A modular cloning toolbox for chloroplast metabolic engineering

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Alexander Carl Pfotenhauer
August 2021

ACKNOWLEDGEMENTS

Thank you to my mentor Scott Lenaghan, who encouraged me to pursue graduate school in a field I initially knew nothing about. His willingness to work with me regardless of my unfamiliarity is something I will always appreciate, and I am grateful to work in science today because of it. I would also like to thank the other members of my committee, Neal Stewart, David White, Curtis Luckett, and a special thanks to Alessandro Occhialini. Alessandro's training and answering of my many questions is responsible for a large amount of my success. Additional thanks is given to all other authors on the paper this thesis is derived from, as disclosed below. I would also like to thank all of my lab members throughout the years for their help with this and other projects. Final thanks goes out to my family, including my girlfriend Alyssa, and two dogs Rowdy and Ramsey for always being home with a smile.

DISCLOSURE

This thesis is largely derived from a published paper that I am an author on.

Occhialini A, Piatek A, Pfotenhauer AC, Frazier TP, Stewart CN, Lenaghan SC (2019) MoChlo: a versatile, modular cloning toolbox for chloroplast biotechnology. Plant Physiology. 179, 943-957.

Funding contribution: The authors acknowledge funding by the Advanced Research Projects Agency-Energy (Award No. DE-AR000660) and the Defense Advanced Research Projects Agency (Award No. D17AC00016).

ABSTRACT

Plant synthetic biology is a rapidly evolving field with new tools constantly emerging to drive innovation. Of particular interest is the application of synthetic biology to chloroplast biotechnology to generate plants capable of producing new metabolites, vaccines, biofuels, and high-value chemicals. Progress made in the assembly of large DNA molecules, composing multiple transcriptional units, has significantly aided in the ability to rapidly construct novel vectors for genetic engineering. In particular, Golden Gate assembly has provided a facile molecular tool for standardized assembly of synthetic genetic elements into larger DNA constructs. In this work a complete modular chloroplast cloning system (MoChlo) was developed and validated for fast and flexible chloroplast engineering in plants. A library of 128 standardized chloroplast-specific parts (47 promoters, 38 5'UTRs, 9 promoter:5'UTR fusions, 10 3' UTRs, 14 genes of interest, and 10 chloroplast-specific destination vectors) were mined from the literature and modified for use in MoChlo assembly, along with chloroplast-specific destination vectors. The strategy was validated by assembling synthetic operons of various sizes, and determining the efficiency of assembly. This method was successfully used to generate chloroplast transformation vectors containing up to 7 transcriptional units in a single vector (~10.6 kb synthetic operon). To enable researchers with limited resources to engage in chloroplast biotechnology, and to accelerate progress in the field, the entire kit, as described, will be available from Addgene at minimal cost. Thus, the MoChlo kit represents a valuable tool for fast and flexible design of heterologous metabolic pathways for plastid metabolic engineering.

TABLE OF CONTENTS

Chapter One Introduction and General Information 1
Chapter Two Literature Review 2
Chapter Three Materials and Methods 4
Design of modules for Golden Gate assembly 4
Golden Gate cloning and transformation in *E. coli* 4
Plasmid extraction from bacterial cells, restriction digestion and PCR 5
Transformation, selection, and screening of potato chloroplasts 5
Chapter Four Results and Discussion 7
Design of the MoChlo kit 7
Validation of the MoChlo strategy for assembly of simple synthetic operons (≤ 4 transcriptional units) 7
Validation of MoChlo strategy for assembly of a large operon (7 transcriptional units) 8
In vivo validation of the MoChlo kit 9
Discussion 9
Chapter Five Conclusions and Recommendations 13
List of References 14
Appendix 21
Vita 84

LIST OF TABLES

Table 1: List of Level-0 parts provided in the MoChlo Kit 23
Table 2: Design of 4 different Level-1 cassettes 24
Table 3: Design of 6 Level-1 cassettes for the *Lux* operon and 2 Level-1 selection cassettes 25
Supplemental Table S1: MoChlo Level-0 modules 26
Supplemental Table S2: Level-1 vectors compatible for the MoChlo Kit 63
Supplemental Table S3: Level-2 end-linkers and Level-2 chloroplast destination vectors 64
Supplemental Table S4: Primers used in this study 65

LIST OF FIGURES

Figure 1: Schematic representation of Golden Gate assembly using the MoChlo kit 66
Figure 2: Level-1 cloning for a 4-gene operon-1 67
Figure 3: Level-2 cloning for operon-1 68
Figure 4: Level-1 cloning for the *Lux*-operon for plastid engineering 69
Figure 5: Level-2 cloning for the 7-gene *Lux* operon 70
Figure 6: An example of *in vivo* validation of two MoChlo-enabled constructs in potato chloroplasts 71
Figure 7: Schematic of MoChlo kit for combinatorial assembly of operons for metabolic engineering 72
Supplemental Figure S1: Level-0 modules and directional cloning into Level-1 cassettes 73
Supplemental Figure S2: Directional cloning of Level-1 cassettes into operons of different lengths 74
Supplemental Figure S3: Level-1 assembly of GOIs 75
Supplemental Figure S4: Examples of Level-1 cassettes synthesized using the MoChlo kit 76
Supplemental Figure S5: Colony screening for Level-2 assembly of operon-1 using *BbsI* in the Golden Gate reaction 77
Supplemental Figure S6: Colony screening for Level-2 assembly of operon-1 using *BpiI* in the Golden Gate reaction 78
Supplemental Figure S7: Level-2 assembly of 2-genes selection cassettes using *BbsI* in the Golden Gate reaction 79
Supplemental Figure S8: Level-2 assembly of 5 different 4-genes operons 80
Supplemental Figure S9: Level-1 assembly of Level-1 *Lux*-cassettes 81
Supplemental Figure S10: Colony screening for Level-2 assembly of *Lux* operon-1 using *BpiI* in the Golden Gate reaction 82
Supplemental Figure S11: Colony screening for Level-2 assembly of *Lux* operon-1 using *BbsI* in the Golden Gate reaction 83

CHAPTER ONE
INTRODUCTION AND GENERAL INFORMATION

In this paper we describe a plastid molecular cloning (MoChlo) kit for plastid genetic engineering in plants. The MoChlo kit builds on previously developed modular cloning kits (Sarrion-Perdigones et al., 2011; Weber et al., 2011; Lampropoulos et al., 2013; Engler et al., 2014; Vafaee et al., 2014). Our goals are to lower the barrier of entry into plastid engineering, to enable troubleshooting for transgene expression, and to provide more flexibility for plastid metabolic engineering. Our hope is that a standardized assembly kit comprised of a library of plastid regulatory elements will accelerate research into metabolic engineering in plastids and benefit the research community as a whole.

CHAPTER TWO
LITERATURE REVIEW

The plastids of land plants represent an underutilized resource with regards to metabolic engineering of heterologous pathways (Maliga and Bock, 2011; Bock, 2013; Jin and Daniell, 2015; Schindel et al., 2018). Plastids have been successfully engineered for expression of valuable carotenoids (Wurbs et al., 2007; Apel and Bock, 2009) and terpenoids (Pasoreck et al., 2016), as well as, for the production of numerous biopharmaceuticals (Davoodi-Semiromi et al., 2010; Boyhan and Daniell, 2011; Lakshmi et al., 2013). Nonetheless there are far fewer instances in which the plastid chassis has been engineered for crop improvement traits. Certainly, a more controlled metabolic engineering approach would facilitate a fundamental change in crop improvement (Schindel et al., 2018). One such example, using plastid engineering, is the metabolic engineering of a complete carbon concentrating mechanism from cyanobacterium into plastids for improving photosynthesis (Lin et al., 2014a; Lin et al., 2014b; Occhialini et al., 2016; Long et al., 2018). Since plastid genomes are relatively conserved among land plants, improving photosynthesis in one species via plastids should be translatable to other species (McNeal et al., 2007).

There are several advantages to engineering plastid genomes, as opposed to nuclear ones, for metabolic engineering in plants. Plastids have small genomes (120 to 217 kb) present in multiple copies (e.g. ~10,000 plastome copies per cell in *Nicotiana tabacum*), lack gene silencing machinery, are capable of extremely high levels of foreign protein production (10-70% total soluble protein), and transgene insertion is achieved through homologous recombination, providing a known landing site for insertion (Kuroda and Maliga, 2001a; Maliga, 2004; Shaver et al., 2006; Verma and Daniell, 2007; Verma et al., 2008; Oey et al., 2009a; Schelkunov et al., 2015). Further, since chloroplasts are prokaryotic in origin, the transcription and translation machinery functions similar to prokaryotes, allowing for coordinated gene expression in operons (Bock, 2013; Lu et al., 2013). For metabolic engineering, the ability to coordinate the expression of multiple genes is essential for controlling complex multigene pathways. Unfortunately, while the regulatory machinery of plastids is prokaryotic in origin, the governing rules/functional motifs used for gene expression are not nearly as well characterized as their bacterial counterparts. Thus, while numerous logic circuits and tunable expression systems have been developed in bacterial systems, few exist for plastids (Verhounig et al., 2010; Emadpour et al., 2015). Instead, most complex pathways expressed in plastids use a single promoter to drive equivalent expression of all transgenes in the pathway (Lu et al., 2013; Fuentes et al., 2016). However, excess production of heterologous protein, even GFP, has had demonstrable negative effects on chloroplast metabolic function and the overall phenotype of the resulting plants (Oey et al., 2009a). Advanced metabolic engineering in plastids requires an understanding of how to tune the stoichiometry of multiple genes within a heterologous pathway to achieve the desired product (Bock, 2013, 2015; Schindel et al., 2018).

Typically heterologous genes engineered into plastids are placed under the control of endogenous regulatory sequences (Tregoning et al., 2003; Lutz et al., 2007; Zhou et al., 2008; Oey et al., 2009b). For effective chloroplast metabolic engineering with tunable expression, the following components must be considered: promoters, 5' untranslated regions (5'UTRs), and 3' untranslated regions (3'UTRs). Precise regulation of gene expression in plastids could be controlled by varying these components to achieve the desired expression level for each transgene or operon. While combinatorial and permutational analysis of these components in a bacterial chassis, such as *Escherichia coli*, are routinely conducted, plastids represent a more challenging system with the plastid toolkit much less developed than that for bacteria. In addition, compared to nuclear engineering in plants, few labs currently conduct plastid engineering.

CHAPTER THREE
MATERIALS AND METHODS

Design of modules for Golden Gate assembly

To generate Level-0 parts the nucleotide sequences of chloroplast regulatory elements (promoters, 5'UTRs and 3'UTRs) and genes (fluorescent/selective markers, *Lux* genes and *GOIs*) were domesticated with both 5' and 3' overhangs using the criteria set forth previously (Weber et al., 2011; Engler et al., 2014) (Supplemental Fig. S1, all tables and figures are located in the appendix). First, all IIS restriction sites (*Bsa*I, *Bpi*I and *Esp*3I) in the native part sequences were removed by single nucleotide substitutions. Next, the parts were flanked by *Bsa*I sites at opposite orientation along with synthetic cut sites that generate 5' and 3' overhangs after restriction digestion. The resulting overhangs guarantee specific directional cloning of Level-0 parts into Level-1 transcriptional units following the order Promoter-5'UTR-GOI-3'UTR (Fig. 1; Supplemental Fig. S1). Compatible overhangs for Level-1 assembly and possible combinations of Level-0 parts are shown in Supplemental Fig. S1. Level-0 modules of the same functional group are equipped with identical overhangs, allowing the generation of different cassettes by choosing specific modules of interest. The 5' overhang of Level-0 promoters and the 3' overhang of 3'UTRs were designed to ligate to compatible ends of Level-1 destination vectors. The overhangs used in the MoChlo kit were specifically designed to be compatible with vectors generated in previous plant MoClo systems (Engler et al., 2014). Finally, Level-0 and homologous sequences of chloroplast destination vectors were chemically synthesized by GeneArt (Invitrogen, Thermo Fisher Scientific) and cloned into *E.coli* pUC based-vectors (kanamycin resistance gene *nptII*). The codon usage of *Lux* genes and *GOI-1*, *2*, *3* were optimized for the *Nicotiana tabacum* chloroplast genome using Optimizer (Evolutionary Genomics Group, Biochemistry and Biotechnology Department Universitat Rovira i Virgili, Tarragona, Spain. http://genomes.urv.es/OPTIMIZER/) and a codon usage table downloaded from the codon usage database (http://www.kazusa.or.jp/codon/). Further, Level-1 destination vectors and Level-2 end-linkers from previous studies were used directly (Weber et al., 2011; Engler et al., 2014).

Golden Gate cloning and transformation in *E. coli*

The one pot-one step reaction for Golden Gate cloning was performed in a total volume of 20 µl adding the following components in the same PCR tube: (I) 25 ng of destination vector either for Level-1 or Level-2 and a variable amount of individual plasmids containing the fragments of interest. One hundred ng of plasmid containing the larger fragment along with variable amounts (ng) of other plasmids with different molecular weights. The amount of all plasmids were adjusted to guarantee the same molar ratio for all components. For Level-2, 50 ng of plasmid containing the correct end-linker were also added. (II) 1 µl (20 units) of the correct IIS restriction enzyme. *Bsa*I (*Bsa*I HF V2, New England Biolabs, NEB) was used for Level-1 and either *Bpi*I (Anza™ 4 *Bpi*I Invitrogen, Thermo Fisher Scientific) or *Bbs*I (*Bbs*I HF, New England Biolabs, NEB) for Level-2. *Bpi*I was used for operons larger than 2 cassettes. (III) 1 µl (400 units)

of T4 DNA ligase (New England Biolabs, NEB). (IV) 2 µl of T4 DNA ligase buffer (10X stock), and the final volume was adjusted to 20 µl using sterile ultra-pure water.

The Golden Gate reactions for both Level-1 and 2 were performed overnight in a thermocycler. A first step of 10 minutes at 37°C and then 10 minutes at 16°C was repeated for 35 cycles. After, the reaction was incubated at 50°C for 30 minutes, then 60 more minutes at 16°C. Before transformation into *E.coli* the enzymes were inactivated for 20 minutes at 65°C. 1 µl of Golden Gate reaction was added to 25 µl of TOP10 Chemically Competent *E.coli* (Invitrogen, Thermo Fisher Scientific) and transformed by heat-shock method (as manufacture protocol). After transformation, 250 µl of SOC media (Invitrogen, Thermo Fisher Scientific) was added and then incubated for 1 hour at 37ºC under agitation. 50 µl of bacteria cultures were plated on a petri dish containing LB-agar (Luria-Bertani media: 10 g/l Tryptone; 10 g/l NaCl; 5 g/l yeast-extract; bacto-agar 15 g/l; pH 7) containing the correct selective agent. Kanamycin and spectinomycin were used at 50 µg/ml and 100 µg/ml, respectively. Liquid bacterial cultures were grown in LB media without bacto-agar.

Plasmid extraction from bacterial cells, restriction digestion and PCR

Pure plasmid preparations were obtained using QIAprep Spin Miniprep Kit (QIAGEN). The restriction digestions using either *Bsa*I or *Bbs*I (New England Biolabs, NEB) were performed in a total volume of 30 µl using 10 units of enzyme, 500 ng of plasmid and 1X Custom Buffer (New England Biolabs, NEB). The restriction digestion using *Bpi*I was performed using the same conditions in Anza™ 1X Buffer (Invitrogen, Thermo Fisher Scientific). PCRs were performed in 25 µl reaction volumes using 1X DreamTaq Green PCR Master Mix (Thermo Fisher Scientific), along with specific pairs of primers and 25 ng of template plasmid. For colony PCR, the reactions were performed using 2 µl of overnight bacteria liquid-cultures. The following combination of forward and reverse primers (FW/RV) were used to amplify the indicated DNA fragments. *GOI-1*: Primers 1/20; *GOI-2*: Primers 2/21; *GOI-3*: Primers 3/22; *GOI-4*: Primers 4/23; *aadA* (full-length): Primers 5/24; Cassette 1A: Primers 6/25; Cassette 1B and 1C: Primers 7/25; Cassette 1D: Primers 8/26; *trnG*/*GOI-1*: Primers 9/20; *GOI-1*/*GOI-2*: Primers 1/21; *GOI-2*/*GOI-3*: Primers 2/22; *GOI-3*/*aadA*: Primers 3/24; *aadA*/*trnfM*: Primers 5/27; *Lux-C*: Primers 10/28; *Lux-D*: Primers 11/29; *Lux-A*: Primers 12/30; *Lux-B*: Primers 13/31; *Lux-E*: Primers 14/32; *Lux-G*: Primers 15/33; *aadA* (fragment): Primers 16/34; *mEmerald* (fragment): Primers 17/35; Cassette-LC1, LD1, LA1, LB1, LE1, LG1 and S1: Primers 18/36; *trnG*/*Lux-C*: Primers 9/28; *Lux-C*/*Lux-D*: Primers 10/29; *Lux-D*/*Lux-A*: Primers 11/30; *Lux-A*/*Lux-B*: Primers 12/31; *Lux-B*/*Lux-E*: Primers 13/32; *Lux-E*/*Lux-G*: Primers 14/33; *Lux-G*/*aadA*: Primers 15/34; *aadA*/*trnfM*: Primers 5/27; Right and left of T-DNA: Primers 19/37. A list of primers used in this work are shown in Supplemental Table S4. Both restriction digestion and PCR products were separated by electrophoresis using 1% agarose gel in 1X TAE buffer (40 mM Tris-acetate; 1 mM EDTA; pH 8.3) containing 0.3 µg/ml ethidium bromide (EtBr). The Hi-Lo™ DNA ladder from 50 bp to 10,000 bp (BIONEXUS) was used as molecular DNA-marker.

Transformation, selection, and screening of potato chloroplasts

Solanum tuberosum (potato) var. Desirée were grown in sterile conditions in Magenta boxes containing MS Reg. media (4.33 g/l Murashige and Skoog (MS) basal salt mixture; 25 g/l sucrose; 100 mg/l myo-inositol; 170 mg/l sodium phosphate monobasic monohydrate; 440 mg/l calcium chloride dihydrate; 0.9 mg/l thiamine-HCl; 2 mg/l glycine; 0.5 mg/l nicotinic acid; 0.5 mg/l pyridoxine-HCl; 1x MS vitamins; 3 g/l phytagel; pH 5.8). Potato plants were kept in a controlled growth room at 16 hours of light (100 µmol/m^2·s) and 8 hours of dark. The temperature was kept at 24°C constant during all light/dark cycle. Transplastomic green callus was obtained from biolistic transformation of potato leaf disks as described previously (Valkov et al., 2014). Briefly, DNA constructs were bound on the surface of 600 nm gold-particles (Seashell Technology), and bombarded into potato leaf disks on M6M media. Selection of transplastomic green callus was obtained by subculturing leaf disks in M6 media (~4 weeks) and Ti media (~6 weeks) with spectinomycin, as described previously (Valkov et al., 2014).

To confirm integration, a CTAB-based procedure was used to extract total genomic DNA from green callus prior to confirmation of integration of the transplastomic cassette by PCR (Lin et al., 2014b). After isolation, 5-10 ng of total genomic DNA was used as template for PCR as described previously. The primers 5/24 and 16/27 were used to detect the full-length *aadA* gene and the *aadA*/*trnfM* integration region, respectively (Supplemental Table S4). Transplastomic green callus was imaged using an Olympus Fv1000 confocal microscope (Olympus). The mEmerald fluorescent protein was excited using at 488 nm and detected at 505-530 nm. Chlorophyll auto-fluorescence was obtained using an excitation at 543 nm and detected at > 560 nm. Digital images were acquired using the Olympus FV10-ASW Viewer software Ver.4.2a (Olympus). Confocal images were processed using ImageJ 1.41o (National Institute of Health, Bethesda, MD, USA).

CHAPTER FOUR
RESULTS AND DISCUSSION

Design of the MoChlo kit

The nomenclature and strategy used for the various levels and components of the MoChlo system were adopted from previous plant MoClo kits (Weber et al., 2011; Engler et al., 2014) (Fig. 1; Supplemental Fig. S1). All individual parts (Level-0) mined from the literature were domesticated using the criteria described in the Materials and Methods. The domesticated MoChlo parts library consists of 47 chloroplast promoters, 38 5' UTRs, 9 fused chloroplast promoter::5'UTRs, and 10 terminators/ 3' UTRs. In addition, the library includes 14 coding genes including selectable markers (e.g. *aadA*) and reporter genes (e.g. *mEmerald* green fluorescent protein). A complete list of all Level-0 modules provided in the MoChlo kit including references to where the sequence was mined is shown in Table 1 and Supplemental Table S1. All 14 Level-1 destination vectors (7 in forward and 7 in reverse orientation) were obtained directly from a previous MoClo kit (Engler et al., 2014) (Supplemental Table S2). With the current components of the MoChlo kit it is possible to assemble up to 7 Level-1 constructs into a single Level-2 construct (Fig. 1; Supplemental Fig. S2). In addition, a set of 7 end-linkers compatible with the 3' overhang of each position of the final operon provides the option to integrate a variable number of cassettes into Level-2 vectors (Supplemental Fig. S2; Supplemental Table S3). The final component of the MoChlo system are the Level-2 chloroplast-specific destination vectors. The Level-2 destination vectors in the MoChlo system were developed for 3 plant species (*Nicotiana tabacum*; *Zea mays*; *Solanum tuberosum*) with homologous arms designed for recombination into 4 different integration sites of the plastome (*trnI-trnA*; *trnG-trnfM*; *trnV-3'rps12*; *atpB-accD*) (Supplemental Table S3). Thus, in the current version of the MoChlo kit, researchers would be able to directly assemble chloroplast transformation cassettes for multiple plant species utilizing a variety of regulatory elements, as well as, selectable marker genes.

Validation of the MoChlo strategy for assembly of simple synthetic operons (≤ 4 transcriptional units)

The first step in validating the MoChlo kit was assembling Level-1 constructs from the Level-0 parts (Fig. 2). Four Level-1 transcriptional units expressing either *GOI-1*, *GOI-2*, *GOI-3* or *aadA* were cloned into 4 Level-1 vectors for position 1, 2, 3 and 4 of the operon, respectively. A description of regulatory sequences (Level-0 parts) used to assemble the aforementioned 1A, 1B, 1C and 1D Level-1 cassette is shown in Table 2. The efficiency of assembly of the Level 1 constructs was 100% as determined by blue-white screening, PCR, and restriction digestion (**Fig. 2**, **Supplemental Fig. S3**). Finally, the correct nucleotide sequence of all Level-1 constructs was verified by Sanger sequencing. Examples of another 14 Level-1 cassettes assembled using the MoChlo kit are shown in Supplemental Fig. S4 and had the same efficiency 100%.

To validate Level 2 assembly, Level-2 operon-1 (~3.5 kb) was assembled in a second assembly step using the 4 Level-1 cassettes (1A, 1B, 1C and 1D) generated previously. For this assembly, a Level-2 destination vector equipped with homologous arms to *trnG-trnfM* from *Solanum tuberosum* was used (Fig. 3). This operon was designed to represent a standard polycistronic operon (3 GOI) along with a monocistronic selection cassette. Using *Bbs*I to release all of the Level-1 cassettes for Level-2 assembly, the efficiency of Level-2 assembly, as determined by PCR, was 22.5% (Supplemental Fig. S5). However, when using the isoschizomer *BpiI* instead, GOI-1 was found to be present in all 40 colonies tested with 37 of the colonies positive for all genes in the cassette (Supplemental Fig. S6). Thus the efficiency of Level 2 assembly was 92.5% using *Bpi*I instead of *Bbs*I, as determined by colony PCR. With a simpler assembly combining only 2 Level-1 cassettes into a Level-2, instead of the 4 previous described, there was no difference in efficiency between assembly using *Bbs*I and *Bpi*I, with both achieving 100% efficiency (Supplemental Fig. S7). Correct assembly of Level-2 operon-1 was further verified by PCR of the individual components (Fig. 3), as well as by PCR confirmation of the junctions between individual Level-1 constructs (*trnG*/*GOI-1*, *GOI-1*/*GOI-2*, *GOI-2*/*GOI-3*, *GOI-3*/*aadA* and *aadA*/*trnfM*) (Fig. 3). Examples of another 4 Level-2 operons assembled by the same 4 Level-1 cassettes are shown in Supplemental Fig. S8.

Validation of MoChlo strategy for assembly of a large operon (7 transcriptional units)

In the current iteration of the MoChlo kit, the maximum number of Level-1 constructs that can be combined into a single Level-2 cassette is 7. To test the efficiency of the maximum assembly, the 6 genes comprising the *Lux* operon (*Lux-C*, *D*, *A*, *B*, *E* and *G*) were domesticated to Level-0 parts, and assembled into Level-1 constructs (LC1, LD1, LA1, LB1, LE1, LG1) with similar efficiency to that described previously (Fig. 4, Table 3, Supplemental Fig. S9). The 7th position in the maximum Level-2 cassette was reserved for selection using either *aadA* or *mEmerald* (Cassette-S1 and S2, respectively). After sequence verification of the Level-1 constructs, Level-2 assembly of the 6 *Lux* cassettes, along with Cassette-S1, was conducted using the *Solanum tuberosum trnG-trnfM* Level-2 chloroplast destination vector (Fig. 5). For this test case, the assembly was designed to express 7 monocistronic transcriptional units (~10.6 kb total), in which the 7 genes are under the control of the same regulatory sequences (Table 3). After double selection with kanamycin/spectinomycin (*nptII* on Level-2 backbone and *aadA* within operon), it was determined 28 of 40 colonies were PCR positive for the presence of *Lux-C*, with 26 of the 28 *Lux-C* positive colonies positive for all genes of the operon (Supplemental Fig. S10). In this test case, the efficiency of assembly of 7 Level-1 constructs into a single Level-2 construct was 65%. In the case of using *Bbs*I instead of *Bpi*I the efficiency was substantially reduced to 1.25% (Supplemental Fig. S11). Correct assembly of the complete Level-2 operon was further validated using single gene and junction PCR (Fig. 5). Thus with the current design of the MoChlo kit, a single chloroplast transformation vector containing up to 7 Level-1 cassettes can be assembled with high efficiency.

In vivo validation of the MoChlo kit

To test functionality of the MoChlo kit and demonstrate the potential of the kit for combinatorial screening, two test operons were assembled using the methods previously. Operon X was driven by the *rrn* promoter and the Synthetic RBS was used as the 5'UTR. For Operon Y, the fused *accD* promoter::5'UTR were used to drive expression of the operon. In both cases, the operons were assembled into Level-2 constructs for integration into the *Solanum tuberosum trnG-trnfM* site (Fig. 6). In both operons, *mEmerald* was used as a visual marker to enable microscopic verification of functionality of the operon, while a monocistronic *aadA* cassette with the fused promoter::5'UTR of *psbA* was used for antibiotic selection of transplastomic tissue. After verification of correct assembly, the Level-2 constructs were bombarded into potato leaf disks and placed under selection with spectinomycin. After ~10 weeks, primary green callus emerged indicating that the monocistronic *aadA* cassette was functional, and molecular analysis confirmed that the operons integrated correctly into the native potato plastome. Further analysis of green callus by confocal microscopy showed that *mEmerald* fluorescence was present and correctly co-localized to the plastids in both operons (Fig. 6). In addition, the fluorescence intensity of *mEmerald* was weaker for Operon Y compared to Operon X, which was expected due to the difference in regulatory elements between the operons. The ability to assemble, transform, and validate operons in ~10 weeks demonstrates the potential for the complete system for combinatorial screening of plastid regulatory elements.

Discussion

In general, chloroplast biotechnology has fallen behind the current trends in synthetic biology, and relies on antiquated systems for vector construction. To facilitate facile assembly of large DNA-constructs necessary for metabolic engineering several molecular cloning methods based on Golden Gate assembly have been developed (Engler et al., 2008; Engler et al., 2009; Weber et al., 2011). Essentially, these methods take advantage of the unique ability of type IIS restriction enzymes to cut DNA outside of their recognition site, generating 4 bp-long non-palindromic 5' or 3' overhangs (Engler et al., 2008; Engler et al., 2009). Cleavage sites can then be designed to produce compatible overhangs able to directionally assemble multiple DNA fragments in a one pot-one step reaction. By adding the same standardized overhangs to a library of DNA basic units, it is possible to generate a number of different combination using the same cloning strategy (Sarrion-Perdigones et al., 2011; Weber et al., 2011; Lampropoulos et al., 2013). Compared with classical cloning techniques, Golden Gate assembly does not require a complicated design or time consuming cloning. In fact the combination of restriction digestion and ligation in a one pot-one step reaction eliminates the need of PCR amplification, fragment digestions and gel purifications, providing a fast and reliable method for assembly of large molecules of DNA (Weber et al., 2011). Golden Gate cloning can be performed in any laboratory equipped for molecular biology by simply purchasing enzymes (IIS restriction enzymes and T4 ligase), together with standardized DNA modules and vectors used for assembly (Marillonnet and Werner, 2015).

At present, only a single study has demonstrated the potential to utilize a modular cloning approach for the generation of chloroplast transformation vectors (Vafaee et al., 2014). In the previous work, GoldenBraid was chosen as the modular cloning strategy to demonstrate the potential for facile modular assembly of chloroplast transformation vectors. While GoldenBraid utilizes a smaller complement of destination/intermediate vectors, the number of parts that can be combined in a single step is more limited than Golden Gate. As such, the MoChlo kit was designed using the Golden Gate MoClo approach. Considering the broad adoption of the plant MoClo kits, genes of interest developed by researchers for standard nuclear expression can be directly used in the MoChlo kit without further domestication. Further, the MoChlo kit developed in this work will be released to the research community (through Addgene) with a complete set of Level-0 parts, including common chloroplast promoters, 5' UTRs, 3' UTRs, and selection genes (118 parts in total). Additionally, the kit contains Level-2 destination vectors specifically designed to enable homologous recombination into the most common sites for tobacco, corn and potato chloroplast engineering. Further, the MoChlo kit has been validated *in vivo,* demonstrating the functionality of the components and strategy. A schematic of a combinatorial study using the MoChlo kit, from assembly of constructs to the generation of transplastomic plants, is shown in Fig. 7.

Of particular importance to chloroplast biotechnology, is the relative dearth of information on precise control of multigene expression. Apart from developing a fast and flexible assembly method for chloroplast transformation vectors, a secondary goal of the MoChlo kit is to accelerate research into controllable multigene expression in plastids. Therefore, promoters and regulatory sequences constitute the majority of the MoChlo kit. Optimal levels of relative protein expression between different components is fundamental to achieve successful metabolic engineering. The expression levels of protein units encoded by both monocistronic DNA assemblies and synthetic operons can be manipulated both at transcriptional and translational level by choosing different promoters and 5' UTRs, respectively. Apart from commonly used strong promoters such as P*rrn*, P*psbA* and P*rbcL* (Shiina et al., 1998; Hayashi et al., 2003; Suzuki et al., 2003), the MoChlo kit has been supplemented with many other characterized and uncharacterized regulatory elements along with modified versions of endogenous promoters. As an example, point mutations within the -10 and -35 consensus sequences of *Prrn*, as well as in between them, have been shown to alter the level of transcription (Suzuki et al., 2003). Moreover, among the group of less characterized chloroplast promoters, several forms of P*atpB* (Xie and Allison, 2002), P*atpH*, P*atpI* (Miyagi et al., 1998), P*psbB* (Tanaka et al., 1987) and P*psbD* (Christopher et al., 1992; Allison and Maliga, 1995) have also been used to achieve variable gene expression. Additionally, promoters of *clpP* and *accD* have shown promise in gene expression in non-green plastids (Zhang et al., 2012; Caroca et al., 2013). Similarly, translation motifs, such as SD (Shine-Dalgarno sequence or ribosome binding site RBS) located in 5'UTRs are important for efficient translation of protein. In the MoChlo kit we provide a library of endogenous 5'UTRs and modified versions that are characterized by different translation efficiencies. Therefore, by using both modified versions of the same 5'UTR

and sequences from different genes, it is possible to manipulate protein translation. For example, deletion of either *psbA* RBS 1 or 2 reduces translation efficiency by ~30%, whereas the disruption of both sequences has a drastic effect on protein translation. In contrast, deletion of the third RBS of the *psbA* 5'UTR has little effect on protein translation (Hirose and Sugiura, 1996). Similarly, mutation of the SD sequence of the *rps2* 5'UTR, resulted in several fold increase in translation efficiency compared to the wild-type sequence (Plader and Sugiura, 2003). Processed forms of 5' UTRs of *psbN* and *psbB* have also been show to significantly affect translation efficiency (Yukawa et al., 2007; Kuroda and Sugiura, 2014). A chimeric synthetic 5'UTR generated from *psbA-rbcL* has also been demonstrated with a 1.5-fold decrease of translation efficiency compared to the wild type sequences (Zou et al., 2003). Considering the goal of many of these earlier works was to maximize transcription and translation of transgenes, many characterized regulatory elements have fallen by the wayside.

A key component for broad adoption of the MoChlo kit, and thus its usefulness to the research community, is its ease of use and efficiency. In our hands, the MoChlo kit provides a robust assembly strategy for chloroplast transformation vectors containing up to 7 different cassettes. A fundamental point of emphasis for achieving high efficiency assembly was the choice of the IIS restriction enzyme. For assembly of chloroplast transformation vectors containing only 2 cassettes, the choice in enzyme between *Bbs*I and *Bpi*I was irrelevant (Supplemental Fig. S7). For Level-2 assembly of 3 or more cassettes a much higher efficiency of assembly was achieved by using *Bpi*I rather than the isoschizomer *Bbs*I. Attempts to assemble 4- and 7-genes operons using *Bbs*I resulted in efficiencies of 22.5% and 1.25% (Supplemental Fig. S5 & S11), compared to 92.5% and 65% using *Bpi*I (Supplemental Fig. S6 & S10). One possible reason for the decrease in efficiency of assembly could be the kinetic properties of the two enzymes. The one pot-one step reaction is performed in T4 DNA ligase buffer and we postulate that *Bpi*I has more favorable kinetics in this suboptimal buffer. Moreover, *Bbs*I has lower thermal stability that may reduce its activity during the Golden Gate reaction. In any event, the MoChlo kit provides a considerable advantage to small labs with limited resources and funding, in that only the genes of interest must be synthesized, and complete assembly can be accomplished using only 2 enzymes.

In its current form, the MoChlo kit was designed to allow assembly of 7 transcriptional units into a single chloroplast transformation vector. Considering the high efficiency obtained for the maximum test case, it would be useful to identify the upper limit for efficient assembly of a chloroplast transformation vector. In a previous study, 11 Level-1 cassettes consisting of 44 Level-0 modules were assembled into a 33 kb DNA fragment using Golden Gate assembly (Weber et al., 2011). In future iterations of the MoChlo kit, the system will be expanded to assemble larger and more complex synthetic operons. For this purpose, a new set of Level-2 end-linkers equipped with a visual marker (e.g. *LacZα*) flanked by different IIS restriction sites (*Bsa*I or *Eps*3I) will be synthesized. The generation of compatible overhangs after restriction digestion will allow an additional Level-3 directional cloning to insert more transcriptional units into pre-assembled Level-2 constructs. Considering the relatively small size of the complete

chloroplast genome of land plants (~100-200 kb), it would theoretically possible to synthesize a complete synthetic chloroplast genome using this approach.

CHAPTER FIVE
CONCLUSIONS AND RECOMMENDATIONS

The main goal of the MoChlo kit developed in this work was to provide a standardized method to assemble multi gene constructs for metabolic engineering by using pre-synthesized chloroplast destination vectors and a wide array of regulatory elements. The library of Level-0 parts developed in this work provides both commonly used and uncharacterized plastid regulatory sequences, allowing for precise modulation of multiple genes necessary to guarantee optimal protein stoichiometry for metabolic engineering. With the modularity of the approach, the MoChlo kit can easily be expanded by adding new regulatory sequences (e.g. prokaryotic inducible promoters), protein tags (e.g. fluorescent proteins and epitopes), species specific integration sites, along with a set of new linkers that could extend the number of transcriptional units (>7) currently allowable in the kit. The substantial decrease in the cost of gene synthesis further facilitates use of the kit by researchers with limited funding and resources. The creation of a common resource of standardized Level-0 parts will provide enormous advantages in terms of collaboration and accessibility to researchers working in chloroplast biotechnology. Furthermore, the MoChlo assembly strategy is highly amenable to automation, which may further drive the potential for combinatorial screening of numerous metabolic engineering cassettes for plastid engineering.

LIST OF REFERENCES

Alieva NO, Konzen KA, Field SF, Meleshkevitch EA, Hunt ME, Beltran-Ramirez V, Miller DJ, Wiedenmann J, Salih A, Matz MV (2008) Diversity and evolution of coral fluorescent proteins. PLoS One 3: e2680

Allison LA, Maliga P (1995) Light-responsive and transcription-enhancing elements regulate the plastid psbD core promoter. EMBO J 14: 3721-3730

Allison LA, Simon LD, Maliga P (1996) Deletion of rpoB reveals a second distinct transcription system in plastids of higher plants. EMBO J 15: 2802-2809

Apel W, Bock R (2009) Enhancement of carotenoid biosynthesis in transplastomic tomatoes by induced lycopene-to-provitamin A conversion. Plant Physiol 151: 59-66

Baecker JJ, Sneddon JC, Hollingsworth MJ (2009) Efficient translation in chloroplasts requires element(s) upstream of the putative ribosome binding site from atpI. Am J Bot 96: 627-636

Bock R (2013) Strategies for metabolic pathway engineering with multiple transgenes. Plant Mol Biol 83: 21-31

Bock R (2015) Engineering plastid genomes: methods, tools, and applications in basic research and biotechnology. Annu Rev Plant Biol 66: 211-241

Bock R, Maliga P (1995) In vivo testing of a tobacco plastid DNA segment for guide RNA function in psbL editing. Mol Gen Genet 247: 439-443

Boyhan D, Daniell H (2011) Low-cost production of proinsulin in tobacco and lettuce chloroplasts for injectable or oral delivery of functional insulin and C-peptide. Plant Biotechnol J 9: 585-598

Bujarski JJ, Dreher TW, Hall TC (1985) Deletions in the 3'-terminal tRNA-like structure of brome mosaic virus RNA differentially affect aminoacylation and replication in vitro. Proc Natl Acad Sci U S A 82: 5636-5640

Caroca R, Howell KA, Hasse C, Ruf S, Bock R (2013) Design of chimeric expression elements that confer high-level gene activity in chromoplasts. Plant J 73: 368-379

Carrer H, Hockenberry TN, Svab Z, Maliga P (1993) Kanamycin resistance as a selectable marker for plastid transformation in tobacco. Mol Gen Genet 241: 49-56

Christopher DA, Kim M, Mullet JE (1992) A novel light-regulated promoter is conserved in cereal and dicot chloroplasts. Plant Cell 4: 785-798

Cubitt AB, Woollenweber LA, Heim R (1999) Understanding structure-function relationships in the Aequorea victoria green fluorescent protein. Methods Cell Biol 58: 19-30

Daniell H, Chebolu S, Kumar S, Singleton M, Falconer R (2005) Chloroplast-derived vaccine antigens and other therapeutic proteins. Vaccine 23: 1779-1783

Daniell H, Datta R, Varma S, Gray S, Lee SB (1998) Containment of herbicide resistance through genetic engineering of the chloroplast genome. Nat Biotechnol 16: 345-348

Davoodi-Semiromi A, Schreiber M, Nalapalli S, Verma D, Singh ND, Banks RK, Chakrabarti D, Daniell H (2010) Chloroplast-derived vaccine antigens confer dual immunity against cholera and malaria by oral or injectable delivery. Plant Biotechnol J 8: 223-242

Eibl C, Zou Z, Beck a, Kim M, Mullet J, Koop HU (1999) In vivo analysis of plastid psbA, rbcL and rpl32 UTR elements by chloroplast transformation: tobacco plastid gene expression is controlled by modulation of transcript levels and translation efficiency. Plant J 19: 333-345

Emadpour M, Karcher D, Bock R (2015) Boosting riboswitch efficiency by RNA amplification. Nucleic Acids Res 43: e66

Engler C, Gruetzner R, Kandzia R, Marillonnet S (2009) Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. PLoS One 4: e5553

Engler C, Kandzia R, Marillonnet S (2008) A one pot, one step, precision cloning method with high throughput capability. PLoS One 3: e3647

Engler C, Youles M, Gruetzner R, Ehnert TM, Werner S, Jones JD, Patron NJ, Marillonnet S (2014) A golden gate modular cloning toolbox for plants. ACS Synth Biol 3: 839-843

Fuentes P, Zhou F, Erban A, Karcher D, Kopka J, Bock R (2016) A new synthetic biology approach allows transfer of an entire metabolic pathway from a medicinal plant to a biomass crop. Elife 5

Hammani K, Cook WB, Barkan A (2012) RNA binding and RNA remodeling activities of the half-a-tetratricopeptide (HAT) protein HCF107 underlie its effects on gene expression. Proc Natl Acad Sci U S A 109: 5651-5656

Hayashi K, Shiina T, Ishii N, Iwai K, Ishizaki Y, Morikawa K, Toyoshima Y (2003) A role of the -35 element in the initiation of transcription at psbA promoter in tobacco plastids. Plant Cell Physiol 44: 334-341

Hirata N, Yonekura D, Yanagisawa S, Iba K (2004) Possible involvement of the 5'-flanking region and the 5'UTR of plastid accD gene in NEP-dependent transcription. Plant Cell Physiol 45: 176-186

Hirose T, Kusumegi T, Sugiura M (1998) Translation of tobacco chloroplast rps14 mRNA depends on a Shine-Dalgarno-like sequence in the 5'-untranslated region but not on internal RNA editing in the coding region. FEBS Lett 430: 257-260

Hirose T, Sugiura M (1996) Cis-acting elements and trans-acting factors for accurate translation of chloroplast psbA mRNAs: development of an in vitro translation system from tobacco chloroplasts. EMBO J 15: 1687-1695

Hirose T, Sugiura M (1997) Both RNA editing and RNA cleavage are required for translation of tobacco chloroplast ndhD mRNA: a possible regulatory mechanism for the expression of a chloroplast operon consisting of functionally unrelated genes. EMBO J 16: 6804-6811

Hirose T, Sugiura M (2004) Multiple elements required for translation of plastid atpB mRNA lacking the Shine-Dalgarno sequence. Nucleic Acids Res 32: 3503-3510

Huang FC, Klaus SM, Herz S, Zou Z, Koop HU, Golds TJ (2002) Efficient plastid transformation in tobacco using the aphA-6 gene and kanamycin selection. Mol Genet Genomics 268: 19-27

Jin S, Daniell H (2015) The engineered chloroplast genome just got smarter. Trends Plant Sci 20: 622-640

Kawamura-Nagaya K, Ishibashi K, Huang YP, Miyashita S, Ishikawa M (2014) Replication protein of tobacco mosaic virus cotranslationally binds the 5' untranslated region of genomic RNA to enable viral replication. Proc Natl Acad Sci U S A 111: E1620-1628

Khan MS, Maliga P (1999) Fluorescent antibiotic resistance marker for tracking plastid transformation in higher plants. Nat Biotechnol 17: 910-915

Kuroda H, Maliga P (2001a) Complementarity of the 16S rRNA penultimate stem with sequences downstream of the AUG destabilizes the plastid mRNAs. Nucleic Acids Res 29: 970-975

Kuroda H, Maliga P (2001b) Sequences downstream of the translation initiation codon are important determinants of translation efficiency in chloroplasts. Plant Physiol 125: 430-436

Kuroda H, Maliga P (2002) Overexpression of the clpP 5'-untranslated region in a chimeric context causes a mutant phenotype, suggesting competition for a clpP-specific RNA maturation factor in tobacco chloroplasts. Plant Physiol 129: 1600-1606

Kuroda H, Sugiura M (2014) Processing of the 5'-UTR and existence of protein factors that regulate translation of tobacco chloroplast psbN mRNA. Plant Mol Biol 86: 585-593

Kuroda H, Suzuki H, Kusumegi T, Hirose T, Yukawa Y, Sugiura M (2007) Translation of psbC mRNAs starts from the downstream GUG, not the upstream AUG, and requires the extended Shine-Dalgarno sequence in tobacco chloroplasts. Plant Cell Physiol 48: 1374-1378

Lakshmi PS, Verma D, Yang X, Lloyd B, Daniell H (2013) Low cost tuberculosis vaccine antigens in capsules: expression in chloroplasts, bio-encapsulation, stability and functional evaluation in vitro. PLoS One 8: e54708

Lampropoulos A, Sutikovic Z, Wenzl C, Maegele I, Lohmann JU, Forner J (2013) GreenGate---a novel, versatile, and efficient cloning system for plant transgenesis. PLoS One 8: e83043

Lee CY, Szittner RB, Meighen EA (1991) The lux genes of the luminous bacterial symbiont, Photobacterium leiognathi, of the ponyfish. Nucleotide sequence, difference in gene organization, and high expression in mutant Escherichia coli. Eur J Biochem 201: 161-167

Liere K, Maliga P (1999) In vitro characterization of the tobacco rpoB promoter reveals a core sequence motif conserved between phage-type plastid and plant mitochondrial promoters. Embo j 18: 249-257

Lin MT, Occhialini A, Andralojc PJ, Devonshire J, Hines KM, Parry MA, Hanson MR (2014a) beta-Carboxysomal proteins assemble into highly organized structures in Nicotiana chloroplasts. Plant J 79: 1-12

Lin MT, Occhialini A, Andralojc PJ, Parry MA, Hanson MR (2014b) A faster Rubisco with potential to increase photosynthesis in crops. Nature 513: 547-550

Long BM, Hee WY, Sharwood RE, Rae BD, Kaines S, Lim Y-L, Nguyen ND, Massey B, Bala S, von Caemmerer S, Badger MR, Price GD (2018) Carboxysome encapsulation of the CO2-fixing enzyme Rubisco in tobacco chloroplasts. Nature Communications 9: 3570

Lu Y, Rijzaani H, Karcher D, Ruf S, Bock R (2013) Efficient metabolic pathway engineering in transgenic tobacco and tomato plastids with synthetic multigene operons. Proc Natl Acad Sci U S A 110: E623-632

Lutz KA, Azhagiri AK, Tungsuchat-Huang T, Maliga P (2007) A guide to choosing vectors for transformation of the plastid genome of higher plants. Plant Physiol 145: 1201-1210

Ma J, Pallett D, Jiang H, Hou Y, Wang H (2015) Mutational bias of Turnip Yellow Mosaic Virus in the context of host anti-viral gene silencing. Virology 486: 2-6

Maliga P (2004) Plastid transformation in higher plants. Annu Rev Plant Biol 55: 289-313

Maliga P, Bock R (2011) Plastid biotechnology: food, fuel, and medicine for the 21st century. Plant Physiol 155: 1501-1510

Marillonnet S, Werner S (2015) Assembly of multigene constructs using Golden Gate cloning. Methods Mol Biol 1321: 269-284

McNeal JR, Kuehl JV, Boore JL, de Pamphilis CW (2007) Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the parasitic plant genus Cuscuta. BMC Plant Biol 7: 57

Miyagi T, Kapoor S, Sugita M, Sugiura M (1998) Transcript analysis of the tobacco plastid operon rps2/atpI/H/F/A reveals the existence of a non-consensus type II (NCII) promoter upstream of the atpI coding sequence. Mol Gen Genet 257: 299-307

Occhialini A, Lin MT, Andralojc PJ, Hanson MR, Parry MA (2016) Transgenic tobacco plants with improved cyanobacterial Rubisco expression but no extra assembly factors grow at near wild-type rates if provided with elevated CO2. Plant J 85: 148-160

Oey M, Lohse M, Kreikemeyer B, Bock R (2009a) Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. Plant J 57: 436-445

Oey M, Lohse M, Scharff LB, Kreikemeyer B, Bock R (2009b) Plastid production of protein antibiotics against pneumonia via a new strategy for high-level expression of antimicrobial proteins. Proc Natl Acad Sci U S A 106: 6579-6584

Pasoreck EK, Su J, Silverman IM, Gosai SJ, Gregory BD, Yuan JS, Daniell H (2016) Terpene metabolic engineering via nuclear or chloroplast genomes profoundly and globally impacts off-target pathways through metabolite signalling. Plant Biotechnol J 14: 1862-1875

Plader W, Sugiura M (2003) The Shine-Dalgarno-like sequence is a negative regulatory element for translation of tobacco chloroplast rps2 mRNA: an additional mechanism for translational control in chloroplasts. Plant J 34: 377-382

Sarrion-Perdigones A, Falconi EE, Zandalinas SI, Juarez P, Fernandez-del-Carmen A, Granell A, Orzaez D (2011) GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules. PLoS One 6: e21622

Schelkunov MI, Shtratnikova VY, Nuraliev MS, Selosse MA, Penin AA, Logacheva MD (2015) Exploring the limits for reduction of plastid genomes: a case study of the mycoheterotrophic orchids Epipogium aphyllum and Epipogium roseum. Genome Biol Evol 7: 1179-1191

Schindel HS, Piatek AA, Stewart CN, Jr., Lenaghan SC (2018) The plastid genome as a chassis for synthetic biology-enabled metabolic engineering: players in gene expression. Plant Cell Rep

Shaver JM, Oldenburg DJ, Bendich AJ (2006) Changes in chloroplast DNA during development in tobacco, Medicago truncatula, pea, and maize. Planta 224: 72-82

Shiina T, Allison L, Maliga P (1998) rbcL Transcript levels in tobacco plastids are independent of light: reduced dark transcription rate is compensated by increased mRNA stability. Plant Cell 10: 1713-1722

Sriraman P, Silhavy D, Maliga P (1998) The phage-type PclpP-53 plastid promoter comprises sequences downstream of the transcription initiation site. Nucleic Acids Res 26: 4874-4879

Staub JM, Maliga P (1992) Long regions of homologous DNA are incorporated into the tobacco plastid genome by transformation. Plant Cell 4: 39-45

Staub JM, Maliga P (1994) Translation of psbA mRNA is regulated by light via the 5'-untranslated region in tobacco plastids. Plant J 6: 547-553

Suzuki JY, Sriraman P, Svab Z, Maliga P (2003) Unique architecture of the plastid ribosomal RNA operon promoter recognized by the multisubunit RNA polymerase in tobacco and other higher plants. Plant Cell 15: 195-205

Svab Z, Maliga P (1993) High-frequency plastid transformation in tobacco by selection for a chimeric aadA gene. Proc Natl Acad Sci U S A 90: 913-917

Tabatabaei I, Ruf S, Bock R (2017) A bifunctional aminoglycoside acetyltransferase/ phosphotransferase conferring tobramycin resistance provides an efficient selectable marker for plastid transformation. Plant Mol Biol 93: 269-281

Tanaka M, Obokata J, Chunwongse J, Shinozaki K, Sugiura M (1987) Rapid splicing and stepwise processing of a transcript from the psbB operon in tobacco chloroplasts: determination of the intron sites in petB and petD. Mol Gen Genet 209: 427-431

Tangphatsornruang S, Birch-Machin I, Newell CA, Gray JC (2011) The effect of different 3' untranslated regions on the accumulation and stability of transcripts of a gfp transgene in chloroplasts of transplastomic tobacco. Plant Mol Biol 76: 385-396

Tregoning JS, Nixon P, Kuroda H, Svab Z, Clare S, Bowe F, Fairweather N, Ytterberg J, van Wijk KJ, Dougan G, Maliga P (2003) Expression of tetanus toxin Fragment C in tobacco chloroplasts. Nucleic Acids Res 31: 1174-1179

Trucco V, de Breuil S, Bejerman N, Lenardon S, Giolitti F (2014) Complete nucleotide sequence of Alfalfa mosaic virus isolated from alfalfa (Medicago sativa L.) in Argentina. Virus Genes 48: 562-565

Vafaee Y, Staniek A, Mancheno-Solano M, Warzecha H (2014) A modular cloning toolbox for the generation of chloroplast transformation vectors. PLoS One 9: e110222

Valkov VT, Gargano D, Scotti N, Cardi T (2014) Plastid transformation in potato: Solanum tuberosum. Methods Mol Biol. 1132: 295-303.

Verhounig A, Karcher D, Bock R (2010) Inducible gene expression from the plastid genome by a synthetic riboswitch. Proc Natl Acad Sci U S A 107: 6204-6209

Verma D, Daniell H (2007) Chloroplast vector systems for biotechnology applications. Plant Physiol 145: 1129-1143

Verma D, Samson NP, Koya V, Daniell H (2008) A protocol for expression of foreign genes in chloroplasts. Nat Protoc 3: 739-758

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

Wurbs D, Ruf S, Bock R (2007) Contained metabolic engineering in tomatoes by expression of carotenoid biosynthesis genes from the plastid genome. Plant J 49: 276-288

Xie G, Allison LA (2002) Sequences upstream of the YRTA core region are essential for transcription of the tobacco atpB NEP promoter in chloroplasts in vivo. Curr Genet 41: 176-182

Yukawa M, Kuroda H, Sugiura M (2007) A new in vitro translation system for non-radioactive assay from tobacco chloroplasts: effect of pre-mRNA processing on translation in vitro. Plant J 49: 367-376

Zhang J, Ruf S, Hasse C, Childs L, Scharff LB, Bock R (2012) Identification of cis-elements conferring high levels of gene expression in non-green plastids. Plant J 72: 115-128

Zhou F, Badillo-Corona JA, Karcher D, Gonzalez-Rabade N, Piepenburg K, Borchers AM, Maloney AP, Kavanagh TA, Gray JC, Bock R (2008) High-level expression of human immunodeficiency virus antigens from the tobacco and tomato plastid genomes. Plant Biotechnol J 6: 897-913

Zhou F, Karcher D, Bock R (2007) Identification of a plastid intercistronic expression element (IEE) facilitating the expression of stable translatable monocistronic mRNAs from operons. Plant J 52: 961-972

Zou Z, Eibl C, Koop HU (2003) The stem-loop region of the tobacco psbA 5'UTR is an important determinant of mRNA stability and translation efficiency. Mol Genet Genomics 269: 340-349

APPENDIX

Table 1: List of Level-0 parts provided in the MoChlo Kit.

Promoters	
P*accD* (-82 to +23)[1]	P*rrn16*[13]
P*accD* (functional)[1]	P*rrn* (*N. tabacum*)[6]
P*atpB* (NEP functional PEP disabled)[2]	P*rrn* (*N. tabacum*) AP*rrn1*^{u}
P*atpB* (NEP)[2]	P*rrn* (*N. tabacum*) AP*rrn2*^{u}
P*atpB* (PEP and NEP functional)[2]	P*rrn* (*N. tabacum*) AP*rrn3*^{u}
P*atpB* (PEP functional NEP disabled)[2]	P*rrn* (*N. tabacum*) AP*rrn4*^{u}
P*atpB* (PEP)[2]	P*rrn* (*N. tabacum*) AP*rrn5*^{u}
P*atpH*[3]	P*rrn* (*N. tabacum*) AP*rrn6*^{u}
P*atpI* (NEP)[3]	P*rrn* (*N. tabacum*) AP*rrn7*^{u}
P*atpI* (PEP Hypothetical)[3]	P*rrn* (*N. tabacum*) AP*rrn8*^{u}
P*clpP*[4]	P*rrn* (*N. tabacum*) AP*rrn9*^{u}
P*clpP* (PEP)[4]	P*rrn* (*N. tabacum*) AP*rrn10*^{u}
IEE[5]	P*rrn* (*N. tabacum*) BP*rrn12*^{u}
P*rrn* (-64 to +17)[6]	P*rrn* (*N. tabacum*) BP*rrn13*^{u}
mP*rrn* (5-fold decrease transcription)[6]	P*rrn* (*N. tabacum*) BP*rrn14*^{u}
P*psbA*[7]	P*rrn* (*N. tabacum*) BP*rrn15*^{u}
P*psbA* (-42 to +9)[8]	P*rrn* (*N. tabacum*) BP*rrn16*^{u}
P*psbB*[9]	P*rrn* (*N. tabacum*) BP*rrn17*^{u}
P*psbD*-leader[10]	P*rrn* (*N. tabacum*) BP*rrn18*^{u}
P*psbD*-PGTbx[10]	P*rrn* (*N. tabacum*) BP*rrn19*^{u}
P*psbD*[10]	P*rrn* (*N. tabacum*) BP*rrn20*^{u}
P*rbcL*[11]	P*rrn* (*N. tabacum*) BP*rrn21*^{u}
P*rbcL* (core -3 to +9)[11]	
P*rbcL* (-110 to +9)[11]	
P*rpoB*[12]	

5' untranslated region (5' UTR)	
atpB 5'UTR[16]	*rbcL* 5'UTR (full)[17]
atpB 5'UTR +10 (functional)[16]	*rbcL* 5'UTR (processed)[17]
atpB 5'UTR +10 processed -90 (functional)[17]	*rbcL* 5'UTR + 14 AA[18]
atpB 5'UTR TCR processed -90 (functional)[18]	Synthetic RBS[26]
atpH 5'UTR[17]	T7G10 leader[26]
atpH 5'UTR processed -45[17]	T7G10 leader + Ec DB[26]
atpI 5'UTR[19]	*rrn16* 5'UTR (native)[13]
psbN 5'UTR[20]	*rrn16* 5'UTR (+37)[13]
psbN 5'UTR processed -39[20]	*psbA* 5'UTR[7]
psbB 5'UTR[17]	Synthetic RBS[10]
psbB 5'UTR processed -54[17]	*psbC* 5'UTR[27]
ndhD 5'UTR[21]	*accD* 5'UTR[28]
rps2 5'UTR[22]	*clpP* 5'UTR[29]
rps2 5'UTR mAAGA[22]	
rps2 5'UTR mCCUC[22]	
rps2 5'UTR mAAGA (-20 to -10 deletion)[22]	
rps2 5'UTR mCCUC (-20 to -10 deletion)[22]	
psbH 5'UTR + start codon + 1AA[23]	
psbA 5'UTR (SNM)[24]	
psbA 5'UTR (SEM)[24]	
psbA 5'UTR (delta SL)[24]	
psbA 5'UTR (delta 5'SL)[24]	
psbA-rbcL chimeric 5'UTR (R delta 5'SL)[24]	
psbA-rbcL chimeric 5'UTR (R'SLWT)[24]	
rps14 5'UTR[25]	

Promoter::5'UTR	Genes of interest (*GOI*)
Zm P*clpP*::5'UTR[14]	*mEmerald*[30]
Zm P*clpP*::Gg10 leader[14]	*pporRFP*[31]
P*accD*::5'UTR[14]	*aadA*[32]
mP*rrn*::RBS (3.9%)[6]	*aphA6*[33]
mP*rrn*::RBS (20%)[6]	*nptII*[34]
mP*rrn*::RBS (46%)[6]	*Myc aadA-GFP*[35]
mP*rrn*::RBS (60%)[6]	*FLARE16-S aadA-GFP*[36]
mP*rrn*::RBS (61%)[6]	*aac6-aph2*[37]
P*rrn*::RBS[15]	*Lux-C* codon opt. *Nt* chloroplasts[38]
	Lux-D codon opt. *Nt* chloroplasts[38]
	Lux-A codon opt. *Nt* chloroplasts[38]
	Lux-B codon opt. *Nt* chloroplasts[38]
	Lux-E codon opt. *Nt* chloroplasts[38]
	Lux-G codon opt. *Nt* chloroplasts[38]

Integration sites	3' untranslated region (3' UTR)
N. tabacum trnI/trnA[45]	*psbA* 3'UTR[39]
N. tabacum trnG/trnfM[46]	*rbcL* 3'UTR[39]
N. tabacum trnV/3'rps12[47]	*petD* 3'UTR[40]
N. tabacum atpB/accD[48]	*rpoA* 3'UTR[40]
Z. mays trnI/trnA[45]	*E.coli rrnB* 3'UTR[40]
Z. mays trnV/3'rps12[47]	*psaC* 3'UTR[21]
S. tuberosum trnI/trnA[45]	TMV 3'UTR[41]
S. tuberosum trnG/trnfM[46]	TYMV 3'UTR[42]
S. tuberosum trnV/3'rps12[47]	BMV 3'UTR[43]
S. tuberosum trnV/3'rps12[47]	ALMV 3'UTR[44]

Table 1 continued

* A detailed description of all Level-0 parts, including nucleotide sequences, is included in Supplemental Table S1.

[1] Hirata et al., 2004; [2] Xie and Allison, 2002; [3] Miyagi et al., 1998; [4] Sriraman et al., 1998; [5] Zhou et al., 2007; [6] Suzuki et al., 2003; [7]: Staub and Maliga, 1994; [8] Hayashi et al., 2003; [9] Tanaka et al., 1987; [10] Allison and Maliga, 1995; [11] Shiina et al., 1998; [12] Liere and Maliga, 1999; [13] Allison et al., 1996; [14] Zhang et al., 2012; [15] Daniell et al., 2005; [16] Hirose and Sugiura, 2004; [17] Yukawa et al., 2007; [18] Kuroda and Maliga, 2001b; [19] Baecker et al., 2009; [20] Kuroda and Sugiura, 2014; [21] Hirose and Sugiura, 1997; [22] Plader and Sugiura, 2003; [23] Hammani et al., 2012; [24] Zou et al., 2003; [25] Hirose et al., 1998; [26] Kuroda and Maliga, 2001a; [27] Kuroda et al., 2007; [28] Caroca et al., 2013; [29] Kuroda and Maliga, 2002; [30] Cubitt et al., 1999; [31] Alieva et al., 2008; [32] Svab and Maliga, 1993; [33] Huang et al., 2002; [34] Carrer et al., 1993; [35] Lutz et al., 2007; [36] Khan and Maliga, 1999; [37] Tabatabaei et al., 2017; [38] Lee et al., 1991; [39] Eibl et al., 1999; [40] Tangphatsornruang et al., 2011; [41] Kawamura-Nagaya et al., 2014; [42] Ma et al., 2015; [43] Bujarski et al., 1985; [44] Trucco et al., 2014; [45] Daniell et al., 1998; [46] Bock and Maliga, 1995; [47] Staub and Maliga, 1992; [48] Lin et al., 2014b.

Table 2: Design of 4 different Level-1 cassettes

Name	Promoter (P)	5'UTR (5')	GOI	3'UTR (3')	Size (kb)	Position (fw)
Cassette 1A	P*accD*	Synth RBS	*GOI-1*	*petD*-3'UTR	0.664	1
Cassette 1B	IEE	Synth RBS	*GOI-2*	*petD*-3'UTR	0.583	2
Cassette 1C	IEE	Synth RBS	*GOI-3*	*petD*-3'UTR	1.114	3
Cassette 1D	P*psbA*	Synth RBS	*aadA*	*psbA*-3'UTR	1.019	4

Table 3: Design of 6 Level-1 cassettes for the *Lux* operon and 2 Level-1 selection cassettes

Name	Promoter (P)	5'UTR (5')	GOI	3'UTR (3')	Size (kb)	Position (fw)
Cassette-LC1	P*rrn*	P*rrn*-5'UTR	*Lux-C*	*rrnB*-3'UTR	1.919	1
Cassette-LD1	P*rrn*	P*rrn*-5'UTR	*Lux-D*	*rrnB*-3'UTR	1.430	2
Cassette-LA1	P*rrn*	P*rrn*-5'UTR	*Lux-A*	*rrnB*-3'UTR	1.547	3
Cassette-LB1	P*rrn*	P*rrn*-5'UTR	*Lux-B*	*rrnB*-3'UTR	1.463	4
Cassette-LE1	P*rrn*	P*rrn*-5'UTR	*Lux-E*	*rrnB*-3'UTR	1.604	5
Cassette-LG1	P*rrn*	P*rrn*-5'UTR	*Lux-G*	*rrnB*-3'UTR	1.187	6
Cassette-S1	P*rrn*	P*rrn*-5'UTR	*aadA*	*rrnB*-3'UTR	1.274	7
Cassette-S2	P*rrn*	Prrn-5'UTR	*mEmerald*	*rrnB*-3'UTR	1.202	7

Supplementary Table S1

Type	Name	Sequence	bp	Resistance	Description	Reference
Promoter	P*accD* (-82 to +23)	GGTCTCAGGAGGACATAGGAGAAGCC GCCCTTTATTAAAAATTATATTATTTTAA ATAATATAAAGGGGGTTCCAACATATTA ATATATAGTGAAGTGTTCCCCCAGATT CAGAACTACTTGAGACC	127	Kanamycin	*accD* promoter, -82 to +23 slight increase from native when transgene attached. Decreased transgene expression if another 5'UTR longer than 129 bp is used in between transgene	(Hirata et al., 2004)
Promoter	P*accD* (functi onal)	GGTCTCAGGAGCGATGCGAATTTGACA CGACATAGGAGAAGCCGCCCTTTATTA AAAATTATATTATTTTAAATAATATAAAG GGGGTTCCAACATATTAATATATAGTG AAGTGTTCCCCCAGATTCAGAACTACT TGAGACC	144	Kanamycin	*accD* promoter, -99 to +23 fully functional when transgene attached. Decreased transgene expression if another 5'UTR longer than 129 bp is used in between transgene	(Hirata et al., 2004)
Promoter	P*atpB* (NEP functio nal PEP disabl ed)	GGTCTCAGGAGAATAGAAATAGAAAAT AAAGTTCAGGTTCGAATTCCATAGAAT AGATAATATGGATGGGATTGTCGCGAA TTGATAGACAATACTTGAGACC	103	Kanamycin	*atpB* promoter, mutation in -10 motif disables PEP function	(Xie and Allison, 2002)
Promoter	P*atpB* (NEP)	GGTCTCAGGAGAATAGAAATAGAAAAT AAAGTTCAGGTTCGAATTCCATAGAAT AGATAATACTTGAGACC	71	Kanamycin	*atpB* promoter; only NEP core sequence (untested)	(Xie and Allison, 2002)
Promoter	P*atpB* (PEP and NEP functio nal)	GGTCTCAGGAGAATAGAAATAGAAAAT AAAGTTCAGGTTCGAATTCCATAGAAT AGATAATATGGATGGGATTGTCTATAAT GATAGACAATACTTGAGACC	102	Kanamycin	*atpB* promoter, functional when fused to RBS leader	(Xie and Allison, 2002)
Promoter	P*atpB* (PEP functio nal NEP disabl ed)	GGTCTCAGGAGAATAAAGTTCAGGTTC GAATTCCATAGAATAGATAATATGGAT GGGATTGTCTATAATGATAGACAATAC TTGAGACC	89	Kanamycin	*atpB* promoter, deletion of GAA-box prevents NEP function	(Xie and Allison, 2002)
Promoter	P*atpB* (PEP)	GGTCTCAGGAGTAGATAATATGGATGG GATTGTCTATAATGATAGACAATACTTG AGACC	60	Kanamycin	*atpB* promoter, only PEP core sequence (untested)	(Xie and Allison, 2002)
Promoter	P*atpH*	GGTCTCAGGAGATAGGATTGAATCTAA GATGAATGCTTGGTTTACGTTATGGAA GAAAGACATGTATATGTGATATTATACT TGAGACC	89	Kanamycin	*atpH* complete promoter (untested)	(Miyagi et al., 1998)
Promoter	P*atpI* (NEP)	GGTCTCAGGAGTCCTTGATTAATAATA ACATAATAAGATATACTTGAGACC	51	Kanamycin	*atpI* promoter NEP	(Miyagi et al., 1998)
Promoter	P*atpI* (PEP Hypot hetical)	GGTCTCAGGAGTCTAGCTATATAAGAA ATCCTTGATTAATAATAACATAATATAC TTGAGACC	63	Kanamycin	Hypothetical *atpI* promoter PEP (untested)	(Miyagi et al., 1998)
Promoter	P*clpP*	GGTCTCAGGAGAGACAATAAAAAAAAT TGTTACGTTTCCACCTCAAAGTGAAATA CTTGAGACC	64	Kanamycin	*clpP* promoter NEP -53	(Sriraman et al., 1998)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
Promoter	P*clpP* (PEP)	GGTCTCAGGAGAAAGAGCTCGGACTTTATTCATTCTGTCTTTCTTTATGAATTTTTATAATCTATGGATAAAATAAATACGATAAAAACCAATATGAATATTATAAAGACAATAAAAAAAATTGTTACTACTTGAGACC	139	Kanamycin	*clpP* promoter PEP -117	(Sriraman et al., 1998)
Promoter	IEE	GGTCTCAGGAGTAGGATCGTTTATTTACAACGGAATGGTATACAAAGTCAACAGATCTCAATACTTGAGACC	72	Kanamycin	Intercistronic expression element for functional cleavage to monocistronic mRNA	(Zhou et al., 2007)
Promoter	P*rrn* (-64 to +17)	GGTCTCAGGAGATGAGAATGGATAAGAGGCTCGTGGGATTGACGTGAGGGGGCAGGGATGGCTATATTTCTGGGAGCGAACTCCGGGCGAATTACTTGAGACC	103	Kanamycin	*rrn* promoter -64 to +17 (fully functional)	(Suzuki et al., 2003)
Promoter	mP*rrn* (5-fold decrease transcription)	GGTCTCAGGAGATTGACGTGAGGGGGCAGGGATGGCTATATTTCTGGGAGCGAACTCCGGGCGAATTACTTGAGACC	77	Kanamycin	*rrn* promoter, 5-fold decrease in transcription from native	(Suzuki et al., 2003)
Promoter	P*psbA*	GGTCTCAGGAGGATCTACATACACCTTGGTTGACACGAGTATATAAGTCATGTTATACTGTTGTACTTGAGACC	74	Kanamycin	*psbA* promoter upstream of transcription start	(Staub and Maliga, 1994)
Promoter	P*psbA* (-42 to +9)	GGTCTCAGGAGAATAAAAAGCCTTCCATTTTCTATTTTGATTTGTAGAAAACTAGTGTGCTTGGGAGTCCCTGATGATTAAATAAACCAAGATTTTACCTACTTGAGACC	110	Kanamycin	*psbA* promoter -42 to +9, functional promoter core when fused to RBS leader	(Hayashi et al., 2003)
Promoter	P*psbB*	GGTCTCAGGAGGAAAATGGGATCAACTCCCCCATTGCGTATTGGTACTTATCGAGTATAGAATAAATCTTACTTGAGACC	80	Kanamycin	*psbB* promoter	(Tanaka et al., 1987)
Promoter	P*psbD* -leader	GGTCTCAGGAGCGAAAGGTTAGAAATCAACAAAAGAAAAAGTAAGTGGACCTGACCTATTGAATCATGACTATATCCGCTATTCTGATATTAAAATTCGATAGAGATGAAATTGGAGCTCTAGAGAATTCAGTTGTAGGGAGGGATCCTACTTGAGACC	159	Kanamycin	Full *psbD* promoter with synthetic leader	(Allison and Maliga, 1995)
Promoter	P*psbD* -PGTbx	GGTCTCAGGAGGAAAAAGTAAGTGGACCTGACCTATTGAATCATGACTATATCCGCTATTCTGATATTAAAATTCGATAGAGATGAAATTGGAGCTCTAGAGAATTCAGTTGTAGGGAGGGATCCTACTTGAGACC	136	Kanamycin	*psbD* promoter, 5-fold decrease in transcription from native	(Allison and Maliga, 1995)
Promoter	P*psbD*	GGTCTCAGGAGCGAAAGGTTAGAAATCAACAAAAGAAAAAGTAAGTGGACCTGACCTATTGAATCATGACTATATCCGCTATTCTGATATTAAAATTCGATAGAGATGAAATTGGAGCTACTTGAGACC	129	Kanamycin	Native *psbD* promoter upstream of transcription start	(Allison and Maliga, 1995)
Promoter	P*rbcL*	GGTCTCAGGAGGGGGGAAGTTCTTATTATTTAGGTTAGTCAGGTATTTCCATTTCAAAAAAAAAAAAAAGTAAAAAAGAAAAATTGGGTTGCGCTATATATATGAAAGAGTATACAATAATGTACTTGAGACC	132	Kanamycin	Native *rbcL* promoter upstream of transcription start	(Shiina et al., 1998)
Promoter	P*rbcL* (core -3 to +9)	GGTCTCAGGAGTTGCGCTATATATATGAAAGAGTATACAATAATGATGTATTTGTACTTGAGACC	65	Kanamycin	Core *rbcL* promoter, fully functional when fused to RBS	(Shiina et al., 1998)
Promoter	P*rbcL* (-110 to +9)	GGTCTCAGGAGGGGGGAAGTTCTTATTATTTAGGTTAGTCAGGTATTTCCATTTCAAAAAAAAAAAAAAGTAAAAAAGAAAAATTGGGTTGCGCTATATATATGAAAGAGTATACAATAATGATGTATTTGTACTTGAGACC	141	Kanamycin	*rbcL* promoter -110 to +9, fully functional when fused to RBS	(Shiina et al., 1998)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
Promoter	P*rpoB*	GGTCTCAGGAGGGTTGGAATGTGTATT ATCATAATAATTACTTGAGACC	49	Kanamycin	*rpoB* promoter NEP -345	(Liere and Maliga, 1999)
Promoter	P*rrn16*	GGTCTCAGGAGATGAGAATGGATAAGA GGCTCGTGGGATTGACGTGAGGGGGC AGGGATGGCTATATTTCTGGGATACTT GAGACC	86	Kanamycin	Native *rrn16* promoter upstream of transcription start	(Allison et al., 1996)
Promoter	P*rrn* (*Nicotiana tabacum*)	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTGACGTGAGGGGGCAGGGATG GCTATATTTCTGGGAGCGAACTCCGGG CGAATACGAAGCGCTTGGATACAGTTG TAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter from *Nicotiana tabacum*	(Suzuki et al., 2003)
Promoter	P*rrn* (*Nicotiana tabacum*) AP*rrn1*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTTACGTGAGGGGGCAGGGATG GCGTATCTTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTTACG), -10(GTATCT)" *Nicotiana tabacum*	(Davis et al., 2011)
Promoter	P*rrn* (*Nicotiana tabacum*) AP*rrn2*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGAGCGGTGTGAGGGGGCAGGGAT GGCTATAATTCTGGGAGCGAACTCCG GGCGAATACGAAGCGCTTGGATACAG TTGTAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(GCGGTG), -10(TATAAT)" *Nicotiana tabacum*	(Davis et al., 2011)
Promoter	P*rrn* (*Nicotiana tabacum*) AP*rrn3*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTTACGTGAGGGGGCAGGGATG GCGAGGATTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTTACG), -10(GAGGAT)" *Nicotiana tabacum*	(Davis et al., 2011)
Promoter	P*rrn* (*Nicotiana tabacum*) AP*rrn4*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTTACGTGAGGGGGCAGGGATG GCTAGGCTTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTTACG), -10(TAGGCT)" *Nicotiana tabacum*	(Davis et al., 2011)
Promoter	P*rrn* (*Nicotiana tabacum*) AP*rrn5*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTTACGTGAGGGGGCAGGGATG GCGATGATTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTTACG), -10(GATGAT)" *Nicotiana tabacum*	(Davis et al., 2011)
Promoter	P*rrn* (*Nicotiana tabacum*) AP*rrn6*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTTACGTGAGGGGGCAGGGATG GCTAGGATTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTTACG), -10(TAGGAT)" *Nicotiana tabacum*	(Davis et al., 2011)
Promoter	P*rrn* (*Nicotiana tabacum*) AP*rrn7*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTTACGTGAGGGGGCAGGGATG GCTAATATTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTTACG), -10(TAATAT)" *Nicotiana tabacum*	(Davis et al., 2011)
Promoter	P*rrn* (*Nicotiana tabacum*) AP*rrn8*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTTACGTGAGGGGGCAGGGATG GCTAAAATTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTTACG), -10(TAAAAT)" *Nicotiana tabacum*	(Davis et al., 2011)
Promoter	P*rrn* (*Nicotiana*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTTACGTGAGGGGGCAGGGATG	157	Kanamycin	*rrn* promoter "-35(TTTACG), -10(TATGAT)"	(Davis et al., 2011)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
	tabacum) AP*rrn9*	GCTATGATTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTTACTTGAGACC			*Nicotiana tabacum*	
Promoter	P*rrn* (*Nicotiana tabacum*) AP*rrn10*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTTACGTGAGGGGGCAGGGATG GCTATAATTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTTACG), -10(TATAAT)" *Nicotiana tabacum*	(Davis et al., 2011)
Promoter	P*rrn* (*Nicotiana tabacum*) BP*rrn12*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTGTCGTGAGGGGGCAGGGATG GCTATATTTCTGGGAGCGAACTCCGGG CGAATACGAAGCGCTTGGATACAGTTG TAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTGTCG), -10(TATATT)"; A to T in -35 3rd pos. (weak destabilize) *Nicotiana tabacum*	Designed in this work
Promoter	P*rrn* (*Nicotiana tabacum*) BP*rrn13*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTGGAGTGAGGGGGCAGGGATG GCTATATTTCTGGGAGCGAACTCCGGG CGAATACGAAGCGCTTGGATACAGTTG TAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTGGAG), -10(TATATT)"; A to G in -35 3rd pos. (strong destabilize) *Nicotiana tabacum*	Designed in this work
Promoter	P*rrn* (*Nicotiana tabacum*) BP*rrn14*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTGAAATGAGGGGGCAGGGATG GCTATATTTCTGGGAGCGAACTCCGGG CGAATACGAAGCGCTTGGATACAGTTG TAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTGAAA), -10(TATATT)"; C to A in -35 2nd pos. (weak destabilize) *Nicotiana tabacum*	Designed in this work
Promoter	P*rrn* (*Nicotiana tabacum*) BP*rrn15*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTGATGTGAGGGGGCAGGGATG GCTATATTTCTGGGAGCGAACTCCGGG CGAATACGAAGCGCTTGGATACAGTTG TAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTGATG), -10(TATATT)"; C to T in -35 2nd pos. (strong destabilize) *Nicotiana tabacum*	Designed in this work
Promoter	P*rrn* (*Nicotiana tabacum*) BP*rrn16*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTGACGTGAGGGGGCAGGGATG GCTATTTTTCTGGGAGCGAACTCCGGG CGAATACGAAGCGCTTGGATACAGTTG TAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTGACG), -10(TATTTT)"; A to T in -10 3rd pos. (weak destabilize) *Nicotiana tabacum*	Designed in this work
Promoter	P*rrn* (*Nicotiana tabacum*) BP*rrn17*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTGACGTGAGGGGGCAGGGATG GCTATGTTTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTGACG), -10(TATGTT)"; A to G in -10 3rd pos. (strong destabilize) *Nicotiana tabacum*	Designed in this work
Promoter	P*rrn* (*Nicotiana tabacum*) BP*rrn18*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTGACGTGAGGGGGCAGGGATG GCTATAGTTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTGACG), -10(TATAGT)"; T to G in -10 2nd pos. (weak destabilize) *Nicotiana tabacum*	Designed in this work
Promoter	P*rrn* (*Nicotiana tabacum*)	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTGACGTGAGGGGGCAGGGATG GCTATACTTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT	157	Kanamycin	*rrn* promoter "-35(TTGACG), -10(TATACT)"; T to C in -10 2nd pos. (strong destabilize)	Designed in this work

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
	BP*rrn19*	GTAGGGAGGGATTTTACTTGAGACC			*Nicotiana tabacum*	
Promoter	P*rrn* (*Nicotiana tabacum*) BP*rrn20*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTGACGTGAGGGGGCAGGGATG GCTATTGTTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTGACG), -10(TATTGT)"; A to T in -10 3rd pos.; T to G in -10 2nd pos. (double weak) *Nicotiana tabacum*	Designed in this work
Promoter	P*rrn* (*Nicotiana tabacum*) BP*rrn21*	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTGACGTGAGGGGGCAGGGATG GCTATGCTTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTTACTTGAGACC	157	Kanamycin	*rrn* promoter "-35(TTGACG), -10(TATGCT)"; A to G in -10 3rd pos.; T to C in -10 2nd pos. (double strong) *Nicotiana tabacum*	Designed in this work
Promoter:: 5'UTR	*Zm* P*clpP*::5'UTR	GGTCTCAGGAGAAAGAGCTCGGATCC TCTATGTATTAATAGAATCTATAGTATT CTTATAGAATAAGAAAAAAAAAATGAAG ATAATAAACTGCGGATTCTTTCTTTCTC TTCCATTCTTACGTTTCCATATTAAAGT GTAGTTTTTTTACTTAAATTTAATAATAT TAATCTACCAATGTGAGACC	187	Kanamycin	*clpP* promoter with 5'UTR, *Zea mays*	(Zhang et al., 2012)
Promoter:: 5'UTR	*Zm* P*clpP*::Gg10 leader	GGTCTCAGGAGAAAGAGCTCGGATCC TCTATGTATTAATAGAATCTATAGTATT CTTATAGAATAAGAAAAAAAAAATGAAG ATAATAAACTGCGGATTCTTTCTTTCTC TTCCATTCTTACGTTTCCATATTAAAGT GTAGTTTTTTTACTTAAATTTAAAGAAG GAGATATACCCAATGTGAGACC	188	Kanamycin	*clpP* promoter with G10 leader, *Zea mays*	(Zhang et al., 2012)
Promoter:: 5'UTR	P*accD*::5'UTR	GGTCTCAGGAGAAAAGAGCTCCGACA TAGGAGAAGCCGCCCTTTATTAAAAAT TATATTATTTTAAATAATATAAAGGGGG TTCCAACATATTAATATATAGTGAAGTG TTCCCCCAGATTCAGAACTTTTTTTCAA TACTCACAATCCTTATTAGTTAATAATC CTAGTGATTGGATTTCTATGCTTAGTCT GATAGGAAATAAGATATTCAAATAAATA ATTTTATAGCGAATGACTATTCATCTAT TGTATTTTCATGCAAATAGGGGGCAAG AAAACTCTATGAATGTGAGACC	298	Kanamycin	*accD* promoter with 5'UTR	(Zhang et al., 2012)
Promoter:: 5'UTR	mP*rrn*::RBS (3.9%)	GGTCTCAGGAGTGAGGGGGCAGGGAT GGCTATATTTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTAATGTGAGACC	105	Kanamycin	*rrn* promoter with RBS 3.9%	(Suzuki et al., 2003)
Promoter:: 5'UTR	mP*rrn*::RBS (20%)	GGTCTCAGGAGTTGACGTGAGGGGGC AGGGATGGCAATATATCTGGGAGCGA ACTCCGGGCGAATACGAAGCGCTTGG ATACAGTTGTAGGGAGGGATTTAATGT GAGACC	111	Kanamycin	*rrn* promoter with RBS 20%	(Suzuki et al., 2003)
Promoter:: 5'UTR	mP*rrn*::RBS (46%)	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTGACGTGAGGGGGCAGGGATG GCAATATATCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTAATGTGAGACC	157	Kanamycin	*rrn* promoter with RBS 46%	(Suzuki et al., 2003)
Promoter:: 5'UTR	mP*rrn*::RBS (60%)	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT GGGATTGACGTGAGGGGGCAGGGATG GCATAATTTCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTAATGTGAGACC	157	Kanamycin	*rrn* promoter with RBS 60%	(Suzuki et al., 2003)
Promoter:: 5'UTR	mP*rrn*::RBS	GGTCTCAGGAGGCTCCCCCGCCGTCG TTCAATGAGAATGGATAAGAGGCTCGT	157	Kanamycin	*rrn* promoter with RBS 61%	(Suzuki et al., 2003)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
	(61%)	GGGATTGACGTGAGGGGGCAGGGATG GCTATTAATCTGGGAGCGAACTCCGG GCGAATACGAAGCGCTTGGATACAGTT GTAGGGAGGGATTTAATGTGAGACC				
Promoter:: 5'UTR	P*rrn*:: RBS	GGTCTCAGGAGTAAGAGGCTCGTGGG ATTGACGTGAGGGGGCAGGGATGGCT ATATTTCTGGGAGCGAACTCCGGGCG AATATGAAGCGCATCGATACAAGTGAG TTGTAGGGAGGGAACCAATGTGAGAC C	132	Kanamycin	*rrn* promoter with RBS	(Daniell et al., 2005)
5' UTR	*atpB* 5'UTR	GGTCTCATACTCAAATGAAAGACTTTCT CAAGATTCTGATTCATCCACTTGAGATT TTAAAATTAAAATAGGTTGGGTGGGCT TGCAAATTCACTCAGTCTCAGTGAATA AGTAAACAATTGAATCGGTTCAATTGC ATGGTGCCAACGAAATCGAGTGCTAAT TCCCATTTTATTGAATTAACCGATCGAC GTGCTAGCGGACATTTATTTTGAATTC GATAATTTTTGCAAAAACATTTCGACAT ATTTATTTATTTTATTATTAATGTGAGAC C	277	Kanamycin	Native *atpB* 5'UTR without start codon	(Hirose and Sugiura, 2004)
5' UTR	*atpB* 5'UTR +10 (functional)	GGTCTCATACTCAAATGAAAGACTTTCT CAAGATTCTGATTCATCCACTTGAGATT TTAAAATTAAAATAGGTTGGGTGGGCT TGCAAATTCACTCAGTCTCAGTGAATA AGTAAACAATTGAATCGGTTCAATTGC ATGGTGCCAACGAAATCGAGTGCTAAT TCCCATTTTATTGAATTAACCGATCGAC GTGCTAGCGGACATTTATTTTGAATTC GATAATTTTTGCAAAAACATTTCGACAT ATTTATTTATTTTATTATTATGAGAATCA AATGTGAGACC	287	Kanamycin	*atpB* 5'UTR with +10 region required for translation	(Hirose and Sugiura, 2004)
5' UTR	*atpB* 5'UTR +10 processed -90 (functional)	GGTCTCATACTGAATTAACCGATCGAC GTGCTAGCGGACATTTATTTTGAATTC GATAATTTTTGCAAAAACATTTCGACAT ATTTATTTATTTTATTATTATGAGAATCA AATGTGAGACC	122	Kanamycin	*atpB* 5'UTR +10 processed -90, processed 5'UTR has more efficient translation	(Yukawa et al., 2007)
5' UTR	*atpB* 5'UTR TCR processed -90 (functional)	GGTCTCATACTGAATTAACCGATCGAC GTGCTAGCGGACATTTATTTTGAATTC GATAATTTTTGCAAAAACATTTCGACAT ATTTATTTATTTTATTATTATGAGAATCA ATCCTACTACTTCTGGTTCTGGGGTTT CCACGAATGTGAGACC	154	Kanamycin	*atpB* 5'UTR TCR processed -90, includes several AA past start codon (neomycin phosphotransferase II accumulated to 7% total soluble protein in leaf)	(Kuroda and Maliga, 2001)
5' UTR	*atpH* 5'UTR	GGTCTCATACTGATATTGACTAGTGCT AGTTATATATGAAATGAACTAAAGATAT ATTTTACTACTCTAGGGGATTCTAATAG ACTAGAAGTCCTTCCGGCCCCTTGTGA CTGTGAATTGAATGAATAAACGGATGA AATCAAGAAATAATTCAACTAACAGTTC GAACCAAGAAATGGAAGAACGAAAGTC GTATGGGTTCACAAAGACTCTGTGGCT AAAAAGTAAAAAAGATATATCGAAGTA GTTCTGATGATTCAATAATCTTATTACT TCAATCCGAAGTTCTTAGTTACTTCGAC TGGATGAGTCCTAGCGAGGGAATAATT AAGTCATAACTCATTGGTTGATTGTATC ATTAACCATTTCTTTTTTTTGGTACGAG GAACTTATCAATGTGAGACC	405	Kanamycin	5'UTR of *atpH*	(Yukawa et al., 2007)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
5' UTR	*atpH* 5'UTR processed -45	GGTCTCATACTATTGTATCATTAACCATTTCTTTTTTTTGGTACGAGGAACTTATCAATGTGAGACC	67	Kanamycin	Processed 5'UTR of *atpH*, no difference in translation from full length	(Yukawa et al., 2007)
5' UTR	*atpI* 5'UTR	GGTCTCATACTAGATAAATAACTTACTTCAGAAATCCCTTATGGAATCGCTTACTATTTCTGAATTTCAAAAAAGAGATAAAAATAGCTGGGGATATTATGTGATTTATTAGTATTCTAAATCTTAGTTGGTATTCAAAATATCCGATTCAAGTAGACAAAGAGATGGTTGAATCAAAAAATTTTGTTTAAAGTTCAATTTTTTCAGAGGGCAAGGCAATAATGTGAGACC	231	Kanamycin	*atpI* 5'UTR, RBS and upstream elements required for translation	(Baecker et al., 2009)
5' UTR	*psbN* 5'UTR	GGTCTCATACTCTATCATAGATCCATTGGGGTCTTGAAATTTTATAAATATGATAAGAAATGTGAGACC	69	Kanamycin	5'UTR of *psbN*	(Kuroda and Sugiura, 2014)
5' UTR	*psbN* 5'UTR processed -39	GGTCTCATACTGATCCATTGGGGTCTTGAAATTTTATAAATATGATAAGAAATGTGAGACC	61	Kanamycin	*psbN* 5'UTR processed -39, 5-fold higher translation efficiency than unprocessed	(Kuroda and Sugiura, 2014)
5' UTR	*psbB* 5'UTR	GGTCTCATACTGCTTCTCTTTGTTCCTACGAACAGAATTGTTCCATTATTACCAACAGAATAGAACACCCTTGTTCGGAAATAATCGACTGAACAAGAGTGGTCCATAGGATAGTCATATTATAGTCTTTTCCAATGCAATAAAGTTACGTAGTGTCTATTTATCTTTGATATAAGGGGTATTTCCAATGTGAGACC	197	Kanamycin	5'UTR of *psbB*	(Yukawa et al., 2007)
5' UTR	*psbB* 5'UTR processed -54	GGTCTCATACTCAATGCAATAAAGTTACGTAGTGTCTATTTATCTTTGATATAAGGGGTATTTCCAATGTGAGACC	76	Kanamycin	*psbB* 5'UTR processed -54, 2-fold higher translation efficiency than full length UTR	(Yukawa et al., 2007)
5' UTR	*ndhD* 5'UTR	GGTCTCATACTTACGTTCGAGAAAACTCTACTTGAATCCATTTAATTTTTTTTACCGACAAACCTGTGCTCGAAAATCACAATATTTTGAGCACGGGTTTTTATGGTCCAAGTGTATCTTGTCTTTACTAATGTGAGACC	140	Kanamycin	5'UTR of *ndhD*	(Hirose and Sugiura, 1997)
5' UTR	*rps2* 5'UTR	GGTCTCATACTTTTTGAAAAAGAATAAAAAAAAAAGGGGGGGTGTGGAGAGAAAATGTGAGACC	64	Kanamycin	5'UTR of *rps2*, SD in inhibitory region	(Plader and Sugiura, 2003)
5' UTR	*rps2* 5'UTR mAAGA	GGTCTCATACTTTTTGAAAAAGAATAAAAAAAAAAGGGGGGGTGTAAGAAGAAAATGTGAGACC	64	Kanamycin	5'UTR mAAGA of *rps2*, several fold increase in translation efficiency over native	(Plader and Sugiura, 2003)
5' UTR	*rps2* 5'UTR mCCUC	GGTCTCATACTTTTTGAAAAAGAATAAAAAAAAAAGGGGGGGTGTCCTCAGAAAATGTGAGACC	64	Kanamycin	5'UTR mCCUC of *rps2*, several fold increase in translation efficiency over native	(Plader and Sugiura, 2003)
5' UTR	*rps2* 5'UTR mAAGA (-20 to -10 deletion)	GGTCTCATACTTTTTGAAAAAGAATAAAAAAAAAGTAAGAAGAAAATGTGAGACC	54	Kanamycin	*rps2* 5'UTR mAGGA -20 to -10 deletion, >20-fold increase in translation efficiency over native	(Plader and Sugiura, 2003)
5' UTR	*rps2* 5'UTR mCCU	GGTCTCATACTTTTTGAAAAAGAATAAAAAAAAAGTCCTCAGAAAATGTGAGACC	54	Kanamycin	*rps2* 5'UTR mCCUC -20 to -10 deletion, >20-fold increase in	(Plader and Sugiura, 2003)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
	C (-20 to -10 deletion)				translation efficiency over native	
5' UTR	*psbH* 5'UTR + start codon + 1AA	GGTCTCATACTACAACGGAATGGTATACAAAGTCAACAGATCTCAATGAATACAATAGGATTTATGGCTAATGTGAGACC	80	Kanamycin	*psbH* 5'UTR + start + 1 AA, start codon included in secondary structure formation	(Hammani et al., 2012)
5' UTR	*psbA* 5'UTR (SNM)	GGTCTCATACTAATAAAAAGCCTTCCAATTGTAATTTTGATTTGTTACAATCTAGTGTGCTTGGGAGTCCCTGATGATTAAATAAACCAAGATTTTACCAATGTGAGACC	110	Kanamycin	5'UTR (SNM) of *psbA*, 1.8-2.5-fold decrease in translation efficiency from native	(Zou et al., 2003)
5' UTR	*psbA* 5'UTR (SEM)	GGTCTCATACTAATAAAAAGCCTTCCATCGTCTACTTTGATTTGTAGACGACTAGTGTGCTTGGGAGTCCCTGATGATTAAATAAACCAAGATTTTACCAATGTGAGACC	110	Kanamycin	5'UTR (SEM) of *psbA*, 1.8-2.5-fold decrease in translation efficiency from native	(Zou et al., 2003)
5' UTR	*psbA* 5'UTR (delta SL)	GGTCTCATACTAATAACTAGTGTGCTTGGGAGTCCCTGATGATTAAATAAACCAAGATTTTACCAATGTGAGACC	75	Kanamycin	5'UTR (delta SL) of *psbA*, 21-fold decrease in mRNA levels and 1.8-2.5-fold decrease in translation efficiency from native	(Zou et al., 2003)
5' UTR	*psbA* 5'UTR (delta 5'SL)	GGTCTCATACTCTAGTGTGCTTGGGAGTCCCTGATGATTAAATAAACCAAGATTTTACCAATGTGAGACC	70	Kanamycin	5'UTR (delta 5'SL) of *psbA*, 6-fold decrease in translation efficiency from native	(Zou et al., 2003)
5' UTR	*psbA-rbcL* chimeric 5'UTR (R delta 5'SL)	GGTCTCATACTATGTATTTGGCAAATCAAATACCATGCTAGTGTGCTTGGGAGTCCCTGATGATTAAATAAACCAAGATTTTACCAATGTGAGACC	96	Kanamycin	Chimeric 5'UTR of *psbA-rbcL* (R delta5'SL), 1.5-fold decrease in translation efficiency from native, increased stability of mRNA in dark	(Zou et al., 2003)
5' UTR	*psbA-rbcL* chimeric 5'UTR (R'SL WT)	GGTCTCATACTATGTATTTGGCAAATCAAATACCATGGTCTAATAATCAAACATTCTGATTAGTTGATAATCTAGTGTGCTTGGGAGTCCCTGATGATTAAATAAACCAAGATTTTACCAATGTGAGACC	130	Kanamycin	Chimeric 5'UTR of *psbA-rbcL* (R'SLWT), 1.5-fold decrease in translation efficiency from native, increased stability of mRNA in dark	(Zou et al., 2003)
5' UTR	*rps14* 5'UTR	GGTCTCATACTTATTTTTAATGTGTGTATCTGCGATAATCTCATTTCTTTCGACGGAGAGGGGGTCCACCTTCTTCTATTTCTACATCTAGGATTCGACTTGTATCATGGATACTAATAGGAATTCAACCATTAATGTGAGACC	144	Kanamycin	5'UTR of *rps14*	(Hirose et al., 1998)
5' UTR	*rbcL* 5'UTR (full)	GGTCTCATACTATGTATTTGGCAAATCAAATACCATGGTCTAATAATCAAACATTCTGATTAGTTGATAATATTAGTATTAGTTGGAAATTTTGTGAAAGATTCCTGTGAAAAGTTTCATTAACACGGAATTCGTGTCGAGTAGACCTTGTTGTTGTGAGAATTCTTAATTCATGAGTTGTAGGGAGGGATTTAATGTGAGACC	204	Kanamycin	Full length 5'UTR of *rbcL*	(Yukawa et al., 2007)
5' UTR	*rbcL* 5'UTR (proce	GGTCTCATACTGTCGAGTAGACCTTGTTGTTGTGAGAATTCTTAATTCATGAGTTGTAGGGAGGGATTTAATGTGAGACC	80	Kanamycin	Processed 5'UTR of *rbcL*	(Yukawa et al., 2007)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
	ssed)					
5' UTR	*rbcL* 5'UTR + 14 AA	GGTCTCATACTGTCGAGTAGACCTTGT TGTTGTGAGAATTCTTAATTCATGAGTT GTAGGGAGGGATTTATGTCACCACAAA CAGAGACTAAAGCAAGTGTTGGATTCA AAAATGTGAGACC	122	Kanamycin	5'UTR +14 AA of *rbcL* (neomycin phosphotransferase II accumulated to 11% total soluble protein in leaf)	(Kuroda and Maliga, 2001)
5' UTR	Synthetic RBS	GGTCTCATACTCTCGAGAATTCAGTTG TAGGGAGGGATCCATGGAATGTGAGA CC	55	Kanamycin	Synthetic RBS	(Kuroda and Maliga, 2001)
5' UTR	T7G10 leader	GGTCTCATACTGGGAGAGCACAACGG TTTCCCACTAGAAATAATTTTGTTTAAC TTTAAGAAGGAGATATACATAATGTGA GACC	85	Kanamycin	Leader from T7G10	(Kuroda and Maliga, 2001)
5' UTR	T7G10 leader + Ec DB	GGTCTCATACTGGGAGAGCACAACGG TTTCCCACTAGAAATAATTTTGTTTAAC TTTAAGAAGGAGATATACATATGGCAA GCATGACTGGTGGACAGAATGTGAGA CC	109	Kanamycin	Leader from T7G10 with Ec downstream box	(Kuroda and Maliga, 2001)
5' UTR	*rrn16* 5'UTR (native)	GGTCTCATACTGCGAACTCCGGGCGA ATACGAAGCGCATGGATACAAGTTATG CCTTGGAATGAAAGACAATTCCGAATC CGCTTTGTCTACGAACAAGGAAGCTAT AAGTAATGCAACTATGAAAATGTGAGA CC	136	Kanamycin	Native 5'UTR of *rrn16*	(Allison et al., 1996)
5' UTR	*rrn16* 5'UTR (+37)	GGTCTCATACTGCGAACTCCGGGCGA ATACGAAGCGCATGGATACAAGAATGT GAGACC	59	Kanamycin	5'UTR (+37) of *rrn16*	(Allison et al., 1996)
5' UTR	*psbA* 5'UTR	GGTCTCATACTAATAAAAAGCCTTCCAT TTTCTATTTTGATTTGTAGAAAACTAGT GTGCTTGGGAGTCCCTGATGATTAAAT AAACCAAGATTTTACCAATGTGAGACC	110	Kanamycin	5'UTR of *psbA*	(Staub and Maliga, 1994)
5' UTR	Synthetic RBS	GGTCTCATACTAAATTCGATAGAGATG AAATTGGAGCTCTAGAGAATTCAGTTG TAGGGAGGGATCCAATGTGAGACC	78	Kanamycin	Synthetic RBS	(Allison and Maliga, 1995)
5' UTR	*psbC* 5'UTR	GGTCTCATACTGTTATTTGTACCAGTAA CCGGTTTATGGATGAGTGCTCTTGGAG TAGTCGGTCTAGCCCTGAACCTACGTG CCTATGACTTCGTTTCTCAGGAAATTC GCGCAGCGGAAGATCCTGAATTTGAG ACTTTCTACACCAAAAATATTCTCTTAA ACGAAGGTATTCGCGCTTGGATGGCG GCTCAAGATCAGCCTCATGAAAACCTT ATATTCCCTGAGGAGGTTCTACCACAA TGTGAGACC	252	Kanamycin	5'UTR of *psbC*, the native transcription start is GUG, but edited to AUG	(Kuroda et al., 2007)
5' UTR	*accD* 5'UTR	GGTCTCATACTAAGTGTTCCCCCAGAT TCAGAACTTTTTTTCAATACTCACAATC CTTATTAGTTAATAATCCTAGTGATTGG ATTTCTATGCTTAGTCTGATAGGAAATA AGATATTCAAATAAATAATTTTATAGCG AATGACTATTCATCTATTGTATTTTCAT GCAAATAGGGGGCAAGAAAACTCTAAT GTGAGACC	202	Kanamycin	5'UTR of *accD*, doubles protein accumulation in chromoplasts	(Caroca et al., 2013)
5' UTR	*clpP* 5'UTR	GGTCTCATACTGTTTCCACCTCAAAGT GAAATATAGTATTTAGTTCTTTCTTTCA TTTAAATGTGAGACC	70	Kanamycin	5'UTR of *clpP*	(Kuroda and Maliga, 2002)
GOI	*mEmerald*	GGTCTCAAATGGTGAGCAAGGGCGAG GAGCTGTTCACCGGGGTGGTGCCCAT CCTGGTCGAGCTGGACGGCGACGTAA ACGGCCACAAGTTCAGCGTGTCCGGC GAGGGCGAGGGCGATGCCACCTACGG CAAGCTGACCCTGAAGTTCATCTGCAC CACCGGCAAGCTGCCCGTGCCCTGGC CCACCCTCGTGACCACCTTGACCTACG	739	Kanamycin	Gene encoding mEmerald fluorescent protein (including stop codon)	(Cubitt et al., 1999)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		GCGTGCAGTGCTTCGCCCGCTACCCC GACCACATGAAGCAGCACGACTTCTTC AAGTCCGCCATGCCCGAAGGCTACGT CCAGGAGCGCACCATCTTCTTCAAGGA CGACGGCAACTACAAGACCCGCGCCG AGGTGAAGTTCGAGGGCGACACCCTG GTGAACCGCATCGAGCTGAAGGGCAT CGACTTCAAGGAGGACGGCAACATCC TGGGGCACAAGCTGGAGTACAACTAC AACAGCCACAAGGTCTATATCACCGCC GACAAGCAGAAGAACGGCATCAAGGT GAACTTCAAGACCCGCCACAACATCGA GGACGGCAGCGTGCAGCTCGCCGACC ACTACCAGCAGAACACCCCCATCGGC GACGGCCCCGTGCTGCTGCCCGACAA CCACTACCTGAGCACCCAGTCCAAGCT GAGCAAAGACCCCAACGAGAAGCGCG ATCACATGGTCCTGCTGGAGTTCGTGA CCGCCGCCGGGATCACTCTCGGCATG GACGAGCTGTACAAGTAAGCTTTGAGA CC				
GOI	*pporRFP*	GGTCTCAAATGGCTCTTTCAAAGCAAA GTGGGGTCAAAGATGTAATGAACACCG AGCTTCATATGGACGGGATCGTCAATG GACACCCCTTTGAGATAAAAGGAAAAG GAAAGGGAAACCCGTACAAGGGTGTG CAGACCATGAAGCTTACAGTCATTAAG GGTGCGCCTTTGCCATTTTCTATTGAC ATTTTGCTGCCTCAACACATGTATGGA AGCAAGCCATTTATTAAGTATCCTGAG AGTATCCCAGACTACATCAAGTTGTCA TTTCCCGAGGGAATCACATGGGAAAG GTCCATGACCTTTGAAGATGGTGCAGT GTGCACTGTCTCTAACGACTCCAGGCT CGATGGCGACTCTTTCATCTACGAAGT CAGGTTTCTTGGCGTGAACTTTCCCCG AGATGGACCTGTTATGCAGAAGAAAAC GCGAGGCTGGGACCCGTCCACAGAGA GACTGTATGAGTGTGGTGGGTGGCAG AGAGGAGATGTCCACATGGCCTTGAA GTTGGAGAACGGTGGCCATTATACGTG CGACTTCAAAACTACTTACAAATCAAAG AAGGGCTTGAAGGTGCCACCGTATCA CTTCGTTGACCACAAACTAGATTTACT GAGCCACAACACCGATGGTGCTACCTT TGAAGAGTTTGAACAACGAGAAATTGC ACATGCACATCTTTCTAACTTACCGGTA GCCTAAGCTTTGAGACC	715	Kanamycin	Gene encoding pporRFP fluorescent protein (including stop codon)	(Alieva et al., 2008)
GOI	*aadA*	GGTCTCAAATGGCAGAAGCGGTGATC GCCGAAGTATCGACTCAACTATCAGAG GTAGTTGGCGTCATCGAGCGCCATCTC GAACCGACGTTGCTGGCCGTACATTTG TACGGCTCCGCAGTGGATGGCGGCCT GAAGCCACACAGTGATATTGATTTGCT GGTTACGGTGACCGTAAGGCTTGATGA AACAACGCGGCGAGCTTTGATCAACGA CCTTTTGGAAACTTCGGCTTCCCCTGG AGAGAGCGAGATTCTCCGCGCTGTAG AAGTCACCATTGTTGTGCACGACGACA TCATTCCGTGGCGTTATCCAGCTAAGC GCGAACTGCAATTTGGAGAATGGCAG CGCAATGACATTCTTGCAGGTATCTTC GAGCCAGCCACGATCGACATTGATCTG GCTATCTTGCTGACAAAAGCAAGAGAA CATAGCGTTGCCTTGGTAGGTCCAGC	811	Kanamycin	Gene encoding streptomycin 3''-adenylyltransferase (including stop codon)	(Svab and Maliga, 1993)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		GGCGGAGGAACTCTTTGATCCGGTTC CTGAACAGGATCTATTTGAGGCGCTAA ATGAAACCTTAACGCTATGGAACTCGC CGCCCGACTGGGCTGGCGATGAGCGA AATGTAGTGCTTACGTTGTCCCGCATT TGGTACAGCGCAGTAACCGGCAAAATC GCGCCGAAGGATGTCGCTGCCGACTG GGCAATGGAGCGCCTGCCGGCCCAGT ATCAGCCCGTCATACTTGAAGCTAGAC AGGCTTATCTTGGACAAGAAGAAGATC GCTTGGCCTCGCGCGCAGATCAGTTG GAAGAATTTGTCCACTACGTGAAAGGC GAGATCACCAAGGTAGTCGGCAAATAA GCTTTGAGACC				
GOI	*aphA6*	GGTCTCAAATGGAATTGCCCAATATTA TTCAACAATTTATCGGAAACAGCGTTTT AGAGCCAAATAAAATTGGTCAGTCGCC ATCGGATGTTTATTCTTTTAATCGAAAT AATGAAACTTTTTTTCTTAAGCGATCTA GCACTTTATATACAGAGACAACATACA GTGTCTCTCGTGAAGCGAAAATGTTGA GTTGGCTCTCTGAGAAATTAAAGGTGC CTGAACTCATCATGACTTTTCAGGATG AGCAGTTTGAATTCATGATCACTAAAG CGATCAATGCAAAACCAATTTCAGCGC TTTTTTTAACAGACCAAGAATTGCTTGC TATCTATAAGGAGGCACTCAATCTGTT AAATTCAATTGCTATTATTGATTGTCCA TTTATTTCAAACATTGATCATCGGTTAA AAGAGTCAAAATTTTTTATTGATAACCA ACTCCTTGACGATATAGATCAAGATGA TTTTGACACTGAATTATGGGGAGATCA TAAAACTTACCTAAGTCTATGGAATGA GTTAACCGAGACTCGTGTTGAAGAAAG ATTGGTTTTTTCTCATGGCGATATCACG GATAGTAATATTTTTATAGATAAATTCA ATGAAATTTATTTTTTAGATCTTGGTCG TGCTGGGTTAGCAGATGAATTTGTAGA TATATCCTTTGTTGAACGTTGCCTAAGA GAGGATGCATCGGAGGAAACTGCGAA AATATTTTTAAAGCATTTAAAAAATGAT AGACCTGACAAAAGGAATTATTTTTTAA AACTTGATGAATTGAATTGAGCTTTGA GACC	799	Kanamycin	Gene encoding 3'-aminoglycoside phosphotransferase type VI (including stop codon)	(Huang et al., 2002)
GOI	*nptII*	GGTCTCAAATGGTTGAACAAGATGGAT TGCACGCAGGTTCTCCGGCCGCTTGG GTGGAGAGGCTATTCGGCTATGACTG GGCACAACAGACAATCGGCTGCTCTG ATGCCGCCGTGTTCCGGCTGTCAGCG CAGGGGCGCCCGGTTCTTTTTGTCAAG ACCGACCTGTCCGGTGCCCTGAATGA ACTGCAGGACGAGGCAGCGCGGCTAT CGTGGCTGGCCACGACGGGCGTTCCT TGCGCAGCTGTGCTCGACGTTGTCACT GAAGCGGGAAGGGACTGGCTGCTATT GGGCGAAGTGCCGGGGCAGGATCTCC TGTCATCTCACCTTGCTCCTGCCGAGA AAGTATCCATCATGGCTGATGCAATGC GGCGGCTGCATACGCTTGATCCGGCT ACCTGCCCATTCGACCACCAAGCGAAA CATCGCATCGAGCGAGCACGTACTCG GATGGAAGCCGGTCTTGTCGATCAGG ATGATCTGGACGAAGAGCATCAGGGG CTCGCGCCAGCCGAACTGTTCGCCAG GCTCAAGGCGCGCATGCCCGACGGCG	814	Kanamycin	Gene encoding neomycin phosphotransferase II (including stop codon)	(Carrer et al., 1993)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		AGGATCTCGTCGTGACTCATGGCGATG CCTGCTTGCCGAATATCATGGTGGAAA ATGGCCGCTTTTCTGGATTCATCGACT GTGGCCGGCTGGGTGTGGCGGACCG CTATCAGGACATAGCGTTGGCTACCCG TGATATTGCTGAAGAGCTTGGCGGCGA ATGGGCTGACCGCTTCCTCGTGCTTTA CGGTATCGCCGCTCCCGATTCGCAGC GCATCGCCTTCTATCGCCTTCTTGACG AGTTCTTCTGAGCTTTGAGACC				
GOI	*Myc aadA-GFP* Fusion	GGTCTCAAATGGCAAGCATGACTGGTG GACAGGCTAGCGAACAAAAACTCATTT CTGAAGAGGACTTGCCTAGCGAAGCG GTGATCGCCGAAGTATCGACTCAACTA TCAGAGGTAGTTGGCGTCATCGAGCG CCATCTCGAACCGACGTTGCTGGCCG TACATTTGTACGGCTCCGCAGTGGATG GCGGCCTGAAGCCACACAGTGATATT GATTTGCTGGTTACGGTGACCGTAAGG CTTGATGAAACAACGCGGCGAGCTTTG ATCAACGACCTTTTGGAAACTTCGGCT TCCCCTGGAGAGAGCGAGATTCTCCG CGCTGTAGAAGTCACCATTGTTGTGCA CGACGACATCATTCCGTGGCGTTATCC AGCTAAGCGCGAACTGCAATTTGGAGA ATGGCAGCGCAATGACATTCTTGCAGG TATCTTCGAGCCAGCCACGATCGACAT TGATCTGGCTATCTTGCTGACAAAAGC AAGAGAACATAGCGTTGCCTTGGTAGG TCCAGCGGCGGAGGAACTCTTTGATC CGGTTCCTGAACAGGATCTATTTGAGG CGCTAAATGAAACCTTAACGCTATGGA ACTCGCCGCCCGACTGGGCTGGCGAT GAGCGAAATGTAGTGCTTACGTTGTCC CGCATTTGGTACAGCGCAGTAACCGG CAAAATCGCGCCGAAGGATGTCGCTG CCGACTGGGCAATGGAGCGCCTGCCG GCCCAGTATCAGCCCGTCATACTTGAA GCTAGACAGGCTTATCTTGGACAAGAA GAAGATCGCTTGGCCTCGCGCGCAGA TCAGTTGGAAGAATTTGTCCACTACGT GAAAGGCGAGATCACCAAGGTAGTGG GCAAAGAACTTGCAGTTGAAGGAAAAT TGGAGGTCGCCATGGCTAGTAAAGGA GAAGAACTTTTCACTGGAGTTGTCCCA ATTCTTGTTGAATTAGATGGTGATGTTA ATGGGCACAAATTTTCTGTCAGTGGAG AGGGTGAAGGTGATGCAACATACGGA AAACTTACCCTTAAATTTATTTGCACTA CTGGAAAACTACCTGTTCCTTGGCCAA CACTTGTCACTACTTTCTCTTATGGTGT TCAATGCTTTTCAAGATACCCAGATCAT ATGAAGCGGCACGACTTCTTCAAGAGC GCCATGCCTGAGGGATACGTGCAGGA GAGGACCATCTCTTTCAAGGACGACG GGAACTACAAGACACGTGCTGAAGTCA AGTTTGAGGGAGACACCCTCGTCAACA GGATCGAGCTTAAGGGAATCGATTTCA AGGAGGACGGAAACATCCTCGGCCAC AAGTTGGAATACAACTACAACTCCCAC AACGTATACATCACGGCAGACAAACAA AAGAATGGAATCAAAGCTAACTTCAAA ATTAGACACAACATTGAAGATGGAAGC GTTCAACTAGCAGACCATTATCAACAA AATACTCCAATTGGCGATGGCCCTGTC	1621	Kanamycin	Gene encoding Myc-aadA-GFP fusion protein (including stop codon)	(Lutz et al., 2007)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		CTTTTACCAGACAACCATTACCTGTCC ACACAATCTGCCCTTTCGAAAGATCCC AACGAAAAGAGAGATCACATGGTCCTT CTTGAGTTTGTAACAGCTGCTGGGATT ACACATGGCATGGATGAACTATACAAA TAAGCTTTGAGACC				
GOI	*FLARE16-SaadA-GFP* fusion	GGTCTCAAATGGGGGCTAGCGAAGCG GTGATCGCCGAAGTATCGACTCAACTA TCAGAGGTAGTTGGCGTCATCGAGCG CCATCTCGAACCGACGTTGCTGGCCG TACATTTGTACGGCTCCGCAGTGGATG GCGGCCTGAAGCCACACAGTGATATT GATTTGCTGGTTACGGTGACCGTAAGG CTTGATGAAACAACGCGGCGAGCTTTG ATCAACGACCTTTTGGAAACTTCGGCT TCCCCTGGAGAGAGCGAGATTCTCCG CGCTGTAGAAGTCACCATTGTTGTGCA CGACGACATCATTCCGTGGCGTTATCC AGCTAAGCGCGAACTGCAATTTGGAGA ATGGCAGCGCAATGACATTCTTGCAGG TATCTTCGAGCCAGCCACGATCGACAT TGATCTGGCTATCTTGCTGACAAAAGC AAGAGAACATAGCGTTGCCTTGGTAGG TCCAGCGGCGGAGGAACTCTTTGATC CGGTTCCTGAACAGGATCTATTTGAGG CGCTAAATGAAACCTTAACGCTATGGA ACTCGCCGCCCGACTGGGCTGGCGAT GAGCGAAATGTAGTGCTTACGTTGTCC CGCATTTGGTACAGCGCAGTAACCGG CAAAATCGCGCCGAAGGATGTCGCTG CCGACTGGGCAATGGAGCGCCTGCCG GCCCAGTATCAGCCCGTCATACTTGAA GCTAGACAGGCTTATCTTGGACAAGAA GAAGATCGCTTGGCCTCGCGCGCAGA TCAGTTGGAAGAATTTGTCCACTACGT GAAAGGCGAGATCACCAAGGTAGTGG GCAAAGAACTTGTTGAAGGAAAATTGG AGCTAGTAGAAGGTCTTAAAGTCGCCA TGGCTAGTAAAGGAGAAGAACTTTTCA CTGGAGTTGTCCCAATTCTTGTTGAATT AGATGGTGATGTTAATGGGCACAAATT TTCTGTCAGTGGAGAGGGTGAAGGTG ATGCAACATACGGAAAACTTACCCTTA AATTTATTTGCACTACTGGAAAACTACC TGTTCCTTGGCCAACACTTGTCACTAC TTTCTCTTATGGTGTTCAATGCTTTTCA AGATACCCAGATCATATGAAGCGGCAC GACTTCTTCAAGAGCGCCATGCCTGAG GGATACGTGCAGGAGAGGACCATCTC TTTCAAGGACGACGGGAACTACAAGAC ACGTGCTGAAGTCAAGTTTGAGGGAGA CACCCTCGTCAACAGGATCGAGCTTAA GGGAATCGATTTCAAGGAGGACGGAA ACATCCTCGGCCACAAGTTGGAATACA ACTACAACTCCCACAACGTATACATCA CGGCAGACAAACAAAAGAATGGAATCA AAGCTAACTTCAAAATTAGACACAACAT TGAAGATGGAAGCGTTCAACTAGCAGA CCATTATCAACAAAATACTCCAATTGGC GATGGCCCTGTCCTTTTACCAGACAAC CATTACCTGTCCACACAATCTGCCCTT TCGAAAGATCCCAACGAAAAGAGAGAT CACATGGTCCTTCTTGAGTTTGTAACA GCTGCTGGGATTACACATGGCATGGAT GAACTATACAAATAAGCTTTGAGACC	1582	Kanamycin	Gene encoding FLARE16S-aadA-GFP fusion protein (including stop codon)	(Khan and Maliga, 1999)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
GOI	*aac6-aph2*	GGTCTCAAATGGAAAATATTGTAGAAA ATGAAATTTGTATTCGAACTTTAATTGA TGATGATTTTCCTTTAATGTTAAAATGG TTAACTGATGAAAGAGTTTTAGAATTTT ATGGAGGTAGAGATAAAAAATATACTTT AGAATCTTTAAAAAAACATTATACTGAA CCTTGGGAAGATGAAGTATTTAGAGTA ATTATTGAATATAATAATGTTCCTATTG GTTATGGACAAATTTATAAAATGTATGA TGAATTATATACTGATTATCATTATCCT AAAACTGATGAAATTGTATATGGAATG GATCAATTTATTGGTGAACCTAATTATT GGTCTAAAGGTATTGGTACTCGATATA TTAAATTAATTTTTGAATTTTTAAAAAAA GAACGAAATGCTAATGCTGTAATTTTA GATCCTCATAAAAATAATCCACGAGCT ATTCGAGCTTATCAAAAATCTGGATTTA GAATTATTGAAGATTTACCTGAACATGA ATTACATGAAGGAAAAAAAGAAGATTG TTATTTAATGGAATATAGATATGATGAT AATGCTACTAATGTTAAAGCTATGAAAT ATTTAATTGAACATTATTTTGATAATTTT AAAGTAGATTCTATTGAAATTATTGGTT CTGGATATGATTCTGTTGCTTATTTAGT TAATAATGAATATATTTTTAAAACTAAAT TTTCTACTAATAAAAAAAAAGGTTATGC TAAAGAAAAAGCTATTTATAATTTTTTAA ATACTAATTTAGAAACTAATGTAAAAAT TCCTAATATTGAATATTCTTATATTTCTG ATGAATTATCTATTTTAGGATATAAAGA AATTAAAGGTACTTTTTTAACTCCTGAA ATTTATTCTACTATGTCTGAAGAAGAAC AAAATTTATTAAAACGAGATATTGCTTC TTTTTTAAGACAAATGCATGGTTTAGAT TATACTGATATTTCTGAATGTACTATTG ATAATAAACAAAATGTATTAGAAGAATA TATTTTATTACGAGAAACTATTTATAAT GATTTAACTGATATTGAAAAAGATTATA TTGAATCTTTTATGGAAAGATTAAATGC TACTACTGTTTTTGAAGGTAAAAAATGT TTATGTCATAATGATTTTTCTTGTAATC ATTTATTATTAGATGGAAATAATAGATT AACTGGAATTATTGATTTTGGTGATTCT GGAATTATTGATGAATATTGTGATTTTA TTTATTTATTAGAAGATTCTGAAGAAGA AATTGGTACTAATTTTGGAGAAGATATT TTACGAATGTATGGAAATATTGATATTG AAAAAGCTAAAGAATATCAAGATATTGT AGAAGAATATTATCCTATTGAAACTATT GTATATGGAATTAAAAATATTAAACAAG AATTTATTGAAAATGGTCGAAAAGAAAT TTATAAAAGAACTTATAAAGATTAAGCT TTGAGACC	1462	Kanamycin	aminoglycoside resistance gene encoding for a bifunctional aminoglycoside acetyltransferase/phosphotransferase (including stop codon)	(Tabatabaei et al., 2017)
GOI	*Lux-C* codon opt for tobacco chl stop	GGTCTCAAATGATTAAAAAAATTCCTAT GATTATTGGAGGAGTAGTACAAAATAC TTCTGGATATGGAATGAGAGAATTAAC TTTAAATAATAATAAAGTAAATATTCCTA TTATTACTCAATCTGATGTAGAAGCTAT TCAATCTTTAAATATTGAAAATAAATTAA CTATTAATCAAATTGTAAATTTTTTATAT ACTGTAGGACAAAAATGGAAATCTGAA ACTTATTCTAGAAGATTAACTTATATTA GAGATTTAATTAAATTTTTAGGATATTC TCAAGAAATGGCTAAATTAGAAGCTAA TTGGATTTCTATGATTTTATGTTCTAAA	1456	Kanamycin	*Lux-C* codon optimized for *Nicotiana tabacum* plastome (including stop codon)	(Lee et al., 1991)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		TCTGCTTTATATGATATTGTAGAAAATG ATTTATCTTCTAGACATATTATTGATGA ATGGATTCCTCAAGGAGAATGTTATGT AAAAGCTTTACCTAAAGGAAAATCTGTA CATTTATTAGCTGGAAATGTACCTTTAT CTGGAGTAACTTCTATTTTAAGAGCTAT TTTAACTAAAAATGAATGTATTATTAAA ACTTCTTCTGCTGATCCTTTTACTGCTA CTGCTTTAGTAAATTCTTTTATTGATGT AGATGCTGAACATCCTATTACTAGATCT ATTTCTGTAATGTATTGGTCCCATTCTG AAGATTTAGCTATTCCTAAACAAATTAT GTCTTGTGCTGATGTAGTAATTGCTTG GGGAGGAGATGATGCTATTAAATGGG CTACTGAACATGCTCCTTCTCATGCTG ATATTTTAAAATTTGGACCTAAAAAATC TATTTCTATTGTAGATAATCCTACTGAT ATTAAAGCTGCTGCTATTGGAGTAGCT CATGATATTTGTTTTTATGATCAACAAG CTTGTTTTTCTACTCAAGATATTTATTAT ATTGGAGATTCTATTGATATTTTTTTTG ATGAATTAGCTCAACAATTAAATAAATA TAAAGATATTTTACCTAAAGGAGAAAGA AATTTTGATGAAAAAGCTGCTTTTTCTT TAACTGAAAGAGAATGTTTATTTGCTAA ATATAAAGTACAAAAAGGAGAATCTCA ATCTTGGTTATTAACTCAATCTCCTGCT GGATCTTTTGGAAATCAACCTTTATCTA GATCTGCTTATATTCATCAAGTAAATGA TATTTCTGAAGTAATTCCTTTTGTACAT AAAGCTGTAACTCAAACTGTAGCTATT GCTCCTTGGGAATCTTCTTTTAAATATA GAGATATTTTAGCTGAACATGGAGCTG AAAGAATTATTGAAGCTGGAATGAATA ATATTTTTAGAGTAGGAGGAGCTCATG ATGGAATGAGGCCTTTACAAAGATTAG TAAATTATATTTCTCATGAAAGACCTTC TACTTATACTACTAAAGATGTATCTGTA AAAATTGAACAAACTAGATATTTAGAAG AAGATAAATTTTTAGTATTTGTACCTTA GGCTTTGAGACC				
GOI	*Lux-D* codon opt for tobacco chl stop	GGTCTCAAATGGAAAATACTCAACATT CTTTACCTATTGATCATGTAATTGATAT TGGAGATAATAGATATATTAGAGTATG GGAAACTAAACCTAAAAATAAAGAAAC TAAAAGAAATAATACTATTGTAATTGCT TCTGGATTTGCTAGAAGAATGGATCAT TTTGCTGGATTAGCTGAATATTTAGCTA ATAATGGATTTAGAGTAATTAGATATGA TTCTTTAAATCATGTAGGATTATCTTCT GGAGAAATTAAACAATTTTCTATGTCTG TAGGAAAACATTCTTTATTAACTGTAAT TGATTGGTTAAAAGAAAGAAATATTAAT AATATTGGATTAATTGCTTCTTCTTTAT CTGCTAGAATTGCTTATGAAGTAGCTG CTGAAATTGATTTATCTTTTTTAATTACT GCTGTAGGAGTAGTAAATTTAAGATCT ACTTTAGAAAAAGCTTTAAAATATGATT ATTTACAAATGGAAGTAAATACTATTCC TGAAGATTTAATTTTTGAAGGACATAAT TTAGGATCTAAAGTATTTGTAACTGATT GTTTTGAAAATAATTGGGATTCTTTAGA TTCTACTATTAATAAAATTTGTGAATTA GATATTCCTTTTATTGCTTTTACTTCTG ATGGAGATGATTGGGTATGTCAACATG	967	Kanamycin	*Lux-D* codon optimized for *Nicotiana tabacum* plastome (including stop codon)	(Lee et al., 1991)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		AAGTAAAACATTTAGTATCTAATGTAAA ATCTGATAAAAAAAAAATTTATTCTTTA GTAGGATCTTCTCATGATTTAGGAGAA AATTTAGTAGTATTAAGAAATTTTTATC AATCTATGACTAAAGCTGCTGTATCTTT AGATAGACAATTAGTAGAATTAGTAGAT GAAATTATTGAACCTAATTTTGAAGATT TAACTGTAATTACTGTAAATGAAAGAAG ATTAAAAAAATAAAATTGAAAATGAAATT ATTAATAGATTAGCTGATAGAGTATTAG CTTCTGTATAAGCTTTGAGACC				
GOI	*Lux-A* codon opt for tobacco chl stop	GGTCTCAAATGAAAATTTCTAATATTTG TTTTTCTTATCAACCTCCTGGAGAATCT CATCAAGAAGTAATGGAAAGATTTATTA GATTAGGAGTAGCTTCTGAAGAATTAA ATTTTGATGGATTTTATACTTTAGAACA TCATTTTACTGAATTTGGAATTACTGGA AATTTATATATTGCTTGTGCTAATATTTT AGGAAGAACTAAAAGAATTCAAGTAGG AACTATGGGAATTGTATTACCTACTGAA CATCCTGCTAGACATGTAGAATCTTTAT TAGTATTAGATCAATTATCTAAAGGAAG ATTTAATTATGGAACTGTAAGAGGATTA TATCATAAAGATTTTAGAGTATTTGGAA CTTCTCAAGAAGATTCTAGAAAAACTG CTGAAAATTTTTATTCTATGATTTTAGAT GCTTCTAAAACTGGAGTATTACATACT GATGGAGAAGTAGTAGAATTTCCTGAT GTAAATGTATATCCTGAAGCTTATTCTA AAAAACAACCTACTTGTATGACTGCTG AATCTTCTGAAACTATTACTTATTTAGC TGAAAGAGGATTACCTATGGTATTATCT TGGATTATTCCTGTATCTGAAAAAGTAT CTCAAATGGAATTATATAATGAAGTAGC TGCTGAACATGGACATGATATTAATAAT ATTGAACATATTTTAACTTTTATTTGTTC TGTAAATGAAGATGGAGAAAAAGCTGA TTCTGTATGTAGAAATTTTTTAGAAAAT TGGTATGATTCTTATAAAAATGCTACTA ATATTTTTAATGATTCTAATCAAACTAG AGGATATGATTATTTAAAAGCTCAATGG AGAGAATGGGTAATGAAAGGATTAGCT GATCCTAGAAGAAGATTAGATTATTCTA ATGAATTAAATCCTGTAGGAACTCCTG AAAGATGTATTGAAATTATTCAATCTAA TATTGATGCTACTGGAATTAAACATATT ACTGTAGGATTTGAAGCTAATGGATCT GAACAAGAAATTAGAGAATCTATGGAA TTATTTATGGAAAAAGTAGCTCCTCATT TAAAAGATCCTCAATAAGCTTTGAGAC C	1084	Kanamycin	*Lux-A* codon optimized for *Nicotiana tabacum* plastome (including stop codon)	(Lee et al., 1991)
GOI	*Lux-B* codon opt for tobacco chl stop	GGTCTCAAATGAATTTTGGATTATTTTT TTTAAATTTTCAACCTGAAGGAATGACT TCTGAAATGGTATTAGATAATATGGTAG ATACTGTAGCTTTAGTAGATAAAGATGA TTATCATTTTAAAAGAGTATTAGTATCT GAACATCATTTTTCTAAAAATGGAATTA TTGGAGAACCTTTAACTGCTATTTCTTT TTTATTAGGATTAACTAAAAGAATTGAA ATTGGATCTTTAAATCAAGTAATTACTA CTCATCATCCTGTAAGAATTGGAGAAC AAACTGGATTATTAGATCAAATGTCTTA TGGAAGATTTGTATTAGGATTATCTGAT TGTGTAAATGATTTTGAAATGGATTTTT TTAAAAGAAAAAGATCTTCTCAACAACA	1000	Kanamycin	*Lux-B* codon optimized for *Nicotiana tabacum* plastome (including stop codon)	(Lee et al., 1991)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		ACAATTTGAAGCTTGTTATGAAATTTTA AATGAAGCTTTAACTACTAATTATTGTC AAGCTGATGATGATTTTTTTAATTTTCC TAGAATTTCTGTAAATCCTCATTGTATT TCTGAAGTAAAACAATATATTTTAGCTT CTTCTATGGGAGTAGTAGAATGGGCTG CTAGAAAAGGATTACCTTTAACTTATAG ATGGTCTGATTCTTTAGCTGAAAAAGA AAAATATTATCAAAGATATTTAGCTGTA GCTAAAGAAAATAATATTGATGTATCTA ATATTGATCATCAATTTCCTTTATTAGT AAATATTAATGAAAATAGAAGAATTGCT AGAGATGAAGTAAGAGAATATATTCAA TCTTATGTATCTGAAGCTTATCCTACTG ATCCTAATATTGAATTAAGAGTAGAAGA ATTAATTGAACAACATGCTGTAGGAAA AGTAGATGAATATTATGATTCTACTATG CATGCTGTAAAAGTAACTGGATCTAAA AATTTATTATTATCTTTTGAATCTATGAA AAATAAAGATGATGTAACTAAATTAATT AATATGTTTAATCAAAAAATTAAAGATA ATTTAATTAAATAAGCTTTGAGACC				
GOI	*Lux-E* codon opt for tobacco chl stop	GGTCTCAAATGTCTACTTTATTAAATAT TGATGCTACTGAAATTAAAGTATCTACT GAAATTGATGATATTATTTTTACTTCTT CTCCTTTAACTTTATTATTTGAAGATCA AGAAAAAATTCAAAAAGAATTAATTTTA GAATCTTTTCATTATCATTATAATCATAA TAAAGATTATAAATATTATTGTAATATTC AAGGAGTAGATGAAAATATTCAATCTAT TGATGATATTCCTGTATTTCCTACTTCT ATGTTTAAATATTCTAGATTACATACTG CTGATGAATCTAATATTGAAAATTGGTT TACTTCTTCTGGAACTAAAGGAGTAAA ATCTCATATTGCTAGAGATAGACAATCT ATTGAAAGATTATTAGGATCTGTAAATT ATGGAATGAAATATTTAGGAGAATTTCA TGAACATCAATTAGAATTAGTAAATATG GGACCTGATAGATTTTCTGCTTCTAAT GTATGGTTTAAATATGTAATGTCTTTAG TACAATTATTATATCCTACTACTTTTACT GTAGAAAATGATGAAATTGATTTTGAAC AAACTATTTTAGCTTTAAAAGCTATTCA AAGAAAAGGAAAAGGAATTTGTTTAATT GGACCTCCTTATTTTATTTATTTATTAT GTCATTATATGAAAGAACATAATATTGA ATTTAATGCTGGAGCTCATATGTTTATT ATTACTGGAGGAGGATGGAAAACTAAA CAAAAAGAAGCTTTAAATAGACAAGATT TTAATCAATTATTAATGGAAACTTTTTCT TTATTTCATGAATCTCAAATTAGAGATA TTTTTAATCAAGTAGAATTAAATACTTG TTTTTTTGAAGATTCTTTACAAAGAAAA CATGTACCTCCTTGGGTATATGCTAGA GCTTTAGATCCTGTAACTTTAACTCCTG TAGAAGATGGACAAGAAGGATTAATGT CTTATATGGATGCTTCTTCTACTTCTTA TCCTACTTTTATTGTAACTGATGATATT GGAATTGTAAGACATTTAAAAGAACCT GATCCTTTTCAAGGAACTACTGTAGAA ATTGTAAGAAGATTAAATACTAGAGAAC AAAAAGGATGTTCTTTATCTATGGCTAC TTCTTTAAAATAAGCTTTGAGACC	1141	Kanamycin	*Lux-E* codon optimized for *Nicotiana tabacum* plastome (including stop codon)	(Lee et al., 1991)
GOI	*Lux-G* codon	GGTCTCAAATGATTTTTAATTGTAAAGT AAAAAAAGTAGAAGCTTCTGATTCTCAT	724	Kanamycin	*Lux-G* codon optimized for	(Lee et al., 1991)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
	opt for tobacco chl stop	ATTTATAAAGTATTTATTAAACCTGATA AATGTTTTGATTTTAAAGCTGGACAATA TGTAATTGTATATTTAAATGGAAAAAAT TTACCTTTTTCTATTGCTAATTGTCCTA CTTGTAATGAATTATTAGAATTACATGT AGGAGGATCTGTAAAAGAATCTGCTAT TGAAGCTATTTCTCATTTTATTAATGCT TTTATTTATCAAAAAGAATTTACTATTGA TGCTCCTCATGGAGATGCTTGGTTAAG AGATGAATCTCAATCTCCTTTATTATTA ATTGCTGGAGGAACTGGATTATCTTAT ATTAATTCTATTTTATCTTGTTGTATTTC TAAACAATTATCTCAACCTATTTATTTAT ATTGGGGAGTAAATAATTGTAATTTATT ATATGCTGATCAACAATTAAAAACTTTA GCTGCTCAATATAGAAATATTAATTATA TTCCTGTAGTAGAAAATTTAAATACTGA TTGGCAAGGAAAAATTGGAAATGTAAT TGATGCTGTAATTGAAGATTTTTCTGAT TTATCTGATTTTGATATTTATGTATGTG GACCTTTTGGAATGTCTAGAACTGCTA AAGATATTTTAATTTCTCAAAAAAAAGC TAATATTGGAAAAATGTATTCTGATGCT TTTTCTTATACTTAAGCTTTGAGACC	*Nicotiana tabacum* plastome (including stop codon)		
3' UTR	*psbA* 3'UTR	GGTCTCAGCTTTCCTGGCCTAGTCTAT AGGAGGTTTTGAAAAGAAAGGAGCAAT AATCATTTTCTTGTTCTATCAAGAGGGT GCTATTGCTCCTTTCTTTTTTTCGCTTG AGACC	115	Kanamycin	3'UTR of *psbA*	(Eibl et al., 1999)
3' UTR	*rbcL* 3'UTR	GGTCTCAGCTTAAACAGTAGACATTAG CAGATAAATTAGCAGGAAATAAAGAAG GATAAGGAGAAAGAACTCAAGTAATTA TCCTTCGTTCTCTTAATTGAATTGCAAT TAAACTCGGCCCAATCTTTTACTAAAAG GATTGAGCCGAATACCGCTTGAGACC	163	Kanamycin	3'UTR of *rbcL*	(Eibl et al., 1999)
3' UTR	*petD* 3'UTR	GGTCTCAGCTTATTTTTAAATTGATTCA ATTGTGAAATAACACGACATGTGTATCT AGGGAATAGTTTCTTCAAAGCGAATTC TCCCTAGATACATCTATTCAATTTAATT CTGAATTTATTTTGAATATATGATATATT AATATATTAATTGTGCTAAAGAGTTTCA ATCTATTTTCACTAAGTAAGTCCAATAG ATCGCTTGAGACC	209	Kanamycin	3'UTR of *petD*	(Tangphatsornruang et al., 2011)
3' UTR	*rpoA* 3'UTR	GGTCTCAGCTTATCTATTGGACTTACTT AGTGAAAATAGATTGAAACTCTTTAGCA CAATTAATATATTAATATATCATATATTC AAAATAAATTCAGAATTAAATTGAATAG ATGTATCTAGGGAGAATTCGCTTTGAA GAAACTATTCCCTAGATACACATGTCG TGTTATTTCACAATTGAATCAATTTAAA AATCGCTTGAGACC	209	Kanamycin	3'UTR of *rpoA*	(Tangphatsornruang et al., 2011)
3' UTR	*E.coli rrnB* 3'UTR	GGTCTCAGCTTGTAGAAACGCAAAAAG GCCATCCGTCAGGATGGCCTTCTGCTT AATTTGATGCCTGGCAGTTTATGGCGG GCGTCCTGCCCGCCACCCTCCGGGCC GTTGCTTCGCAACGTTCAAATCCGCTC CCGGCGGATTTGTCCTACTCAGGAGA GCGTTCACCGACAAACAACAGATAAAA CGAAAGGCCCAGTCTTTCGACTGAGC CTTTCGTTTTATTTGATGCGCTTGAGAC C	242	Kanamycin	3'UTR of *rrnB* from *E.coli*	(Tangphatsornruang et al., 2011)
3' UTR	*psaC* 3'UTR	GGTCTCAGCTTTGATACGTTCGAGAAA ACTCTACTTGAATCCATTTAATTTTTTTT ACCGACAAACCTGTGCTCGAAAATCAC	119	Kanamycin	3'UTR of *psaC*	(Hirose and Sugiura, 1997)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		AATATTTTGAGCACGGGTTTTTATGCG CTTGAGACC				
3' UTR	TMV 3'UTR	GGTCTCAGCTTATAATAAATAACGGATT GTGTCCGTAATCACACGTGGTGCGTAC GATAACGCATAGTGTTTTTCCCTCCAC TTAAATCGAAGGGTTGTGTCTTGGATC GCGCGGGTCAAATGTATATGGTTCATA TACATCCGCAGGCACGTAATAAAGCGA GGGGTTCGAATCCCCCCGTTACCCCC GGTAGGGGCCCACGCTTGAGACC	212	Kanamycin	3'UTR of tobacco mosaic virus (TMV)	(Kawamura-Nagaya et al., 2014)
3' UTR	TYMV 3'UTR	GGTCTCAGCTTGTTCTCGATCTTTAAAA TCGTTAGCTCGCCAGTTAGCGAGGTCT GTCCCCACACGACAGATAATCGGGTG CAACTCCCGCCCCTTTTCCGAGGGTCA TCGGAACCCGCTTGAGACC	127	Kanamycin	3'UTR of turnip yellow mosaic virus (TYMV)	(Ma et al., 2015)
3' UTR	BMV 3'UTR	GGTCTCAGCTTGGTGCCTTTGAGAGTC TACTTTTGCTCTCTTCGGAAGAACCCTT AGGGGTTCGTGCATGGGCTTGCATAG CAAGTCTAGATGCGGGTACCGTACAGT GTTGAAAAACACTGTAAATCTCTAAAAG AAACCACGCTTGAGACC	153	Kanamycin	3'UTR of brome mosaic virus (BMV)	(Bujarski et al., 1985)
3' UTR	ALMV 3'UTR	GGTCTCAGCTTCGTTTTCTTGACATAA GTCAAATTGCCAACCTCCACTGGGTGG ATTAAGGTTGAGGTATAGAATCCTATTC GCTCCTGATAGGAGAAATTCTATATTG CTTATATATGTGCTTATGCACATATATA AATGCTCATGCAAAACTGCATGAATGC CCCTAAGGGATGCCGCTTGAGACC	188	Kanamycin	3'UTR of almpiwar virus (ALMV)	(Trucco et al., 2014)
Integration site	*Nt trnI/trn A*	GGGCCTTGTACACACCGCCCGTCACA CTATGGGAGCTGGCCATGCCCGAAGT CGTTACCTTAACCGCAAGGAGGGGGA TGCCGAAGGCAGGGCTAGTGACTGGA GTGAAGTCGTAACAAGGTAGCCGTACT GGAAGGTGCGGCTGGATCACCTCCTT TTCAGGGAGAGCTAATGCTTGTTGGGT ATTTTGGTTTGACACTGCTTCACACCC CCAAAAAAAAGAAGGGAGCTACGTCTG AGTTAAACTTGGAGATGGAAGTCATCT TTCCTTTCTCGACGGTGAAGTAAGACC AAGCTCATGAGCTTATTATCCTAGGTC GGAACAAGTTGATAGGACCCCCTTTTT TACGTCCCCATGTTCCCCCCGTGTGGC GACATGGGGGCGAAAAAAGGAAAGAG AGGGATGGGGTTTCTCTCGCTTTTGGC ATAGCGGGCCCCCAGTGGGAGGCTCG CACGACGGGCTATTAGCTCAGTGGTA GAGCGCGCCCCTGATAATTGCGTCGTT GTGCCTGGGCTGTGAGGGCTCTCAGC CACATGGATAGTTCAATGTGCTCATCG GCGCCTGACCCTGAGATGTGGATCAT CCAAGGCACATTAGCATGGCGTACTCC TCCTGTTCGAACCGGGGGTTTGAAACCA AACTCCTCCTCAGGAGGATAGATGGG GCGATTCGGGTGAGATCCAATGTAGAT CCAACTTTCGATTCACTCGTGGGATCC GGGCGGTCCGGGGGGGGACCACCACG GCTCCTCTCTTCTCGAGAATCCATACA TCCCTTATCAGTGTATGGACAGCTATC TCTCGAGCACAGGTTTAGCAATGGGAA AATAAAATGGAGCACCTAACAACGCAT CTTCACAGACCAAGAACTACGAGATCG CCCCTTTCATTCTGGGGTGACGGAGG GATCGTACCATTCGAGCCGTTTTTTTCT TGACTCGAAATGGGAGCAGGTTTGAAA AAGGATCTTAGAGTGTCTAGGGTTGGG	2838	Kanamycin	Tobacco plastome integration site, *trnI/trnA* (*Nicotiana tabacum*)	Designed in this work base on integration site used by (Daniell et al., 1998)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		CCAGGAGGCTCTCTTAACGCCTTCTTT TTTCTTCTCATCGGAGTTATTTCACAAA GACTTGCCAGGGTAAGGAAGAAGGGG GGAACAAGCACACTTGGAGAGCGCAG TACAACGGAGAGTTGTATGCTGCGTTC GGGAAGGATGAATCGCTCCCGAAAAG GAATCTATTGATTCTCTCCCAATTGGTT GGACCGTAGGTGCGATGATTTACTTCA CGGGCGAGCTCTCTGGTTCAAGTCCA GGATGGCCCAGCTGCGAATTCGTTTAA ACTGCCTTGTCTTCATGGTGAGCAAGG GCGAGGAGCTGTTCACCGGGGTGGTG CCCATCCTGGTCGAGCTGGACGGCGA CGTAAACGGCCACAAGTTCAGCGTGTC CGGCGAGGGCGAGGGCGATGCCACC TACGGCAAGCTGACCCTGAAGTTCATC TGCACCACCGGCAAGCTGCCCGTGCC CTGGCCCACCCTCGTGACCACCTTGA CCTACGGCGTGCAGTGCTTCGCCCGC TACCCCGACCACATGAAGCAGCACGA CTTCTTCAAGTCCGCCATGCCCGAAGG CTACGTCCAGGAGCGCACCATCTTCTT CAAGGACGACGGCAACTACAAGACCC GCGCCGAGGTGAAGTTCGAGGGCGAC ACCCTGGTGAACCGCATCGAGCTGAA GGGCATCGACTTCAAGGAGGACGGCA ACATCCTGGGGCACAAGCTGGAGTAC AACTACAACAGCCACAAGGTCTATATC ACCGCCGACAAGCAGAAGAACGGCAT CAAGGTGAACTTCAAGACCCGCCACAA CATCGAGGACGGCAGCGTGCAGCTCG CCGACCACTACCAGCAGAACACCCCC ATCGGCGACGGCCCCGTGCTGCTGCC CGACAACCACTACCTGAGCACCCAGTC CAAGCTGAGCAAAGACCCCAACGAGA AGCGCGATCACATGGTCCTGCTGGAG TTCGTGACCGCCGCCGGGATCACTCT CGGCATGGACGAGCTGTACAAGTAAG AAGACAAGGGAGTTTAAACAAGCTTCT CCACTTGGCTCGGGGGGATATAGCTC AGTTGGTAGAGCTCCGCTCTTGCAATT GGGTCGTTGCGATTACGGGTTGGATG TCTAATTGTCCAGGCGGTAATGATAGT ATCTTGTACCTGAACCGGTGGCTCACT TTTTCTAAGTAATGGGGAAGAGGACCG AAACGTGCCACTGAAAGACTCTACTGA GACAAAGATGGGCTGTCAAGAACGTA GAGGAGGTAGGATGGGCAGTTGGTCA GATCTAGTATGGATCGTACATGGACGG TAGTTGGAGTCGGCGGCTCTCCCAGG GTTCCCTCATCTGAGATCTCTGGGGAA GAGGATCAAGTTGGCCCTTGCGAACA GCTTGATGCACTATCTCCCTTCAACCC TTTGAGCGAAATGCGGCAAAAGAAAAG GAAGGAAAATCCATGGACCGACCCCAT CATCTCCACCCCGTAGGAACTACGAGA TCACCCCAAGGACGCCTTCGGCATCC AGGGGTCACGGACCGACCATAGAACC CTGTTCAATAAGTGGAACGCATTAGCT GTCCGCTCTCAGGTTGGGCAGTCAGG GTCGGAGAAGGGCAATGACTCATTCTT AGTTAGAATGGGATTCCAACTCAGCAC CTTTTGAGTGAGATTTTGAGAAGAGTT GCTCTTTGGAGAGCACAGTACGATGAA AGTTGTAAGCTGTGTTCGGGGGGGAG				

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		TTATTGTCTATCGTTGGCCTCTATGGTA GAATCAGTCGGGGGACCTGAGAGGCG GTGGTTTACCCTGCGGCGGATGTCAG CGGTTCGAGTCCGCTTATCTCCAACTC GTGAACTTAGCCGATACAAAGCT				
Integration site	*Nt trnG/trnfM*	AGGAGAGAGAGGGATTCGAACCCTCG ATAGTTCTTTGTTAAAACTATACCGGTT TTCAAGACCGGGGCTATCAACCGCTCA GCCATCTCTCCGAAAGACTATTTTTATT TTATTCCTCCGAATAGAACATGGCCAT AGGGGTGGATACCCCCACTATCTGTAC TATCTGTAAAAAGATCTCAGGTGCGAA TCCACCGGTCGATCTATCTATCCGTAT ATAGATATATGATCTAGCATGCCCATTT GTGAAATAAAAAATAAAATTCCATTTCC CCCCACTCCATGTACGAATAAAGTGCG AAAGGGGGAGTAGTAATAAGTCATATA GAATCAATGGATTCATGATAAAGTAAAA TCCCTCGATGACATATTTTATCACAATT AATATTTTTTGGCTGATAGAGGGATCA AATGGTATATAGTTCATTTGTTGGTAGC TTGGAGGATTAAAAGCATGACTCTTGC TTTCCAATTGGCTGTTTTTGCATTAATT GCTACTTCATTAATCTTATTGATTAGCG TACCCGTTGTATTTGCTTCTCCTGATG GCTGGTCAAGTAACAAAAATGTTGTAT TTTCTGGTACATCCTTATGGATTGGATT AGTCTTTCTGGTGGGTATCCTTAATTCT CTCATCTCTTGAACCTATTCGTCGCAG ACCCAAAACCAAAATGACCCCCCTAAT TTTTCTCGGTTGTGAGACACATTAAATT GGAATCTAAGTCCCCAAAGAAAACGCA AATCAAATAAAGAAAACAAAAAAATTAG AGGGGGGTCAAACTTCTTGAATAAAAA GAATACAATTAAAAAAATAATTGGAATC GTTCCGAAGAGAATATGTGTCCCGGCA CTGCACAAAAAAGATCCGGTTATATAT CATATATGTGGGTACATATTGTGTATCA AGAACAAAAAAATGCGGATATGGTCGA ATGGTAAAATTTCTCTTTGCCAAGGAG AAGATGCGGGTTCGATTCCCGCTATCC GCCCAAGATCCAAGATAAAGTAATTTT ATTACTATTTATTTATTATTTAATTTCAT AAATAGCATTAAATATATCCTTAAATTA AGGATTTGGTATAGTTGGCCGTGATAG TGTAGTGATTCTATCCCTCCCCTACGT TTTCTTTTTCCTTCCACCCCCAAAAAGC GAAAGGCGGGAATTAATTACTAGGAAT TCGTTTAAACTGCCTTGTCTTCATGGT GAGCAAGGGCGAGGAGCTGTTCACCG GGGTGGTGCCCATCCTGGTCGAGCTG GACGGCGACGTAAACGGCCACAAGTT CAGCGTGTCCGGCGAGGGCGAGGGC GATGCCACCTACGGCAAGCTGACCCT GAAGTTCATCTGCACCACCGGCAAGCT GCCCGTGCCCTGGCCCACCCTCGTGA CCACCTTGACCTACGGCGTGCAGTGC TTCGCCCGCTACCCCGACCACATGAA GCAGCACGACTTCTTCAAGTCCGCCAT GCCCGAAGGCTACGTCCAGGAGCGCA CCATCTTCTTCAAGGACGACGGCAACT ACAAGACCCGCGCCGAGGTGAAGTTC GAGGGCGACACCCTGGTGAACCGCAT CGAGCTGAAGGGCATCGACTTCAAGG AGGACGGCAACATCCTGGGGCACAAG	3827	Kanamycin	Tobacco plastome integration site, *trnG/trnfM* (*Nicotiana tabacum*)	Designed in this work base on integration site used by (Bock and Maliga, 1995)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		CTGGAGTACAACTACAACAGCCACAAG GTCTATATCACCGCCGACAAGCAGAAG AACGGCATCAAGGTGAACTTCAAGACC CGCCACAACATCGAGGACGGCAGCGT GCAGCTCGCCGACCACTACCAGCAGA ACACCCCCATCGGCGACGGCCCCGTG CTGCTGCCCGACAACCACTACCTGAG CACCCAGTCCAAGCTGAGCAAAGACC CCAACGAGAAGCGCGATCACATGGTC CTGCTGGAGTTCGTGACCGCCGCCGG GATCACTCTCGGCATGGACGAGCTGTA CAAGTAAGAAGACAAGGGAGTTTAAAC AAGCTTAACCCTAAAATAGTTTGGCAA AACAAGATGTTGCGGAGACAGGATTTG AACCCGTGACCTCAAGGTTATGAGCCT TGCGAGCTACCAAACTGCTCTACCCCG CGCCGAAGATAAGAACTGAAAACTAAT AGATAAACAAGGATTAAATGCGCCCCT CCACCCTATCTGTACAAATAGAATAGC CCATTTATACAGAATGGTAAAGGGGCT TCTATGATCATCGACCATAGAAATAGA AATGAAGCGTTAATCCTTACCAACTTG ATCTTGTTGCTCCTGGCAACAAACATG CATGAACCATTTCACGAAGTATGTGTC CGGATAGTCCAAAGTCTCGATAGTTAG CTCTCGGCCTTCCGGTCAAAAAACAAC GTCGATGAAGGCGTGTAGGTGCACTAT TCCGTGGTGGGGATTGTAACTTTCCAT AAATTTCCCATTTGTCACTCAACGACG GAACCTTGCTTATTTCTTTCTTTGAGGA TCGACGAATCGAATGATATTTCTGTTC CAATTTTTGCCTCTTCTTCTCCCTCTGA ATCAAACTTTTCCTTGCCATAATGGTTG AGTTCCTATTAGTATCCATGATACAAGT CGAATCCTAGATGTAGAAATAGAAGAA GGTGGACCCCCTCTCCGTCGAAAGAA ATGAGATTATCGCAGATACACACATTA AAAATATTAACCAAATTTGCCCGACGTA GAGGCAATCAAGAAAGCCGCATAAGT GAATATATAACCTACAGAAAAGTGAGC TAATCCAACCAATCTTGCTTGTACAATG GAAAGGGCCACTGGTTTATCTCTCCAG CGAATCAAATTGGCCAAAGGTGTGCGT TCATGAGCCCATGCTAAAGTTTCAATC AATTCCTGCCAATATCCACGCCAAGAA ATTAAGAACATAAATCCAGTAGCCCAA ACAAGATGTCCAAATAAGAACATCCAT GCCCAAACCGATAAACTATTCATACCA AAAGGATTATATCCGTTGATAAGTTGT GAAGAGTTTAACCATAAATAATCCCTTA ACCAGCCCATCAAATAAGTGGAAGATT CATTAAACTGTGAAACGTTACCCTGCC ATAATGTGATGTGCTTCCAATGCCAAT AAAAAGTAACCCATCCAATAGTATTTAA CATCCAAAAAACTGCCAAATAAAACGC GTCCCATGCCGAAATATCACAAGTACC GCCTCGTCCTGGGCCATCGCACGGAA AACTATAACCGAAATCCTTTTTATCTGG CATTAACTTGGAACCACGTGCATCTAA AGCACCTTTTACTAAGATCAATGTAGTT GTATGTAAACCAAGAGCAATAGCATGA TGAACCAAAAAGTCTCCAGGACCTATT GTTAAAAATAATGAATTACTATTTTCATT AACAGCATTTAACCAACCCGGCAACCA GATGCTTCGACCCGCATTGAATGCTGG				

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		ACCACTCGTTGAAGATAAAAGTACATC GAACCCATATGAAGTTTTACCATGAGC GGATTGTATCCATTGAGCAAATATAGG TTCAATCAAGATTTGCTTCTCCGGAGT GCCAAAGGCAAGCATGACATCATTATG AACATAAAGTCCCAGGGTATGGAATCC CAGAAAGAGGCTGGCCCAACTTAAATG AGATATGATAGCTTCTTTATGCTCTAAC ATTCTTGCCAATACATTATCTTCATTTT GCTCCGGATTGTAATCTCTAATGAAAA ATATAGCTCCATGAGCAAAAGCTCCTG TCATGATGAATCCTGCGATATATTGGT GGTGGGTATATAATGCAGCTTGAGTAG TAAAGTCTTGTGCTATGAATGCATAAG CAGGTAAAGAGTACATGTGTTGAGCTA CCAAAGAAGTAATAACCCCTAAAGAAG CTAGAGCAAGGC				
Integration site	*Nt trnV/3' rps12*	AATTCACCGCCGTATGGCTGACCGGC GATTACTAGCGATTCCGGCTTCATGCA GGCGAGTTGCAGCCTGCAATCCGAAC TGAGGACGGGTTTTTGGGGTTAGCTCA CCCTCGCGGGATCGCGACCCTTTGTC CCGGCCATTGTAGCACGTGTGTCGCC CAGGGCATAAGGGGCATGATGACTTG ACGTCATCCTCACCTTCCTCCGGCTTA TCACCGGCAGTCTGTTCAGGGTTCCAA ACTCAACGATGGCAACTAAACACGAGG GTTGCGCTCGTTGCGGGACTTAACCCA ACACCTTACGGCACGAGCTGACGACA GCCATGCACCACCTGTGTCCGCGTTC CCGAAGGCACCCCTCTCTTTCAAGAGG ATTCGCGGCATGTCAAGCCCTGGTAAG GTTCTTCGCTTTGCATCGAATTAAACCA CATGCTCCACCGCTTGTGCGGGCCCC CGTCAATTCCTTTGAGTTTCATTCTTGC GAACGTACTCCCCAGGCGGGATACTTA ACGCGTTAGCTACAGCACTGCACGGG TCGATACGCACAGCGCCTAGTATCCAT CGTTTACGGCTAGGACTACTGGGGTAT CTAATCCCATTCGCTCCCCTAGCTTTC CTCTCTCAGTGTCAGTGTCGGCCCAGC AGAGTGCTTTCGCCGTTGGTGTTCTTT CCGATCTCTACGCATTTCACCGCTCCA CCGGAAATTCCCTCTGCCCCTACCGTA CTCCAGCTTGGTAGTTTCCACCGCCTG TCCAGGGTTGAGCCCTGGGATTTGAC GGCGGACTTAAAAAGCCACCTACAGAC GCTTTACGCCCAATCATTCCGGATAAC GCTTGCATCCTCTGTATTACCGCGGCT GCTGGCACAGAGTTAGCCGATGCTTAT TCCCCAGATACCGTCATTGCTTCTTCT CCGGGAAAAGAAGTTCACGACCCGTG GGCCTTCTACCTCCACGCGGCATTGCT CCGTCAGGCTTTCGCCCATTGCGGAAA ATTCCCCACTGCTGCCTCCCGTAGGAG TCTGGGCCGTGTCTCAGTCCCAGTGT GGCTGATCATCCTCTCGGACCAGCTAC TGATCATCGCCTTGGTAAGCTATTGCC TCACCAACTAGCTAATCAGACGCGAGC CCCTCCTCGGGCGGATTCCTCCTTTTG CTCCTCAGCCTACGGGGTATTAGCAGC CGTTTCCAGCTGTTGTTCCCCTCCCAA GGGCAGGTTCTTACGCGTTACTCACCC GTCCGCCACTGGAAACACCACTTCCC GTCCGACTTGCATGTGTTAAGCATGCC	3726	Kanamycin	Tobacco plastome integration site, *trnV/3'rps12* (*Nicotiana tabacum*)	Designed in this work base on integration site used by (Staub and Maliga, 1992)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		GCCAGCGTTCATCCTGAGCCAGGATC GAACTCTCCATGAGATTCATAGTTGCA TTACTTATAGCTTCCTTGTTCGTAGACA AAGCGGATTCGGAATTGTCTTTCATTC CAAGGCATAACTTGTATCCATGCGCTT CATATTCGCCCGGAGTTCGCTCCCAGA AATATAGCCATCCCTGCCCCCTCACGT CAATCCCACGAGCCTCTTATCCATTCT CATTGAACGACGGCGGGGGAGCTTTC GAGGCCTCGAAATCCAACTAGAAAAAC TCACATTGGGCTTAGGGATAATCAGGC TCGAACTGATGACTTCCACCACGTCAA GGTGACACTCTACCGCTGAGTTATATC CCTTCCCCGCCCCATCGAGAAATAGAA CTGACTAATCCTAAGTCAAAGGGTCGA GAAACTCAACGCCACTATTCTTGAACA ACTTGGAGCCGGGCCTTCTTTTCGCAC TATTACGGATATGAAAATAATGGTCAAA ATCGGATTCAATTGTCAAGAATTCGTTT AAACTGCCTTGTCTTCATGGTGAGCAA GGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGC GACGTAAACGGCCACAAGTTCAGCGT GTCCGGCGAGGGCGAGGGCGATGCC ACCTACGGCAAGCTGACCCTGAAGTTC ATCTGCACCACCGGCAAGCTGCCCGT GCCCTGGCCCACCCTCGTGACCACCT TGACCTACGGCGTGCAGTGCTTCGCC CGCTACCCCGACCACATGAAGCAGCA CGACTTCTTCAAGTCCGCCATGCCCGA AGGCTACGTCCAGGAGCGCACCATCT TCTTCAAGGACGACGGCAACTACAAGA CCCGCGCCGAGGTGAAGTTCGAGGGC GACACCCTGGTGAACCGCATCGAGCT GAAGGGCATCGACTTCAAGGAGGACG GCAACATCCTGGGGCACAAGCTGGAG TACAACTACAACAGCCACAAGGTCTAT ATCACCGCCGACAAGCAGAAGAACGG CATCAAGGTGAACTTCAAGACCCGCCA CAACATCGAGGACGGCAGCGTGCAGC TCGCCGACCACTACCAGCAGAACACC CCCATCGGCGACGGCCCCGTGCTGCT GCCCGACAACCACTACCTGAGCACCC AGTCCAAGCTGAGCAAAGACCCCAAC GAGAAGCGCGATCACATGGTCCTGCT GGAGTTCGTGACCGCCGCCGGGATCA CTCTCGGCATGGACGAGCTGTACAAGT AAGAAGACAAGGGAGTTTAAACAAGCT TCTGCCCCTATCGGAAATAGGATTGAC TACCGATTCCGAAGGAACTGGAGTTAC ATCTCTTTTCCATTCAAGAGTTCTTATG CGTTTCCACGCCCCTTTAAGACCCCGA AAAATGGACAAATTCCTTTTCTTAGGAA CACATACAAGATTCGTCACTACAAAAA GGATAATGGTAACCCTACCATTAACTA CTTCATTTATGAATTTCATAGTAATAGA AATACATGTCCTACCGAGACAGAATTT GGAACTTGCTATCCTCTTGCCTAGCAG GCAAAGATTTACCTCCGTGGAAAGGAT GATTCATTCGGATCGACATGAGAGTCC AACTACATTGCCAGAATCCATGTTGTAT ATTTGAAAGAGGTTGACCTCCTTGCTT CTCTCATGGTACACTCCTCTTCCCGCC GAGCCCCTTTTCTCCTCGGTCCACAGA GACAAAATGTAGGACTGGTGCCAACAA				

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		TTCATCAGACTCACTAAGTCGGGATCA CTAACTAATACTAATCTAATATAATAGT CTAATATATCTAATATAATAGAAAATAC TAATATAATAGAAAAGAACTGTCTTTTC TGTATACTTTCCCCGGTTCCGTTGCTA CCGCGGGCTTTACGCAATCGATCGGA TTAGATAGATATCCCTTCAACATAGGTC ATCGAAAGGATCTCGGAGAGCCACCA AAGTACGAAAGCCAGGATCTTTCAGAA AACGGATTCCTATTCAAAGAGTGCATA ACCGCATGGATAAGCTCACACTAACCC GTCAATTTGGGATCCAAATTCGAGATT TTCCTTGGGAGGTATCGGGAAGGATTT GGAATGGAATAATATCGATTCATACAG AAGAAAAGGTTCTCTATTGATTCAAACA CTGTACCTAACCTATGGGATAGGGATC GAGGAAGGGGAAAAACCGAAGATTTC ACATGGTACTTTTATCAATCTGATTTAT TTCGTACCTTTCGTTCAATGAGAAAATG GGTCAAATTCTACAGGATCAAACCTAT GGGACTTAAGGAATGATATAAAAAAAA GAGAGGGAAAATATTCATATTAAATAAA TATGAAGTAGAAGAACCCAGATTCCAA ATGAACAAATTCAAACTTGAAAAGGAT CTTCCTTATTCTTGAAGAATGAGGGGC AAAGGGATTGATCAAGAAAGATC				
Integration site	*Nt atpB/accD*	CACTATCTCGACCTTGAACTACCAGAG CGTTATAAATATTCGGCATCTTGCCCG GGGGAAAGGCTACATCTAGTACCGGA CCGATGATTTGGACGACACGCCCCGG GTTTTTTTTTTCAAGCGTGGAAACCCCA GAACCAGAAGTAGTAGGATTGATTCTC ATAATAATAAAATAAATAAATATGTCGA AATGTTTTTGCAAAAATTATCGAGTTCA AAATAAATGTCCGCTAGCACGTCGATC GGTTAATTCAATAAAATGGGAATTAGC ACTCGATTTCGTTGGCACCATGCAATT GAACCGATTCAATTGTTTACTTATTCAC TGAGACTGAGTGAATTTGCAAGCCCAC CCAACCTATTTTAATTTTAAAATCTCAA GTGGATGAATCAGAATCTTGAGAAAGT CTTTCATTTGTCTATCATTATAGACAAT CCCATCCATATTATCTATTCTATGGAGT TCGAACCTGAACTTTATTTTCTATTTCT ATTACGATTCATTATTTGTATCTAATTG GCTCCTCTTCTTATTTATTTTTGATTTCA ATTTCAGCATATCGATTTATGCCTAGCC TATTCTTTTCTTTGTGTTTTTCTTTCTTT TTTATACCTTTCATAGATTCATAGAGGA GTTCCGTATATTTTCACATCTAGGATTT ACATATACAACATATACCACTGTCAAG GGGGAAGTTCTTATTATTTAGGTTAGT CAGGTATTTCCATTTCAAAAAAAAAAAA AGTAAAAAAGAAAAATTGGGTTGCGCT ATATATATGAAAGAGTATACAATAATGA TGTATTTGGCAAATCAAATACCATGGT CTAATAATCAAACATTCTGATTAGTTGA TAATATTAGTATTAGTTGGAAATTTTGT GAAAGATTCCTGTGAAAAGTTTCATTAA CACGGAGTTCGTGTCGAGTAGACCTTG TTGTTGTGAGAGTTCTTAATTCATGAGT TGTAGGGAGGGATTTGAATTCGTTTAA ACTGCCTTGTCTTCATGGTGAGCAAGG GCGAGGAGCTGTTCACCGGGGTGGTG CCCATCCTGGTCGAGCTGGACGGCGA	2752	Kanamycin	Tobacco plastome integration site, *atpB/accD* (*Nicotiana tabacum*)	(Lin et al., 2014)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		CGTAAACGGCCACAAGTTCAGCGTGTC CGGCGAGGGCGAGGGCGATGCCACC TACGGCAAGCTGACCCTGAAGTTCATC TGCACCACCGGCAAGCTGCCCGTGCC CTGGCCCACCCTCGTGACCACCTTGA CCTACGGCGTGCAGTGCTTCGCCCGC TACCCCGACCACATGAAGCAGCACGA CTTCTTCAAGTCCGCCATGCCCGAAGG CTACGTCCAGGAGCGCACCATCTTCTT CAAGGACGACGGCAACTACAAGACCC GCGCCGAGGTGAAGTTCGAGGGCGAC ACCCTGGTGAACCGCATCGAGCTGAA GGGCATCGACTTCAAGGAGGACGGCA ACATCCTGGGGCACAAGCTGGAGTAC AACTACAACAGCCACAAGGTCTATATC ACCGCCGACAAGCAGAAGAACGGCAT CAAGGTGAACTTCAAGACCCGCCACAA CATCGAGGACGGCAGCGTGCAGCTCG CCGACCACTACCAGCAGAACACCCCC ATCGGCGACGGCCCCGTGCTGCTGCC CGACAACCACTACCTGAGCACCCAGTC CAAGCTGAGCAAAGACCCCAACGAGA AGCGCGATCACATGGTCCTGCTGGAG TTCGTGACCGCCGCCGGGATCACTCT CGGCATGGACGAGCTGTACAAGTAAG AAGACAAGGGAGTTTAAACAAGCTTAA ACAGTAGACATTAGCAGATAAATTAGC AGGAAATAAAGAAGGATAAGGAGAAAG AACTCAAGTAATTATCCTTCGTTCTCTT AATTGAATTGCAATTAAACTCGGCCCA ATCTTTTACTAAAAGGATTGAGCCGAAT ACAACAAAGATTCTATTGCATATATTTT GACTAAGTATATACTTACCTAGATATAC AAGATTTGAAATACAAAATCTAGAAAAC TAAATCAAAATCTAAGACTCAAATCTTT CTATTGTTGTCTTGGATCCACAATTAAT CCTACGGATCCTTAGGATTGGTATATT CTTTTCTATCCTGTAGTTTGTAGTTTCC CTGAATCAAGCCAAGTATCACACCTCT TTCTACCCATCCTGTATATTGTCCCCTT TGTTCCGTGTTGAAATAGAACCTTAATT TATTACTTATTTTTTTATTAAATTTTAGA TTTGTTAGTGATTAGATATTAGTATTAG ACGAGATTTTACGAAACAATTATTTTTT TATTTCTTTATAGGAGAGGACAAATCTC TTTTTTCGATGCGAATTTGACACGACAT AGGAGAAGCCGCCCTTTATTAAAAATT ATATTATTTTAAATAATATAAAGGGGGT TCCAACATATTAATATATAGTGAAGTGT TCCCCCAGATTCAGAACTTTTTTTCAAT ACTCACAATCCTTATTAGTTAATAATCC TAGTGATTGGATTTCTATGCTTAGTCTG ATAGGAAATAAGATATTCAAATAAATAA TTTTATAGCGAATGACTATTCATCTATT GTATTTTCATGCAAATAGGGGGCAAGA AAACTCTATGGAAAGATGGTGGTTTAA TTCGATGTTGTTTAAGAAGGAGTTCGA ACGCAGGTGTGGGCTAAATAAATCAAT GGGCAGTCTTGGTCCTATTGAAAATAC CAATGAAGATCCAAATCGAAAAGTGAA AAACATTCATAGTTGGAGGAATCGTGA CAATTCTAGTTGCAGTAATGTTGATTAT TT				
Integration site	*Zm trnI/trn*	AAGTCTTCTTTCGTTTAGGGTGAAGTA AGACCAAGCTCATGAGCTTATTATCCT	3479	Kanamycin	Corn plastome integration site,	Designed in this work

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
	A	AGGTCGGAACAAATTAGTTGATAGTGA TAGGATCCCCTTTTTGACGTCCCCATG TCCCCCCGTGTGGCGGCATGGGGATG TCAAAAGGAAAGGGATGGAGTTTTTCT CGCTTTTGGCGTAGCGGCCTCCCTTTG GGAGGCCGCGCGACGGGCTATTAGCT CAGTGGTAGAGCGCGCCCCTGATAATT CGTCGTTGTGCCTCGGCTGTGAGGGC TCTCAGCCACATGGATAGTTCAATGTG CTCATCAGCGCCTGACCCGAAGATGT GGATCATCCAAGGCACATTAGCATGGC GTACTCCTCCTGTTTGAATCGGAGTTT GAAACCAAACAAACTTCTCCTCAGGAG GATAGATGGGGCGATTCAGGTGAGAT CCCATGTAGATCGAACTTTCTATTCACT CGTGGGATCCGGGCGGTCCGGGGGG GGGCCACCGGGGCTCCTCTCTTCTCG AGAATCCATACATCCCTTATCAGTGTAT GGAGAGCTATCTCTCGAGCACAGGTT GAGGTTCGTCCTCAATGGGAAAATGGA GCACCTAACAACGCATCTTCACAGACC AAGAACTACGAGATCACCCCTTTCATT CTGGGGTGACGGAGGGATCGTACCAT TCGAGCCTTTTTTTTCATGCTTTTCCCGG CGGTCTGGAGAAAGCAGTAATCAATAG GACTTCCCTAATCCTCCCTTCCTGAAA GAAGAACGTGAAATTCTTTTTCCTTTCT GCAGGGACCAGGAGATTGGATCTAGC CATAAGAGGAATGCTTGGTATAAATAA GCCACTTCTTGGTCCTCGACCCCCTAA GTCACTACGAGCGCCCCCGATCAGTG CAATGGGATGTGGCTATTTATCTATCT CTTGACTCGAAATGGGAGCAGAGCAG GTTTGAAAAAGGATCTTAGAGTGTCTA GGGTTGGGCCAGGAGGATCTCTTAAC CCCTTCTTTTTTCTGCCCATCGGAGTTA TTTCCCAAGGACTTGCCGTGGTAAGGG GGAGAAGGGGGAAGAAGCACACTTGA AGAGCGCAGTACAACGGGGAGTTGTA TGCTGCGTTCGGGAAGGATGGATCGC TCCCGAAAAGGAGTCTATTGATTCTCT CCCAATTGGTTGGATCGTAGGGGCGA TGATTTACTTCACGGGCGAGCTCTCTG GTTCAAGTCCAGGATGGCCCAGCTGC GAATTCGTTTAAACTGCCTTGTCTTCAT GGTGAGCAAGGGCGAGGAGCTGTTCA CCGGGGTGGTGCCCATCCTGGTCGAG CTGGACGGCGACGTAAACGGCCACAA GTTCAGCGTGTCCGGCGAGGGCGAGG GCGATGCCACCTACGGCAAGCTGACC CTGAAGTTCATCTGCACCACCGGCAAG CTGCCCGTGCCCTGGCCCACCCTCGT GACCACCTTGACCTACGGCGTGCAGT GCTTCGCCCGCTACCCCGACCACATG AAGCAGCACGACTTCTTCAAGTCCGCC ATGCCCGAAGGCTACGTCCAGGAGCG CACCATCTTCTTCAAGGACGACGGCAA CTACAAGACCCGCGCCGAGGTGAAGT TCGAGGGCGACACCCTGGTGAACCGC ATCGAGCTGAAGGGCATCGACTTCAAG GAGGACGGCAACATCCTGGGGCACAA GCTGGAGTACAACTACAACAGCCACAA GGTCTATATCACCGCCGACAAGCAGAA GAACGGCATCAAGGTGAACTTCAAGAC CCGCCACAACATCGAGGACGGCAGCG			*trnI/trnA* (*Zea Mays*)	base on integration site used by (Daniell et al., 1998)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		TGCAGCTCGCCGACCACTACCAGCAG AACACCCCCATCGGCGACGGCCCCGT GCTGCTGCCCGACAACCACTACCTGA GCACCCAGTCCAAGCTGAGCAAAGAC CCCAACGAGAAGCGCGATCACATGGT CCTGCTGGAGTTCGTGACCGCCGCCG GGATCACTCTCGGCATGGACGAGCTG TACAAGTAAGAAGACAAGGGAGTTTAA ACAAGCTTCTCCACTTGGCTCGGGGG GGATATAGCTCAGTTGGTAGAGCTCCG CTCTTGCAATTGGGTCGTTGCGATTAC GGGTTGGCTGTCTAATTGTCCAGGCG GTAATGATAGTATCTTGTACCTGAACC GGTGGCTCACTTTTTCTAAGTAATGGG GAAGAGGACTGAAACATGCCACTGAAA GACTCTACTGAGACAAAAAGATGGGCT GTCAAAAAGGTAGAGGAGGTAGGATG GGCAGTTGGTCAGATCTAGTATGGATC GTACATGGACGATAGTTGGAGTCGGC GGCTCTCCTAGGCTTCCCTCATCTGGG ATCCCTGGGGAAGAGGATCAAGTTGG CCCTTGCGAATAGCTTGATGCACTATC TCCCTTCAACCCTTTGAGCGAAATGTG GCAAAAGGAAGGAAAATCCATGGACC GACCCCATTGTCTCCACCCCGTAGGAA CTACGAGATCACCCCAAGGACGCCTTC GGTGGGTCTATCGGACCGACCATAGA TCCTGTTCAATAAGTGGAACACAATAG CCGTCCGCTCTCCGGTTGGGCAGTAA GGGTCGGAGAAGGGCAATCACTCGTT CTTAAAACCAGCATTCTTAAGTTAAGAT CAAAGAGTCGGGCGGAAAAAGGGGAG AGCTCCCCGTTCCTGGTTCTCCTGTAG CTGGATTCCCCGGAACCACAAGAATCC TTAGAATGGGATTCCAACTCAGCACCT TTTGTTTTGGGATTTTGAGAAGAGTTG CTCTTTGGAGAGCACAGTACGATGAAA GTTGTAAGCTGTGTTCGGGGGGGAGT TATTGCCTATCGTTGTCCTCTATGGTA GAACCCGTCGGGGAGGCCTGAGAGGC GGTGGTTTACCCTGTGGCGGATGTCA GCGGTTCGAGTCCGCTTATCTCCAGCC CGTGAACTTAGCGGATACTATGATAGC ACCGAAGGTTGCCAATTCGTCAGTTCG ATCTATGATTTCGCATTCATGGACGTT GATAAGATCCTTCCATTTAGTAGCACC TTAGGATGGCATAGCCTTAACGTTAAT GGCGAGGTTCAAAAGAGGAAAGGCTT GCGGTGGATACCTAGGCACCCAGAGG CGAGGAAGAGGCGTAGCAAGCGACGA AATGCTTCGGGGAGTTGAAAATAAGCA TAGATCCGGAGATTCCCAAATAGGTCA ACCTTTTGAACTGCCTGCCGAATCCAT GAGCAGGCAAGAGACAACCTGGCGAA CTGAAACATCTTAGTAGCCAGAGGAAA AGAAAGCAAAAGCGATTCCCGTAGTAG CGGCGAGCGAAATGGGAGCAGCCTAA ACCGTGAAAACGGGGTTGTGGGTGAG CAATACAAGCGTTGTGCTGCTAGGCGA AGCGGTTGAGTGCCGCACCCTAGATG GCTATAGTCCAGTAGCCGATAGCATCA CTGCTTACGCTCTGACCCGAGTAGCAT GGGGCACGTGGAATCCCGTGTGAATC AGCAAGGACCACCTTGCAAGGC				
Integration	*Zm*	CGGTGTGTACAAGGCCCGGGAACGGA	4251	Kanamycin	Corn plastome	Designed in

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
site	*trnV/3' rps12*	TTCGCCGCCGTATGGCTGACCGGCGA TTACTAGCGATTCCTGCTTCATGCAGG CGAGTTGCAGCCTGCAATCCGAACTGA GGACGGGTTTTTGGAGTTAGCTCACCC TCGCGAGATCGCGACCCTTTGTCCCG CCCATTGTAGCACGTGTGTCGCCCAG GGCATAAGGGGCATGATGACTTGGCC TCATCCTCTCCTTCCTCCGGCTTAACA CCGGCGGTCTGTTCAGGGTTCCAAACT CATAGTGGCAACTAAACACGAGGGTTG CGCTCGTTGCGAGACTTAACCCAACAC CTTACGGCACGAGCTGACGACAGCCA TGCACCACCTGTGTCCGCGTTCCCGA GGGCACCCCTCTCTTTCAAGAGGATTC GCGGCATGTCAAGCCCTGGTAAGGTT CTTCGCCTTGCATCGAATTAAACCACA TGCTCCACCGCTTGTGCGGGCCCCCG TCAATTCCTTTGAGTTTCATTCTTGCGA ACGTACTCCCCAGGCGGGATACTTAAC GCGTTAGCTACAGCACTGCACGGGTC GAGTCGCACAGCACCTAGTATCCATCG TTTACGGCTAGGACTACTGGGATCTCT AATCCCATTTGCTCCCCTAGCTTTCGT CGCTCAGTGTCAGTGTCGGCCCAGCA GAGTGCTTTCGCCGTTGGTGTTCTTTC CGATCTCAATGCATTTCACCGCTCCAC CGGAAATTCCCTCTGCCCCTACCGTAC TCCAGCTTGGTAGTTTCCACCGCCTGT CCAGGGTTGAGCCCTGGGATTTGACG GCGGACTTGAAAAGCCACCTACAGAC GCTTTACGCCCAATCATTCCGGATAAC GCTTGCATCCTCTGTCTTACCGCGGCT GCTGGCACAGAGTTAGCCGATGCTTAT TCCTCAGATACCGTCATTGTTTCTTCTC CGAGAAAAGAAGTTGACGACCCGTAG GCCTTCCACCTCCACGCGGCATTGCTC CGTCAGGCTTTCGCCCATTGCGGAAAA TTCCCCACTGCTGCCTCCCGTAGGAGT CTGGGCCGTGTCTCAGTCCCAGTGTG GCTGATCATCCTCTCGGACCAGCTACT GATCATCGCCTTGGTAAGCTATTGCCT CACCAACTAGCTAATCAGACGCGAGCC CCTCCTTGGGCGGATTTCTCCTTTTGC TCCTCAGCCTACGGGGTATTAGCAACC GTTTCCAGTTGTTGTTCCCCTCCCAAG GGCAGGTTCTTACGCGTTACTCACCCG TTCGCCACTGGAAACACCACTTCCCGT TCGACTTGCATGTGTTAAGCATGCCGC CAGCGTTCATCCTGAGCCAGGATCGAA CTCTCCATGAGATTCATAGTTGCATTAC TTATAGCTTCCTTATTCGTAGACAAAGC GGATTCGGAATTGTCTTTCCTTCCAAG GATAACTTGTATCCATGCGCTTCAGAT TATTAGCCTGGAGTTCGCCACCAGCAG TATAGCCAACCCTACCCTATCACGTCA ATCCCACAAGCCTCCTATCCATTCCCG TTCGATCGTGGTGGGGGAGTAAGTCA AAATAGAAAAAACTCACATTAGGTTTAG GGATAATCAGGCTCGAACTGATGACTT CCACCACGTCAAGGTGACACTCTACCG CTGAGTTATATCCCTTCCCCGTCCCAT CGAGAAAGAGAATTAACGAATCCTAAG GCAAAGGGGCGAGAAACGCAAGGCCA CTCTTCCTCCGGGCTTTCTTTCCGCAC TATTATGGACAGTCAAATAATGGGAAA			integration site, *trnV/3'rps12* (*Zea Mays*)	this work base on integration site used by (Staub and Maliga, 1992)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		AATTGGATTCAATTGTCAAGAATTCGTT TAAACTGCCTTGTCTTCATGGTGAGCA AGGGCGAGGAGCTGTTCACCGGGGTG GTGCCCATCCTGGTCGAGCTGGACGG CGACGTAAACGGCCACAAGTTCAGCG TGTCCGGCGAGGGCGAGGGCGATGC CACCTACGGCAAGCTGACCCTGAAGTT CATCTGCACCACCGGCAAGCTGCCCG TGCCCTGGCCCACCCTCGTGACCACC TTGACCTACGGCGTGCAGTGCTTCGC CCGCTACCCCGACCACATGAAGCAGC ACGACTTCTTCAAGTCCGCCATGCCCG AAGGCTACGTCCAGGAGCGCACCATC TTCTTCAAGGACGACGGCAACTACAAG ACCCGCGCCGAGGTGAAGTTCGAGGG CGACACCCTGGTGAACCGCATCGAGC TGAAGGGCATCGACTTCAAGGAGGAC GGCAACATCCTGGGGCACAAGCTGGA GTACAACTACAACAGCCACAAGGTCTA TATCACCGCCGACAAGCAGAAGAACG GCATCAAGGTGAACTTCAAGACCCGCC ACAACATCGAGGACGGCAGCGTGCAG CTCGCCGACCACTACCAGCAGAACAC CCCCATCGGCGACGGCCCCGTGCTGC TGCCCGACAACCACTACCTGAGCACC CAGTCCAAGCTGAGCAAAGACCCCAA CGAGAAGCGCGATCACATGGTCCTGC TGGAGTTCGTGACCGCCGCCGGGATC ACTCTCGGCATGGACGAGCTGTACAA GTAAGAAGACAAGGGAGTTTAAACAAG CTTGATTCGAGCCATAGCACATGGTTT CATAAAATCTGTACGATTTTCCCGATCT AAATCGAGTGGGTTTCCATGAAGAAGA TCTTGTTCAGCATGTTCTATTCGATACT GGTAGGAGAAGAACCCGACTCGGTAT TCTTAAAAAAAAGAGGGGAAGCAGAACC AAGTCAAGATGATATGGATCGCCCCTT CTTCTTGCGCCAAAGATCTTACCATTTC CGAAGGAACTGGGGCTACATTTCCTTT CCATTTCCATTCCAGAGTTTCTATCTGT TTCCACGCCCTTTTTTTGAAACCTCGAA ACATGAAATGGACAAATTCCTTCTCTTA GGAACACATACAAGAAAAAGGATAATG GTAGCCCCCCCATTAACTAATTTATGA ATTTCATAGTAATAGAAATCCATGTCCT ACCGAGACAGAATTTCGAACTTGCTAT CCTCTTGCCTAATAGGCAAAGATTGAC CTCTGTAGAAAGACTGATTCATTCGGA TCGATATGAGGACCCAACTCCGTTGCA TTGCCAGAATCCATGTTCCATATTTGAA GCGGGTTGACCTCTGTGCTTCTCTCAT GGTACAATCCTCTTCCTGCTGAGCCCC CTTTCTCCTCGGTCCACAGAGAAAAAA TGTAGGACTGGTGCCGACAGTTCATCA CGGAAGAAAGAACTCACAGAGCCGGG GATCGCTAACTAATAGAATAGTACTACT AACTAATACTAATATATAGATAACTAAT ACTAATATATAGATATTAATACTAATATA TAGATATATAGAAATAGATATCTAGAAA TAGAAACGAACTAATATATAGATAATCG AAATCGAAAAGAACTGTCTTTTCTGTAT ACTTTCCCCGTTCTATGCTACCGCGGG TCTTATGCAATCGATCGGATCATATAG ATATCCCTTCAACACAACATAGGTCAT CGAAAGGATCTCGGACGACTCACCAAA				

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		GCACGAAAGCCAGTTAGAAAATGGATT CCTATTTGAAGAGTGCCTAACCGCATG GATAAGCTCACATTAACCCGTCAATTTT GGATCCAATTCGGGATTTTTCTTGGGA AGTTTCGGGAAGAAATTGGAATAATAT CGATTCATACAGAGGAAAAAGTTCTCT ATTGATGCAAACGCTGTACCTCAGAGG ATAGGGATAGAGGAAGAGGGAAAAAT CGAAATGAAATAAATAAAGAATAAAGC CAAAAAAAATAAGTCGAAGATAGAAGAG CCCAGATTCAAAATGAAGAAATGGAAA CTCGAAAAGGATCCTTCTGATTCTCAA AGAATGAGGGGCAAGGGGATTGATAC CGAGAAAGATTTCTTCTTATTATAAGAC GTGATTTGATCCGCATATGTTTGGTAA AAGAACAATCTTCTCCTTTAATCATAAA TGGAAAGTGTTCAATTAGAACATGAAA ACGTGACTCAATTGGTCTTAGTTAGCC TTCGGGACGGAGTGGAAGAAGGGCGG AGACTCTCGAACGAGGAAAAGGATCC CTTCGAAAGAATTGAACGAGGAGCCGT ATGAGGTGAAAATCTCATGTACGGTTC TGTAGAGGGACAGTAAGGATGACTTAT CTGTCGACTTTTCCACTATCAACCCCA AAAAACCCAACTCTGCCTTACGTAAAG TTGCCAGAGTACGATTAACCTCTGGAT TTGAAATCACTGCTTATATACCTGGTAT TGGCCA				
Integration site	*St trnI/trnA*	GTACACACCGCCCGTCACACTATGGG AGCTGGCCATGCCCGAAGTCGTTACCT TAACCGCAAGGAGGGGGATGCCGAAG GCAGGGCTAGTGACTGGAGTGAAGTC GTAACAAGGTAGCCGTACTGGAAGGT GCGGCTGGATCACCTCCTTTTCAGGGA GAGCTAATGCTTGTTGGGTATTTTGGT TTGACACTGCTTCACACCCCCAAAAAA AAGAAGGGAGCTACGTCTGAGTTAAAC TTGGAGATGGAAGTCGTCTTTCCTTTC TCGACGGTGAAGTAAGACCAAGCTCAT GAGCTTATTATCCTAGGTCGGAACAAG TTGATAGGATCCCCTTTTTTACGTCCCT ATGTTCCCCCGTGTGGCGACATGGGG GCGAAAAAAGGAAAGAGAGGGATGGG GTTTCTCTCGCTTTTGGCATAGCGGGC CCCCAGTGGGAGGCTCGCACGACGGG CTATTAGCTCAGTGGTAGAGCGCGCC CCTGATAATTGCGTCGTTGTGCCTGGG CTGTGAGGGCTCTCAGCCACATGGATA GTTCAATGTGCTCATCGGCGCCTGACC CTGAGATGTGGATCATCCAAGGCACAT TAGCATGGCGTACTCCTCCTGTTCGAA CCGGGGTTTGAAACCAAACTCCTCCTC AGGAGGATAGATGGGGCGATTCGGGT GAGATCCAATGTAGATCCAACTTTCGA TTCACTCGTGGGATCCGGGCGGTCCG GGGGGGACCACCACGGCTCCTCTCTT CTCGAGAATCCATACATCCCTTATCAG TGTATGGACAGCTATCTCTCGAGCACA GGTTTAGGTTCGGCCTCAATGGGAAAA TAAAATGGAGCACCTAACAACGCATCT TCACAGACCAAGAACTACGAGATCACC CCTTTCATTCTGGGGTGACGGAGGGAT CGTACCATTCGAGCCGTTTTTTTCTTGA CTCGAAATCGAAATGGGAGCAGGTTTG AAAAAGGATCTTAGAGTGTCTAGGGTT	2993	Kanamycin	Potato plastome integration site, *trnI/trnA* (*Solanum tuberosum*)	Designed in this work base on integration site used by (Daniell et al., 1998)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		GGGCCAGGAGGATCTCTTAACGCCTT CTTTTTTCTTCTCATCGGAGTTATTTCA CAAAGACTTGCCAGGGTAAGGAAGAA GGGGGGAACAAGCACACTTGGAGAGC GCAGTACAACGGAGAGTTGTATGCTGC GTTCGGGAAGGATGAATCGCTCCCGA AAAGGAATCTATTGATTCTCTCCCAATT GGTTGGACCGTAGGTGCGATGATTTAC TTCACGGGCGAGATCTCTGGTTCAAGT CCAGGATGGCCCAGCTGCGAATTCGT TTAAACTGCCTTGTCTTCATGGTGAGC AAGGGCGAGGAGCTGTTCACCGGGGT GGTGCCCATCCTGGTCGAGCTGGACG GCGACGTAAACGGCCACAAGTTCAGC GTGTCCGGCGAGGGCGAGGGCGATG CCACCTACGGCAAGCTGACCCTGAAG TTCATCTGCACCACCGGCAAGCTGCCC GTGCCCTGGCCCACCCTCGTGACCAC CTTGACCTACGGCGTGCAGTGCTTCG CCCGCTACCCCGACCACATGAAGCAG CACGACTTCTTCAAGTCCGCCATGCCC GAAGGCTACGTCCAGGAGCGCACCAT CTTCTTCAAGGACGACGGCAACTACAA GACCCGCGCCGAGGTGAAGTTCGAGG GCGACACCCTGGTGAACCGCATCGAG CTGAAGGGCATCGACTTCAAGGAGGA CGGCAACATCCTGGGGCACAAGCTGG AGTACAACTACAACAGCCACAAGGTCT ATATCACCGCCGACAAGCAGAAGAAC GGCATCAAGGTGAACTTCAAGACCCG CCACAACATCGAGGACGGCAGCGTGC AGCTCGCCGACCACTACCAGCAGAAC ACCCCCATCGGCGACGGCCCCGTGCT GCTGCCCGACAACCACTACCTGAGCA CCCAGTCCAAGCTGAGCAAAGACCCC AACGAGAAGCGCGATCACATGGTCCT GCTGGAGTTCGTGACCGCCGCCGGGA TCACTCTCGGCATGGACGAGCTGTACA AGTAAGAAGACAAGGGAGTTTAAACAA GCTTCTCCACTTGGCTCGGGGGGATAT AGCTCAGTTGGTAGAGCTCCGCTCTTG CAATTGGGTCGTTGCGATTACGGGTTG GATGTCTAATTGTCCAGGCGGTAATGA TAGTATCTTGTACCTGAACCGGTGGCT CACTTTTTCTAAGTAATGGGGAAGAGG ACCGAAACGTGCCACTGAAAGACTCTA CTGAGACAAAGATGGGCTGTCAAGAAC GTAGAGGAGGTAGGATGGGCAGTTGG TCAGATCTAGTATGGATCGTACATGGA CGGTAGTTGGAGTCGGCGGCTCTCCC AGGGTTCCCTCATCTGAGATCTCTGGG GAAGAGGATCAAGTTGGCCCTTGCGA ACAGCTTGATGCACTATCTCCCTTCAA CCCTTTGAGCGAAATGCGGCAAAAGAA AAGGAAGGAAAATCCATGGACCGACC CCATCATCTCCACCCCGTAGGAACTAC GAGATCACCCCAAGGACGCCTTCGGC ATCCAGGGGTCACGGACCGACCATAG AACCCTGTTCAATAAGTGGAACGCATT AGCTGTCCGCTCTCAGGTTGGGCAGT CAGGGTCGGAGAAGGGCAATGACTCA TTCTTAAAACCAGCGTTCTTAAGACCAA AGAGTCGGGCGGAAGGGGGGGAAAG CCCTCCGTTCCTGGTTCTCCTGTAGTT GGATCCTCCGGAACCACAAGAATCCTT				

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		AGTTAGAATGGGATTCCAACTCAGCAC CTTTTGAGTGAGATTTTGAGAAGAGTT GCTCTTTGGAGAGCACAGTACGATGAA AGTTGTAAGCTGTGTTCGGGGGGGAG TTATTGTCTATCGTTGGCCTCTATGGTA GAATCAGTCGGGGGACCTGAGAGGCG GTGGTTTACCCTGCGGCGGATGTCAG CGGTTCGAGTCCGCTTATCTCCAACTC GTGAACTTAGCCGATACAAGGCTTTAT GATAGCACCCAATTTTTCCGATTCGGC GGTTCGATCTATGAT				
Integration site	*St trnG/trnfM*	CCACTAAAGAAAAAAATCTATTCAATTA GCAAAAAAGGAGAGAGAGGGATTCGA ACCCTCGATAGTTCTTTGTTCAAAACTA TACCGGTTTTCAAGACCGGGGCTATCA ACCACTCAGCCATCTCTCCGAAAGACT ATTTTTATTTTATTCCTCCGAATAGAAC ATGGCTATAGGGGTGGATACCCCCACT ATCTGTAGAAAGATCTCAGGTGCGAAT CCACCGGTCGATCTATCTATCCGTATA TAGATATATGATCTAGCATGTCCATTTG TGAAATAAAAAAGAAAATTCCATTTCCC CCCACCCCACTCCATGTACGAATAAAG TGCGAAAGGGGGAGTAGTAATAAGTCA TATAGAATCAATGGATTCATGATAAAGT AAAATCCCTCGATGACATATTTTATCAC AATTAATATTTTTTGGCTGATAGAGGGA TCAAATGGTATATAGTTCATTTGTTGGT AGCTTGGAGGATTAAAAGCATGACTCT TGCTTTCCAATTGGCTGTTTTTGCATTA ATTGCTACTTCATTAATCTTATTGATTA GCGTACCCGTTGTATTTGCTTCTCCTG ATGGCTGGTCAAGTAACAAAAATGTTG TATTTTCTGGTACATCCTTATGGATTGG ATTAGTCTTTCTGGTGGGTATCCTTAAT TCTCTCATCTCTTGAACCTATTCGTCGC AGACCCAAAACCAAAATGACCCCCCTA ATTTTTCTCGGTTGTGAGACACATTAAA TTTGAATCTAAGTCCCCAAAGAAAATG CAAATCAAATAAAGAAAACAAAAAAATT AGAGGGGGGTCAAACTTCTTGAATAAA AAGAATATAATTAACAAAATAATTGGAA TCGTTCCGAAGAGAATATGTGTCCCGG CACTGCACAAATAAGATCCGGTTATAT ATCATATATGTGGGTACATATTGTGTAT CAAGAACAAAAAAATGCGGATATGGTC GAATGGTAAAATTTCTCTTTGCCAAGG AGAAGATGCGGGTTCGATTCCCGCTAT CCGCCCAAGATCCAAGATAAAGTAATT TTTTTACTATTTATTTTTATTATTGAATT TCATAAATAGCATTAAATATATCCTTAA ATTAAGGAGTTGGTATAGTTGGCCGTG ATAGTGTAGTGATTCTATCCCTCCCCT ACGTTTTCGAATTCGTTTAAACTGCCTT GTCTTCATGGTGAGCAAGGGCGAGGA GCTGTTCACCGGGGTGGTGCCCATCC TGGTCGAGCTGGACGGCGACGTAAAC GGCCACAAGTTCAGCGTGTCCGGCGA GGGCGAGGGCGATGCCACCTACGGCA AGCTGACCCTGAAGTTCATCTGCACCA CCGGCAAGCTGCCCGTGCCCTGGCCC ACCCTCGTGACCACCTTGACCTACGGC GTGCAGTGCTTCGCCCGCTACCCCGA CCACATGAAGCAGCACGACTTCTTCAA GTCCGCCATGCCCGAAGGCTACGTCC	3372	Kanamycin	Potato plastome integration site, *trnG/trnfM* (*Solanum tuberosum*)	Designed in this work base on integration site used by (Bock and Maliga, 1995)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		AGGAGCGCACCATCTTCTTCAAGGACG ACGGCAACTACAAGACCCGCGCCGAG GTGAAGTTCGAGGGCGACACCCTGGT GAACCGCATCGAGCTGAAGGGCATCG ACTTCAAGGAGGACGGCAACATCCTG GGGCACAAGCTGGAGTACAACTACAA CAGCCACAAGGTCTATATCACCGCCGA CAAGCAGAAGAACGGCATCAAGGTGA ACTTCAAGACCCGCCACAACATCGAGG ACGGCAGCGTGCAGCTCGCCGACCAC TACCAGCAGAACACCCCCATCGGCGA CGGCCCCGTGCTGCTGCCCGACAACC ACTACCTGAGCACCCAGTCCAAGCTGA GCAAAGACCCCAACGAGAAGCGCGAT CACATGGTCCTGCTGGAGTTCGTGACC GCCGCCGGGATCACTCTCGGCATGGA CGAGCTGTACAAGTAAGAAGACAAGG GAGTTTAAACAAGCTTCTTACCAGAATA AGATGTTGCGGAGACAGGATTTGAACC CGTGACCTCAAGGTTATGAGCCTTGCG AGCTACCAAACTGCTCTACCCCGCGCC AAAGATAAGAACTGAAAACTAATAGATA AACAAGGATTAAATGCGCCCCTCCACC CTATCTGTACAAATAGAATAGCCCATTT ATACAGAATGGTAAAGGGGCTTCTATG ATCATCGACCATAGAAATAGAAATGAA GCGTTAATCCTTACCAACTCGATCTTG TTGCTCCTGGCAACAAACATGCATGAA CCATTTCACGAAGTATGTGTCCGGATA GGCCAAAGTCTCGATAGTTAGCTCTCG GCCTTCCGGTCAAAAAACAACGTCGAT GAAGGCGTGTAGGTGCACTATTCCGT GGTAGGGATTGTAACTTTCCATAAATTT CCCATTTGTCACTCAACGACGGAACCT TGCTTATTTCTTTTTTTGAGGATCGACG AATCGAATGATATTTCTGTTCCAATTTT TGCCTCTTCTTCTCCCTCTGAATCAAAC TTTTCCTTGCCATAATGGTTGAGTTCCT ATTAGTATCCATGATACAAGTCGAATC CTAGATGTAGAAATAGAAGAAGGTGGA CCCCCTTCTCCGTCGAAAGAAATGAGA TTATCGCGAATACACACATTAATTAACC AAATTTGCCCGACGTAGAGGCAATCAA GAAAGCCGCATAAGTGAATATATAACC TACAGAAAAGTGAGCTAATCCAACCAA TCTTGCTTGTACAATAGAAAGGGCCAC TGGTTTATCTCTCCATCGAATCAAATTG GCCAAAGGTGTGCGTTCGTGAGCCCA TGCTAAAGTTTCAATCAATTCCTGCCAA TATCCACGCCAAGAAATTAAGAACATA AATCCAGTAGCCCAAACAAGATGTCCA AATAAGAACATCCATGCCCAAACCGAT AAACTATTCATACCAAAAGGGTTATATC CATTGATAAGTTGTGAAGAGTTTAACC ATAAATAATCCCTTAACCAGCCCATCAA ATAAGTGGAAGATTCATTAAACTGTGA AACGTTACCTTGCCATAATGTGATGTG CTTCCAATGCCAATAAAAAGTAACCCA TCCAATAGTATTTAACATCCAAAAAACT GCCAAATAAAACGCGTCCCATGCCGAA ATATCACAAGTACCGCCTCGTCCTGGG CCATCGCACGGAAAACTATAACCGAAA TCCTTTTTATCTGGCATTAACTTGGAAC CACGTGCATCTAAAGCACCTTTTACTA AGATCAATGTAGTTGTATGTAAACCAA				

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		GGGCAATAGCATGATGAACCAAAAAGT CTCCAGGACCTATTGTTAAGAATAATG AATTACTATTTTCATTAACAGCATTTAA CCAACCCGGCAACCAGATGCTTCGAC CCGCATTGAATGCTGGGCCAGTCGTT GAAGATAAAAG				
Integration site	*St trnV/3' rps12*	AACGAGTTCACCGCCGTATGGCTGAC CGGCGATTACTAGCGATTCCGGCTTCA TGCAGGCGAGTTGCAGCCTGCAATCC GAACTGAGGACGGGTTTTTGGGGTTA GCTCACCCTCGCGGGATCGCGACCCT TTGTCCCGGCCATTGTAGCACGTGTGT CGCCCAGGGCATAAGGGGCATGATGA CTTGACGTCATCCTCACCTTCCTCCGG CTTATCACCGGCAGTCTGTTCAGGGTT CCAAACTCAACGATGGCAACTAAACAC GAGGGTTGCGCTCGTTGCGGGACTTA ACCCAACACCTTACGGCACGAGCTGA CGACAGCCATGCACCACCTGTGTCCG CGTTCCCGAAGGCACCCCTCTCTTTCA AGAGGATTCGCGGCATGTCAAGCCCT GGTAAGGTTCTTCGCTTTGCATCGAAT TAAACCACATGCTCCACCGCTTGTGCG GGCCCCCGTCAATTCCTTTGAGTTTCA TTCTTGCGAACGTACTCCCCAGGCGG GATACTTAACGCGTTAGCTACAGCACT GCACGGGTCGATACGCACAGCGCCTA GTATCCATCGTTTACGGCTAGGACTAC TGGGGTATCTAATCCCATTCGCTCCCC TAGCTTTCGGCTCTCAGTGTCAGTGTC GGCCCAGCAGAGTGCTTTCGCCGTTG GTGTTCTTTCCGATCTCTACGCATTTCA CCGCTCCACCGGAAATTCCCTCTGCCC CTACCGTACTCCAGCTTGGTAGTTTCC ACCGCCTGTCCAGGGTTGAGCCCTGG GATTTGACGGCGGACTTAAAAAGCCAC CTACAGACGCTTTACGCCCAATCATTC CGGATAACGCTTGCATCCTCTGTATTA CCGCGGCTGCTGGCACAGAGTTAGCC GATGCTTATTCCCCAGATACCGTCATT GCTTCTTCTCCGGGAAAAGAAGTTCAC GACCCGTGGGCCTTCTACCTCCACGC GGCATTGCTCCGTCAGGCTTTCGCCCA TTGCGGAAAATTCCCCACTGCTGCCTC CCGTAGGAGTCTGGGCCGTGTCTCAG TCCCAGTGTGGCTGATCATCCTCTCGG ACCAGCTACTGATCATCGCCTTGGTAA GCTATTGCCTCACCAACTAGCTAATCA GACGCGAGCCCCTCCTCGGGCGGATT CCTCCTTTTGCTCCTCAGCCTACGGGG TATTAGCAGCCGTTTCCAGCTGTTGTT CCCCTCCCAAGGGCAGGTTCTTACGC GTTACTCACCCGTCCGCCACTGGAAAC ACCGTTTCCCGTGTTTCCCGTCCGACT TGCATGTGTTAAGCATGCCGCCAGCGT TCATCCTGAGCCAGGATCGAACTCTCC ATGAGATTCATAGTTGCATTACTTATAG CTTCCTTGTTCGTAGACAAAGCGGATT CGGAATTGTCTTTCATTCCAAGGCATA ACTTGTATCCATGCGCTTCATATTCGC CCGGAGTTCGCTCCCAGAAATATAGTC ATCCCTGCCCCCTCACGTCAATCCCAC GAGCCTCTTATCCATTCTCATTGAACG ACGGCGGGGGAGCAAATCCAACTAGA AAAACTCACATTGGGCTTAGGGATAAT	3726	Kanamycin	Potato plastome integration site, *trnV/3'rps12* (*Solanum tuberosum*)	Designed in this work base on integration site used by (Staub and Maliga, 1992)

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		CAGGCTCGAACTGATGACTTCCACCAC GTCAAGGTGACACTCTACCGCTGAGTT ATATCCCTTCCCCGCCCCATCGAGAAA TAGAACTGACTAATCCTAAGTCAAAGG GTCGAGAAACTCAACGCCACTATTCTT GAACAACTTGGAGCCGGGCCTTCTTTT CGCACTATTACGGATATGAAAATAATG GTCAAAATCGGATTCAATTGTCAAGAA TTCGTTTAAACTGCCTTGTCTTCATGGT GAGCAAGGGCGAGGAGCTGTTCACCG GGGTGGTGCCCATCCTGGTCGAGCTG GACGGCGACGTAAACGGCCACAAGTT CAGCGTGTCCGGCGAGGGCGAGGGC GATGCCACCTACGGCAAGCTGACCCT GAAGTTCATCTGCACCACCGGCAAGCT GCCCGTGCCCTGGCCCACCCTCGTGA CCACCTTGACCTACGGCGTGCAGTGC TTCGCCCGCTACCCCGACCACATGAA GCAGCACGACTTCTTCAAGTCCGCCAT GCCCGAAGGCTACGTCCAGGAGCGCA CCATCTTCTTCAAGGACGACGGCAACT ACAAGACCCGCGCCGAGGTGAAGTTC GAGGGCGACACCCTGGTGAACCGCAT CGAGCTGAAGGGCATCGACTTCAAGG AGGACGGCAACATCCTGGGGCACAAG CTGGAGTACAACTACAACAGCCACAAG GTCTATATCACCGCCGACAAGCAGAAG AACGGCATCAAGGTGAACTTCAAGACC CGCCACAACATCGAGGACGGCAGCGT GCAGCTCGCCGACCACTACCAGCAGA ACACCCCCATCGGCGACGGCCCCGTG CTGCTGCCCGACAACCACTACCTGAG CACCCAGTCCAAGCTGAGCAAAGACC CCAACGAGAAGCGCGATCACATGGTC CTGCTGGAGTTCGTGACCGCCGCCGG GATCACTCTCGGCATGGACGAGCTGTA CAAGTAAGAAGACAAGGGAGTTTAAAC AAGCTTCTGCCCCTATCGGAAATAGGA TTGACTACCGATTCCGAAGGAACTGGA GTTACATCTCTTTTCCATTCAAGAGTTC TTATGCGTTTCCACGCCCCTTTAAGAC CCCGAAAAATGGACAAATTCCTTTTCTT AGGAACACATACAAGATTCGTCACTAC AAAAAGGATAATGGGAACCCTACCATT AACTACTTCATTTATGAATTTCATAGTA ATAGAAATACATGTCCTACCGAGACAG AATTTGGAACTTGCTACCCTCTTGCCT AGCAGGCAAAGATTTACCTCCGTGGAA AGGATGATTCATTCGGATCGACATGAG AGTCCAACTACATTGCCAGAATCCATG TTGTATATTTGAAAGAGGTTGACCTCCT TGCTTCTCTCATGGTACACTCCTCTTC CCGCCGAGCCCCTTTTCTCCTCGGTCC ACAGAGACAAAATGTAGGACTGGTGCC AACAGTTCATCACGGAAGAAAGGACTC ACTAAGTCGGGATCACTAACTAATACT AATCTAATATAATAGTCTAATATATCTA ATATAATAGAAAATACTAATATAATAGA AAAGAACTGTCTTTTCTGTATACTTTCC CCGGTTCCGTTGCTACCGCGGGCTTTA CGCAATCGATCGGATTAGATAGATATC CCTTCAACATAGGTCATCGAAAGGATC TCGGATACCCACCAAAGTACGAAAGCC AGGATCTTTCAGAAAACGGATTCCTAT TCAAAGAGTGCATAACCGCATGGATAA				

Supplementary Table S1 continued

Type	Name	Sequence	bp	Resistance	Description	Reference
		GCTCACACTAACCCGTCAATTTGGGAT CCAAATTAGAGATTTTCCTTGGGAGGT ATCGGGAAGGATTTGGAATGGAATAAT ATCGATTCATACAGAAGAAAAGGTTCT CTATTGATTCAAACACTGTACCTATGG GATAGGGATCGAGGAAGGGGAAAAAC CGAAGATTTCACATGGTACTTTTATCAA TCTGATTTATTTCGTACCTTTCGTTCAA TGAGAAAATGGGTCAAATTCTACAGGA TCAAACCTATGGGACTTAAGGAATGAT ATAAAAAAAAGAGAGGGAAAATATTCA TATTAAATAAATATGAAGTAGAAGAACC CAGATTCCAAATGAACAAATTCAAACTT GAAAAGGATCTTCCTTATTCTTGAAGAA TGAGGGGCAAAGGGATTGATCAAGAA				

Supplemental Table S2

(A) - Level-1 vectors (forward orientation)

Vector	position	Plate id.	Selection	Reference
pICH47732	1	B3	Carbenicillin	(Engler et al., 2014)
pICH47742	2	C3	Carbenicillin	(Engler et al., 2014)
pICH47751	3	D3	Carbenicillin	(Engler et al., 2014)
pICH47761	4	E3	Carbenicillin	(Engler et al., 2014)
pICH47772	5	F3	Carbenicillin	(Engler et al., 2014)
pICH47781	6	G3	Carbenicillin	(Engler et al., 2014)
pICH47791	7	H3	Carbenicillin	(Engler et al., 2014)

(B) - Level-1 vectors (reverse orientation)

Vector	position	Plate id.	Selection	Reference
pICH47802	1	A4	Carbenicillin	(Engler et al., 2014)
pICH47811	2	B4	Carbenicillin	(Engler et al., 2014)
pICH47822	3	C4	Carbenicillin	(Engler et al., 2014)
pICH47831	4	D4	Carbenicillin	(Engler et al., 2014)
pICH47841	5	E4	Carbenicillin	(Engler et al., 2014)
pICH47852	6	F4	Carbenicillin	(Engler et al., 2014)
pICH47861	7	G4	Carbenicillin	(Engler et al., 2014)

Supplemental Table S3

(A) – Example of Level-2 End-Linkers compatible with the kit

Vector	Position	Plate id.	Selection	Reference
pICH41722	1	B5	Spectinomycin	(Engler et al., 2014)
pICH41744	2	C5	Spectinomycin	(Engler et al., 2014)
pICH41766	3	D5	Spectinomycin	(Engler et al., 2014)
pICH41780	4	E5	Spectinomycin	(Engler et al., 2014)
pICH41800	5	F5	Spectinomycin	(Engler et al., 2014)
pICH41822	6	G5	Spectinomycin	(Engler et al., 2014)
pICH50866	7	H5	Spectinomycin	(Engler et al., 2014)

(B) - Level-2 Chl-Vectors

Vector	Species	Integration site	Selection	Reference
242	Nt	*trnI-trnA*	Kanamycin	this work
243	Nt	*trnG-trnfM*	Kanamycin	this work
244	Nt	*trnV-3'rps12*	Kanamycin	this work
245	Nt	*atpB-accD*	Kanamycin	this work
246	Zm	*trnI-trnA*	Kanamycin	this work
247	Zm	*trnV-3'rps12*	Kanamycin	this work
248	St	*trnI-trnA*	Kanamycin	this work
249	St	*trnG-trnfM*	Kanamycin	this work
250	St	*trnV-3'rps12*	Kanamycin	this work

Supplemental Table S4

N.	Primer Name	Sequence 5'-3'	Target
1	GOI-1-Chl-FW	ATGCCTATTGCTGTAGGAAT	Full-length *GOI-1*
2	GOI-2-Chl-FW	ATGAGAATTGCTAAAGTAAGAG	Full-length *GOI-2*
3	GOI-3-Chl-FW	ATGTCTGCTTCTTTACCTGC	Full-length *GOI-3*
4	GOI-4-Chl-FW	ATGTCTGCTTCTTTACCTGC	Portion of *GOI-4*
5	aadA-full-FW	ATGGCAGAAGCGGTGATC	Full-length *aadA*
6	PaccD-FW	GGAGCGATGCGAATTTGACA	*accD* promoter
7	IEE-FW	GGAGTAGGATCGTTTATTTACA	intercistronic expression element (IEE)
8	PpsbA-FW	GGAGAATAAAAAGCCTTCCATT	*psbA* promoter
9	Int-St-trnG-trnfM-FW	CTATCCGCCCAAGATCCAAGATAAAG	*trnG* of plastome
10	Lux-C-FW	CTGAACATGCTCCTTCTCATGC	portion of *Lux-C*
11	Lux-D-FW	TGGATCATTTTGCTGGATTAGCTG	portion of *Lux-D*
12	Lux-A-FW	CAACCTCCTGGAGAATCTCATCA	portion of *Lux-A*
13	Lux-B-FW	CTTCTCAACAACAACAATTTGAAGC	portion of *Lux-B*
14	Lux-E-FW	GGGACCTGATAGATTTTCTGCTTC	portion of *Lux-E*
15	Lux-G-FW	GGAGGATCTGTAAAAGAATCTGCT	portion of *Lux-G*
16	aadA-frag-FW	GTGCACGACGACATCATTCC	portion of *aadA*
17	mEmerald-frag-FW	GTGACCACCTTGACCTACGG	portion of *mEmerald*
18	Prrn-Nic-FW	TGGATAAGAGGCTCGTGGGA	*rrn* promoter
19	RB-FW	GTTTACCCGCCAATATATCC	Right border of T-DNA
20	GOI-1-Chl-RV	TTACATTCTAAATTGTTCTACAGC	Full-length *GOI-1*
21	GOI-2-Chl-RV	TTAAGAATGTTCTCTTTTATCATA	Full-length *GOI-2*
22	GOI-3-Chl-RV	TCATTGATCATCTCTAGGATTAG	Full-length *GOI-3*
23	GOI-4-Chl-RV	AACTTCAGGGTCAGCTTGC	Portion of *GOI-4*
24	aadA-full-RV	TTATTTGCCGACTACCTTGGT	Full-length *aadA*
25	petD-3UTR-RV	AGCGATCTATTGGACTTACTTA	3'UTR of *petD*
26	psbA-3UTR-RV	AGCGAAAAAAAGAAAGGAGCAATAG	3'UTR of *psbA*
27	Int-St-trnG-trnfM-RV	TCTATTTGTACAGATAGGGTGGAGGG	*trnfM* of plastome
28	Lux-C-RV	CCAAGGAGCAATAGCTACAGTTT	Portion of *Lux-C*
29	Lux-D-RV	ACCCAATCATCTCCATCAGAAGT	Portion of *Lux-D*
30	Lux-A-RV	CAGCAGTCATACAAGTAGGTTGTT	Portion of *Lux-A*
31	Lux-B-RV	CTTTTCCTACAGCATGTTGTTCAAT	Portion of *Lux-B*
32	Lux-E-RV	AAAGCTCTAGCATATACCCAAGGA	Portion of *Lux-E*
33	Lux-G-RV	AGACATTCCAAAAGGTCCACAT	Portion of *Lux-G*
34	aadA-frag-RV	TCGCCTTTCACGTAGTGGAC	Portion of *aadA*
35	mEmerald-frag-RV	TCTCGTTGGGGTCTTTGCTC	Portion of *mEmerald*
36	rrnB-3UTR-RV	TGTTGTTTGTCGGTGAACGC	3'UTR of *rrnB*
37	LB-RV	TGGCAGGATATATTGTGGTG	Left border of T-DNA

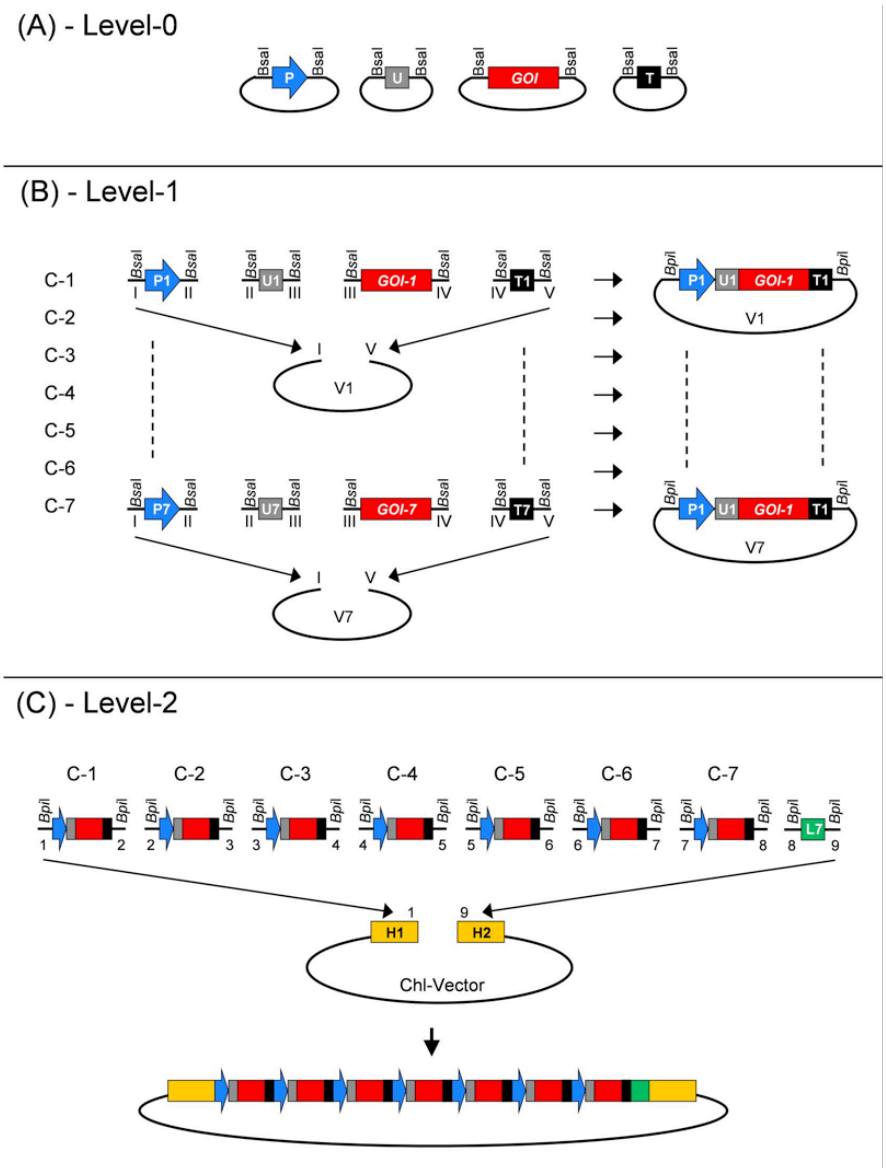


Figure 1: Schematic representation of Golden Gate assembly using the MoChlo kit. (A) 4 vectors containing different Level-0 modules flanked by *Bsa*I sites: P: promoter (blue); U: 5' untranslated region (UTR; gray); GOI: gene of interest (red); T: terminator/ 3' untranslated region (UTR; black). (B) Level-1 golden gate assembly. The *Bsa*I restriction digestion releases Level-0 parts (P1-7, U1-7, GOI1-7 and T1-7) equipped with 5' and 3' overhangs (I-V) used for directional cloning. The 5' and 3' overhangs (I and V, respectively) of P-U-GOI-T cassettes (C1-7) are compatible with Level-1 vectors (V1-7). The Level-1 assembly results in seven Level-1 vectors containing seven cassettes (C1-7) flanked by *Bpi*I sites. (C) Level-2 Golden Gate assembly. The *Bpi*I restriction digestion of Level-1 vectors releases cassettes (C1-7) equipped with 5' and 3' overhangs (1-8) able to provide directional cloning from C1 to C7. An end-linker (L7, green) equipped with compatible 5' and 3' overhangs (8 and 9, respectively) allows integration of the Level-2 operon in between the homologous sequences (H1 and H2) of a chloroplast transformation vector (Chl-Vector: destination vector Level-2).

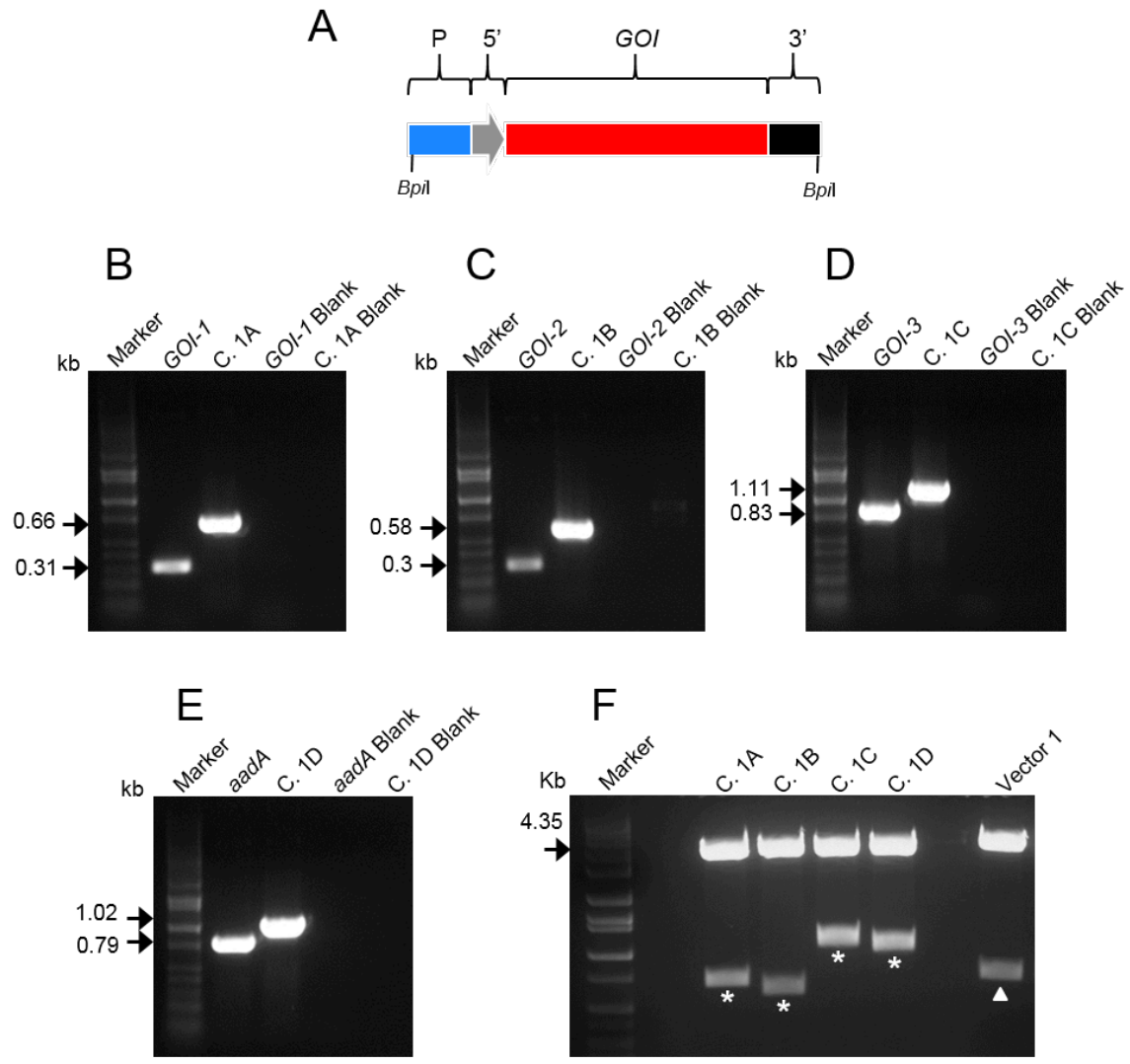


Figure 2: Level-1 cloning for a 4-gene operon-1. (A) Schematic representation of a Level-1 cassette. Promoter (P: blue); 5' untranslated region (5': gray); gene of interest (*GOI 1-3* or *aadA*: red); terminator/ 3'untranslated region (3': black). *Bpi*I sites flanking the Level-1 cassette are indicated. (B-E) PCR to confirm Level 1 assembly. Specific primers for the internal gene (B: *GOI-1*; C: *GOI-2*; D: *GOI-3*; E: *aadA*) and the promoter/3'UTR of cassettes 1A, 1B, 1C and 1D (B: C. 1A; C: C. 1B; D: C. 1C; E: C. 1D) have been used. DNA bands at the predicted molecular weight (kb) confirm correct Level-1 assembly of all cassettes. Negative controls (blanks) of PCR B-E are indicated. (F) *Bpi*I restriction digestion of Level-1 vectors. The 4 Level-1 cassettes released after digestion are indicated with asterisks (size indicated in Table 1). The empty vector 1 (vector position 1 forward) is used as positive control. The *LacZα* fragment (0.616 kb) release after digestion of vector 1 is indicated with a triangle. The molecular weight (kb) of backbone vector is indicated with black arrows. DNA markers are shown in the pictures.

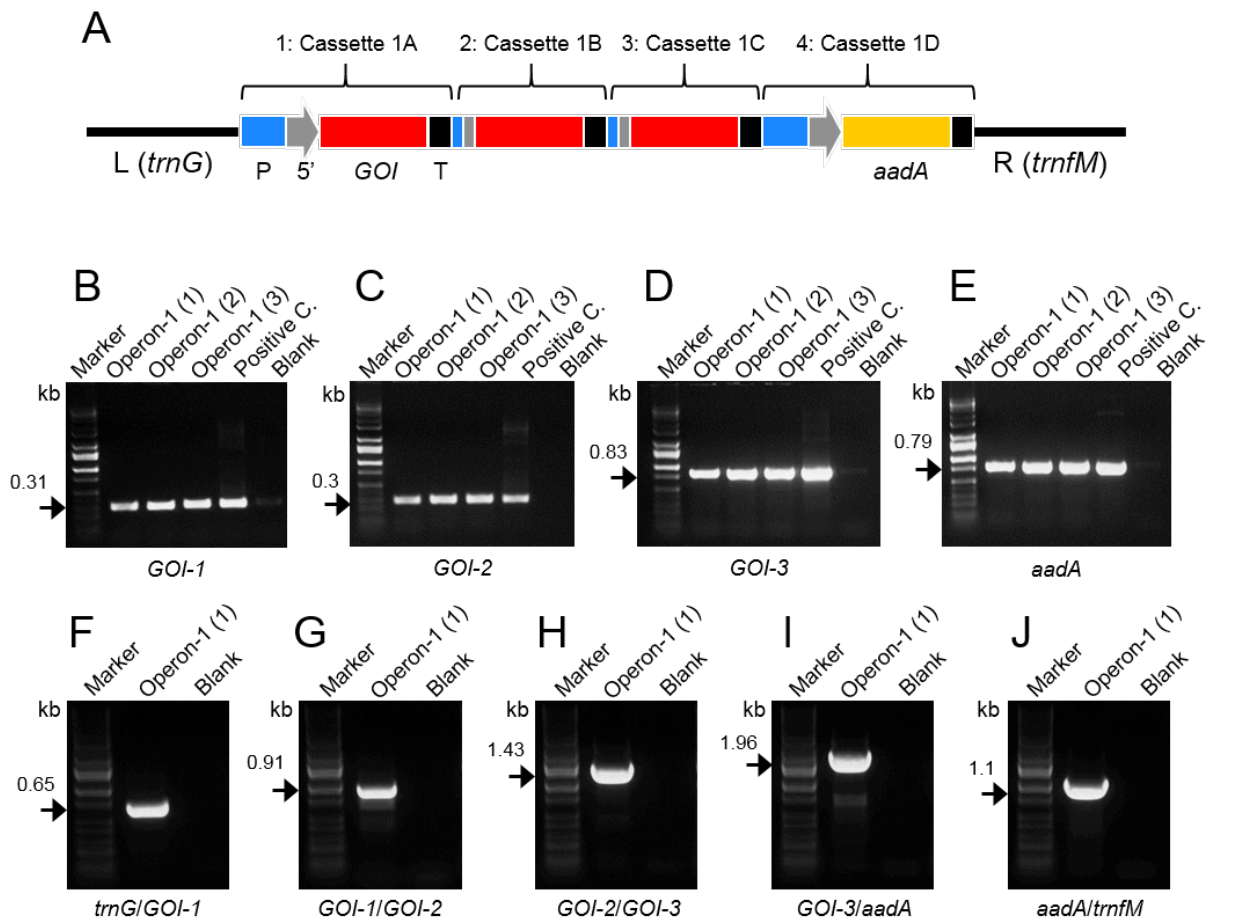


Figure 3: Level-2 cloning for operon-1. (A) Schematic representation of Level-2 operon-1. Level-1 cassettes 1A, 1B, 1C and 1D are assembled in position 1, 2, 3 and 4 of operon-1, respectively. Promoter (P: blue); 5' untranslated region (5': gray); gene of interest (*GOI 1-3*: red); gene for selection (*aadA*: yellow); terminator/ 3' untranslated region (3': black). The operon-1 is cloned in a Level-2 vector between left (L; 1.17 kb) and right (R; 1.45 kb) arms homologous to the *trnG/trnfM* site of potato plastome. (B-E) PCR to confirm the presence of all genes of operon-1 (3 purified vectors 1-3). Specific primers for *GOI-1* (B), *GOI-2* (C), *GOI-3* (D) and *aadA* (E) have been used. DNA bands at the predicted molecular weight (kb, black arrows B-E) in both purified plasmid (operon-1) and positive controls (positive C.) confirm correct Level-2 assembly of operon-1. (F-J) PCR to confirm the correct order of 4 genes in the operon-1. Pair or primers specific for *trnG*/*GOI-1* (F), *GOI-1*/*GOI-2* (G), *GOI-2*/*GOI-3* (H), *GOI-3*/*aadA* (I) and *aadA*/*trnfM* (J) have been used. DNA bands at the predicted molecular weight confirm the correct order (kb, black arrows F-J). Negative controls (blanks) for PCR and DNA markers are shown in panels (B-J).

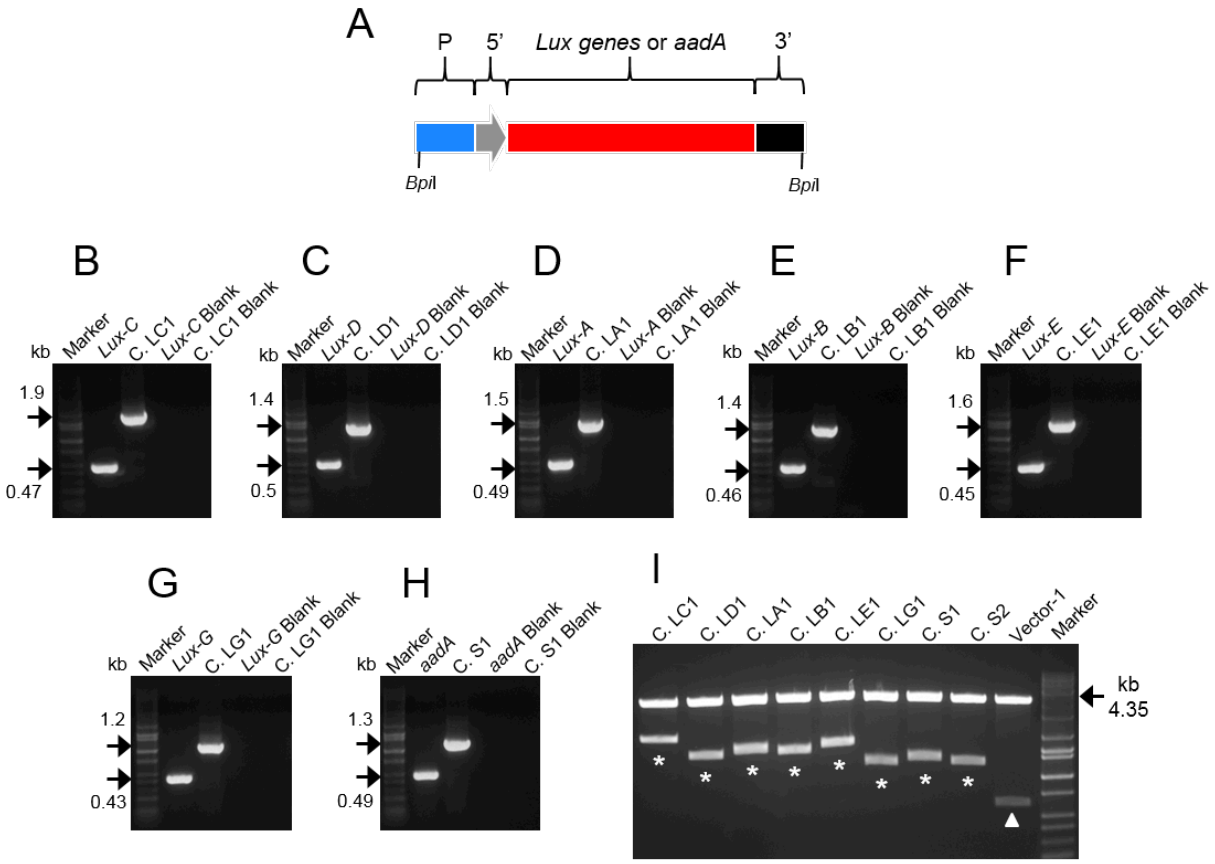


Figure 4: Level-1 cloning for the *Lux*-operon for plastid engineering. (A) Schematic representation of a Level-1 cassette. Promoter (P: blue); 5' untranslated region (5': gray); *Lux* (*C*, *D*, *A*, *B*, *E* or *G*) or *aadA* gene for selection (red); terminator/ 3' untranslated region (3': black). *Bpi*I sites flanking the Level-1 cassette are indicated. (B-H) PCR to confirm Level 1 assembly. Specific primers for *Lux* genes (B: *Lux-C*; C: *Lux-D*; D: *Lux-A*; E: *Lux-B*; F: *Lux-E*; G: *Lux-G*), gene for selection (H: *aadA*) and the promoter/3'UTR of cassettes LC1, LD1, LA1, LB1, LE1, LG1 and S1 (B: C. LC1; C: C. LD1; D: C. LA1; E: C. LB1; F: C. LE1; G: C. LG1; H: C. S1) have been used. DNA bands at the predicted molecular weight (kb, black arrows) confirm correct Level-1 assembly of all cassettes. Negative controls (blanks) of PCR B-H are indicated. (J) *Bpi*I restriction digestion of Level-1 vectors. Level-1 cassettes released after digestion are indicated with asterisk (cassette size Table 2). The empty vector 1 (position 1 forward) is used as positive control. The *LacZα* fragment release after digestion of vector-1 is indicated with triangle (0.616 kb). The molecular weight (kb) of the backbone vector is indicated with black arrows. DNA markers are shown in the pictures.

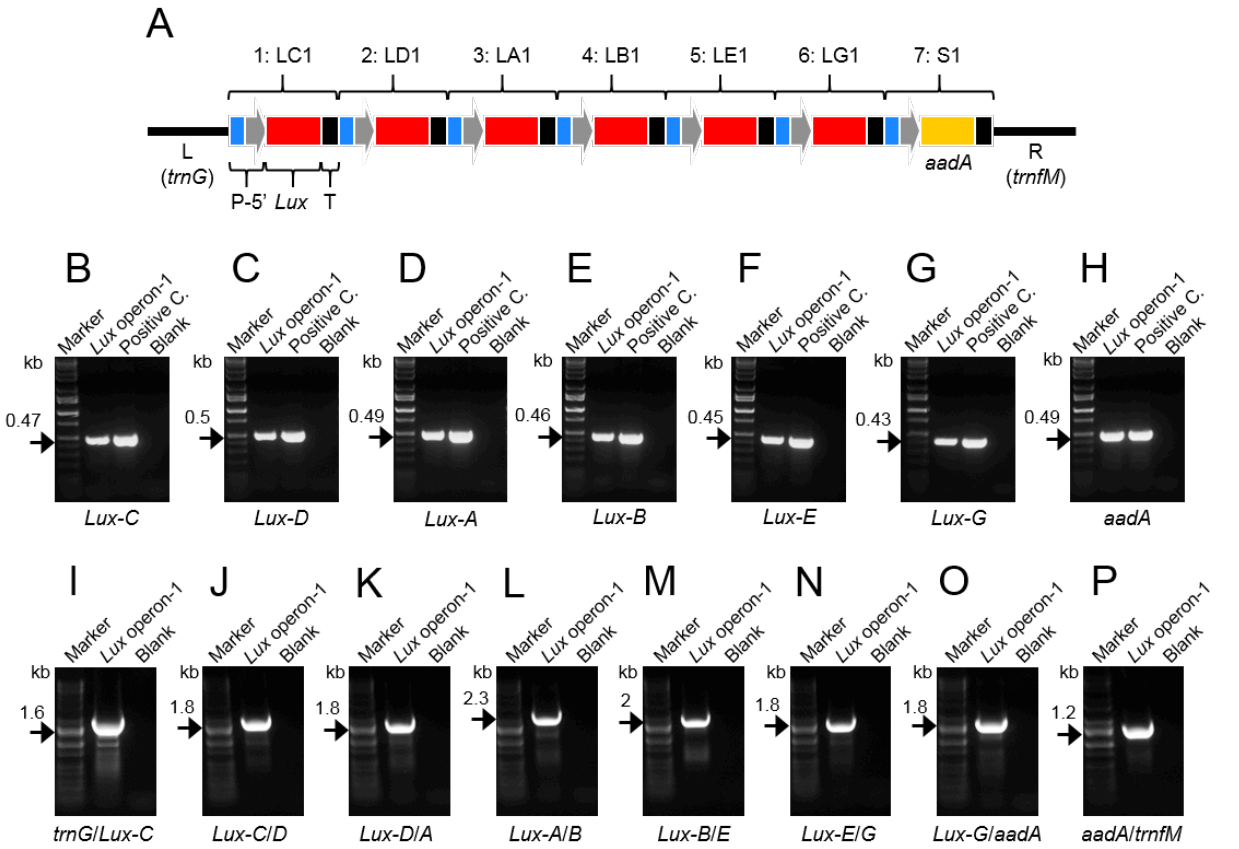


Figure 5: Level-2 cloning for the 7-gene *Lux*-operon. (A) Schematic representation of Level-2 assembly of *Lux* operon-1. Level-1 cassettes LC1, LD1, LA1, LB1, LE1, LG1 and S1 are assembled in operon position 1, 2, 3, 4, 5, 6 and 7, respectively. Promoter (P: blue); 5' untranslated region (5': gray); *Lux* (*C*, *D*, *A*, *B*, *E* or *G*: Red); gene for selection (*aadA*: yellow); terminator/ 3' untranslated region (3': black). The *Lux* operon-1 is cloned in Level-2 vector between left (L; 1.17 kb) and right (R; 1.45 kb) arms homologous to the *trnG/trnfM* site of potato plastome. (B-H) PCR of all genes of the *Lux* operon-1. Primers specific for *Lux-C* (B), *Lux-D* (C), *Lux-A* (D), *Lux-B* (E), *Lux-E* (F), *Lux-G* (G) and *aadA* (H) have been used. DNA bands at the predicted molecular weight (kb, black arrows B-H) for both purified plasmid (*Lux* operon-1) and positive controls (positive C.) confirmed the presence of all genes. (F-J) PCR to check correct gene order in the *Lux* operon-1. Pair or primers specific for *trnG*/*Lux-C* (I), *Lux-C*/*Lux-D* (J), *Lux-D*/*Lux-A* (K), *Lux-A*/*Lux-B* (L), *Lux-B*/*Lux-E* (M), *Lux-E*/*Lux-G* (N), *Lux-G*/*aadA* (O) and *aadA*/*trnfM* (P) have been used. DNA bands (black arrows I-P) at the predicted molecular weight (kb) confirm correct gene order. Negative controls (blanks) of PCR and DNA markers are shown in the pictures (B-P).

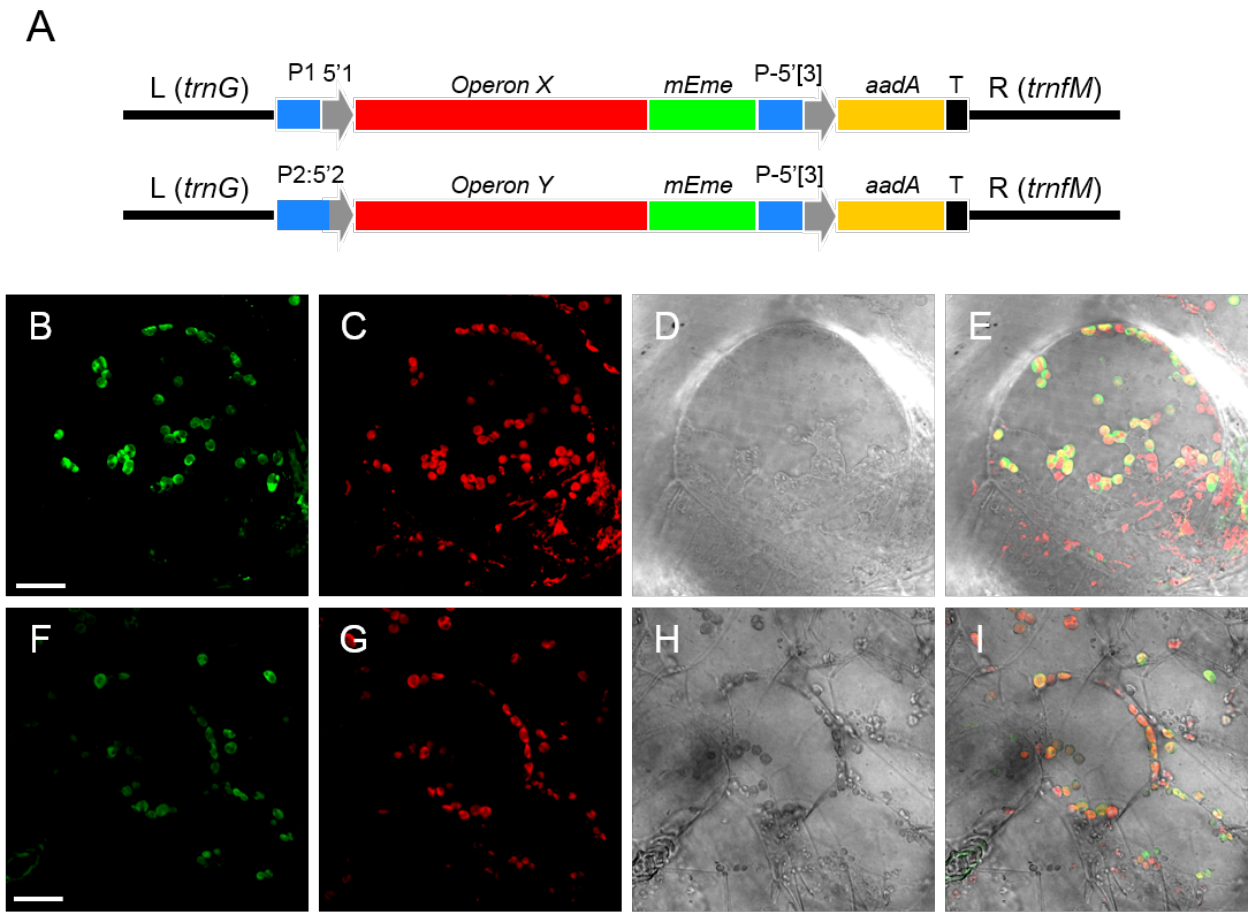


Figure 6: An example of *in vivo* validation of two MoChlo-enabled constructs in potato chloroplasts. (A) Schematic representation of test operons controlled by two examples of promoter::5'UTR combinations. The operon (red) is controlled by the *rrn* promoter (P1) along with a synthetic ribosome binding site as 5'UTR (5'1; blue and gray: Operon X) or the promoter::5'UTR fusion of *accD* gene (P2:5'2; blue and gray: Operon Y). The expression of *mEmerald* (*mEme*; green) is under the control of the operon promoter, while the selection gene, *aadA* (yellow), is under the control of both promoter::5'UTR (P3-5'; blue and gray) and terminator/ 3'UTR (T: black) of the *psbA* gene. (B-I) Confocal images of transplastomic green callus transformed with Operon X (B-E) or Operon Y (F-I). In both cases the mEmerald fluorescent signal is localized to chloroplasts. (B and F) mEmerald signal; (C and G) chlorophyll autofluorescence; (D and H) bright-field; (E and I) merged images. Scale bars = 20 µm (B-E and F-I).

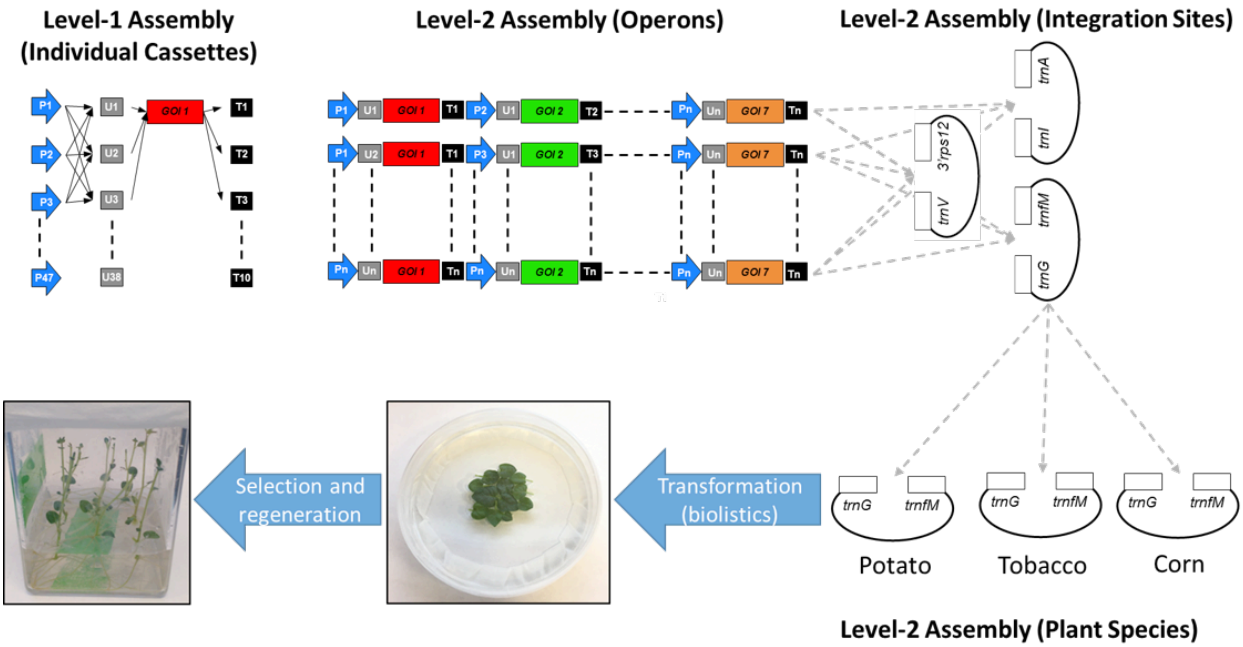


Figure 7: Schematic of MoChlo kit for combinatorial assembly of operons for metabolic engineering. In the current iteration of the kit, combinatorial assembly of individual cassettes (Level-1 constructs) can be conducted using 47 promoters, 38 5'UTRs, 9 promoter:5'UTR fusions and 10 terminators/ 3'UTRs. In addition, the user has the ability to assemble up to 7 individual cassettes, each with their own regulatory elements and genes of interest (GOI). Further, during Level-2 assembly, the user can choose from 3 different integration sites (*trnI-trnA*; *trnV-3'rps12*; *trnG-trnfM*) from 3 different plant species (*Nicotiana tabacum; Zea mays; Solanum tuberosum*). After completing assembly, the user would then bombard the Level-2 constructs into leaf disks, select, and regenerate transplastomic plants.

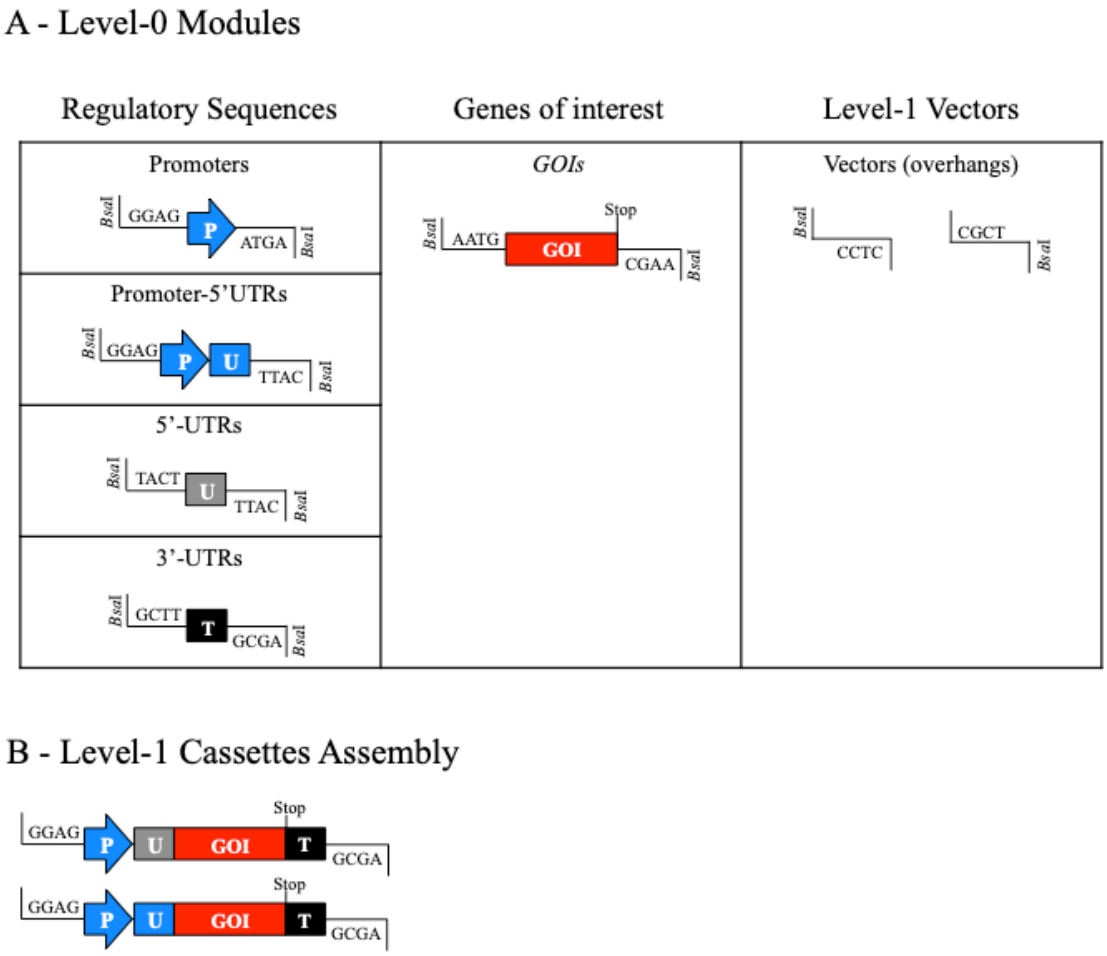


Figure S1: Level-0 modules and directional cloning into Level-1 cassettes. (A) List of Level-0 modules available in the MoChlo kit. P: promoter (blue); P-U: promoter fused with 5' untranslated region UTR (blue); U: 5' untranslated region UTR (gray); T: terminator/ 3' untranslated region UTR (black); GOI: gene of interest. Both 5' and 3' overhangs of level-0 modules and Level-1 destination vectors are indicated. *Bsa*I sites are shown at both ends of Level-0 modules and vectors. (B) Possible Level-1 cassettes originated by directional cloning of different combinations of Level-0 modules (regulatory sequences along with GOIs). For each cassette overhangs compatible with Level-1 vectors are indicated.

A - Level-1 Cassettes and End-Linkers

Position-1 (Fw or Rv)	End-Linker position-1	Level-2 Chl-vectors (overhangs)
BpiI TGCC **C-1** CGTT *BpiI*	*BpiI* GCAA **L1** CCCT *BpiI*	*BpiI* ACGG GGGA *BpiI*
Position-2 (Fw or Rv)	End-Linker position-2	
BpiI GCAA **C-2** TGAT *BpiI*	*BpiI* ACTA **L2** CCCT *BpiI*	
Position-3 (Fw or Rv)	End-Linker position-3	
BpiI ACTA **C-3** AATG *BpiI*	*BpiI* TTAC **L3** CCCT *BpiI*	
Position-4 (Fw or Rv)	End-Linker position-4	
BpiI TTAC **C-4** GTCT *BpiI*	*BpiI* CAGA **L4** CCCT *BpiI*	
Position-5 (Fw or Rv)	End-Linker position-5	
BpiI CAGA **C-5** ACAC *BpiI*	*BpiI* TGTG **L5** CCCT *BpiI*	
Position-6 (Fw or Rv)	End-Linker position-6	
BpiI TGTG **C-6** CTCG *BpiI*	*BpiI* GAGC **L6** CCCT *BpiI*	
Position-7 (Fw or Rv)	End-Linker position-7	
BpiI GAGC **C-7** ACGG *BpiI*	*BpiI* TGCC **L7** CCCT *BpiI*	

B - Level-2 Operons Assembly

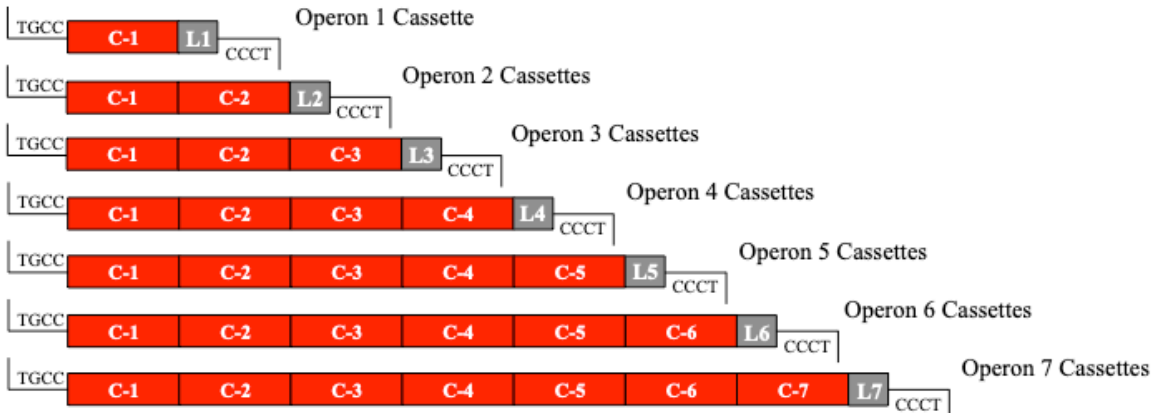


Figure S2: Directional cloning of Level-1 cassettes into operons of different lengths. (A) Table resuming Level-1 cassettes (C1-7) equipped with specific overhangs for directional cloning of Level-2 operons. Cassettes are either in forward or reverse orientation depending on the vectors used in Level-1 assembly. End-linkers for each position in the operon equipped with specific overhangs are indicated. Overhangs of a chloroplast destination vector Level-2 are indicated. *Bpi*I sites at both ends of cassettes, end-linkers and vector are shown. (B) Assembly of cassettes and end-linkers forming Level-2 operons of different lengths (operons composed by 1 to 7 cassettes). 5' and 3' overhangs compatible with chloroplast destination vectors are indicated.

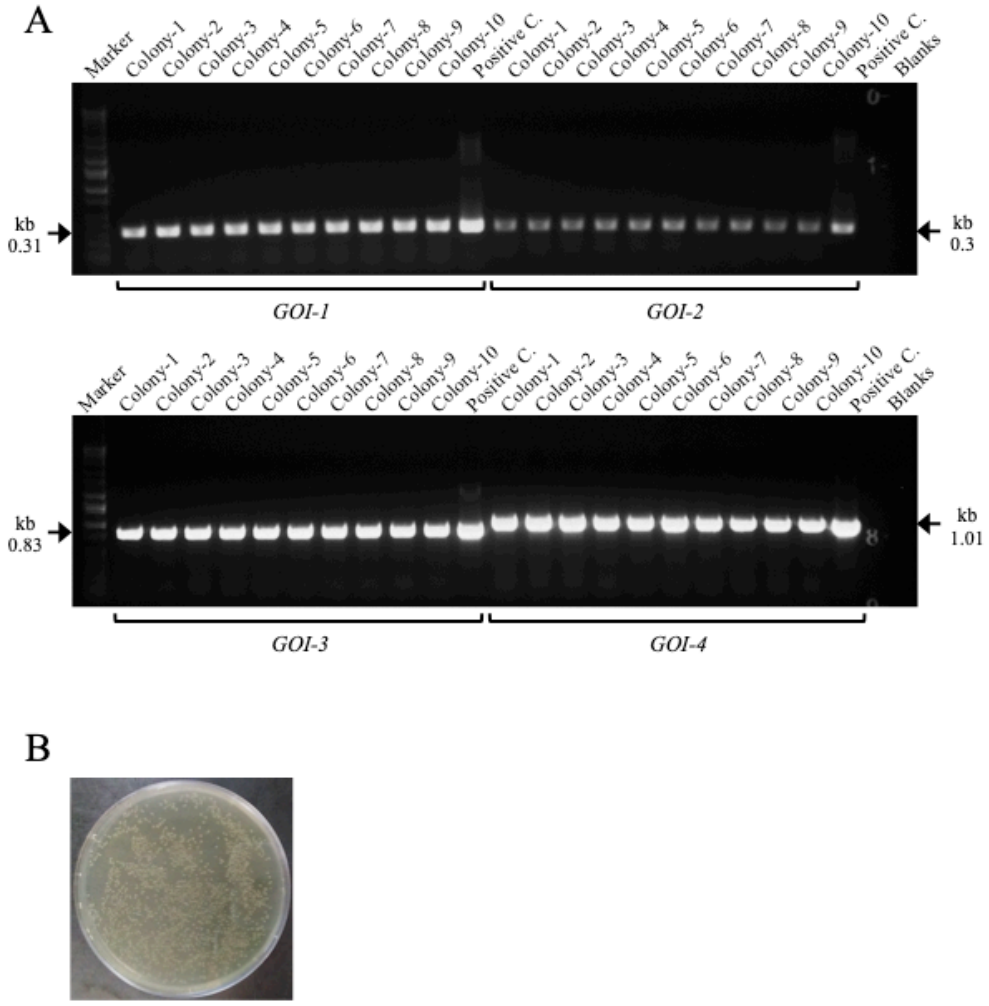


Figure S3: Level-1 assembly of GOIs. (A) Example of PCR reactions on bacteria colonies transformed with ligation products of different Level-1 constructs. 10 colonies per each Level-1 cassettes have been tested. Specific primers for the internal genes (*GOI-1*; *GOI-2*; *GOI-3*; *GOI-4*) have been used to check the presence of Level-1 cassettes. All tested bacteria colonies are positives for the presence of Level-1 cassettes confirmed by DNA bands at the predicted molecular weight (kb). Positive and negative controls (blanks) have been used to support the PCR reactions. DNA markers are shown in the picture. (B) Example of Blue-White screen of bacteria transformed with ligation products of Level-1 constructs. All bacteria colonies are white confirming high efficiency of Level-1 assembly.

A

Name	Promoter	5'UTR	GOI	3'UTR	Size (kb)	Position (fw)
Cassette 2A	mP*rrn* (46%)-RBS		*GOI-1*	*petD*-3'UTR	0.640	1
Cassette 3A	mP*rrn* (46%)-RBS		*GOI-2*	*petD*-3'UTR	0.631	1
Cassette 4A	mP*rrn* (46%)-RBS		*GOI-3*	*petD*-3'UTR	1.162	1
Cassette 5A	mP*rrn* (60%)-RBS		*GOI-1*	*petD*-3'UTR	0.640	1
Cassette 6A	mP*rrn* (60%)-RBS		*GOI-2*	*petD*-3'UTR	0.631	1
Cassette 7A	mP*rrn* (60%)-RBS		*GOI-3*	*petD*-3'UTR	1.162	1
Cassette 8A	mP*rrn* (20%)-RBS		*GOI-1*	*petD*-3'UTR	0.594	1
Cassette 9A	P*rrn*-RBS		*GOI-1*	*petD*-3'UTR	0.615	1
Cassette 10A	P*rrn*-RBS		*GOI-2*	*petD*-3'UTR	0.606	1
Cassette 11A	P*rrn*-RBS		*GOI-3*	*petD*-3'UTR	1.137	1
Cassette 2B	mP*rrn* (46%)-RBS		*GOI-2*	*petD*-3'UTR	0.631	2
Cassette 2C	IEE	Synth RBS	*GOI-4*	*petD*-3'UTR	1.870	3
Cassette 3C	IEE	Synth RBS	*GOI-5*	*petD*-3'UTR	2.659	3
Cassette 4C	mP*rrn* (20%)-RBS		*GOI-3*	*petD*-3'UTR	1.116	3

B

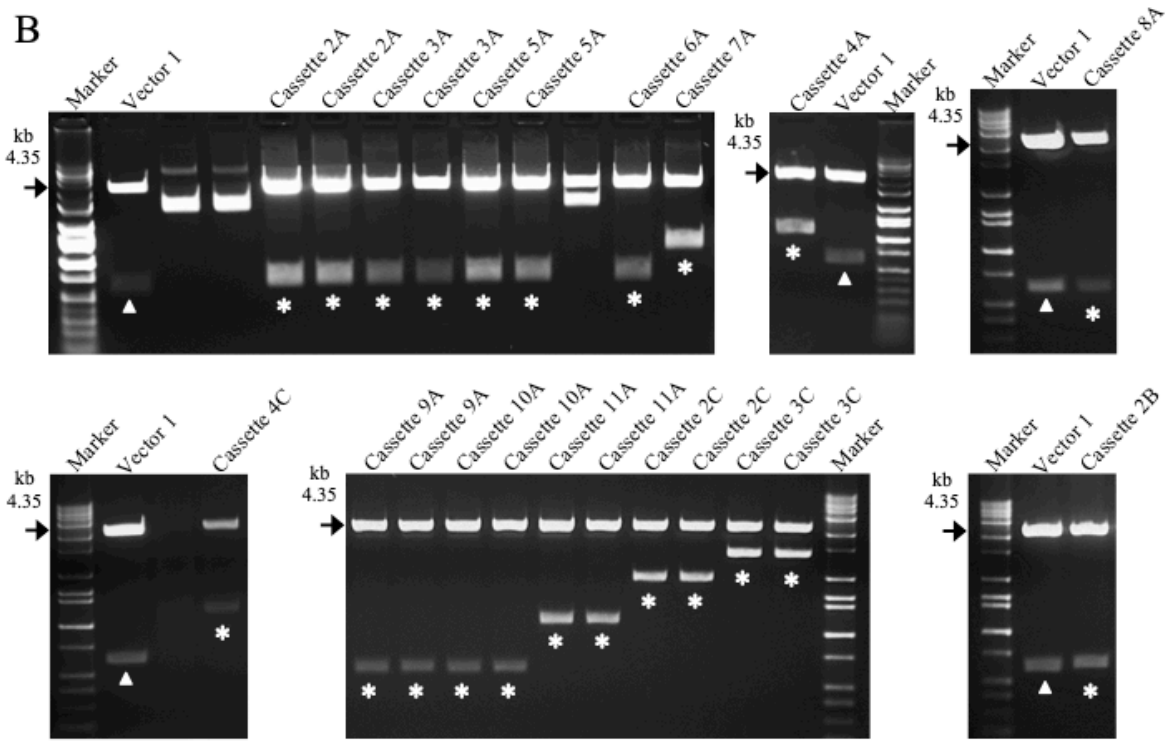


Figure S4: Examples of Level-1 cassettes synthesized using the MoClo kit. (A) Table describing Level-1 cassettes cloned in 4 different Level-1 vectors (forward position 1-4). P*rrn*-RBS (promoter of *rrn* gene along with a ribosome binding site); mP*rrn* (20-46-60%)-RBS (promoter of *rrn* gene 20, 46 or 60% active along with a ribosome binding site); IEE (intercistronic expression element); synth RBS (synthetic ribosome binding site); *GOI-1-5* (gene of interest 1-5); *petD*-3'UTR (3'UTR of *petD* gene). (B) Restriction digestion (*Bpi*I) of aforementioned Level-1 vectors. The Level-1 cassettes released after digestion are indicated with asterisk (size kb are indicated in the table). The empty vector-1 (vector position 1 forward) is used as positive control. The *LacZα* operon release after digestion of vector 1 is indicated with triangle (0.616 kb). The molecular weight (kb) of backbone vector is indicated with black arrows. DNA markers are shown in the pictures.

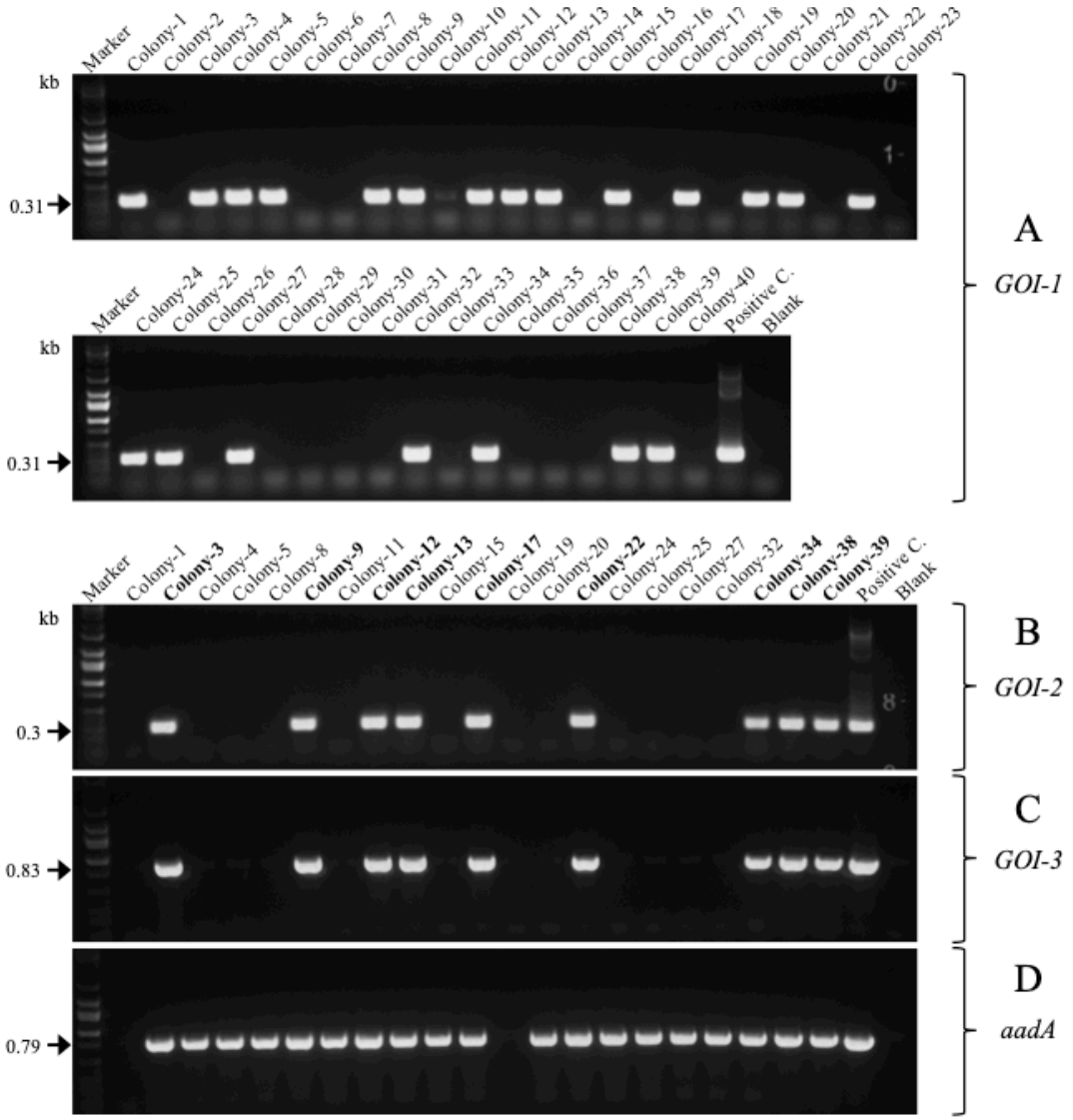


Figure S5: Colony screening for Level-2 assembly of operon-1 using *Bbs*I in the Golden Gate reaction. (A) PCR reactions on bacteria colonies (1-40) transformed with the ligation product of Level-2 assembly of operon-1. In the first screen, a pair of primers were used to check the presence of *GOI-1*. DNA bands at the predicted molecular weight (kb; black arrow) confirmed the presence of *GOI-1* in 21 colonies. (B-D) PCR reactions on *GOI-1* positive colonies to check the presence of the other 3 genes. Reactions using primers specific for *GOI-2*, *GOI-3* and *aadA* are shown in panel B, C and D, respectively. DNA bands at the predicted molecular weight (kb; black arrow) confirmed the presence of *GOI-2*, *GOI-3* and *aadA* genes in 9 *GOI-1* positive colonies (indicated in bold). The efficiency of Level-2 assembly of a 4-genes operon using *Bbs*I is about 22.5%. Positive and negative controls (blanks) along with DNA markers are shown in the pictures.

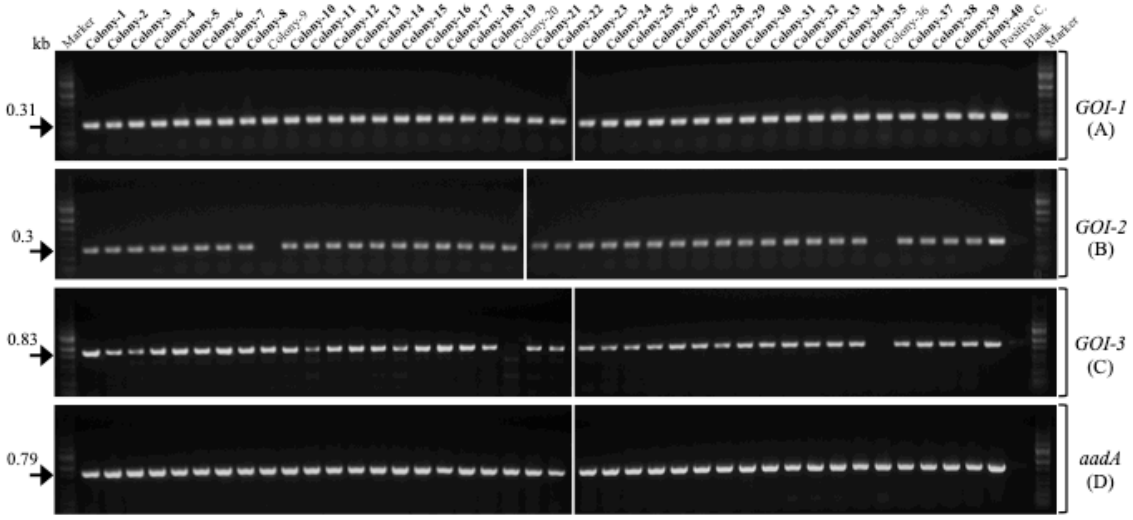
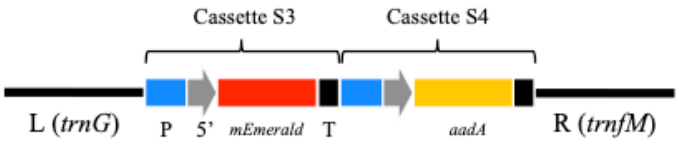


Figure S6: Colony screening for Level-2 assembly of operon-1 using *Bpi*I in the Golden Gate reaction. (A) PCR reactions on bacteria colonies (1-40) transformed with the ligation product of Level-2 assembly of operon-1. In the first screen, a pair of primers has been used to check the presence of *GOI-1*. DNA bands at the predicted molecular weight (kb; black arrow) confirmed the presence of *GOI-1* in all 40 colonies. (B-D) PCR reactions on *GOI-1* positive colonies to check the presence of the other 3 genes. Reactions using primers specific for *GOI-2*, *GOI-3* and *aadA* are shown in panel B, C and D, respectively. DNA bands at the predicted molecular weight (kb; black arrow) confirmed the presence of *GOI-2*, *GOI-3* and *aadA* in 37 positive colonies for *GOI-1* (indicated in bold). The efficiency of Level-2 assembly of a 4-genes operon using *Bpi*I is about 92.5%. Positive and negative controls (blanks) along with DNA markers are shown in the pictures.

A



B

Name	Promoter (P)	5'UTR (5')	GOI	3'UTR (3')	Size (kb)	Position (fw)
Cassette-S3	P*rrn*	P*rrn*-5'UTR	*mEmerald*	*psbA*-3'UTR	1.096	1
Cassette-S4	P*psbA*	P*psbA*-5'UTR	*aadA*	*psbA*-3'UTR	1.058	2

C

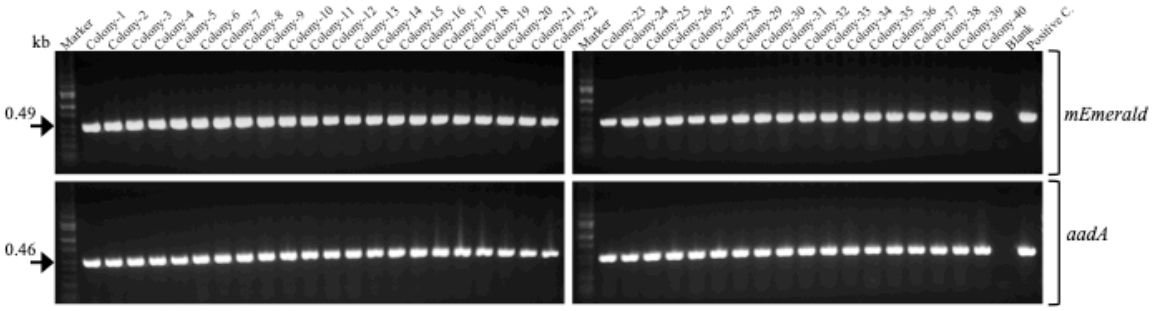


Figure S7: Level-2 assembly of 2-genes selection cassettes using *Bbs*I in the Golden Gate reaction. (A) Schematic representation of Level-2 selection operon-1. Level-1 cassettes S3 and S4 are assembled in position 1 and 2, respectively. Promoter (P: blue); 5' untranslated region (5': gray); *mEmerald* gene (red); *aadA* gene for selection (yellow); terminator/ 3' untranslated region (3': black). The operon is cloned in a Level-2 vector between left (L; 1.17 kb) and right (R; 1.45 kb) arms homologous to the *trnG/trnfM* site of potato plastome. (B) Table describing level-1 cassette S3 and S4. Promoter (P); 5'UTR (5' untranslated region); gene of interest (*GOI*); terminator/ 3'UTR (3' untranslated region); cassette size (kb) and position in the operon (forward position 1-2) are indicated. P*rrn* (promoter of *rrn* gene); P*psbA* (promoter of *psbA* gene); P*rrn*-5'UTR (5'UTR of *rrn* promoter); P*psbA*-5'UTR (5'UTR of *psbA* promoter); *mEmerald* (gene encoding mEmerald fluorescent protein); *aadA* gene for selection (encoding the 3"-adenylyltransferase); *psbA*-3'UTR (3'UTR of *psbA* gene). (C) PCR reactions on bacteria colonies (1-40) transformed with the ligation product of Level-2 assembly of selection operon-1. A pair of primers specific either for *mEmerald* or *aadA* were used. DNA bands at the predicted molecular weight (kb, black arrow) confirm the presence of both genes in all 40 colonies, indicating 100% efficiency of Level-2 assembly. Positive and negative controls (blanks) along with DNA markers are shown in the pictures.

A

Name	Position-1	Position-2	Position-3	Position-4
Operon-1	Cassette-1A	Cassette-1B	Cassette-1C	Cassette-1D
Operon-2	Cassette-2A	Cassette-1B	Cassette-1C	Cassette-1D
Operon-3	Cassette-5A	Cassette-1B	Cassette-1C	Cassette-1D
Operon-4	Cassette-5A	Cassette-2B	Cassette-2C	Cassette-1D
Operon-5	Cassette-9A	Cassette-2B	Cassette-2C	Cassette-1D

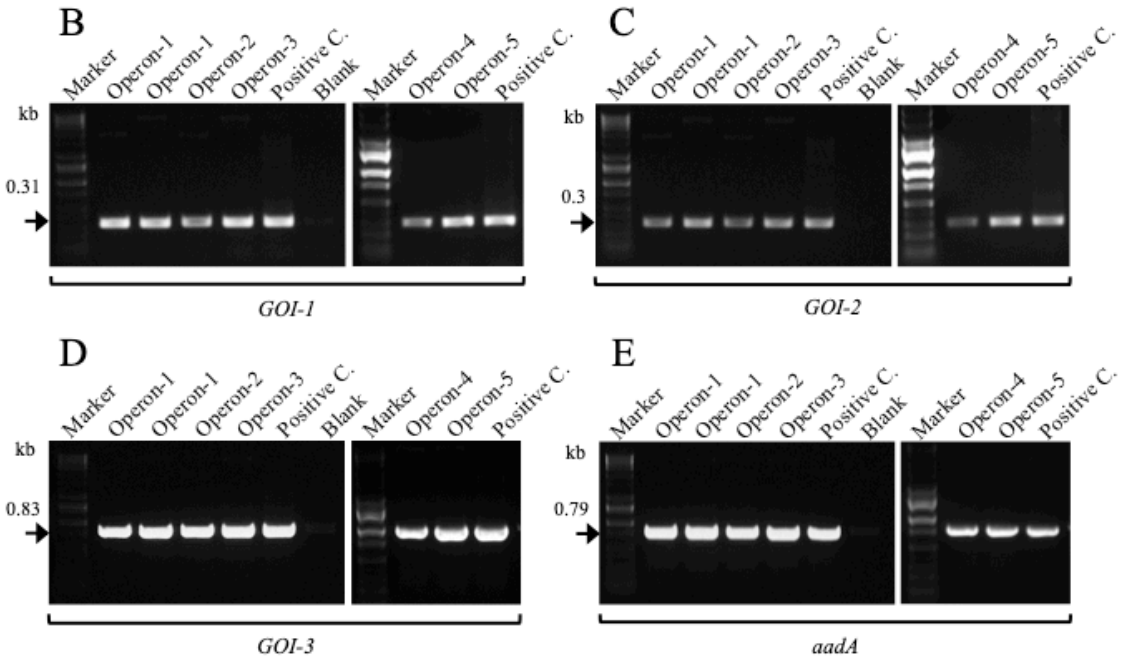


Figure S8: Level-2 assembly of 5 different 4-genes operons. (A) Table describing different Level-2 operons (from 1 to 5). For each operon the Level-1 cassettes and positions are indicated. A description of Level-1 cassettes is shown in Table 1 and Supplemental Figure S4, A. (B-E) PCR reactions using primers specific for *GOI-1* (B), *GOI-2* (C), *GOI-3* (D) and *aadA* (E). The presence of DNA bands at the predicted molecular weight (kb, black arrows) confirm the presence of 4 genes in all operons: operon-1 (1,2); operon-2; operon-3; operon-4; and operon-5. Positive and negative (blanks) controls along with DNA markers are shown in the pictures (B-E).

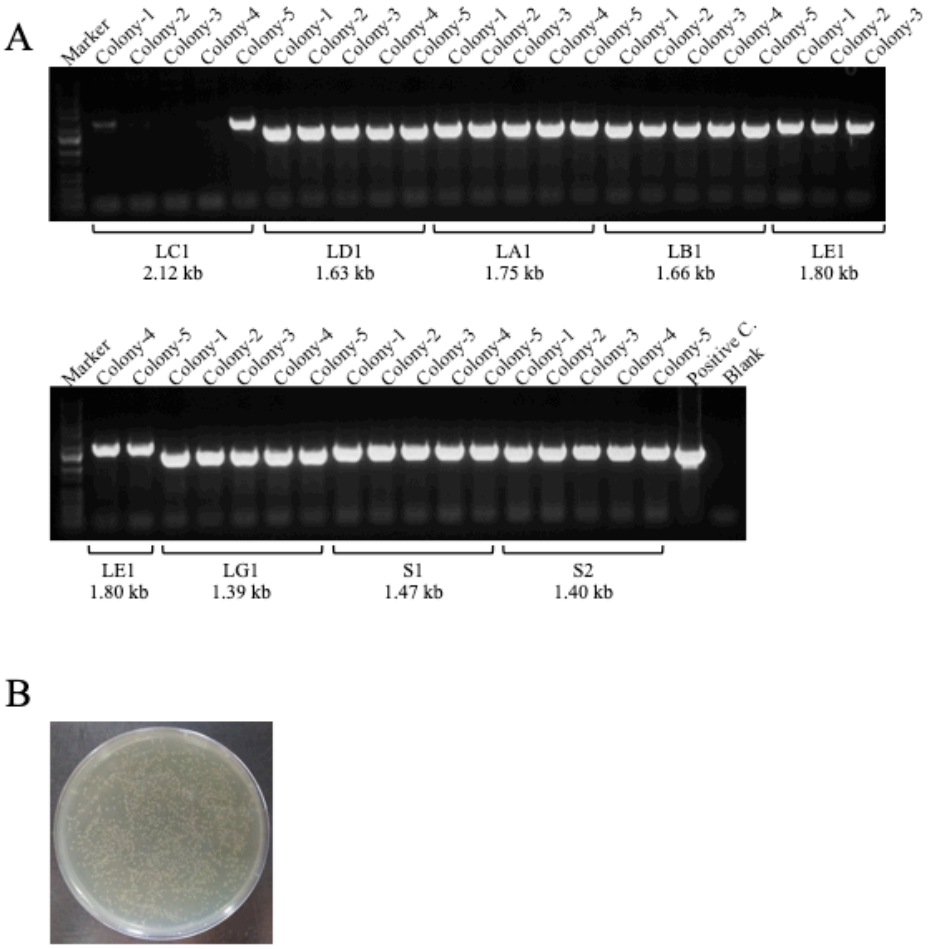


Figure S9: Level-1 assembly of Level-1 *Lux*-cassettes. (A) Example of PCR reactions on bacteria colonies transformed with ligation products of different Level-1 constructs for *Lux* operon (Cassette-LC1, LD1, LA1, LB1, LE1, LG1) and selection cassettes (Cassette-S1, S2). 5 colonies (colony 1-5) per each constructs have been tested. A pair of primers specific for RB/LB located at both sites of the cassettes have been used to check the presence of all Level-1 cassettes. The presence of DNA bands at the predicted molecular weight (kb) confirms that all tested bacteria colonies, except three for LC1, are positives. Positive and negative controls (blanks) along with DNA markers are shown in the picture. (B) Example of Blue-White screen of bacteria transformed with ligation products of Level-1 constructs indicated above. All bacteria colonies are white confirming high efficiency of Level-1 assembly.

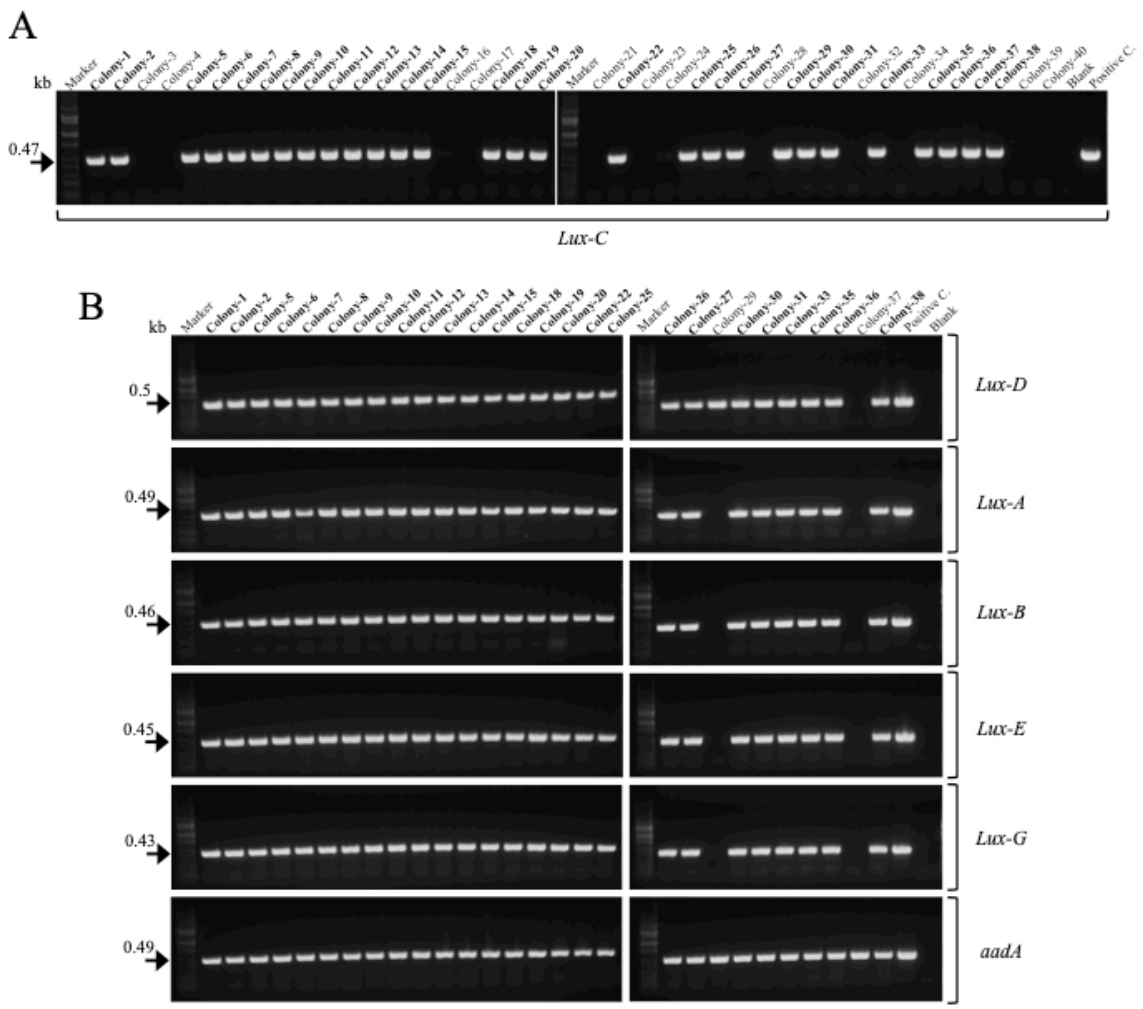


Figure S10: Colony screening for Level-2 assembly of *Lux* operon-1 using *Bpi*I in the Golden Gate reaction. (A) PCR reactions on bacteria colonies (1-40) transformed with the ligation product of Level-2 assembly of *Lux* operon-1. A pair of primers specific for *Lux-C* has been used for the first PCR screen. DNA bands at the predicted molecular weight (kb, black arrow) confirmed the presence of *Lux-C* in 28 colonies. (B) PCR reactions on *Lux-C* positive colonies to check the presence of the other 5 *Lux* genes along with *aadA*. Specific primers for *Lux-D*, *Lux-A*, *Lux-B*, *Lux-E*, *Lux-G* and *aadA* have been used for reactions indicated panels. DNA bands at the predicted molecular weight (kb, black arrow) confirm the presence of all *Lux* genes together with *aadA* in 26 colonies (indicated in bold). The efficiency of Level-2 assembly of a 7-genes operon using *Bpi*I is about 65%.Positive and negative controls (blanks) along with DNA markers are shown in the pictures.

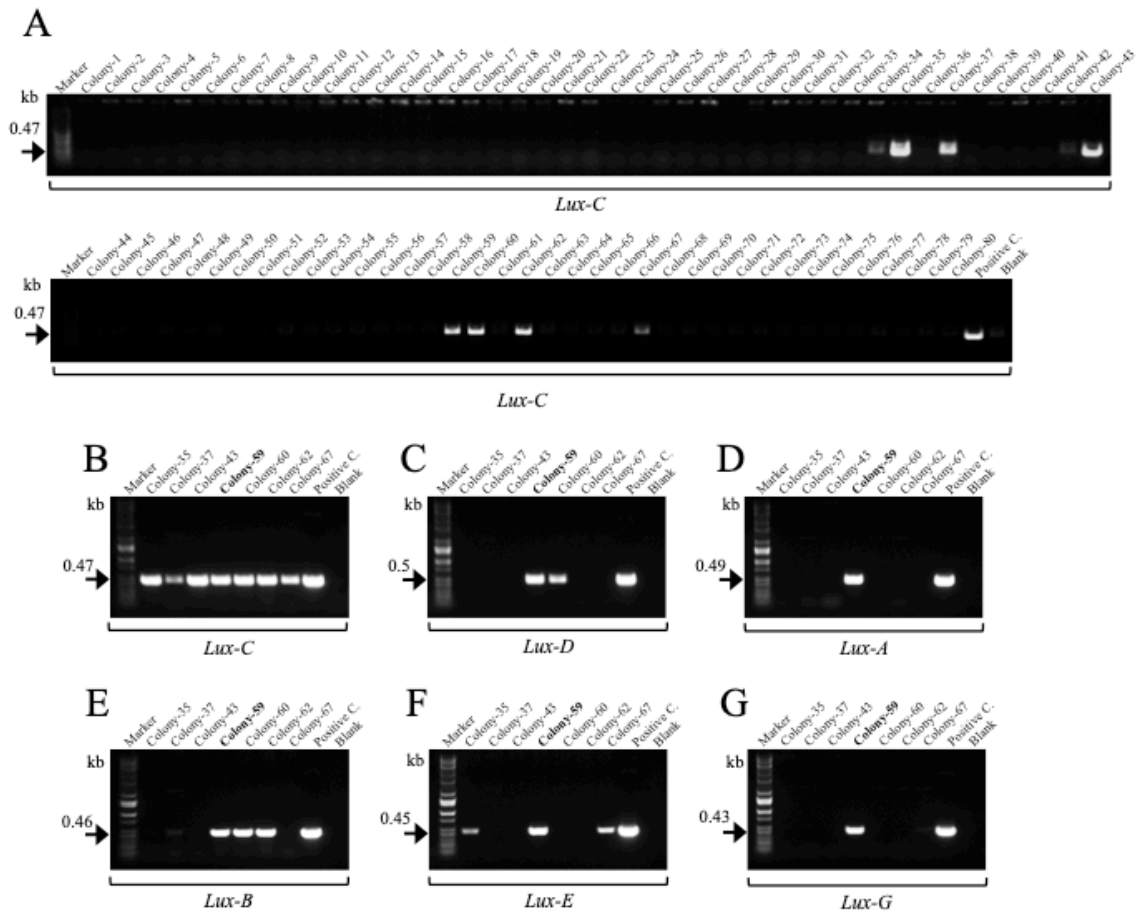


Figure S11: Colony screening for Level-2 assembly of *Lux* operon-1 using *Bbs*I in the Golden Gate reaction. (A) PCR reactions on bacteria colonies (1-80) transformed with the ligation product of Level-2 assembly of *Lux* operon-1. All 80 colonies have been grown on spectinomycin confirming the presence of *aadA* gene. A pair of primers specific for *Lux-C* has been used for the first PCR screen. DNA bands at the predicted molecular weight (kb, black arrow) confirmed the presence of *Lux-C* in 7 colonies. (B-G) PCR reactions on *Lux-C* positive colonies to check the presence of the other 5 *Lux* genes. Specific primers for *Lux-C*, *Lux-D*, *Lux-A*, *Lux-B*, *Lux-E* and *Lux-G* have been used for reactions shown in panel B, C, D, E, F and G, respectively. DNA bands at the predicted molecular weight (kb, black arrow) confirm the presence of all *Lux* genes in one colony (indicated in bold). The efficiency of Level-2 assembly of a 7-genes operon using *Bbs*I is about 1.25%. Positive and negative controls (blanks) along with DNA markers are shown in the pictures.

VITA

Alex Pfotenhauer was born in Nashville, Tennessee. He received his B.S. in Food Science from the University of Tennessee in 2017. From 2017-2021 he remained in the Food Science department for grad school, but shifted focus into synthetic biology. His mentors Scott Lenaghan and Neal Stewart opened the Center for Agricultural Synthetic Biology during this time, where Alex worked on a variety of projects related to biotechnology and synthetic biology. He is on track to graduate with his PhD in summer 2021. In his free time, he enjoys spending time outdoors, with his dogs and girlfriend Alyssa, or working in the garage on small engine repair and other mechanical projects.