Metabolic Engineering To Enhance The Value of Plants As Green Factories

Ling Yuan¹ and Erich Grotewold^{2*}

¹Department of Plant and Soil Sciences, University of Kentucky, 1401 University Drive, Lexington, KY 40546. ²Center for Applied Plant Sciences (CAPS), Department of Molecular Genetics and Department of Horticulture and Crop Science, 012 Rightmire Hall, 1060 Carmack Rd, The Ohio State University, Columbus, OH 43210

*Corresponding author: grotewold.1@osu.edu

ABSTRACTS

The promise of plants to serve as the green factories of the future is ever increasing. Plants have been used traditionally for construction, energy, food and feed. Bioactive compounds primarily derived from specialized plant metabolism continue to serve as important scaffold molecules for pharmaceutical drug production. Yet, the past few years have witnessed a growing interest on plants as the ultimate harvesters of carbon and energy from the sun, providing carbohydrate and lipid biofuels that would contribute to balancing atmospheric carbon. How can the metabolic output from plants be increased even further, and what are the bottlenecks? Here, we present what we perceive to be the main opportunities and challenges associated with increasing the efficiency of plants as chemical factories. We offer some perspectives on when it makes sense to use plants as production systems because the amount of biomass needed makes any other system unfeasible. However, there are other instances in which plants serve as great sources of biological catalysts, yet are not necessarily the best-suited systems for production. We also present emerging opportunities for manipulating plant genomes to make plant synthetic biology a reality.

Keywords

Plant metabolic engineering; specialized metabolism; synthetic biology; commodity chemicals

1. Introduction

Some of the main challenges humanity will face in the next 50-100 years include insufficient (healthy) food, climate change resulting from increased greenhouse gas accumulation, inadequate power for the needs of a growing and increasingly energy-dependent population, and a need for new medicines to treat existing and emerging diseases. Plants are clearly part of the solution to all these challenges. Terrestrial plants capture 121.7×10^9 metric tons of carbon from the atmosphere each year (Beer et al., 2010) using solar light as the sole energy source. Photosynthesized carbon is then chemically converted into a very large number of chemical compounds. Some of these compounds are among the most abundant on earth; others are found in low quantities in a limited number of plant species. Generally (but not always), highly abundant compounds are derived from primary (central) metabolism and include the aromatic polymer lignin, triacylglycerols (vegetable oils) and the polysaccharides cellulose and starch. Providing just one example of the impact that these plant compounds have on humans, the world's annual production of starch exceeds 75 million tons (Data for 2012, <http://www.zuckerforschung.at/>).

In contrast, most of the less abundant plant compounds derive from specialized metabolism, an aspect of the plant metabolic activity that results in the production of a very large number [estimated to be in excess of 200,000 (Dixon and Strack, 2003)] of compounds which plants use, as sessile organisms, in their interactions with the environment (biotic and abiotic). A large number [20-30% (Raskin et al., 2002)] of pharmaceutical drugs currently in the market were originally obtained from plants, where they often accumulate in very small quantities. For example, 10,000 kg of Pacific yew bark yields less than 1 kg of the potent anti-cancer compound paclitaxel (taxol) (Vidensek et al., 1990).

Thus, while plants have the ability to produce some of the most abundant compounds on Earth (e.g. lignin and cellulose), they also produce other chemicals of high utility for humankind in minute quantities. Strategies directed at engineering the metabolism of plant compounds must take this into consideration, as the approaches, tools and business propositions are very different for both (Figure 1). It is obvious that plants will continue to be the major producers of carbohydrates, vegetable oils and cellulose, and there are

significant opportunities to engineer the respective plant pathways to alter the chemical properties or the levels of such compounds. But as we learn more about how plants synthesize specialized metabolites, the reconstitution of these pathways in microorganisms, either *de novo* or from privileged scaffolds using synthetic biology approaches, is becoming increasingly attractive from both economical and experimental perspectives (Mitchell, 2011; Xu et al., 2012). The promise of synthetic biology for the production of plant-derived specialized metabolites that are expensive and difficult to obtain is exemplified by the production of the antimalarial drug artemisinin in *Saccharomyces cerevisiae* (Paddon and Keasling, 2014; Paddon et al., 2013), and by the high-level production of precursors for the anti-cancer drug taxol in *E. coli* (Ajikumar et al., 2010). Application of synthetic biology in plants is still at an early stage, but as described in subsequent sections, new breakthroughs in genome engineering offer new opportunities. Plant metabolic engineering can also be used to increase plant performance, for example by increasing resistance to pathogens or tolerance to abiotic stress conditions. However, in this review, we will primarily focus on metabolic engineering as a way to produce plant chemical compounds.

2. Understanding plant metabolism essential for predictive engineering

For many years, plant metabolic engineering consisted of modifying a regulatory or biosynthetic step in a pathway with subsequent evaluation of the outcome, with the consequence that the results were often very different from what was expected (Carrari et al., 2003). Predictive metabolic engineering consists of tinkering with metabolism in ways that result in anticipated and desired outcomes (Chappell and Grotewold, 2008), which requires an exceptional knowledge of plant biosynthetic pathways and plant systems biology (Curran and Alper, 2012). It is not our intention to recapitulate here the wealth of knowledge on plant metabolism, since there are some excellent recent reviews on the topic (Anarat-Cappillino and Sattely, 2014; Farre et al., 2014; Shumskaya and Wurtzel, 2013; Vogt, 2010; Ziegler and Facchini, 2008). But there are a few aspects of plant metabolism that have emerged over the past few years which must be considered in any successful metabolic engineering activity, and are therefore worth highlighting.

2.1. Enzymes can be promiscuous

A general assumption has been that enzymes are “specialists”, selected to catalyze just one chemical reaction. While this appears to be true for many catalysts from central metabolism, it is becoming evident that a vast majority of the (specialized) metabolic enzymes in microbes (Nam et al., 2012) and plants (Bar-Even and Tawfik, 2013) are promiscuous (generalists), meaning that they can act on a number of different (often, but not always, related) substrates. This broad substrate specificity of specialized metabolism biosynthetic enzymes provides an outstanding opportunity for evolutionary innovations (Grotewold, 2005; Tawfik, 2010; Weng et al., 2012). The ability of specialized metabolism enzymes to utilize multiple substrates means that many specialized pathways [*e.g.*, phenylpropanoid biosynthesis (Bonawitz and Chapple, 2010)] are structured as complex grids, rather than as straightforward linear pathways (*e.g.*, glycolysis).

2.2. Enzymes involved in specialized metabolism are inefficient

Many of the genes encoding enzymes involved in specialized metabolism appear to have arisen from genes of central metabolism, for example by the process of gene duplication and neofunctionalization (Ober, 2005; Ober, 2010). Interestingly, however, specialized metabolism enzymes commonly possess significantly lower k_{cat} values compared to their corresponding central metabolism counterparts. This could be a consequence of reduced selective pressures, when compared to enzymes of core metabolism, offering significant opportunities for enzyme optimization (Bar-Even and Tawfik, 2013).

2.3. Cluster organization of genes encoding metabolic enzymes

In bacteria, loci belonging to functional groups are often arranged in operons, which are clusters of genes with a related function and under the control of a single regulatory signal that are co-transcribed into a common mRNA. Indeed, it is thought that the organization of genes into co-regulated clusters (or operons) plays a major role in enabling bacteria to quickly adapt to changing environmental conditions, with classical examples such as the *lac* and *trp* operons, that have had a major influence on how we perceive genetic regulatory mechanisms in all biological systems (Lawrence, 2002).

Operons, as described in bacteria, are not found in plants (Osbourn and Field, 2009), but there is increasing evidence for the presence of gene clusters, particularly for the biosynthesis of specialized metabolites [*e.g.*, oat avenacins (Qi et al., 2004); the maize benzoaxazinoid DIMBOA (2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one) (Frey et al., 1997), the *Papaver somniferum* noscapine alkaloid (Winzer et al., 2012), and the Solanaceae steroidal glycoalkaloids (Itkin et al., 2013)]. These clusters are up to several hundred kb in length (Chu et al., 2011), and can harbor from two to more than ten biosynthesis enzymes (Chu et al., 2011; Winzer et al., 2012).

As of April 2014, approximately 100 plant genomes have been sequenced to a reasonable depth, providing an unparalleled opportunity for identifying enzymes that fit the requirements (*e.g.*, specificity, kinetic parameters) of metabolic engineers. The Plant Metabolic Network (PMN, <http://www.plantcyc.org/>) attempts to catalog enzymes and catalyzed reactions from a number of plants, but limitations in the amount of biochemical information available makes the resource of modest utility. This is a great first step towards developing genome-scale models of metabolism. However, for the manipulation of metabolic pathways in plants, significant bottlenecks exist that have slowed down the translation of successful endeavors in model systems, such as *Arabidopsis* or *Nicotiana*, to production plants. Chief among these bottlenecks are difficulties associated with introducing genes into plants, and the limitations to fully assessing the metabolic consequences of gene manipulation at the cellular and subcellular level.

3. Overcoming metabolic flux bottlenecks

Bottlenecks in metabolic pathways are often not limited to a single enzymatic step, as predicted by metabolic control analysis theory (Kacser, 1995), but rather are distributed over multiple enzymes. While it is not uncommon to express multiple enzymes simultaneously [*e.g.*, 20+ for the degradation, uptake and metabolism of alginate (Wargacki et al., 2012)] in bacteria, there are still very significant challenges associated with expressing multiple genes in plants. Below, we briefly revisit the challenges and the emerging solutions.

3.1. Challenges of plant transformation

3.1.1 GMO issues solved with clean technologies

In the US, “transgenic plants” that the US Department of Agriculture (USDA) is authorized to regulate, are those containing transformed genes from plant pests, non-plant species or invasive plants. In other words, USDA oversight is required when field testing or producing plants that are developed through genetic engineering involving donor, recipient or vectors of plant pest or non-plant origins. The transgenic plants now available in the market are usually distinguished by detectable fingerprints of the foreign genes, such as the cauliflower mosaic virus 35S promoter (p35S), *Bacillus thuringiensis* toxin genes (*BT* genes), and border sequences of *Agrobacterium tumefaciens* T-DNA (transfer-DNA) vectors. A comprehensive list of transgenic crops that received regulatory approval as of 2004 is available as part of the Handbook of Plant Biotechnology (Christou and Klee, 2004).

Recently, engineered plants free of these fingerprints have become available through a number of new, “clean” technologies. In most of these new technologies, a genome modification system (present for example in a T-DNA) is first transformed into a crop plant via conventional gene transformation. Once the intended genotypic outcome is achieved (*e.g.*, site-specific mutation or insertion), the integrated T-DNA can be segregated out, thus leaving only the altered gene. These engineered plants lack any transgene fingerprint, because no foreign DNA sequence is left behind in the genome. In other technologies, such as intragenic or cisgenic transformation [for review see (Holme et al., 2013)], all transformed genetic materials are from native plants. Accelerated plant breeding can be achieved through early flowering using the dominant *flowering locus T* gene (*FT*) (Kobayashi et al., 1999). At the end of the breeding process, the *FT* gene T-DNA can be removed by segregation. The resulting plants therefore possess no transgenic fingerprint. Several products developed using the new generation of technologies have received a deregulation green light from USDA as they fall out of the agency’s regulatory authority (Waltz, 2012). These recent technologies that have the potential to significantly change the GMO and breeding landscape are briefly described below.

3.1.1.1. Site-specific genome editing

Different from animals or fungi, plant genome engineering (site-directed DNA sequence modification) has been previously significantly complicated by the very low frequency of homologous recombination during transgene integration. However, several promising approaches have been recently developed for the engineering of plant genomes (Curtin et al., 2012). Precise genome editing with engineered nucleases is a versatile tool for both basic biology and applied biotechnology. Four families of designer nucleases have so far proven to be effective, and include engineered meganuclease, zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and the clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 system (Esvelt and Wang, 2013; Fichtner et al., 2014). The customized molecular ‘DNA scissors’ create specific double-stranded breaks at chosen sites in the genome, and when the endogenous DNA repairing machinery repairs the induced break, it commonly induces mutations, deletions, or insertions at the target site in the genome (Curtin et al., 2011; Jiang et al., 2013a; Shan et al., 2013a; Shan et al., 2013c; Shukla et al., 2009; Sommer et al., 2014).

Designer nucleases have been used successfully to modify plant metabolic pathways in a number of plant species, including *Arabidopsis*, rice, sorghum, maize and barley (Chen and Gao, 2013). Recently, TALENs and CRISPR technologies have become more widely used as tools for site-specific genome modification (Table I). Genes involved in carotenoid biosynthesis and amino acid metabolism, including *phytoene desaturase* and *betaine aldehyde dehydrogenase*, have been knocked-out using TALENs and CRISPR (Shan et al., 2013a; Shan et al., 2013c). CRISPR has been used to target the photosynthetic pathway and gravity signaling pathways by mutating *CHLOROPHYLL A OXYGENASE 1* and *LAZY1*, respectively (Miao et al., 2013). Several phytohormone related genes, including *BRASSINOSTEROID INSENSITIVE 1*, *JASMONATE-ZIM-DOMAIN PROTEIN 1*, *GIBBERELIC ACID INSENSITIVE*, and other genes, have also been disrupted using CRISPR (Feng et al., 2013).

These revolutionary genome modification tools have their unique advantages and disadvantages. Multiplexing or targeting several sites within one genome simultaneously can be achieved by the CRISPR/Cas9 system. By comparison, multiplexing is more difficult to achieve using ZFNs or TALENs. Unlike other described tools, the DNA-binding domains of most meganucleases are not distinctly separated from the catalytic

domains, making meganuclease engineering procedures more complex and error prone (Smith et al., 2006; Szczepek et al., 2007). Relative high cost and complex design are limiting factors for ZFNs and TALENs. Simple design and inexpensive assembly make the CRISPR/Cas9 system uniquely attractive; however, non-specific binding, or off-target hits due to high tolerance for mismatches, are major concerns (Fichtner et al., 2014). The newly developed CRISPR-*FokI* system, based on the similar principle of TALEN-*FokI*, has significantly improved the precision of its genome editing (Guilinger et al., 2014b; Tsai et al., 2014). All of these technology platforms are being improved through bioinformatics, directed evolution technologies and sequence specificity profiling (Fu et al., 2014; Guilinger et al., 2014a).

3.1.1.2. Intragenic or cisgenic transformation

The concept of intragenic transformation describes the process of isolating specific genetic elements from a plant, recombining these elements *in vitro* and inserting the resulting expression cassettes into a plant belonging to the same sexual compatibility group (Rommens, 2004). In cisgenesis transformation, the cisgene is an identical copy of the endogenous gene including the promoter, introns and the terminator in the normal-sense orientation. As successful examples, the native apple scab resistance genes, *HcrVf1* and *HcrVf2*, including their native promoter, coding and terminator sequences, have been isolated and transformed to apple to achieve scab resistance (Joshi et al., 2011). The recently commercialized herbicide-tolerant bluegrass was generated by introducing genetic elements from corn, rice and *Arabidopsis* using biolistic transformation. Such a product falls out of the classic definition of “transgenic plants” as it does not contain genes from plant pests, non-plant species or invasive plants.

3.1.1.3. Genome-wide association mapping

The goal of genome-wide association (GWA) mapping is to dissect quantitative trait variants and to identify replication in different populations. GWA mapping is based on the principle that genes or mutant alleles that control a trait can retain genetic linkage to surrounding sequences from the original trait donor. GWA mapping scans the entire genome for significant associations between a panel of SNPs and a particular phenotype;

such associations are often based on the detection of linkage disequilibrium. Once significant associations are made, their relative contribution and linkage to a quantitative trait locus (QTL) that contributes to the trait of interest need to be independently verified. Therefore, performing GWA mapping often requires extensive knowledge of SNPs within the genome of the species of interest. It is thus not surprising that the most significant work with plants has been with *Arabidopsis* and maize, which both have abundant inbred lines (Atwell et al., 2010; Riedelsheimer et al., 2012). However, with the completion of reference genome sequences, and the ever advancing genomics technologies, other economically valuable plants will be subject to GWA analysis (Huang and Han, 2013). GWA studies can also be useful for metabolic engineering. In a large-scale study, close to 300 field-grown inbred maize lines were genotyped using over 56,000 SNPs and assayed for their chemical compositions in leaves (Riedelsheimer et al., 2012). The association of genetic variants and concentrations of more than 100 metabolites was used to connect specific metabolite concentrations with limited genetic loci and to correlate metabolites with agronomic traits. Focusing on lignin biosynthesis, this study suggests correlations between lignin content, plant height, biomass yield and five metabolites, are strongly associated with a specific region of chromosome 9. Dissecting complex links between genotype and phenotype, through genome-wide metabolic profiling, identifies genetic control points that can be used to guide metabolic engineering.

3.1.1.4. FT for accelerated breeding

Backcrossing is commonly used in cultivar development to transfer one or a few genes to desired genetic backgrounds. The duration necessary to complete such trait conversions is largely dependent upon generation times. Constitutive overexpression of the *Arabidopsis FT* gene induces early flowering in many plants with a dwarf phenotype that produces normal seeds. *FT* and its orthologs exist in all plants. In *FT*-assisted backcrossing, an *FT* transgene insertion and the trait(s) of interest at each backcross generation are selected, except for the last. In the final generation, selection is conducted for the trait(s) of interest but against *FT*, to generate the backcross-derived trait conversion and a plant with normal growth phenotype. When tobacco was used as a

demonstration model, the *FT*-containing plants flowered in an average of 39 days, in comparison with 87–138 days for non-*FT* plants. A seven-generation *FT*-assisted backcrossing is reduced to 604 days from the convention breeding time of 1,150 days (Lewis and Kernodle, 2009). Again, because the last generation of backcrossing is against the *FT* (selected against the dwarf phenotype), the final selection contains no *FT* T-DNA.

3.2. *Enzymes versus transcription factors*

There is a wealth of successful examples of utilizing pathway enzymes and their coding genes for metabolic engineering. Alternatively, the use of transcription factors (TFs) for metabolic engineering is sometimes more effective (Braun et al., 2001; Broun, 2004; Grotewold, 2008; Xie et al., 2006). When a large number of genes reported to be associated with domestication of crop plants (the first metabolic engineering by mankind!) is analyzed, almost 70% of them are transcription regulatory genes while over 20% encode enzymes (Meyer and Purugganan, 2013). This is not surprising because TFs commonly control multiple enzymes within a metabolic pathway, and are thus believed to have profound impacts on phenotypic and metabolic outcomes. In metabolic engineering, the “push” strategy refers to using TF genes or early pathway enzyme genes to push the flux downstream, whereas the “pull” strategy refers to application of late pathway genes to draw the substrates toward a particular final product. The best engineering will likely combine the two strategies (Vanhercke et al., 2014).

3.2.1 Enzymes to bypass feedback regulation

Almost all metabolic pathways are subject to feedback regulation by which the product influences its production, often by controlling activities of pathway enzymes. Negative feedback regulation of enzymes can result from product inhibition or allosteric inhibition, in which the regulatory molecule interacts at an allosteric site other than the active site, to inhibit the enzyme activity. Two approaches have been shown to be effective in overcoming negative feedback regulation. The first approach is to introduce into plants a feedback-insensitive enzyme. Genes encoding such enzymes can sometimes be found in microbes or plants with unique phenotypes. When natural mutant enzymes

are unavailable, engineered feedback-insensitive enzymes can be developed through technologies such as directed evolution [*e.g.*, to remove allosteric feedback inhibition of tomato coumarate:CoA ligase (Alberstein et al., 2012)]. Expression of an engineered bacterial feedback-insensitive enzyme gene of the shikimate pathway in *Arabidopsis* results in an increase in the number of specialized metabolites derived from aromatic amino acids, revealing a complex connection between primary and secondary metabolism which had not previously been observed (Tzin et al., 2012).

3.2.2 Natural or synthetic transcription factors

The successful application of natural TFs in metabolic engineering also inspired the development of synthetic TFs. Synthetic TFs, that regulate gene expression by targeting natural or synthetic promoters in a spatio-temporally tunable fashion, offer an on-command tool for engineering valuable metabolic traits. The first generation of synthetic TFs have been developed as fusion proteins in which either a broad-function trans-activator or repressor domain (Matsui et al., 2004; Mitsuda et al., 2011) is fused with a natural DNA-binding protein. In recent years, the ability to synthesize DNA-binding modules has been significantly advanced by developing *in vivo* screening of tunable artificial TFs with minimal off-target specificity (McIsaac et al., 2013). A light-controlled gene expression switch that can be instantaneously turned on and off by red light, or by far red light-supplemented white light, provides an elegant example of a synthetic transcriptional regulation system with high spatio-temporal resolution (Muller et al., 2014). While most synthetic TF designs are based on protein-DNA interaction (as discussed above), a RNA-guided gene activation system also has been developed (Perez-Pinera et al., 2013). In this system, the inactivated Cas9 protein is fused to a transcriptional activation domain and recruited to the target gene promoter by the guide RNA to activate the gene expression. The same design principle can be applied to synthesize a target promoter-specific repressor that has been demonstrated in plant using TALEN (Mahfouz et al., 2012). CRISPR-interference (CRISPRi) is another reversible repression system in which the inactivated Cas9, targeted to a specific promoter by the guide RNA, interferes with transcription (Qi et al., 2013). Such guide-RNA based

activation/repression systems are simple to construct and potentially capable of targeting multiple promoters simultaneously.

3.3. Impact of post-transcriptional and post-translational modifications

Metabolic flux can be profoundly impacted by control of protein stability of regulators and enzymes post-transcriptionally and/or post-translationally. Post-transcriptional modifications by small RNAs and post-translational modifications by phosphorylation and proteasomal degradation play important roles in fine tuning the activities of TFs (Filtz et al., 2014). Regulated degradation of TFs by the 26S ubiquitin proteasomal system (UPS) is a known control mechanism involved in metabolic pathways of plants. Gene transcripts in the anthocyanin pathway in Arabidopsis appear to be post-transcriptionally affected by microRNAs, such as miRNA156 (Gou et al., 2011) and miRNA828 (Hsieh et al., 2009; Rajagopalan et al., 2006). This metabolic flux control by miRNA suggests the possibility of re-directing flavonoid biosynthesis to targeted tissues using synthetic gene expression vectors.

Loss of UPS function results in the accumulation of anthocyanins, pointing to a probable role of proteasomal degradation in control of anthocyanin biosynthesis (Smalle et al., 2002). Recent studies demonstrate that Arabidopsis and apple TF complexes are subjected to degradation by UPS (Li et al., 2012; Maier et al., 2013; Patra et al., 2013a). In addition to controlling TF protein stability, UPS also regulates the turnover of metabolic enzymes [*e.g.* those in the phenylpropanoid pathway (Zhang et al., 2013)]. These examples suggest strategies by which regulatory or enzyme proteins can be engineered to become degradation resistant. Precise control of the accumulation of such degradation resistant regulator/enzymes in target tissues or developmental stages, using, for example, spatial/temporal-specific or inducible promoters for their gene expression, will allow the controlled production of valuable metabolites in predetermined locations and with specific timing.

One approach to engineer UPS-degradation resistant proteins is the use of “lysine scanning”. UPS-mediated degradation is highly lysine-dependent (lysine being the primary site for polyubiquitin chain decoration). When the likely ubiquitinated lysine residues are mutated, the bHLH TFs, GL3 and EGL3, are resistant to UPS-degradation,

resulting in increased anthocyanin production (Patra et al., 2013b). Similar approaches, in which lysine residues in targeted proteins are scanned by *in vitro* or *in vivo* site-specific mutagenesis to determine their involvement in UPS-degradation, are feasible engineering strategies. Posttranslational modification by phosphorylation of regulatory proteins is often involved in the regulation of metabolic pathways [*e.g.*, phosphorylation of a WRKY TF in phytoalexin biosynthesis (Mao et al., 2011)]. Targeting phosphorylation sites (*e.g.*, serine or threonine residues) of regulator or enzyme proteins is likely to be an effective method to turn these proteins on and off (Kurepa et al., 2014).

4. Metabolite trafficking and sequestration

It is inconceivable that a factory would consider increasing the production of a particular product without first taking into consideration how the increased amount of product will be stored and distributed. To a large extent, plant metabolic engineering has so far managed, with mixed success, to proceed with a very primitive knowledge of the site of localization of most metabolites present in plants, and the trafficking pathways responsible for getting them there (Grotewold, 2004). When plant metabolites accumulate in the wrong subcellular compartment, negative feedback mechanisms often kick-in to prevent further damage to the plant (Goodman et al., 2004). Thus, understanding plant metabolite trafficking and sequestration, and being able to engineer this in concert with the manipulation of the respective metabolic pathway, is very likely to significantly increase the success of plant metabolic engineering efforts.

4.1. Detecting metabolites

An important challenge continues to be establishing in which cells particular metabolites accumulate, and within a cell, in which compartment or compartments. This aspect of metabolism is particularly relevant when multiple cell types are part of a single metabolic pathway. The biosynthesis of alkaloids provides multiple examples of cellular specialization for performing particular enzymatic steps (De Luca and St. Pierre, 2000), indicating a need for routes that permit pathway intermediates to exit one cell and enter another. Within a cell, the biosynthesis of monoterpene indole alkaloids involves various compartments including the chloroplast, the vacuole, the endoplasmic reticulum and the

cytosol (Facchini, 2001; Verma et al., 2012). While cellular/subcellular fractionation followed by metabolite analysis continues to be a valuable tool to determine where metabolites accumulate, a new generation of approaches is rapidly developing. For example, fluorescence resonance energy transfer (FRET)-based nanosensors have been successfully used to measure subcellular levels of a number of metabolites and other small molecules, including maltose, glucose, sucrose, ammonium and nitrate in plants and other organisms (De Michele et al., 2013; Fehr et al., 2002; Fehr et al., 2004; Fehr et al., 2005; Frommer et al., 2009; Ho and Frommer, 2014; Jones et al., 2013). The sensitivity of nanosensors is in the nano- to micromolar range, and they provide significant advantages over other quantitative reporters of small molecule concentration inside cells (Okumoto et al., 2012). The recent deployment of FRET sensors specific to the plant hormone abscisic acid (ABA) provides a good example of the biological questions that can now be investigated (Jones et al., 2014). The intrinsic autofluorescence of plants cells, however, can sometimes become a significant limitation for using fluorescence-based biosensors. Mass spectrometry imaging (MSI) is emerging as a powerful complement to establish the spatial distribution of metabolites between and within cells (Horn and Chapman, 2014; Lee et al., 2010; Matros and Mock, 2013; Muller et al., 2011). Autofluorescent properties of some specialized metabolites have provided a powerful tool to investigate trafficking, as shown for phenylpropanoids controlled by the P1 locus of maize (Grotewold et al., 1998; Lin et al., 2003), for coumarins (Tattini et al., 2014) and for anthocyanins (Gomez et al., 2011; Poustka et al., 2007).

4.2. Moving metabolites across membranes

Membrane fusion has been proposed as a mechanism by which non-polar metabolites can move across membranes. For example, the excretion of terpenoids by the appendix of *Sauromatum guttatum* (voodoo lily) involves the fusion of endoplasmic reticulum pockets with the plasma membrane (Skubatz et al., 1995). More recently, membrane hemifusion between the plastid outer envelope and the membrane of the endoplasmic reticulum has been proposed as a possible mechanism by which non-polar plastid-synthesized metabolites can be modified and transported to other compartments (Mehrshahi et al., 2014).

However, most metabolites cannot freely diffuse across membranes, and the importance of transporters in moving metabolites across membranes has been known for many years (Kunze et al., 2002). Organellar transporters that are involved in the intracellular movement of intermediates of primary metabolism (*e.g.*, phosphoenolpyruvate, pyruvate, malate) have been characterized in *Arabidopsis* and other plant species (Weber and Brautigam, 2013). Sucrose efflux from mesophyll cells is carried out by SWEET transporters (Chen et al., 2012), while phloem loading (necessary for the transport of sucrose from source to sink tissues) is carried out by the SUT1/SUC2 group of transporters (Riesmeier et al., 1994). Multidrug and toxin extrusion (MATE), ATP-binding cassette (ABC), purine uptake permease (PUP) and nitrate and peptide transport (NRT/PTR) proton-dependent transporters have been associated with the movement of specialized metabolites across various cellular membranes (Nour-Eldin and Halkier, 2013). Recently, blockage of import of glucosinolates to the seed of *Arabidopsis* was obtained by mutating two glucosinolate transporters from the NRT/PRT family (Nour-Eldin et al., 2012). This illustrates that transport engineering [*i.e.*, engineering of transport processes (Nour-Eldin and Halkier, 2013)] can be used to eliminate anti-nutritional compounds from specific tissues.

4.3. From site of synthesis to site of accumulation

Largely based on knowledge gained on the transport of flavonoids (Grotewold and Davis, 2008; Zhao and Dixon, 2010), at least two non-exclusive mechanisms appear to be involved in the movement of specialized metabolites from the site of synthesis to the membrane harboring transporters, or directly to the site of accumulation. The first mechanism considers the role of specific proteins (ligandins) that escort metabolites from the final enzymatic step to the corresponding transporter. One group of ligandins corresponds to glutathione *S*-transferases (GSTs), enzymes that have been classically associated with the transfer of glutathione to a diversity of chemical compounds, but that more recently are emerging as candidate escorts of unstable metabolites (Dixon et al., 2011; Dixon et al., 2010; Mueller et al., 2000). The second mechanism involves the endomembrane system and vesicles that sequester, and thus protect, unstable metabolites. Alkaloid vesicles have long been known (Roberts et al., 1983), and recent evidence

suggests that different kinds of flavonoids transiently accumulate in the cytoplasm surrounded by membranes (Gomez et al., 2011; Hsieh and Huang, 2007; Poustka et al., 2007). While engineering GSTs to increase the stability and accumulation levels of intermediates or final metabolic products is currently feasible, modifying the endomembrane system to enhance metabolite transport will require a much deeper understanding of intracellular trafficking pathways.

As discussed above, different plant cells often cooperate in the formation of a particular metabolite. However, the knowledge that, similar to humans (Human Microbiome Project Consortium, 2012), only a fraction of all the cells in the plant are of plant origin (the others being of microbial origin) (Lundberg et al., 2012; Penuelas and Terradas, 2014) raises the question to what extent the metabolism attributed to a plant is solely of plant origin, and does not involve one or multiple enzymatic steps occurring in a microbe associated with the plant. This type of metabolic exchange is very common in microbial communities (Little et al., 2008), and is likely an important reason why only a small fraction ($< 1\%$) of the $10^6 - 10^7$ microbial species proposed to exist (Morris et al., 2012) can be isolated in pure culture (Kaeberlein et al., 2002). It is well established that the interaction of plants with a large diversity of microbes (*e.g.*, bacteria, fungi) is essential for plant fitness. The nodule-forming legume-rhizobia interaction (that results in nitrogen fixation) (Hirsch et al., 2001; Yang et al., 2010) is a well-known example. Plant-microbe interactions occur in the rhizosphere (soil space immediately adjacent to the roots) as well as in the phyllosphere (leaf surface). The rhizosphere is characterized by the presence of a very large number of plant-secreted compounds [root exudates (Walker et al., 2003)], with estimates suggesting that up to 20% of all the carbon fixed by the plant is secreted to the rhizosphere (Walker et al., 2003). Root exudates include sugars, amino acids, organic acids and a vast array of specialized metabolites, including phenylpropanoids and flavonoids. A vast majority ($>80\%$) of all plants participate in mutualistic relationships with mycorrhizal fungi (Bonanomi et al., 2009). Thus, exchange metabolism as part of the interaction of plants with bacteria and fungi offers another untapped opportunity for the engineering of plant metabolic pathways.

5. Synthetic biology for plant metabolic engineering

The knowledge and application of restriction nucleases in the 1970's brought about the new concept of synthetic biology in which genes are rearranged for the purposes of designing useful biological devices, constructing novel metabolic pathways, and elucidating gene functions. In the last decade, key technological advances, including high-throughput and low-cost DNA sequencing, whole genome sequences, genetic network modeling, affordable DNA synthesis, and real-time, quantitative measurement of biological variables, promise a new era of synthetic biology. More and more DNA sequences corresponding to different functional traits have been characterized and synthesized as building blocks that are used to construct biological devices such as genetic circuits, redesigned enzymes, biosensors, synthetic chromosomes, and synthetic cells or cellular organelles. With regard to plant metabolic engineering, genetic building blocks are being moved within a cell from one compartment to the other, relocated from plant cells to microbial cells or vice versa, mostly for the purpose of producing valuable products at large scale. Recent successful examples of how genes can be transferred from one species to another with the goal of engineering metabolism at a large scale are the redesign of lignin in poplar into what has become to be known as Zip-lignin, in which ester linkages are introduced, making it easier to deconstruct (Wilkerson et al., 2014), and the introduction of a faster D-ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) enzyme from cyanobacteria into multicellular plants (Lin et al., 2014).

Plastids, including chloroplasts, function to generate metabolic energy, contain their own genetic system, and replicate by division. The plastid genome has evolutionarily undergone a great size reduction while maintaining the machinery for replication; it is thus a platform for understanding and design of minimum genomes, a major aim of synthetic biology. In addition, plastids are rich in metabolic precursors, making them an ideal site for production of green compounds. To demonstrate feasibility of metabolic engineering in plastids using synthetic multigene operons, the vitamin E biosynthetic pathway, consisting of three key pathway genes, was designed, synthesized, and integrated into the plastid genomes of tobacco and tomato (Lu et al., 2013). Up to 10-fold greater accumulation of tocopherol has been achieved in leafy and non-leafy (fruit) tissues. Pathways that are normally present in the cytosol also can be relocated to plastids, taking advantage of their metabolic sink, by nuclear transformation of synthetic

genes encoding metabolic enzymes that are targeted to plastids (Kumar et al., 2011). The production of sesquiterpenes, normally produced by the cytosolic mevalonic acid pathway, can be increased more than 1,000-fold by expressing the synthetic genes that lead to accumulation of the pathway enzymes in chloroplasts, where the methyl-D-erythritol 4-phosphate pathway supplies significantly more precursors for terpene biosynthesis (Wu et al., 2006).

Synthetic biology enables construction of valuable plant metabolic pathways in fast growing microbes. Reprograming of microbes, using a complete or partial plant metabolic pathway, is exemplified by the production in yeast of the precursors for the antimalarial drug, artemisinin, and the subsequent redesign for the high-level production of the drug using synthetic pathways and semi-chemical synthesis (Paddon et al., 2013; Ro et al., 2006; Westfall et al., 2012). From metabolic engineering perspectives, it is equally exciting to introduce non-plant genes that catalyze unique reactions into plants, resulting in novel modifications of valuable plant compounds. Pharmaceuticals naturally made by plants often require modification to increase or broaden their biological activities. When the bacterial chlorination biosynthetic pathway was introduced into the medicinal plant *Catharanthus roseus*, known for the production of a wide variety of terpenoid indole alkaloids, chlorinated alkaloids were produced in a predictable and regioselective manner (Runguphan et al., 2010).

The creation of genetic circuits can now be achieved *in vivo* using precision genome modification technologies such as TALEN and CRISPR. The most straightforward design strategy involves the targeted inactivation of branches of a specific pathway, aiming to channel the intermediate precursors in a desired direction. While a knockdown approach by RNAi has achieved impressive pathway branch inactivation, it does not generate permanent mutations and the T-DNA cannot be removed from the inserted genome. The new generation of genome editing technology is not only superior in these aspects, it also permits fine-tuned metabolic pathway regulation (de Lange et al., 2014). For example, specific gene promoter sequences can be targeted by CRISPR and/or TALEN (Jiang et al., 2013b). These technologies enable the delivery of synthetic, fine-tunable regulatory elements to the targeted promoters, thereby reconstructing a pathway that is responsive to a pre-determined signal, a feedback loop, or a master transcription

factor. Furthermore, Transcription Activator-Like Effectors (TALE), with repression or activation activity, can be developed to specifically bind to the synthetic regulatory elements, creating another level of “wired” logic control (Blount et al., 2012).

6. Forward looking prospects

Plant metabolic engineering is full of promise, but success cases continue to be rare. Yet, the past few years have witnessed significant progress in the development of analytical methods that permit the sensitive detection and molecular weight identification of plant metabolites in complex mixtures. This is further enhanced by the ability to identify metabolites *in situ*, in individual cells or at the subcellular level by MSI or FRET-based nanosensors. While metabolite structural characterization still relies on NMR techniques that demand larger amounts of pure compounds, the future looks bright in terms of our ability to establish when, where and how much of a particular compound there is. Exponential progress also has been made in the area of genomics, metagenomics and metatranscriptomics, which now permit the straightforward and inexpensive identification of a large number of enzyme gene sequences from either a single or complex organismal mixtures. Synthetic biology approaches that rely on gene synthesis and the rational assembly of promoter-protein sequence cassettes can be used to reconstitute and optimize entire pathways. This has thus far largely been limited to microbes (bacteria, fungi), but the advent of powerful plant genome engineering technologies create opportunities for similar approaches in plants. However, concerns about GMOs persist, despite ample evidence that the benefits they provide are several orders of magnitude larger than the potential risks, and so far largely unwarranted, that have been attributed to them. Nevertheless, industrial applications of plant metabolic engineering are rich in possibilities, including exploiting natural variation in cleverly designed breeding programs.

Acknowledgments

Research in LY’s lab is supported by grants from the US Department of Energy (DE-AR0000204) and the Kentucky Tobacco Research and Development Center. Research in EG’s lab is supported by grants from the National Science Foundation (IOS-1125620 and

MCB-1048847). We thank Drs. Barunava Patra, Sanjay Singh, from the University of Kentucky, and Dr. Katrina Cornish from the Ohio State University for their contribution and critical comments to the manuscript.

References

- Human Microbiome Project Consortium, 2012. A framework for human microbiome research. *Nature*. 486, 215-21.
- Ajikumar, P. K., et al., 2010. Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science*. 330, 70-4.
- Alberstein, M., et al., 2012. Removing allosteric feedback inhibition of tomato 4-coumarate:CoA ligase by directed evolution. *Plant J*. 69, 57-69.
- Anarat-Cappillino, G. I., Sattely, E. S., 2014. The chemical logic of plant natural product biosynthesis. *Current Opinion in Plant Biology*. 19, 51-58.
- Atwell, S., et al., 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature*. 465, 627-31.
- Bar-Even, A., Tawfik, D. S., 2013. Engineering specialized metabolic pathways - is there room for enzyme improvements? *Current Opinion in Biotechnology*. 24, 310-319.
- Beer, C., et al., 2010. Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science*. 329, 834-8.
- Blount, B. A., et al., 2012. Rational diversification of a promoter providing fine-tuned expression and orthogonal regulation for synthetic biology. *PLoS One*. 7, e33279.
- Bonanomi, G., et al., 2009. The role of natural products in plant-micorbe interactions. In: Osbourn, A. E., Lanzotti, V., Eds.), *Plant-derived natural products*. Springer, Dodrecht Heidelberg London New York, pp. 301-320.
- Bonawitz, N. D., Chapple, C., 2010. The genetics of lignin biosynthesis: connecting genotype to phenotype. *Annu Rev Genet*. 44, 337-63.
- Braun, E. L., et al., 2001. Transcription factors and metabolic engineering: Novel applications for ancient tools. *Recent. Adv. Phytochem*. 35, 79-109.
- Broun, P., 2004. Transcription factors as tools for metabolic engineering in plants. *Curr Opin Plant Biol*. 7, 202-9.
- Carrari, F., et al., 2003. Engineering central metabolism in crop species: learning the system. *Metab Eng*. 5, 191-200.
- Chappell, J., Grotewold, E., 2008. Plant biotechnology - Predictive, green and quantitative. *Curr Opin Biotechnol*.
- Chen, K., Gao, C., 2013. Targeted genome modification technologies and their applications in crop improvements. *Plant Cell Reports*. 1-9.
- Chen, L. Q., et al., 2012. Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. *Science*. 335, 207-11.
- Christou, P., Klee, H. Eds.), 2004. *Hanbook of Plant Biotechnology*. John Wiley & Sons Ltd, Chichester, England.
- Chu, H. Y., et al., 2011. From hormones to secondary metabolism: the emergence of metabolic gene clusters in plants. *Plant J*. 66, 66-79.
- Curran, K. A., Alper, H. S., 2012. Expanding the chemical palate of cells by combining systems biology and metabolic engineering. *Metab Eng*. 14, 289-97.

- Curtin, S. J., et al., 2012. Genome Engineering of Crops with Designer Nucleases. *Plant Gen.* 5, 42-50.
- Curtin, S. J., et al., 2011. Targeted Mutagenesis of Duplicated Genes in Soybean with Zinc-Finger Nucleases. *Plant Physiology*. 156, 466-473.
- de Lange, O., et al., 2014. From dead leaf, to new life: TAL effectors as tools for synthetic biology. *Plant J.* In Press.
- De Luca, V., St. Pierre, B., 2000. The cell and developmental biology of alkaloid biosynthesis. *Trends Plant Sci.* 5, 168-173.
- De Michele, R., et al., 2013. Fluorescent sensors reporting the activity of ammonium transceptors in live cells. *Elife*. 2, e00800.
- Dixon, D. P., et al., 2011. The Arabidopsis phi class glutathione transferase AtGSTF2: binding and regulation by biologically active heterocyclic ligands. *Biochem J.* 438, 63-70.
- Dixon, D. P., et al., 2010. Roles for glutathione transferases in plant secondary metabolism. *Phytochemistry*. 71, 338-50.
- Dixon, R. A., Strack, D., 2003. Phytochemistry meets genome analysis, and beyond. *Phytochemistry*. 62, 815-6.
- Esvelt, K. M., Wang, H. H., 2013. Genome-scale engineering for systems and synthetic biology. *Molecular Systems Biology*. 9.
- Facchini, P. J., 2001. Alkaloid biosynthesis in plants: biochemistry, cell biology, molecular regulation, and metabolic engineering applications. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 52, 29-66.
- Farre, G., et al., 2014. Engineering complex metabolic pathways in plants. *Annu Rev Plant Biol.* 65, 187-223.
- Fehr, M., et al., 2002. Visualization of maltose uptake in living yeast cells by fluorescent nanosensors. *Proc Natl Acad Sci U S A.* 99, 9846-51.
- Fehr, M., et al., 2004. Live imaging of glucose homeostasis in nuclei of COS-7 cells. *J Fluoresc.* 14, 603-9.
- Fehr, M., et al., 2005. Evidence for high-capacity bidirectional glucose transport across the endoplasmic reticulum membrane by genetically encoded fluorescence resonance energy transfer nanosensors. *Mol Cell Biol.* 25, 11102-12.
- Feng, Z., et al., 2013. Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res.* 23, 1229-1232.
- Fichtner, F., et al., 2014. Precision genetic modifications: a new era in molecular biology and crop improvement. *Planta*. 1-19.
- Filtz, T. M., et al., 2014. Regulation of transcription factor activity by interconnected post-translational modifications. *Trends Pharmacol Sci.* 35, 76-85.
- Frey, M., et al., 1997. Analysis of a chemical plant defense mechanism in grasses. *Science*. 277, 696-9.
- Frommer, W. B., et al., 2009. Genetically encoded biosensors based on engineered fluorescent proteins. *Chem Soc Rev.* 38, 2833-41.
- Fu, Y., et al., 2014. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotech.* advance online publication.
- Gomez, C., et al., 2011. In vivo grapevine anthocyanin transport involves vesicle-mediated trafficking and the contribution of anthoMATE transporters and GST. *Plant J.* 67, 960-70.

- Goodman, C. D., et al., 2004. A multidrug resistance-associated protein involved in anthocyanin transport in *Zea mays*. *Plant Cell*. 16, 1812-1826.
- Gou, J. Y., et al., 2011. Negative regulation of anthocyanin biosynthesis in *Arabidopsis* by a miR156-targeted SPL transcription factor. *Plant Cell*. 23, 1512-22.
- Grotewold, E., 2004. The challenges of moving chemicals within and out of cells: insights into the transport of plant natural products. *Planta*. 219, 906-909.
- Grotewold, E., 2005. Plant metabolic diversity: A regulatory perspective. *Trends Plant Sci*. 10, 57-62.
- Grotewold, E., 2008. Transcription factors for predictive plant metabolic engineering: are we there yet? *Curr Opin Biotechnol*. 19, 138-44.
- Grotewold, E., et al., 1998. Engineering secondary metabolism in maize cells by ectopic expression of transcription factors. *Plant Cell*. 10, 721-40.
- Grotewold, E., Davis, K., 2008. Trafficking and sequestration of anthocyanins. *Nat Prod Comm*. 3, 1251-1258.
- Guilinger, J. P., et al., 2014a. Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity. *Nat Meth*. advance online publication.
- Guilinger, J. P., et al., 2014b. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotechnol*. 32, 577-82.
- Gurushidze, M., et al., 2014. True-breeding targeted gene knock-out in barley using designer TALE-nuclease in haploid cells. *PLoS One*. 9, e92046.
- Hirsch, A. M., et al., 2001. What makes the rhizobia-legume symbiosis so special? *Plant Physiol*. 127, 1484-92.
- Ho, C. H., Frommer, W. B., 2014. Fluorescent sensors for activity and regulation of the nitrate transceptor CHL1/NRT1.1 and oligopeptide transporters. *Elife*. 3, e01917.
- Holme, I. B., et al., 2013. Intragenesis and cisgenesis as alternatives to transgenic crop development. *Plant Biotechnology Journal*. 11, 395-407.
- Horn, P. J., Chapman, K. D., 2014. Lipidomics in situ: Insights into plant lipid metabolism from high resolution spatial maps of metabolites. *Prog Lipid Res*. 54C, 32-52.
- Hsieh, K., Huang, A. H., 2007. Tapetosomes in *Brassica* tapetum accumulate endoplasmic reticulum-derived flavonoids and alkanes for delivery to the pollen surface. *Plant Cell*. 19, 582-96.
- Hsieh, L. C., et al., 2009. Uncovering small RNA-mediated responses to phosphate deficiency in *Arabidopsis* by deep sequencing. *Plant Physiol*. 151, 2120-32.
- Huang, X., Han, B., 2013. Natural Variations and Genome-Wide Association Studies in Crop Plants. *Annu Rev Plant Biol*. 65, 531-551.
- Itkin, M., et al., 2013. Biosynthesis of Antinutritional Alkaloids in Solanaceous Crops Is Mediated by Clustered Genes. *Science*. In Press.
- Jiang, W., et al., 2013a. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. *Nucleic Acids Research*. 41, e188.
- Jiang, W., et al., 2013b. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. *Nucleic Acids Res*. 41, e188.

- Jones, A. M., et al., 2014. Absciscic acid dynamics in roots detected with genetically encoded FRET sensors. *Elife*. 3, e01741.
- Jones, A. M., et al., 2013. In vivo biochemistry: applications for small molecule biosensors in plant biology. *Curr Opin Plant Biol*. 16, 389-95.
- Joshi, S. G., et al., 2011. Functional analysis and expression profiling of HcrVf1 and HcrVf2 for development of scab resistant cisgenic and intragenic apples. *Plant Molecular Biology*. 75, 579-591.
- Kacser, H., 1995. Recent developments beyond metabolic control analysis. *Biochem Soc Trans*. 23, 387-91.
- Kaeberlein, T., et al., 2002. Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science*. 296, 1127-9.
- Kobayashi, Y., et al., 1999. A pair of related genes with antagonistic roles in mediating flowering signals. *Science*. 286, 1960-2.
- Kumar, S., et al., 2011. Remodeling the isoprenoid pathway in tobacco by expressing the cytoplasmic mevalonate pathway in chloroplasts. *Metab Eng*. 14, 19-28.
- Kunze, R., et al., 2002. Metabolic engineering of plants: the role of membrane transport. *Metab Eng*. 4, 57-66.
- Kurepa, J., et al., 2014. Ectopic expression of the phosphomimic mutant version of Arabidopsis response regulator 1 promotes a constitutive cytokinin response phenotype. *BMC Plant Biol*. 14, 28.
- Lawrence, J. G., 2002. Shared strategies in gene organization among prokaryotes and eukaryotes. *Cell*. 110, 407-13.
- Lee, D. Y., et al., 2010. Mass spectrometry-based metabolomics, analysis of metabolite-protein interactions, and imaging. *Biotechniques*. 49, 557-65.
- Lewis, R. S., Kernodle, S. P., 2009. A method for accelerated trait conversion in plant breeding. *Theor Appl Genet*. 118, 1499-508.
- Li, Y. Y., et al., 2012. MdCOP1 ubiquitin E3 ligases interact with MdMYB1 to regulate light-induced anthocyanin biosynthesis and red fruit coloration in apple. *Plant Physiol*. 160, 1011-22.
- Lin, M. T., et al., 2014. A faster Rubisco with potential to increase photosynthesis in crops. *Nature*. 513, 547-50.
- Lin, Y., et al., 2003. Sub-cellular trafficking of phytochemicals using auto-fluorescent compounds in maize cells. *BMC Plant Biology*. 3, 10.
- Little, A. E., et al., 2008. Rules of engagement: interspecies interactions that regulate microbial communities. *Annu Rev Microbiol*. 62, 375-401.
- Lu, Y., et al., 2013. Efficient metabolic pathway engineering in transgenic tobacco and tomato plastids with synthetic multigene operons. *Proc Natl Acad Sci U S A*. 110, E623-32.
- Lundberg, D. S., et al., 2012. Defining the core Arabidopsis thaliana root microbiome. *Nature*. 488, 86-90.
- Ma, N., et al., 2013. Transcription activator-like effector nuclease (TALEN)-mediated gene correction in integration-free beta-thalassemia induced pluripotent stem cells. *J Biol Chem*. 288, 34671-9.
- Mahfouz, M. M., et al., 2012. Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein. *Plant Mol Biol*. 78, 311-21.

- Maier, A., et al., 2013. Light and the E3 ubiquitin ligase COP1/SPA control the protein stability of the MYB transcription factors PAP1 and PAP2 involved in anthocyanin accumulation in Arabidopsis. *Plant J.* 74, 638-51.
- Mao, G., et al., 2011. Phosphorylation of a WRKY transcription factor by two pathogen-responsive MAPKs drives phytoalexin biosynthesis in Arabidopsis. *Plant Cell.* 23, 1639-53.
- Matros, A., Mock, H. P., 2013. Mass spectrometry based imaging techniques for spatially resolved analysis of molecules. *Front Plant Sci.* 4, 89.
- Matsui, K., et al., 2004. Suppression of the biosynthesis of proanthocyanidin in Arabidopsis by a chimeric PAP1 repressor. *Plant Biotechnol J.* 2, 487-93.
- McIsaac, R. S., et al., 2013. Synthetic gene expression perturbation systems with rapid, tunable, single-gene specificity in yeast. *Nucleic Acids Res.* 41, e57.
- Mehrshahi, P., et al., 2014. Redefining the metabolic continuity of chloroplasts and ER. *Trends Plant Sci.*
- Meyer, R. S., Purugganan, M. D., 2013. Evolution of crop species: genetics of domestication and diversification. *Nature Reviews Genetics.* 14, 840-852.
- Miao, J., et al., 2013. Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res.* 23, 1233-1236.
- Mitchell, W., 2011. Natural products from synthetic biology. *Curr Opin Chem Biol.* 15, 505-15.
- Mitsuda, N., et al., 2011. CRES-T, an effective gene silencing system utilizing chimeric repressors. *Methods Mol Biol.* 754, 87-105.
- Morris, J. J., et al., 2012. The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. *MBio.* 3.
- Mueller, L. A., et al., 2000. AN9, a petunia glutathione S-transferase required for anthocyanin sequestration, is a flavonoid-binding protein. *Plant Physiol.* 123, 1561-1570.
- Muller, K., et al., 2014. A red light-controlled synthetic gene expression switch for plant systems. *Mol Biosyst.*
- Muller, T., et al., 2011. Direct plant tissue analysis and imprint imaging by desorption electrospray ionization mass spectrometry. *Anal Chem.* 83, 5754-61.
- Nam, H., et al., 2012. Network context and selection in the evolution to enzyme specificity. *Science.* 337, 1101-4.
- Nour-Eldin, H. H., et al., 2012. NRT/PTR transporters are essential for translocation of glucosinolate defence compounds to seeds. *Nature.* 488, 531-4.
- Nour-Eldin, H. H., Halkier, B. A., 2013. The emerging field of transport engineering of plant specialized metabolites. *Curr Opin Biotechnol.* 24, 263-70.
- Ober, D., 2005. Seeing double: gene duplication and diversification in plant secondary metabolism. *Trends Plant Sci.* 10, 444-9.
- Ober, D., 2010. Gene duplications and the time thereafter - examples from plant secondary metabolism. *Plant Biol (Stuttg).* 12, 570-7.
- Okumoto, S., et al., 2012. Quantitative imaging with fluorescent biosensors. *Annu Rev Plant Biol.* 63, 663-706.
- Osborn, M. J., et al., 2013. TALEN-based gene correction for epidermolysis bullosa. *Mol Ther.* 21, 1151-9.
- Osbourn, A. E., Field, B., 2009. Operons. *Cell Mol Life Sci.* 66, 3755-3775.

- Paddon, C. J., Keasling, J. D., 2014. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat Rev Microbiol.*
- Paddon, C. J., et al., 2013. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature.* 496, 528-32.
- Patra, B., et al., 2013a. Proteolytic degradation of the flavonoid regulators, TRANSPARENT TESTA8 and TRANSPARENT TESTA GLABRA1, in Arabidopsis is mediated by the ubiquitin/26Sproteasome system. *Plant Signal Behav.* 8, doi: 10.4161/psb.25901.
- Patra, B., et al., 2013b. Ubiquitin protein ligase 3 mediates the proteasomal degradation of GLABROUS 3 and ENHANCER OF GLABROUS 3, regulators of trichome development and flavonoid biosynthesis in Arabidopsis. *Plant J.* 74, 435-47.
- Penuelas, J., Terradas, J., 2014. The foliar microbiome. *Trends Plant Sci.*
- Perez-Pinera, P., et al., 2013. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods.* 10, 973-6.
- Poustka, F., et al., 2007. A trafficking pathway for anthocyanins overlaps with the endoplasmic reticulum-to-vacuole protein sorting route in Arabidopsis and contributes to the formation of vacuolar inclusions. *Plant Physiol.* 145, 1323-1335.
- Qi, L. S., et al., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell.* 152, 1173-83.
- Qi, X., et al., 2004. A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. *Proc Natl Acad Sci U S A.* 101, 8233-8.
- Rajagopalan, R., et al., 2006. A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. *Genes Dev.* 20, 3407-25.
- Raskin, I., et al., 2002. Plants and human health in the twenty-first century. *Trends Biotechnol.* 20, 522-31.
- Riedelsheimer, C., et al., 2012. Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proc Natl Acad Sci U S A.* 109, 8872-7.
- Riesmeier, J. W., et al., 1994. Evidence for an essential role of the sucrose transporter in phloem loading and assimilate partitioning. *Embo J.* 13, 1-7.
- Ro, D. K., et al., 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature.* 440, 940-3.
- Roberts, M. F., et al., 1983. Localization of enzymes and alkaloidal metabolites in *Papaver latex*. *Arch. Biochem. Biophys.* 222, 599-609.
- Rommens, C. M., 2004. All-native DNA transformation: a new approach to plant genetic engineering. *Trends Plant Sci.* 9, 457-64.
- Rungtaphan, W., et al., 2010. Integrating carbon-halogen bond formation into medicinal plant metabolism. *Nature.* 468, 461-4.
- Schwank, G., et al., 2013. Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell.* 13, 653-8.
- Shan, Q., et al., 2013a. Rapid and Efficient Gene Modification in Rice and Brachypodium Using TALENs. *Molecular Plant.* 6, 1365-1368.
- Shan, Q., et al., 2013b. Rapid and efficient gene modification in rice and Brachypodium using TALENs. *Mol Plant.* 6, 1365-8.

- Shan, Q., et al., 2013c. Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotech.* 31, 686-688.
- Shukla, V. K., et al., 2009. Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature.* 459, 437-441.
- Shumskaya, M., Wurtzel, E. T., 2013. The carotenoid biosynthetic pathway: thinking in all dimensions. *Plant Sci.* 208, 58-63.
- Skubatz, H., et al., 1995. Pathway of terpene excretion by the appendix of *Sauromatum guttatum*. *Proc. Natl. Acad. Sci. USA.* 92, 10084-10088.
- Smalle, J., et al., 2002. Cytokinin growth responses in *Arabidopsis* involve the 26S proteasome subunit RPN12. *Plant Cell.* 14, 17-32.
- Smith, J., et al., 2006. A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Research.* 34, e149.
- Sommer, D., et al., 2014. Efficient genome engineering by targeted homologous recombination in mouse embryos using transcription activator-like effector nucleases. *Nat Commun.* 5.
- Szczepek, M., et al., 2007. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotech.* 25, 786-793.
- Tattini, M., et al., 2014. Esculetin and esculin (esculetin 6-O-glucoside) occur as inclusions and are differentially distributed in the vacuole of palisade cells in *Fraxinus ornus* leaves: A fluorescence microscopy analysis. *J Photochem Photobiol B.* 140C, 28-35.
- Tawfik, D. S., 2010. Messy biology and the origins of evolutionary innovations. *Nat Chem Biol.* 6, 692-6.
- Tsai, S. Q., et al., 2014. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol.*
- Tzin, V., et al., 2012. Expression of a bacterial feedback-insensitive 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase of the shikimate pathway in *Arabidopsis* elucidates potential metabolic bottlenecks between primary and secondary metabolism. *New Phytol.* 194, 430-9.
- Vanhercke, T., et al., 2014. Metabolic engineering of biomass for high energy density: oilseed-like triacylglycerol yields from plant leaves. *Plant Biotechnology Journal.* 12, 231-239.
- Verma, P., et al., 2012. Emerging trends in research on spatial and temporal organization of terpenoid indole alkaloid pathway in *Catharanthus roseus*: a literature update. *Protoplasma.* 249, 255-68.
- Vidensek, N., et al., 1990. Taxol Content in Bark, Wood, Root, Leaf, Twig, and Seedling from Several *Taxus* Species. *Journal of Natural Products.* 53, 1609-1610.
- Vogt, T., 2010. Phenylpropanoid biosynthesis. *Mol Plant.* 3, 2-20.
- Walker, T. S., et al., 2003. Root exudation and rhizosphere biology. *Plant Physiol.* 132, 44-51.
- Waltz, E., 2012. Tiptoeing around transgenics. *Nature Biotechnology.* 30, 215-217.
- Wang, Y., et al., 2014. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat Biotechnol.* 32, 947-51.
- Wargacki, A. J., et al., 2012. An engineered microbial platform for direct biofuel production from brown macroalgae. *Science.* 335, 308-13.

- Weber, A. P., Brautigam, A., 2013. The role of membrane transport in metabolic engineering of plant primary metabolism. *Curr Opin Biotechnol.* 24, 256-62.
- Weng, J. K., et al., 2012. The rise of chemodiversity in plants. *Science.* 336, 1667-70.
- Westfall, P. J., et al., 2012. Production of amorphadiene in yeast, and its conversion to dihydroartemisinin acid, precursor to the antimalarial agent artemisinin. *Proc Natl Acad Sci U S A.* 109, E111-8.
- Wilkerson, C. G., et al., 2014. Monolignol ferulate transferase introduces chemically labile linkages into the lignin backbone. *Science.* 344, 90-3.
- Winzer, T., et al., 2012. A *Papaver somniferum* 10-gene cluster for synthesis of the anticancer alkaloid noscapine. *Science.* 336, 1704-8.
- Wu, S., et al., 2006. Redirection of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. *Nat Biotechnol.* 24, 1441-7.
- Xie, D. Y., et al., 2006. Metabolic engineering of proanthocyanidins through co-expression of anthocyanidin reductase and the PAP1 MYB transcription factor. *Plant J.* 45, 895-907.
- Xie, F., et al., 2014. Seamless gene correction of beta-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. *Genome Res.* 24, 1526-33.
- Xu, P., et al., 2012. Engineering plant metabolism into microbes: from systems biology to synthetic biology. *Curr Opin Biotechnol.* 24, 291-9.
- Yang, S., et al., 2010. R gene-controlled host specificity in the legume-rhizobia symbiosis. *Proc Natl Acad Sci U S A.* 107, 18735-40.
- Zhang, X., et al., 2013. Arabidopsis Kelch repeat F-box proteins regulate phenylpropanoid biosynthesis via controlling the turnover of phenylalanine ammonia-lyase. *Plant Cell.* 25, 4994-5010.
- Zhao, J., Dixon, R. A., 2010. The 'ins' and 'outs' of flavonoid transport. *Trends Plant Sci.* 15, 72-80.
- Zheng, Q., et al., 2014. Precise gene deletion and replacement using the CRISPR/Cas9 system in human cells. *Biotechniques.* 57, 115-24.
- Ziegler, J., Facchini, P. J., 2008. Alkaloid biosynthesis: metabolism and trafficking. *Annu Rev Plant Biol.* 59, 735-69.

Figure Legends

Figure 1: Outline of the steps and challenges associated with rational metabolic engineering. Bottlenecks are shaded red; the intensity of the color reflects to what extent each aspect is a bottleneck.

Tables

	TALEN	CRISPR/Cas9	CRISPR/<i>FokI</i>*
Target DNA binding mode	Protein-DNA	RNA-DNA	RNA-DNA
Design and construction	Moderate	Simple	Simple
Cost	Moderate	Low	Low
Editing applications	1. Gene knockout ^{a,b,c} 2. Point mutation ^{b,c} 3. Site-directed mutagenesis ^{g,h} 4. Gene/promoter addition and replacement ^c 5. Simultaneous editing of multiple genes ^c (less straightforward)	1. Gene knockout ^c 2. Point mutation ^c 3. Site-directed mutagenesis ^{i,j} 4. Gene/promoter addition and replacement ^d 5. Simultaneous editing of multiple genes ^c	1. Gene knockout ^{e,f} 2. Point mutation 3. Site-directed mutagenesis 4. Gene/promoter addition and replacement 5. Simultaneous editing of multiple genes (less straightforward)
Editing efficiency	High	High	High
Off-target rate	Low	High	Low
Remarks	High editing specificity is resulted from the dual DNA-binding protein domains flanking the target site.	Lower editing specificity is resulted from the single, short binding guide-RNA sequence that tolerates a small number of mismatch	High editing specificity is resulted from the dual DNA-binding guide-RNA flanking the target site.

Table I: Characteristics of various genome-editing strategies. ^a(Gurushidze et al., 2014), ^b(Shan et al., 2013b), ^c(Wang et al., 2014), ^d(Zheng et al., 2014), ^e(Tsai et al., 2014), ^f(Guilinger et al., 2014b), ^g(Osborn et al., 2013), ^h(Ma et al., 2013), ⁱ(Xie et al., 2014), ^j(Schwank et al., 2013). *No published demonstration yet of this strategy in a plant system.

Figure