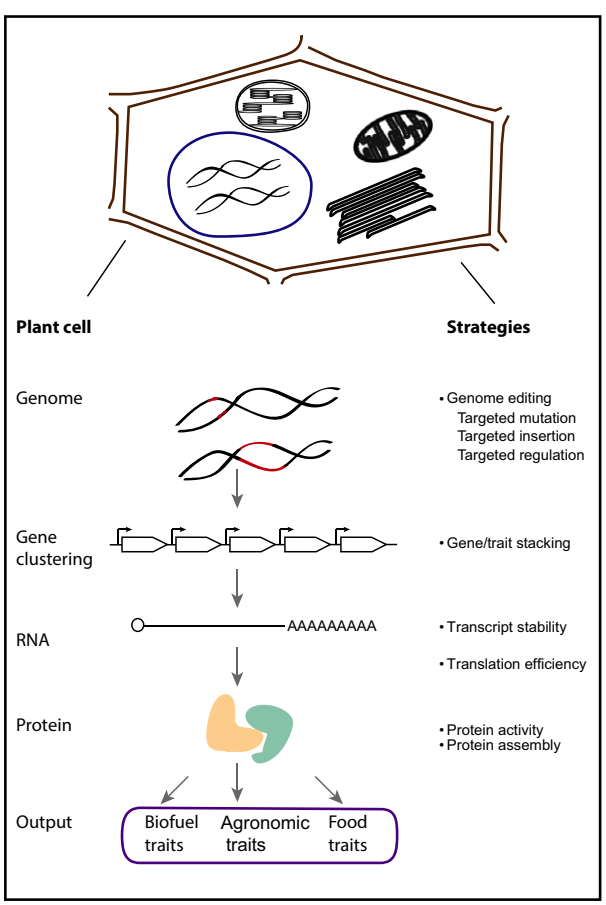

Figure 1. Summary of synthetic biology tools developed for the manipulation of DNA to proteins to optimize bioenergy traits in plants.

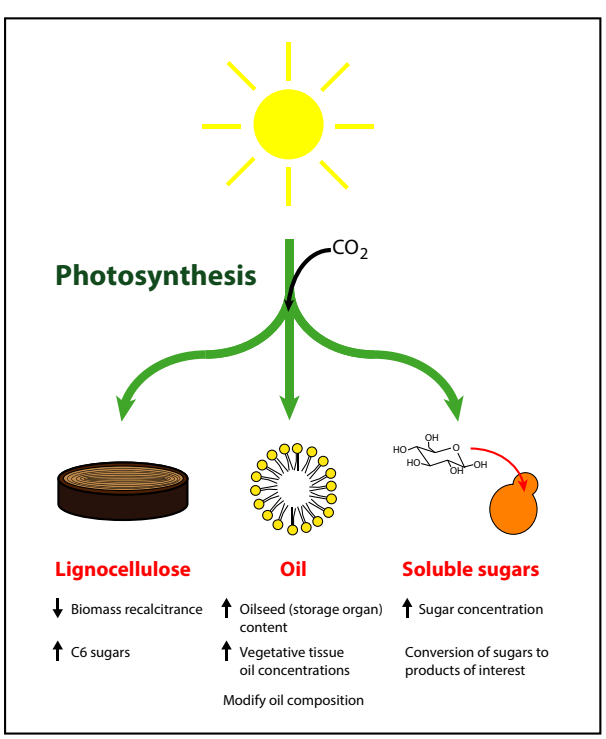

Figure 2. Application of synthetic biology to improve and engineer bioenergy traits.

*Fok*I, which cleaves DNA double strands upon dimerization. In contrast with mature technologies in genome editing – such as meganucleases, ZFNs and TALENs, which have been developed over a decade and are already deployed to modify commercial crops (examples of crop genome editing are highlighted in Table 1) – RNA-guided genome editing is still in its infancy with promising application potentials. This recently emerged CRISPR (clustered, regularly interspaced, short palindromic repeats)/CAS (CRISPR-associated) double-stranded DNA nuclease utilizes RNA-based DNA recognition (reviewed in Puchta, 2016). This CRISPR/CAS editing system is used with a synthetic-guide RNA (sgRNA), that contains an ~20 nucleotide (nt) sequence complementary to that of the target locus, allowing direction of the nuclease (CAS protein) to a precise location in the genome. The CRISPR/CAS9 system, a RNA-guided nuclease system, has been widely adopted for genome editing because of its efficacy and efficiency in various plant species. Within the last 2 years, genome editing with CRISPR/CAS has been demonstrated in over 13 plant species, including biofuel species *Zea mays* (Liang *et al.*, 2014), *Sorghum bicolor* (Jiang *et al.*, 2013), *Glycine max* (Jacobs *et al.*, 2015), and *Populus tomentosa* (Fan *et al.*, 2015). Several reviews (Baltes and Voytas, 2015; Bortesi and Fischer, 2015; Hsu *et al.*, 2014; Schaeffer and Nakata, 2015) have discussed the basic principle, technical details, and comparisons among different genome editing tools.

Outside of CAS9, Cpf1 was recently identified as another RNA-guided double-stranded DNA nuclease, offering new potential for genome editing since it generates 5′ staggered ends (Zetsche *et al.*, 2015). Various RNA-guided nucleases provide options for selecting proper tools to meet different needs for genome editing. For example, off-target effects in most cases need to be avoided but sometimes could be utilized to simultaneously target multiple homologous or paralogous genes (Endo *et al.*, 2015). Targeting efficiency and specificity are not only influenced by RNA-guided nucleases but are also determined by the target sequence, sgRNA structure, and genomic context (Bortesi and Fischer, 2015). Thus, these aspects need to be taken into consideration before making a choice among nucleases. For example, Cpf1 expands the genome targeting sequences to AT-rich regions and offers a complementary editing approach to the CAS9 system. With bioinformatics tools, it is possible that more RNA-guided endonucleases will be identified. With a library of various nucleases and their preferred target sites, users can chose the best fit for their target gene or site for modification.

In addition to advances in RNA-guided nucleases, progress has been made in increasing the efficiency in delivery and expression of genome-editing tools. Direct transformation of plant protoplast cells with preassembled complexes of purified CAS9 protein and guide RNA is intriguing because DNA-free genome editing avoids genomic insertion of encoding elements of double-stranded DNA nuclease complexes into target plants (Woo *et al.*, 2015). Protoplast-based genome editing might not be directly applicable for bioenergy crops, which cannot be easily regenerated from protoplasts. However, the same concept may be achieved given that several crop plants can be regenerated from tissue culture and are amenable to biolistic transformation, and that biolistic delivery of protein and nucleic acid into plant cells has been shown to be possible (Martin-Ortigosa *et al.*, 2012; Martin-Ortigosa and Wang, 2014).

Table 1 Synthetic biotechnology tools applied in plants

Synthetic tools	Representative examples
Genome editing	
Meganuclease	Targeted molecular trait stacking in cotton (D'Halluin *et al.*, 2013)
ZFN	Targeted insertion in Zea mays (Shukla *et al.*, 2009)
TALEN	Targeted mutagenesis in potato (Clasen *et al.*, 2016)
CRISPR/CAS	Targeted mutagenesis in wheat (Wang *et al.*, 2014)
Transgene regulation at the transcriptional level	
Synthetic promoter based on endogenous *cis*-elements	Nematode-responsive synthetic promoter (Liu *et al.*, 2014)
De novo synthetic promoter	Transcription activator-like effector-activated promoter (Bruckner *et al.*, 2015)
Orthogonal promoter	T7 polymerase-directed transcription (Nguyen *et al.*, 2004)
Synthetic *cis*-elements	dCAS9-based transcription activators and repressors (Lowder *et al.*, 2015; Piatek *et al.*, 2015)
Transgene regulation at the post-transcriptional level	
mRNA stability at the splicing step	OsL5-mediated exon skipping system (Hickey *et al.*, 2012; Gonzalez *et al.*, 2015)
mRNA stability regulated by RNase	CSY4-mediated transcript repression (Qi *et al.*, 2012; Borchardt *et al.*, 2015). Not tested in plants
Regulation at the translational level	Thiamine pyrophosphate riboswitch (Ramundo *et al.*, 2013)
Regulation at the post-translational level	2A peptide for coordinated protein expression (Halpin *et al.*, 1999)
	Thermoregulation of xylanase via self-splicing bacterial inteins (Shen *et al.*, 2012)

Use of RNA-guided nucleases for complex genome editing

The simple structure of the gRNA not only facilitates the design of a single gRNA but also enables the construction and delivery of multiple functional sgRNAs simultaneously into plant cells. Genome editing with a single CAS9 protein and multiple sgRNAs has been shown to have many potential applications, including increasing the rate of mutagenesis in a single target gene (Li *et al.*, 2013; Ma *et al.*, 2015), targeting multiple genes within a gene family (Mao *et al.*, 2013; Xing *et al.*, 2014; Zhou *et al.*, 2014; Ma *et al.*, 2015), and targeting of multiple genes in a pathway simultaneously (Ma *et al.*, 2015). With sgRNAs targeting two sites on the same chromosome, deletion of a gene fragment or gene clusters was also achieved (Mao *et al.*, 2013; Zhou *et al.*, 2014). Such multiplexed genome editing has been demonstrated in transient expression systems as well as Arabidopsis and rice stable transgenic lines.

HOMOLOGOUS RECOMBINATION-DEPENDENT GENE TARGETING (HRGT)

In addition to NHEJ, homologous recombination can be triggered at the site of DNA double-stranded breaks if donor DNA with homology to the genome editing site is delivered to target cells (Bibikova *et al.*, 2001, 2003). HRGT allows sequence insertion, deletion, or replacement at the target site based on the sequence of the donor DNA. Thus, it enables precise genome modifications such as modification of the activity centre of an enzyme or replacement of specific regulatory element of a gene (e.g. promoter, transcription factor DNA binding domain). In addition to directed genome manipulations, HRGT presents the opportunity to developing landing pads for gene stacks, ultimately increasing expression predictability of transgenes and significantly reducing the downstream process of selecting the 'best' transgenic lines (Yau *et al.*, 2013; Fogg *et al.*, 2014; Baltes and Voytas, 2015). Despite the general low rate of HRGT events, the use of plant tissues with high division and regeneration potencies, as well as increasing the availability of donor DNA template in target cells, have been promising strategies to increase the success rate of HRGT events (Chan *et al.*, 2011; Schiml *et al.*, 2014; Cermak *et al.*, 2015).

EXPRESSION CONTROL OF TRANSGENES *IN PLANTA*

One defining emphasis of synthetic biology is the designed control of gene expression. Through various approaches, expression control of genes can be controlled at the DNA, RNA, or protein level, depending on the strategy or application. Many efforts in microbial systems have involved tight expression regulation of inducible systems, many times for the intended purpose of building genetic circuits. Much progress has been made in building synthetic circuits in microbes; however, the focus of many of these efforts has revolved around temporal and expression levels. One emphasis of plant synthetic biology will be new tools and approaches that not only address temporal but also spatial expression regulation, based on the complexity and needs of working with multicellular organisms. In the following paragraphs we outline various strategies and potential future efforts in improving transgene expression regulation *in planta* (Figure 1).

Expression control at the transcriptional level

The most straightforward way to control gene expression is the selection of the promoter. A subfield of synthetic biology has focused on the characterization and modification of promoter elements to engineer them for more precise and tuneable expression levels. These synthetic promoters have a great deal of potential when translated into crop plants. This includes tweaks on endogenous plant promoters versus the completely synthetic and orthogonal approach of introducing entirely synthetic gene networks with synthetic promoters. It is also expected that this will increase promoter diversity and engineering options with the potential to reduce gene silencing, which can be caused by either high constitutive expression of genes or the expression of multiple genes under the same promoter.

One approach to characterize and design promoters is through the identification of *cis*-elements within endogenous plant promoter regions. A combination of co-expression analysis and motif searches within upstream promoter regions has enabled computational predictions of important *cis*-elements that regulate spatiotemporal expression or induction by various environmental stimuli. This approach has been successful in designing and building cyst nematode-responsive synthetic promoters (Liu *et al.*, 2014) as well as auxin-responsive promoter like DR5 (Ulmasov *et al.*, 1997). Thus, by screening within one plant species, one can identify conserved motif elements that may enable the redesign of synthetic promoters composed of reshuffled *cis*-elements. Potential limitations of this approach include the necessity to have access to transcriptome data for co-expression analysis. Furthermore, one could generate many different synthetic promoters; however, each promoter would have to be validated experimentally for each set of synthetic promoters desired for each specific condition of inducibility.

Another approach to designing synthetic promoters is the complete introduction of an orthogonal gene network. This eliminates the requirement to predict pre-existing motifs for any given stimuli or tissue specificity. Synthetic promoters designed by entirely heterologous and orthogonal systems have been implemented for almost 2 decades in plants, essentially hijacking other transcriptional systems such as the GAL4 and LacI systems (Moore *et al.*, 1998; Johnson *et al.*, 2005). To this extent, programmable synthetic transcription factors have been developed to bind

and induce expression of genes downstream of synthetic promoters in eukaryotic systems (Khalil *et al.*, 2012). In an attempt to design even more orthogonal systems to decrease any potential of interfering with the endogenous transcriptional machinery, orthogonal polymerases have been engineered into microbes (Temme *et al.*, 2012). For example, the bacteriophage T7 RNA polymerase was fused with a nuclear localization signal and expressed under plant tissue-specific promoters; transgenes under the control of T7 expression signal showed tissue-specific expression with enhanced expression strength (Nguyen *et al.*, 2004).

A modified version of CAS9 has been developed in which both of the catalytic sites are mutated to generate 'dead' CAS9 (dCAS9). dCAS9 loses its nuclease activity but maintains its sgRNA-mediated DNA-binding capability. Using dCAS9 as a platform to fuse to transcriptional activators or repressors, targeted activation or repression respectively of external reporters or endogenous genes has been achieved in bacteria (Qi *et al.*, 2013), yeast (Gilbert *et al.*, 2013), human cells (Gilbert *et al.*, 2013; Qi *et al.*, 2013), and plants (Lowder *et al.*, 2015; Piatek *et al.*, 2015). Unlike ZFN- or TALEN-based activation/repression systems, CRISPR/dCAS9-based targeting is not sensitive to DNA methylation and can be used to activate transcriptionally silenced genes (Lowder *et al.*, 2015). CRISPR/dCAS9-mediated repression could be used to turn off gene expression in a reversible manner and in a tissue-specific manner, in contrast with RNAi, which often moves from cell to cell (Brosnan and Voinnet, 2011). Furthermore, multiplexed targeting with multiple sgRNAs is also applicable for targeted gene expression control. The use of multiple sgRNA targeting the same promoter has been shown to enhance activation and repression effects compared with single sgRNA targeting, which provides an option for precise transcriptional regulation of the targeted promoters (Piatek *et al.*, 2015).

The increasing number of innovative approaches for transgene expression regulation at transcriptional level is of great value to support precise engineering multicellular organisms, such as crop plants, and optimize traits outputs while minimizing potential side effects. Versatile and universally fitted transcriptional regulation tools, together with post-translational regulation tools as discussed below, serve as a basis for the design of complex genetic circuit (Brophy and Voigt, 2014).

Manipulation of mRNA stability at the splicing step with an exon skipping system

The application of inducible promoters or spatial and temporal specific promoters to drive transgene expression is the main strategy to optimize metabolic flux through a synthetic or optimized pathway and to prevent undesired phenotypes that are often observed when transgenes are constitutively expressed (Padidam, 2003). However, basal level expression due to leakiness of promoter activities can still cause undesired phenotypes for genes encoding toxic proteins or those that function at very low expression levels (Padidam, 2003). Conditional splicing systems have been constructed to regulate gene expression in budding yeast (Weigand and Suess, 2007) and mammalian cells (Kim *et al.*, 2008; Culler *et al.*, 2010). In many plant lineages, the expression of transcription factor IIIA is regulated by the P5SM RNA element, which is composed of an exon and two flanking introns. Exon skipping by complete splicing of P5SM requires ribosomal protein L5 and allows the expression of transcription factor IIIA. A hybrid version of P5SM (HyP5SM) was synthesized that combined the features of rice L5 and Arabidopsis L5 proteins. HyP5SM can be effectively spliced by the rice (monocot) L5 protein (OsL5) but not with the OsL5 orthologous protein from Arabidopsis or tobacco (dicots) allowing the use of HyP5SM to establish a OsL5-dependent transgene expression in dicot plants (HyP5SM is properly spliced only in the presence of OsL5 protein). This was the case when a p35S-driven eGFP reporter carrying HyP5SM showed traceless expression in the absence of OsL5 (Hickey *et al.*, 2012). In plant immune responses, plant resistant proteins (e.g., receptor BS2) interact with pathogen effectors (e.g., avrBS2) and activate the hypersensitive response, which is one of the most sensitive phenotypes with a very low threshold requirement for the presence of resistant proteins and corresponding effectors (Gonzalez *et al.*, 2015). Gonzalez and colleagues demonstrated the use of the exon skipping system in a combination of transcriptional controls to achieve non-leaky transgene expression in the absence of the OsL5 splicing component. This expression switch requires that both promoters drive the expression of OsL5 and the transgene harbouring the HyP5SM cassette are active simultaneously and in the same cell to produce the protein encoded by the transgene. Not only does control of exon skipping prevent transgene expression, it also controls its induction stringently in a temporal and spatial manner (Gonzalez *et al.*, 2015).

Manipulation of mRNA stability (RNase-mediated transcript repression)

While the exon skipping system represents a 'switch on' control for conditional expression, a 'switch off' control will extend the flexibility of transgene expression control at the post-transcriptional level. We previously proposed that endoribonucleases like CSY4 from *Pseudomonas aeruginosa* may serve as 'switch off' controls (Eudes *et al.*, 2014). In *P. aeruginosa*, CSY4 is a component of the CRISPR/CAS immune system and functions in crRNA maturation by cleaving repetitive units in CRISPR arrays (Haurwitz *et al.*, 2010). The cognition sequence for CSY4 cleavage is a 28-bp unique sequence. In a synthetic biology application, CSY4

cleavage was utilized to dissociate transgenes from their linked genetic elements at the transcript level in *E. coli* and yeast (Qi *et al.*, 2012). In mammalian cells, CSY4 cleavage, together with other RNA regulatory strategies, enabled expression of protein and gRNAs from a single RNA transcript (Nissim *et al.*, 2014). In *E. coli*, yeast, and mammalian cell studies, effective cleavage at one or more CSY4 cognition sites in one transcript was demonstrated. Because of the broad functionality of CSY4 in these organisms, it can be expected that site-targeted endoribonucleases can be fashioned to work in plants by adding CSY4's recognition cognition sequence in either the coding sequence or untranslated regulatory regions (Qi *et al.*, 2012; Eudes *et al.*, 2014; Nissim *et al.*, 2014; Borchardt *et al.*, 2015). Such an RNase-based system can be used to generate context-free genetic elements, resulting in predictable expression regulation of one or multiple genes.

Control of translation activity (riboswitch-dependent translation)

Riboswitches are mRNA elements that regulate gene expression in response to ligand binding. Upon ligand binding at the 'aptamer' domain, conformation changes are transmitted to the 'expression platform' domain of the riboswitch, resulting in inhibition or activation of the regulated mRNA expression (Serganov and Nudler, 2013). Regulation of gene expression by riboswitches is found in all domains of life, but is most prevalent in bacteria (Rodionov *et al.*, 2002; Kubodera *et al.*, 2003; Sudarsan *et al.*, 2003; Winkler and Breaker, 2005; Croft *et al.*, 2007). Natural riboswitches regulate expression of genes that are involved in the biosynthesis, catabolism, signalling, or transport of the riboswitch ligand, maintaining the proper physiological level of the ligand (Serganov and Nudler, 2013). The modular structure of riboswitches facilitates their engineering and adaptation for transgene expression regulation. Potential applications of natural and engineered riboswitches have been extensively studied in bacteria (Wittmann and Suess, 2012). Riboswitch-based transgene regulation in higher organisms is more challenging but still holds lots of promise. The challenge partially stems from the complexity of the post-transcriptional gene regulation. The thiamine pyrophosphate (TPP) riboswitch is a naturally occurring riboswitch that functions in preserving metabolite homeostasis (Bocobza and Aharoni, 2014) and is widely distributed in plant species (Wachter *et al.*, 2007). The feasibility of riboswitch-based transgene regulation is further supported by Arabidopsis studies, in which the expression of YFP was tuned by a TPP riboswitch in response to *in vitro* TPP concentration (Bocobza and Aharoni, 2014). Higher organisms have specified cellular and subcellular functional units. Functional compartmentation may limit the accessibility and availability of certain metabolites to mRNA molecules. Nevertheless, this obstacle may be circumvented by using riboswitches to sense metabolite intermediates in a specific pathway, rather than in the final product. The flexibility of riboswitch control comes from the RNA aptamer, which in theory can be engineered to bind to any kind of molecule, including nucleic acid, amino acids, and metabolites (Berens *et al.*, 2015), as well as from the expression platforms that execute gene expression regulation at diverse levels (Henkin, 2008). Two recent studies exemplified sophisticated transgene control by riboswitches in plants. In *Chlamydomonas reinhardtii* chloroplasts, inducible repression of two essential genes *Rps12* and *RpoA* was obtained by down-regulation of their upstream regulator Nac2, which is controlled by a TPP riboswitch (Ramundo *et al.*, 2013; Auslander *et al.*, 2014). In another example, T7 RNA polymerase was fused with a theophylline-activated riboswitch and induced robust transgene expression in the presence of theophylline in tobacco chloroplasts. Combining a riboswitch control and secondary effectors (like the nuclear *Nac2* gene or RNA polymerase), the two studies achieved robust and conditional transgene expression in plastids. Plant plastids, in which the gene expression mechanism is more closely related to bacterial systems, serve as a convenient system for the study and application of riboswitch-mediated regulation (Bocobza and Aharoni, 2014) in plants.

Emerging tools for tight transgene expression regulation are necessary to expand the portfolio of synthetic biology tools for metabolic engineering of plants. Many of these tools are still too premature to be fully deployed. For example, exon-skipping-based transgene expression regulation has been developed for dicot plants, but its uses in monocot plants demand the constructions of a new synthetic P5SM that is not recognized by the monocot L5 protein. The structure details in the interaction model of P5SM and L5 (Hammond *et al.*, 2009), as well as the experience in the designing of HyP5SM (Hickey *et al.*, 2012) may shed light on optimized P5SM. Achievement of the full potential of riboswitches relies on the capability of constructing custom riboswitches. Encouragingly, the upcoming techniques and strategies, including information-intensive rational design (Auslander *et al.*, 2014) and high-throughput screening *in vivo* and analysis (Townshend *et al.*, 2015), will speed up the construction of active custom riboswitches for plants. Lastly, the proper selection and creative integration of synthetic regulators with cellular machinery may be critical for the successful application of the synthetic tools.

Protein expression regulation at the post-translational level

Synthetic biology approaches to pathway regulation at the protein level have been tailored for a wide variety of applications. Although most of these methods have been developed for microbial systems, many can be adopted for

tighter activity regulation of transgene products in crop plants for multigene expression, tighter control on protein levels, or regulation of protein activity. Tight spatiotemporal regulation may be necessary for either tissue-specific or environmentally responsive engineering efforts in future bioenergy crops.

One of the targets of plant synthetic biology is to provide the ability to express multiple proteins simultaneously and at stoichiometry. Because plants do not have the luxury of utilizing operons for multigene expression, it is often technically cumbersome to build constructs for expression of more than one gene. To address this issue, a system adopting the self-cleaving 2A peptide has been successfully used in plants to express multiple genes from one transcript (Halpin *et al.*, 1999). The 20-amino acid peptide is placed between two protein-encoding sequences to generate a single transcript. During translation of the 2A peptide sequence, the first protein moiety splits from the second half during translation, allowing two separate proteins to be produced from one transcript. Engineering fusion proteins with the 2A peptide will facilitate many future efforts in the simultaneous expression of proteins from a single promoter; however, the 2A peptide will remain on the C-terminus of the first protein, which may affect protein activity.

Regulation at the protein level offers a more direct and quicker response than going through signalling pathways to change transcriptional expression. Therefore, tight regulation of protein activity or expression in response to environmental stimuli can sometimes be more attractive. Proteins have been engineered to be activated by light (Nihongaki *et al.*, 2015) in mammalian cells, temperature (Shen *et al.*, 2012) *in planta*, and small molecule ligands as biosensors in *E. coli* (Lopez and Anderson, 2015). For example, thermoregulation of xylanases has been developed to degrade plant cell wall material via self-splicing bacterial inteins (Shen *et al.*, 2012). Tight activity regulation of cell wall degrading enzymes is necessary because expression *in planta* may result in many pleiotropic effects detrimental to overall plant fitness. With an increase in temperature >59°C, the intein would self-splice and reconstitute a functional xylanase. Regulation of transgene products at the protein level provides an additional layer for precise activity/expression control of heterologous proteins needed to not only improve crop yields but also for downstream processing of bioenergy crops.

Protein modifications have also been used for the synthetic regulation of metabolic flux. The enzymes that constitute a specific metabolic pathway have been spatially brought together to increase metabolic channelling between concurrent enzymatic reactions using synthetic protein scaffolds. Dueber and colleagues successfully implemented this strategy by heterologously expressing various interaction domain/ligand pairs from metazoans in yeast (Dueber *et al.*, 2009). Scaffolds of the interaction domains recruited enzyme–ligand fusions within much closer proximity to one another to increase flux through the pathways of interest, resulting in higher yields. With a clear hope and desire to express whole synthetic metabolic pathways in plants for various uses, regulation and increased flux through these pathways through protein modifications present another key example of how some of the many synthetic biology tools that have been developed for microbes may be leveraged by plant biologists.

GENE STACKING IN CROP PLANTS

Most plant engineering efforts have been based on the targeted overexpression of one transgene. In contrast, efforts using simpler unicellular systems have explored the expression of multiple genes simultaneously for introduction of multi-enzyme metabolic pathways, genetic circuits, and complex traits. Just as breeders will stack multiple traits into a single cultivar, molecular biology provides the means to stack multiple genes that will be transferred simultaneously into a single plant. The stacking of multiple genes/traits through DNA assembly is integral to many synthetic biology projects, as more intricate and complex approaches often require more parts. The ability to efficiently stitch strands of DNA together may seem straightforward, but in practice there are many approaches and challenges, given the size and scope of a given project. Thus, many groups have recognized the challenges involved in DNA assembly, and multiple tools have been developed. Plant molecular biologists are now catching up to their microbiologist counterparts and have adopted many of these newer strategies to manipulate plant binary vectors. These methods will allow plant biologists to transfer DNA material into plant genomes more easily and effectively.

One reason why bacterial systems are easier to work with is the ability to control multiple genes under one promoter by designing operons. One promoter can be used to drive several genes, with ribosomal binding sites separating each coding sequence. Ribosomal binding sites in many bacterial systems have been predicted and verified experimentally, allowing the expression level of each gene within the operon to be tuned to certain expression strengths. Eukaryotic systems, like plants and yeast, do not have the luxury to easily manipulate multiple genes simultaneously and need each coding sequence to be driven by a promoter and terminator element, thus making DNA assembly a more tedious and complex process.

One possible way to express multiple transgenes simultaneously is to cross plants expressing different transgenes. This method is not ideal for plant synthetic biologists for several reasons. Transgenes are randomly inserted into the genome and sometimes in multiple copies, and thus there may be positional effects on the

expression of the transgene. This variation forces plant biologists to screen multiple transformants and verify the number of transgene insertions, to determine whether and to what level the transgene is being expressed – not a trivial task. Consequently, it is more desirable to stack genes within one T-DNA and introduce all transgenes of interest simultaneously. Gene stacking permits all the transgenes to be inserted in the same locus within the genome, lessening the complications of positional effects that could affect the different transgenes.

Another challenge biologists face is the robust ability of plants to silence transgenes. This process is primarily associated with defence against potential pathogens, especially virus infections, as a key mechanism by which plants destroy transcripts and suppress expression of foreign genes. Unfortunately for plant molecular biologists, the introduction of genes of interest into plants may sometimes trigger this response. Even a single gene expressed constitutively may run the risk of being silenced after a few generations. Thus, expression of more than one gene constitutively will have an even greater chance of running into silencing issues. When stacking genes in plants, it is important to avoid gene silencing issues through the use of different promoters, each driving a different coding sequence. Traditionally, plant molecular biologists clone a single gene into a multiple cloning site in a binary vector for plant transformation. However, one can easily imagine how stacking the expression of even two genes simultaneously becomes difficult and cumbersome when using traditional large binary vectors that have been primarily and originally intended for the expression of one single gene. A key criterion for facilitating interchangeable and universal DNA parts (i.e., promoters, coding sequences, terminators, etc.) is standardizing assembly methods and DNA components (Patron *et al.*, 2015). By doing so, various DNA parts can be freely exchanged between researchers. In the following paragraphs, we outline a few existing methods for gene stacking in plant systems (Table 2).

Table 2 Summary of DNA assembly methods

Method	Description
Traditional (MCS)	Traditional plant binary vectors (e.g. pCAMBIA) allow for integration of a single coding sequence into a multiple cloning site by choosing from several restriction enzymes
BioBrick	Standardized parts that allow for an iterative assembly of parts. Restriction enzyme dependent. Iterations allow for DNA assembly, but only two parts can be added together at a time – thus time consuming (Philips and Silver, 2006; Anderson *et al.*, 2010; Boyle *et al.*, 2012)
Gateway	Circumvents restriction enzyme cloning. Various Gateway-compatible binary vectors have been already made (Earley *et al.*, 2006). DNA assembly is limited by the number of recombination sites (5 maximum). Proprietary (Life Technologies)
Gibson	Assembly of multiple parts without scars. Largely PCR-based, therefore needs to be thoroughly resequenced after cloned into destination vector (Gibson *et al.*, 2009)
Golden Gate	Multi-part assembly from standardized parts. Type II restriction enzyme dependent (Rebatchouk *et al.*, 1996). Hierarchical Golden Gate cloning is implemented for plant binary vectors: GoldenBraid and MoClo systems (Weber *et al.*, 2011; Sarrion-Perdigones *et al.*, 2013)

The BioBrick DNA assembly standard is a large concerted effort to develop a widely accepted method for DNA assembly in synthetic biology. Iterative assembly of various vectors housing DNA parts could be digested with compatible sticky ends from different conventional restriction enzymes. This ultimately allows vectors with various DNA parts to be deposited into a central repository to facilitate the open-source free exchange of parts between different laboratories and institutions. One limitation of the BioBrick system is its limited assembly capability, as only two parts can be fused together at one time. This only allows for iterative assembly, in comparison with other methods (described in this section), which may permit the assembly of multiple parts at one time. Another drawback of the BioBrick system is that the use of four restriction enzymes to generate compatible sticky ends necessitates that all DNA parts are void of these restriction enzyme cut sites. Furthermore, since the inception of the BioBrick method, numerous variations of the assembly with different combinations of different restriction enzymes have been developed (Philips and Silver, 2006; Anderson *et al.*, 2010). Thus, not all BioBrick parts in the repository are compatible. Nonetheless, plant BioBrick vectors have been developed (Boyle *et al.*, 2012). Although there are shortcomings with the BioBrick assembly, the open-source spirit has remained a core component of the synthetic biology community.

A method that has gained and retained popularity within the plant molecular biology community is the Gateway system. By using proprietary enzymes from the company Life Technologies, this recombination-based method eliminates the need for specific restriction enzyme cut sites. Although many plant binary vectors have been developed to easily shuffle genes into different Gateway-compatible destination vectors (Earley *et al.*, 2006), this method leaves larger scars in contrast with restriction-based DNA assembly, and the majority of these vectors are still designed for a very limited number of genes to be expressed *in planta.* Using the Gateway system, multiple DNA parts can be assembled into a destination vector in one reaction, however unlike the BioBrick assembly this is not an iterative process and more genes cannot be subsequently added into the destination vector. Therefore, the Gateway system may face challenges as a standardized DNA assembly method for the plant synthetic biology community.

Recent methods that have become more widely used and have excelled in multi-fragment DNA assembly have explored the use of type IIs restriction enzymes. Type IIs restriction enzymes cut at a defined distance away from their recognition site, thus enabling researchers to design 'custom' sticky ends and providing directionality of assembled DNA parts. Twenty years ago, NOMAD (nucleic acid ordered assembly with directionality), the birth of standardized assembly method exploiting the use of type IIs restriction enzymes, was described (Rebatchouk *et al.*, 1996). Since this time, additional methods have been developed, notably the Golden Gate assembly method, which allows assembly of more than 10 parts simultaneously (Weber *et al.*, 2011). Variations of hierarchical Golden Gate assembly have been described for plant binary vectors using a standardized syntax for stitching together promoters, coding sequences, and terminators. The nascent plant synthetic biology community is working to build international recognition of certain syntax to be used for iterative Golden Gate assembly for plant synthetic biology. Two widely used methods are the MoClo (Weber *et al.*, 2011) and Golden Braid systems (Sarrion-Perdigones *et al.*, 2013).

The ability to introduce multiple genes simultaneously into crop plants opens the door to more sophisticated and targeted manipulations of plant genomes. New traits and metabolic pathways that may make crop plants more amenable to societal and bioenergy needs will be pivotal for the future of plant biotechnology. It is not clear what limitations on size and content exist for delivery of DNA material into plants for more extensive and complex engineering efforts. The development of synthetic plant chromosomes has been used as a platform to address some of these issues; however, there have been some logistical pitfalls, and much more basic understanding of the system is necessary before widespread application of this technology is feasible (Birchler, 2015). Until this time, traditional methods utilizing *Agrobacterium*- or biolistic-mediated transformations that can leverage the earlier-mentioned methods will provide conventional approaches to delivering multiple genes into plant genomes.

ENGINEERING BIOENERGY TRAITS IN CROPS

Synthetic biology has been used to manipulate and introduce metabolic pathways into various hosts. To this end, microbes such as yeast and *E. coli* have been engineered to make molecules of interest to serve societal needs (Keasling, 2010), ranging from pharmaceutical to bioenergy applications (Ro *et al.*, 2006; Bokinsky *et al.*, 2011). Different bioenergy crops are grown for different biofuel feedstocks materials. Thus, disparate crops will require different emphases in the traits to be metabolically engineered. The three major areas in which bioenergy crops will play a pertinent role as feedstocks are: (i) lignocellulose production, (ii) oil production, and (iii) soluble sugar production (Figure 2).

Lignocellulose biomass crops

The majority of carbon fixed by plants ends up in plant cell walls stored as sugar polymers, thus making plant cell walls an attractive feedstock for microbial production of biofuels. Although simpler engineering efforts using the expression of either single genes or RNAi constructs have been shown to improve yields for lignocellulosic-derived biofuels, the introduction of more complex metabolic pathways may improve the yield of sugar from cellulose or decrease the recalcitrance of lignin, improving the extraction of cell wall sugars and reducing the overall cost of the entire process (Klein-Marcuschamer *et al.*, 2010; Loqué *et al.*, 2015).

A major obstacle in lignocellulosic-based biofuels is the difficulty of freeing sugars from cell wall polysaccharides because of sequestration by lignin polymers. Early studies have shown that lignin content could be decreased by targeting lignin biosynthesis genes with RNAi (Reddy *et al.*, 2005; Chen and Dixon, 2007; Smith *et al.*, 2013). However, many times this approach resulted in dwarf plants, highlighting the structural significance of lignin (Bonawitz and Chapple, 2013). The introduction of metabolic pathways that compete away substrate from the lignin biosynthetic pathway or produce enzymatic inhibitors has also been shown to be an effective method for decreasing lignin content without resulting in dwarfed plants (Eudes *et al.*, 2015, 2016).

Another approach that has only recently been explored is manipulating the sugar composition of the plant cell wall. Specifically, it has been proposed that increasing the C6 to C5 sugar ratio (hexoses and pentoses, respectively) would ultimately yield a higher titre of biofuels, as C6 sugars are usually more efficiently fermented into ethanol than C5 sugars (van Vleet and Jeffries, 2009; Young *et al.*, 2010). Some early studies have shown that plant cell walls can be engineered with reduced xylan content (Petersen *et al.*, 2012) as well as manipulated to enrich them with galactan (a C6-sugar polymer; Gondolf *et al.*, 2014) or mixed-linkage glucan (a C6-sugar polymer; Vega-Sanchez *et al.*, 2015). All approaches resulted in an overall increase in C6 sugars, and no deleterious growth effect was observed.

The next expected engineering steps would be to translate these approaches to bioenergy crops and stack both the low lignin traits with the enrichment of C6 sugar ratio traits to gain a cumulative benefit of each trait without affecting plant development. This goal could be seen as extremely long and challenging, but with the emergence of the necessary synthetic biology tools, it could be achieved in a short time. For example, all the genes used to develop these straits could be rapidly assembled within a single T-DNA, and various promoter-encoding sequence

combinations and gene stack combinations could be generated simultaneously. Furthermore, synthetic promoters could be utilized to synchronize and express a series of genes at the same time or in specific tissues. More precise engineering could be further developed using molecular switches. For example, they could be implemented to turn off the expression of transgenes involved in lignin reduction under external stresses, such as drought stress, to retain high hydrophobicity of the vessels to reduce risks of embolism. Alternatively, or in combination with synthetic transcriptional activators, cell wall hydrophobicity could be enhanced by boosting vessel lignification. In this particular example, synthetic transcriptional activators would be designed to enhance expression of multiple genes from the lignin biosynthesis pathway or a transcription factor such as AtMyb58 (Zhou *et al.*, 2009) to increase the metabolic flux through this pathway.

It is important to note that although the plant research community has started to build an armamentarium of tools for more complex plant metabolic engineering, many aspects of plant metabolism (e.g., feedback regulations) are not yet fully understood. In the case of lignocellulosic bioenergy crops, we do not completely understand the physiological ramifications of manipulating cell walls, as they play a role in many other aspects of plant physiology (e.g., turgor pressure, water potential, and plant–pathogen response). Synthetic biology will not only allow us to engineer plants with more complex pathways to test the boundaries of plant cell wall modifications, but synthetic biology can leverage our basic understanding of plant cell wall properties, and ultimately enable more precise and targeted approaches in future engineering efforts. This basic research will provide the foundation needed to manipulate plant cell walls and prevent the appearance of pleiotropic and detrimental effects on overall fitness.

Oil-rich crops

Bioenergy crops that are grown for the eventual production of biodiesel require an entirely different strategy in metabolic engineering, as oils from these crops are primarily harvested from their seeds. Two areas of focus that have predominated efforts in metabolic engineering of oil seed crops are: (i) increasing the oil content, and (ii) manipulating the lipid composition.

Increasing the oil content of crops is of biotechnological relevance to both bioenergy and agricultural needs. Because these crops (e.g., canola and camelina) accumulate large amounts of oil in their seeds, most engineering efforts have emphasized increasing oil production in seeds. Expression of transcriptional regulators (e.g., *WRI1*, *LEC1*, and *LEC2*) and key triacylglycerol biosynthesis enzymes (e.g., *DGAT* and *MGAT*) has been successfully manipulated to further enrich the oil content of seeds (Napier *et al.*, 2014). Furthermore, decreasing flux to competing carbon sinks, such as starch, through knockdown approaches that target important metabolic forks has also yielded higher seed oil content and overall oil production (Slocombe *et al.*, 2009; Kelly *et al.*, 2013). More recently, studies have shown how gene stacking has enabled more sophisticated metabolic engineering to additively increase the oil yields in comparison with earlier efforts that overexpressed a single gene (van Erp *et al.*, 2014; Vanhercke *et al.*, 2014). With efficient editing tools such as CRISPR/CAS9-based approaches (Puchta, 2016), it will become possible to engineer oil crops and conditionally or developmentally repress endogenous competitive carbon sinks during the seed filling period.

Another strategy has been to increase and expand oil/triacylglycerol production in vegetative tissue, because the majority of the plant is composed of vegetative tissue. Leafy crop plants, such as *Nicotiana tabacum*, have been metabolically engineered by stacking genes to increase triacylglycerol production in leaf tissue to levels that are competitive with those seen for oilseed crops (Vanhercke *et al.*, 2014). Another impressive study demonstrated how metabolic engineering of sugarcane could be used to accumulate significant amounts of triacylglycerols in the vegetative biomass (Zale *et al.*, 2016). Furthermore, increases in triacylglycerol (TAG) in Arabidopsis leaf tissue have been achieved with alternate approaches that do not require the overexpression of any genes, but rather through mutations in genes such as TRIGALACTOSYLDIACYLGLYCEROL1 (*TGD1*) and SUGAR-DEPENDENT1 (*SDP1*), in order to decrease fatty acid turnover and block rates of fatty acid beta-oxidation (Fan *et al.*, 2014). Metabolic engineering feats in high-impact bioenergy crops, such as sugarcane, provide an example of the role metabolic engineering and synthetic biology may play in the future of bioenergy crops.

Beyond increasing oil content, metabolic engineering has been used successfully to manipulate the fatty acid composition of plants. Through the genetic manipulation of crop species such as soybean and canola, fatty acid saturation composition has been modified to improve nutritional value or economic relevance. For example, studies have shown that the oleic acid content can be significantly increased in soybean oil to levels that will improve the nutritional value and stability of the oil (Pham *et al.*, 2010). More complex fatty acid modifications have also been applied to seed oils by introducing and expressing up to nine transgenes into *Brassica juncea* seeds to address the human health benefits of very-long-chain polyunsaturated fatty acids (Wu *et al.*, 2005). As in studies that manipulate fatty acid composition for nutritional benefits, the optimal biodiesel will benefit from precise manipulation of the lipid pathway in oil crops. Biodiesels from different crop plants have varying chemical properties that may make it difficult to use them in varying geographical regions. For example, the cloud point, which refers to the temperature at which a

given fuel begins to thicken, of biodiesels derived from different crops can vary widely. Palm biodiesel typically has a cloud point of 17°C, versus canola biodiesel which has a cloud point of about 0°C (Mittelbach and Remschmidt, 2004; Moser, 2008). This is in comparison with the standard diesel, D2, which has a cloud point of −12°C. Thus, palm biodiesel cannot be used in cold climates unless the chemical composition of the TAG is modified, which can be achieved through genetic manipulation. Similar to the isoprenoid pathway engineering in plants, synthetic biology can be leveraged to manipulate TAG composition in various crops, as well as fine tune and optimize expression of genes of interest in order to optimize metabolic flux through a given pathway (Reviewed in Fesenko and Edwards, 2014; Kempinski *et al.*, 2015; Yuan and Grotewold, 2015).

There are some immediate shortcomings in the current market for biodiesel. Namely, conversion of all American vegetable oil production to biodiesel would replace only 10% of current diesel consumed. Currently, the devotion of some of this sector to biodiesel has already contributed to increased vegetable oil prices, and such a drastic shift in diverting plant oils to biodiesel would have unintended consequences on other sectors of the economy, especially food prices. Thus, production of biodiesel from seed oil is far from being a panacea to the biofuel problem, but a combination of improved yields through biotechnology – possibly with the development of increased vegetative oil production (Xu and Shanklin, 2016) – in conjunction with prudent policies that are aimed at slowly mitigating unintended effects on other sectors may enable a sustainable solution through biodiesel.

Soluble sugars-rich crops

Soluble sugars make up a large and commercially important sector of the existing bioenergy feedstock market. Currently, sugarcane is the most successful bioenergy crop, as large amounts of its fermentable sugars are water soluble and are extracted directly from the plant very efficiently and at low cost. The remaining lignocellulosic material is burned for downstream fermentation processing or saccharified to generate an additional fermentable sugar stream. However, even with such an 'ideal' bioenergy crop, there are further modifications that plant metabolic engineering may offer in improving soluble sugar yields in sugarcane that could be also applied in crops that accumulate less soluble sugars, like sweet sorghum. Sugars and sugar phosphates are core to central carbon metabolism, complicating the ability to significantly modify and perturb their levels. Nonetheless, there are two complementary approaches that may provide strategies for increasing sugar levels: increasing flux on the source side or on the sink side. By pushing or pulling on carbon, one can potentially increase the theoretical amount of photosynthate that can be produced or redirect sugar biosynthesis towards other products of interest to pull more flux through the pathway.

Efforts have been made to increase carbon fixation on the source side to eventually push higher yields of sugar through improving the photosynthetic capacity of crop plants. Although this in itself warrants its own review (Ort *et al.*, 2015), we highlight just a few strategies that have been considered for improving photosynthesis for the eventual increase in biomass, growth, or overall carbon fixation. At the heart of many decades of research into improving carbon fixation is the key enzyme RuBisCO. Improving the kinetics of RuBisCO has been resistant to molecular engineering efforts for decades, however expression of heterologous RuBisCO from cyanobacteria in crop plants may provide a way (Lin *et al.*, 2014b), especially given the improvements that have been made in improving cyanobacterial RuBisCO (Durão *et al.*, 2015). Other possible ways to improve carbon fixation involve addressing the carbon lost to photorespiration. Various metabolic engineering efforts have been proposed to decrease the inefficiencies of RuBisCO by addressing photorespiration through the introduction of bacterial metabolic pathways (Kebeish *et al.*, 2007; Shih *et al.*, 2014). Finally, potential enzyme bottlenecks in the Calvin–Benson cycle have been identified with metabolic flux studies and validated experimentally with overexpression studies, providing hints that metabolic engineering to optimize photosynthesis may be possible (Miyagawa *et al.*, 2001). All these efforts would ultimately yield increased flux through sugars and may eventually be used to increase soluble sugar yields in crops such as sugarcane.

Complementary approaches in metabolic engineering will focus on the sink-side strategies on pulling more carbon through sugar biosynthesis into other products of interest. Because sucrose accumulation will inhibit photosynthesis, it is important to draw the sugar into a different sink, ideally another carbon source or sugar that will not induce a negative feedback on the photosynthetic machinery. Pulling on sucrose by diverting its biosynthesis into the sugar isomers isomaltulose and trehalulose has been successfully engineered into sugarcane to increase overall soluble sugar content in the plant (Wu and Birch, 2007; Hamerli and Birch, 2011). Other limitations that would need to be considered are the sugar transporters that are integral in the shuttling of sugars from source to sink tissue (Chen *et al.*, 2015; Rossi *et al.*, 2015). For example, the SWEET transporter family has been often highlighted to control carbohydrate translocation into sink tissues such as SWEET4 in rice and maize seeds (Sosso *et al.*, 2015) and SWEET9 in *Brassica* and *Nicotiana* nectarines (Lin *et al.*, 2014a). Moreover in an effort to increase the amount of galactose – and ultimately galactan deposition – in plant biomass, Gondolf *et al.* (2014) simultaneously overexpressed galactan synthase and the upstream epimerase

needed to divert flux from UDP-glucose to UDP-galactose in Arabidopsis. Despite the cell wall galactose content increase, they further suggested that the transport of UDP-galactose into the Golgi lumen is the next limiting step (Gondolf *et al.*, 2014). This suggestion would be in agreement with the increase in galactosylated cell wall components when the human UDP-galactose transporter 1 gene (*hUGT1*) and the Arabidopsis UDP-Rha/UDP-galactose transporter 1 (*URGT1*) are overexpressed in tobacco and Arabidopsis respectively (Rautengarten *et al.*, 2014; Abedi *et al.*, 2016).

It is important to note that even if bioenergy crops that accumulate large amounts of soluble fermentable sugars seem to be ideal crops, precise conditions are required to collect these sugars. After harvest, this biomass cannot be stored and needs to be processed immediately to avoid spoilage due to the easy accessibility of its fermentable sugars to various opportunistic microbes. Therefore, it could be more suitable to engineer these crops to be capable of converting these sugars into products that would be used as biofuel directly, such as triacylglycerols (Zale *et al.*, 2016) or biofuel precursors such as pinene or limonene (Dunlop *et al.*, 2011). Some of these approaches need significant metabolic remodelling, and thus require manipulating expression of several genes, such as induction and repression of native genes, as well as introduction of new genes to support the development of foreign pathways. A decade ago these sorts of projects may have seemed like a significant undertaking, however with recent advances and the support of metabolic engineering and synthetic biology these projects are becoming more common and routine (Napier *et al.*, 2014; Zale *et al.*, 2016).

The metabolic engineering of feedstock crops is not limited to modifying bioenergy traits. Rather, it will also be important in the future to improve abiotic and biotic stress resistance in various crops. For example, if cold tolerance traits can be easily engineered in sugarcane that would push its geographical/climate-dependence boundaries, it would boost the availability of easily extractable and fermentable sugars, and increase cost-competitiveness of biofuels. Thus, metabolic engineering will play an important role in introducing new traits into crop plants in the future.

CONCLUSION

As discussed in this review, advances have been made in developing synthetic biology tools to support plant metabolic engineering, including the optimization of some of the bioenergy crops using the first generation of tools. Beyond these bioenergy traits, it is very important to develop new or to optimize current agronomical traits and transfer these to the most promising bioenergy crops to grow these crops sustainably while increasing product yield. The consideration of: (i) how fast technologies related to analytic methods are progressing, allowing faster and cheaper analysis of complex samples, (ii) how sequencing and DNA synthesis costs are decreasing, (iii) the development of high-throughput and miniaturization assays supporting low cost of DNA assembly and enzymatic assays, and (iv) the development of new algorithms for large data analysis, metabolic engineering and synthetic biology have a bright future and will be of great support in tackling important challenges related to food, energy, and sustainability.

DISCLOSURE

DL has financial conflicts of interest in Afingen Inc., USA.

ACKNOWLEDGEMENTS

This work was part of the DOE Early Career Award and the DOE Joint BioEnergy Institute (http://www.jbei.org) supported by the United States Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the US Department of Energy. The US Government retains, and the publisher by accepting the article for publication acknowledges, that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide licence to publish or reproduce the published form of this manuscript, or allow others to do so, for US Government purposes. P.M.S. was supported by the Gordon and Betty Moore Foundation through Grant GBMF 2550.04 to the Life Sciences Research Foundation.

REFERENCES

Abedi, T., Khalil, M.F., Asai, T., Ishihara, N., Kitamura, K., Ishida, N. and Tanaka, N. (2016) UDP-galactose transporter gene hUGT1 expression in tobacco plants leads to hyper-galactosylated cell wall components. *J. Biosci. Bioeng.* **121**, 573–583.

Anderson, J., Dueber, J.E., Leguia, M., Wu, G.C., Goler, J.A., Arkin, A.P. and Keasling, J.D. (2010) BglBricks: a flexible standard for biological part assembly. *J. Biol. Eng.* **4**, 1–12.

Auslander, S., Stucheli, P., Rehm, C., Auslander, D., Hartig, J.S. and Fussenegger, M. (2014) A general design strategy for protein-responsive riboswitches in mammalian cells. *Nat. Methods*, **11**, 1154–1160.

Baltes, N.J. and Voytas, D.F. (2015) Enabling plant synthetic biology through genome engineering. *Trends Biotechnol.* **33**, 120–131.

Berens, C., Groher, F. and Suess, B. (2015) RNA aptamers as genetic control devices: the potential of riboswitches as synthetic elements for regulating gene expression. *Biotechnol. J.* **10**, 246–257.

Bibikova, M., Carroll, D., Segal, D.J., Trautman, J.K., Smith, J., Kim, Y.G. and Chandrasegaran, S. (2001) Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol. Cell. Biol.* **21**, 289–297.

Bibikova, M., Golic, M., Golic, K.G. and Carroll, D. (2002) Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics*, **161**, 1169–1175.

Bibikova, M., Beumer, K., Trautman, J.K. and Carroll, D. (2003) Enhancing gene targeting with designed zinc finger nucleases. *Science*, **300**, 764.

Birchler, J.A. (2015) Promises and pitfalls of synthetic chromosomes in plants. *Trends Biotechnol.* **33**, 189–194.

Bocobza, S.E. and Aharoni, A. (2014) Small molecules that interact with RNA: riboswitch-based gene control and its involvement in metabolic regulation in plants and algae. *Plant J.* **79**, 693–703.

Bokinsky, G., Peralta-Yahya, P.P., George, A. *et al.* (2011) Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered Escherichia coli. *Proc. Natl Acad. Sci. USA*, **108**, 19949–19954.

Bonawitz, N.D. and Chapple, C. (2013) Can genetic engineering of lignin deposition be accomplished without an unacceptable yield penalty? *Curr. Opin. Biotechnol.* **24**, 336–343.

Borchardt, E.K., Vandoros, L.A., Huang, M., Lackey, P.E., Marzluff, W.F. and Asokan, A. (2015) Controlling mRNA stability and translation with the CRISPR endoribonuclease Csy4. *RNA*, **21**, 1921–1930.

Bortesi, L. and Fischer, R. (2015) The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnol. Adv.* **33**, 41–52.

Boyle, P.M., Burrill, D.R., Inniss, M.C. *et al.* (2012) A BioBrick compatible strategy for genetic modification of plants. *J. Biol. Eng.* **6**, 1–8.

Brophy, J.A. and Voigt, C.A. (2014) Principles of genetic circuit design. *Nat. Methods*, **11**, 508–520.

Brosnan, C.A. and Voinnet, O. (2011) Cell-to-cell and long-distance siRNA movement in plants: mechanisms and biological implications. *Curr. Opin. Plant Biol.* **14**, 580–587.

Bruckner, K., Schafer, P., Weber, E., Grutzner, R., Marillonnet, S. and Tissier, A. (2015) A library of synthetic transcription activator-like effector-activated promoters for coordinated orthogonal gene expression in plants. *Plant J.* **82**, 707–716.

Cermak, T., Baltes, N.J., Cegan, R., Zhang, Y. and Voytas, D.F. (2015) High-frequency, precise modification of the tomato genome. *Genome Biol.* **16**, 232.

Chan, F., Hauswirth, W.W., Wensel, T.G. and Wilson, J.H. (2011) Efficient mutagenesis of the rhodopsin gene in rod photoreceptor neurons in mice. *Nucleic Acids Res.* **39**, 5955–5966.

Chen, F. and Dixon, R.A. (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nat. Biotechnol.* **25**, 759–761.

Chen, L.Q., Cheung, L.S., Feng, L., Tanner, W. and Frommer, W.B. (2015) Transport of sugars. *Annu. Rev. Biochem.* **84**, 865–894.

Clasen, B.M., Stoddard, T.J., Luo, S. *et al.* (2016) Improving cold storage and processing traits in potato through targeted gene knockout. *Plant Biotechnol. J.* **14**, 169–176.

Croft, M.T., Moulin, M., Webb, M.E. and Smith, A.G. (2007) Thiamine biosynthesis in algae is regulated by riboswitches. *Proc. Natl Acad. Sci. USA*, **104**, 20770–20775.

Culler, S.J., Hoff, K.G. and Smolke, C.D. (2010) Reprogramming cellular behavior with RNA controllers responsive to endogenous proteins. *Science*, **330**, 1251–1255.

D'Halluin, K., Vanderstraeten, C., Van Hulle, J. *et al.* (2013) Targeted molecular trait stacking in cotton through targeted double-strand break induction. *Plant Biotechnol. J.* **11**, 933–941.

Dueber, J.E., Wu, G.C., Malmirchegini, G.R., Moon, T.S., Petzold, C.J., Ullal, A.V., Prather, K.L.J. and Keasling, J.D. (2009) Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* **27**, 753–759.

Dunlop, M.J., Dossani, Z.Y., Szmidt, H.L., Chu, H.C., Lee, T.S., Keasling, J.D., Hadi, M.Z. and Mukhopadhyay, A. (2011) Engineering microbial biofuel tolerance and export using efflux pumps. *Mol. Syst. Biol.* **7**, 487.

Durão, P., Aigner, H., Nagy, P., Mueller-Cajar, O., Hartl, F.U. and Hayer-Hartl, M. (2015) Opposing effects of folding and assembly chaperones on evolvability of RuBisCO. *Nat. Chem. Biol.* **11**, 148–155.

Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K. and Pikaard, C.S. (2006) Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J.* **45**, 616–629.

Endo, M., Mikami, M. and Toki, S. (2015) Multigene knockout utilizing off-target mutations of the CRISPR/Cas9 system in rice. *Plant Cell Physiol.* **56**, 41–47.

van Erp, H., Kelly, A.A., Menard, G. and Eastmond, P.J. (2014) multigene engineering of triacylglycerol metabolism boosts seed oil content in Arabidopsis. *Plant Physiol.* **165**, 30–36.

Eudes, A., Liang, Y., Mitra, P. and Loque, D. (2014) Lignin bioengineering. *Curr. Opin. Plant Biol.* **26**, 189–198.

Eudes, A., Sathitsuksanoh, N., Baidoo, E.E.K., George, A., Liang, Y., Yang, F., Singh, S., Keasling, J.D., Simmons, B.A. and Loqué, D. (2015) Expression of a bacterial 3-dehydroshikimate dehydratase reduces lignin content and improves biomass saccharification efficiency. *Plant Biotechnol. J.* **13**, 1241–1250.

Eudes, A., Pereira, J.H., Yogiswara, S., Wang, G., Teixeira Benites, V., Baidoo, E.E., Lee, T.S., Adams, P.D., Keasling, J.D. and Loque, D. (2016) Exploiting the Substrate Promiscuity of Hydroxycinnamoyl-CoA: Shikimate Hydroxycinnamoyl Transferase to reduce lignin. *Plant Cell Physiol.* **57**, 568–579.

Fan, J., Yan, C., Roston, R., Shanklin, J. and Xu, C. (2014) Arabidopsis lipins, PDAT1 acyltransferase, and SDP1 triacylglycerol lipase synergistically direct fatty acids toward beta-oxidation, thereby maintaining membrane lipid homeostasis. *Plant Cell*, **26**, 4119–4134.

Fan, D., Liu, T., Li, C., Jiao, B., Li, S., Hou, Y. and Luo, K. (2015) Efficient CRISPR/Cas9-mediated targeted mutagenesis in Populus in the first generation. *Sci. Rep.* **5**, 12217.

Fesenko, E. and Edwards, R. (2014) Plant synthetic biology: a new platform for industrial biotechnology. *J. Exp. Bot.* **65**, 1927–1937.

Fogg, P.C., Colloms, S., Rosser, S., Stark, M. and Smith, M.C. (2014) New applications for phage integrases. *J. Mol. Biol.* **426**, 2703–2716.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A. 3rd and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, **6**, 343–345.

Gilbert, L.A., Larson, M.H., Morsut, L. *et al.* (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, **154**, 442–451.

Gondolf, V.M., Stoppel, R., Ebert, B., Rautengarten, C., Liwanag, A.J., Loque, D. and Scheller, H.V. (2014) A gene stacking approach leads to engineered plants with highly increased galactan levels in Arabidopsis. *BMC Plant Biol.* **14**, 344.

Gonzalez, T.L., Liang, Y., Nguyen, B.N., Staskawicz, B.J., Loque, D. and Hammond, M.C. (2015) Tight regulation of plant immune responses by combining promoter and suicide exon elements. *Nucleic Acids Res.* **43**, 7152–7161.

Gorbunova, V. and Levy, A.A. (1997) Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. *Nucleic Acids Res.* **25**, 4650–4657.

Halpin, C., Cooke, S.E., Barakate, A., Amrani, A.E. and Ryan, M.D. (1999) Self-processing 2A-polyproteins–a system for co-ordinate expression of multiple proteins in transgenic plants. *Plant J.* **17**, 453–459.

Hamerli, D. and Birch, R.G. (2011) Transgenic expression of trehalulose synthase results in high concentrations of the sucrose isomer trehalulose in mature stems of field-grown sugarcane. *Plant Biotechnol. J.* **9**, 32–37.

Hammond, M.C., Wachter, A. and Breaker, R.R. (2009) A plant 5S ribosomal RNA mimic regulates alternative splicing of transcription factor IIIA pre-mRNAs. *Nat. Struct. Mol. Biol.* **16**, 541–549.

Haurwitz, R.E., Jinek, M., Wiedenheft, B., Zhou, K. and Doudna, J.A. (2010) Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science*, **329**, 1355–1358.

Henkin, T.M. (2008) Riboswitch RNAs: using RNA to sense cellular metabolism. *Genes Dev.* **22**, 3383–3390.

Hickey, S.F., Sridhar, M., Westermann, A.J., Qin, Q., Vijayendra, P., Liou, G. and Hammond, M.C. (2012) Transgene regulation in plants by alternative splicing of a suicide exon. *Nucleic Acids Res.* **40**, 4701–4710.

Hsu, P.D., Lander, E.S. and Zhang, F. (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, **157**, 1262–1278.

Jacobs, T.B., LaFayette, P.R., Schmitz, R.J. and Parrott, W.A. (2015) Targeted genome modifications in soybean with CRISPR/Cas9. *BMC Biotechnol.* **15**, 16.

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B. and Weeks, D.P. (2013) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res.* **41**, e188.

Johnson, A.A.T., Hibberd, J.M., Gay, C., Essah, P.A., Haseloff, J., Tester, M. and Guiderdoni, E. (2005) Spatial control of transgene expression in rice (*Oryza sativa* L.) using the GAL4 enhancer trapping system. *Plant J.* **41**, 779–789.

Keasling, J.D. (2010) Manufacturing molecules through metabolic engineering. *Science*, **330**, 1355–1358.

Kebeish, R., Niessen, M., Thiruveedhi, K., Bari, R., Hirsch, H.-J., Rosenkranz, R., Stabler, N., Schonfeld, B., Kreuzaler, F. and Peterhansel, C. (2007) Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in Arabidopsis thaliana. *Nat. Biotechnol.* **25**, 593–599.

Kelly, A.A., van Erp, H., Quettier, A.-L., Shaw, E., Menard, G., Kurup, S. and Eastmond, P.J. (2013) The SUGAR-DEPENDENT1 lipase limits triacylglycerol accumulation in vegetative tissues of Arabidopsis. *Plant Physiol.* **162**, 1282–1289.

Kempinski, C., Jiang, Z., Bell, S. and Chappell, J. (2015) Metabolic engineering of higher plants and algae for isoprenoid production. *Adv. Biochem. Eng. Biotechnol.* **148**, 161–199.

Khalil, A.S., Lu, T.K., Bashor, C.J., Ramirez, C.L., Pyenson, N.C., Joung, J.K. and Collins, J.J. (2012) A synthetic biology framework for programming eukaryotic transcription functions. *Cell*, **150**, 647–658.

Kim, D.S., Gusti, V., Dery, K.J. and Gaur, R.K. (2008) Ligand-induced sequestering of branchpoint sequence allows conditional control of splicing. *BMC Mol. Biol.* **9**, 23.

Klein-Marcuschamer, D., Oleskowicz-Popiel, P., Simmons, B.A. and Blanch, H.W. (2010) Technoeconomic analysis of biofuels: a wiki-based platform for lignocellulosic biorefineries. *Biomass Bioenergy*, **34**, 1914–1921.

Kubodera, T., Watanabe, M., Yoshiuchi, K., Yamashita, N., Nishimura, A., Nakai, S., Gomi, K. and Hanamoto, H. (2003) Thiamine-regulated gene expression of *Aspergillus oryzae* thiA requires splicing of the intron containing a riboswitch-like domain in the 5′-UTR. *FEBS Lett.* **555**, 516–520.

Li, T., Huang, S., Jiang, W.Z., Wright, D., Spalding, M.H., Weeks, D.P. and Yang, B. (2011) TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.* **39**, 359–372.

Li, J.F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M. and Sheen, J. (2013) Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.* **31**, 688–691.

Liang, Z., Zhang, K., Chen, K. and Gao, C. (2014) Targeted mutagenesis in Zea mays using TALENs and the CRISPR/Cas system. *J. Genet Genomics* **41**, 63–68.

Lin, I.W., Sosso, D., Chen, L.Q. *et al.* (2014a) Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. *Nature*, **508**, 546–549.

Lin, M.T., Occhialini, A., Andralojc, P.J., Parry, M.A.J. and Hanson, M.R. (2014b) A faster RuBisCO with potential to increase photosynthesis in crops. *Nature*, **513**, 547–550.

Liu, W., Mazarei, M., Peng, Y., Fethe, M.H., Rudis, M.R., Lin, J., Millwood, R.J., Arelli, P.R. and Stewart, C.N. (2014) Computational discovery of soybean promoter cis-regulatory elements for the construction of soybean cyst nematode-inducible synthetic promoters. *Plant Biotechnol. J.* **12**, 1015–1026.

Lopez, G. and Anderson, J.C. (2015) Synthetic auxotrophs with ligand-dependent essential genes for a BL21(DE3) biosafety strain. *ACS Synth. Biol.* **4**, 1279–1286.

Loqué, D., Scheller, H.V. and Pauly, M. (2015) Engineering of plant cell walls for enhanced biofuel production. *Curr. Opin. Plant Biol.* **25**, 151–161.

Lowder, L.G., Zhang, D., Baltes, N.J., Paul, J.W. 3rd, Tang, X., Zheng, X., Voytas, D.F., Hsieh, T.F., Zhang, Y. and Qi, Y. (2015) A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. *Plant Physiol.* **169**, 971–985.

Ma, X., Zhang, Q., Zhu, Q. *et al.* (2015) A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol. Plant*, **8**, 1274–1284.

Mao, Y., Zhang, H., Xu, N., Zhang, B., Gou, F. and Zhu, J.K. (2013) Application of the CRISPR-Cas system for efficient genome engineering in plants. *Mol. Plant*, **6**, 2008–2011.

Martin-Ortigosa, S. and Wang, K. (2014) Proteolistics: a biolistic method for intracellular delivery of proteins. *Transgenic Res.* **23**, 743–756.

Martin-Ortigosa, S., Valenstein, J.S., Lin, V.S.Y., Trewyn, B.G. and Wang, K. (2012) Gold functionalized mesoporous silica nanoparticle mediated protein and DNA codelivery to plant cells via the biolistic method. *Adv. Funct. Mater.* **22**, 3576–3582.

Mittelbach, M. and Remschmidt, C. (2004) *Biodiesel: The Comprehensive Handbook*. Graz, Austria: Martin Mittelbach.

Miyagawa, Y., Tamoi, M. and Shigeoka, S. (2001) Overexpression of a cyanobacterial fructose-1,6-/sedoheptulose-1,7-bisphosphatase in tobacco enhances photosynthesis and growth. *Nat. Biotechnol.* **19**, 965–969.

Moore, I., Gälweiler, L., Grosskopf, D., Schell, J. and Palme, K. (1998) A transcription activation system for regulated gene expression in transgenic plants. *Proc. Natl Acad. Sci. USA*, **95**, 376–381.

Moser, B.R. (2008) Influence of blending canola, palm, soybean, and sunflower oil methyl esters on fuel properties of biodiesel. *Energy Fuels*, **22**, 4301–4306.

Napier, J.A., Haslam, R.P., Beaudoin, F. and Cahoon, E.B. (2014) Understanding and manipulating plant lipid composition: metabolic engineering leads the way. *Curr. Opin. Plant Biol.* **19**, 68–75.

Nguyen, H.T., Leelavathi, S. and Reddy, V.S. (2004) Bacteriophage T7 RNA polymerase-directed, inducible and tissue-specific over-expression of foreign genes in transgenic plants. *Plant Biotechnol. J.* **2**, 301–310.

Nihongaki, Y., Yamamoto, S., Kawano, F., Suzuki, H. and Sato, M. (2015) CRISPR-Cas9-based photoactivatable transcription system. *Chem. Biol.* **22**, 169–174.

Nissim, L., Perli, S.D., Fridkin, A., Perez-Pinera, P. and Lu, T.K. (2014) Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells. *Mol. Cell*, **54**, 698–710.

Ort, D.R., Merchant, S.S., Alric, J. *et al.* (2015) Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proc. Natl Acad. Sci. USA*, **112**, 8529–8536.

Padidam, M. (2003) Chemically regulated gene expression in plants. *Curr. Opin. Plant Biol.* **6**, 169–177.

Patron, N.J., Orzaez, D., Marillonnet, S. *et al.* (2015) Standards for plant synthetic biology: a common syntax for exchange of DNA parts. *New Phytol.* **208**, 13–19.

Petersen, P.D., Lau, J., Ebert, B. *et al.* (2012) Engineering of plants with improved properties as biofuels feedstocks by vessel-specific complementation of xylan biosynthesis mutants. *Biotechnol. Biofuels*, **5**, 84.

Pham, A.-T., Lee, J.-D., Shannon, J.G. and Bilyeu, K. (2010) Mutant alleles of FAD2-1A and FAD2-1B combine to produce soybeans with the high oleic acid seed oil trait. *BMC Plant Biol.* **10**, 195.

Philips, I. and Silver, P. (2006) A new BioBrick assembly strategy designed for facile protein engineering. MIT SBWG Tech Rep. (http://hdl.handle.net/1721.1/32535)

Piatek, A., Ali, Z., Baazim, H., Li, L., Abulfaraj, A., Al-Shareef, S., Aouida, M. and Mahfouz, M.M. (2015) RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. *Plant Biotechnol. J.* **13**, 578–589.

Puchta, H. (2016) Using CRISPR/Cas in three dimensions: towards synthetic plant genomes, transcriptomes and epigenomes. *Plant J.* doi: 10.1111/tpj.13100.

Qi, L., Haurwitz, R.E., Shao, W., Doudna, J.A. and Arkin, A.P. (2012) RNA processing enables predictable programming of gene expression. *Nat. Biotechnol.* **30**, 1002–1006.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P. and Lim, W.A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, **152**, 1173–1183.

Ramundo, S., Rahire, M., Schaad, O. and Rochaix, J.D. (2013) Repression of essential chloroplast genes reveals new signalling pathways and regulatory feedback loops in *Chlamydomonas*. *Plant Cell*, **25**, 167–186.

Rautengarten, C., Ebert, B., Moreno, I. *et al.* (2014) The Golgi localized bifunctional UDP-rhamnose/UDP-galactose transporter family of Arabidopsis. *Proc. Natl Acad. Sci. USA*, **111**, 11563–11568.

Rebatchouk, D., Daraselia, N. and Narita, J.O. (1996) NOMAD: a versatile strategy for *in vitro* DNA manipulation applied to promoter analysis and vector design. *Proc. Natl Acad. Sci. USA*, **93**, 10891–10896.

Reddy, M.S.S., Chen, F., Shadle, G., Jackson, L., Aljoe, H. and Dixon, R.A. (2005) Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (*Medicago sativa* L.). *Proc. Natl Acad. Sci. USA*, **102**, 16573–16578.

Ro, D.-K., Paradise, E.M., Ouellet, M. *et al.* (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, **440**, 940–943.

Rodionov, D.A., Vitreschak, A.G., Mironov, A.A. and Gelfand, M.S. (2002) Comparative genomics of thiamin biosynthesis in procaryotes. New genes and regulatory mechanisms. *J. Biol. Chem.* **277**, 48949–48959.

Rossi, M., Bermudez, L. and Carrari, F. (2015) Crop yield: challenges from a metabolic perspective. *Curr. Opin. Plant Biol.* **25**, 79–89.

Sarrion-Perdigones, A., Vazquez-Vilar, M., Palací, J., Castelijns, B., Forment, J., Ziarsolo, P., Blanca, J., Granell, A. and Orzaez, D. (2013) GoldenBraid 2.0: a comprehensive dna assembly framework for plant synthetic biology. *Plant Physiol.* **162**, 1618–1631.

Schaeffer, S.M. and Nakata, P.A. (2015) CRISPR/Cas9-mediated genome editing and gene replacement in plants: transitioning from lab to field. *Plant Sci.* **240**, 130–142.

Schiml, S., Fauser, F. and Puchta, H. (2014) The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. *Plant J.* **80**, 1139–1150.

Serganov, A. and Nudler, E. (2013) A decade of riboswitches. *Cell*, **152**, 17–24.

Shen, B., Sun, X., Zuo, X. *et al.* (2012) Engineering a thermoregulated intein-modified xylanase into maize for consolidated lignocellulosic biomass processing. *Nat. Biotechnol.* **30**, 1131–1136.

Shih, P.M., Zarzycki, J., Niyogi, K.K. and Kerfeld, C.A. (2014) Introduction of a synthetic CO2-fixing photorespiratory bypass into a cyanobacterium. *J. Biol. Chem.* **289**, 9493–9500.

Shukla, V.K., Doyon, Y., Miller, J.C. *et al.* (2009) Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature*, **459**, 437–441.

Slocombe, S.P., Cornah, J., Pinfield-Wells, H., Soady, K., Zhang, Q., Gilday, A., Dyer, J.M. and Graham, I.A. (2009) Oil accumulation in leaves directed by modification of fatty acid breakdown and lipid synthesis pathways. *Plant Biotechnol. J.* **7**, 694–703.

Smith, R.A., Schuetz, M., Roach, M., Mansfield, S.D., Ellis, B. and Samuels, L. (2013) Neighboring parenchyma cells contribute to Arabidopsis xylem lignification, while lignification of interfascicular fibers is cell autonomous. *Plant Cell*, **25**, 3988–3999.

Sosso, D., Luo, D., Li, Q.B. *et al.* (2015) Seed filling in domesticated maize and rice depends on SWEET-mediated hexose transport. *Nat. Genet.* **47**, 1489–1493.

Sudarsan, N., Barrick, J.E. and Breaker, R.R. (2003) Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA*, **9**, 644–647.

Temme, K., Hill, R., Segall-Shapiro, T.H., Moser, F. and Voigt, C.A. (2012) Modular control of multiple pathways using engineered orthogonal T7 polymerases. *Nucleic Acids Res.* **40**, 8773–8781.

Townshend, B., Kennedy, A.B., Xiang, J.S. and Smolke, C.D. (2015) High-throughput cellular RNA device engineering. *Nat. Methods*, **12**, 989–994.

Ulmasov, T., Murfett, J., Hagen, G. and Guilfoyle, T.J. (1997) Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell*, **9**, 1963–1971.

Vanhercke, T., El Tahchy, A., Liu, Q. *et al.* (2014) Metabolic engineering of biomass for high energy density: oilseed-like triacylglycerol yields from plant leaves. *Plant Biotechnol. J.* **12**, 231–239.

Vega-Sanchez, M.E., Loque, D., Lao, J. *et al.* (2015) Engineering temporal accumulation of a low recalcitrance polysaccharide leads to increased C6 sugar content in plant cell walls. *Plant Biotechnol. J.* **13**, 903–914.

van Vleet, J.H. and Jeffries, T.W. (2009) Yeast metabolic engineering for hemicellulosic ethanol production. *Curr. Opin. Biotechnol.* **20**, 300–306.

Wachter, A., Tunc-Ozdemir, M., Grove, B.C., Green, P.J., Shintani, D.K. and Breaker, R.R. (2007) Riboswitch control of gene expression in plants by splicing and alternative 3′ end processing of mRNAs. *Plant Cell*, **19**, 3437–3450.

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C. and Qiu, J.L. (2014) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.* **32**, 947–951.

Weber, E., Engler, C., Gruetzner, R., Werner, S. and Marillonnet, S. (2011) A modular cloning system for standardized assembly of multigene constructs. *PLoS ONE*, **6**, e16765.

Weigand, J.E. and Suess, B. (2007) Tetracycline aptamer-controlled regulation of pre-mRNA splicing in yeast. *Nucleic Acids Res.* **35**, 4179–4185.

Weinthal, D., Tovkach, A., Zeevi, V. and Tzfira, T. (2010) Genome editing in plant cells by zinc finger nucleases. *Trends Plant Sci.* **15**, 308–321.

Winkler, W.C. and Breaker, R.R. (2005) Regulation of bacterial gene expression by riboswitches. *Annu. Rev. Microbiol.* **59**, 487–517.

Wittmann, A. and Suess, B. (2012) Engineered riboswitches: expanding researchers' toolbox with synthetic RNA regulators. *FEBS Lett.* **586**, 2076–2083.

Woo, J.W., Kim, J., Kwon, S.I., Corvalan, C., Cho, S.W., Kim, H., Kim, S.G., Kim, S.T., Choe, S. and Kim, J.S. (2015) DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat. Biotechnol.* **33**, 1162–1164.

Wu, L. and Birch, R.G. (2007) Doubled sugar content in sugarcane plants modified to produce a sucrose isomer. *Plant Biotechnol. J.* **5**, 109–117.

Wu, G., Truksa, M., Datla, N., Vrinten, P., Bauer, J., Zank, T., Cirpus, P., Heinz, E. and Qiu, X. (2005) Stepwise engineering to produce high yields of very long-chain polyunsaturated fatty acids in plants. *Nat. Biotechnol.* **23**, 1013–1017.

Xing, H.L., Dong, L., Wang, Z.P., Zhang, H.Y., Han, C.Y., Liu, B., Wang, X.C. and Chen, Q.J. (2014) A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol.* **14**, 327.

Xu, C. and Shanklin, J. (2016) Triacylglycerol metabolism, function, and accumulation in plant vegetative tissues. *Annu. Rev. Plant Biol.* **67**, 13.1–13.28.

Yau, Y., Easterling, M. and Stewart, N. Jr (2013) Precise transgene stacking in planta through the combined use of TALENs and unidirectional site-specific recombination systems. *OA Biotechnology*, **2**, 24.

Young, E., Lee, S.M. and Alper, H. (2010) Optimizing pentose utilization in yeast: the need for novel tools and approaches. *Biotechnol. Biofuels*, **3**, 24.

Yuan, L. and Grotewold, E. (2016) Metabolic engineering to enhance the value of plants as green factories. *Metab. Eng.* **27**, 83–91.

Zale, J., Jung, J.H., Kim, J.Y. *et al.* (2016) Metabolic engineering of sugarcane to accumulate energy-dense triacylglycerols in vegetative biomass. *Plant Biotechnol. J.* **14**, 661–669.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O. *et al.* (2015) Cpf1 Is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, **163**, 759–771.

Zhang, Y., Zhang, F., Li, X., Baller, J.A., Qi, Y., Starker, C.G., Bogdanove, A.J. and Voytas, D.F. (2013) Transcription activator-like effector nucleases enable efficient plant genome engineering. *Plant Physiol.* **161**, 20–27.

Zhou, J., Lee, C., Zhong, R. and Ye, Z.H. (2009) MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in Arabidopsis. *Plant Cell*, **21**, 248–266.

Zhou, H., Liu, B., Weeks, D.P., Spalding, M.H. and Yang, B. (2014) Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res.* **42**, 10903–10914.