Recent Advances and Current Challenges in Synthetic Biology of the Plastid Genetic System and Metabolism

Christian R. Boehm and Ralph Bock

Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany

ORCID IDs: 0000-0002-6633-7998 (C.R.B.); 0000-0001-7502-6940 (R.B.)

Building on recombinant DNA technology, leaps in synthesis, assembly, and analysis of DNA have revolutionized genetics and molecular biology over the past two decades (Kosuri and Church, 2014). These technological advances have accelerated the emergence of synthetic biology as a new discipline (Cameron et al., 2014). Synthetic biology is characterized by efforts targeted at the modification of existing and the design of novel biological systems based on principles adopted from information technology and engineering (Andrianantoandro et al., 2006; Khalil and Collins, 2010). As in more traditional engineering disciplines such as mechanical, electrical and civil engineering, synthetic biologists utilize abstraction, decoupling and standardization to make the design of biological systems more efficient and scalable. To facilitate the management of complexity, synthetic biology relies on an abstraction hierarchy composed of multiple levels (Endy, 2005): DNA as genetic material, “parts” as elements of DNA encoding basic biological functions (e.g. promoter, ribosome-binding site, terminator sequence), “devices” as any combination of parts implementing a human-defined function, and “systems” as any combination of devices fulfilling a predefined purpose. Parts are designated to perform predictable and modular functions in the context of higher-level devices or systems, which are successively refined through a cycle of designing, building, and testing.

Within the past two decades, the synthetic biology approach has produced several notable successes, especially in microbial systems. These include, for example, the design of a minimal bacterial genome (Hutchinson et al., 2016) and a highly modified yeast genome (Richardson et al., 2017), as well as the metabolic engineering of yeast for the biosynthesis of the antimalarial drug precursor artemisinic acid (Ro et al., 2006) and the opioid compounds thebaine and hydrocodone (Galanie et al., 2015). Compared to synthetic biology in bacteria and yeast, synthetic biology in algae and plants is still lagging behind. While the potential of photoautotrophic organisms for environmentally sustainable bioproduction has long been recognized (Georgianna and Mayfield, 2012; Fesenko and Edwards, 2014; Liu and Stewart, 2015; Boehm et al., 2017), their relatively slow growth, scarcely available tools for genetic manipulation, and the physiological as well as genomic complexity of plant systems have delayed their widespread adoption as synthetic biology chassis. However, especially the small genome of the plastid (chloroplast) represents a highly promising platform for engineering the sophisticated metabolism and physiology of the eukaryotic cell it is embedded in (Fig. 1).

The chloroplast originated through the endosymbiotic uptake of a cyanobacterium by a heterotrophic eukaryote more than a billion years ago (Palmer, 2003). Following this event, the endosymbiont evolved mechanisms for facilitated exchange of metabolites with the host cell, underwent radical streamlining of its genome (by gene loss and large-scale transfer of genes to the host nuclear genome) and established an import machinery for the uptake of nucleus-encoded proteins. The resulting organelle serves as the major biosynthetic compartment in photoautotrophic organisms, and has been exploited as a platform for metabolic engineering and molecular farming since the successful development of transformation technologies in the late 1980s (Boynton et al., 1988; Svab et al., 1990). Compared to nuclear genetic engineering, plastid transformation offers several notable advantages relevant to plant biotechnology. These include (1) the

**ADVANCES**

- The function of most plastid genes is known.
- Plastid transformation vectors can be cloned using a modular toolbox.
- Inducible plastid gene expression has been demonstrated in the absence of nuclear transgenes.
- Heterologous metabolic pathways composed of up to nine enzymes have been implemented in plastids.
- Synthetic plastomes can be designed, assembled and amplified.
- Engineered plastomes can be introduced into non-transformable species using horizontal genome transfer.
high precision of genetic engineering enabled by efficient homologous recombination, (2) the possibility of transgene stacking in synthetic operons, (3) the potential for high-level expression of gene products, (4) the absence of epigenetic transgene silencing, and (5) the reduced risk of unwanted transgene transmission due to maternal inheritance of plastid DNA (Bock, 2015).

In this article, we provide an update on tools and technologies available for extending the synthetic biology approach to plastids and highlight key challenges to be addressed through future research. Guided by an abstraction hierarchy of biological design, we identify a scarcity of well-characterized genetic parts, tightly controlled expression devices, and quantitative knowledge of plastid gene expression as current key limitations to plastid synthetic biology. We highlight recent technological developments narrowing the existing complexity gap between bacterial and plastid synthetic biology and provide an outlook to the implementation of complex systems such as synthetic metabolic feedback loops, designer subcompartments and tailor-made genomes in chloroplasts.

**Parts**

The Registry of Standard Biological Parts (http://parts.igem.org) currently contains over 20,000 genetic elements which can be requested by researchers for use in synthetic biology applications. From this collection, approximately 100 parts each have been designed for use in the unicellular green alga *Chlamydomonas reinhardtii* and in multicellular plants (e.g. the seed plants *Nicotiana tabacum* and *Arabidopsis thaliana*, the moss *Physcomitrella patens* and the liverwort *Marchantia polymorpha*). The majority of these parts are designated for nuclear engineering, with only about two dozen suitable for gene expression from the chloroplast genome. One explanation for the relative paucity of plastid genetic elements in the Registry of Standard Biological Parts lies in the half-year timeframe of projects pursued as part of the international Genetically Engineered Machine (iGEM) competition (Smolke, 2009), which is barely compatible with the generation and characterization of stable plastid-engineered (transplastomic) organisms. Beyond iGEM, the repertoire of regulatory sequences routinely used for transgene expression in plastids has remained similarly small: it is comprised of a few preferred promoters (e.g. from the plastid rRNA operon, *PrRN*; the gene for the large subunit of Rubisco, *PrbcL*; and the gene for the D1 protein of photosystem II, *PpsbA*) and a handful of 5′- and 3′-UTRs (Jin and Daniell, 2015). In addition, the bacterial hybrid promoter *Ptrc* (Newell et al., 2003) and several bacteriophage-derived expression elements (McBride et al.,...
et al., 2013), strong terminators (Chen et al., 2013) and part activity, standardized flanking sequences (Mutalik et al., 2016). To reduce the context dependence of activities (Canton et al., 2008; Kelly et al., 2009; Rudge et al., 2016). Consequently, part behavior is, by default, poorly insulated from its specific genetic context: both upstream promoters and downstream antisense promoters may significantly affect the expression level of a target gene (Quesada-Vargas et al., 2005; Sharwood et al., 2011). However, some sequences such as the endogenous tRNA genes trnS and trnH (Stem and Gruissem, 1987) or the heterologous E. coli Thr attenuator (thr; Chen and Orozco, 1988) have been shown to terminate plastid transcription with at least 85% efficiency. Use of insulators based on these parts or new synthetic terminators can potentially enhance the robustness of gene expression levels generated by plastid synthetic biology devices.

Third, plastid transgene expression has been shown to be primarily determined by posttranscriptional control and protein stability rather than by the accumulation of mRNA (Eberhard et al., 2002; Birch-Machin et al., 2002; Rott et al., 1996; Legen et al., 2002; Shi et al., 2016). Consequently, part behavior is, by default, poorly insulated from its specific genetic context: both upstream promoters and downstream antisense promoters may significantly affect the expression level of a target gene (Quesada-Vargas et al., 2005; Sharwood et al., 2011). However, some sequences such as the endogenous tRNA genes trnS and trnH (Stem and Gruissem, 1987) or the heterologous E. coli Thr attenuator (thr; Chen and Orozco, 1988) have been shown to terminate plastid transcription with at least 85% efficiency. Use of insulators based on these parts or new synthetic terminators can potentially enhance the robustness of gene expression levels generated by plastid synthetic biology devices.

Second, the plastome exhibits abundant read-through transcription due to inefficient termination (Stern and Gruissem, 1987; Rott et al., 1996; Legen et al., 2002; Shi et al., 2016). Consequently, part behavior is, by default, poorly insulated from its specific genetic context: both upstream promoters and downstream antisense promoters may significantly affect the expression level of a target gene (Quesada-Vargas et al., 2005; Sharwood et al., 2011). However, some sequences such as the endogenous tRNA genes trnS and trnH (Stem and Gruissem, 1987) or the heterologous E. coli Thr attenuator (thr; Chen and Orozco, 1988) have been shown to terminate plastid transcription with at least 85% efficiency. Use of insulators based on these parts or new synthetic terminators can potentially enhance the robustness of gene expression levels generated by plastid synthetic biology devices.

Third, plastid transgene expression has been shown to be primarily determined by posttranscriptional control and protein stability rather than by the accumulation of mRNA (Eberhard et al., 2002; Birch-Machin et al., 2002; Rott et al., 1996; Legen et al., 2002; Shi et al., 2016). Consequently, part behavior is, by default, poorly insulated from its specific genetic context: both upstream promoters and downstream antisense promoters may significantly affect the expression level of a target gene (Quesada-Vargas et al., 2005; Sharwood et al., 2011). However, some sequences such as the endogenous tRNA genes trnS and trnH (Stem and Gruissem, 1987) or the heterologous E. coli Thr attenuator (thr; Chen and Orozco, 1988) have been shown to terminate plastid transcription with at least 85% efficiency. Use of insulators based on these parts or new synthetic terminators can potentially enhance the robustness of gene expression levels generated by plastid synthetic biology devices.

Third, plastid transgene expression has been shown to be primarily determined by posttranscriptional control and protein stability rather than by the accumulation of mRNA (Eberhard et al., 2002; Birch-Machin et al., 2002; Rott et al., 1996; Legen et al., 2002; Shi et al., 2016). Consequently, part behavior is, by default, poorly insulated from its specific genetic context: both upstream promoters and downstream antisense promoters may significantly affect the expression level of a target gene (Quesada-Vargas et al., 2005; Sharwood et al., 2011). However, some sequences such as the endogenous tRNA genes trnS and trnH (Stem and Gruissem, 1987) or the heterologous E. coli Thr attenuator (thr; Chen and Orozco, 1988) have been shown to terminate plastid transcription with at least 85% efficiency. Use of insulators based on these parts or new synthetic terminators can potentially enhance the robustness of gene expression levels generated by plastid synthetic biology devices.

Third, plastid transgene expression has been shown to be primarily determined by posttranscriptional control and protein stability rather than by the accumulation of mRNA (Eberhard et al., 2002; Birch-Machin et al., 2002; Rott et al., 1996; Legen et al., 2002; Shi et al., 2016). Consequently, part behavior is, by default, poorly insulated from its specific genetic context: both upstream promoters and downstream antisense promoters may significantly affect the expression level of a target gene (Quesada-Vargas et al., 2005; Sharwood et al., 2011). However, some sequences such as the endogenous tRNA genes trnS and trnH (Stem and Gruissem, 1987) or the heterologous E. coli Thr attenuator (thr; Chen and Orozco, 1988) have been shown to terminate plastid transcription with at least 85% efficiency. Use of insulators based on these parts or new synthetic terminators can potentially enhance the robustness of gene expression levels generated by plastid synthetic biology devices.

Third, plastid transgene expression has been shown to be primarily determined by posttranscriptional control and protein stability rather than by the accumulation of mRNA (Eberhard et al., 2002; Birch-Machin et al., 2002; Rott et al., 1996; Legen et al., 2002; Shi et al., 2016). Consequently, part behavior is, by default, poorly insulated from its specific genetic context: both upstream promoters and downstream antisense promoters may significantly affect the expression level of a target gene (Quesada-Vargas et al., 2005; Sharwood et al., 2011). However, some sequences such as the endogenous tRNA genes trnS and trnH (Stem and Gruissem, 1987) or the heterologous E. coli Thr attenuator (thr; Chen and Orozco, 1988) have been shown to terminate plastid transcription with at least 85% efficiency. Use of insulators based on these parts or new synthetic terminators can potentially enhance the robustness of gene expression levels generated by plastid synthetic biology devices.

Third, plastid transgene expression has been shown to be primarily determined by posttranscriptional control and protein stability rather than by the accumulation of mRNA (Eberhard et al., 2002; Birch-Machin et al., 2002; Rott et al., 1996; Legen et al., 2002; Shi et al., 2016). Consequently, part behavior is, by default, poorly insulated from its specific genetic context: both upstream promoters and downstream antisense promoters may significantly affect the expression level of a target gene (Quesada-Vargas et al., 2005; Sharwood et al., 2011). However, some sequences such as the endogenous tRNA genes trnS and trnH (Stem and Gruissem, 1987) or the heterologous E. coli Thr attenuator (thr; Chen and Orozco, 1988) have been shown to terminate plastid transcription with at least 85% efficiency. Use of insulators based on these parts or new synthetic terminators can potentially enhance the robustness of gene expression levels generated by plastid synthetic biology devices.
Plant Physiol. Vol. 179, 2019

Genetic Circuits

Genetic circuits mimic logical functions commonly found in their electronic counterparts. A genetic circuit can be used to control the activity of other devices (such as the gene expression devices or metabolic devices discussed above) in response to external stimuli. A wide range of genetic circuits implementing Boolean logic functions such as yes, not, and, or, nand, nor, xor and n-imply has been reported for bacteria, yeast and mammalian cells (Miyamoto et al., 2013). In plastids, only the simplest logic function yes has been implemented in the form of chemically inducible transgene expression.

Chloroplast transcription is natively controlled by two different types of RNA polymerases in seed plants. The nucleus-encoded RNA polymerase (NEP) is a chloroplast-targeted bacteriophage-type single subunit enzyme, while the plastid-encoded RNA polymerase (PEP) is a eubacteria-type multisubunit enzyme (Barkan, 2011; Börner et al., 2015). The promoter specificity of PEP is modulated by nucleus-encoded and plastid-targeted sigma factors in response to light, hormones and biotic as well as abiotic stresses. However, due to their important role in plant growth, development and survival (and the pervasive transcription of essentially all plastid genes), NEP and PEP are poorly suited as stringent controllers of synthetic genetic circuits in plastids.

As an alternative to transgene control by the endogenous transcription machineries, plastid transgene expression has been controlled through nucleus-encoded and plastid-targeted bacteriophage RNA polymerases or processing factors that are responsive to chemical inducers such as salicylic acid (Magee et al., 2004), ethanol (Lössl et al., 2005), copper (Surzycki et al., 2007) or theophylline (Craig et al., 2008). While heterologous pathways composed of 20 genes or more have been expressed in bacteria and yeast (Temme et al., 2012; Galanée et al., 2015; Li et al., 2018), no more than seven transgenes have to date been simultaneously expressed from the plastome (Krichevsky et al., 2010). The complexity of plastid-based metabolic devices has primarily been limited by a scarcity of available expression signals (see Gene Expression Devices) rather than by the physical size of the introduced DNA (Adachi et al., 2007). Recently, the complexity and number of pathway variants accessible to experimental interrogation has been expanded through combinatorial supertransformation of transplastomic recipient lines (COST-REL). Using this approach, an up to 77-fold increase in artemisinic acid production has been demonstrated in transplastomic tobacco plants combinatorially supertransformed by five additional nuclear transgenes (Fuentes et al., 2016). There is no in-principle limitation to the number of transgenes that can be simultaneously introduced into the plant nucleus using combinatorial transformation (Naqvi et al., 2009). However, handling hundreds to thousands of plants resulting from combinatorial transformation with several dozen transgenes will require an effective screening pipeline.

In plastid-based metabolic devices containing multicistronic operons, intercistronic expression elements (IEEs) can be used to facilitate correct processing of polycistronic transcripts into monocistronic mRNAs and their efficient translation (Fig. 2A; Zhou et al., 2007). To avoid defects in mRNA stabilization upon repeated use of the same IEE, more complex future metabolic devices may feature a variety of different such elements and/or additionally overexpress their cognate RNA-binding proteins (Legen et al., 2018).
(Verhounig et al., 2010; Emadpour et al., 2015) have been shown to be functional, yet fall short of binary behavior due to the pronounced transcriptional leakiness present in plastids (see Gene Expression Devices). To achieve a signal-to-noise ratio sufficient for the implementation of more complex logic gates, future plastid-based genetic circuits may employ synthetic RNA-binding proteins of the PPR class (see Gene Expression Devices; Coquille et al., 2014; Gully et al., 2015) to selectively control the maturation of target mRNAs in the chloroplast (Fig. 2B; Stern et al., 2010; Barkan and Small, 2014).

**Systems**

Beyond hard-wired logic gates, synthetic biologists have explored dynamic feedback mechanisms to enhance the efficiency of engineered metabolic pathways in bacteria and yeast (Venayak et al., 2015; Del Vecchio et al., 2016). Translation of this approach to plastids is currently hampered by our limited quantitative understanding of chloroplast gene expression, though new tools for analysis of the metabolic network shared between the chloroplast and its host cell are emerging (Gloaguen et al., 2017). Metabolic engineering in plastids may further be supported by expression of synthetic subcompartments for substrate concentration, metabolite channeling and the prevention of unwanted reactions between subcompartmentalized and endogenous plastid metabolites and enzymes (Winkel, 2004; Ort et al., 2015; Hanson et al., 2016). Synthetic subcompartments have already been introduced in bacteria and yeast (Bonacci et al., 2012; Lau et al., 2018), and carboxysomal shell proteins transiently expressed in leaves of *Nicotiana benthamiana* have been shown to be capable of assembling into carboxysome-like structures within chloroplasts (Lin et al., 2014), encouraging further efforts in this area.

Among the most complex systems proposed for implementation in plastids are entire synthetic genomes, inspired by recent successes in microbial synthetic genomics (Hutchison et al., 2016; Richardson et al., 2017). A minimum-size plastid genome composed of the smallest possible number of components will be of great value for two reasons: it will advance our understanding of the regulatory network underlying plastid function, and it will serve as a template for engineering synthetic plastomes to be used in biotechnological applications. We have previously proposed a design for a synthetic minimal plastome of *N. tabacum* that is free of all genes nonessential under heterotrophic growth conditions (Fig. 3), intergenic spacers, introns, and isoaccepting tRNA genes that are dispensable or become dispensable after genome-wide modification of codon usage (Scharff and Bock, 2014). Such a synthetic chloroplast genome can be assembled from linear DNA fragments in yeast (O’Neill et al., 2012) and, prior to plant transformation, can be amplified in vitro using rolling circle amplification (Jansen et al., 2005). The major hurdle to the successful implementation of fully synthetic plastomes *in planta* is the high probability of homologous recombination between the (largely nonrecodeable) rRNA and tRNA genes and their counterparts in the resident plastid genome, leading to chimeric genomes of unpredictable structure and function (O’Neill et al., 2012). In addition, the effects of
synthetic lethality (i.e. the combined knock-out of two nonessential genes being lethal; e.g. Ehrnthaler et al., 2014) cannot currently be excluded to occur in a synthetic minimal plastome.

Despite numerous technical advances made over the past 30 years, the number of algal and plant species whose plastids can reliably be transformed has remained small (Bock, 2015). Transplantation of transgenic plastids
from a species amenable to transformation to a species recalcitrant to transformation represents an attractive alternative to painstakingly developing specialized transformation protocols for the latter. Plastomes can be horizontally transferred across graft junctions with relative ease (Stegemann and Bock, 2009; Stegemann et al., 2012; Thyssen et al., 2012; for review, see Bock, 2017) and this process has been exploited for transplanting a plastid-encoded synthetic metabolic device into a currently nontransformable species (Lu et al., 2017). The graft-mediated horizontal transfer of transgenic plastid genomes may not be feasible between distantly related species due to the close coevolution of nuclear and plastid genomes, and the probability of nuclear-cytoplasmic incompatibilities that increases with phylogenetic distance and can cause deleterious phenotypes (Schmitz-Linneweber et al., 2005; Greiner and Bock, 2013). However, the transfer will certainly facilitate the expansion of transplastomic technologies from model species and cultivars used in research to related species and elite cultivars grown commercially.

**LITERATURE CITED**


Copyright (c) 2020 American Society of Plant Biologists. All rights reserved.

Downloaded on January 2, 2021. - Published by https://plantphysiol.org