METABOLIC ENGINEERING TO ENHANCE BIOFUEL AND BIOPRODUCTS PRODUCTION IN MICROORGANISMS

A Dissertation

by

MAN LI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Joshua S. Yuan
Committee Members,	Peter M. Rentzepis
	Elizabeth (Betsy) Pierson
	Susie S. Dai
Head of Department,	Leland S. Pierson

August 2020

Major Subject: Plant Pathology

ABSTRACT

The negative impact of burning fossil fuels and oil supply concerns have prompted researchers to look for alternative renewable fuel sources. Sustainable lignocellulosic and photosynthetic biofuels are promising renewable energy resources that reduce greenhouse gas emission and pose least risks to the food supply. However, the metabolic pathways involved in lignocellulosic biomass utilization and subsequent biofuel or bioproduct production in microorganisms are still unclear, which hinders high yield biofuel production through genetic engineering. In the present study, systems biology guided genetic engineering technologies and strategies have been systematically investigated to pave a biological route to produce renewable fuels and bioproducts. Two effective cell factories, photosynthetic *Synechococcus elongatus* and ligninolytic *Pseudomonas putida*, have been designed and evaluated to produce the sustainable limonene and polyhydroxyalkanoates (PHAs), respectively. First, proteomics, transcriptomics and metabolomics technologies have been employed to find the key limiting steps in the limonene biosynthesis pathway (MEP). The comprehensive proteomics analysis and subsequent genetic engineering results highlighted that regulations in the MEP were not the rate limiting factors for limonene productivity in cyanobacteria. Further metabolomics analysis suggested that the carbon partitions between the primary and secondary metabolites could have a strong impact on limonene productivity. By mutating the sucrose biosynthesis pathway in *S. elongatus* PCC7942, the limonene productivity was significantly increased. Interestingly, the increased glycogen content in sucrose phosphate synthase mutated strain led to the hypothesis of redirecting the carbon flux to the MEP pathway by mutating the glycogen biosynthesis pathway. This hypothesis was supported by an increased yield of 14.9 mg/L by the glucose-1-phosphate

adenylyltransferase(glgC) mutated strain under continuous growing conditions. Second, to produce high yields of PHAs in heterotrophic *Pseudomonas putida* from lignin derivatives, metabolic engineering strategies with substrate optimization were designed. The lignin with high biological reactivity was prepared with an innovative fractionation strategy. Comparative proteomics between several lignin substrate was conducted for a better understanding of lignin degradation mechanism and PHA biosynthesis in *Pseudomonas putida*. Further genetic engineering strategies based on the pathway findings from proteomics promoted the cell growth, lignin consumption and PHA accumulations. In conclusion, with deeper understanding of metabolic pathways regulations from multidenominational omics study, we successfully designed new metabolic engineering strategies to get most effective strains and hence improve the production of the limonene and PHAs in microorganisms.

DEDICATION

This dissertation is dedicated to my parents, Yuming Li and Wei Zou who taught me that life presents many obstacles, and the strongest are not those who never fall, but those who do, and learn how to get back up.

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my committee chair, Dr.Yuan for his invaluable guidance and suggestions, his constant encouragement and understanding in my pursuit of Ph.D. degree. I would like to thank my committee members, Dr. Dai, Dr. Pierson and Dr. Rentzepis for their guidance and support throughout the course of this research.

Thanks also go to all my lab members, colleagues, especially Dr. Zhi-Hua Liu and Dr. Xin Wang for their unconditional help in both my researches and life during my Ph. D. study. I also would like to thank the faculties, staff and students of the Department of Plant Pathology and Microbiology for making my time at Texas A&M University.

Finally, thanks to all of my family members, my mother and father for their encouragement, endless support, and love during my Ph.D. studies in the United States.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Joshua S. Yuan, Susie Y. Dai, and Elizabeth (Betsy) Pierson of the Department of Plant Pathology and Microbiology and Professor Peter M. Rentzepis of the Department of Electrical & Computer Engineering. All other work conducted for the thesis (or) dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a fellowship from China Scholarship Council and a research assistantship sponsored from National Science Foundation grant number 505670 and Department of Energy grant number 230268.

NOMENCLATURE

GC-MS	Gas Chromatography-Mass Spectrometry
LC-LTQ	Liquid Chromatography-Linear Quadrupole Ion Trap
OD	Optical Density
MudPIT	Multidimensional Protein Identification Technology
ATP	Adenosine triphosphate
NADPH	Nicotinamide adenine dinucleotide phosphate

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

1.1 Biofuel development overview

Increasing attention to energy and environmental issues gradually requires the development of renewable feedstocks for chemicals and fuels. The development of sustainable energy shows the promise to reduce greenhouse gas emission and our dependence on fossil fuels, eventually solve the challenges of global warming (Fatih Demirbas, 2009; Rosgaard et al., 2012). Biofuels derived from biological materials including plants and microorganisms provide a superior solution to address the current challenges of human society. First, second and third generation biofuels are named based on source and processing technology (Demirbas, 2010; Sims et al., 2010) (Figure 1.1). First-generation biofuels are made from sugar, starch or vegetable oils and they are produced by fermentation, distillation and transesterification. However, the source for first-generation biofuels are also a food source such as sugarcane or corn. Compared to first-generation biofuels, second-generation biofuels are sourced from the lignocellulosic biomass and the residual non-food parts of crops, which do not compete with food production (Liew et al., 2014).Third-generation biofuel are produced by algal and cyanobacteria biomass(Lee and Lavoie, 2013). The biomass from algae and cyanobacteria have lower hemicellulose and zero lignin content compared to the source of second-generation biofuels(Saqib et al., 2013) and therefore biofuel fermentation processes are much easier. Despite these efforts, the production of biofuels has not been

industrially implemented due to the high cost of process. In regard to sustainable bioeconomy, the development of efficient biofuel production technologies is of pivotal importance.

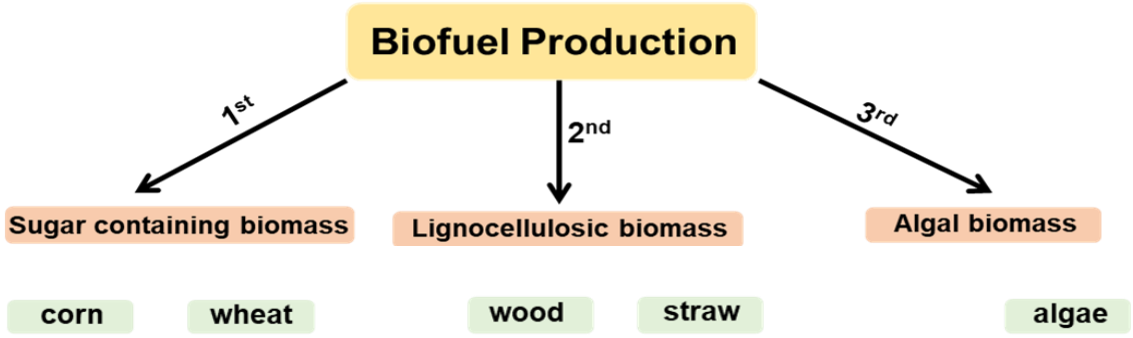

Figure 1.1 Schematic representation of first, second and third generation biofuel.

1.2 Second-generation biofuel developments

1.2.1 Pretreatments to overcome lignocellulose biomass recalcitrance

Lignocellulosic biomass has been considered one of the most important sources of renewable energy(Yuan et al., 2008; Hu and Ragauskas, 2012). Perennial feedstock like switchgrass could provide environmental and economic advantages over the current corn-based ethanol, considering the higher net energy gain and potential marginal land usage (Yuan et al., 2011) However, lignocellulosic biomass conversion is more challenging than that for corn, considering the recalcitrant nature of lignocellulosic biomass and the need to disrupt secondary cell wall structure to release carbohydrate and lignin(Isikgor and Becer, 2015; Balch et al., 2017).

To enable the cellulose for efficient hydrolysis to release glucose, many pretreatment methods including steam explosion, diluted acid, hot water, and organic solvents have been developed to

overcome the recalcitrance(Lau et al., 2009; Zhao et al., 2009; Lima et al., 2013; Zhang et al., 2016; Zhuang et al., 2016). However, most of these pretreatment processes were developed by solely considering the hydrolysis of carbohydrates, while lignin was considered a waste stream. Such a strategy poses a significant challenge in the sustainability and economic feasibility of lignocellulosic biorefinery. For lignocellulosic biorefinery to be viable, all cell wall components need to be processed into valuable products in a similar way to a corn ethanol refinery or a petroleum refinery. For example, corn-ethanol biorefineries produce multiple product streams, including ethanol, dried distillers' grains (DDG), and corn oil. Similarly, petroleum refineries convert the entire crude oil feedstock into multiple maximum-value products. Thus, a fundamentally different biomass processing strategy is required to co-optimize the processability for both carbohydrates and lignin to improve net energy gain, reduce carbon emissions, enhance overall sustainability, match products to demand, and increase profitability.(Liu et al., 2018a; Liu et al., 2019a) However, even though lignin is a major component of cell walls and has higher energy content than cellulose, it has a recalcitrant complex structure with plentiful aromatic moieties, posing significant challenges for its usage as a valuable bioenergy and biomaterial resource.(Matsakas et al., 2018) As a consequence, an efficient biomass process should not only improve the fermentable sugar release, but also maximize the lignin processability for utilization(Xie et al., 2017; Xu et al., 2019).

1.2.2 Metabolic engineering in *Pseudomonas putida* KT2440 to enhance lignin valorization and polyhydroxyalkanoates (PHAs)production from pretreated biomass

Lignin, as an aromatic heteropolymer, is made up from three hydroxycinnamyl monomers: guaiacyl alcohol (G unit), p-coumaryl alcohol (H unit), and syringyl alcohol (S unit). Lignin constitutes approximately 30% of the organic carbon on the earth(Ragauskas et al., 2014b). However, the heterogeneity and complex structure in lignin impede the conversion of lignin into high value products (Chen and Wan, 2017a).

Polyhydroxyalkanoates (PHA) are polyesters produced in nature by various microorganisms. Due to its biodegradability and the potential for making bioplastics, much attention has been directed to producing high titer PHA in microorganisms as a bio renewable resource(Sudesh et al., 2000). Bacteria from the genus *Pseudomonas* have the machinery to synthesize PHA as carbon and energy storage from different carbon sources(Poblete-Castro et al., 2012a). For decades, metabolic engineering research into *Pseudomonas putida* have been made to investigate PHA production using different carbon sources. Under nitrogen limited conditions, 81% of the cell's dry weight consist of PHA in *Pseudomonas putida* using fatty acids as the carbon source(Poblete-Castro et al., 2012b). Additionally, up-regulation of the β-oxidation fatty acid pathway improved PHA titer about one fold in *P. putida*(Lin et al., 2016a) on lignin carbon source. Mutating the glycerol-3-phosphate regulon repressor increased the PHA titer nearly two fold on glycerol carbon sources(Escapa et al., 2013). Since the identification of the first PHA synthase gene, significant metabolic engineering achievements have been made in bioproduction of PHA, such as increasing

the precursor carbon sources(Sung-Jin et al., 2007) and pathway optimization(Hoffmann et al., 2000; O'Leary et al., 2001). Recent developments in systems biological analysis of metabolic pathways and regulatory networks makes it possible for a comprehensive understanding of the PHA accumulating mechanisms especially from lignin derivatives.

1.3 Cyanobacteria are used as efficient cellular factories for third-generation biofuel production

Currently, bioethanol generated from corn or sugarcane are the most widely used biofuels(Fortman et al., 2008). However, the high production cost from the bioethanol fermentation by fungi weakens the biofuels competitiveness against the traditional oil refining(McAloon et al., 2000). Moreover, the rigidity of plants and the irregular structure of cellulose, hemicellulose, and lignin increases the cost of biofuel production(Madhavan et al., 2012). Cyanobacteria and algae have attracted the attention of researchers as an alternative cell factory and feedstock for biofuel production of third-generation biofuels while having low hemicellulose content and less land requirements than plants(Behera et al., 2015). Taking advantage of less nutrient requirements, less land usage for their growth, and higher lipid and carbohydrate content, cyanobacteria and microalgae are competitive to crop derived biofuels both environmentally and economically(Parmar et al., 2011).

The rapid development of genetic manipulation and modification tools have enabled cyanobacteria to be an efficient microbial cell factory. Cyanobacteria generates ATP, NADPH, and condenses carbon dioxide through oxygenic photosynthesis and subsequent Calvin-Benson-Basshan (CBB) Cycle(Badger and Price, 2003). Compared to heterotrophic fermentative platforms such as ethanol fermentation by yeast or butanol production by *E.coli*(Koppolu and Vasigala, 2016), cyanobacteria directly absorbs solar energy instead of massive usage of high cost sugars. Impressive efforts have been made to make cyanobacteria a strong cell factory compared to heterotrophic systems such as engineering cyanobacteria to produce butanol(Lan and Liao, 2011), fatty acids(Liu et al., 2011), or isoprene(Lindberg et al., 2010).

1.4 Metabolic engineering in cyanobacteria for biochemical production

In cyanobacteria, oxygenic photosynthesis converts sunlight and carbon dioxide into chemical energy in the form of ATP, NADPH, and biomass. Metabolic engineering with synthetic and system biological tools improves the photosynthetic chemical yields from the aspects of light energy harvesting, CO_2 fixation, and carbon repartitions(Peers, 2014). The model cyanobacteria, *Synechococcus elongatus* PCC7942 has been considered as an efficient cell factory to for the production of chemicals including isobutanol, isoprene, fatty acids, and limonene (Figure 1.2).

Cyanobacteria photosystem I&II (P700 and P680) capture sunlight and pass the energy to chlorophyll A, exciting electrons, and then transferring electrons to electron transfer chains

(Vermaas, 2001). The CBB cycle fixes atmospheric CO_2 into 3-phosphoglyceric acid(3PG) by ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). The ATP and NADPH produced during the oxygenic photosynthesis are used in the CBB cycle for the synthesis of G3P and pyruvate which serve as key building blocks for all cellular constituents and growth(Hagemann and Bauwe, 2016).

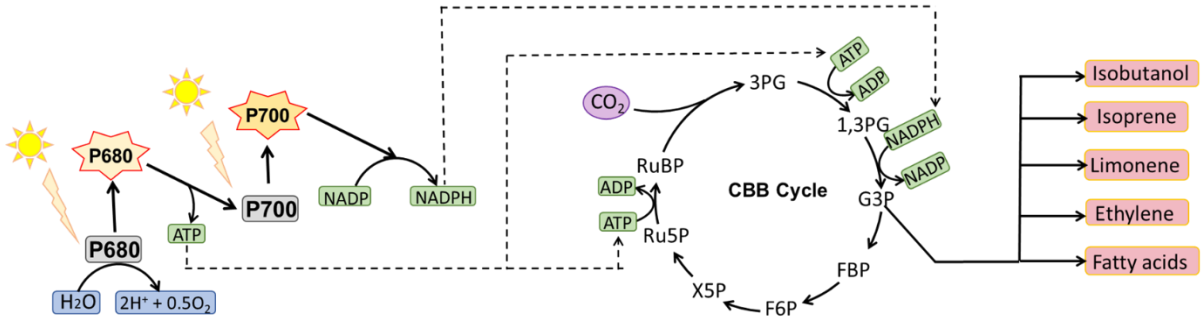

Figure 1.2 Schematic metabolism representation of oxygenic photosynthesis and CBB cycle. CBB cycle, Calvin-benson-basshan Cycle; P680, photosystem II; P700, photosystem I; 3PG, 3-phosphoglyceric acid; 1,3PG, 1,3-bisphosphoglyceric acid; G3P, glyceraldehyde 3-phosphate; FBP,fructose-1,6-bisphosphate; F6P, fructose-6-phosphate; X5P, xylulose-5-phosphate; Ru5p, ribulose-5-phosphate; RuBP, ribulose-1,5-bisphosphate.

Through genetic engineering, the carbon fixed in the CBB cycle could be redirected to synthesize valuable biofuels and commodity chemicals from atmospheric CO_2 and sunlight. Moreover, the cyanobacterial biomass is also fermentable by microorganisms for ethanol production.

1.4.1 Isobutanol production in cyanobacteria

Butanol, a four-carbon alcohol, is considered a potential biofuel and has a higher energy density compared to ethanol and almost as much as gasoline(Atsumi and Liao, 2008). Currently, butanol is mainly produced from the acetone-butanol-ethanol pathway of *Clostridium acetobutylicum* or engineered keto-acid pathways in *E. coli*(Machado and Atsumi, 2012) by using sugars as the carbon source.

Atsumi et al. has successfully engineered *Synechococcus elongates* PCC7942 to produce isobutanol and isobutyraldehyde directly from atmospheric CO_2 and sunlight(Atsumi et al., 2009). A 2-ketoacid-based pathway was engineered in PCC7942 to convert 2-ketoacid into alcohols in which 2-ketoacids. It was achieved with the overexpression of 2-ketoacid decarboxylase and alcohol dehydrogenase. Moreover, the strains with the engineered isobutyraldehyde synthesis pathway reached a yield of 723mg/L for 12 days. Moreover, they also engineered an enhanced Rubisco, the key enzyme in CO_2 fixation into strain and increased the yield to 1.1g/L in 8 days. These results showed that the CO_2 fixation reaction could be a limiting step in the CBB cycle and could be further used in other metabolic engineering works. With the downstream engineering, isobutyraldehyde was converted to isobutanol in cyanobacteria in a yield of 450mg/L isobutanol in 6 days.

Glycogen, as a major form of the fixed carbon in cyanobacteria (Shinde et al., 2020), could also limit the target chemical productions in cyanobacteria. Removal of the glycogen storage sink could

redirect carbon to other metabolites. Li et al. successfully built a glycogen deficiency mutant with the mutation of *glgC* which encodes the first enzyme in glycogen biosynthesis: glucose-1-phosphate adenylyltransferase in *S. elongatus* PCC7942 and reached a yield of 500 mg/L and 52% of the total fixed carbon was directed into isobutanol production compared to 23% in the same overexpression gene strain without *ΔglgC* (Li et al., 2014)(Table 1.1).

Synechocystis sp. PCC6803 is another model species for cyanobacterial metabolic engineering. Compared to PCC7942, PCC6803 can grow in both photoautotrophic and mixotrophic systems. Varman el al. also engineered two key genes in isobutanol biosynthesis pathways: *kivd* and *adhA*, encoding alpha-ketoisovalerate decarboxylase and alcohol dehydrogenase respectively. The engineered strain can produce isobutanol at a yield of 114 mg/L under mixotrophic growth and 27mg/L solely on glucose (Table 1.1). Moreover, continuous removal of isobutanol to decrease the cell toxicity brought the yield to 298 mg/L under mixotrophic conditions and 240 mg/L under autotrophic systems (Table 1.1).

Overall, through the expression of a heterologous isobutanol pathway, cyanobacteria were able to produce isobutanol at a satisfactory level. With further engineering works in the pathway such as increasing the carbon fixation rate, culturing conditions such as additional CO_2 added in the medium, and timely removal of the products could further increase the isobutanol productions in cyanobacteria.

Species	Productivity	Metabolic engineered genes	Carbon Source	Reference
S. elongates PCC7942	450 mg/L	*alsS, ilvC, ilcD, kivd, rbcL, rbcS*	CO_2, 0.5M $NaHCO_3$	(Atsumi et al., 2009)
Synechocystis sp. PCC6803	298 mg/L	*kivd, adhA*	50 mM $NaHCO_3$ and 0.5% Glucose	(Varman et al., 2013)
S. elongates PCC7942	500 mg/L	*ΔglgC, alsS, ilvC, ilcD, kivd, vqhD,*	CO_2, 50mM $NaHCO_3$	(Li et al., 2014)

Table 1.1 Metabolic engineering works for isobutanol production in cyanobacteria. AlsS, acetolactate synthase; IlvC, ketol-acid reductoisomerase; IlvD, dihydroxy-acid dehydratase; Kivd; alpha-ketoisovalerate decarboxylase; YqhD, alcohol dehydrogenase from *E. coli*; AdhA, alcohol dehydrogenase from *Lactococcus lactis*; RbcL, ribulose-bisphosphate carboxylase large chain; RbcS, ribulose-bisphosphate carboxylase small chain; GlgC, glucose-1-phosphate adenylyltransferase.

1.4.2 Fatty acid and hydrocarbon production in cyanobacteria

Biodiesel is a form of diesel fuel derived from the chemical reaction between chart chain alcohols and lipids into the corresponding fatty acid esters(Ma and Hanna, 1999). Microbial production of free fatty acids has also attracted attention as an alternative tool for renewable energy production. A large variety of enzymes for fatty acid and hydrocarbon production in microorganisms have been discovered and well-studied such as the acy-ACP reductase and aldehyde-deformylating oxygenase in hydrocarbon producing cyanobacteria *S. elongatus* PCC7942(Schirmer et al., 2010). Schirmer et al. engineered *E. coli* with a pair of the two enzymes which led to the production of

alkanes and proved the strong ability of microorganisms to produce fatty acids and hydrocarbons with suitable metabolic engineering strategies. The advantages of cyanobacteria fatty acid lies in the direct precursor source fatty acid synthesis-acetyl-CoA, which comes from the Cavin-benson-bassham cycle in cyanobacteria compared to sugars in *E. coli*(Lem and Stumpf, 1984).

Ruffing successfully engineered the *Synechococcus sp.* PCC7002 for free fatty acid(FFA) production with a yield of 130mg/L(Schirmer et al., 2010). Gene knockout of long-chain-fatty-acid CoA ligase (encoded by *aas*), and overexpression of thioesterase for fatty acid (FA) cleavage from (ACP) (encoded by *fat*) enables the production of FFA in *Synechocystis sp.* PCC6803. Ruffing's employment of the same strategies in PCC7942 only resulted in less than 6mg/L FFA production. When it was transferred to strain PCC7002, an FAA tolerant strain, FAA yield increased to 41mg/L. Overexpression of non-native RubisCO subunits (encoded by *rbcL*) led to a threefold increase of FFA in PCC7002 and reached a yield of 130mg/L.

Overexpressing codon optimized thioesterase genes and weakening of cell wall boundaries in *Synechocystis sp.* PCC6803 resulted in a yield of 197mg/L secreted fatty acid(Liu et al., 2011). Cyanobacterial cell walls are composed of a peptidoglycan layer and an outer membrane(Hoiczyk and Hansel, 2000). Liu et al first overexpressed gene *tesA* encoding acy-cayl carrier protein (ACP) thioesterase, gene *accA,B,C,D* encoding the fatty acid synthesis controlling enzymes deleted fatty acid activation genes *aas* for FFA production in PCC6803. When the possible surface proteins and peptidoglycan assembly proteins (PBP2) were deleted in order to increase the FFA secretion from the cells, FFA production increased from 83.6mg/L to the final yield of 197mg/L.

These results demonstrate that cyanobacteria could be an efficient cell factory for fatty acid production and. Increasing the precursors by enhancing the related enzyme activity increased the targeted bioproducts in cyanobacteria. Meanwhile, toxicity from the high concentration of targeted bioproducts also could be a limiting factor in increasing the productivity. Lifting the toxic stress by changing the cell membrane property would further increase FFA yield in cyanobacteria.

1.4.3 Terpene and terpenoids production in cyanobacteria

Terpenes and terpenoids represent a large variety of bioactive molecules and some are excellent targets for biofuels and pharmaceutical chemicals. In cyanobacteria and many other microorganisms, terpenes are synthesized from the methyl-erythritol-4-phosphate (MEP) pathway. Isoprene(C_5H_8), a five-carbon diene, is a key component in rubber and latex industry. Currently, the demand of isoprene is fitted by the petrochemical source(Whited et al., 2010).

Producing isoprene from microorganisms has been extensively studied. Cyanobacteria synthesize isoprene from two precursors, G3P and pyruvate, which come from the CBB cycle while the mevalonate (MVA) pathway uses acetyl-CoA(Miziorko, 2011). Bentley et al. first heterologously expressed MVA pathways in *Synechocystis sp*. PCC6803 for isoprene which is commonly used for isoprene production in *E. coli* and achieved isoprene production in PCC6803 at a yield of 0.25mg/L(Bentley et al., 2014)(Table 1.2). Gao et al. achieved photosynthetic isoprene production in *Synechocystis sp*. PCC6803 by engineering the MEP pathway int cyanobacteria (Gao et al.,

2016). Compared to the previous mevalonate pathway in E. coli for isoprene, the pyruvate pool in *Synechocystis sp*. had a much higher intracellular concentration than acetyl-CoA in E. coli. It demonstrated the *Synechocystis sp*.'s great potential for isoprene production. By overexpressing the i*spS* (encodes isoprene synthase) and *IDI* (encodes isopentenyl-diphosphate delta isomerase) in *Synechocystis sp*. PCC6803, a yield of 3mg/L isoprene was produced. IDI is the isomerase catalyzing the isomerization of two precursors for isoprene synthesis. With further metabolic engineering and cultivation optimization, a yield of 1.26g/L isoprene was reached (Table 1.2).

Plant terpenoid limonene($C_{10}H_{16}$) is the precursor for a series of industrial chemicals. It is primarily produced in animals, plants, and insects and are important components in hormone and pigment biosynthesis, cell membranes, and essential oils(Degenhardt et al., 2009). Halfmann et al. engineered a nitrogen-fixing cyanobacterium *Anabanea sp.* PCC7120 for limonene production. Overexpressions of *lims* (encodes limonene synthase) and several MEP pathway genes *dxs-ipphp-gpps* achieved a production rate of $3.6\mu gL^{-1}OD^{-1}h^{-1}$. With further enhancement of photosystem II by increasing the light intensity, a yield of 0.52mg/L was reached(Halfmann et al., 2014)(Table 1.2). Wang et al. used computational modeling and found that the limonene synthase activity was the speeding limiting factor in the limonene production pathway. With a different combination of promoters, ribosome binding sites, and best fit limonene gene sequence, a yield of $0.89mgL^{-1}OD^{-1}day^{-1}$ was observed(Wang et al., 2016)(Table 1.2).

Target Chemicals	Species	Productivity	Metabolic engineered genes	Reference
Isoprene	*Synechocystis sp.* PCC6803	0.35mg/L	*ispS*	(Bentley and Melis, 2012)
Isoprene	*Synechocystis sp.* PCC6803	0.25 mg/L	*ispS, hmgS, hmgR, fni, mk, pmd, pmk*	(Bentley et al., 2014)
Isoprene	*Synechocystis sp.* PCC6803	1.26 g/L	*idi-ispS fusion protein, ispG, dxs*	(Gao et al., 2016)
Limonene	*Anabanea sp.* PCC7120	0.52mg/L	*limS, dxs-ipphp-gpps*	(Halfmann et al., 2014)
Limonene	*Synechocystis sp.* PCC6803	1mg/L	*limS, dxs, crtE, ipi*	(Kiyota et al., 2014)
Limonene	*S. elongatus* PCC7942	0.89mg/L	*limS*	(Wang et al., 2016)

Table 1.2 Metabolic engineering works for terpene production in cyanobacteira. ispS, isoprene synthase; *hmgS*, hmg-CoA synthase; *hmgR*,hmg-CoA reductase; *fni*, IPP isomerase; *mk*,mevalonic acid kinase; *pmd*, mevalonic acid 5-diphoshate decarboxylase; *pmk*, mevalonic acid 5-phosphate kinase; dxs, 1-deoxy-D-xylulose 5-phosphate synthase, ispG, 4-hydroxy-3-methylbut-2-enyl-diphosphate synthase; *crtE*, GPP synthase gene, *ipi*, isopentenyl pyrophosphate.

1.5 Developments of photosynthetic terpene production in cyanobacteria as a source of third generation biofuel

Due to rapid growth, the ability to fix CO_2**,** and no need for fermentative sugars or cultivation lands, cyanobacteria has great potential for biofuel production. Improving photosynthesis efficiency and optimization of genetic parts by tuning MEP pathways are the most pertinent strategies to increase the flux from "source" to "sink" in advancing cyanobacterial terpene production.

1.5.1 Improve photosynthesis efficiency to enhance the 'source' pool for terpene production

Cyanobacteria captures solar energy and then converts CO_2 into organic carbons through light reactions and the CBB cycle. To increase the terpene yield, attention was given to increasing the "source" pool by increasing the efficiency of photosynthesis(Evans, 2013). Ribulose-1,5-biophosphate carboxylase/oxygenase (RuBisCO), the enzyme that catalyzes CO_2 fixation and oxygenation reactions, has attracted significant attention to increase system efficiency. C_4 plants have a higher efficiency in converting sunlight into biomass than C_3 plants due to the higher CO_2 concentration around the RuBisCO, reducing its oxygenase activity(Sage, 2002; Ainsworth and Long, 2005). Optimizing RuBisCO has been a long-term target in the agriculture industry for increasing photosynthesis in plants. Overexpression of *RuBisCO* improved photosynthesis and growth in *Synechocystis sp.* PCC6803(Liang and Lindblad, 2017) and borrowing the Rubisco from

Chlamydomonas reinhardtii increased photosynthesis and free fatty acid production in cyanobacteria(Liang and Lindblad, 2017).

Moreover, efforts have been made to increase photosynthesis efficiency through introducing carboxysome complex(Lin et al., 2014) and bicarbonate transporters(Price et al., 2012) in higher plants. Overall, increasing the 'source' pool from photosynthesis may lead to more carbon flux in the terpene 'sink'.

1.5.2 Pathway tuning to increase the 'sink' pool for terpene production

In cyanobacteria, the MEP pathway begins with the condensation of pyruvate and glyceraldehyde 3-phosphate (G3P) by 1-deoxy-D-xylose 5-phosphate synthase (encoded by *dxs*) and produces precursors for terpene synthesis after seven enzymatic reactions(Kirby and Keasling, 2009). Efforts have been made to tune MEP pathways to increase the carbon flux to increase the precursors DMAPP and IPP pools for terpene production in *E. coli*. It is now widely accepted that Dxs and other MEP enzymes are rate limiting factors in terpene production. Overexpression of gene *dxs* encoding the first enzyme in MEP pathway improved the lycopene production 2 to 3-fold (Matthews and Wurtzel, 2000). Through the chromosomal promoter replacement of the key MEP genes *dxs, ispD, ispF*, *idi*, and *ispB*, the engineered E. coli achieved 6mg/g dry cell weight carotene production(Yuan et al., 2006). Moreover, the transcriptional controls within the MEP pathway also regulated terpene production. The essential oil production in tea tree leaves was found to be

regulated through the transcript abundance of genes in the MEP pathway. It was found that the variation in MEP gene expressions accounts for 87% of the variation in terpene concentrations in plant leaves(Webb et al., 2013).

When overcoming the pathway bottleneck, it is important to integrate the different modules in a fine-tuned manner to reduce cell toxicity and feedback inhibitions. Overall, there is a great development prospect in photosynthetic cyanobacteria for terpene production. Increase the 'source' to 'sink' pool will advance the development of creating a photosynthetic cellular factory for terpene synthesis.

1.6 System biology tools advance metabolic engineering for efficient cell factory

System biology studies the components of biological systems and their dynamic relationship and regulations by integrating information from different disciplines and levels (Kitano, 2002). Basic system biology workflow includes four steps. First, build a preliminary system model by studying all of components in biological system, and understanding their genetic interaction networks and metabolic pathways. Second, change the internal components or external growth conditions and then observe the corresponding changes in components or structures including gene expression, protein expression, and interacting metabolic pathways. Third, compare the data obtained through experiments with the data predicted by the model and revise the preliminary model. Forth, based on the prediction of the revised model, integrate experimental biology and computational biology

to get an ideal model so that the prediction can reflect the authentic of biological systems(Alon, 2019). Recent advances in system biology enable us to perform metabolic engineering at whole cell level, thus optimize the efficiency in microorganisms for biofuels and biochemicals production.

Nowadays, many tools are available for system metabolic engineering by understanding the cellular metabolic regulations and bottlenecks under different conditions. Omics analysis including genomic, metabolomics, transcriptomics and proteomics could provide system-wide information of cellular components and their regulations(Cho et al., 2015). For example, by analyzing the metabolomic flux changes in *S. cerevisiae*, low glycolysis in xylose metabolism was identified as the limiting step for ethanol production in *S. cerevisiae*(Wasylenko and Stephanopoulos, 2015). The combination of transcriptomics and proteomics also helps us to get further insights of the gene regulations in the cellular systems. With different omics analysis, a high yield succinic acid producing *Mannheimia* strain was developed with the optimization of the central metabolism(Lee et al., 2016). Successful examples for enhancing biochemicals production based on system biology studies are increasing available. And the development of more convenient, efficient and robust systems metabolic engineering tool is promoting the researches in developing efficient microbial cell factories.

1.7 Objective of this study

The increasing demand for renewable fuels accelerates the development of biofuel production. However, the limited understanding of the metabolic pathways in microorganisms impedes the high yield production. In this study, system biology tools were used to study the metabolic pathways and the regulations within these pathways to find the rate limiting factors in producing limonene and PHAs in *Synechococcus elongatus* and *Pseudomonas putida*, respectively.

First, photosynthetic fuel production is an attractive route to harness sunlight energy and reduce carbon dioxide or sugar into fuel molecules. However, the interactions between photosynthesis and other metabolic pathways for the synthesis of a wide variety of chemicals are still under investigation and limits the capacity for high yield production of the targeted compounds. This study focuses on implementation of a photosynthetic platform to produce hydrocarbons in cyanobacteria. By using system biology tools such as metabolomics and proteomics**,** understanding of carbon partitions will be advanced to identify bottlenecks that limit the terpene yield and carbon partition in cyanobacteria. Previous work has identified downstream terpene synthase as a key metabolic bottleneck and overcame the bottleneck to achieve record limonene yield(Wang et al., 2016). Based on this strain, comparative metabolomics, transcriptomics and proteomics is used to analyze key regulator factors controlling the carbon repartition of carbon flux between primary metabolism and secondary terpene metabolism to eventually achieve a high terpene yield in cyanobacteria.

Second, the biofuel and bioproduct production from lignin substrates have great potential to improve the sustainability of biorefineries due to its low cost, high abundance properties. Biological lignin valorization is an effective way to overcome lignin heterogeneity by converting lignin and aromatics to target products. However, the low yield of target products presents the barriers to the utilization of lignin. In this study, several pretreatment methods for lignocellulosic biomass will be system biology guided genetic engineering was developed to increasing the lignin conversion to polyhydroxyalkanoates (PHA).

CHAPTER II

CARBON REPARTITION ENHANCED CO_2 TO TERPENE CONVERSION IN CYANOBACTERIA

2.1 Summary

Photosynthetic terpene production represents one of the most carbon and energy-efficient routes for the conversion of CO2 into hydrocarbon. Previous metabolic engineering efforts, center on the heterologous expression of MEP related enzymes, have resulted in limited limonene production in cyanobacteria. To produce high titers of terpenes, carbon partition from the Calvin-Benson-Bassham cycle was rewired to MEP pathway to increase the carbon flux by mutating the sucrose biosynthesis pathway. Proteomics and further related heterologous overexpression of several enzymes revealed MEP pathway limitations for increasing the terpene production. Metabolomics study showed a huge accumulation of sucrose after several days of culturing accompanied by a decreased limonene productivity. By blocking the sucrose biosynthesis to rewire the carbon flux from sucrose storage to the MEP pathway, we reached a record yield of 1100μg/L/day/OD730 in *S. elongatus* PCC7942. These results indicate that rewiring the carbon from sugar storage to the Calvin-Benson-Bassham cycle, and then to the MEP pathways increases the carbon flux for limonene production. The strategies of balancing carbon between primary and secondary metabolites developed in this study provides the perspectives for terpene production in other

organisms and the strain developed in this study could also serve as host for other photosynthetic hydrocarbons.

2.2 Introduction

Terpenes are a large class of natural products produced by plants and bacteria with both important biological functions and broad commercial applications. In particular, terpenes play essential roles in energy harvesting (e.g. chlorophylls), membrane stability (e.g. steroids), and multi-trophic signaling (Yuan et al., 2009; Pasoreck et al., 2016). Many terpenes are also valuable chemicals with broad applications in pharmaceutical, nutraceutical, cosmetics and biotech industries (Wang et al., 2015). The past decade has witnessed a rapid increase in the levels of atmospheric CO_2 due to fossil fuel combustion and deforestation. Terpene production in phototrophs represents one of the most energy and carbon-efficient routes for converting CO_2 to fungible fuels and products(Song, 2006). Photosynthetic terpene production uniquely addresses this problem by directly converting CO_2 into hydrocarbon as 'drop-in' biofuels, which could mitigate the global climate changes by both reducing the fossil fuel utilization and enabling sustainable carbon capture and utilization (CCU). Moreover, the efficient conversion of CO_2 to high-value terpene products could improve the economics of algal biomanufacturing industries, enabling the market transition and achieving economically feasible CCU (Bohlmann and Keeling, 2008; Choi et al., 2016).

Considering the importance, significant efforts have been invested in improving the efficiency of photosynthetic terpene production. Cyanobacteria has recently emerged as a major model system for such efforts due to the rapid growth and readily available genetic toolbox(Ducat et al., 2011). In cyanobacteria, terpenes are synthesized from two C_5 precursors dimethylallyl pyrophosphate (DMAPP) and isopentenyl pyrophosphate (IPP) by various terpene synthases. IPP and DMAPP are derived in the methylerythritol phosphate (MEP) pathway through seven enzymatic steps by converting pyruvate and the photosynthate glyceraldehyde-3-phosphate (G3P), which directly connects the photosynthetic outputs with reduction and hydrocarbon production(Pattanaik and Lindberg, 2015). Recent genetic engineering efforts successfully validated kinetics modeling prediction and identified terpene synthases as the key metabolic bottleneck in enhancing photosynthetic carbon partition into terpene biosynthesis. In our previous study, *limS* gene encoding limonene synthase from *Mentha spicata* was integrated into the genome of *Synechococcus elongatus* PCC7942 under the driven by a pea *psb*A promoter, and achieved limonene productivity of 885.1 μg/L/OD/d by enhancing *limS* expression through synthetic promoter design(Wang et al., 2016). Despite the progress, the productivity is far from reaching a commercially relevant level.

The fundamental challenge for enhancing the terpene productivity in a photosynthetic system remains to be the low carbon partition. As a secondary metabolic pathway, carbon flux into the MEP pathway is low, and the estimated photosynthetic carbon partition for isoprene synthesis is less than 1% in cyanobacteria (Melis, 2013). Considering of the low carbon partition of photosynthate, it is critical to enhance carbon flux toward terpenes biosynthesis to achieve a high

yield. Previous research has established that enhancing downstream biosynthesis of a secondary or non-natural metabolites could synergize either higher photosynthesis rate or more carbon partition(Ducat et al., 2012; Ducat and Silver, 2012).In fact, we have observed increased expression of carbon fixation enzymes in the engineered strain with high limonene productivity, even though the oxygen evolution was actually not increased. An alternative strategy to increase target products is to re-partition the carbon. It is thus critical to understand additional metabolic and biochemical bottlenecks in the overall CO_2 to the terpene conversion process.

In this study, we employed systems biology approaches to uncover limitations impeding photosynthetic carbon partitioning to terpene biosynthesis in the cyanobacterium *S. elongatus* PCC 7942. Our study has revealed the active flux competition between primary metabolisms and terpene biosynthesis. Enhance terpene production was observed by tuning down the sucrose biosynthesis pathways. The systems biology guided engineering design in this study could be applied to future engineering efforts to seek higher productivity of other secondary metabolites.

2.3 Materials and methods

2.3.1 *S. elongatus* PCC7942 cultivation conditions

The limonene-producing *S. elongatus* strain L1118(Wang et al., 2016) and engineered strains were grown in BG11 medium (Sigma, Chicago, IL) supplied with 10 mM N-[Tris(hydroxymethyl)methyl]-2-aminoethanesulfonic acid (TES, pH=8.2). The strain was first

cultured in 20 mL BG11 medium, in a 250 mL flask under the light intensity of 75 μmol photons $m^{-2}s^{-1}$, and at 30 °C and then transferred into a 1-L rectangle bottle with continuous 5% CO_2 bubbling and under the light intensity of 100 μmol photons $m^{-2}s^{-1}$. The medium was supplied with appropriate antibiotics such as kanamycin (5 μg/L) and gentamycin (2 μg/L).

2.3.2 Plasmid and strain constructions

Strains and plasmids used in this study were summarized in Table 2.1. The related plasmids were constructed through Gibson Assembly (NEB, Ipswich, MA) following the manufacture's instruction. The related genes were amplified from the specific genomic DNA (Table 2.1) using Phusion High-Fidelity DNA Polymerase (NEB, Ipswich, MA) and ligated to the proper plasmids. Strains that express the specific genes were construction by insertion of an expression cassette into the specific neutral site in the genome of *S. elongatus* PCC7942, L1118. L1118 is a previous published limonene producing strain with the overexpression of *limS* under a pPsbA promoter at the neutral site I in the wild type PCC7942 chromosome.

Transformation of the *S. elongatus* was carried out by chromosomal recombination. Cyanobacterial transformants with targeting vector were selected on BG11 agar plated with appropriate antibiotics kanamycin (5 μg/L) and gentamycin (2 μg/L). The result of transformation was confirmed by segregation PCR.

Strains that overexpress *ispG*, *gap2*, and *pgk* were constructed by insertion of expression cassette into neutral site II in *S. elongatus* PCC7942 L1118. For construction of strains L1218, L1219, L1221, and L2211, the genes *ispG*, *gap2* and *pgk* were cloned into the neutral site II targeting vector, pAM1579 (purchased from Addgene) under an isopropyl β- d-1-thiogalactopyranoside (IPTG) pLacO promoter. The coding region of *ispG*, *gap2*, *pgk* were amplified from the genome DNA from the specified genome DNA listed in Table 2.2. The resulting plasmid were named pWX1218, pWX1219, pWX1221, pLM2211. The genome DNA was isolated followed the protocol from previous publications(Morin et al., 2010). The plasmids were then transformed into L1118 and the resulting strains were named L1218, L1219, L1221, L2211 respectively.

Stain with *sps* (encodes sucrose phosphate synthase) mutation were constructed by chromosomal recombination. For the generation of the mutagenesis construct pLM21, 1000 bp DNA fragment prior to the *sps* coding region, a gentamycin resistant gene cassette, 800 bp DNA fragment after the *sps* coding region were ligated to pUC19 (purchased from Addgene). The *sps* coding region were replaced with the gentamycin resistant gene cassette by homologous recombination. The plasmid pLM21 was then transformed into L1118 and the resulting strain was named Lsps with the *sps* mutation.

Strain or plasmid	Relevant characteristics	Source or reference
Strains		
S. elogatus PCC 7942	Wild type	S. Golden
L1118	*limS* integrated at NSI of *S. elogatus* PCC7942	X. Wang
Lsps	*sps* insertional mutation by gentamycin resistant gene in the genome of L1118	This study
L1218	*ispG* (from Synechocystis sp. PCC 6803) at NSII of L1118	This study
L1219	*ispG* (from *Botryococcus braunii*) at NSII of L1118	This study
L1221	*ribB* (Codon optimized for PCC7942) and *ispG* (from *S. elogatus* PCC7942) at NSII of L1118	This study
L2211	*gap2* and *pgk* (from Synechocystis sp. PCC 6803) at NSII of L1118	This Study
Plasmids		
pAM2991	Targeting PCC7942 NSI; Ptrc; Sp/Sm^R	SS. Golden
pAM1579	Targeting PCC7942 NSII; $P_L$lacO, Km^R	CR. Andersson
pLM21	Targeting PCC7942 *sps*, Gm^R	This study
pWX1218	*ispG* (from *Synechocystis sp*. PCC 6803 in pAM1579, Km^R	This study
pWX1219	*ispG*(from *Botryococcus braunii* in pAM1579, Km^R	This study
pWX1221	*ribB* (Codon optimized for PCC7942) and *ispG*(from *Botryococcus braunii* in pAM1579, Km^R	This study
pLM2211	*gap2* and *pgk* (from *Synechocystis sp*. PCC 6803) in pAM1579, Km^R	This study

Table 2.1 Strains and plasmids used in this study. *limS*: encodes limonene synthase; *sps*: encodes sucrose phosphate synthase; *ispG*: encodes 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase; *ribB*: encodes 3,4-dihydroxy-2-butanone 4-phosphate synthase; *gap2*: encodes glyceraldehyde-3-phosphate dehydrogenase; *pgk*: encodes phosphoglycerate kinase.

2.3.3 Limonene collections and quantification by GC-MS

S. elongatus strain L1118 and other engineered strains were grown in triplicates in 500 mL BG11 medium in a 1 L rectangle bottle coupled with a self-made HayeSep polymer trap. The vaporized

limonene was collected each day by eluting with 1 mL hexane containing 10 µg/mL cedrene as the internal standard. The sample was then analyzed by gas chromatography-mass spectrometry (GC-MS). GC-MS was performed on a GCMS-QP2010SE (Shimadzu Scientific Instruments, Inc. JP). In detail, 1 µL eluted sample was injected into a Shimadzu SH-Rxi-5Sil column (30 m × 250 µm × 0.25 µm). Helium was used as the carrier gas at the flow rate of 1.0 mL/min. The temperature was held at 50 °C for 3 min, and then increased to 140 °C at a rate of 20 °C/min. Mass spectral peak identification and quantification was performed with GCMSolution software Ver. 2.6 against a limonene standard curve with the internal standard.

2.3.4 Quantitative reverse transcription PCR (RT-qPCR)

S. elongatus strain L1118 cells were grown in triplicates in 500 mL fresh BG11 (Sigma, USA) in a 1-L Roux bottle with 5% CO_2 bubbling and 100 µmol photons$m^{-2}s^{-1}$for 7 days. 10mL of the cyanobacteria cells were harvested at day 2, day 7. The cultures were centrifuged at 5000 rpm for 10 min at 4 ºC followed by total RNA extraction using the TRIzol reagent. Two µg of RNA was first treated with DNase I (Sigma, USA) following the manufacturer's instruction. The cDNA synthesis was performed in a 20-µL reaction using the OneStep RT-PCR kit (Qiagen, USA) following the manufacturer's instruction.

All of the MEP pathway genes and the limonene synthesis gene were selected to compare their expression levels between two growth stages at day 2 and day 7, i.e. exponential and stationary growth stage (MEP pathway genes are shown in Figure 2.1, primers used for qRT-PCR are listed

in Table 2.1). The sigma factor RpoD gene was used as the internal control. The synthesized cDNA was diluted 500 times and used as the template for qPCR using the 2 × SYBR green master mix (Bio-Rad, USA). For each gene, 2 technical replicates and 3 biological replicates were included and the real-time PCR was carried out on a Bio-Rad Real-time PCR system (Bio-Rad, USA). For gene expression comparison, ΔCt was calculated as the cycle difference between target gene and the *rpoD* control. ΔΔCt was further calculated by subtracting each ΔCt with the average ΔCt of one of the biological replicates for each gene. A ΔΔCt value of close to 0 thus means minimal change between replicates. And the gene expression differences between two growth stages were compared using the Log_2 value of ΔΔCt.

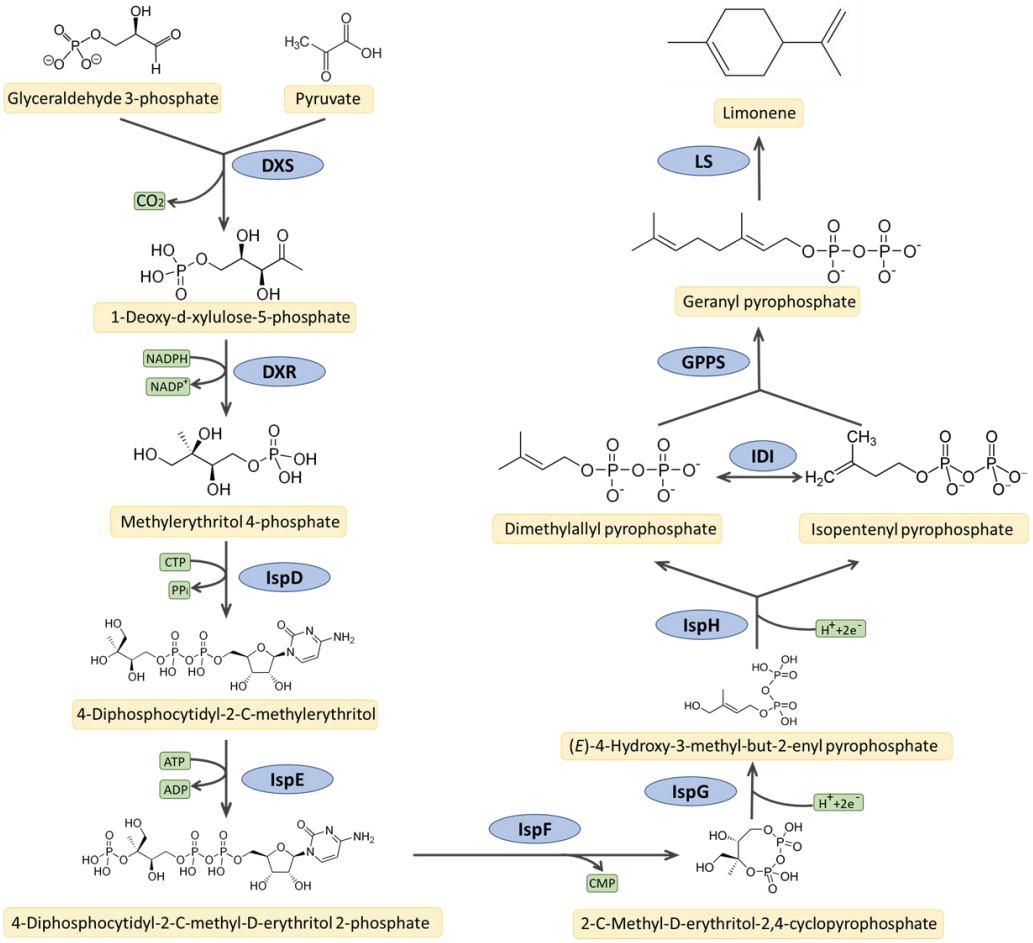

Figure 2.1 Pathway map for limonene biosynthesis in *S. elogatus* PCC 7942 L1118. DXS: 1-deoxy-D-xylulose-5-phosphate synthase; DXR: 1-deoxy-D-xylulose 5-phosphate reductoisomerase; IspD: 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase; IspE: 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase; IspF: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; IspG: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase; IspH: 4-hydroxy-3-methylbut-2-enyl diphosphate reductase; IDI: Isopentenyl-diphosphate delta-isomerase; GPPS: Geranyl pyrophosphate synthase; LS: Limonene synthas

Primer	5’-3’	Application
qPCR_ispD-F	ACACCGATTCCTGGCGTTCT	qPCR for *ispD*
qPCR_ispD-R	GGTTCTGGCCTGGACCCTAC	qPCR for *ispD*
qPCR_ispE-F	TCGAAGCGGTGGTCTTACCC	qPCR for *ispE*
qPCR_ispE-R	AGCTTCAGGTCGGGATCAGC	qPCR for *ispE*
qPCR_ispF-F	ACACTCCCTTGGGCTGGATG	qPCR for *ispF*
qPCR_ispF-R	GGATCCGTCGGCGGAAAGTA	qPCR for *ispF*
qPCR_ispG-F	AGCAAATTCGCGACACGCTT	qPCR for *ispG*
qPCR_ispG-R	CTCTTCGCAGAGTCGCAGGA	qPCR for *ispG*
qPCR_ispH-F	CTTGCGCCAGAATGGCTACC	qPCR for *ispH*
qPCR_ispH-R	GATGCTGGTTGACCGAGGGA	qPCR for *ispH*
qPCR_dxs-F	TTCTGCTGCCCTCGGTATGG	qPCR for *dxs*
qPCR_dxs-R	ACGTTGGGCGAGATCGACAT	qPCR for *dxs*
qPCR_ispC-F	CGAGGTGATCGAAGCCCACT	qPCR for *ispC*
qPCR_ispC-R	GAGAGGGCGTAGAGCAAGGG	qPCR for *ispC*
16S_qPCR-R	GTTTGTCACCGGCAGTTTCT	qPCR for *16S*
rpoD1_qPCR-F	GCGGGGGCTGAGAGTGCTAAG	qPCR for *rpoD1*
rpoD1_qPCR-R	CGTAAGCCCACCGAGGAAGAGA	qPCR for *rpoD1*
qPCR_idi-F	CTAGGGACGCAGTTCTTGGG	qPCR for *idi*
qPCR_idi-R	ATGGCGAGCCGATACTTCTG	qPCR for *idi*
qPCR_limS-F	GTTGGCGTTCCGACTCCTG	qPCR for *limS*
qPCR_limS-R	CGTATCGTCCGACAGGCTCT	qPCR for *limS*

Table 2.2 Primers used for qRT-PCR

2.3.5 Total intercellular protein extraction

50 mL cyanobacteria culture was harvested by centrifugation at 8000 rpm at day 2 and day 7, 4°C for 10 min. For each strain, 3 biological replicates were used for total intercellular protein extraction. The pellets were washed with 10 mL 0.9% NaCl. Pellets were resuspended in 2 mL 50 mM Tris·HCl buffer (pH=7.6). 20 µL protease inhibitor cocktail (Sigma, USA) was added into each sample. Cells were lysed by tip-probe sonication using 20 cycles of 15 s. The lysates were centrifuged at 12000 rpm at 4 °C for 10 min. Supernatants were harvested and stored at -80 °C for future proteomics use. The protein concentration was measured through Pierce Bradford Protein Assay (Thermo Fisher, USA) with BSA standards.

2.3.6 MudPIT-based proteomics analysis and follow up data analysis

100 µg protein from each sample were denatured under conditions of 8 M urea and 5mM DTT and incubated at 37 °C for 1 hour. The denatured proteins were then treated with iodoacetamide to a final concentration of 15 mM followed by an incubation period for 15 min at room temperature in dark. The samples were diluted 4 times to a final concentration of 2 M urea. 1 g trypsin (Promega, USA) was added into each sample and incubated at 37°C overnight. The digested peptides were desalted through a sep-pak C18 column (Waters, USA). A SpeedVac (Brinkmann Instruments, USA) was used to dry these samples. The dried digested peptides were dissolved in 0.1% FA and centrifuged for 10min at 12000rmp 4°C. The supernatants were harvested and stored at -80°C for future use. The desalted samples were loaded into a self-pack 150 µm capillary column with 3 cm SCX and 5 cm C18 beads. Then the samples were analyzed by LC-MS/MS (Orbitrap velos pro) (Thermo, USA). The mass spectra were extracted from raw files and then converted into an MS2

file and were searched against a composite database of *S. elongatus* 7942, common contaminants, and reversed sequences using ProLuCID algorithm(Zhang et al., 2012). And the sqt file was filtered through PatternLab4.0 for the following data analysis.

2.4 Results

2.4.1 Proteomic analysis indicates the limitation of substrate input in enhancing MEP pathway carbon flux

Previous study identified limonene synthase as a rate controlling step for limonene production in *S. elongatus*. By creating a strong limonene sink through high limonene synthase expression, the strain L1118, with insertion of *limS* (encodes limonene synthase) under the driven of psbA2 promoter in the neutral site I of wild type *S. elongatus* PCC7942, serves as a powerful platform to further investigate limitations in the upstream metabolic pathways (Wang et al., 2016).

In this study, the limonene production from L1118 were collected daily to evaluate the limonene productivity along the cell growth. When L1118 were grown under continuous light for 7 days, it was noticed that the limonene productivity changed across the growth phases. The highest limonene productivity was found in the exponential phase (day 2) whereas productivity decreased significantly at the stationary phase (day 7) (Figure 2.1 A).

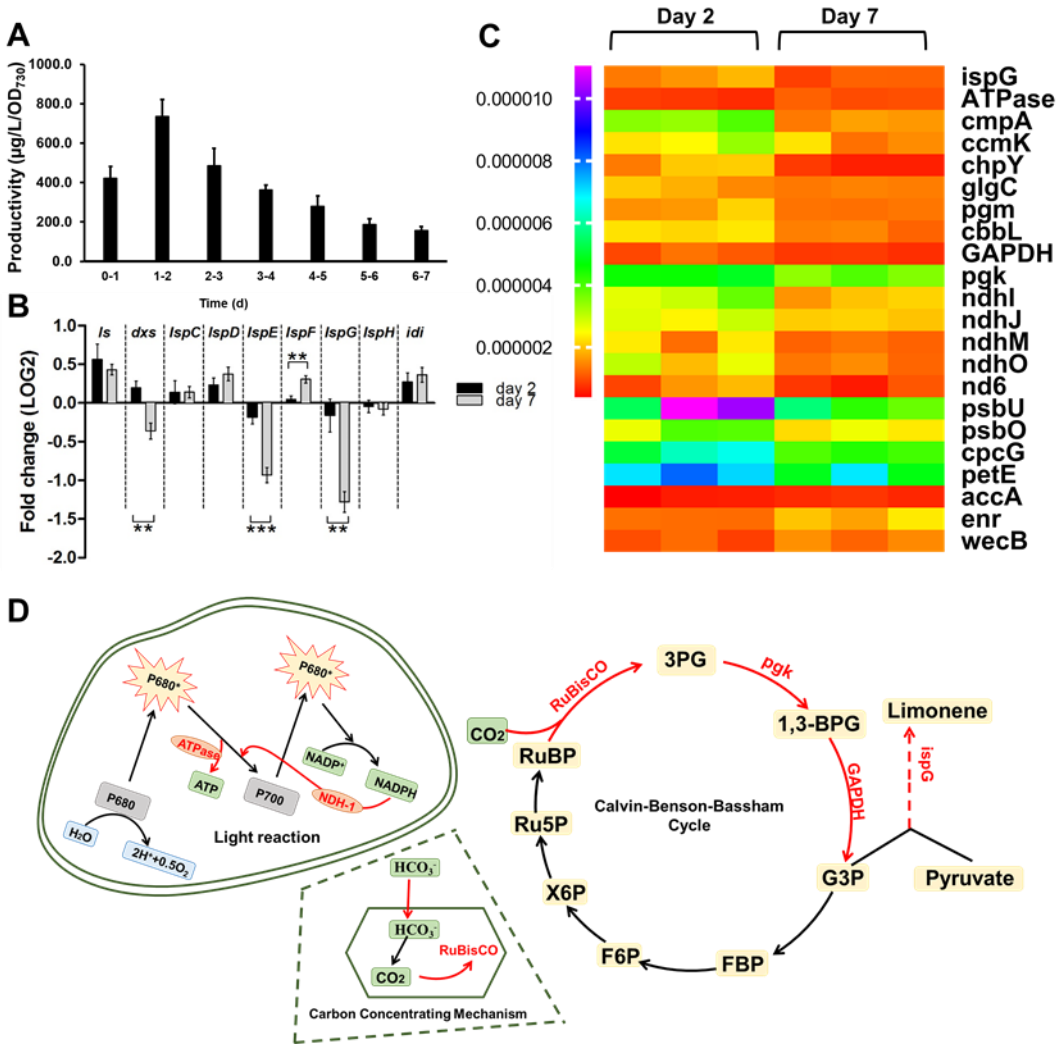

Figure 2.2 Comparative proteomics and transcriptomics in L1118. A: Limonene productivity decrease after 2 days cultivation, B: Real-time PCR showing MEP gene transcriptional change, C: Proteomics analysis of day 2, day 7 in L1118. Proteomics analysis was done with three biological replicates. Experiments at each stage were done in three biological replicates. All proteins showed in the figure have a p-value<0.1. D. Scheme of pathway changes in L1118 in the stationary phase. Red color refers to the related enzyme expression level decrease. NDH-1: NAD(P)H-quinone oxidoreductase subunit I,J,M,O, ribulose-1,5-bisphosphate carboxylase/oxygenase; PGK, phosphoglycerate kinase; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; IspG: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase.

The initial hypothesis was that low expression of MEP enzymes could lead to the decrease of terpene productivity. To identify which steps are potential bottlenecks, L1118 cultures were harvested during high and low productivity phases to examine the expressions of both *limS* and

MEP pathway genes (Figure 2.1B&C). Firstly, quantitative reverse transcription PCR (RT-qPCR) analysis was carried out to evaluate the transcriptional levels of these genes. Compared to early growth stage, the expression of *ispE* and *ispG* decreased significantly in the late growth stage. Interestingly, the *ispF* expression was significantly higher in the late growth stage when the limonene productivity was low (Figure 2.1B). The expression of other MEP genes and *limS* showed similar levels at both phases. 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, encoded by *ispE*, catalyzes the phosphoralation of the position 2 hydroxy group of 4-diphosphocytidyl-2-C-methyl-D-erythritol. 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, encoded by *ispF*, involves in the biosynthesis of isopentenyl diphosphate (IPP) and it's isomer dimethylallyl diphosphate (DMAPP). 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase, encodes by *ispG*, converts 2C-methyl-D-erythritol 2,4-cyclodiphosphate (ME-2,4cPP) into 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate (Figure 2.2).

Then proteomic analysis was carried out based on the two stages of L1118 cells, i.e. exponential and stationary phase (day2 and day7). Similar to the qPCR results, the LS enzyme levels were found to be similar between these two growth stages. In addition, IspG was the only MEP enzyme that showed an abundance difference between these two phases with statistical significance (p-value<0.5). IspG is an essential enzyme for terpene synthesis in the MEP pathway (Eisenreich et al., 2004). Both the RT-qPCR and proteomics results suggest that IspG might be a MEP pathway bottleneck step that hindered further increase of limonene flux in L1118 cells. We thus overexpressed the *ispG* gene under the control of the LacO-1 promoter in the neutral site II of L1118 genome. Both *S. elongatus* 6803 and *B. braunii ispG* genes were engineered and generated strains L1218 and L1219, respectively. Although overexpression of *ispG* did lead to a slightly

slower decrease of limonene productivity when cell enters late growth stage, it did not lead to an increase of limonene productivity in the early growth (Figure 2.3). These results indicated either sufficient IspG activities in actively growing cells or a limited role in enhancing MEP flux by IspG alone. Another possibility is that the expression level of IspG might not be optimal to efficiently channel the carbon flux to limonene. The accumulation of IPP/DMAPP due to *ispG* overexpression could have caused a feedback inhibition to the 1-deoxy-D-xylulose-5-phosphate synthase (DXS), the first enzyme in the MEP pathway (Banerjee et al., 2013). A previous study reported the isoprene production in *S. elongatus*, in which the co-expression of several MEP enzymes (encodes by *dxs-ispG-idi*) did increase the MEP carbon flux and lead to increased isoprene production (Gao et al., 2016). However, other MEP pathway enzymes do not seem to be limiting steps in the case of limonene production.

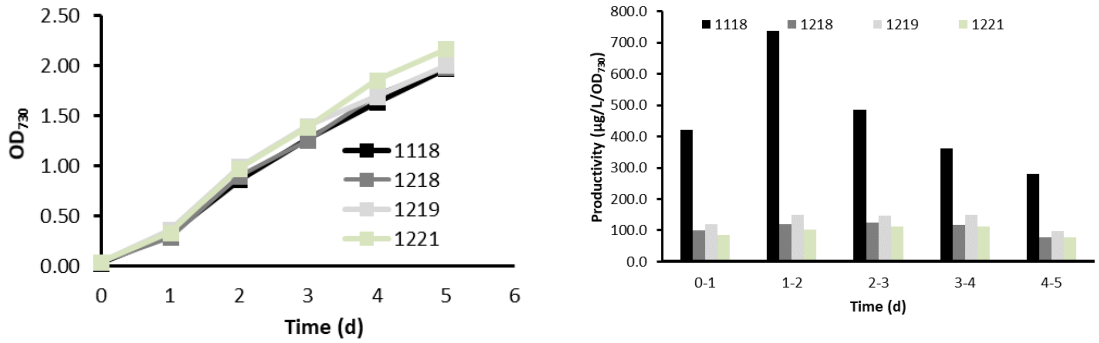

Figure 2.3 Limonene productivity of engineered *IspG* strain. L1218: Engineered with SelispG, *ispG* gene was cloned from *Synechocystis sp.* PCC6803; L1219: Engineered with BbispG, *ispG* gene was cloned from *Botryococcus braunii*, L1221: Engineered with *rib* and *IspG* gene, *ribB* gene was cloned from *E. coli* and *ispG* gene was cloned from *Botryococcus braunii*. The strains were complete segregated and confirmed with PCR. Each strain was grown in three biological replicates.

As *ispG* overexpression did not lead to a further increase of limonene productivity, it was hypothesized that substrates input (G3P and pyruvate) into the MEP pathway might play a bigger role in enhancing MEP carbon flux than enzyme kinetics. The role of substrate levels and enzyme kinetics on pathway efficiency was illustrated in a previous study, indicating that substrates concentrations could play bigger roles in the overall efficiency of target pathways(Hackett et al., 2016). In this study, the proteomics data indicates the upstream limitation in supplying G3P and/or pyruvate into the MEP pathway (Figure 2.1D). When comparing the early and late growth stage cells, several enzymes involved in CO_2 assimilation and the Calvin-Benson-Bassham cycle showed decreased levels at the late growth stage. The levels of CO_2 hydration protein ChpX and bicarbonate-binding protein CmpA showed 2.5-fold and 2.8-fold decrease (Table 2.1), respectively. These two enzymes play an essential role in the CO_2 concentrating process. CmpA belongs to the ATP binding cassette (ABC) transporter family and is specific for HCO_3^- affinity which plays an important role in CO_2 capture process in cyanobacteria(Badger and Price, 2003). Moreover, the expression of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) also showed significantly decreased levels at the late growth stage (Figure 2.1C). The decreased expression of GAPDH thus indicates a lower photosynthesis carbon (G3P) output in the late stage of growth in *S. elongatus*. Coincidently, several proteins involved in the light reactions also showed decreased level at the late growth stage. Compared to protein levels in the early stage, PSII extrinsic protein (psbU) and NAD(P)H-quinone oxidoreductase subunits I,J,M,O of the NDH-I complex showed 2.3-fold decrease and 2-fold decrease in the late growth stage (Table 2.1). PSII extrinsic proteins catalyze the photosynthetic oxidation of water and provides the electrons for photosynthesis (Bricker et al., 2012). NDH-I complex also plays an essential role in the cyclic electron flow of photosynthesis(Zhao et al., 2018). Together, both light reactions and the Cavin-Benson cycle

suggest a decreased photosynthesis efficiency in the late growth stage of growth, thus limiting the overall carbon output from photosynthesis.

Metabolism	Enzymes	Relative Abundance Day2 *vs.* Day 7	p-value
Light reaction	PS II 12kDa extrinsic protein(PsbU)	2.27	**
	NAD(P)H-Qunione oxidoreductase SubunitsI, J, M, O	~2	***
CO_2 assimilation & CBB cycle	CO_2 hydration protein	2.53	**
	Bicarbonate-binding protein	2.89	*
	Glyceraldehyde-3-phosphate dehydrogenase	1.73	**
Glycogen biosynthesis	1,4-alpha-glucan branching enzyme	2.61	**

Table 2.3 Enzyme expression level changes in early and late growth stage in L1118. *: p-value<0.1, **: p-value<0.05, ***: p-value<0.01. Cells at each stage for proteomics were harvested in three biological replicates.

To validate whether increasing substrates input could enhance MEP flux, we conducted a substrate feeding experiment in L1118 cells. Both pyruvate (200mg/L) and glycerol (252mg/L) (the closest substrate that lead to G3P)(Farmer and Liao, 2001) were supplemented in the growth medium and limonene production was measured. Surprisingly, neither pyruvate nor glycerol feeding led to the increase of limonene production in the L1118 cells (Figure 2.4). One possible explanation is that these substrates were not imported into the cell very efficiently. However, the hydrophobicity of glycerol determines that cell transport should not be a limiting factor even without a specific transporter protein. Another possibility is that these enhanced substrate levels could have caused an adverse effect on overall efficiency of photosynthesis due to feedback inhibition, thus decreasing the net carbon input into MEP pathway. These observations call for an *in vivo* strategy that can enhance carbon input into the MEP pathway.

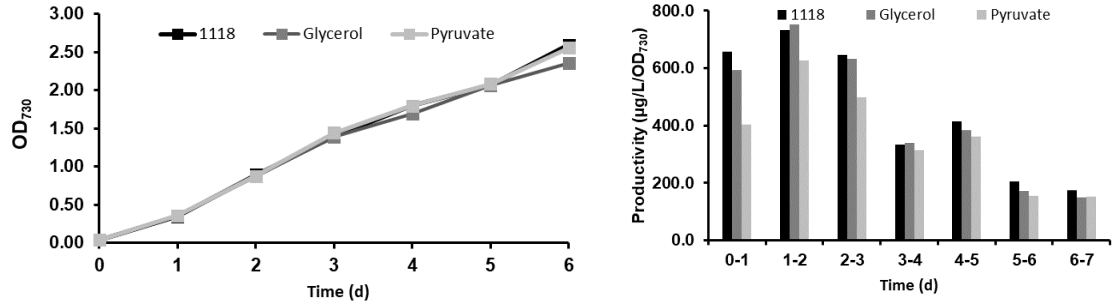

Figure 2.4 Substrate feeding impact on the productivity and growth of L1118. In glycerol group, 252mg/L glycerol was added in the BG11 medium and in pyruvate, 200mg/L pyruvate was added in the BG11 medium. Each feeding and the control experiments were done in three biological replicates.

The proteomic analysis showed decreased expression of several enzymes in the CBB cycle in the late growth stage cells, including Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and phosphoglycerate kinase (PGK) (Figure 2.1C). We thus constructed strain L2211 in which GAPDH and PGK were inserted into the neutral site II under the driven of a pLacO promoter of L1118 to evaluate their effects on limonene production. However, the overexpression of GAPDH and PGK did not further increase the limonene yield (Figure 2.5). These results indicate either the inefficiency of enhancing carbon output in the Calvin-Benson-Bassham cycle by simply overexpressing its enzymes, or the weak capacity of the MEP pathway serving as a carbon sink for photosynthesis. We thus set out to evaluate the native metabolite distribution among metabolic pathways to understand their capacity to serve as photosynthetic carbon sinks.

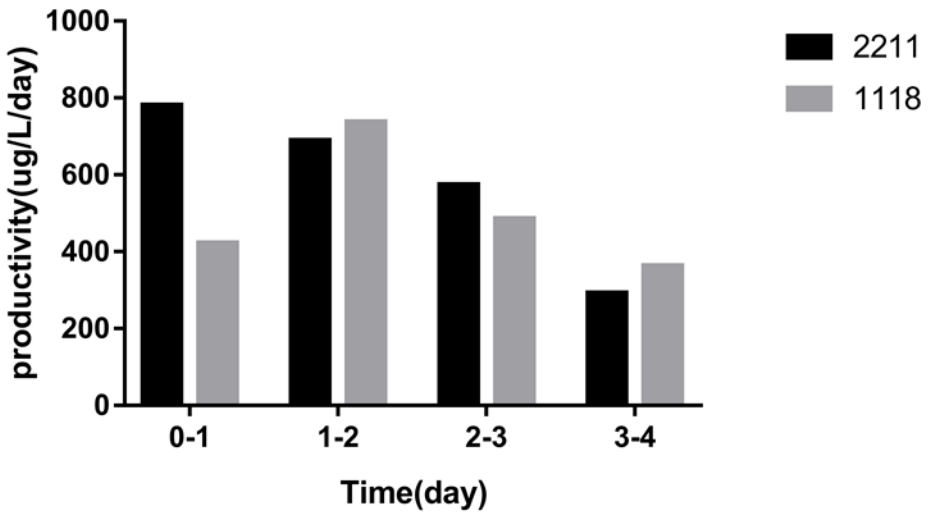

Figure 2.5 Productivity of strain L2211 engineered *GAPDH* and *pgk*. L2211: *GAPDH* and *pgk* from *Synechocystis sp.* PCC6803 were insert into NSII with Km^R in L1118. The strains were complete segregated and confirmed with PCR. Each strain was grown in three biological replicates.

2.4.2 Metabolomics revealed the active competition on substrates between the MEP pathway and carbon storage pathways

To investigate metabolites allocations in *S. elongatus* L1118, we conducted a metabolomics study by sampling three time points during the growth (Figure 2.6A). Metabolomics analysis showed that there was a significant increase of amino acids such as tyrosine and tryptophan when cells enter late growth stage (Figure 2.6 A). These aromatic amino acids are synthesized from erythrose 4-phosphate (E4P) and phosphoenolpyruvate (PEP)(Hall et al., 1982). E4P directly participates in the Calvin-Benson-Bassham cycle whereas PEP can be derived from pyruvate(Savakis and Hellingwerf, 2015). The accumulation of these amino acids could indicate a competition with the MEP pathway for metabolite precursors in the late growth stage and thus led to the limonene productivity decrease.

Second, metabolomics analysis revealed the carbon partition competition between sugar metabolism and terpene metabolism. Another group of metabolites that accumulated in higher amount during late growth stage of growth are sugars and sugar phosphates molecules (Figure 2.6A). The level of glucose-6-phosphate (G6P), fructose-6-phosphate (F6P), and sucrose increased significantly along the growth (Figure 2.6B). The increased levels of these metabolites along the cyanobacterial growth stages indicate the photosynthate relocation toward storage carbons when cells approach high densities. Interestingly, pyruvate levels were kept similar during these growth stages (Figure 2.6B), suggesting that it might not be a limiting substrate for the MEP pathway. The fact partially explains the feeding experiment result where pyruvate supplementation failed to enhance limonene production (Figure 2.4). Overall, the comparative metabolomics suggests an activate competition for photosynthate output between the MEP pathway and other photosynthate-consuming pathways including both carbon storage and amino acid biosynthesis.

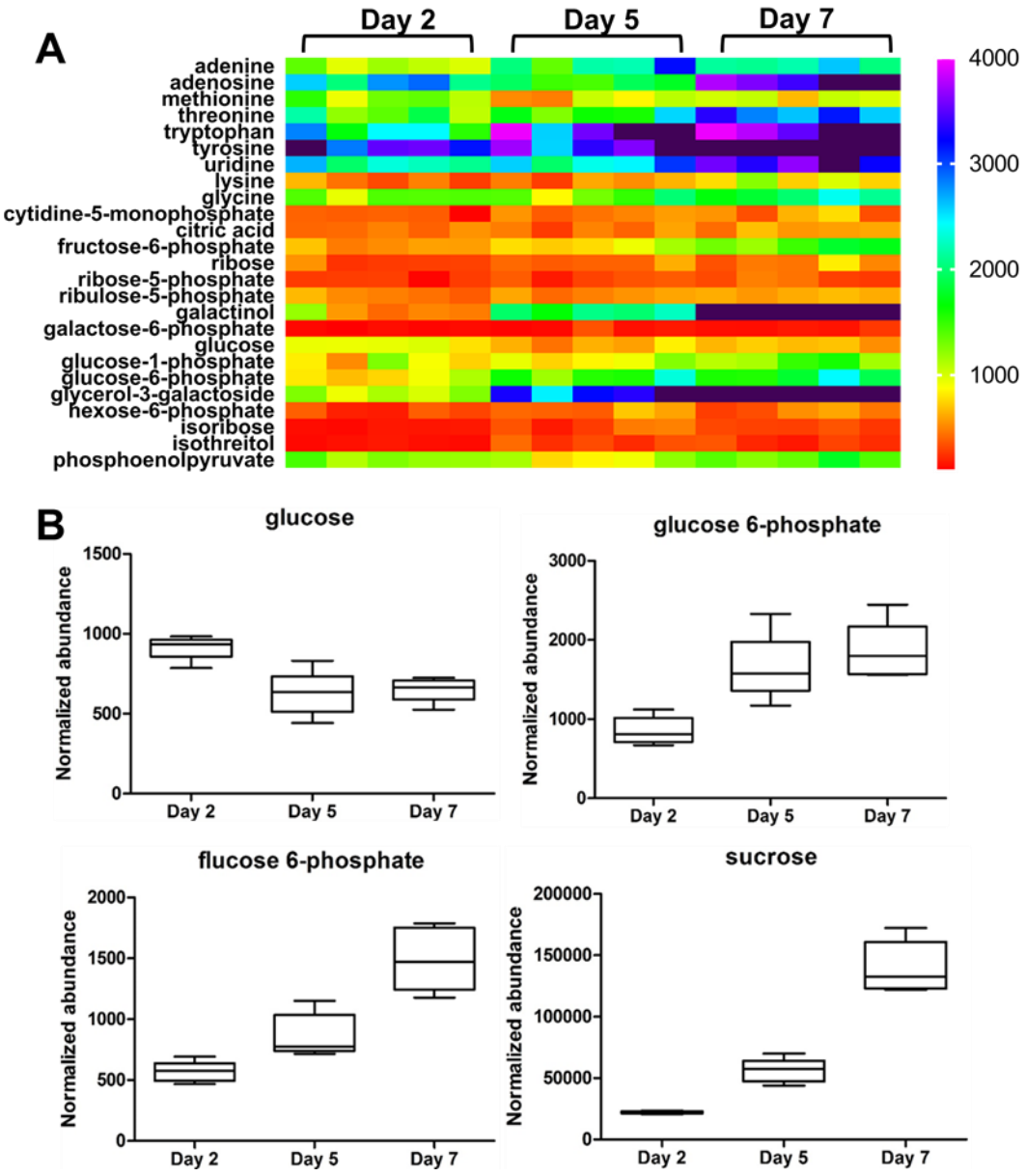

Figure 2.6 Metabolomics analysis reveal the MEP flux limitations. (A) Overview of metabolomics at day 2, day 5 and day 7. (B) Detailed metabolites changes at day2, day5 and day 7. Five biological replicates at each growth stage were collected for metabolomics analysis.

2.4.3 Carbon repartitioning from carbon storage to terpene biosynthesis significantly enhanced limonene production

Proteomic and metabolomic analyses indicate a tight regulation in carbon flux distribution between the MEP pathway and other central pathways. The inferior ability of MEP pathway in serving as a photosynthetic carbon sink calls for engineering strategies to reallocate carbon flux from central metabolic processes to terpene biosynthesis. Metabolomic results showed sucrose accumulation when cells grow to high density.

Sucrose is synthesized from the photosynthates through two enzymatic steps (Figure 2.7). In the first step, UDP-glucose and F6P are first converted into sucrose-6-phosphate catalyzed by sucrose-phosphate synthase (encodes by *sps*). The sucrose phosphatase then converts sucrose-6-phosphate into sucrose (Porchia and Salerno, 1996)(Figure 2.7).

Figure 2.7 Sucrose biosynthesis pathway in *S. elongatus* PCC7942.

To achieve carbon repartition from sucrose to terpene biosynthesis, we generated a mutant by insertional mutating the gene *sps* from the L1118 genome, creating strain Lsps. The clean mutant was verified through antibiotic selection and segregation PCR (Figure 2.8).

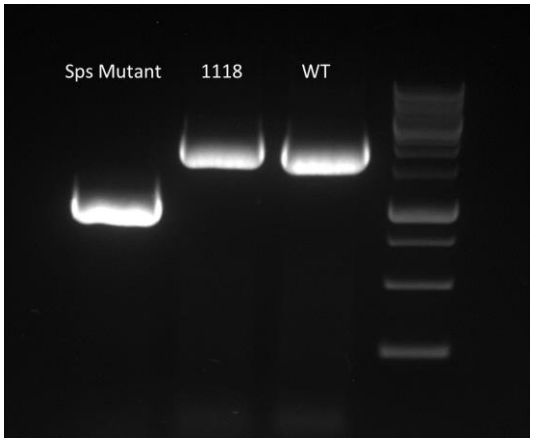

Figure 2.8 Insertional inactivation of *sps* gene in *S. elongatus* 7942. DNA fragments amplified by PCR showing complete segregation of the inactivated gene. In Lsps, the sucrose phosphate synthase gene locus was inserted and replaced with a gentamycin resistance gene and resulted in a PCR product of 1255bp.

The mutant Lsps effectively increases limonene productivity. To confirmed that engineering carbon repartition from primary metabolism to terpene metabolism, we have compared both terpene yield and sucrose accumulation. Compared to strain L1118, the sucrose mutant Lsps showed slightly lower biomass accumulation (Figure 2.9 B). However, both limonene productivity and the final titer is significantly higher(p-value:0.0034<0.05) in the Lsps strain (Figure 2.9 C, D). Overall, the productivity of Lsps increased to 1100µg/L/OD/day as compared to that of L1118, and the tier reached about 4 mg/L after six days of cultivation (Figure. 2.9D).

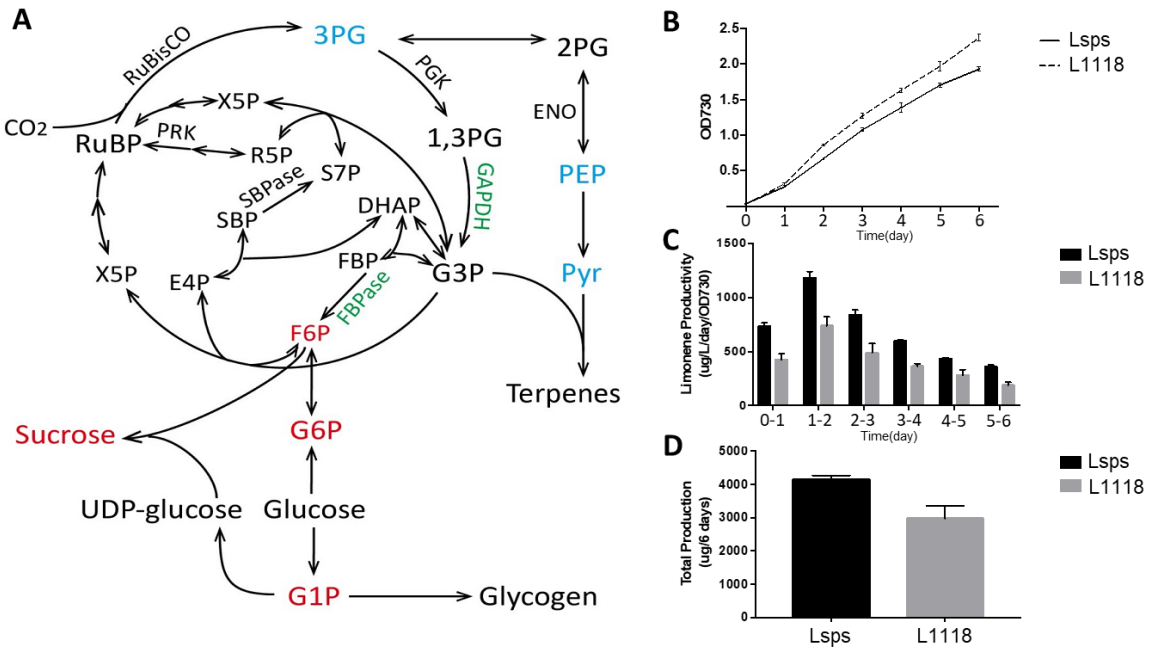

Figure 2.9 Sucrose mutant increased the limonene productivity. A. Overview of metabolomics and proteomics change. (Red color: Increase, Blue color: Decrease). B. Growth curve of Lsps and L1118. C. Daily limonene productivity of Lsps and L1118. D. Total production of limonene in Lsps and L1118.Each strain were grown in three biological replicates.

In order to further validate the carbon repartition from sucrose to terpene, we measured the sucrose content of the Lsps and L1118 cells using HPLC. No detectable amount of sucrose was found in the Lsps (Figure 2.10). The results confirmed that the deletion of *sps* gene efficiently blocked the sucrose biosynthesis, and effectively redirected carbon flux from sucrose to limonene biosynthesis. The engineering of carbon repartition between sugar and terpene metabolism thus represents an effective approach to improve terpene yield.

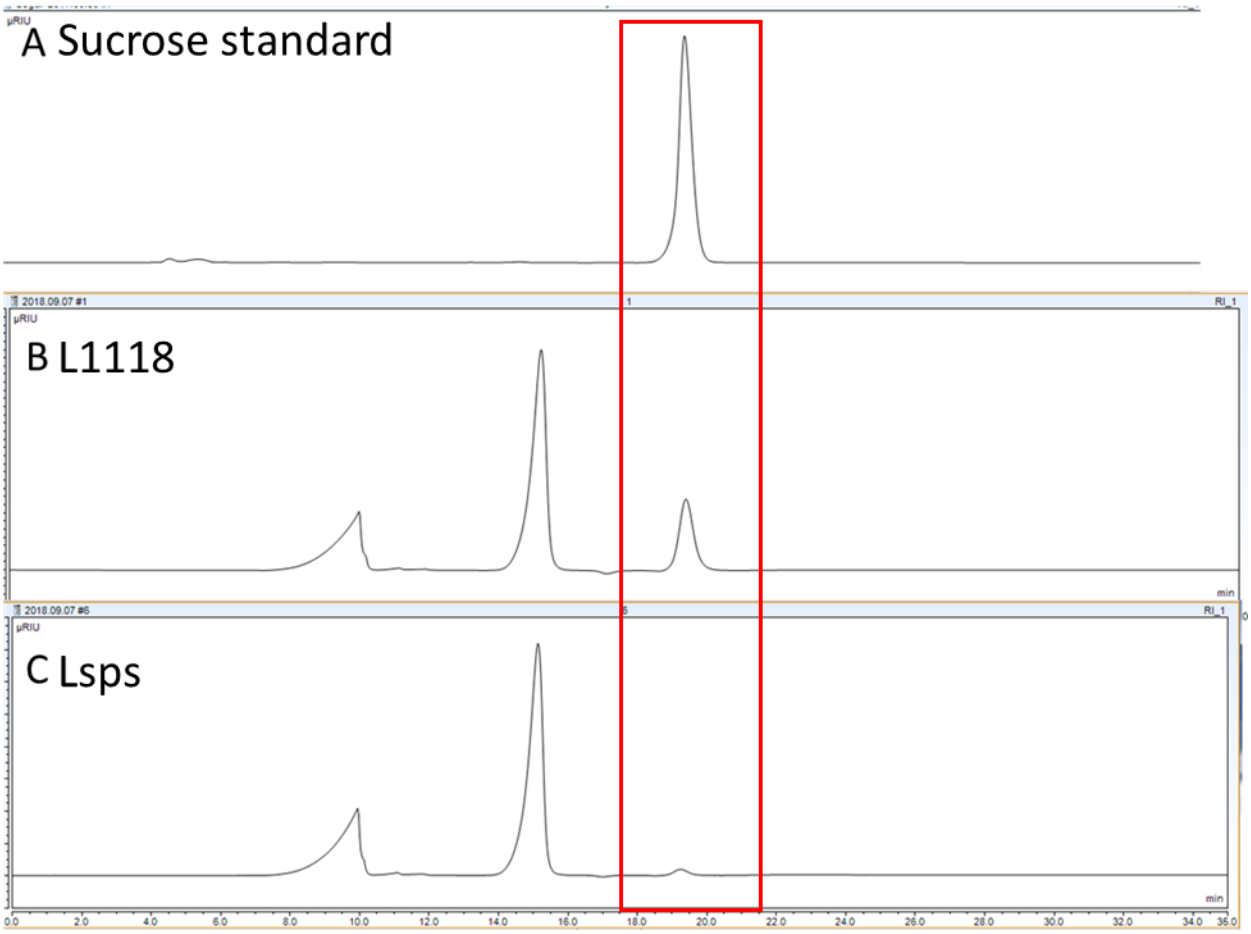

Figure 2.10 Representative chromatogram of sucrose chromatogram of L1118 and Lsps. A: Sucrose Standard B: 1118 Strain C: Lsps Strain. The sucrose retention time was 19.23min. And the peak was found in 10g/L sucrose standard solution. And L1118 containing the corresponding peaks at RT 19.23 were collected and used for further calculation.

2.4.4 A rewired carbon metabolism supports the enhanced limonene production

To further understand the carbon repartition between sugar metabolism and terpene biosynthesis, we carried out another proteomics study to analyze the protein expression profile between Lsps and L1118 strains. Compared to L1118, several enzymes related to photosynthesis were downregulated in the Lsps (Table 2.2), including photosystem II reaction center CP47, CP42, D2 and lipoprotein. Moreover, the reduced expression of ATPase and ATP synthase, NAD(P)H-

quinone oxidoreductase subunits(Ndh1)K, M, N suggested a slower light reaction rate in Lsps. These decreases could lead to the decreased biomass accumulation in the Lsps (Figure 2.9B).

Sucrose is synthesized from glucose-1-phosphate(G1P). Phosphoglucomutase catalyzes the reaction between G1P to glucose-6-phosphate(G6P) (Figure 2.11). In Lsps, phosphoglucomutase was showed a decreased expression level. By blocking the sucrose biosynthesis pathway in Lsps, there would be an excessive accumulation of G1P and thus inhibited the expression of phosphoglucomutase. Moreover, glucose-1-phosphate adenylyltransferase from glycogen biosynthesis pathways showed increased expression level in Lsps.

Enzyme involved in fatty acid biosynthesis pathway (Figure 2.11) was also upregulated including 3-oxoacyl-(Acyl-carrier protein) reductase (FabG), and the 3-oxoacyl- (Acyl-carrier protein) (ACP) synthase was downregulated correspondingly. FabG catalyzed the first reductive step in fatty acid biosynthesis by the reduction of 3-oxoacyl-ACP to 3-hydroxyacyl-ACP(Sheldon et al., 1990). The increase of 3-oxyacyl-reductase and decrease of synthase suggested an increased fatty acid biosynthesis in Lsps. Isocitrate dehydrogenase, first enzyme in tricarboxylic acid cycle (TCA), was also upregulated suggesting an enhanced carbon flux of citrate in TCA cycle. With the blocking of sucrose biosynthesis pathway in Lsps, photosynthesis outputs were partially redirected from sucrose biosynthesis to limonene synthesis and resulted in higher limonene yield in Lsps. Other metabolisms including fatty acid and glycogen biosynthesis, TCA cycle inputs were also upregulated in Lsps.

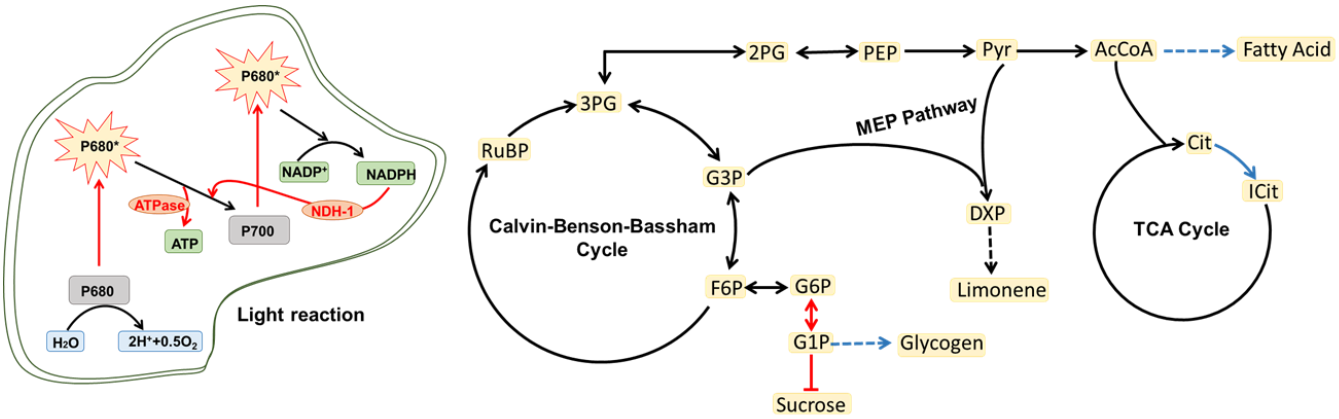

Figure 2.11 Scheme of metabolic pathway changes in Lsps comparing to L1118. Blue arrow: related enzyme expression level increase, orange arrow: related enzyme expression level decrease

Metabolism	Enzymes	Relative Abundance Lsps *vs.* L1118	p-value
Light reaction	NAD(P)H-quinone oxidoreductase subunits K, M, N	0.33,0.29,0.47	**,***,**
	ATP synthase subunit beta	0.36	**
	ATPase	0.23	***
	Photosystem II reaction center CP47, CP43, lipoprotein	0.40,0.39,0.44	***,***,*
	Photosystem II D2 protein	0.36	**
	Photosystem I reaction center subunit IV	0.22	***
TCA Cycle	Isocitrate dehydrogenase [NADP]	2.01	**
Sugar metabolism	Glucose-1-phosphate adenylyltransferase	1.44	**
	Phosphoglucomutase	0.41	**
Fatty acid synthesis	3-oxoacyl-(Acyl-carrier protein) reductase	2.05	***
	3-oxoacyl-(acyl-carrier-protein] synthase	0.19	***

Table 2.4 Enzyme expression level changes in L1118 and Lsps. *: p-value<0.1, **: p-value<0.05, ***: p-value<0.01. Strains L1118 and Lsps for proteomics were harvested at day 5 in three biological replicates.

2.5 Discussion

The results in this study highlighted that carbon repartition from primary metabolism to secondary metabolism could be an effective approach to enhance the productivity of secondary metabolites. Since the discovery of MEP pathways(Eisenreich et al., 2001), enzymes involved in condensation of pyruvate and G3P has been identified gradually(Eisenreich et al., 2001). Intensive research has been carried out for understanding the carbon partitions between the different metabolic pathways. In this study, systems biology-based approaches were taken to thoroughly understand carbon partitions during limonene production. By measuring the daily limonene product from L1118, it was found that the limonene productivity was associate with the cell growth stages. The limonene productivity was showed highest in the log phase and decreased in the stationary phase. From the proteomics result, we found decreased enzyme expression involving in the CBB cycle, photosynthesis light reactions, and the MEP pathways. Enzymes involved in acetyl-coA metabolism, fatty acid synthesis and amino sugar synthesis were upregulated during the stationary growth stage in L1118. These differential expressions indicate the changes of carbon partitions between the MEP pathway and other major metabolites. To further demonstrate our hypothesis for the carbon partitions, the major metabolites in L1118 were quantified and analyzed. In the stationary phase, the accumulation of amino acids, sugar and sugar phosphates along with the decreased limonene productivity verified the previous hypothesis about the carbon partitions between the primary and secondary metabolites. Consequently, further metabolic engineering might lie in the carbon reparations to channel the carbon flux from primary metabolism to MEP pathway. Moreover, the optimization of cultivation technologies to keep the cell in the exponential and late exponential growth would also help to improve the MEP flux and limonene yield.

The genetic and metabolic regulations with the MEP pathway has been discussed broadly but the deeps insights from the 'source' to 'sink' limitations are still missing(Banerjee and Sharkey, 2014). The 'source' G3P and pyruvate were considered key limitations for the carbon flux in MEP pathways. Genetical modifications to increase G3P and pyruvate level in *E. coli* enhanced the isoprene production(Liu et al., 2013). To increase the G3P and pyruvate level in L1118, a feeding experiment was carried out. However, the limonene productivity was not increased correspondently in our case. Moreover, overexpression of *GAPDH* and *pgk* to increase the CBB cycle efficiency and the 'source' did not increase limonene production in L1118. Then we hypothesized that the 'source' limitation was not the key factor for limonene production in cyanobacteria, especially in the late growth stage. With further understanding of the regulations in CBB cycle and MEP pathway from proteomics results, we then focus on increasing the MEP pathway efficiency to increase the 'sink' by overexpressing enzymes involved in the MEP pathways. IspG, a [4Fe-4S] cluster-containing protein(Xiao et al., 2009) was downregulated in the stationary phase. However, several strategies to overexpress *ispG* failed to increase the limonene productivity as we expected. These results indicated the 'sink' limitations also not controlling the carbon partition for limonene biosynthesis. Inspired by the previous results of sucrose accumulation in the stationary phase, rechanneling of the carbons from sucrose to MEP pathway was employed to enhance limonene production. The strategy by deleting the sucrose phosphate synthase led to significantly improved limonene productivity and yield. Moreover, subsequent proteomics was done to understand the metabolisms changes under the sucrose deficiency. The fatty acid biosynthesis, TCA cycle and glycogen biosynthesis were upregulated correspondingly with the increased carbon flux by sucrose deficiency. Glycogen is a major carbon storage source other than sucrose, and the reduction in sucrose carbon flux could also led to the increase of

glycogen accumulation in addition to enhanced terpene biosynthesis. Such a change in glycogen content could have led to the differential expression of the enzymes involved in glycogen metabolism. As Lsps could not store the carbon into the form of sucrose, the photosynthate generated from CBB cycle was partially redirected into pyruvate and resulted in higher limonene production in Lsps. Pyruvate could be involved into TCA cycle and fatty acid synthesis by converted into acetyl-coA. The rewired carbon flux from sucrose to fatty acid, and glycogen also provides instructions for further genetical engineering to increase the terpene yield.

The plasticity of cyanobacterial carbon metabolism supports the sustainable production of photosynthetic terpenes(Xiong et al., 2017). The flexible of carbon flux distribution and adaptable to genetical modifications makes them an ideal cell factory for targeted chemicals(Machado and Atsumi, 2012). However, the low carbon partitions to MEP pathway impedes the terpene production(Melis, 2013). The MEP pathway limitations were broadly considered as the bottleneck for enhancing carbon flux towards the terpene biosynthesis. Previous study proved the MEP 'sink' limitation by unleashing the 'sink' with strong limonene synthase activity(Wang et al., 2016). Increase the MEP efficiency with synthetic biology design increased photosynthetic isoprene production(Gao et al., 2016). Although our strategies with *ispG* overexpression failed to increase the terpene carbon flux, several enzymes including the IspE and IspH might be the limiting factors considering their differential expressions along with limonene productivity decrease. Moreover, the enhancement of 'source' also plays an important role in photosynthetic terpene production. Engineered heterologous xylose utilization enhanced the ethylene production in *Synechocystis sp.* PCC6803(Lee et al., 2015). Our results proved the carbon partitions between the 'source' and 'sink' also plays an important role in photosynthetic terpene production. Further metabolic

engineering could be employed in the aspect of increasing the MEP pathway capacity by enzyme modifications. Increasing the carbon fixation efficiency by modification of light reactions, CCM and CBB metabolic would also help to lift the ‘source’ limitations. Additional strategies need to be developed to further increase limonene yield to enable commercial application, in particular, for the biofuel industry. Ultimately, the efficient conversion of CO_2 to terpene will provide an effective solution for ‘drop-in’ biofuel production and biomanufacturing of high-value products.

CHAPTER III

REDIRECTING THE CARBON FLUX FROM GLYCOGEN STORAGE TO TERPENE IN CYANOBACTERIA

3.1 Summary

Photosynthetic terpene production represents one of the most energy and carbon-efficient route for the production of 'drop-in' hydrocarbon biofuel from CO_2. Despite the potential, the inherent metabolic rigidity hinders the partitioning of carbon into secondary metabolites like terpenes. In this study, we redirected the carbon from the primary metabolite, glycogen by blocking the glycogen biosynthesis pathway in cyanobacteria, and with optimized continuous growth conditions, the mutant strain LglgC reached a yield of 14.9mg/L limonene production in *S. elongatus* PCC7942.

3.2 Introduction

The demands for renewable energy have become urgent in recent years due to greenhouse gas emission and various environmental concerns associated with fossil fuel consumption(Himmel et al., 2007). Moreover, increasing concerns of global warming promote the development of alternative renewable energy sources. First and second-generation production of biofuel such as bioethanol are using the released carbohydrates from crop plants through both enzymatic hydrolysis and microorganism fermentation processes(Naik et al., 2010). Despite the potential of crop plants, lignocellulosic biomass conversion is more challenging considering the recalcitrant

nature of biomass and the need to disrupt secondary cell wall structures to release carbohydrates and lignin for microorganism fermentation(Isikgor and Becer, 2015; Balch et al., 2017). The pretreatment prior to microorganism processing has been developed to overcome the lignocellulosic biomass recalcitrance such as steam explosion, diluted acid, hot water, and organic solvents have been developed. However, the processing cost and chemical usages raise economic and environmental questions about these practices(Lau et al., 2009; Zhao et al., 2009; Lima et al., 2013; Zhang et al., 2016; Zhuang et al., 2016). Therefore there is a growing interest in using photosynthetic microorganisms for biofuel and biochemical production(Rosgaard et al., 2012).

Cyanobacteria is a phototrophic microorganism, only requiring CO_2, sunlight, and essential minerals for their growth and possesses great potential for desired biofuel and biochemical productions(Lu, 2010). The strong ability to fix CO_2 into cellular composites using water and light, and thriving in diverse environments has attracted researchers' attention to their potentials as cell factories for renewable biofuels and biochemical production(Rosgaard et al., 2012). Moreover, cyanobacteria cell walls are composed of a peptidoglycan layer and an outer membrane instead of the recalcitrant lignocellulosic biomass in plants which make them more economically and environmentally friendly for bioethanol fermentations than plants(Hoiczyk and Hansel, 2000). Currently, biofuels and desired chemical have been produced through heterologous expression of genes and other metabolic engineering in cyanobacteria, including ethanol(Dexter and Fu, 2009; Gao et al., 2012), isopropanol(Kusakabe et al., 2013; Hirokawa et al., 2015), 1-butanol(Lan and Liao, 2012a; Lan et al., 2013), fatty acids (Liu et al., 2011; Ruffing and Jones, 2012) and ethylene(Takahama et al., 2003; Jindou et al., 2014).

Terpenes are a diverse group of hydrocarbons mainly synthesized by microorganisms, plants and insects as toxins, repellents, or attractants(Gershenzon and Dudareva, 2007). The increased demand of terpenes in the field of nutraceuticals, antioxidants, cosmetics, and drugs urges the research of biosynthetic terpenes in microorganisms(Ajikumar et al., 2010). In addition, the competitive energy density of terpenes compared to gasoline makes them a promising candidate for biofuels. The photosynthetic terpenes from cyanobacteria not only address the needs of terpenes production in microorganisms but also act as an environmentally friendly source compared to the sugar consuming heterotrophic systems. Terpenes are generated from C_5 precursors isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate DMAPP. In cyanobacteria, IPP and DMAPP are synthesized from the methylerythritol phosphate (MEP) pathway (Chatzivasileiou et al., 2019). The MEP pathway starts from pyruvate and glucose-3-phosphate and condenses the photosynthates from Calvin-Benson-Bassham-cycle eight enzymatic steps.

Limonene, a C-10 isoprenoid, is primarily produced from the essential oils of citrus fruits. It is commercially used as a fragrance in perfumes and now is considered to be a “drop-in” replacement for jet fuels(Chuck and Donnelly, 2014). The low boiling point and high volatility make them much easier to collect and less toxic to cells during biosynthesis in cyanobacteria. However, the competing carbon pool between major metabolites such as sugars and the secondary metabolites such as terpenes are still not well studied. Glycogen, a multibranched polysaccharide of glucose is a major carbon sink for energy and carbon storage from the CBB cycle in cyanobacteria and may take up to 50% of the cellular dry weight(Aikawa et al., 2014). The glycogen metabolism makes

important contributions for cellular activity in diurnal cycles when the photosynthesis activities are suspended.

In this study, by comparing the metabolic changes in the limonene producing strain and wild type, we identified the carbon flux to glycogen storage that limit their partitions to MEP and we successfully engineered a glycogen deficient strain in *Synechococcus elongatus* PCC7942 for high limonene production.

3.3 Materials and methods

3.3.1 Plasmid and strain constructions

Previously established strain *Synechococcus elongatus* PCC7942 L1118(with the overexpression of *limS* under a pPsbA promoter at the neutral site I in the wild type PCC7942 chromosome) and wild type PCC7942 were used as platforms in this study. For insertion mutagenesis of *glgC* in PCC7942, plasmid pΔ0603 obtained from Dr. James W. Golden's lab was used for for the mutation of *glgC* which encodes glycose-1-phosphate adenylyltransferase. *glgC* coding region was replaced by a kanamycin resistance cassette through homologous recombination. The clean mutant strain LglgC with kanamycin resistance cassette insertion was verified with two pairs of segregation primers together with PCC7942 wild type.

For overexpression of isopentenyl-diphosphate delta-isomerase (encodes by *idi)* in LglgC, the origin region from pAM2991(purchased from addgene) was used for the *idi* overexpression plasmid pLM232 backbone construction. Neutral site III upstream and downstream sequences were cloned from PCC7942 genomic DNA for the homologous recombination and insertion of

target and antibiotic resistance genes. Then the *idi* sequence from *Saccharomyces cerevisiae* was used for *idi* overexpression. The codon optimized *idi* sequence together with a pLacI promoter was inserted between the neutral site III upstream and downstream sequence resulting the plasmid pLM232. All the DNA fragments used in this study were amplified with Phusion High-Fidelity DNA Polymerase (New England Biolabs), and purified using Zymoclean Gel DNA recovery Kits (Zymo Research).

Transformation of the *S. elongatus* was carried out by chromosomal recombination. 200μg of each plasmid was mixed with the 200 μl specific strain cultures in dark for 24 hours. For the construction of LglgC, pΔ0603 was mixed with L1118. For the construction of LglgC-IDI, pLM232 was mixed with LglgC. And for LglgC-sps, pLM21(chapter 2) was mixed with LglgC. The cells were plated and the transformants were seleted on BG11 agar with appropriate antibiotics kanamycin (5 μg/L) and gentamycin (2 μg/L). The result of completely transformation in all copies of chromosome was confirmed by segregation PCR.

3.3.2 *S. elongatus* PCC7942 cultivation conditions

S. elongatus PCC7942 wild type and engineered strains were grown in a BG11 medium (Sigma), supplied with10 mM N-[Tris(hydroxymethyl)methyl]-2-aminoethanesulfonic acid (TES, pH=8.2) (Sigma) at 30 °C and 37 °C. The strains were first cultured in 30 mL BG11 medium suppled with appropriate antibiotics in 250 ml flask with 75 μmol photons $m^{-2}s^{-1}$ for genotype verification and then transferred into 500 or 300 mL BG11 medium and cultured a 1-l Roux bottle with continuous 5% CO_2 bubbling and 100 or 150 μmol photons $m^{-2}s^{-1}$. For continuous growing for high limonene

yield, 100ml culture was removed with the addition of 100 mL fresh medium, and doubled light intensity was provided after 3 days from the initial inoculation.

3.3.3 Limonene collections and quantification by GC-MS

S. elongatus PCC7942 limonene producing strains were grown in triplicates in a 1-L Roux bottle equipped with a HayeSeq polymer trap on the cap of the bottle. The vaporized limonene was collected each day and eluted with 1mL hexane containing 50 µg/ mL cedrene as the internal standard. The recovery rate of limonene from the bottle was calculated based on the limonene collected from 300 mL fresh BG11 with 100 µg limonene. Limonene standards with series concentrations of 62.5, 125, 250, 500 µg/mL limonene, and 50 µg/ mL were used for establishing a standard curve. The samples were analyzed by gas chromatography-mass spectrometry (GC-MS). GC-MS was performed on a GCMS-QP2010SE (Shimadzu Scientific Instruments). For each run. 1 µL sample was injected into a Shimazu SH-Rxi-5Sil column (30 m × 250 µm × 0.25 µm) (Shimadzu Scientific Instruments) with1 mL/min helium as carrier gas. The temperature was held at 50 °C for 3 min, and then increased to 140 °C at 20 °C/min. Mass spectral peak quantification was performed using GCMSolution software Ver. 2.6.

3.3.4 Estimation of glycogen content in *S. elongatus* PCC7942

50 mL cultures of PCC7942 wild and engineered strains were harvest and centrifuged at 8,000rpm, 4 °C for 10 mins. The pellets were washed twice with 50 mL 0.9% NaCl and then lyophilized at -55 °C with 0.05 mbar for 24 h using a 4.5 L Benchtop Freeze-Dry System (Labconco Corp., USA). The lyophilized cells were weighed for further normalization. The lyophilized cells were digested

with 1ml of 30% KOH at 100 °C for 20 min. Then the samples were cooled in an ice bath. 1.25 mL of 95% ethanol was added into the sample. The mixtures were vortexed foo 30 s. The samples were then centrifuged at 4 °C for 15 min at 13,000 rpm and the supernatant was discarded. The pellets were dissolved in 1 mL of distilled water and then reprecipitated with 1mL of 95% ethanol as stated before. The precipitates were then dissolved in 5ml distilled water. 10ml of prechilled 0.2% anthrone reagent (0.2% anthrone in 95% sulphuric acid) was added to each sample and incubated at 100°C for 10 min. The contents were cooled at room temperature and the color was read at 680 nm. The content of glycogen was calculated against glucose standards with series concentrations of 25, 50, 100, 150, 250 mg/ mL glucose and normalized with dry cell weight.

3.3.5 ATP determination in *S. elongatus* PCC7942

1.5 mL cultures of PCC7942 engineered strains were collected and harvested through centrifuge at 13,000rpm for 5mins. The pellets were washed with 0.9% NaCl twice. 100µL prechilled 1% TCA was added to each sample. The samples were vortexed for 30s and then centrifuged at 4°C for 10 mins. 100µL of the supernatant was collected and neutralized with 100µL 1M Tris-acetate pH=7.8. 800µL distilled water was then added to each tube. ATP Determination Kit (Molecular Probes) was used for ATP measurement. 10µL of the previously prepared sample was added to each well of a Corning 96-Well Black Plates (Sigma Aldrich) to reduce the crossing noise. 10 µL ATP standards with a concentration of 50, 100, 250, 500, 1000 nM/mL were measured each time together with the samples. The standard reaction solution containing luciferase and D-luciferin was prepared following the manufactures instruction. 100µl reaction solution was added to each

well and gently mixed with the samples. The luminescence was collected, and ATP concentration was calculated on SpectraMax i3x (Molecular Devices).

3.4 Results and discussion

3.4.1 Reduced glycogen in limonene production strain *S. elongatus* PCC7942 L1118

L1118 was a previously built limonene producing strain from Dr. Xin Wang's engineering work in increasing the limonene synthase activity(Wang et al., 2016). By employing synthetic biology tools such as codon optimization, synthetic ribosome binding site, and strong promoter psbA, they achieved a record limonene yield which indicated the strong ability of L1118 for limonene production. However, the carbon pool to secondary metabolites such as terpenes in cyanobacteria is still low compared to the major metabolites. Glycogen can make up to 50% of the cell dry weight in cyanobacteria as an energy bank in cells(Cano et al., 2018). To further investigate the potential in cyanobacteria for terpene production, we first analyzed the glycogen content between L1118 and wild type strain (Figure 3.1).

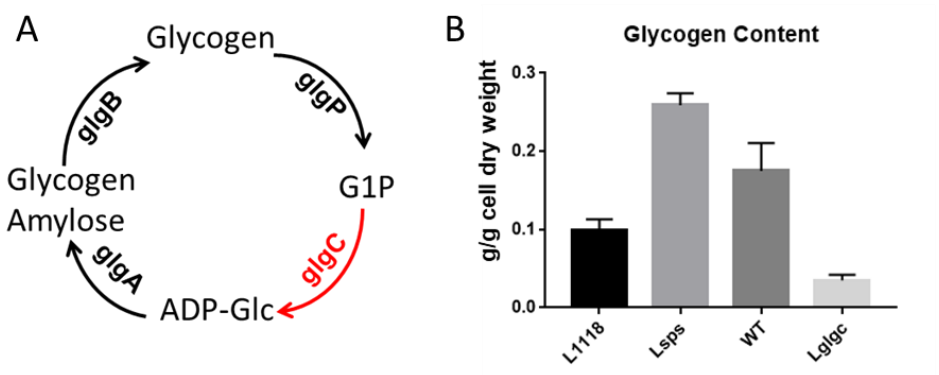

Figure 3.1 Glycogen synthesis pathways in *S. elongatus* PCC7942(A), and glycogen content comparation between L1118, Lsps, LglgC and wild type PCC7942. G1P, glucose 1-phosphate; *glgC*, glucose 1-phosphate adenylyltransferase; *glgA*, glycogen synthase; *glgB*, 1,4-alpha-glucan branching enzyme.

Compared to wild type strain, L1118 showed a reduced glycogen content (Figure 3.1). This decrease was due to part of the carbon fixed in the CBB cycle were directed into the MEP pathway for limonene synthesis. We also compared the glycogen content in Lsps which was a sucrose deficient strain built in Chapter3. When a major carbon storage pool-sucrose was blocked, the glycogen content increased dramatically in Lsps compared to L1118 and wild type. However, the regulations between energy storages and MEP carbon flux in cyanobacteria remains unclear. We then hypothesized that glycogen deficiency in L1118 would help in redirecting the carbon into MEP.

Cyanobacteria uses glucose 1-phosphate (G1P) as the precursors, and ADP-glucose (ADP-Glc) as an intermediate for glycogen synthesis. Three enzymes are involved in glycogen biosynthesis including glucose 1-phosphate adenylyltransferase (encoded by *glgC*), converting G1P to ADP-

Glc; glycogen synthase encoded by (*glgA*), converting ADP-Glc to glycogen amylose and 1,4-alpha-glucan branching enzye (encoded by *glgB*)(Preiss, 2006) (Figure 3.1 A). By disrupting the *glgC* gene in L1118 genome with the insertion of a kanamycin resistant gene, the glycogen synthesis pathway would be blocked at the first step. Then the glycogen content was compared between mutated strain LglgC and other strains (Figure 3.1). We found a decreased glycogen content in LglgC. As the glycogen was measured in form of glucose, some free glucose was still detected in LglgC. The reduced total carbohydrates content in LglgC promotes the hypothesis that the decreased of total carbohydrates would increase the terpene yield in LglgC.

3.4.2 Enhanced limonene productivity in the glycogen deficient strain

LglgC and L1118 was cultured at same time for the limonene productivity comparation. The photosynthetic growth of LglgC was found to be a little slower compared to L1118. As shown in Figure 3.2, the growth of both strains measured from inoculation at a low cell density showed little difference from day 0 to day 2. However, at day 3, the OD_{730} in L1118 reached 2.05 comparing to 1.63 in LglgC. Glycogen is the major energy balancing bricks for cyanobacteria cell growth especially when the photosynthetic activity was paused in dark conditions(Cano et al., 2018). When the cell density was high after two days, the cyanobacteria cells were unable to access enough light for their photosynthesis. In L1118, the energy from glycogen hydrolysis could partially make up for continuous cell growth while the glycogen deficient strain did not have enough energy for cell growth and then resulted in a slower growth at day 3.

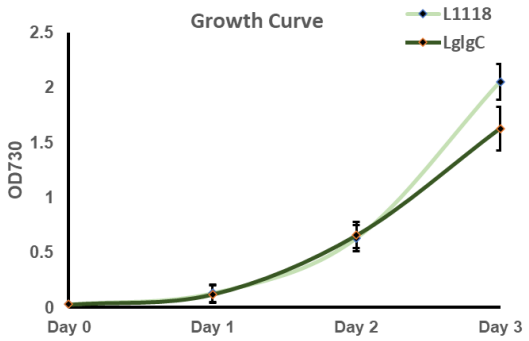

Figure 3.2 Slower growth of L1118 and LglgC. Strains cultured in 500 mL BG11 medium under conditions of 5% CO_2 and 100 µmol photons m^{-2} s^{-1} at 37°C.

Limonene was collected daily in L1118 and LglgC. As shown in Figure 3.3, there was a significant increase of limonene production over three days with a p-value of 0.003. In the first day, the limonene productivity in LglgC was 802±38µg/day/L which is 2.46 folds of the L1118. From day 1 to day 2, the limonene productivity in LglgC was 1560±125µg/day/L compared to L1118 1028±80µg/day/L. And at day three, the limonene productivity is 2626±310µg/day/L which is 2.37 folds of L1118. The increased limonene productivity in LglgC proved the feasibility of redirecting the carbon flux from glycogen storage to limonene production in cyanobacteria. Although the growth was delayed in LglgC, the 3-day total limonene production was 4989±732µg/L compared to 2460±732µg/L in L1118.

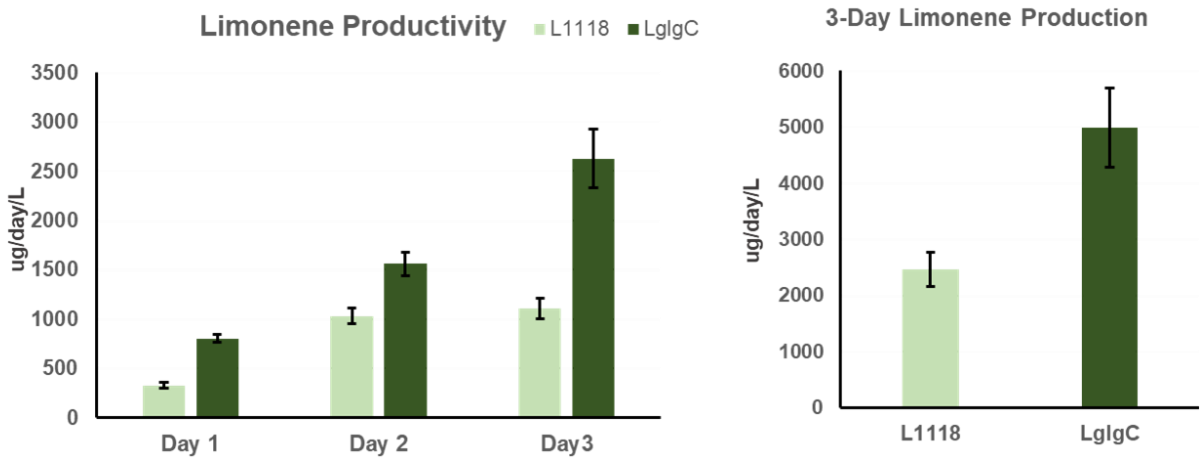

Figure 3.3 Increased limonene productivity in L1118 and LglgC.

The glycogen deficiency in LglgC did not show strong growth delays under the standard growth conditions, which indicated glycogen biosynthesis was not essential for continuous growth in cyanobacteria when the cells are under continuous light. More importantly, by blocking the glycogen biosynthesis in LglgC, we successfully redirected the carbon flux from primary metabolism into the secondary metabolite-limonene.

3.4.3 Impacts on the high-energy metabolites ATP in glycogen deficienct strain.

During the preparation of seed cultures for large scale limonene collection, we observed a color change in the growing cyanobacteria cultures (Figure 3.4). In L1118 cultures, the color changed from blue-green to yellow green when the cell density was continuously increasing. However, in LglgC, the culture remained blue-green even when the cell density was high. Previous studies reported the delayed degradation of phycobilisome rod protein contributed to remaining the blue-green color change in glycogen deficient *Synechocystis sp.* PCC6803(Carrieri et al., 2017). However, the mechanism of the delayed degradation of the phycobilisome is still unclear.

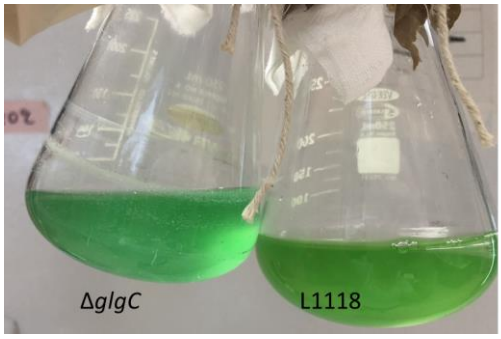

Figure 3.4 Comparation of growing culture color between LglgC and L1118. The strains were cultured with, 75 µmol photons m^{-2} s^{-1} at 37°C in 50 mL BG11 medium.

To further investigate the impact on the glycogen deficiency and limonene production increase in LglgC, we measured the relative ATP levels in LglgC and L1118. (Figure 3.5). In cyanobacteria, ATP is generated in the photosynthetic electron transport chain by converting solar energy and releasing the energy into chemical bonds which is used for cellular activities.

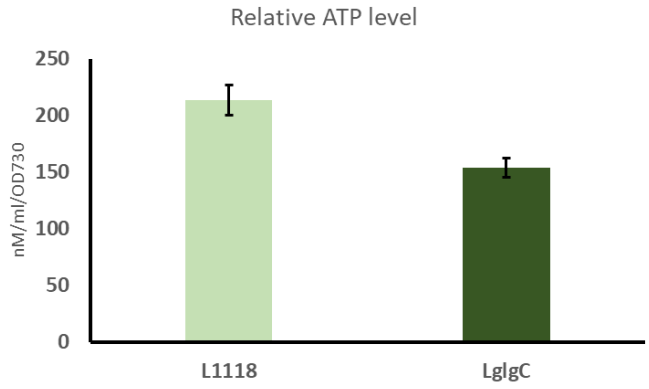

Figure 3.5 Changes in ATP level in LglgC and L1118 strain. ATP concentration was normalized with OD730. The strains were cultured with 5% CO_2, 150µmol photons m^{-2} s^{-1} at 37°C in 300 mL BG11 medium.

Compared to L1118, we saw a decrease of relative ATP level in the LglgC strain. ATP is consumed in the MEP pathway when the CDP-ME is converted to CDP-MEP with IspE (Figure 2.2). The

increased limonene production in LglgC would also increase the consumption of ATP. Also, the decreased growth was delayed in LglgC could also be partially explained by the decreased ATP level for sufficient cellular activities. To provide sufficient ATP for cellular activities, the delayed degradation of phycobilisome is required for supporting sufficient light reactions in LglgC. Lan and Liao reported that additional consumption of ATP was helpful in 1-butanol production in cyanobacteria by artificially engineering an ATP consumption pathway and showed the ATP concentration is also a factor in alternating metabolic flux (Lan and Liao, 2012b). It was hypothesized that the decreased ATP in LglgC might also involve in the carbon partition from glycogen to limonene biosynthesis.

3.4.4 Balancing IPP and DMAPP in LglgC for higher limonene productivity

As mentioned before, limonene was synthesized from two C_5 precursors IPP and its isomer DMAPP. However, IPP and DMAPP are at a ratio of 1:6 in *Synechocystis sp.* PCC6803(Barkley et al., 2004). We then hypothesized balancing IPP and DMAPP level by overexpression of isomerase encoded by *idi* might further increase the limonene productivity in cyanobacteria. *Saccharomyces cerevisiae idi* sequence was codon optimized against PCC7942 preference and then was integrated into at neutral site III in LglgC genome establishing the strain LglgC-IDI. Shown in Figure 3.6, the limonene productivity in LglgC-IDI was lower than LglgC in the first two days. The result revealed that the carbon pool for terpene synthesis was not limited by IPP and DMAPP ratio when light was sufficient in glycogen deficient strains. As reported by Wang et al, limonene synthase activity is the key limiting step in MEP limonene production. The excessive IPP and DMAPP produced by *idi* overexpression in the first two days might cannot be efficiently

converted to limonene by limited limonene synthesis and then resulted in a feedback inhibition to limonene production in LglgC-IDI. After two days of cultivation, the cell density was high and not enough light was accessible for cyanobacteria cells, and the total carbon pool from photosynthesis was decreased. Increased limonene productivity was observed in LglgC-IDI. Overall, when the carbon flux was not limited in the early growth stage, the balance between IPP and DMAPP is not the limiting factor to increase the limonene production in cyanobacteria.

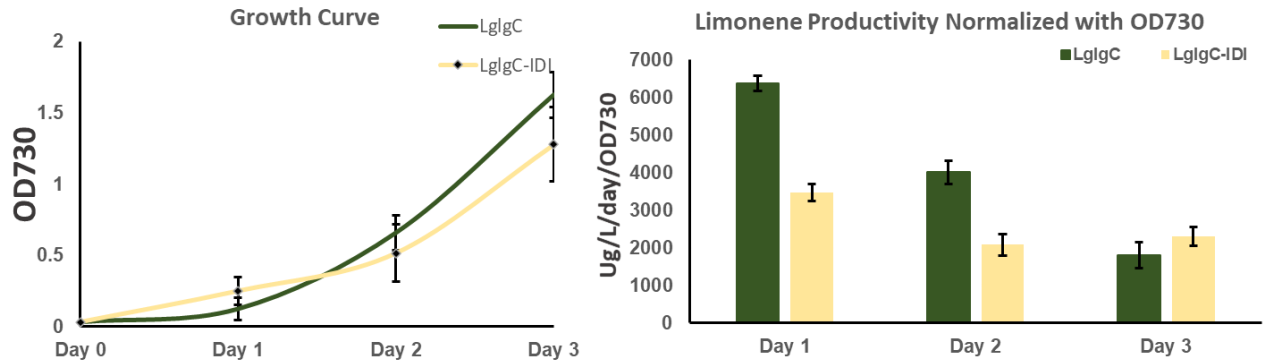

Figure 3.6 Growth(A) and limonene productivity(B) comparation between LglgC and LglgC-IDI. The limonene productivity was normalized with OD730 to compare the cellular productivity. The strains were cultured with 5% CO_2, 100μmol photons m^{-2} s^{-1} at 37°C in 500 mL BG11 medium.

3.4.5 Further increase carbon flux to limonene production by sucrose deficiency in LglgC

To further increase the limonene productivity and unearth the most potential in LglgC, we also tried to employ the sucrose deficiency strategy discussed in chapter 3 within the LglgC strain and built a double mutated strain LglgC-sps. As shown in Figure 3.7A and Figure 3.8A, the growth was dramatically decreased in LglgC-sps compared to LglgC. After three days of cultivation, the

OD730 in LglgC strain was 1.55-fold of the LglgC-sps. However, the deficiency in sucrose storage did not increase the limonene productivity in LglgC-sps in the first two days while the light was still sufficient in the strains for photosynthetic activity and subsequent CO_2 fixation process. On day 3, when the cell density was high and the light was limited, the limonene productivity in LglgC-sps showed a slightly increase which indicates the carbon flux was redirected from sucrose to limonene in LglgC-sps while the carbon 'source' pool was limited. As the growth was dramatically decreased in LglgC-sps, three-day limonene production in LglgC-sps was 2.6mg/L compared to 5.6mg/L in LglgC. Although the sucrose deficiency can redirect partial carbon flux to terpene production, the decreased growth still hinders higher limonene production in LglgC-sps.

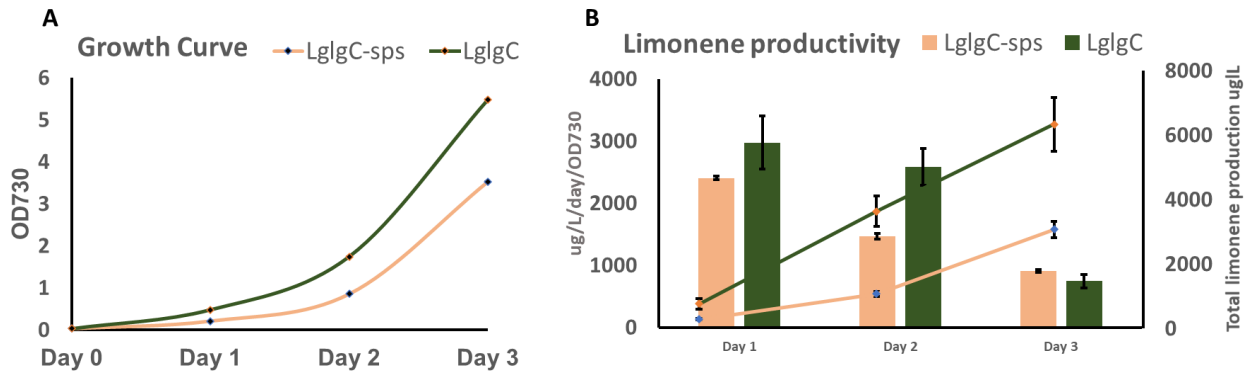

Figure 3.7 Growth(A) and limonene productivity(B) comparation of LglgC-sps and LglgC. Total limonene production was calculated and shown in Figure 4.7B. The strains were cultured with 5% CO_2, 150µmol photons m^{-2} s^{-1} at 37°C in 300 mL BG 11 medium.

To determine the reason for the decreased growth and decreased limonene productivity in LglgC-sps, the relative ATP levels were measured.

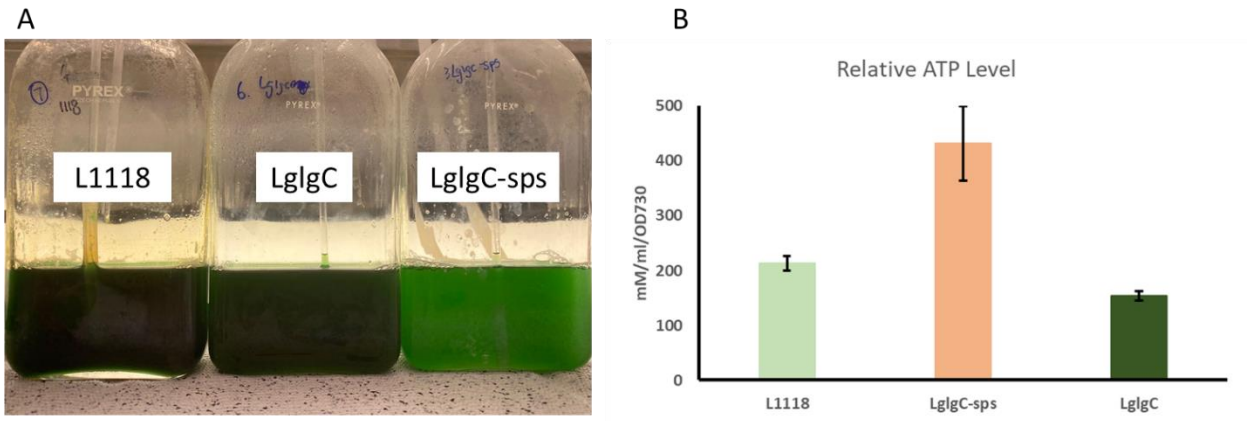

Figure 3.8 Growth(A) and relative ATP level(B) comparation between LglgC-sps, LglgC and L1118. Total ATP concentration was normalized with OD730. The strains were cultured with 5% CO_2, 150µmol photons m^{-2} s^{-1} at 37°C in 300 mL BG 11 medium.

Compared to L1118 and LglgC, the ATP concentrations in LglgC-sps were dramatically increased. In cyanobacteria, high energy metabolites ATP and NADPH are produced during photosynthesis and stored in the form of organic carbons. It was reported that the energy charge (ratio of ATP over ADP+ATP) in cyanobacteria is tightly linked to growth and carbon partitions(Cano et al., 2018). The deficiency in glycogen and sucrose synthesis blocked the storage of high energy compound ATP, and thus impede the growth of LglgC-sps.

Moreover, the growing conditions including light intensity, CO_2 accessibility, and temperature are also important limiting factors for both the growth and limonene production in cyanobacteria. Increasing the temperature from 30°C to 37°C shortened the time for L1118 to grow OD730 to 2 from 7 days to 3 days (Figure 2.1&3.2). Moreover, compared to the previously used conditions, 100 µmol photons m^{-2} s^{-1} in 500 mL BG11 medium, increasing the light to 150µmol m^{-2} s^{-1} photons and reducing the total culture volume to 300 mL increased the three-day OD730 to two to three folds comparing to previous light and CO_2 limiting culture conditions(Figure 3.6 and 3.7).

To further tap the potential of limonene production in LglgC, a continuous growth strategy was employed. Under previous conditions, the strain showed a very limited limonene production after three days of cultivation because of fast cell death under light limiting conditions (Figure 2.1A). To overcome this growth bottleneck, one third of the old culture was replaced with fresh BG11 medium every day after day 3. Under this condition, both L1118 and LglgC were able to maintain a high limonene productivity (Figure 3.9). Compared to the previous 100µmol photons m^{-2} s^{-1} and 30°C culturing conditions, the 7-day limonene yield increased from 2.97mg/L to 7.44mg/L in L1118. The optimized growing condition were able to further improve the limonene production in cyanobacteria. More importantly, the limonene yield in LglgC reached a record yield at 14.9mg/L and it was one-fold increase compared to L1118. During the long-term production, limonene was produced at a maximum rate in the mutant strain LglgC and the fixed CO_2 was most directed toward limonene production at a high rate. This feather is ideal for industrial applications using high-density cyanobacteria as efficient cell factories.

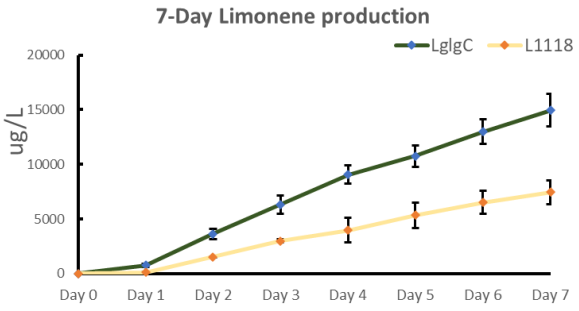

Figure 3.9 Cumulative limonene production in LglgC and L1118. The strains were cultured with 5% CO_2, 150µmol photons m^{-2} s^{-1} at 37°C in 300 mL BG 11 medium.

3.5 Conclusion

Cyanobacteria express the MEP pathway, *via* which they synthesize a wide variety of terpenoids(Pattanaik and Lindberg, 2015). However, the carbon flux to MEP pathway was estimated to be 4-5% of the total fixed carbons(Lindberg et al., 2010). In this study, we enhanced the terpene production in cyanobacteria by redirecting carbon flux from glycogen storage to MEP terpene synthesis, which resulted in a production yield of 2626±310µg/day/L limonene production from CO_2. The *glgC* mutant strain (LglgC) has lower total carbohydrate content and ~1-fold increase of limonene production compared to previously established limonene producing strain L1118. While a previous study reported that deficiency of glycogen storage would increase the ATP concentrations and delay the growth in cyanobacteria(Cano et al., 2018), we found there was a slight decrease of ATP in LglgC which could be explained by the consumption of ATP in MEP for limonene synthesis. However, cellular ATP regulations are still unclear and needs further studies. Moreover, overexpression of downstream enzyme *idi* did not increase the limonene production which indicates downstream force to drive the carbon flux into MEP is not the key limiting factor comparing to the competitions with other primary metabolites. With the optimization of culture conditions, the strains were able to grow in high density continuously and were able to produce the limonene at a record yield of 14.9mg/L. The metabolic engineering and cultivation strategies in this study could also be employed for other bioproducts production in photosynthetic systems.

CHAPTER IV

CREATING BIOLOGICAL ROUTE FOR LIGNIN VALORIZATION THROUGH THE ENHANCEMENT OF POLYHYDROXYALKANOATES (PHAS) SYNTHESIS

4.1 Summary

The biological valorization of the lignin polymer shows the potential to enable the profitability of biorefineries. In this study, a promising microbial lignin conversion route to polyhydroxyalkanoates (PHAs) had been created through employing process engineering design, systems biology analysis and metabolic engineering, and fermentation development. Results show that understanding the biological metabolic mechanisms of lignin degradation and conversion in ligninolytic strains facilitated the enhancement of PHAs biosynthesis and was also of importance for their eventual industrial application. First, an innovative fractionation strategy has been proposed by integrating sulfuric acid pretreatment with enzymatic and alkaline treatment to process lignin and tune its biological reactivity from process engineering concept. The improved biological reactivity was confirmed for microbial conversion to PHAs with the increased cell growth and PHAs accumulation using lignin as carbon sources. Second, the system biology analysis promoted further understanding the lignin degradation metabolism of *P. putida* KT2440 and guided the redesign the metabolic pathways to provide a renewable route for the production of PHAs. Third, the strategy of strain engineering was then developed by overexpressing enoyl-coenzyme A hydratase of β-oxidation pathway and poly(3-hydroxyalkanoic acid) synthase of PHAs synthesis pathway in *P. putida KT2440*, which significantly promoted the PHA synthesis using SSL as carbon sources and modify the monomer composition of PHAs to directly define its properties and applications. Finally, the fermentation strategies were exploited by optimizing

substrate type and concentration, nitrogen source, fermentation mode and parameter control. This strategy promoted the cell growth, lignin consumption and PHAs accumulation. Taken these together, a cell dry weight of 6.4 g/l was produced using SSL as carbon sources by *P. putida*, and a comparable level of PHAs concentration was obtained correspondingly. Overall, a promising microbial conversion route for lignin valorization have been created to improve the lignin utilization and PHAs yield, which holds promise for the production of valuable products from lignin derivatives and the improvement of lignin valorization profitability.

4.2 Introduction

Biorefineries have great potential to reduce fossil fuel usage and decrease greenhouse gas emission and eventually mitigate climate change by producing biofuel and renewable coproducts (Ragauskas et al., 2014a; Rinaldi et al., 2016). However, there are still many technical barriers in developing a sustainable biorefinery. As intrinsic biomass recalcitrance are result from lignin-carbohydrate complex (LCC), the inherent value of lignocellulosic biomass (LCB) is highly dependent on deconstructing and co-valorizing lignin polymer in a biorefinery, and not just exploiting more uniform carbohydrates(Ragauskas et al., 2014a; Rinaldi et al., 2016; Gillet et al., 2017). Lignin is the largest renewable aromatic carbon source and the second most abundant terrestrial polymer after cellulose(Ragauskas et al., 2014a; Liu et al., 2019c). The amount of lignin annually produced in the pulping industry can be estimated to be approximately 130 million tons(Abdelaziz et al., 2016; Rinaldi et al., 2016; Bajwa et al., 2019), most of which has been seen as a waste and primarily burnt for the generation of heating and electric power. Meanwhile, the generation of lignin waste will soar with the commercial scale-up of lignocellulosic biorefinery plant(Ragauskas et al., 2014a; Liu et al., 2019c). Therefore, lignin valorization has become an

urgent and necessary issue. Recently, biological processing of lignin to valuable products has been developed as an efficient approach for both lignin valorization and economically feasible biorefineries due to mild conditions employed and effective tools for specific product synthesis. Despite the significant progress, lignin remains as an untapped resource and biological lignin valorization is relatively unexplored because of the heterogeneity of diverse lignin substrates and the unsatisfactory performance of the bioconversion.

Some microorganisms have evolved powerful enzymatic toolbox, including laccase (Lac), lignin peroxidase (LiP), manganese peroxidase (MnP), versatile peroxidase (VP), as well as others, to broken down lignin polymers to smaller fragments or aromatic compounds(Fuchs et al., 2011; Beckham et al., 2016; Chen and Wan, 2017b). These aromatic oligomers can be further assimilated as carbon and energy sources by ligninolytic strains, through at least four known catabolic paradigms(Fuchs et al., 2011; Liu et al., 2019c). Lignin and its derived aromatics are catalyzed through upper pathways in bacteria, which act as a "biological funnel" to convert heterogeneous lignin derived molecules to central metabolites. Protocatechuic acid and catechol are typical intermediates for both extradiol and intradiol cleavages in bacteria, and these compounds are key nodes in aromatic degradation. The β-ketoadipate pathway is well documented and is one of catabolic paradigms for further degradation of these intermediates, by which protocatechuic acid and catechol can undergo ring cleavage and be further transformed into metabolites for the tricarboxylic acid cycle (central pathway), ultimately resulting in acetyl-CoA for targeted product synthesis.(Fuchs et al., 2011; Liu et al., 2019c)

Polyhydroxyalkanoates (PHAs) are synthesized and serves as energy storage compounds in numerous bacteria and they are generally accumulated as granules in the cytoplasm of cells(Zhao

et al., 2003; Chen, 2009). PHAs possess some promising properties, such as biodegradation, biocompatibility, low toxicity, which make them interesting to the fields of tissue engineering, drug delivery, nanomaterial, etc. *P. putida* is a metabolically versatile saprophytic soil bacterium that has been certified as a promising lignin metabolizers and a good candidate for PHA production due to high secretion of lignin-degrading enzymes and easy genetic engineering operation(Nelson et al., 2002). Despite the recent effort, the lignin conversion to PHAs by *P. putida* remains a significant challenge.

First, lignin substrates used as carbon source of fermentation are highly heterogonous due to its cross-linked structural complexity, hindering the utilization efficiency and make the engineering of conversion route more difficult. The lignin macromolecule is polymerized primarily from three monolignols, *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol, though different ether and C-C bonds, including β-O-4-aryl ether, β-β, β-5, 5-5, and α-O-4-aryl ether linkages. It also carries various functional groups, including phenolic hydroxyl, aliphatic hydroxyl, carbonyl, and benzyl alcohol groups. Most important, the relative proportions of the subunits, linkages and functional groups in lignin vary depending on LCB source types and fractionation employed(Abdelaziz et al., 2016; Liu et al., 2018b). These chemical properties contribute to the heterogeneity of lignin, define its processibility, and thus have a significant effect on its fractionation and affordable valorization.

Second, the chemistry and biological reactivity of technical lignin significantly depend on fractionation approaches employed, and in turn will affect the valorization efficiency. Various fractionation approaches toward lignin depolymerization and fractionation have been developed in biorefineries, which are generally grouped into three categories: (i) release the carbohydrates in

the liquid stream and retain most of the lignin polymer in the solid residue; (ii) extract lignin as a precipitate; (iii) dissolve lignin in the liquid stream as a depolymerized product mixture(Ragauskas et al., 2014a; Varman et al., 2016). These approaches alter the structure of LCB and yield a specific lignin stream with different chemistries and processibilities to define the biological valorization. (Liu et al., 2017; Liu et al., 2019b) To date the relatively poor solubility of lignin in fractionation solvents remains one of the biggest challenges in biological lignin valorization.

Third, ligninolytic bacteria should have substrate specificities and posse different capacity of lignin catabolism and product synthesis. The technical lignin with the broader distribution of the molecular weight and high content of hydrophilic groups has been confirmed to facilitate its biological processability(Lin et al., 2016b; Liu et al., 2017). Ligninolytic bacteria such as *P. putida* can simultaneously convert lignin-derived aromatic molecules and carbohydrate-derived products in complex mixtures like alkaline pretreated liquor (APL) to PHAs, but it exhibit substrate specificities on ferulic acid, *p*-coumaric acid, glucose, and acetate(Linger et al., 2014; Salvachua et al., 2015). Despite, few studies have been conducted to compare the microbial conversion of various lignin substrates and their catabolic mechanisms especially for PHA. Taken together, the relatively narrow substrate specificities elucidated to date for ligninolytic bacteria, coupled to the potentially broad distribution of structurally and chemically distinct lignin products from fractionation, prompted us to reveal the catabolic mechanisms for targeted product synthesis and thus guide us to design the potential alternative biological pathway of lignin valorization.

The aim of this study is to create the biological pathway for lignin valorization and elucidate the mechanism of the broad lignin substrate specificity for PHA synthesis by *P. putida*. Process

engineering design, systems biology and metabolic engineering, and fermentation development have been integrated to create a lignin biological valorization route. In detail, an effective fractionation approach of lignin was designed to improve the lignin depolymerization and biological processibility. The optimization of fermentation technologies has been then conducted using lignin as carbon source to enhance PHA yield. Systems biology analysis and genetic engineering were combined to improve the capacity of PHA synthesis by *P. putida* KT 2440. Taken these together, these approaches would gain a deeper understanding of the lignin metabolism and PHA synthesis using *P. putida* KT2440 by enabling it to survive on various lignin-derived substrates.

4.3 Material and methods

4.3.1 Fractionation process for lignin substrate preparation

Alkaline pretreated liquor (APL) was prepared from corn stover biomass by the fractionation using sodium hydroxide pretreatment. During the fractionation, 50 g corn stover (dry weight, dw) was treated in a 1.0 L screw bottle with solid loading percentage of 10% and 1% (w/w) sodium hydroxide was added. The sealed bottle was heated to 121 ℃ in a steam sterilizer (Amsco® LG 250) (Steris, USA) and maintained at 121 ℃ for 60 min. APL from the pretreated slurry was separated from insoluble solid fraction by filtration and collected for further analysis and use. The other fractionation employed sulfuric acid followed by sodium. In details, 50 g corn stover (dw) was treated in a 1.0 L screw bottle with solid loading percentage of 10% and 1% sulfuric acid, and then was heated to 121 °C and maintained for 30 min. The solid fraction was separated by filtration. Cellic CTec2 and HTec2 was used for enzymatically hydrolyzation of the solid fraction.

Enzymatic hydrolysis was carried out at 50 °C for 168 h in 0.05 M citrate buffer (pH 4.8), supplied with a 10 FPU/g solid CTec2 loading. HTec2 was added in a volumetric ratio 10:1 of CTec2:HTec2. The insoluble hydrolyzed residues containing lignin and residual sugar was separated and then treated by 1% sodium hydroxide at 121 °C for 60 min. The liquid stream containing soluble lignin was separated by filtration and named as sulfuric acid and sodium hydroxide treated liquor (SSL). Kraft lignin (KL) purchased from Sigma Aldrich (USA). These lignin substrates were used as the carbon source in PHA fermentation and components were provided in Table 4.1.

Components (g/l)	KL	APL	CEL
Acid insoluble lignin (AIL)	44.1 (1.8)	10.4 (1.5)	17.6 (1.3)
Acid soluble lignin (ASL)	0.3 (0)	1.7 (0.1)	1.9 (0)
Glucose	0.3 (0)	4.2 (0.1)	9.9 (0.2)
Xylose	0.2 (0)	3.4 (0)	2.6 (0.1)
Others	3.5 (0.1)	25.5 (1.2)	26.2 (0.7)
Total SSC	48.4 (1.9)	45.1 (2.1)	58.2 (2.4)

Table 4.1 Components of different lignin substrates used as carbon sources by *P. putida* KT2440 for the production of PHAs. Three different lignin liquors, Kraft lignin (KL), alkaline pretreated liquor (APL), and chemical-enzymatic treated liquor (CEL) were used as carbon sources for PHA fermentation; SSC represents soluble substrate concentration; All data in the table are mean values of duplicate experiments; standard deviations are shown in parentheses.

4.3.2 Strains and seed culture preparation

Pseudomonas putida KT2440 wild type was purchased from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany). For seed culture preparation, single colonies of *P. putida* KT2440 were inoculated in 10 ml LB medium at 28 °C for 12 h. *P. putida KT2440* cultures was then inoculated in 100 ml of seed medium at 28 °C with

for 24 h to reach an OD_{600}4.0. The seed medium was prepared as follows. 100 mL seed medium contains 10 g/L glucose, 10 mL sterilized 10 × Basal salts, and 1 mL sterilized 100× Goodies mixture. Basal salts contain 30 g/L KH_2PO_4 and 60 g/L $NaHPO_4$. 1L 100× Goodies mixture contains 500 ml stock salt solution, 3.009 g $MgSO_4$, and 25 mL 1% $FeSO_4$.

4.3.3 Plasmid and strain construction

For construction of engineered *P.putida* KT2440 strain, *phaC*, encodes poly(3-hydroxyalkanoic acid) synthase an *phaJ* enoyl-coenzyme A (CoA) hydratase was were amplified from *P. putida* KT2440 genomic DNA using Phusion High-Fidelity DNA Polymerase (NEB, Ipswich, MA). These gene fragments were then ligated into pPROBE(Gm^R) plasmid (obtained from Dr. Elizabeth Pierson's lab) using Gibson Assembly (NEB, Ipswich, MA). The 100ng of the plasmid was transformed into *P. putida* KT2440 wild type strain by electroporation in electrocuvettes using the following settings: 2.5kV/cmm 600 Ω, and 50 μF. Pulsed cells were immediately diluted with 1 mL of LB and then were incubated at 28°C for 4 h before they were plated on gentamycin selection medium(50 μg/L)

TF-Cra, encodes catabolite repressor activator, *TF-TET*, encodes Tet repressor protein, TetR, *TF-RNAP*, encodes RNA polymerase binding protein, *phaG*, encodes 3-hydroxydecanoyl-ACP:CoA transacylase, *fadD*, encodes fatty acyl coenzyme A synthetase, *pcaF-I*, encodes beta-ketoadipyl-CoA thiolase were amplified from *P. putida* KT2440 genomic DNA using Phusion High-Fidelity DNA Polymerase (NEB, Ipswich, MA). The genes were then ligated into pPROBE (Km^R) plasmid (obtained from Dr. Elizabeth Pierson's lab) using Gibson Assembly (NEB, Ipswich, MA) following the manufacture's instruction. The 100ng of each plasmid was transformed into the

engineered *P. putida* KT2440 strain by electroporation. Pulsed cells were immediately diluted with 1 mL of LB and then were incubated at 28°C for 4 h before they were plated on selection medium with 50 μg/L kanamycin and 10 μg/L gentamycin.

4.3.4 Polyhydroxyalkanoates (PHAs) fermentation

The lignin substrate was used as carbon source in fermentation for polyhydroxyalkanoates (PHAs) production by *P. putida* KT2440. For the preparation of lignin medium, pH of lignin liquor was adjusted to 7.0 using 1.0 M H_2SO_4 and then dissolved to desired soluble substrate concentrations (SSC) with addition of sterilized ddH_2O. 100 mL lignin medium contains 10 ml sterilized 10× Basal salts and 1 mL sterilized 100× Goodies mixture. The medium was then transferred into 250-mL Erlenmeyer flasks with a working volume of 100 ml. *P. putida* KT2440 were collected by centrifuging the seed culture at 4,000 rpm for 10 min. The cell pellets were then inoculated in toe lignin medium. PHAs fermentation was conducted at pH 7.0, 28°C, and 200 rpm for 18 h. Fed-batch fermentation employed in fermentation was conducted as follows. Fed-batch fermentation cycle 1 was carried out a 20 g/l SSC with 1.0 g/l NH_4Cl to facilitate the cell growth, while fed-batch fermentation cycle 2 was conducted at 40 g/l SSC to improve the formation of PHA (Table 4.2).

Experiment	Strains	Lignin cosubstrate	Fermentation conditions
Case 1	*P. putida* KT2440	KL	Fed-batch fermentation cycle 1, 20 g/l SSC; Fed-batch fermentation cycle 2, 40 g/l SSC; pH 7.0, 28°C, and 200 rpm for 18 h.
Case 2	*P. putida* KT2440	APL	
Case 3	Engineered *P. putida* KT2440	APL	
Case 4	*P. putida* KT2440	APL with glucose	
Case 5	*P. putida* KT2440	CEL	
Case 6	Engineered *P. putida* KT2440	CEL	

Table 4.2 Fermentation strategies for the production of PHAs from lignin substrates. Note: KL represents Kraft lignin; APL represents alkaline pretreated liquor; CEL represents chemical-enzymatic treated liquor; SSC represent soluble substrate concentration, g/l.

4.3.5 PHA analysis using GC-MS method

The fermented cultures were centrifuged at 8,000 rpm for 10 min at 4 °C. The pellets were collected and washed twice using 20 mL 0.9% NaCl solution, and then lyophilized at -55 °C with 0.05 mbar for 24 h using a 4.5 L Benchtop Freeze-Dry System (Labconco Corp., USA). The lyophilized cell biomass was weighed to calculate the cell dry weight. For PHA polymer analysis, 5-10 mg freeze-dried cells was subjected to methanolysis with addition of 2 ml of 15% (v/v) sulfuric acid in methanol and 2 ml chloroform. The samples were incubated at 100 °C for 140 min in a closed glass vial. The PHA units inside the cells are to methyl esters during the methanolysis process. After cooling at room temperature, 1 ml of ddH_2O was added into the mixture and vortexed for 30 s. The lower chloroform organic phase containing the methyl esters was then separated and collected by centrifugation at 4000 rpm for 10 min.

GC-MS method was used to analyze the PHA polymer. In detail, 0.5 ml lower chloroform phase containing methyl ester was mixed with 0.5 ml 0.1% caprylic acid as the internal standard. Then the mixture was filtered with 0.22-μm polytetrafluorethylene (PTFE) membrane before loaded into

GC-MS. GC-MS analysis was performed on a GC-MS-QP2010SE (Shimadzu Scientific Instruments, Inc.) with a Shimadzu SH-Rxi-5Sil column (30 m × 250 µm × 0.25 µm). Helium was used as carrier gas at a flow rate of 1.0 ml/min. The temperature was maintained at 50°C for 3 min and then increased to 300°C at a rate of 10°C/min. Mass spectra were recorded by a 70-eV electron beam at an ionization current of 40 µA. Mass spectral peak quantification was performed using GCMSsolution software Ver. 2.6. All experiments were performed in 3 biological replicates.

4.3.6 Proteomic analysis

100 ml of culture cells were harvest by centrifugation with 8000 rpm for 10 min at 4°C and washes twice with 50 mL phosphate buffered saline (PBS) solution (pH 7.4). Pellets were then suspended in 2 ml 50 mM TrisHCl buffer (pH 7.6) with addition of 20 µl protease inhibitor cocktail (Sigma Aldrich, USA). Cells were then lysed by tip-probe sonication using 20 cycles of 15s and centrifuged at 12000 rpm for 10 min at 4°C. Supernatants were harvested for proteomic analysis. The protein concentration was measured through Pierce Bradford Protein Assay (Thermo Fisher) with standards.

MudPIT-based proteomic analysis and follow-up data processing were conducted as follows: 100 µg proteins of each sample were denatured under conditions of 8 M urea and 5 mM dithiothreitol (DTT) and incubated at 37 °C for 1 hour. The denatured proteins were then treated with iodoacetamide in a final concentration of 15 mM iodoacetamide for 15 min at room temperature in dark. The samples were diluted 3 times to a final concentration of 2 M urea. 1 µg trypsin (Promega, Medison, WI) was added into each sample and then incubated at 37°C overnight. The digested peptide was then desalted through a sep-pak C18 column (Waters, Milford, MA), and

then dried by a SpeedVac (Brinkmann Instruments, Westbury, NY). The digest peptides were dissolved in 300 µl 0.1% formic acid and centrifuged with 12000 rpm for 10 min at 4°C, and then harvested for proteomic analysis.

The desalted samples were loaded into a self-pack 150 um capillary column with 3 cm SCX (strong cation exchange) and 5 cm C18 beads within a pressure cell and followed by two dimensional peptide separation within a 100um C18 capillary column.(Zhang et al., 2013) The peptide fractions were analyzed by LTQ Velos pro mass spectrometer (Thermo Finnegan, San Jose, CA). The 5 most abundant peaks were subjected for MS/MS analysis. The mass spectra were extracted from raw files to convert into a MS2 file and then searched against a composite database of *Pseudomonas putida* KT2440, common contaminants, and reversed sequences using ProLuCID algorithm. The sqt file was filtered through PatternLab4 for data analysis with the parameters, including Delta mass PPM 30.0, DeltaCN 0.0010, acceptable false discovery rate: sepctra 3.00%, peptide 2.00%, protein 1.00%.

4.3.7 Composition analysis

Composition analysis of the lignin liquor produced from each fractionation was conducted following the Laboratory Analysis Protocol (LAP) of the National Renewable Energy Laboratory (NREL), Golden, CO, USA. The sugars analysis were carried on an Ultimate 3000 HPLC System (Thermo Scientific, USA) equipped with an Aminex HPX-87P carbohydrate analysis column (Bio-Rad Laboratories, CA) and a refractive index detector using HPLC grade water as the mobile

phase at a flow rate of 0.6 ml/min. Error bars in the tables and figures represented the standard deviation of the replicates.

4.4 Results and discussion

4.4.1 Process engineering design of biological lignin valorization to PHA

Lignin is the largest renewable aromatic carbon source comprised of phenylpropanoid monomers on Earth. Various types of technical lignin have been produced in lignocellulosic biorefinery and wood chemical pulping, and the amount of lignin will dramatically increase with the intensive development of biorefineries as well as paper pulp and paper industry. Among these technical lignins, Kraft lignin is one of most abundant and inexpensive aromatic resources available in the market, while alkaline pretreated lignin (APL) is a potential feedstock for microbial conversion due to its high solubility and reactivity(Linger et al., 2014; Bajwa et al., 2019). However, technical lignin is typically underused as a feedstock in a biorefinery to produce chemicals and materials because it carries on specific chemistries and structures compared with native lignin and its inherent heterogeneity and recalcitrance make it difficult to selectively valorize. Therefore, novel lignin upgrading strategies are needed to produce commodities effectively and make a profitable biorefinery.

The ability of ligninolytic *P. putida* KT2440 to convert lignin had hence been examined by evaluating cell growth and the production of a biodegradable polymer, PHAs. Different types of technical lignin substrate were first used as carbon sources (Figure 4.1). Results suggested that the fermentation performance significantly depended on the lignin sources used, and *P. putida*

KT2440 had strong capacity to consume lignin selectively for PHA synthesis. Kraft lignin is one of most abundant and inexpensive aromatic resources available in the market(Linger et al., 2014; Bajwa et al., 2019). However, *P. putida* KT2440 hardly grew on Kraft lignin medium to accumulate PHAs possibly due to its high molecular weight and sulfur content (1.0-3.0%) after extraction process, which inhibited the strain growth and thus reduced its biological reactivity(Wei et al., 2015; Bajwa et al., 2019; Kumar et al., 2019). Alkaline pretreated lignin (APL) from biorefineries is a potential feedstock for microbial conversion due to its high solubility and reactivity(Linger et al., 2014; Bajwa et al., 2019). A lignin-rich stream was prepared by alkaline pretreatment from corn stover biomass and used as carbon source for PHA synthesis. Results showed that *P. putida* KT2440 grew well on APL medium and accumulated PHAs to 0.33 g/l, which was 11 times higher than that on KF medium (Figure 4.1). *P. putida* KT2440 have strong capacity to grow on APL medium likely due to its lower molecular weight, higher aromatic compound content, and better solubility.

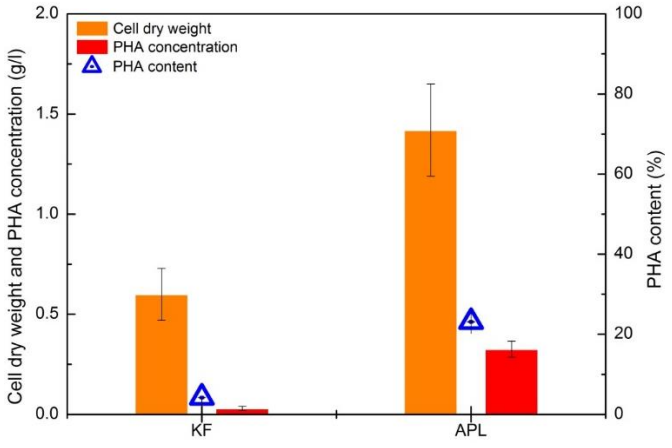

Figure 4.1 Microbial fermentation of Kraft lignin (KF) and alkaline pretreated liquor (APL) for the production of polyhydroxyalkanoates (PHAs) by wild-type *P. putida* KT2440. Fermentation conditions: inoculum OD1.0, 10 g/l of soluble substrate concentration (SSC), 1.0 g/l NH_4Cl, pH 7.0, 28°C, and 200 rpm.

To further increase the accumulation of PHAs in *P. putida* KT2440 using lignin as carbon sources, three aspects regarding to fractionation and depolymerization of lignin, aromatic metabolism and PHA synthesis, and fermentation technology have been investigated from the 'upstream' to 'downstream' of lignin bioconversion (Figures 4.2 and 4.3).

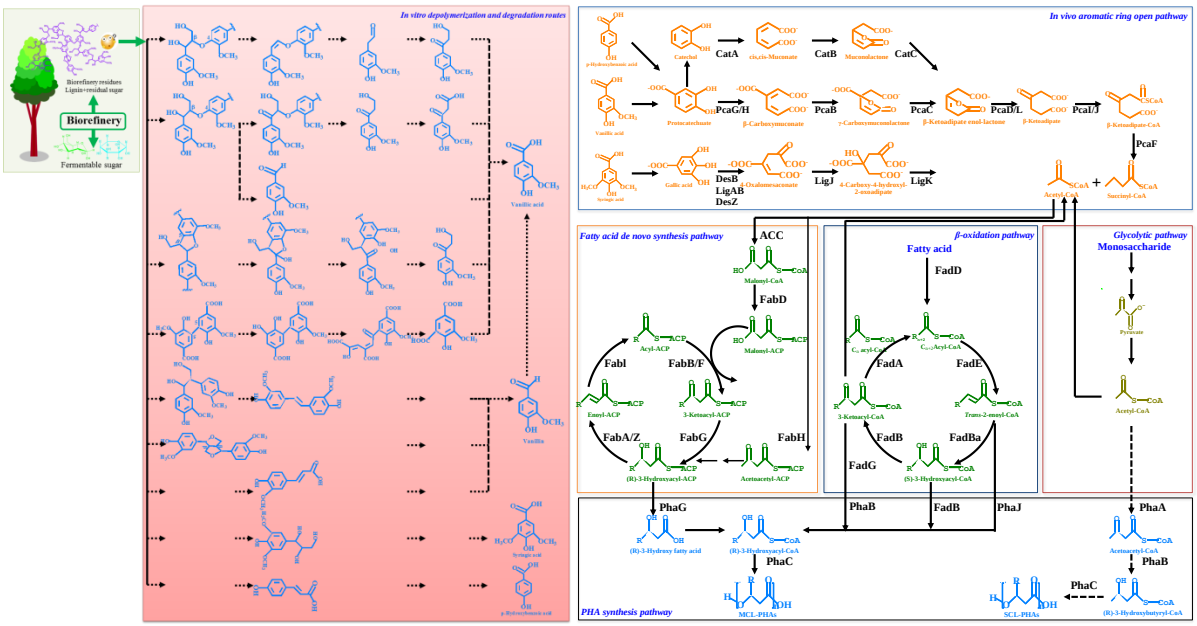

Figure 4.2 Potential route of biological lignin valorization to produce PHAs. Lignin deploymerization with alkaline and biological processing (A) and aromatic degradation and catabolic pathways in some ligninolytic bacteria (B). Key enzymes, genes and metabolic intermediates are provided and highlighted.

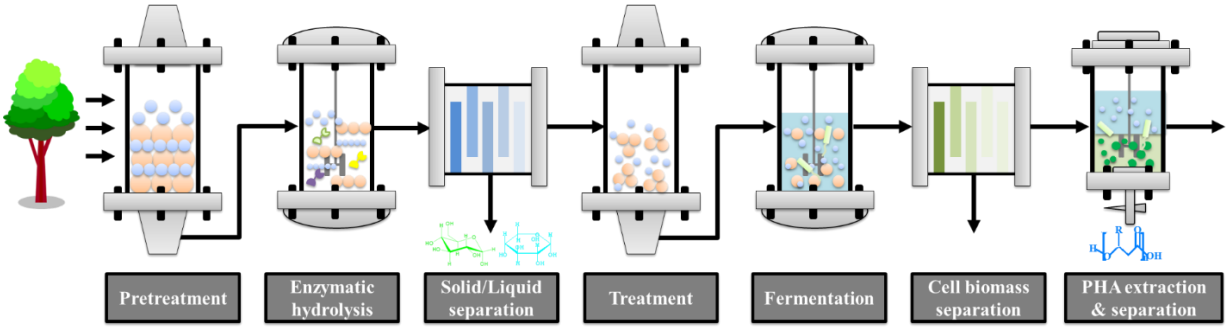

Figure 4.3 Schematic diagram comprising the entire process for the production of PHAs using waste lignin from corn stover biomass by *P. putida* KT2440 in a biorefinery concept. Pretreatment used in this biorefinery scenario is dilute sulfuric acid; enzymatic hydrolysis and alkaline treatment was further employed to depolymerize the hydrolyzed solid residue; Organic solvent extraction method was employed to extract the PHAs from cell biomass.

Generally, LCC and lignin recalcitrance hinder the fractionation and depolarization of lignin, and thus its microbial conversion(Chen and Liu, 2015; Liu et al., 2015; Kim et al., 2017; Tian et al., 2017). Above results (Figure 4.1) indicated that lignin biological valorization significantly depended on lignin source used. Generally, the biological reactivity of lignin could be improved to facilitate its conversion by employing the fractionation to decrease the lignin molecular weight, break down the linkages, and enrich aromatic compounds, which have been confirmed in the previous studies(Li et al., 2007; Liu et al., 2017). Therefore, an innovative lignin processing has been designed from process engineering concept of biorefinery to depolymerize lignin polymer, and tune its biological reactivity, and produce a suitable lignin stream to facilitate microbial conversion (Figures 4.2 and 4.3). In detail, dilute sulfuric acid pretreatment was employed to deconstruct LCC structure of corn stover by dissolving hemicellulose, depolymerizing lignin, and removing non-structural components. The insoluble solid fraction was then enzymatically hydrolyzed to release fermentation sugars and modify lignin. After removing the hydrolysates, enzymatic residues were further treated by alkaline treatment to maximize the depolymerization of lignin by cleaving lignin linkages, reducing molecular weight, and generating aromatic compounds.

To evaluate the biological reactivity of lignin produced from this innovative processing strategy, microbial conversion of lignin to PHAs had been hence examined by ligninolytic *P. putida* KT2440 (Figure 4.4). Results showed that this fractionation configuration could significantly improve the biological reactivity of lignin to facilitate the growth of *P. putida* KT2440 and the accumulation of PHAs. The cell dry weight and PHA concentration using CEL was carbon sources reached to 1.8 g/l and 0.73 g/l, respectively, which was 3 and 28 times higher than that on KF

medium. The improved biological reactivity of lignin was supported by the results of molecular weight analysis. As compared with corn stover native lignin, CEL possessed lower number-average molecular weight (Mn) and weight-average molecular weight (Mw), which decreased by 77.8% and 81.5%, respectively. The M_n and M_w of CEL also significantly decreased as compared with that of KF and APL. Results suggested that this innovative configuration designed in the present study depolymerized the lignin severely to produce the lignin with small molecular weight, which indicated better biological reactivity. As a result, the improved biological reactivity promoted the microbial consumption and conversion of lignin to synthesis PHAs.

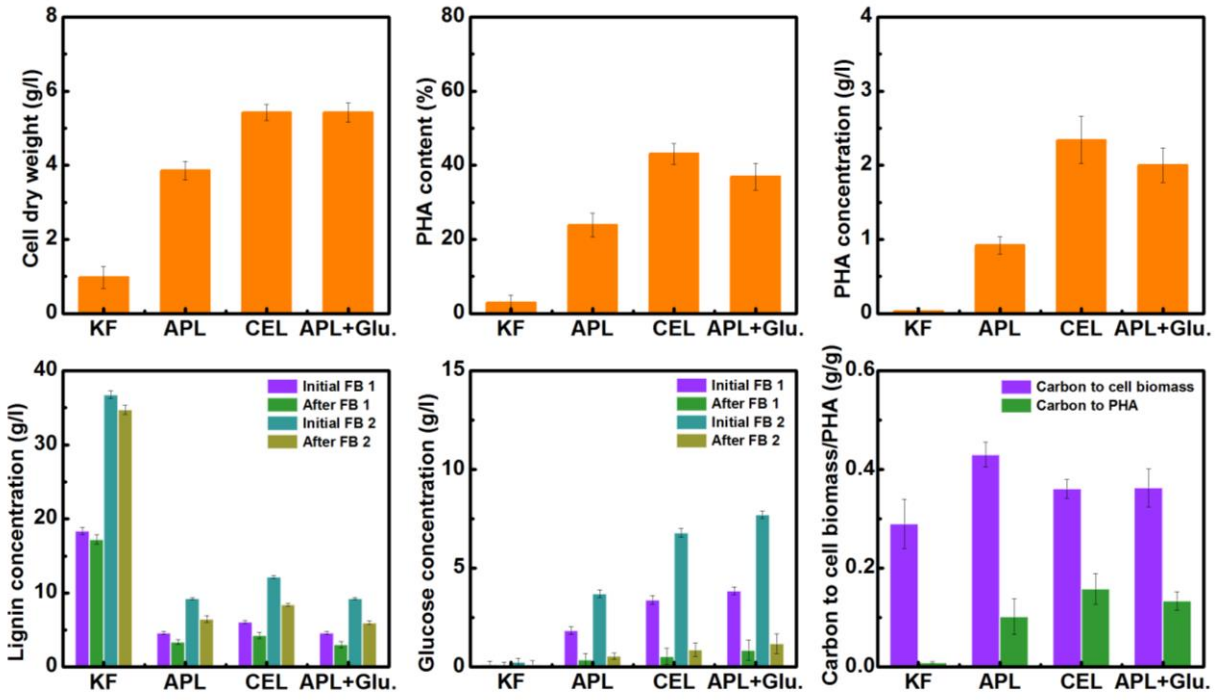

Figure 4.4 Fermentation for the production of polyhydroxyalkanoates (PHAs) by wild-type *P. putida* KT2440 grown on various lignin cosubstrates. KL represents Kraft lignin; APL represents alkaline pretreated liquor; CEL represents chemical-enzymatic treated liquor. Glu represents glucose.

4.4.2 Fermentation strategies for enhancing PHA accumulation

As PHAs comprise a large class of polyesters synthesized from various monomers, the properties of PHAs and its accumulation in bacteria will be not only influenced by the strain and substrate used, but also significantly depended on fermentation conditions and technologies employed. The fermentation strategies were thus exploited to improve the accumulation of PHAs in *P. putida* KT2440 by optimizing fermentation mode, substrate type and concentration, and nitrogen source.

As an effective operational technique in biotechnological processes, a fed-batch fermentation scenario was developed to reduce the potential inhibition effects of substrate and improve the cell growth and the PHA titer. A high substrate concentration is also necessary to make a profitable fermentation by producing a high titer of target product and reducing the separation cost. However, a high SSC of lignin stream generated from the fractionation may contain more degradation products to inhibit cell growth due to the inhibition effect and rheology behavior change(Liu et al., 2014; Jonsson and Martin, 2016; Rasmussen et al., 2017). As suggested by Figure 4.1, technical lignin prepared from different processes could carry specific chemistries and hence biological reactivities. Four types of lignin cosubstrates were used as carbon sources for the production of PHA in fed-batch fermentation at high SSC to investigate the selective capacity of lignin consumption by *P. putida* KT2440. Besides, bacteria usually require a certain deficiency condition to synthesize PHAs, such as the excess supply of carbon sources with a threshold concentration of nitrogen source(Faccin et al., 2009; La Rosa et al., 2014). Thus, sufficient nitrogen source was fed to the medium in cycle 1 of fed-batch fermentation to multiply cell biomass and low nitrogen source was maintained in cycle 2 of fed-batch fermentation to force the accumulation of PHA in *P. putida* KT2440.

As expected, the microbial conversion performance of lignin definitely depended on the fermentation strategies employed and these combined fermentation scenarios obviously promoted the cell growth, the lignin utilization, and the PHA accumulation (Figure 4.4).

First, results showed that the cell growth and PHA accumulation at high SSC was determined by the lignin substrate type used, which is consistent with the results at low SSC (Figure 4.1). *P. putida* KT2440 exhibited the selective capacity of lignin consumption and PHAs accumulation when it grew on various lignin mediums. Compared with KF and APL the substrate of CEL produced the highest levels of cell dry weight and PHA accumulation in *P. putida* KT 2440, corresponding to 5.3 g/l and 43%, respectively. The PHA concentration generating from CEL reached to 2.4 g/L, which was 2.5 times higher than that from APL. The results suggested that CEL produced from the innovative fractionation of lignin processing could also have better biological reactivity at high SSC to promote the cell growth and PHAs formation as compared with KF and APL.

Second, the preformation of PHA fermentation was also closely related to the composition of lignin cosubstrates. Figure 4.4 showed that CEL contained more lignin and residual sugars as compared with APL. Results suggested that besides the properties of lignin itself, residual sugar in CEL could improve the fermentation performance by promoting the lignin consumption and facilitating the cell growth and PHA accumulation. To validate the effects of residual sugar on fermentation, limited glucose was fed to APL medium at high SSC. Results show that the addition of glucose in APL obviously facilitated lignin consumption, and promoted the cell growth and PHA accumulation (Figure 4.4 and Table 4.1). However, APL medium with the addition of

glucose exhibited lower the cell dry weight and PHA concentration compared with CEL medium, indicating that CEL could have high biological reactivity of lignin.

Third, fed-batch fermentation with two cycles was an effective scenario for the production of PHAs using lignin medium at high SSC in *P. putida* KT2440 (Figures 4.1 and 4.4). Cell dry weight from fed-batch fermentation mode using APL and CEL was 2.7 and 2.9 times higher than that from batch mode, respectively. Correspondingly, PHA concentration from fed-batch mode using APL and CEL was 2.8 and 3.2 times as high as that from batch mode, respectively. As a result, PHA yield reached to 0.10 and 0.16 g/g carbon source when cell grew on APL and CEL medium, respectively.

The monomer constitution will define the properties of PHAs and thus their applications. The monomer constitution of PHAs was completely different among these lignin mediums employed (Figure 4.6A). Results showed that the HHx, HO, and HDD fraction in PHAs produced from APL medium were 35.2%, 29.2% and 24.3%, respectively. With the addition of glucose in APL medium, the HHx, HO, and HDD fraction changed to 19.9%, 59.8% and 13.4%, respectively. However, CEL medium possessed 21.2% HHx, 64.9% HO, and 7.2% HDD fraction in PHAs. Results suggested that the types and compositions of lignin cosubstrates significantly influenced the monomers and thus the properties of PHAs, indicating that the biological reactivity of lignin is a key factor to determine the formation and property of PHAs. Overall, by employing combined fermentation scenarios, comparable levels of cell dry weight, PHA concentration, and PHA yield have been obtained especially when CEL was used as carbon sources by *P. putida* KT2440.

To further enhance the PHA biosynthesis from lignin derivatives, genetic engineering strategy was developed by improving the expression levels of enoyl-coenzyme A (CoA) hydratase (encoded by phaJ) of β-oxidation pathway and poly(3-hydroxyalkanoic acid) synthase (encoded by *phaC*) of PHA synthesis pathway in *P. putida* KT244 (Figures 4.2 and 4.5). Results showed that the engineered *P. putida* KT244 significantly promoted the PHA synthesis. As compared with wild type strain, engineered *P. putida* KT2440 produced higher PHA content, which increased by 58% and 29% using APL and CEL as carbon sources, respectively. Correspondingly, it produced more than 67% and 45% PHA concentration on APL and CEL medium, respectively. The highest level of PHA concentration was 3.4 g/l, achieving by engineered *P. putida* KT244 with fed-batch fermentation mode using CEL as carbon sources (Figure 4.5). As expected, residual sugar had been completely consumed in APL and CEL, while more CEL had been consumed by engineered *P. putida* KT2440. Meanwhile, higher PHA yield was obtained by engineered *P. putida* KT2440 grown on APL and CEL medium compared with wild type strain.

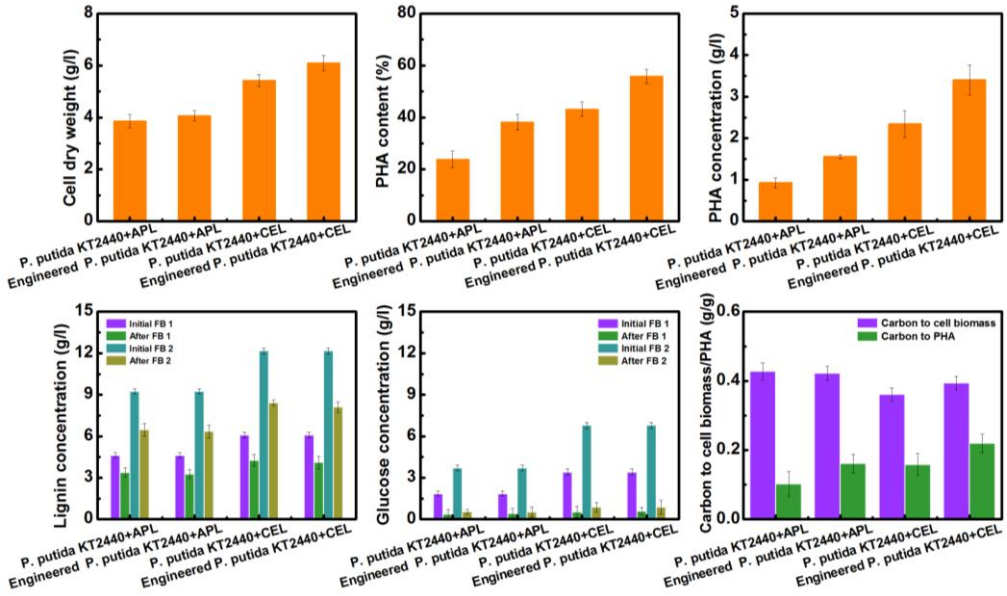

Figure 4.5 Improved fermentation for the production of PHAs by wild-type or engineered *P. putida* KT2440 grown on various lignin cosubstrate. APL represents alkaline pretreated liquor; CEL represents chemical-enzymatic treated liquor.

The PHA monomer composition was completely different between wild type and engineered *P. putida* KT2440 when these strains grew on APL and CEL (Figure 4.6). The HHx, HO, and HDD fraction in PHA produced by wild type strain was 35.2%, 29.2% and 24.3% from APL medium and 21.2%, 64.9% and 7.21% from CEL medium, respectively. Results supported that the monomer composition of PHA depended on the lignin type used and CEL facilitated the formation of PHAs containing more HOs. The HHx, HO, and HDD fraction in PHA produced by engineered *P. putida* KT2440 was 27.9%, 51.0% and 14.7% from APL and 15.6%, 67.2% and 10.8%, respectively. PHAs produced by engineered *P. putida* KT2440 contained much higher content of HOs compared with wild type one. As the monomer composition could determine the properties of PHAs, the genetic engineering strategy and the lignin medium option should have potential to alter the synthesis of PHAs and thus its properties and applications.

Overall, the engineering of *P. putida* KT2440 rechanneled the β-oxidation pathway and PHA synthesis pathway to enhance the lignin consumption and PHA synthesis. A higher level of PHA concentration was achieved using CEL as carbon sources in fed-batch fermentation.

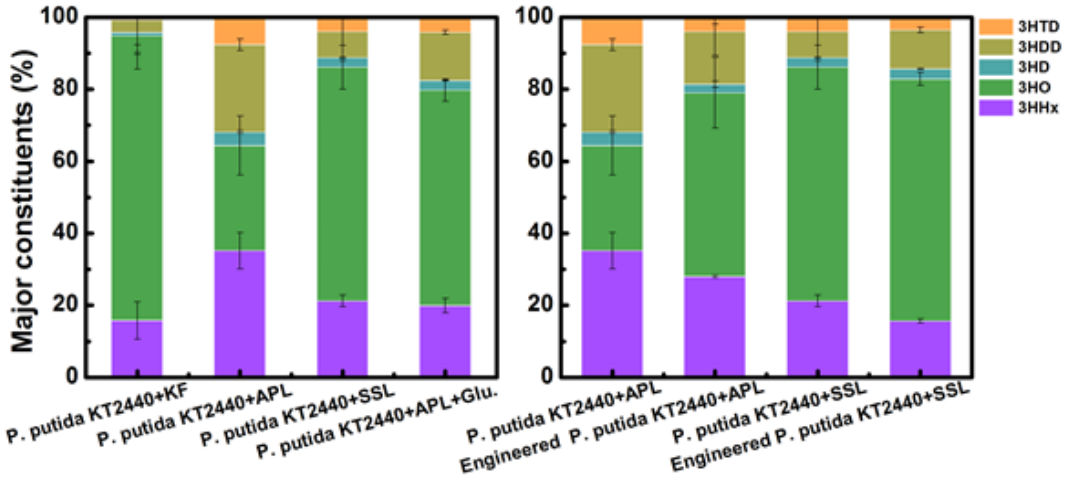

Figure 4.6 Monomer component of polyhydroxyalkanoates (PHAs) produced from *P. putida* KT2440 during fermentation on various lignin cosubstrate. 3HHx represents 3-hydroxyhexanoate; 3HO represents 3-hydroxyoctanoate; 3HD represents 3-hydroxydecanoate; 3HDD represents 3-hydroxydodecanoate; 3HTD represents 3-hydroxytetradecanoate; 3HHD, 3-Hydroxyhexadecanoate; ND, not detected.

4.4.3 Proteomics analysis reveals the mechanism of lignin metabolism and PHA synthesis

To further understand the mechanisms of lignin metabolism and PHA synthesis, proteomics profile of *P. putida* KT2440 grown on various substrates (glucose, ferulic acid, APL or CEL) were compared (Figure 4.7). Results showed that the differentially expressed proteins were mostly those involved in aromatic ring open pathway, β-oxidation pathway, fatty acid de novo synthesis pathway, and PHA synthesis pathway (Figure 4.7).

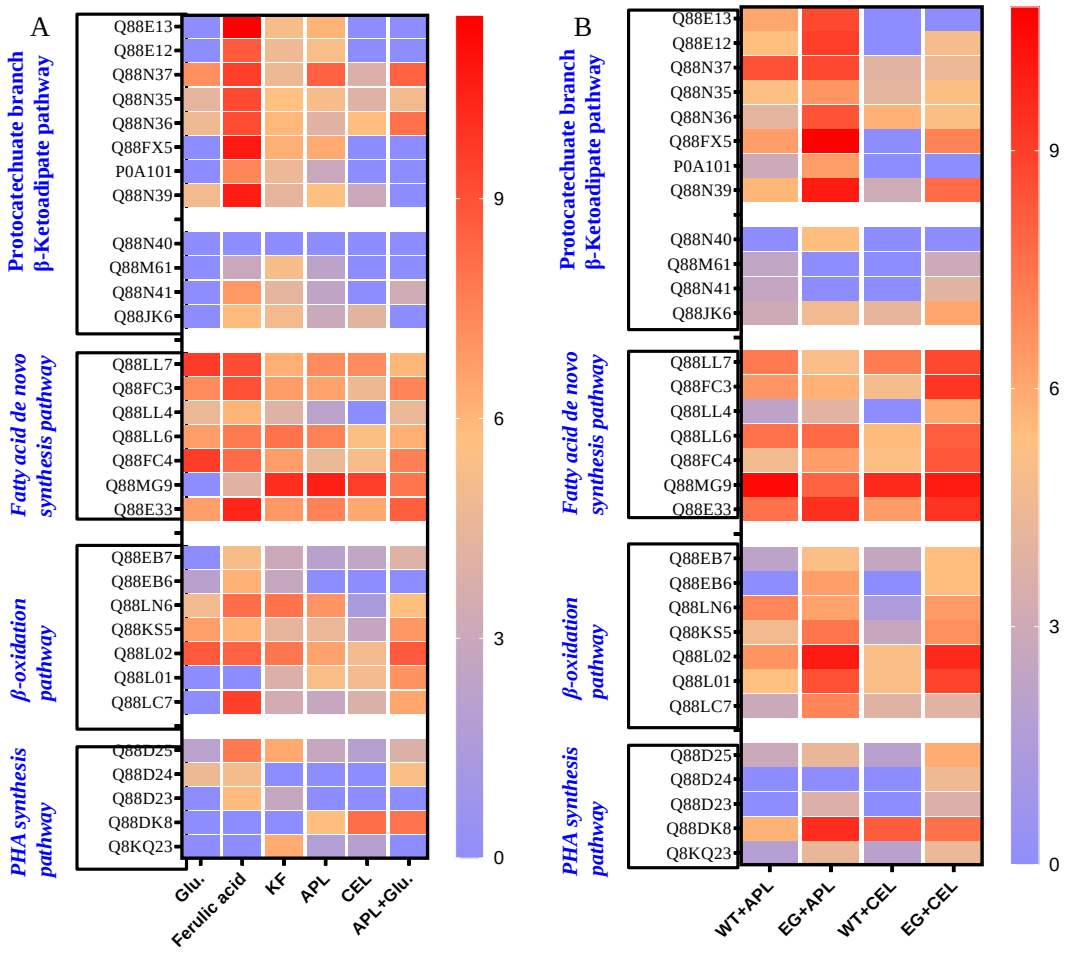

Figure 4.7 Expression levels of proteins *P. putida* KT2440 grown on different carbon sources. KL represents Kraft lignin; APL represents alkaline pretreated liquor; CEL represents chemical-enzymatic treated liquor. Glu. represents the glucose.

P. putida KT2440 can metabolize diverse lignin substrates and exhibited the specificity of lignin utilization. Besides, even the growth of this strain was observed to be good on these mediums, the PHA titer and monomer component varied with the lignin medium used (Figures 4.1, 4.4, and 4.5). These phenomena could be elaborated by the proteomics analysis, which provided the expression levels of proteins related with lignin metabolism and PHA synthesis (Figure 4.7). As expected, when grown on glucose medium only, *P. putida* KT2440 exhibited low expression levels of proteins especially these involved in aromatic ring open pathway, fatty acid de novo synthesis

pathway, and PHA synthesis pathway, suggesting the low capacity of PHA accumulation (Figure 4.7A).

When grown on ferulic acid medium only, *P. putida* KT2440 presented high expression levels of proteins involved in aromatic ring open pathway, β-oxidation pathway, fatty acid de novo synthesis pathway, and PHA synthesis pathway compared with that on glucose medium only (Figure 4.7B). Abundance of proteins involved in aromatic ring open pathway, especially protocatechuate 3,4-dioxygenase alpha chain, 3-oxoadipate CoA-transferase subunit A, and β-ketoadipyl-CoA thiolase, indicated their role in the degradation of the aromatic ring of ferulic acid. As a result, ferulic acid could be sufficiently transformed into acetyl-CoA, a key intermediate metabolites of lignin derivatives for PHA synthesis. The proteomics analysis revealed that enzymes participating in the fatty acid de novo synthesis pathway displayed high expression levels. Most of proteins involved in β-oxidation pathway presented also a high expression level, which may facilitate the PHA synthesis. the proteomics information also showed that the enzymes related to PHA synthesis pathway from ferulic acid had a higher high expression level as compared with that from glucose medium, which may explain the high PHA yield from ferulic acid.

P. putida KT2440 grown on Kraft lignin medium possessed obviously lower expression levels of proteins related to aromatic ring open pathway, *β*-oxidation pathway, and PHA synthesis pathway than that on ferulic acid medium, which supported less lignin degradation and PHA accumulation (Figure 4.7A). High concentration of Kraft lignin in medium may be acted as a high-stress environment to induce the expression of genes related to the degradation of aromatics.

The proteomics profiles showed that the expression levels of most proteins related to aromatic ring open pathway, especially protocatechuate 3,4-dioxygenase alpha chain, 3-carboxy-cis,cis-muconate cycloisomerase, and β-ketoadipyl-CoA thiolase, were much higher when *P. putida* KT2440 was grown on APL compared with Kraft lignin (Figure 4.7A). Abundance of proteins in fatty acid de novo synthesis pathway could indicate the better cell growth on APL medium. Interestingly, the expression levels of Protocatechuate 3,4-dioxygenase alpha chain, protocatechuate 3,4-dioxygenase beta chain, and 3-oxoadipate CoA-transferase subunit B from CEL were not as high as that from Kraft lignin and APL. It is likely due to that the presence of residual sugar in CEL medium may reduce the stress environment of lignin. Results showed that the expression levels of proteins related to fatty acid de novo synthesis pathway on CEL were similar to those on APL. Besides, abundance of proteins involved in PHA synthesis pathway on CEL supported the improved PHA accumulation. Therefore, the proteomics profiles revealed the mechanism of aromatic metabolism and PHA synthesis.

Figure 4.5B showed the proteomics profile of wild type and engineered *P. putida* KT2440 grown on APL and CEL medium, respectively. Abundance of proteins involved in PHA synthesis especially enoyl-coenzyme A (CoA) hydratase and poly (3-hydroxyalkanoic acid) synthase were obtained from engineered *P. putida* KT2440. Interestingly, the expression levels of most proteins related to aromatic ring open pathway, fatty acid de novo synthesis pathway, and β-oxidation pathway had been enhanced in engineered *P. putida* KT2440. Results suggested that the engineered *P. putida* KT2440 cell grew well on CEL medium possibly due to that it could catalyze more lignin derivatives and aromatics, produce more intermediate metabolites, and thus enhance

the PHA accumulation by expressing the related proteins, which was consistent with the results of cell growth and PHA yield.

4.4.4 Enhanced PHA synthesis with the engineering of ligninolytic *P. putida* KT2440

The information provided by the proteomics analysis could help to further guide the engineering design of *P. putida* KT2440 to improve lignin metabolism and PHA synthesis. Besides, based on the proteomics analysis genetic engineering could rechannel the metabolism pathway to modify the monomer distribution of PHAs for its specific application in *P. putida.* First, proteomics analysis showed that high expression levels of beta-ketoadipyl-CoA thiolase (encoded by *pcaF*) involved in aromatic ring open pathway may be related to high degradation degree of aromatics, which may improve the production of acetyl-CoA and thus facilitate the PHA accumulation. The overexpression of *pcaF* in *P. putida* KT2440 was conducted and results showed that even PHA content hardly increased after fermentation, the long chain of monomer composition (HDD and HTD) in PHAs was improved, suggesting the tuned properties of PHAs. The enhancements of β-oxidation pathway may facilitate the PHA synthesis as proteomics analysis indicated that the overexpression of *phaJ* improved the PHA accumulation. Therefore, the gene of *FadD* involved in β-oxidation pathway was overexpressed to enhance the degradation of fatty acid for improving PHA synthesis. Interestingly, the *P. putida* KT2440 cell grows well on CEL medium, while the increase of PHA accumulation was not observed. This genetic engineering strategy modified the PHA properties by increasing the long chain of monomer composition (HDD and HTD). *P. putida* KT2440 grown on non-fatty acid substrates is highly dependent on the expression levels of *phaG*, encoding a transacylase that converts ACP thioesters to acyl-CoAs. Results showed that the genetic engineering strategy of the overexpression of *phaG* not only led to the better cell growth,

but also promoted the PHA accumulation. As a result, a record PHA concentration was obtained using CEL as carbon sources with fed-batch fermentation mode.

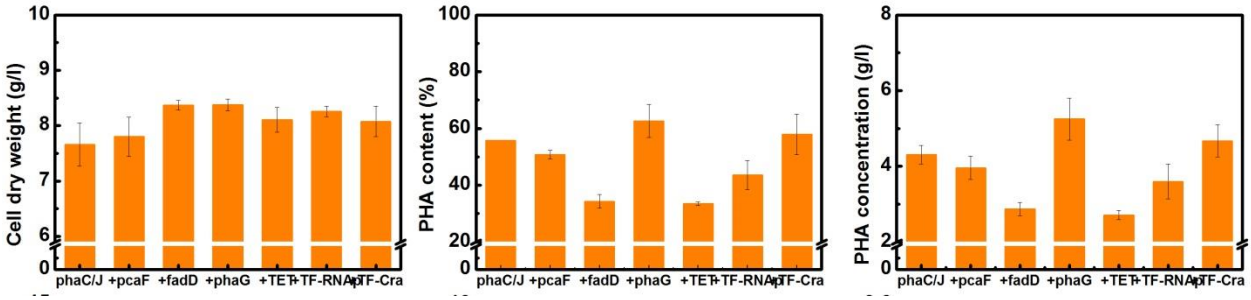

Figure 4.8 Improved fermentation for the production of PHAs by engineered *P. putida* KT2440 grown on CEL cosubstrate. CEL represents chemical-enzymatic treated liquor.

Besides, there transcription factors encoded by *TET*, *TF-RNAp, TF-Cra* was overexpressed to regulate gene expression and transcription. All these three genetic engineering strategies facilitated the cell growth, while the overexpression of TF-RNAp slightly improved the PHA accumulation. These strategies also changed the monomer distribution of PHAs by increasing the 3HTD and 3HDD content. Overall, the proteomics analysis helped to further guide the engineering design of *P. putida* KT2440 to improve PHA synthesis and tune PHA constitutes.

4.4.5 Biological lignin valorization route for the improvement of PHA synthesis

Overall, an effective biological route for lignin valorization has been created through processing engineering design, systems biology analysis and genetic engineering, and fermentation development. First, an innovative fractionation of lignin processing has been designed to tune the biological reactivity of lignin from process engineering concept of biorefinery (Figures 4.2 and 4.3).

This innovative process employed chemical-enzymatic treatment technology for lignin fractionation by sequentially deconstructing LCC structure of corn stover, enzymatically modifying and depolymerizing the lignin polymer through cleaving lignin linkages, reducing molecular weight, and generating aromatic compounds. The improved biological reactivity of lignin had been confirmed by improving cell growth and PHA accumulation in the microbial conversion of lignin. Second, the systems biology analysis helps the understanding the metabolism of ligninolytic *P. putida* KT2440 and guided the redesign the metabolic pathways for the production of PHAs currently sourced from lignin derivatives. The genetic engineering scenarios were then developed by overexpressing several genes involved in the lignin metabolism and PHA synthesis (Figures 4.2, 4.5, and 4.9).

As compared with wild type strain, the engineering of *P. putida* KT244 significantly increased the cell growth and promoted the PHA synthesis when it was grown on CEL medium in fed-batch fermentation. Much higher content of HTD and HDD in PHAs was produced, which could directly define the properties of PHAs. The genetic engineering and lignin medium option shows the promising potential to rechanneled the pathways for lignin metabolism and PHA synthesis. Third, the fermentation strategies were exploited to improve the PHA accumulation from lignin medium by optimizing substrate type and concentration, nitrogen source, fermentation mode and parameter control. The fed-batch fermentation with two cycles was carried out at high concentration of CEL by *P. putida*. This fermentation scenario promoted the cell growth, lignin consumption and PHA accumulation. Taken these together, a comparable level of PHA concentration was obtained correspondingly. Overall, a promising biological conversion route of lignin valorization have been

created to facilitate the lignin utilization and the target product yield, which holds promise for improving the profitability of microbial lignin conversion and biorefinery.

4.5 Conclusion

A promising microbial conversion route for lignin valorization to PHAs has been created by improving the lignin utilization and PHA yield. An innovative fractionation was first proposed to improve its biological reactivity. The systems biology analysis was then carried out to help the understanding the lignin metabolism and PHA synthesis, which guided the redesign the metabolic pathways to provide a renewable route for the production of PHAs. By enhancing the expression of proteins involved in the PHA synthesis, the PHA yield was significantly increased. By optimizing fermentation technology, a comparable level of PHA concentration was obtained Therefore, the biological lignin valorization holds the potential to improve the lignin utilization and could enable the profitability of biorefineries.

CHAPTER V

CONCLUSIONS AND PERSEPCTIVES

Biofuel and bioproducts from microorganism have attracted massive attention as renewable sources of energy and valuable chemicals. Significant advances have been made toward the development of metabolic engineering for higher productivity in microorganisms. However, due to the complex regulations of the metabolic pathways, a comprehensive understanding of the carbon flux between primary and secondary metabolites in microorganisms is necessary. Also, the mechanism of efficient utilization of lignocellulosic biomass by microorganisms is still needs to be well studied. In this dissertation research, based the combinations of metabolomics and proteomics study, higher limonene and PHA production in *S. elongatus* and *P. putida* were achieved with appropriate genetic engineering strategies.

For limonene production in *S. elongatus*, comparative proteomics and transcriptomics was done between the high and low limonene productivity strains. Several downregulated genes in the CBB cycle showed deficient photosynthesis in the low limonene productivity strains, indicating a limited carbon pool. However, further engineering of these genes' overexpression did not increase limonene productivity. Results showed that reduced carbon pool cannot be made up by the increasing activity of several enzymes. To get a further understanding of the carbon flux during the late cell growth stage (low limonene productivity), a metabolomic study was used to study the major carbon flux variations. As major carbon and energy storages, sucrose and glycogen were found an increased level in the low limonene productivity strain. With blocking sucrose and glycogen biosynthesis pathways, the mutant strains reached yields of 4.1mg/L and 4.9mg/L respectively. Although a considerable increase in limonene production was attained in this study,

the regulations between the primary and MEP pathways are still unclear. Further studies are needed to engineer the MEP pathway efficiency by identification of the rate-limiting step and replacement of inefficient enzymes with more efficient enzymes or the overexpression of the enzyme.

For PHA production in *P. putida*, the processibility of the lignin substrate limits the target product yield. Thus, several strategies were employed to increase the lignin utilization efficiency. Firstly, a fractionation strategy of lignin processing by integrating dilute acid pretreatment, enzymatic and alkaline treatment was developed for lignin biological reactivity. Then a comparative proteomics study between the strains culturing on different lignin substrate was used for finding key enzymes involved in efficient lignin utilization and PHA biosynthesis pathways. With further genetic engineering work based on the finding in proteomics study, a cell dry weight of 6.4g/L with 58% PHA content was reached in the engineered strain. We also saw the changes in the PHA monomer contents between the engineering strains. The mechanisms of these changes are still under investigation. To further increase the PHA production from *P.putida*, there are three main perspectives. First, the optimization of the carbon source, lignin, studies could focus on the utilization of lignin fragments in *P. putida* with deep analysis of the aromatic compound consumption. And specific pretreatments could be developed for generating the favored aromatic compounds. Moreover, the metabolic pathway involved in the specific aromatic compounds could be further enhanced. Second, the cultivation conditions could be further optimized. In our study, fed-batch experiments were used for high tier PHA production. In the first batch, *P. putida* were grown on glucose medium to reach a high OD yield. Other substrates favored by *P. putida* could be used to in the first batch to further tune the strains for high PHA production. Third, with a deeper understanding of acetyl-CoA metabolisms, rechanneling the carbon flux from other

metabolites such as fatty acid would help to increase the PHA yield. Overall, understanding the regulations between the PHA biosynthesis pathway and substrate utilization pathway, together with the optimization of cultivation conditions would further increase PHA yield in *P. putida.*

REFERENCES

Abdelaziz, O.Y., Brink, D.P., Prothmann, J., Ravi, K., Sun, M.Z., Garcia-Hidalgo, J., et al. (2016). Biological valorization of low molecular weight lignin. *Biotechnology Advances* 34(8)**,** 1318-1346.

Aikawa, S., Nishida, A., Ho, S.-H., Chang, J.-S., Hasunuma, T., and Kondo, A. (2014). Glycogen production for biofuels by the euryhaline cyanobacteria *Synechococcus sp.* strain PCC 7002 from an oceanic environment. *Biotechnology for Biofuels* 7(1)**,** 88-95.

Ainsworth, E.A., and Long, S.P. (2005). What have we learned from 15 years of free-air CO2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO2. *New Phytologist* 165(2)**,** 351-372.

Ajikumar, P.K., Xiao, W.-H., Tyo, K.E.J., Wang, Y., Simeon, F., Leonard, E., et al. (2010). Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. *Science* 330(6000)**,** 70-80.

Atsumi, S., Higashide, W., and Liao, J.C. (2009). Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nature Biotechnology* 27(12)**,** 1177-1180.

Atsumi, S., and Liao, J.C. (2008). Metabolic engineering for advanced biofuels production from *Escherichia coli*. *Current Opinion in Biotechnology* 19(5)**,** 414-419.

Badger, M.R., and Price, G.D. (2003). CO2 concentrating mechanisms in cyanobacteria: Molecular components, their diversity and evolution. *Journal of Experimental Botany* 54(383)**,** 609-622.

Bajwa, D.S., Pourhashem, G., Ullah, A.H., and Bajwa, S.G. (2019). A concise review of current lignin production, applications, products and their environmental impact. *Industrial Crops and Products* 139-150.

Balch, M.L., Holwerda, E.K., Davis, M.F., Sykes, R.W., Happs, R.M., Kumar, R., et al. (2017). Lignocellulose fermentation and residual solids characterization for senescent switchgrass fermentation by *Clostridium thermocellum* in the presence and absence of continuous in situ ball-milling. *Energy & Environmental Science* 10(5)**,** 1252-1261.

Banerjee, A., and Sharkey, T.D. (2014). Methylerythritol 4-phosphate (MEP) pathway metabolic regulation. *Natural Product Reports* 31(8)**,** 1043-1055.

Barkley, S.J., Desai, S.B., and Poulter, C.D. (2004). Type II isopentenyl diphosphate isomerase from *Synechocystis sp.* strain PCC 6803. *Journal of Bacteriology* 186(23)**,** 8156-8164.

Beckham, G.T., Johnson, C.W., Karp, E.M., Salvachua, D., and Vardon, D.R. (2016). Opportunities and challenges in biological lignin valorization. *Current Opinion in Biotechnology* 42**,** 40-53.

Behera, S., Singh, R., Arora, R., Sharma, N.K., Shukla, M., and Kumar, S. (2015). Scope of algae as third generation biofuels. *Frontiers in Bioengineering and Biotechnology* 2**,** 90-98.

Bentley, F.K., and Melis, A. (2012). Diffusion-based process for carbon dioxide uptake and isoprene emission in gaseous/aqueous two-phase photobioreactors by photosynthetic microorganisms. *Biotechnology and Bioengineering* 109(1)**,** 100-109.

Bentley, F.K., Zurbriggen, A., and Melis, A. (2014). Heterologous expression of the mevalonic acid pathway in cyanobacteria enhances endogenous carbon partitioning to isoprene. *Molecular Plant* 7(1)**,** 71-86.

Bohlmann, J., and Keeling, C.I. (2008). Terpenoid biomaterials. *The Plant Journal* 54(4)**,** 656-669.

Bricker, T.M., Roose, J.L., Fagerlund, R.D., Frankel, L.K., and Eaton-Rye, J.J. (2012). The extrinsic proteins of photosystem II. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1817(1)**,** 121-142.

Cano, M., Holland, S.C., Artier, J., Burnap, R.L., Ghirardi, M., Morgan, J.A., et al. (2018). Glycogen synthesis and metabolite overflow contribute to energy balancing in cyanobacteria. *Cell Reports* 23(3)**,** 667-672.

Carrieri, D., Lombardi, T., Paddock, T., Cano, M., Goodney, G.A., Nag, A., et al. (2017). Transcriptome and proteome analysis of nitrogen starvation responses in Synechocystis 6803 ΔglgC, a mutant incapable of glycogen storage. *Algal Research* 21**,** 64-75.

Chatzivasileiou, A.O., Ward, V., Edgar, S.M., and Stephanopoulos, G. (2019). Two-step pathway for isoprenoid synthesis. *Proceedings of the National Academy of Sciences* 116(2)**,** 506-514.

Chen, G.Q. (2009). A microbial polyhydroxyalkanoates (PHA) based bio- and materials industry. *Chemical Society Reviews* 38(8)**,** 2434-2446.

Chen, H.Z., and Liu, Z.H. (2015). Steam explosion and its combinatorial pretreatment refining technology of plant biomass to bio‐based products. *Biotechnology Journal* 10(6)**,** 866-885.

Chen, Z., and Wan, C. (2017a). Biological valorization strategies for converting lignin into fuels and chemicals. *Renewable and Sustainable Energy Reviews* 73**,** 610-621.

Cho, C., Choi, S.Y., Luo, Z.W., and Lee, S.Y. (2015). Recent advances in microbial production of fuels and chemicals using tools and strategies of systems metabolic engineering. *Biotechnology Advances* 33(7)**,** 1455-1466.

Choi, S.Y., Lee, H.J., Choi, J., Kim, J., Sim, S.J., Um, Y., et al. (2016). Photosynthetic conversion of CO2 to farnesyl diphosphate-derived phytochemicals (amorpha-4,11-diene and squalene) by engineered cyanobacteria. *Biotechnology for Biofuels* 9(1)**,** 202-220.

Chuck, C., and Donnelly, J. (2014). The compatibility of potential bioderived fuels with Jet A-1 aviation kerosene. *Applied Energy* 118**,** 83-91.

Degenhardt, J., Köllner, T.G., and Gershenzon, J. (2009). Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* 70(15)**,** 1621-1637.

Demirbas, A. (2010). Use of algae as biofuel sources. *Energy Conversion and Management* 51(12)**,** 2738-2749.

Dexter, J., and Fu, P. (2009). Metabolic engineering of cyanobacteria for ethanol production. *Energy & Environmental Science* 2(8)**,** 857-864.

Ducat, D.C., Avelar-Rivas, J.A., Way, J.C., and Silver, P.A. (2012). Rerouting carbon flux to enhance photosynthetic productivity. *Applied and Environmental Microbiology* 78(8)**,** 2660-2673.

Ducat, D.C., and Silver, P.A. (2012). Improving carbon fixation pathways. *Current Opinion in Chemical Biology* 16(3)**,** 337-344.

Ducat, D.C., Way, J.C., and Silver, P.A. (2011). Engineering cyanobacteria to generate high-value products. *Trends in Biotechnology* 29(2)**,** 95-103.

Eisenreich, W., Bacher, A., Arigoni, D., and Rohdich, F. (2004). Biosynthesis of isoprenoids via the non-mevalonate pathway. *Cellular and Molecular Life Sciences CMLS* 61(12)**,** 1401-1426.

Eisenreich, W., Rohdich, F., and Bacher, A. (2001). Deoxyxylulose phosphate pathway to terpenoids. *Trends in Plant Science* 6(2)**,** 78-84.

Escapa, I.F., del Cerro, C., García, J.L., and Prieto, M.A. (2013). The role of GlpR repressor in *Pseudomonas putida* KT2440 growth and PHA production from glycerol. *Environmental Microbiology* 15(1)**,** 93-110.

Evans, J.R. (2013). Improving photosynthesis. *Plant Physiology* 162(4)**,** 1780-1790.

Faccin, D.J.L., Martins, I., Cardozo, N.S.M., Rech, R., Ayub, M.A.Z., Alves, T.L.M., et al. (2009). Optimization of C : N ratio and minimal initial carbon source for poly(3-hydroxybutyrate) production by *Bacillus megaterium*. *Journal of Chemical Technology and Biotechnology* 84(12)**,** 1756-1761.

Farmer, W.R., and Liao, J.C. (2001). Precursor balancing for metabolic engineering of lycopene production in *Escherichia coli*. *Biotechnology Progress* 17(1)**,** 57-61.

Fatih Demirbas, M. (2009). Biorefineries for biofuel upgrading: A critical review. *Applied Energy* 86**,** S151-S161.

Fortman, J.L., Chhabra, S., Mukhopadhyay, A., Chou, H., Lee, T.S., Steen, E., et al. (2008). Biofuel alternatives to ethanol: Pumping the microbial well. *Trends in Biotechnology* 26(7)**,** 375-381.

Fuchs, G., Boll, M., and Heider, J. (2011). Microbial degradation of aromatic compounds - from one strategy to four. *Nature Reviews Microbiology* 9(11)**,** 803-816.

Gao, X., Gao, F., Liu, D., Zhang, H., Nie, X., and Yang, C. (2016). Engineering the methylerythritol phosphate pathway in cyanobacteria for photosynthetic isoprene production from CO2. *Energy & Environmental Science* 9(4)**,** 1400-1411.

Gao, Z., Zhao, H., Li, Z., Tan, X., and Lu, X. (2012). Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria. *Energy & Environmental Science* 5(12)**,** 9857-9865.

Gershenzon, J., and Dudareva, N. (2007). The function of terpene natural products in the natural world. *Nature Chemical Biology* 3(7)**,** 408-414.

Gillet, S., Aguedo, M., Petitjean, L., Morais, A.R.C., Lopes, A.M.D., Lukasik, R.M., et al. (2017). Lignin transformations for high value applications: Towards targeted modifications using green chemistry. *Green Chemistry* 19(18)**,** 4200-4233.

Hackett, S.R., Zanotelli, V.R.T., Xu, W., Goya, J., Park, J.O., Perlman, D.H., et al. (2016). Systems-level analysis of mechanisms regulating yeast metabolic flux. *Science* 354(6311)**,** 2786-2794.

Hagemann, M., and Bauwe, H. (2016). Photorespiration and the potential to improve photosynthesis. *Current Opinion in Chemical Biology* 35**,** 109-116.

Halfmann, C., Gu, L., and Zhou, R. (2014). Engineering cyanobacteria for the production of a cyclic hydrocarbon fuel from CO2 and H2O. *Green Chemistry* 16(6)**,** 3175-3185.

Hall, G.C., Flick, M.B., Gherna, R.L., and Jensen, R.A. (1982). Biochemical diversity for biosynthesis of aromatic amino acids among the cyanobacteria. *Journal of Bacteriology* 149(1)**,** 65-78.

Himmel, M.E., Ding, S.-Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., Brady, J.W., et al. (2007). Biomass recalcitrance: Engineering plants and enzymes for biofuels production. *Science* 315**,** 804-807.

Hirokawa, Y., Suzuki, I., and Hanai, T. (2015). Optimization of isopropanol production by engineered cyanobacteria with a synthetic metabolic pathway. *Journal of Bioscience and Bioengineering* 119(5)**,** 585-590.

Hoffmann, N., Steinbüchel, A., and Rehm, B.H.A. (2000). Homologous functional expression of cryptic phaG from *Pseudomonas oleovorans* establishes the transacylase-mediated polyhydroxyalkanoate biosynthetic pathway. *Applied Microbiology and Biotechnology* 54(5)**,** 665-670.

Hoiczyk, E., and Hansel, A. (2000). Cyanobacterial cell walls: News from an unusual prokaryotic envelope. *Journal of Bacteriology* 182(5)**,** 1191-1203.

Hu, F., and Ragauskas, A. (2012). Pretreatment and lignocellulosic chemistry. *BioEnergy Research* 5(4)**,** 1043-1066.

Isikgor, F.H., and Becer, C.R. (2015). Lignocellulosic biomass: A sustainable platform for the production of bio-based chemicals and polymers. *Polymer Chemistry* 6(25)**,** 4497-4559.

Jindou, S., Ito, Y., Mito, N., Uematsu, K., Hosoda, A., and Tamura, H. (2014). Engineered platform for bioethylene production by a cyanobacterium expressing a chimeric complex of plant enzymes. *ACS Synthetic Biology* 3(7)**,** 487-496.

Jonsson, L.J., and Martin, C. (2016). Pretreatment of lignocellulose: Formation of inhibitory by-products and strategies for minimizing their effects. *Bioresource Technology* 199**,** 103-112.

Kirby, J., and Keasling, J.D. (2009). Biosynthesis of plant isoprenoids: Perspectives for microbial engineering. *Annual Review of Plant Biology* 60(1)**,** 335-355.

Kiyota, H., Okuda, Y., Ito, M., Hirai, M.Y., and Ikeuchi, M. (2014). Engineering of cyanobacteria for the photosynthetic production of limonene from CO2. *Journal of Biotechnology* 185**,** 1-7.

Koppolu, V., and Vasigala, V.K.R. (2016). Role of *Escherichia coli* in biofuel production. *Microbiology Insights* 9**,** 10878-10890.

Kumar, P., Maharjan, A., Jun, H.B., and Kim, B.S. (2019). Bioconversion of lignin and its derivatives into polyhydroxyalkanoates: Challenges and opportunities. *Biotechnology and Applied Biochemistry* 66(2)**,** 153-162.

Kusakabe, T., Tatsuke, T., Tsuruno, K., Hirokawa, Y., Atsumi, S., Liao, J.C., et al. (2013). Engineering a synthetic pathway in cyanobacteria for isopropanol production directly from carbon dioxide and light. *Metabolic Engineering* 20**,** 101-108.

La Rosa, R., de la Pena, F., Prieto, M.A., and Rojo, F. (2014). The Crc protein inhibits the production of polyhydroxyalkanoates in *Pseudomonas putida* under balanced carbon/nitrogen growth conditions. *Environmental Microbiology* 16(1)**,** 278-290.

Lan, E.I., and Liao, J.C. (2011). Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide. *Metabolic Engineering* 13(4)**,** 353-363.

Lan, E.I., and Liao, J.C. (2012a). ATP drives direct photosynthetic production of 1-butanol in cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America* 109(16)**,** 6018-6023.

Lan, E.I., Ro, S.Y., and Liao, J.C. (2013). Oxygen-tolerant coenzyme A-acylating aldehyde dehydrogenase facilitates efficient photosynthetic n-butanol biosynthesis in cyanobacteria. *Energy & Environmental Science* 6(9)**,** 2672-2681.

Lau, M.W., Gunawan, C., and Dale, B.E. (2009). The impacts of pretreatment on the fermentability of pretreated lignocellulosic biomass: A comparative evaluation between ammonia fiber expansion and dilute acid pretreatment. *Biotechnology for Biofuels* 2(1)**,** 30-45.

Lee, J.W., Yi, J., Kim, T.Y., Choi, S., Ahn, J.H., Song, H., et al. (2016). Homo-succinic acid production by metabolically engineered *Mannheimia succiniciproducens*. *Metabolic Engineering* 38**,** 409-417.

Lee, R.A., and Lavoie, J.-M. (2013). From first- to third-generation biofuels: Challenges of producing a commodity from a biomass of increasing complexity. *Animal Frontiers* 3(2)**,** 6-11.

Lee, T.-C., Xiong, W., Paddock, T., Carrieri, D., Chang, I.-F., Chiu, H.-F., et al. (2015). Engineered xylose utilization enhances bio-products productivity in the cyanobacterium *Synechocystis* sp. PCC 6803. *Metabolic Engineering* 30**,** 179-189.

Lem, N.W., and Stumpf, P.K. (1984). In vitro fatty acid synthesis and complex lipid metabolism in the cyanobacterium *Anabaena variabilis*: I. Some characteristics of fatty acid synthesis. *Plant Physiology* 74(1)**,** 134-138.

Li, J.B., Henriksson, G., and Gellerstedt, G. (2007). Lignin depolymerization/repolymerization and its critical role for delignification of aspen wood by steam explosion. *Bioresource Technology* 98(16)**,** 3061-3068.

Li, X., Shen, C.R., and Liao, J.C. (2014). Isobutanol production as an alternative metabolic sink to rescue the growth deficiency of the glycogen mutant of *Synechococcus elongatus* PCC 7942. *Photosynthesis Research* 120(3)**,** 301-310.

Liang, F., and Lindblad, P. (2017). *Synechocystis* PCC 6803 overexpressing RuBisCO grow faster with increased photosynthesis. *Metabolic Engineering Communications* 4**,** 29-36.

Liew, W.H., Hassim, M.H., and Ng, D.K.S. (2014). Review of evolution, technology and sustainability assessments of biofuel production. *Journal of Cleaner Production* 71**,** 11-29.

Lima, M.A., Lavorente, G.B., da Silva, H.K., Bragatto, J., Rezende, C.A., Bernardinelli, O.D., et al. (2013). Effects of pretreatment on morphology, chemical composition and enzymatic digestibility of eucalyptus bark: A potentially valuable source of fermentable sugars for biofuel production – part 1. *Biotechnology for Biofuels* 6(1)**,** 75-90.

Lin, L., Cheng, Y., Pu, Y., Sun, S., Li, X., Jin, M., et al. (2016a). Systems biology-guided biodesign of consolidated lignin conversion. *Green Chemistry* 18(20)**,** 5536-5547.

Lin, M.T., Occhialini, A., Andralojc, P.J., Parry, M.A.J., and Hanson, M.R. (2014). A faster Rubisco with potential to increase photosynthesis in crops. *Nature* 513(7519)**,** 547-550.

Lindberg, P., Park, S., and Melis, A. (2010). Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. *Metabolic Engineering* 12(1)**,** 70-79.

Linger, J.G., Vardon, D.R., Guarnieri, M.T., Karp, E.M., Hunsinger, G.B., Franden, M.A., et al. (2014). Lignin valorization through integrated biological funneling and chemical catalysis. *Proceedings of the National Academy of Sciences of the United States of America* 111(33)**,** 12013-12018.

Liu, H., Sun, Y., Ramos, K.R.M., Nisola, G.M., Valdehuesa, K.N.G., Lee, W.-K., et al. (2013). Combination of entner-doudoroff pathway with MEP increases isoprene production in engineered Escherichia coli. *Plos One* 8(12)**,** e83290-e83290.

Liu, X., Sheng, J., and Curtiss Iii, R. (2011). Fatty acid production in genetically modified cyanobacteria. *Proceedings of the National Academy of Sciences* 108(17)**,** 6899-6907.

Liu, Z.H., Le, R.K., Kosa, M., Yang, B., Yuan, J.S., and Ragauskas, A.J. (2019). Identifying and creating pathways to improve biological lignin valorization. *Renewable & Sustainable Energy Reviews* 105**,** 349-362.

Liu, Z.H., Olson, M.L., Shinde, S., Wang, X., Hao, N.J., Yoo, C.G., et al. (2017). Synergistic maximization of the carbohydrate output and lignin processability by combinatorial pretreatment. *Green Chemistry* 19(20)**,** 4939-4955.

Liu, Z.H., Qin, L., Li, B.Z., and Yuan, Y.J. (2015). Physical and chemical characterizations of corn stover from leading pretreatment methods and effects on enzymatic hydrolysis. *Acs Sustainable Chemistry & Engineering* 3(1)**,** 140-146.

Liu, Z.H., Xie, S.X., Lin, F.R., Jin, M.J., and Yuan, J.S. (2018). Combinatorial pretreatment and fermentation optimization enabled a record yield on lignin bioconversion. *Biotechnology for Biofuels,* 11-21.

Lu, X. (2010). A perspective: Photosynthetic production of fatty acid-based biofuels in genetically engineered cyanobacteria. *Biotechnology Advances* 28(6)**,** 742-746.

Machado, I.M.P., and Atsumi, S. (2012). Cyanobacterial biofuel production. *Journal of Biotechnology* 162(1)**,** 50-56.

Madhavan, A., Srivastava, A., Kondo, A., and Bisaria, V.S. (2012). Bioconversion of lignocellulose-derived sugars to ethanol by engineered *Saccharomyces cerevisiae*. *Critical Reviews in Biotechnology* 32(1)**,** 22-48.

Matsakas, L., Nitsos, C., Raghavendran, V., Yakimenko, O., Persson, G., Olsson, E., et al. (2018). A novel hybrid organosolv: Steam explosion method for the efficient fractionation and pretreatment of birch biomass. *Biotechnology for Biofuels* 11(1)**,** 160-171.

Matthews, P.D., and Wurtzel, E.T. (2000). Metabolic engineering of carotenoid accumulation in *Escherichia coli* by modulation of the isoprenoid precursor pool with expression of deoxyxylulose phosphate synthase. *Applied Microbiology and Biotechnology* 53(4)**,** 396-400.

Melis, A. (2013). Carbon partitioning in photosynthesis. *Current Opinion in Chemical Biology* 17(3)**,** 453-456.

Miziorko, H.M. (2011). Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Archives of Biochemistry and Biophysics* 505(2)**,** 131-143.

Morin, N., Vallaeys, T., Hendrickx, L., Natalie, L., and Wilmotte, A. (2010). An efficient DNA isolation protocol for filamentous cyanobacteria of the genus *Arthrospira*. *Journal of Microbiological Methods* 80(2)**,** 148-154.

Naik, S.N., Goud, V.V., Rout, P.K., and Dalai, A.K. (2010). Production of first and second generation biofuels: A comprehensive review. *Renewable and Sustainable Energy Reviews* 14(2)**,** 578-597.

Nelson, K.E., Weinel, C., Paulsen, I.T., Dodson, R.J., Hilbert, H., dos Santos, V.A.P.M., et al. (2002). Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environmental Microbiology* 4(12)**,** 799-808.

O'Leary, N.D., O'Connor, K.E., Duetz, W., and Dobson, A.D.W. (2001). Transcriptional regulation of styrene degradation in *Pseudomonas putida* CA-3 the GenBank accession number for the sequence determined in this work is AF257095. *Microbiology* 147(4)**,** 973-979.

Parmar, A., Singh, N.K., Pandey, A., Gnansounou, E., and Madamwar, D. (2011). Cyanobacteria and microalgae: A positive prospect for biofuels. *Bioresource Technology* 102(22)**,** 10163-10172.

Pasoreck, E.K., Su, J., Silverman, I.M., Gosai, S.J., Gregory, B.D., Yuan, J.S., et al. (2016). Terpene metabolic engineering via nuclear or chloroplast genomes profoundly and globally impacts off-target pathways through metabolite signalling. *Plant Biotechnology Journal* 14(9)**,** 1862-1875.

Pattanaik, B., and Lindberg, P. (2015). Terpenoids and their biosynthesis in cyanobacteria. *Life (Basel, Switzerland)* 5(1)**,** 269-293.

Peers, G. (2014). Increasing algal photosynthetic productivity by integrating ecophysiology with systems biology. *Trends in Biotechnology* 32(11)**,** 551-555.

Poblete-Castro, I., Becker, J., Dohnt, K., dos Santos, V.M., and Wittmann, C. (2012a). Industrial biotechnology of *Pseudomonas putida* and related species. *Applied Microbiology and Biotechnology* 93(6)**,** 2279-2290.

Poblete-Castro, I., Escapa, I.F., Jäger, C., Puchalka, J., Chi Lam, C.M., Schomburg, D., et al. (2012b). The metabolic response of *P. putida* KT2442 producing high levels of polyhydroxyalkanoate under single- and multiple-nutrient-limited growth: Highlights from a multi-level omics approach. *Microbial Cell Factories* 11(1)**,** 34-42.

Porchia, A.C., and Salerno, G.L. (1996). Sucrose biosynthesis in a prokaryotic organism: Presence of two sucrose-phosphate synthases with remarkable differences compared with the plant enzymes. *Proceedings of the National Academy of Sciences* 93(24)**,** 13600-13611.

Price, G.D., Pengelly, J.J.L., Forster, B., Du, J., Whitney, S.M., von Caemmerer, S., et al. (2012). The cyanobacterial CCM as a source of genes for improving photosynthetic CO2 fixation in crop species. *Journal of Experimental Botany* 64(3)**,** 753-768.

Ragauskas, A.J., Beckham, G.T., Biddy, M.J., Chandra, R., Chen, F., Davis, M.F., et al. (2014a). Lignin valorization: Improving lignin processing in the biorefinery. *Science* 344(6185)**,** 709-715.

Rasmussen, H., Tanner, D., Sorensen, H.R., and Meyer, A.S. (2017). New degradation compounds from lignocellulosic biomass pretreatment: Routes for formation of potent oligophenolic enzyme inhibitors. *Green Chemistry* 19(2)**,** 464-473.

Rinaldi, R., Jastrzebski, R., Clough, M.T., Ralph, J., Kennema, M., Bruijnincx, P.C.A., et al. (2016). Paving the way for lignin valorisation: Recent advances in bioengineering, biorefining and catalysis. *Angewandte Chemie-International Edition* 55(29)**,** 8164-8215.

Rosgaard, L., de Porcellinis, A.J., Jacobsen, J.H., Frigaard, N.-U., and Sakuragi, Y. (2012). Bioengineering of carbon fixation, biofuels, and biochemicals in cyanobacteria and plants. *Journal of Biotechnology* 162(1)**,** 134-147.

Ruffing, A.M., and Jones, H.D.T. (2012). Physiological effects of free fatty acid production in genetically engineered *Synechococcus elongatus* PCC 7942. *Biotechnology and Bioengineering* 109(9)**,** 2190-2199.

Sage, R.F. (2002). Variation in the kcat of Rubisco in C3 and C4 plants and some implications for photosynthetic performance at high and low temperature. *Journal of Experimental Botany* 53(369)**,** 609-620.

Salvachua, D., Karp, E.M., Nimlos, C.T., Vardon, D.R., and Beckham, G.T. (2015). Towards lignin consolidated bioprocessing: Simultaneous lignin depolymerization and product generation by bacteria. *Green Chemistry* 17(11)**,** 4951-4967.

Saqib, A., Rizwan, M., Rashid, U., Ibrahim, M., Gill, S., and Mehmood, M. (2013). Marine macro algae Ulva : A potential feed-stock for bio-ethanol and biogas production. *Asian Journal of Agriculture and Biology* 1**,** 155-163.

Savakis, P., and Hellingwerf, K.J. (2015). Engineering cyanobacteria for direct biofuel production from CO2. *Current Opinion in Biotechnology* 33**,** 8-14.

Schirmer, A., Rude, M.A., Li, X., Popova, E., and del Cardayre, S.B. (2010). Microbial biosynthesis of alkanes. *Science* 329(5991)**,** 559-562.

Sheldon, P.S., Kekwick, R.G., Sidebottom, C., Smith, C.G., and Slabas, A.R. (1990). 3-Oxoacyl-(acyl-carrier protein) reductase from avocado (*Persea americana*) fruit mesocarp. *The Biochemical Journal* 271(3)**,** 713-720.

Shinde, S., Zhang, X., Singapuri, S.P., Kalra, I., Liu, X., Morgan-Kiss, R.M., et al. (2020). Glycogen metabolism supports photosynthesis start through the oxidative pentose phosphate pathway in cyanobacteria. *Plant Physiology* 182(1)**,** 507-517.

Sims, R.E.H., Mabee, W., Saddler, J.N., and Taylor, M. (2010). An overview of second generation biofuel technologies. *Bioresource Technology* 101(6)**,** 1570-1580.

Song, C. (2006). Global challenges and strategies for control, conversion and utilization of CO2 for sustainable development involving energy, catalysis, adsorption and chemical processing. *Catalysis Today* 115(1)**,** 2-32.

Sudesh, K., Abe, H., and Doi, Y. (2000). Synthesis, structure and properties of polyhydroxyalkanoates: Biological polyesters. *Progress in Polymer Science* 25(10)**,** 1503-1555.

Sung-Jin, J.O., Ken'ichiro, M., Ring, L.C., Ooi, T., and Taguchi, S. (2007). Improvement of poly(3-Hydroxybutyrate) [P(3HB)] production in corynebacterium glutamicum by codon optimization, point mutation and gene dosage of P(3HB) biosynthetic genes. *Journal of Bioscience and Bioengineering* 104(6)**,** 457-463.

Takahama, K., Matsuoka, M., Nagahama, K., and Ogawa, T. (2003). Construction and analysis of a recombinant cyanobacterium expressing a chromosomally inserted gene for an ethylene-forming enzyme at the psbAI locus. *Journal of Bioscience and Bioengineering* 95(3)**,** 302-305.

Tian, D., Chandra, R.P., Lee, J.S., Lu, C.H., and Saddler, J.N. (2017). A comparison of various lignin-extraction methods to enhance the accessibility and ease of enzymatic hydrolysis of the cellulosic component of steam-pretreated poplar. *Biotechnology for Biofuels*, 10-18.

Varman, A.M., He, L., Follenfant, R., Wu, W.H., Wemmer, S., Wrobel, S.A., et al. (2016). Decoding how a soil bacterium extracts building blocks and metabolic energy from ligninolysis provides road map for lignin valorization. *Proceedings of the National Academy of Sciences of the United States of America* 113(40)**,** E5802-E5811.

Wang, X., Liu, W., Xin, C., Zheng, Y., Cheng, Y., Sun, S., et al. (2016). Enhanced limonene production in cyanobacteria reveals photosynthesis limitations. *Proceedings of the National Academy of Sciences* 113(50)**,** 14225-14237.

Wang, X., Ort, D.R., and Yuan, J.S. (2015). Photosynthetic terpene hydrocarbon production for fuels and chemicals. *Plant Biotechnology Journal* 13(2)**,** 137-146.

Wasylenko, T.M., and Stephanopoulos, G. (2015). Metabolomic and (13)C-metabolic flux analysis of a xylose-consuming *Saccharomyces cerevisiae* strain expressing xylose isomerase. *Biotechnology and Bioengineering* 112(3)**,** 470-483.

Webb, H., Lanfear, R., Hamill, J., Foley, W.J., and Külheim, C. (2013). The yield of essential oils in *Melaleuca alternifolia* (Myrtaceae) is regulated through transcript abundance of genes in the MEP pathway. *Plos One* 8(3)**,** e60631-e60641.

Wei, Z., Zeng, G.M., Huang, F., Kosa, M., Huang, D.L., and Ragauskas, A.J. (2015). Bioconversion of oxygen-pretreated Kraft lignin to microbial lipid with oleaginous *Rhodococcus opacus* DSM 1069. *Green Chemistry* 17(5)**,** 2784-2789.

Whited, G.M., Feher, F.J., Benko, D.A., Cervin, M.A., Chotani, G.K., McAuliffe, J.C., et al. (2010). Technology update: Development of a gas-phase bioprocess for isoprene-monomer production using metabolic pathway engineering. *Industrial Biotechnology* 6(3)**,** 152-163.

Xiao, Y., Zahariou, G., Sanakis, Y., and Liu, P. (2009). IspG enzyme activity in the deoxyxylulose phosphate pathway: Roles of the iron−sulfur cluster. *Biochemistry* 48(44)**,** 10483-10485.

Xie, S., Sun, Q., Pu, Y., Lin, F., Sun, S., Wang, X., et al. (2017). Advanced chemical design for efficient lignin bioconversion. *ACS Sustainable Chemistry & Engineering* 5(3)**,** 2215-2223.

Xiong, W., Cano, M., Wang, B., Douchi, D., and Yu, J. (2017). The plasticity of cyanobacterial carbon metabolism. *Current Opinion in Chemical Biology* 41**,** 12-19.

Xu, Z., Lei, P., Zhai, R., Wen, Z., and Jin, M. (2019). Recent advances in lignin valorization with bacterial cultures: Microorganisms, metabolic pathways, and bio-products. *Biotechnology for Biofuels* 12(1)**,** 32-42.

Yuan, J.S., Himanen, S.J., Holopainen, J.K., Chen, F., and Stewart, C.N. (2009). Smelling global climate change: Mitigation of function for plant volatile organic compounds. *Trends in Ecology & Evolution* 24(6)**,** 323-331.

Yuan, J.S., Tiller, K.H., Al-Ahmad, H., Stewart, N.R., and Stewart, C.N. (2008). Plants to power: Bioenergy to fuel the future. *Trends in Plant Science* 13(8)**,** 421-429.

Yuan, J.S., Wang, X., and Stewart, C.N. (2011). Biomass feedstock: Diversity as a solution. *Biofuels* 2(5)**,** 491-493.

Yuan, L.Z., Rouvière, P.E., LaRossa, R.A., and Suh, W. (2006). Chromosomal promoter replacement of the isoprenoid pathway for enhancing carotenoid production in *E. coli*. *Metabolic Engineering* 8(1)**,** 79-90.

Zhang, Y., Liu, S., Dai, S.Y., and Yuan, J.S. (2012). Integration of shot-gun proteomics and bioinformatics analysis to explore plant hormone responses. *BMC Bioinformatics* 13(15)**,** S8-S17.

Zhang, Y.X., Gao, P., Xing, Z., Jin, S.M., Chen, Z.D., Liu, L.T., et al. (2013). Application of an improved proteomics method for abundant protein cleanup: Molecular and genomic mechanisms study in plant defense. *Molecular & Cellular Proteomics* 12(11)**,** 3431-3442.

Zhang, Z., Harrison, M.D., Rackemann, D.W., Doherty, W.O.S., and O'Hara, I.M. (2016). Organosolv pretreatment of plant biomass for enhanced enzymatic saccharification. *Green Chemistry*, 18(2)**,** 360-381.

Zhao, K., Deng, Y., Chen, J.C., and Chen, G.Q. (2003). Polyhydroxyalkanoate (PHA) scaffolds with good mechanical properties and biocompatibility. *Biomaterials* 24(6)**,** 1041-1045.

Zhao, X., Cheng, K., and Liu, D. (2009). Organosolv pretreatment of lignocellulosic biomass for enzymatic hydrolysis. *Applied Microbiology and Biotechnology* 82(5)**,** 815-822.

Zhuang, X., Wang, W., Yu, Q., Qi, W., Wang, Q., Tan, X., et al. (2016). Liquid hot water pretreatment of lignocellulosic biomass for bioethanol production accompanying with high valuable products. *Bioresource Technology* 199**,** 68-75.

APPENDIX

Table A-1: Strains and plasmids used in this study

Strain or plasmid	Relevant characteristics	Source or reference
Strains		
S. elogatus PCC 7942	Wild type	S. Golden
L1118	*ls* at NSI of *S. elogatus* PCC7942	X. Wang
Lsps	Δ*sps* of L1118	This study
L1218	*IspG*(from Synechocystis sp. PCC 6803) at NSII of L1118	This study
L1219	*IspG*(from *Botryococcus braunii*) at NSII of L1118	This study
L1221	*RibB* (Codon optimized for PCC7942) and *IspG*(from *S. elogatus* PCC7942) at NSII of L1118	This study
L2211	*Gap2* and *PGK* (from Synechocystis sp. PCC 6803) at NSII of L1118	This study
LglgC	ΔglgC of L1118	This study
LglgC-IDI	ΔglgC, idi (from *Saccharomyces cerevisiae*) at NSIII	This study
LglgC-sps	ΔglgC and Δ*sps* of L1118	This Study
Plasmids		
pAM2991	Targeting PCC7942 NSI; Ptrc; Sp/Sm^R	SS. Golden
pAM1579	Targeting PCC7942 NSII; $P_L$lacO, Km^R	CR. Andersson
pLM21	Targeting PCC7942 *sps*, Gm^R	This study
pWX1218	*IspG*(from *Synechocystis sp*. PCC 6803 in pAM1579, Km^R	This study
pWX1219	*IspG*(from *Botryococcus braunii* in pAM1579, Km^R	This study
pWX1221	*RibB* (Codon optimized for PCC7942) and *IspG*(from *S. elogatus* PCC7942 in pAM1579, Km^R	This study
pLM2211	*Gap2* and *PGK* (from Synechocystis sp. PCC 6803) in pAM1579, Km^R	This study
Δ0603	Targeting PCC7942 glgC, Km^R	James Golden
pLM232	*Idi* (from *Saccharomyces cerevisiae*) ligated to origin of pAM2991 and NSIII, Cm^R	This Study
pLMPHA1	*FadD* insertion in pPROBE, (Km^R)	This study
pLMPHA2	*PcaF-I* insertion in pPROBE, (Km^R)	This study
pLMPHA3	*phaG* insertion in pPROBE, (Km^R)	This study
pLMPHA4	*TET* insertion in pPROBE, (Km^R)	This study
pLMPHA5	*RNAp* insertion in pPROBE, (Km^R)	This study
pLMPHA6	*Cra* insertion in pPROBE, (Km^R)	This study

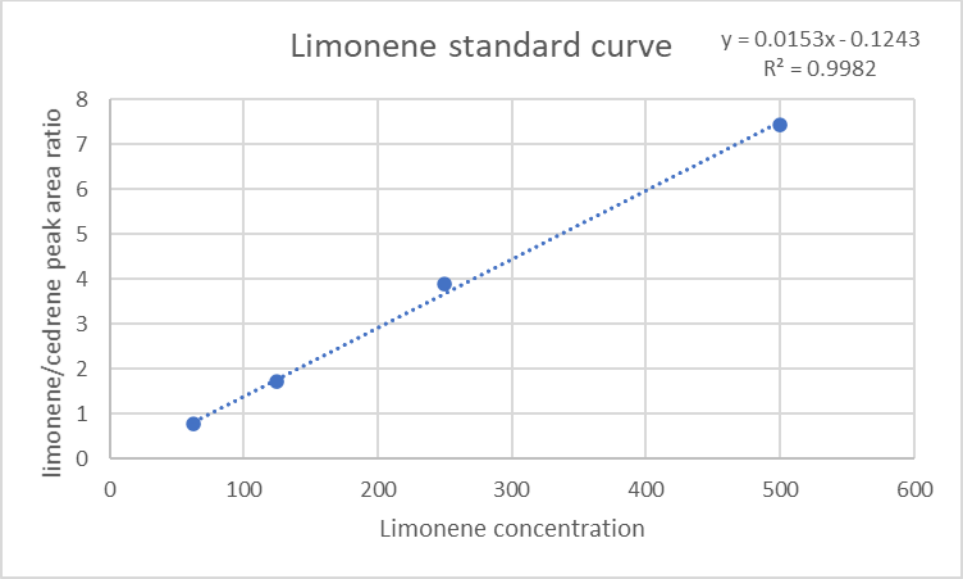

Figure A-1 Standard curve for limonene. A regression trendline was made based on theconcentration and the peak height ratio of limonene and internal standard cedrene

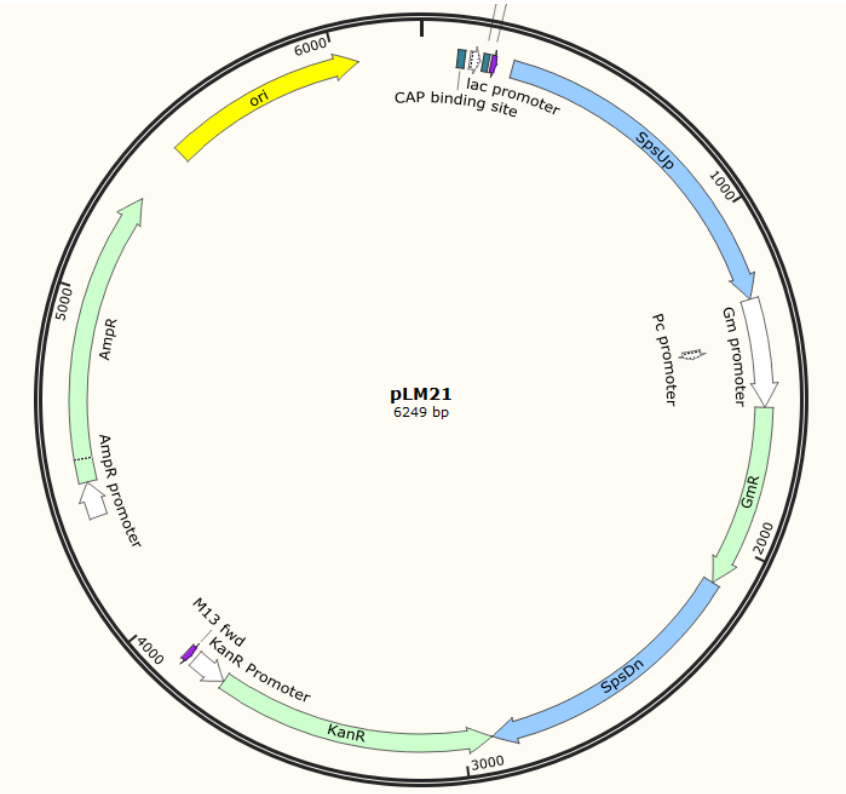

Figure A-2: Plasmid map and sequence of pLM21

Sequences:
GCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTG

GCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGA
GTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTT
GTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATT
ACGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCACAAGCTGTTTTACGCA
ATTTTTGCGACACTGCCCGATTTGGTGACGGAACTGGTGGAGGGTATTCCGCCCGAG
GCAAAGTACCGGTTTTCGGCACCGGTGGTGAAGGCGCAAGAGTTTCGGCTAGATGG
ATTGCTGGAGCCGATCGAAGAAGGAAGCGGCTATCCAACAGTGTTTTTAGAAGCGC
AGATGCAGAGCGATCGCGGCTTTTACAGTCGCTACTTTGCAGAGTTATTTGGTTACA
TCCGTCAGTATCCCGAGGCAGCCAATTGGAGAGGGCTATTGCTGATCAGGGCGCGA
TCGCTGGACTTAGGTGAAGAGTCGAGTTTTACGGAACTGCTGCAGGGACGAGTGCA
GCGGCTGTATCTAGAGGATTTGATCGGGCGATCGCTGAATTCAGTGCGATTGCAGTT
GCTGAAGCTATTGGTCGTACCAGTGGCAGCGATAGGAGAGACGGCGAGAGCGGTGT
TGTTGTCAGCAGAGGATGAGGATGCGTTTGAGCAGCTACTCGAACTGGTGGAGGCT
ATAGTGGCAAGTAAGTTGCCTCAGCTTGAGATTGAGGAGATTCATCAAATGCTCGGT
ATTGAAGTCTCCGACTTAAAACAAACCCGCCTCTATCAAAGCGCGGTGGAAGAAGG
TGAGAAAAAGGGACGCCAAGAAGGTGAGCTGGCTCTGGTCTTGAGGCAGCTGCAGC
GGCGATTTGAGAACTTAACGGAAGAGCAGGAGCAGCAGATTCGATCGCTGTCATTG
GAGGCGATCGAAGCCTTGAGTCTTGATTGGCTCGACTTTGAGACTGTCGCTGATTTA
AATCGCTGGCTCCAAAAACATTAAATGCATGGGATTGCAAAAAAGACCTTAAGCCC
ACGCGCTTCTCTCAGCCCGTCGTGAAAAAGACTGCAATATCGCACTGAATACTTTGG
CTGCCCGCTCAAGTGGAATGGTCAAGATTGGCAATCGCTACCTATTGCAGCAACGCG
ATCGCCCAATCTTTAACTGAAAAGGCGTGGAGACCGAAACCTTGCGCTCGTTCGCCA
GCCAGGACAGAAATGCCTCGACTTCGCTGCTGCCCAAGGTTGCCGGGTGACGCACA
CCGTGGAAACGGATGAAGGCACGAACCCAGTTGACATAAGCCTGTTCGGTTCGTAA
ACTGTAATGCAAGTAGCGTATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAAC
GCAGCGGTGGTAACGGCGCAGTGGCGGTTTTCATGGCTTGTTATGACTGTTTTTTTGT
ACAGTCTATGCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTCGATGTTT
GATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGT
TAGGTGGCTCAAGTATGGGCATCATTCGCACATGTAGGCTCGGCCCTGACCAAGTCA
AATCCATGCGGGCTGCTCTTGATCTTTTCGGTCGTGAGTTCGGAGACGTAGCCACCT
ACTCCCAACATCAGCCGGACTCCGATTACCTCGGGAACTTGCTCCGTAGTAAGACAT
TCATCGCGCTTGCTGCCTTCGACCAAGAAGCGGTTGTTGGCGCTCTCGCGGCTTACG
TTCTGCCCAAGTTTGAGCAGCCGCGTAGTGAGATCTATATCTATGATCTCGCAGTCT
CCGGCGAGCACCGGAGGCAGGGCATTGCCACCGCGCTCATCAATCTCCTCAAGCAT
GAGGCCAACGCGCTTGGTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACGAT
CCCGCAGTGGCTCTCTATACAAAGTTGGGCATACGGGAAGAAGTGATGCACTTTGAT
ATCGACCCAAGTACCGCCACCTAAAAGCGGCCAGTCTTGCTCCAACAGCTTTTTTAG
CTTGATCCGCCCCAGAGAATGCCCTGTGAAAATTAGCGGTACATTCAACCAGCGACT
CAGCAGTGATCCCACTTGGCCAGCATCAGCATAGTGGGCCTGAATCCAAGTCGGGG
TGCGCTTTTGCTGAGCCAGATATTGGAGAATTGCATCCGCAAAGGTGTAGAGATGGG
GCCAAAGCAGCTCTTTACGGAGGTAGCGTTTGGGGCCAAAAGGCAAACGGACAATC
CGACCTTTGGGCGCAAAGGGTTCGATCGCCTGACTGTAACCAACACTGACGCGGGG
GTCGGTGATTTGGCGGGTGATGATGTCGACTTGTTGGACTTGTGGGGATTTAGCTTG
GGCTTGAGCCAGTTCTAAGACGTACTTGGTCTGCCCGCCGGTGTCGGCATCTCGCCC
CAGTTCCAAGTTCTGCCCTCGCAGCAGACCATGGGTCTGAATGTGCAGAATGTAGAG
ATTTTGAGCTGCCACGCGCTAGTCAGCCTCACAGAAAACGCCCCCATCCTAGTCTCC

GATCCGAAAAGGTGGGGTGGTAGCGTAGGGGCGTTGCATGGCTGGTATCGATGTTTT
CGAGCCGCTCACCAAGGATTATGGGCCCGCGCCGACCCGCGATCGCGATCGCCCTC
GGTCTGCTGCTGTTAGCTTTCCTGATCTTGGTGGGGTTGAGTCTTGGCTCCTCAACAA
GTCTGATGTCTCACGATGAGGGCTATTACGCCCTGCAGGCCCGTTGGATTGTGGAAA
CGGGTGATTGGGTAACGCCGCGCTGGTGGTTAGAAAAACTCATCGAGCATCAAATG
AAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTC
TGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTA
TCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCA
AAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAA
TGGCAAAAGCTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTC
GTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGC
GAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCA
ACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATT
CTTCTAATACCTGGAATGCTGTTTTCCCGGGGATCGCAGTGGTGAGTAACCATGCAT
CATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGC
CAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTT
TCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAATCGATAGATTGTCGCACCTG
ATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGG
AATTTAATCGCGGCCTCGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCC
TTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATATTTTTATCT
TGTGCAATGTAACATCAGAGATTTTGAGACACAACGTGGGAGCTCGAATTCACTGGC
CGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCT
TGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATC
GCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTTC
TCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTG
CTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCC
CTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGG
GAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGACGAAAGG
GCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGAC
GTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAA
ATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAA
TATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTT
TTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAG
ATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGC
GGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTT
AAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTC
GGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAA
AAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCAT
GAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGC
TAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAAC
CGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCA
ATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGG
CAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTC
GGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTC
TCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTAT
CTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGA

TAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATAC
TTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTT
TGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGA
CCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGC
TGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGA
GCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATAC
TGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCC
TACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTC
GTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGG
GCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAA
CTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAA
GGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAG
CTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGA
CTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCC
AGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCT
TTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA
TACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCG
GAAGA

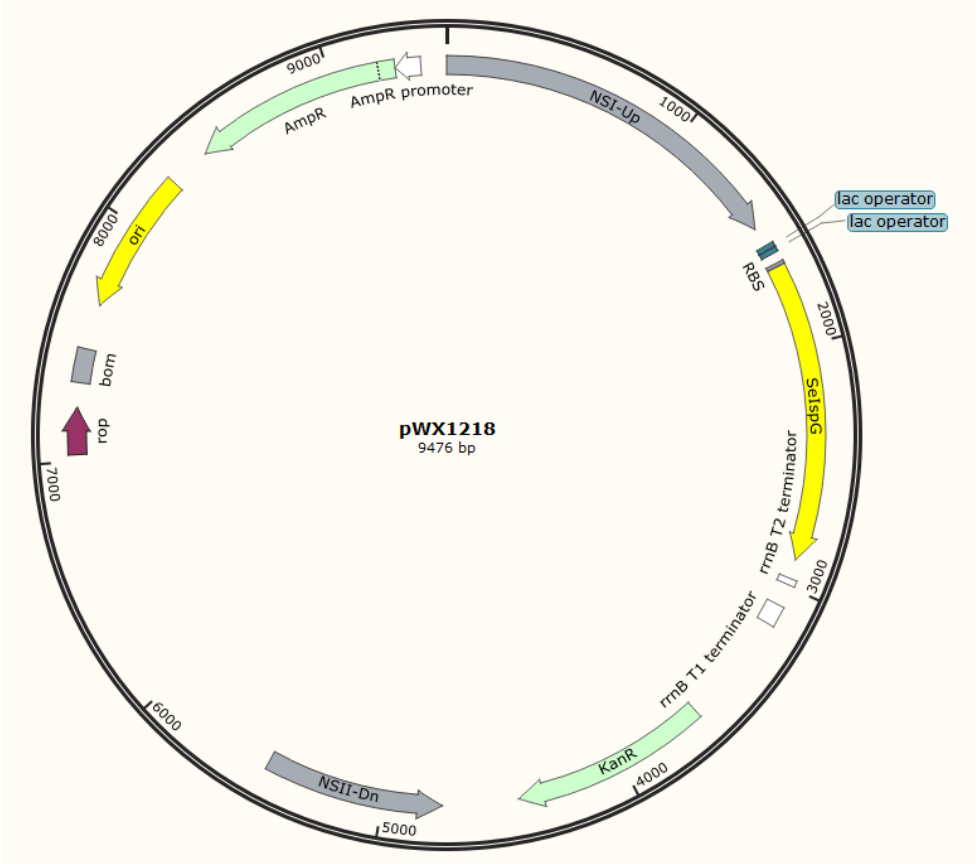

Figure A-3 Plasmid map and sequence of pWX1218

Sequence:

AGCTTGTCATCTGCCGGATGAGGCAAAACCCTGCCTACGGCGCGATTACATCGTCCC
AGCGCGATCGCTCTTACTGTTGATGGCTCGTGCTTAAAAACAATGCAAACTTCACCG
TTTCAGCTGGTGATTTTCGACTGTGATGGTGTGCTTGTTGATAGCGAACGCATCACTA
ATCGCGTCTTTGCAGACATGCTCAATGAACTGGGTCTGTTGGTGACTTTGGATGACA
TGTTTGAGCAGTTTGTGGGTCATTCCATGGCTGACTGTCTCAAACTAATTGAGCGAC
GGTTAGGCAATCCTCCACCCCCTGACTTTGTTCAGCACTATCAACGCCGTACCCGTA
TCGCGTTAGAAACGCATCTACAAGCCGTTCCTGGGGTTGAAGAGGCTTTGGATGCTC
TTGAATTGCCCTACTGTGTTGCGTCCAGTGGTGATCATCAAAAGATGCGAACCACAC
TGAGCCTGACGAAGCTCTGGCCACGATTTGAGGGACGAATCTTCAGCGTGACTGAA
GTACCTCGCGGCAAGCCATTTCCCGATGTCTTTTTGTTGGCCGCCGATCGCTTCGGGG
TTAATCCTACGGCCTGCGCTGTGATCGAAGACACCCCCTTGGGAGTAGCGGCAGGC
GTGGCGGCAGGAATGCAAGTGTTTGGCTACGCGGGTTCCATGCCCGCTTGGCGTCTG
CAAGAAGCCGGTGCCCATCTCATTTTTGACGATATGCGACTGCTGCCCAGTCTGCTC
CAATCGTCGCCAAAAGATAACTCCACAGCATTGCCCAATCCCTAACCCCTGCTCGCG
CCGCAACTACACACTAAACCGTTCCTGCGCGATCGCTCTTACTGTTGATGGCTCGTG
CTTAAAAACAATGCAACCCTAACCGTTTCAGCTGGTGATTTTCGGACGATTTGGCTT
ACAGGGATAACTGAGAGTCAACAGCCTCTGTCCGTCATTGCACACCCATCCATGCAC
TGGGGACTTGACTCATGCTGAATCACATTTCCCTTGTCCATTGGGCGAGAGGGGAGG
GGAATCTTCTGGACTCTTCACTAAGCGGCGATCGCAGGTTCTTCTACCCAAGCAGTG
GCGATCGCTTGATTGCAGTCTTCAATGCTGGCCTCTGCAGCCATCGCCGCCACCAAA
GCATCGTAGGCGGGACGTTGTTGCTCCAGTAAAGTCTTCGCCCGTAACAATCCCCAG
CGACTGCGTAAATCCGCTTCGGCAGGATTGCGATCGAGTTGCCGCCACAGTTGTTTC
CACTGGGCGCGATCGTCAGCTCCCCCTTCCACGTTGCCGTAGACCAGTTGCTCTGCC
GCTGCACCGGCCATCAACACCTGACACCACTGTTCCAGCGATCGCTGACTGAGTTGC
CCCTGTGCGGCTTCGGCTTCTAGCGCAGCTGCTTGGAACTGCACACCCCCGCGACCA
GGTTGTCCTTGGCGCAGCGCTTCCCACGCTGAGAGGGTGTAGCCCGTCACGGGTAAC
CGATTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAG
GCCCTTTCGTCTTCACCTCGAGAATTGTGAGCGGATAACAATTGACATTGTGAGCGG
ATAACAAGATACTGAGCACATCAGCAGGACGCACTGACCGAATTCATTAAAGAGGA
GAAAGATATCATGCAGACCCTCTCCACCCCCAGCACAACTGCCACTGAGTTTGATAC
CGTCATCCACCGGCGCCCTACGCGATCGGTTCGGGTCGGTGATATCTGGATTGGCAG
CCGCCATCCCGTCGTTGTCCAGTCGATGATCAACGAGGACACCCTCGACATCGATGG
GTCTGTTGCCGCGATCCGCCGCCTGCATGAGATTGGCTGCGAGATCGTCCGCGTCAC
GGTGCCCAGTCTTGGCCATGCCAAAGCGGTCGGTGACATTAAAAAGAAACTGCAAG
ACACCTATCGCGACGTGCCCTTGGTTGCCGACGTGCACCACAACGGTATGAAGATCG
CGCTGGAAGTCGCCAAGCACGTTGACAAAGTGCGGATCAATCCTGGTCTCTACGTCT
TTGAGAAGCCCGATCCGAATCGTCAGGGCTACACACCAGAAGAATTTGAGCGAATT
GGCAAGCAAATTCGCGACACGCTTGAGCCCTTAGTCACCAGCCTGCGCGAGCAGGA
CAAGGCCATGCGGATTGGCGTCAACCACGGCTCTCTAGCCGAGCGGATGCTGTTCAC
CTACGGCGACACGCCGGAAGGCATGGTGGAATCGGCGTTGGAATTCCTGCGACTCT
GCGAAGAGATGGACTTCCGTAATCTCGTCATCTCCATGAAAGCGAGCCGAGCGCCG
GTGATGATGGCGGCCTATCGCCTGATGGCCAAACGGATGGATGATCTGGGCATGGA
TTATCCGCTCCACTTGGGTGTGACCGAAGCCGGTGATGGTGACTATGGCCGGATCAA
ATCGACGGTTGGGATTGGCACGCTGCTAGCGGAAGGGATTGGCGATACGATTCGCG
TTTCGCTCACGGAAGCCCCCGAAAACGAAATTCCGGTTTGCTATTCGATTTTGCAAG
CCCTCGGCCTCCGCAAAACCATGGTTGAGTATGTCGCCTGTCCAAGTTGCGGACGGA

CACTCTTCAATCTGGAAGAGGTCCTTCACAAAGTGCGGGCTGCAACCAATCATTTGG
TTGGCTTGGATATTGCTGTCATGGGTTGCATCGTCAATGGCCCTGGCGAAATGGCTG
ATGCTGACTACGGCTACGTTGGCAAAACCCCAGGAACGATCGCCCTTTACCGAGGA
CGAGATGAAATCAAACGAGTTCCCGAAGAGCAAGGTGTAGAAGAACTGATCAACCT
GATCAAAGCAGATGGACGCTGGGTTGAACCGGAACCGATTGCCTAAGTCGACAGGC
CTCTAGACCCGGGCTCGAGCTAGCAAGCTTGGCCGGATCCGGCCGGATCCGGAGTTT
GTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAATTTGATGCCTGG
CAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGGGCCGTTGCTTCGCAACGTTCA
AATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAACAG
ATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCA
GTTCCCTACTCTCGCATGGGGAGACCCCACACTACCATCGGCGCTACGGCGTTTCAC
TTCTGAGTTCGGCATGGGGTCAGGTGGGACCACCGCGCTACTGCCGCCAGGCAAATT
CTGTTTTATCAGCCGTTACCCCACCTACTAGCTAATCCCATCTGGGCACATCCGATG
GCAAGAGGCCCGAAGGTCCCCCTCTTTGGTCTTGCGACGTTATGCGGTATTAGCTAC
CGTTTCCAGTAGTTATCCCCCTCCATCAGGCAGTTTCCCAGACATTACTCACCCGTCC
GCCACTCGTCAGCAAAGAAGCAAGCTTAGATCGACCTGCAGGGGGGGGGGGGAAA
GCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT
CATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGC
CATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGGATGCT
GATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATC
TATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGT
AGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTT
ATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCA
CCACTGCGATCCCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAG
GTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGT
TTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGA
ATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCT
GTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTC
GTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATA
GGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATC
CTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAAT
ATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTT
TTTCTAATCAGAATTGGTTAATTGGTTGTAACACTGGCAGAGCATTACGCTGACTTG
ACGGGACGGCGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGATCAGATC
ACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAAATCAC
CAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCT
GGATGATGGGGCGATTCAGGCCTGGTATGAGTCAGCAACACCTTCTTCACGAGGCA
GACCTCAGCGCCCCCCCCCCCCTGCAGGTCGATCTGGTAACCCCAGCGCGGTTGCTA
CCAAGTAGTGACCCGCTTCGTGATGCAAAATCCGCTGACGATATTCGGGCGATCGCT
GCTGAATGCCATCGAGCAGTAACGTGGCACCCCGCCCCTGCCAAGTCACCGCATCC
AGACTGAACAGCACCAAGAGGCTAAAACCCAATCCCGCCGGTAGCAGCGGAGAACT
ACCCAGCATTGGTCCCACCAAAGCTAATGCCGTCGTGGTAAAAATCGCGATCGCCGT
CAGACTCAAGCCCAGTTCGCTCATGCTTCCTCATCTAGGTCACAGTCTTCGGCGATC
GCATCGATCTGATGCTGCAGCAAGCGTTTTCCATACCGGCGATCGCGCCGTCGCCCT
TTCGCTGCCGTGGCCCGCTTACGAGCTCGTTTATCGACCACGATCGCATCCAAATCC
GCGATCGCTTCCCAGTCCGGCAATTCAGTCTGGGGCGTCCGTTTCATTAATCCTGATC

AGGCACGAAATTGCTGTGCGTAGTATCGCGCATAGCGGCCAGCCTCTGCCAACAGC
GCATCGTGATTGCCTGCCTCAACAATCTGGCCGCGCTCCATCACCAAGATGCGGCTG
GCATTACGAACCGTAGCCAGACGGTGAGCAATGATAAAGACCGTCCGTCCCTGCAT
CACCCGTTCTAGGGCCTCTTGCACCAAGGTTTCGGACTCGGAATCAAGCGCCGAAGT
CGCCTCATCCAGAATTAAAATGCGTGGATCCTCTACGCCGGACGCATCGTGGCCGGC
ATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGG
GGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGT
GGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGCATGCACCATTCCT
TGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATGCAGGA
GTCGCATAAGGGAGAGCGTCGATCGACCGATGCCCTTGAGAGCCTTCAACCCAGTC
AGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTC
TTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCATTTTCGGCGAG
GACCGCTTTCGCTGGAGCGCGACGATGATCGGCCTGTCGCTTGCGGTATTCGGAATC
TTGCACGCCCTCGCTCAAGCCTTCGTCACTGGTCCCGCCACCAAACGTTTCGGCGAG
AAGCAGGCCATTATCGCCGGCATGGCGGCCGACGCGCTGGGCTACGTCTTGCTGGC
GTTCGCGACGCGAGGCTGGATGGCCTTCCCCATTATGATTCTTCTCGCTTCCGGCGG
CATCGGGATGCCCGCGTTGCAGGCCATGCTGTCCAGGCAGGTAGATGACGACCATC
AGGGACAGCTTCAAGGATCGCTCGCGGCTCTTACCAGCCTAACTTCGATCACTGGAC
CGCTGATCGTCACGGCGATTTATGCCGCCTCGGCGAGCACATGGAACGGGTTGGCAT
GGATTGTAGGCGCCGCCCTATACCTTGTCTGCCTCCCCGCGTTGCGTCGCGGTGCAT
GGAGCCGGGCCACCTCGACCTGAATGGAAGCCGGCGGCACCTCGCTAACGGATTCA
CCACTCCAAGAATTGGAGCCAATCAATTCTTGCGGAGAACTGTGAATGCGCAAACC
AACCCTTGGCAGAACATATCCATCGCGTCCGCCATCTCCAGCAGCCGCACGCGGCGC
ATCTCGGGCAGCGTTGGGTCCTGGCCACGGGTGCGCATGATCGTGCTCCTGTCGTTG
AGGACCCGGCTAGGCTGGCGGGGTTGCCTTACTGGTTAGCAGAATGAATCACCGAT
ACGCGAGCGAACGTGAAGCGACTGCTGCTGCAAAACGTCTGCGACCTGAGCAACAA
CATGAATGGTCTTCGGTTTCCGTGTTTCGTAAAGTCTGGAAACGCGGAAGTCAGCGC
CCTGCACCATTATGTTCCGGATCTGCATCGCAGGATGCTGCTGGCTACCCTGTGGAA
CACCTACATCTGTATTAACGAAGCGCTGGCATTGACCCTGAGTGATTTTTCTCTGGTC
CCGCCGCATCCATACCGCCAGTTGTTTACCCTCACAACGTTCCAGTAACCGGGCATG
TTCATCATCAGTAACCCGTATCGTGAGCATCCTCTCTCGTTTCATCGGTATCATTACC
CCCATGAACAGAAATCCCCCTTACACGGAGGCATCAGTGACCAAACAGGAAAAAAC
CGCCCTTAACATGGCCCGCTTTATCAGAAGCCAGACATTAACGCTTCTGGAGAAACT
CAACGAGCTGGACGCGGATGAACAGGCAGACATCTGTGAATCGCTTCACGACCACG
CTGATGAGCTTTACCGCAGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCT
GACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGC
AGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCAT
GACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGA
GCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAA
GGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCT
CGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTAT
CCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAA
GGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC
TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC
TATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGA
CCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTC

TCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGG
CTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG
TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGC
CTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAG
TTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTA
GCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAG
AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTT
AAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT
AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTT
ACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCAT
AGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGG
CCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC
AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCG
CCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTA
ATAGTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTCACGCTCGTCGTT
TGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCC
CATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAA
GTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC
ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGA
GAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACC
GCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA
AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA
CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACA
GGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATAC
TCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAG
CGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATT
TCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTA
TAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAAGAATT

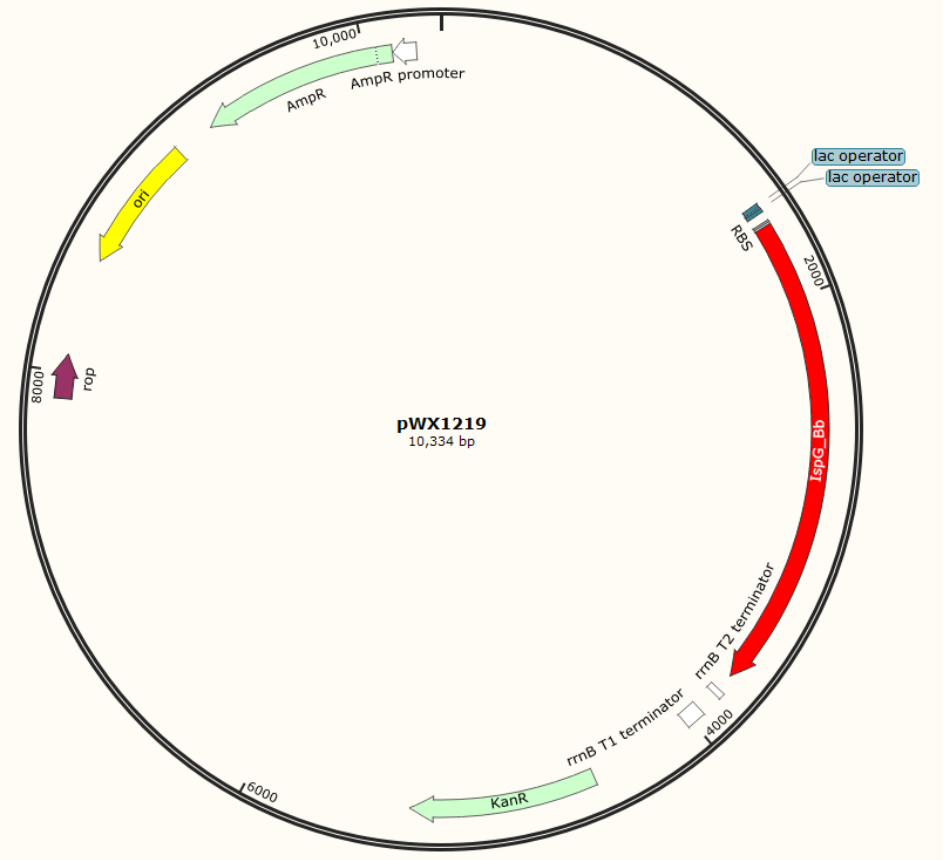

Figure A-4 Plasmid map and sequence of pWX1219

Sequence:

AGCTTGTCATCTGCCGGATGAGGCAAAACCCTGCCTACGGCGCGATTACATCGTCCC
AGCGCGATCGCTCTTACTGTTGATGGCTCGTGCTTAAAAACAATGCAAACTTCACCG
TTTCAGCTGGTGATTTTCGACTGTGATGGTGTGCTTGTTGATAGCGAACGCATCACTA
ATCGCGTCTTTGCAGACATGCTCAATGAACTGGGTCTGTTGGTGACTTTGGATGACA
TGTTTGAGCAGTTTGTGGGTCATTCCATGGCTGACTGTCTCAAACTAATTGAGCGAC
GGTTAGGCAATCCTCCACCCCCTGACTTTGTTCAGCACTATCAACGCCGTACCCGTA
TCGCGTTAGAAACGCATCTACAAGCCGTTCCTGGGGTTGAAGAGGCTTTGGATGCTC
TTGAATTGCCCTACTGTGTTGCGTCCAGTGGTGATCATCAAAAGATGCGAACCACAC
TGAGCCTGACGAAGCTCTGGCCACGATTTGAGGGACGAATCTTCAGCGTGACTGAA
GTACCTCGCGGCAAGCCATTTCCCGATGTCTTTTTGTTGGCCGCCGATCGCTTCGGGG
TTAATCCTACGGCCTGCGCTGTGATCGAAGACACCCCCTTGGGAGTAGCGGCAGGC
GTGGCGGCAGGAATGCAAGTGTTTGGCTACGCGGGTTCCATGCCCGCTTGGCGTCTG
CAAGAAGCCGGTGCCCATCTCATTTTTGACGATATGCGACTGCTGCCCAGTCTGCTC
CAATCGTCGCCAAAAGATAACTCCACAGCATTGCCCAATCCCTAACCCCTGCTCGCG
CCGCAACTACACACTAAACCGTTCCTGCGCGATCGCTCTTACTGTTGATGGCTCGTG
CTTAAAAACAATGCAACCCTAACCGTTTCAGCTGGTGATTTTCGGACGATTTGGCTT
ACAGGGATAACTGAGAGTCAACAGCCTCTGTCCGTCATTGCACACCCATCCATGCAC
TGGGGACTTGACTCATGCTGAATCACATTTCCCTTGTCCATTGGGCGAGAGGGGAGG
GGAATCTTCTGGACTCTTCACTAAGCGGCGATCGCAGGTTCTTCTACCCAAGCAGTG
GCGATCGCTTGATTGCAGTCTTCAATGCTGGCCTCTGCAGCCATCGCCGCCACCAAA

GCATCGTAGGCGGGACGTTGTTGCTCCAGTAAAGTCTTCGCCCGTAACAATCCCCAG
CGACTGCGTAAATCCGCTTCGGCAGGATTGCGATCGAGTTGCCGCCACAGTTGTTTC
CACTGGGCGCGATCGTCAGCTCCCCCTTCCACGTTGCCGTAGACCAGTTGCTCTGCC
GCTGCACCGGCCATCAACACCTGACACCACTGTTCCAGCGATCGCTGACTGAGTTGC
CCCTGTGCGGCTTCGGCTTCTAGCGCAGCTGCTTGGAACTGCACACCCCCGCGACCA
GGTTGTCCTTGGCGCAGCGCTTCCCACGCTGAGAGGGTGTAGCCCGTCACGGGTAAC
CGATTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAG
GCCCTTTCGTCTTCACCTCGAGAATTGTGAGCGGATAACAATTGACATTGTGAGCGG
ATAACAAGATACTGAGCACATCAGCAGGACGCACTGACCGAATTCATTAAAGAGGA
GAAAgatatcATGGCTACTGCCACACAAGATCCCACACTTGTTCTTGATCCATCTAAGAT
TGTGCTGCCTAAATATTGCGAGAGCACCTACAAGACTGTCCGGCGGCCAACACGGA
CGATCTATATTGGCAAAGTCCCAGTAGGATCGGAGCACCCTATCCGCTTGCAGACAA
TGACCACAACAGACACCAGGAATGTGGAAGCAACTGTTGATCAAGTAAAGAAGTGT
GTCGATGTCGGTGCGGACTTCGTGCGCATCACCGTGCAGGGCAAGAAGGAGGCAGA
TGCTTGCATGCAGATCAGGGAGCGCCTCTTCAAAGACAGGTATGACGTACCTCTTGT
CGCAGACATCCACTTCCAGCCCAAGGTTGCCCTAGCTGTTGCAGAAGCCTTCGAGAA
GATCCGCATCAACCCAGGCAATTTTGCCGATGGTCGGAAGTCGTTCGAGGTCATCAA
CTACGATGATCCTGCCCAATTCACCCGAGAGCAGGAGGAAATCCGTGAGCTGTTCA
CCCCGCTAGTGGAGAAGTGCAAGTCTTTGAACCGTGCTATCCGTATTGGGACCAACC
ATGGCTCTCTGTCTGCCCGCATCCTGTCGTACTATGGGGACACCCCCAGAGGAATGG
TAAATTCTGCCTTCGAATTTGCTGAGATTTGCCGGGACTTGGACTTCCACAACTTCTT
GTTCAGCATGAAGGCCAGCAACCCTCTGGTCATGGTCCAGGCATACCGCCTCCTCGC
TGAGGAAATGTACAACAAGGGCTGGGACTACCCACTGCATCTGGGTGTTACTGAGG
CTGGTGAAGGAGAGGATGGCCGAATGAAGTCTGCCATTGGTATTGGATCTCTTTTAA
TGGATGGCCTGGGTGACACAATCCGTGTCTCCCTCACTGAGGACCCTGAGTTTGAGA
TTGTTCCTTGCGGGCAATTGGCCAACTTTGGGGCGCGGGCAGCCAAAGAGAACTGG
GGTGTTGCGCCCTTCACAGAGGTCTACAGGGACACACACTCTTTCGCCCGCCGCACA
GGCCGCCTGCCAGAGCAGCGGGAAGGTGACCAGTTGGATTTCCGCTCCCTGCTGCAT
CGCGATGGATCCGTGTTGTCCTTCGTACCTCTGGAGGAGCTGAAGCAGCCAGAGCTC
TTGTACCGCAAACTGGGAGCCAAGCTTGTGGTTGGCATGCCTTTCAAGGACCTGGCT
ACGTCTGACACTTTGATGCTGCCTGAGGTGCCGCCATCTTCCGATGTGGAGGGCCGC
CGCGCTCTGGCCCGCCTCCAGGAGGTAGGAGTGCATGTTGTCGCACCGCTGGCAGCC
TTGGCAAAGGACCCACTGCCCAACGCAGTGGCATTGGTCTCCCTTCGGGATTATGCC
GCAAAGGGAGTGACCCTCCCGGAGGGTGCTGCTAGGTTTGCCATCACAGCGGATGG
TACTGAATCGGACGCTGAGGTGGAAGCCTTGAAGGACTCCAAAGCAGTCATGGTGC
TGCTGGACGTGAAGAAGGGTGTCAGTCGGGTGCATGCTTCTCGCCACTTCTTCGAGT
TGTTGCGGTTCTATGACATTGACCTCCCTGTCATCCACACTCGCCGTTTCCCTGCAGG
ATTGACCAGGGATGAGATTGTGCTGACATGCGGAGCGGAGATTGGAGCCCTGTTGG
TGGATGGCCTCGGCGATGGTGTTCTATTGGACTGCCCCACAGAGGACCTAGACTTCA
TTCGGACAATGTCTTTCGGTCTCCTGCAGGGCTCACGCATGAGGAACACAAAGACAG
AGTACGTATCATGCCCAAGCTGCGGGCGCACGCTGTTCGACCTCCAGGAGGTGACT
GAGCAGATCAGGGTACGCACCGGCCACCTGCCAGGTGTTTCCATCGCCATTATGGGC
TGCATTGTCAACGGCCCTGGTGAAATGGCGGATGCTGACTTCGGCTATGTAGGAGGC
CGCCCCGGCCTCATCGACCTGTATGTTGGCAAAGAGGTTGTGAGGAAGAATATCCCC
ATGGAAGATGCCACCAATCAGCTTATTGAGCTCATCAAGGAATATGGCCGCTGGGT
GGACAAGGAGGAGGAGGAGGTTGCTCAAGAACTGCAGGTTGCATAAGTCGACAGG

CCTCTAGACCCGGGCTCGAGCTAGCAAGCTTGGCCGGATCCGGCCGGATCCGGAGT
TTGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAATTTGATGCCT
GGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGGGCCGTTGCTTCGCAACGTT
CAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAAC
AGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGG
CAGTTCCCTACTCTCGCATGGGGAGACCCCACACTACCATCGGCGCTACGGCGTTTC
ACTTCTGAGTTCGGCATGGGGTCAGGTGGGACCACCGCGCTACTGCCGCCAGGCAA
ATTCTGTTTTATCAGCCGTTACCCCACCTACTAGCTAATCCCATCTGGGCACATCCGA
TGGCAAGAGGCCCGAAGGTCCCCCTCTTTGGTCTTGCGACGTTATGCGGTATTAGCT
ACCGTTTCCAGTAGTTATCCCCCTCCATCAGGCAGTTTCCCAGACATTACTCACCCGT
CCGCCACTCGTCAGCAAAGAAGCAAGCTTAGATCGACCTGCAGGGGGGGGGGGGGA
AAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATC
ATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGA
GCCATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGGATG
CTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAA
TCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAG
GTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAA
TTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTAC
TCACCACTGCGATCCCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATT
CAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTC
CTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATC
ACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTG
GCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTC
AGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTA
ATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCC
ATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAA
AATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATG
AGTTTTTCTAATCAGAATTGGTTAATTGGTTGTAACACTGGCAGAGCATTACGCTGA
CTTGACGGGACGGCGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGATCA
GATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAAA
TCACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTTCT
GGCTGGATGATGGGGCGATTCAGGCCTGGTATGAGTCAGCAACACCTTCTTCACGAG
GCAGACCTCAGCGCCCCCCCCCCCCTGCAGGTCGATCTGGTAACCCCAGCGCGGTTG
CTACCAAGTAGTGACCCGCTTCGTGATGCAAAATCCGCTGACGATATTCGGGCGATC
GCTGCTGAATGCCATCGAGCAGTAACGTGGCACCCCGCCCCTGCCAAGTCACCGCAT
CCAGACTGAACAGCACCAAGAGGCTAAAACCCAATCCCGCCGGTAGCAGCGGAGA
ACTACCCAGCATTGGTCCCACCAAAGCTAATGCCGTCGTGGTAAAAATCGCGATCGC
CGTCAGACTCAAGCCCAGTTCGCTCATGCTTCCTCATCTAGGTCACAGTCTTCGGCG
ATCGCATCGATCTGATGCTGCAGCAAGCGTTTTCCATACCGGCGATCGCGCCGTCGC
CCTTTCGCTGCCGTGGCCCGCTTACGAGCTCGTTTATCGACCACGATCGCATCCAAA
TCCGCGATCGCTTCCCAGTCCGGCAATTCAGTCTGGGGCGTCCGTTTCATTAATCCTG
ATCAGGCACGAAATTGCTGTGCGTAGTATCGCGCATAGCGGCCAGCCTCTGCCAAC
AGCGCATCGTGATTGCCTGCCTCAACAATCTGGCCGCGCTCCATCACCAAGATGCGG
CTGGCATTACGAACCGTAGCCAGACGGTGAGCAATGATAAAGACCGTCCGTCCCTG
CATCACCCGTTCTAGGGCCTCTTGCACCAAGGTTTCGGACTCGGAATCAAGCGCCGA
AGTCGCCTCATCCAGAATTAAAATGCGTGGATCCTCTACGCCGGACGCATCGTGGCC

GGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGAT
GGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATG
GTGGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGCATGCACCATTC
CTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATGCAG
GAGTCGCATAAGGGAGAGCGTCGATCGACCGATGCCCTTGAGAGCCTTCAACCCAG
TCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCT
TCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCATTTTCGGCG
AGGACCGCTTTCGCTGGAGCGCGACGATGATCGGCCTGTCGCTTGCGGTATTCGGAA
TCTTGCACGCCCTCGCTCAAGCCTTCGTCACTGGTCCCGCCACCAAACGTTTCGGCG
AGAAGCAGGCCATTATCGCCGGCATGGCGGCCGACGCGCTGGGCTACGTCTTGCTG
GCGTTCGCGACGCGAGGCTGGATGGCCTTCCCCATTATGATTCTTCTCGCTTCCGGC
GGCATCGGGATGCCCGCGTTGCAGGCCATGCTGTCCAGGCAGGTAGATGACGACCA
TCAGGGACAGCTTCAAGGATCGCTCGCGGCTCTTACCAGCCTAACTTCGATCACTGG
ACCGCTGATCGTCACGGCGATTTATGCCGCCTCGGCGAGCACATGGAACGGGTTGG
CATGGATTGTAGGCGCCGCCCTATACCTTGTCTGCCTCCCCGCGTTGCGTCGCGGTG
CATGGAGCCGGGCCACCTCGACCTGAATGGAAGCCGGCGGCACCTCGCTAACGGAT
TCACCACTCCAAGAATTGGAGCCAATCAATTCTTGCGGAGAACTGTGAATGCGCAA
ACCAACCCTTGGCAGAACATATCCATCGCGTCCGCCATCTCCAGCAGCCGCACGCGG
CGCATCTCGGGCAGCGTTGGGTCCTGGCCACGGGTGCGCATGATCGTGCTCCTGTCG
TTGAGGACCCGGCTAGGCTGGCGGGGTTGCCTTACTGGTTAGCAGAATGAATCACCG
ATACGCGAGCGAACGTGAAGCGACTGCTGCTGCAAAACGTCTGCGACCTGAGCAAC
AACATGAATGGTCTTCGGTTTCCGTGTTTCGTAAAGTCTGGAAACGCGGAAGTCAGC
GCCCTGCACCATTATGTTCCGGATCTGCATCGCAGGATGCTGCTGGCTACCCTGTGG
AACACCTACATCTGTATTAACGAAGCGCTGGCATTGACCCTGAGTGATTTTTCTCTG
GTCCCGCCGCATCCATACCGCCAGTTGTTTACCCTCACAACGTTCCAGTAACCGGGC
ATGTTCATCATCAGTAACCCGTATCGTGAGCATCCTCTCTCGTTTCATCGGTATCATT
ACCCCCATGAACAGAAATCCCCCTTACACGGAGGCATCAGTGACCAAACAGGAAAA
AACCGCCCTTAACATGGCCCGCTTTATCAGAAGCCAGACATTAACGCTTCTGGAGAA
ACTCAACGAGCTGGACGCGGATGAACAGGCAGACATCTGTGAATCGCTTCACGACC
ACGCTGATGAGCTTTACCGCAGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACC
TCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGG
AGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGC
CATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCA
GAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGT
AAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCG
CTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTT
ATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAA
AAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC
CCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG
ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCC
GACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCT
TTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTG
GGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT
CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGT
AACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTG
GCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCC

AGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG
GTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTC
AAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC
GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAA
ATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACA
GTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATC
CATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATC
TGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATC
AGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTAT
CCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAG
TTAATAGTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTCACGCTCGTC
GTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATC
CCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAG
TAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACT
GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCT
GAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAAT
ACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG
CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGT
GCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAA
ACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAA
TACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATG
AGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCAC
ATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAAC
CTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAAGAATT

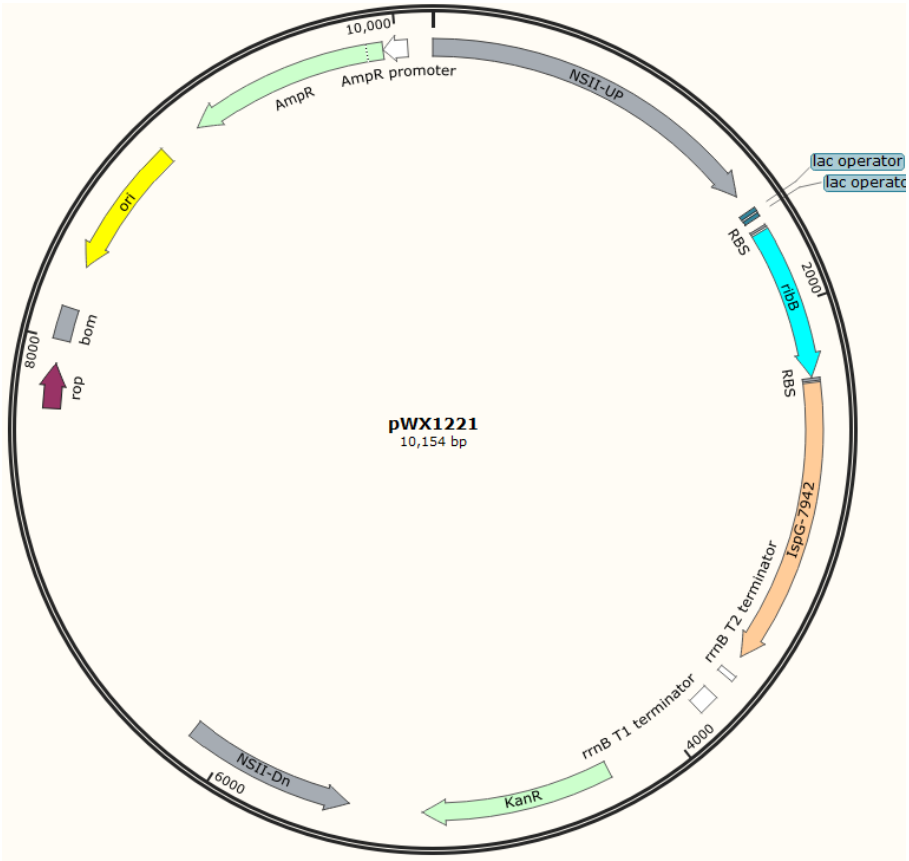

Figure A-5 Plasmid map and sequence of pWX1221

Sequence:

AGCTTGTCATCTGCCGGATGAGGCAAAACCCTGCCTACGGCGCGATTACATCGTCCC
AGCGCGATCGCTCTTACTGTTGATGGCTCGTGCTTAAAAACAATGCAAACTTCACCG
TTTCAGCTGGTGATTTTCGACTGTGATGGTGTGCTTGTTGATAGCGAACGCATCACTA
ATCGCGTCTTTGCAGACATGCTCAATGAACTGGGTCTGTTGGTGACTTTGGATGACA
TGTTTGAGCAGTTTGTGGGTCATTCCATGGCTGACTGTCTCAAACTAATTGAGCGAC
GGTTAGGCAATCCTCCACCCCCTGACTTTGTTCAGCACTATCAACGCCGTACCCGTA
TCGCGTTAGAAACGCATCTACAAGCCGTTCCTGGGGTTGAAGAGGCTTTGGATGCTC
TTGAATTGCCCTACTGTGTTGCGTCCAGTGGTGATCATCAAAAGATGCGAACCACAC
TGAGCCTGACGAAGCTCTGGCCACGATTTGAGGGACGAATCTTCAGCGTGACTGAA
GTACCTCGCGGCAAGCCATTTCCCGATGTCTTTTTGTTGGCCGCCGATCGCTTCGGGG
TTAATCCTACGGCCTGCGCTGTGATCGAAGACACCCCCTTGGGAGTAGCGGCAGGC
GTGGCGGCAGGAATGCAAGTGTTTGGCTACGCGGGTTCCATGCCCGCTTGGCGTCTG
CAAGAAGCCGGTGCCCATCTCATTTTTGACGATATGCGACTGCTGCCCAGTCTGCTC
CAATCGTCGCCAAAAGATAACTCCACAGCATTGCCCAATCCCTAACCCCTGCTCGCG
CCGCAACTACACACTAAACCGTTCCTGCGCGATCGCTCTTACTGTTGATGGCTCGTG
CTTAAAAACAATGCAACCCTAACCGTTTCAGCTGGTGATTTTCGGACGATTTGGCTT
ACAGGGATAACTGAGAGTCAACAGCCTCTGTCCGTCATTGCACACCCATCCATGCAC
TGGGGACTTGACTCATGCTGAATCACATTTCCCTTGTCCATTGGGCGAGAGGGGAGG
GGAATCTTCTGGACTCTTCACTAAGCGGCGATCGCAGGTTCTTCTACCCAAGCAGTG

GCGATCGCTTGATTGCAGTCTTCAATGCTGGCCTCTGCAGCCATCGCCGCCACCAAA
GCATCGTAGGCGGGACGTTGTTGCTCCAGTAAAGTCTTCGCCCGTAACAATCCCCAG
CGACTGCGTAAATCCGCTTCGGCAGGATTGCGATCGAGTTGCCGCCACAGTTGTTTC
CACTGGGCGCGATCGTCAGCTCCCCCTTCCACGTTGCCGTAGACCAGTTGCTCTGCC
GCTGCACCGGCCATCAACACCTGACACCACTGTTCCAGCGATCGCTGACTGAGTTGC
CCCTGTGCGGCTTCGGCTTCTAGCGCAGCTGCTTGGAACTGCACACCCCCGCGACCA
GGTTGTCCTTGGCGCAGCGCTTCCCACGCTGAGAGGGTGTAGCCCGTCACGGGTAAC
CGATTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAG
GCCCTTTCGTCTTCACCTCGAGAATTGTGAGCGGATAACAATTGACATTGTGAGCGG
ATAACAAGATACTGAGCACATCAGCAGGACGCACTGACCGAATTCATTAAAGAGGA
GAAAGATATCATGAACCAGACTTTGCTCAGCAGTTTTGGGACACCATTCGAACGAGT
GGAAAATGCGCTGGCGGCATTGCGGGAAGGACGGGGCGTCATGGTGCTCGATGATG
AGGATCGCGAAAATGAAGGTGACATGATCTTTCCTGCCGAAACCATGACCGTGGAG
CAAATGGCGTTGACCATTCGCCATGGGTCCGGTATTGTGTGTTTGTGCATCACTGAG
GATCGCCGGAAGCAGCTGGATTTGCCGATGATGGTCGAAAACAACACGTCGGCCTA
CGGCACGGGTTTTACAGTCACGATTGAAGCTGCGGAAGGTGTGACCACAAGCGTCT
CTGCGGCTGATCGTATTACAACCGTGCGCGCCGCCATCGCAGATGGTGCGAAGCCCT
CCGATCTGAATCGACCTGGTCACGTGTTCCCTCTGCGTGCCCAAGCGGGCGGGGTGC
TGACACGCGGCGGGCACACTGAAGCCACAATCGACCTGATGACGCTCGCCGGATTT
AAACCCGCGGGAGTTCTGTGCGAGCTGACGAATGACGATGGTACTATGGCCCGCGC
CCCCGAGTGTATTGAGTTCGCTAACAAACACAATATGGCACTGGTGACAATTGAAG
ATCTGGTCGCGTACCGCCAAGCACACGAGCGAAAAGCATCTTAAGTCGACAAAGAG
GAGAAAgatatcATGCAGACCCTCTCCACCCCCAGCACAACTGCCACTGAGTTTGATAC
CGTCATCCACCGGCGCCCTACGCGATCGGTTCGGGTCGGTGATATCTGGATTGGCAG
CCGCCATCCCGTCGTTGTCCAGTCGATGATCAACGAGGACACCCTCGACATCGATGG
GTCTGTTGCCGCGATCCGCCGCCTGCATGAGATTGGCTGCGAGATCGTCCGCGTCAC
GGTGCCCAGTCTTGGCCATGCCAAAGCGGTCGGTGACATTAAAAAGAAACTGCAAG
ACACCTATCGCGACGTGCCCTTGGTTGCCGACGTGCACCACAACGGTATGAAGATCG
CGCTGGAAGTCGCCAAGCACGTTGACAAAGTGCGGATCAATCCTGGTCTCTACGTCT
TTGAGAAGCCCGATCCGAATCGTCAGGGCTACACACCAGAAGAATTTGAGCGAATT
GGCAAGCAAATTCGCGACACGCTTGAGCCCTTAGTCACCAGCCTGCGCGAGCAGGA
CAAGGCCATGCGGATTGGCGTCAACCACGGCTCTCTAGCCGAGCGGATGCTGTTCAC
CTACGGCGACACGCCGGAAGGCATGGTGGAATCGGCGTTGGAATTCCTGCGACTCT
GCGAAGAGATGGACTTCCGTAATCTCGTCATCTCCATGAAAGCGAGCCGAGCGCCG
GTGATGATGGCGGCCTATCGCCTGATGGCCAAACGGATGGATGATCTGGGCATGGA
TTATCCGCTCCACTTGGGTGTGACCGAAGCCGGTGATGGTGACTATGGCCGGATCAA
ATCGACGGTTGGGATTGGCACGCTGCTAGCGGAAGGGATTGGCGATACGATTCGCG
TTTCGCTCACGGAAGCCCCCGAAAACGAAATTCCGGTTTGCTATTCGATTTTGCAAG
CCCTCGGCCTCCGCAAAACCATGGTTGAGTATGTCGCCTGTCCAAGTTGCGGACGGA
CACTCTTCAATCTGGAAGAGGTCCTTCACAAAGTGCGGGCTGCAACCAATCATTTGG
TTGGCTTGGATATTGCTGTCATGGGTTGCATCGTCAATGGCCCTGGCGAAATGGCTG
ATGCTGACTACGGCTACGTTGGCAAAACCCCAGGAACGATCGCCCTTTACCGAGGA
CGAGATGAAATCAAACGAGTTCCCGAAGAGCAAGGTGTAGAAGAACTGATCAACCT
GATCAAAGCAGATGGACGCTGGGTTGAACCGGAACCGATTGCCTAAGTCGACAGGC
CTCTAGACCCGGGCTCGAGCTAGCAAGCTTGGCCGGATCCGGCCGGATCCGGAGTTT
GTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAATTTGATGCCTGG

CAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGGGCCGTTGCTTCGCAACGTTCA
AATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAACAG
ATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCA
GTTCCCTACTCTCGCATGGGGAGACCCCACACTACCATCGGCGCTACGGCGTTTCAC
TTCTGAGTTCGGCATGGGGTCAGGTGGGACCACCGCGCTACTGCCGCCAGGCAAATT
CTGTTTTATCAGCCGTTACCCCACCTACTAGCTAATCCCATCTGGGCACATCCGATG
GCAAGAGGCCCGAAGGTCCCCCTCTTTGGTCTTGCGACGTTATGCGGTATTAGCTAC
CGTTTCCAGTAGTTATCCCCCTCCATCAGGCAGTTTCCCAGACATTACTCACCCGTCC
GCCACTCGTCAGCAAAGAAGCAAGCTTAGATCGACCTGCAGGGGGGGGGGGGAAA
GCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCAT
CATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGC
CATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGGATGCT
GATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATC
TATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGT
AGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTT
ATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCA
CCACTGCGATCCCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAG
GTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGT
TTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGA
ATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCT
GTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTC
GTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATA
GGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATC
CTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAAT
ATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTT
TTTCTAATCAGAATTGGTTAATTGGTTGTAACACTGGCAGAGCATTACGCTGACTTG
ACGGGACGGCGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGATCAGATC
ACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAAATCAC
CAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCT
GGATGATGGGGCGATTCAGGCCTGGTATGAGTCAGCAACACCTTCTTCACGAGGCA
GACCTCAGCGCCCCCCCCCCCCTGCAGGTCGATCTGGTAACCCCAGCGCGGTTGCTA
CCAAGTAGTGACCCGCTTCGTGATGCAAAATCCGCTGACGATATTCGGGCGATCGCT
GCTGAATGCCATCGAGCAGTAACGTGGCACCCCGCCCCTGCCAAGTCACCGCATCC
AGACTGAACAGCACCAAGAGGCTAAAACCCAATCCCGCCGGTAGCAGCGGAGAACT
ACCCAGCATTGGTCCCACCAAAGCTAATGCCGTCGTGGTAAAAATCGCGATCGCCGT
CAGACTCAAGCCCAGTTCGCTCATGCTTCCTCATCTAGGTCACAGTCTTCGGCGATC
GCATCGATCTGATGCTGCAGCAAGCGTTTTCCATACCGGCGATCGCGCCGTCGCCCT
TTCGCTGCCGTGGCCCGCTTACGAGCTCGTTTATCGACCACGATCGCATCCAAATCC
GCGATCGCTTCCCAGTCCGGCAATTCAGTCTGGGGCGTCCGTTTCATTAATCCTGATC
AGGCACGAAATTGCTGTGCGTAGTATCGCGCATAGCGGCCAGCCTCTGCCAACAGC
GCATCGTGATTGCCTGCCTCAACAATCTGGCCGCGCTCCATCACCAAGATGCGGCTG
GCATTACGAACCGTAGCCAGACGGTGAGCAATGATAAAGACCGTCCGTCCCTGCAT
CACCCGTTCTAGGGCCTCTTGCACCAAGGTTTCGGACTCGGAATCAAGCGCCGAAGT
CGCCTCATCCAGAATTAAAATGCGTGGATCCTCTACGCCGGACGCATCGTGGCCGGC
ATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGG
GGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGT

GGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGCATGCACCATTCCT
TGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATGCAGGA
GTCGCATAAGGGAGAGCGTCGATCGACCGATGCCCTTGAGAGCCTTCAACCCAGTC
AGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTC
TTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCATTTTCGGCGAG
GACCGCTTTCGCTGGAGCGCGACGATGATCGGCCTGTCGCTTGCGGTATTCGGAATC
TTGCACGCCCTCGCTCAAGCCTTCGTCACTGGTCCCGCCACCAAACGTTTCGGCGAG
AAGCAGGCCATTATCGCCGGCATGGCGGCCGACGCGCTGGGCTACGTCTTGCTGGC
GTTCGCGACGCGAGGCTGGATGGCCTTCCCCATTATGATTCTTCTCGCTTCCGGCGG
CATCGGGATGCCCGCGTTGCAGGCCATGCTGTCCAGGCAGGTAGATGACGACCATC
AGGGACAGCTTCAAGGATCGCTCGCGGCTCTTACCAGCCTAACTTCGATCACTGGAC
CGCTGATCGTCACGGCGATTTATGCCGCCTCGGCGAGCACATGGAACGGGTTGGCAT
GGATTGTAGGCGCCGCCCTATACCTTGTCTGCCTCCCCGCGTTGCGTCGCGGTGCAT
GGAGCCGGGCCACCTCGACCTGAATGGAAGCCGGCGGCACCTCGCTAACGGATTCA
CCACTCCAAGAATTGGAGCCAATCAATTCTTGCGGAGAACTGTGAATGCGCAAACC
AACCCTTGGCAGAACATATCCATCGCGTCCGCCATCTCCAGCAGCCGCACGCGGCGC
ATCTCGGGCAGCGTTGGGTCCTGGCCACGGGTGCGCATGATCGTGCTCCTGTCGTTG
AGGACCCGGCTAGGCTGGCGGGGTTGCCTTACTGGTTAGCAGAATGAATCACCGAT
ACGCGAGCGAACGTGAAGCGACTGCTGCTGCAAAACGTCTGCGACCTGAGCAACAA
CATGAATGGTCTTCGGTTTCCGTGTTTCGTAAAGTCTGGAAACGCGGAAGTCAGCGC
CCTGCACCATTATGTTCCGGATCTGCATCGCAGGATGCTGCTGGCTACCCTGTGGAA
CACCTACATCTGTATTAACGAAGCGCTGGCATTGACCCTGAGTGATTTTTCTCTGGTC
CCGCCGCATCCATACCGCCAGTTGTTTACCCTCACAACGTTCCAGTAACCGGGCATG
TTCATCATCAGTAACCCGTATCGTGAGCATCCTCTCTCGTTTCATCGGTATCATTACC
CCCATGAACAGAAATCCCCCTTACACGGAGGCATCAGTGACCAAACAGGAAAAAAC
CGCCCTTAACATGGCCCGCTTTATCAGAAGCCAGACATTAACGCTTCTGGAGAAACT
CAACGAGCTGGACGCGGATGAACAGGCAGACATCTGTGAATCGCTTCACGACCACG
CTGATGAGCTTTACCGCAGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCT
GACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGC
AGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCAT
GACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGA
GCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAA
GGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCT
CGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTAT
CCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAA
GGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC
TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC
TATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGA
CCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTC
TCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGG
CTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG
TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGC
CTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAG
TTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTA
GCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAG

AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTT
AAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT
AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTT
ACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCAT
AGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGG
CCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC
AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCG
CCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTA
ATAGTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTCACGCTCGTCGTT
TGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCC
CATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAA
GTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC
ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGA
GAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACC
GCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA
AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA
CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACA
GGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATAC
TCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAG
CGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATT
TCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTA
TAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAAGAATT

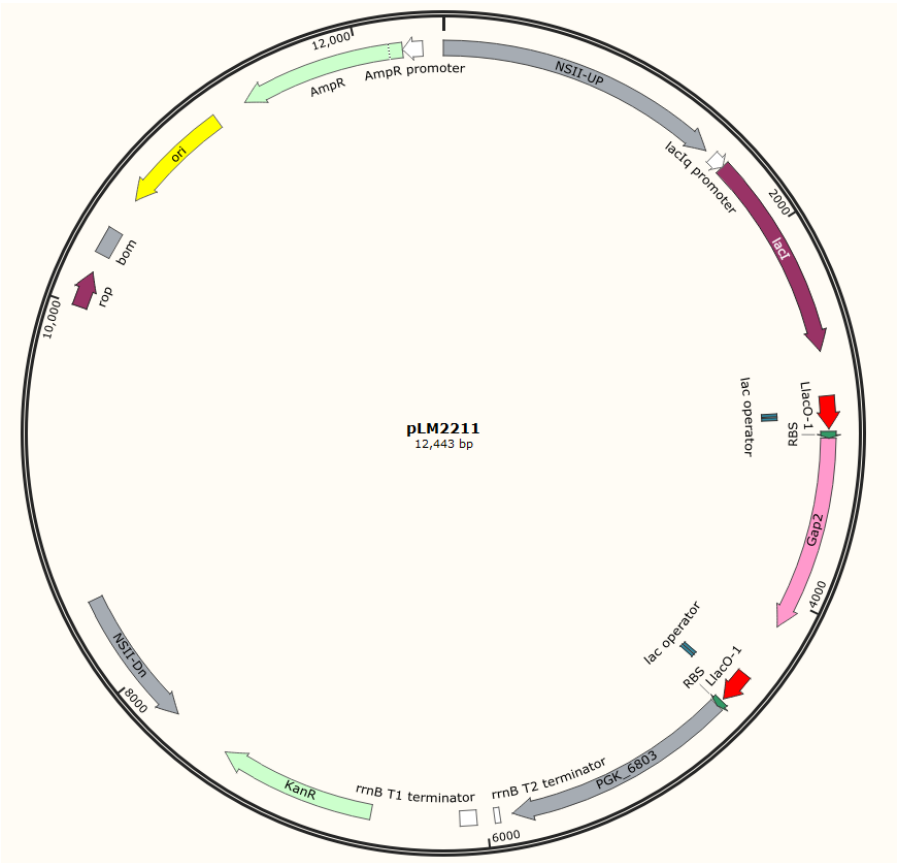

Figure A-6 Plasmid map and sequence of pLM2211

Sequence:

AGCTTGTCATCTGCCGGATGAGGCAAAACCCTGCCTACGGCGCGATTACATCGTCCC
AGCGCGATCGCTCTTACTGTTGATGGCTCGTGCTTAAAAACAATGCAAACTTCACCG
TTTCAGCTGGTGATTTTCGACTGTGATGGTGTGCTTGTTGATAGCGAACGCATCACTA
ATCGCGTCTTTGCAGACATGCTCAATGAACTGGGTCTGTTGGTGACTTTGGATGACA
TGTTTGAGCAGTTTGTGGGTCATTCCATGGCTGACTGTCTCAAACTAATTGAGCGAC
GGTTAGGCAATCCTCCACCCCCTGACTTTGTTCAGCACTATCAACGCCGTACCCGTA
TCGCGTTAGAAACGCATCTACAAGCCGTTCCTGGGGTTGAAGAGGCTTTGGATGCTC
TTGAATTGCCCTACTGTGTTGCGTCCAGTGGTGATCATCAAAAGATGCGAACCACAC
TGAGCCTGACGAAGCTCTGGCCACGATTTGAGGGACGAATCTTCAGCGTGACTGAA
GTACCTCGCGGCAAGCCATTTCCCGATGTCTTTTTGTTGGCCGCCGATCGCTTCGGGG
TTAATCCTACGGCCTGCGCTGTGATCGAAGACACCCCCTTGGGAGTAGCGGCAGGC
GTGGCGGCAGGAATGCAAGTGTTTGGCTACGCGGGTTCCATGCCCGCTTGGCGTCTG
CAAGAAGCCGGTGCCCATCTCATTTTTGACGATATGCGACTGCTGCCCAGTCTGCTC
CAATCGTCGCCAAAAGATAACTCCACAGCATTGCCCAATCCCTAACCCCTGCTCGCG
CCGCAACTACACACTAAACCGTTCCTGCGCGATCGCTCTTACTGTTGATGGCTCGTG
CTTAAAAACAATGCAACCCTAACCGTTTCAGCTGGTGATTTTCGGACGATTTGGCTT
ACAGGGATAACTGAGAGTCAACAGCCTCTGTCCGTCATTGCACACCCATCCATGCAC
TGGGGACTTGACTCATGCTGAATCACATTTCCCTTGTCCATTGGGCGAGAGGGGAGG
GGAATCTTCTGGACTCTTCACTAAGCGGCGATCGCAGGTTCTTCTACCCAAGCAGTG

GCGATCGCTTGATTGCAGTCTTCAATGCTGGCCTCTGCAGCCATCGCCGCCACCAAA
GCATCGTAGGCGGGACGTTGTTGCTCCAGTAAAGTCTTCGCCCGTAACAATCCCCAG
CGACTGCGTAAATCCGCTTCGGCAGGATTGCGATCGAGTTGCCGCCACAGTTGTTTC
CACTGGGCGCGATCGTCAGCTCCCCCTTCCACGTTGCCGTAGACCAGTTGCTCTGCC
GCTGCACCGGCCATCAACACCTGACACCACTGTTCCAGCGATCGCTGACTGAGTTGC
CCCTGTGCGGCTTCGGCTTCTAGCGCAGCTGCTTGGAACTGCACACCCCCGCGACCA
GGTTGTCCTTGGCGCAGCGCTTCCCACGCTGAGAGGGTGTAGCCCGTCACGGGTAAC
CAGATCAATTCGCGCGCGAAGGCGAAGCGGCATGCATTTACGTTGACACCATCGAA
TGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAG
GGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTC
TTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCG
GGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCGCGTGGCAC
AACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCC
TGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTG
CCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCG
GTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGAT
GACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTT
GATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACG
CGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCG
GGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTC
ACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTC
CGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCT
GGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGC
TGCGCGTTGGTGCGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCA
TGTTATATCCCGCCGTCAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACC
AGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCT
GTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCG
CCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGAC
TGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCGCGAATTGATCTG
GTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCAGCC
ATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAAATCACTGCATAATTCGTGTCGCT
CAAGGCGCACTCCCGTTCTGGATAATGTTTTTTGCGCCGACATCATAACGGTTCTGG
CAAATATTCTGAAATGAGCTGTAAGAAACCATTATTATCATGACATTAACCTATAAA
AATAGGCGTATCACGAGGCCCTTTCGTCTTCACCTCGAGAATTGTGAGCGGATAACA
ATTGACATTGTGAGCGGATAACAAGATACTGAGCACATCAGCAGGACGCACTGACC
GAATTCATTAAAGAGGAGAAAGATATCGTCGACTTTACCGTTCCCAAAAATAAAGA
AGGAGGAACAGCATGACTAGAGTAGCAATTAACGGATTTGGACGGATCGGACGCAA
CTTTCTCCGTTGCTGGCTGGGGCGCACCGATAGCCAGTTAGAAGTAGTCGGTATCAA
CGACACCTCTGATCCCAGAACCAATGCTCACCTTTTGCGCTACGACTCCATGTTGGG
TAAGTTGGATGCGGACATCAGTGCCGACGAAAACTCCATTACCGTCAATGGCAAGA
CTATTAAATGTGTTTCCGACCGGAATCCCCTCAATTTGCCCTGGGCAGAATGGAATG
TAGATCTAGTCATCGAAGCCACCGGTGTTTTTGTTACCCATGAAGGGGCCACCAAGC
ACGTTCAAGCCGGAGCCAAAAAAGTTTTAATCACTGCTCCTGGCAAGGGGCCCAAC
ATCGGCACCTATGTGGTGGGGGTCAATGCCCACGAATATAAGCACGAAGAATACGA
AGTAATTAGTAACGCTAGTTGTACTACCAACTGCCTCGCCCCGATCGCCAAAGTAAT
CAACGACAATTTTGGCATCATCAAAGGCACCATGACCACCACCCACAGCTACACCG

GAGACCAACGGATCCTCGATGCTAGCCACCGGGATCTACGCCGGGCCCGGGCTGCT
GCCGTTAACATCGTGCCCACCTCCACCGGAGCTGCCAAAGCGGTGGCCCTGGTAATT
CCTGAACTGCAAGGCAAATTGAACGGTATTGCCCTGCGGGTGCCCACCCCCAACGTT
TCCGTGGTGGATTTGGTAGTACAAGTAGAGAAAAACACCATCGCTGAACAGGTTAA
CGGAGTGCTCAAAGAAGCAGCCAACACCAGCCTTAAAGGAGTGTTGGAATACACCG
ATTTGGAATTGGTTTCCAGCGACTTCCGGGGCACGGATTGCTCTTCCACTGTGGATG
GTAGTCTGACCATGGTAATGGGCGGTGACATGGTTAAAGTCATTGCTTGGTACGACA
ACGAATGGGGCTATTCCCAACGGGTGGTGGACTTGGCTGAAATTGTGGCTAAAAAC
TGGAAATAGGTCGACAGGCCTCTAGACCCGGGCTCGAGCTAGCAAGCTTGGCCGGA
TCCGGCCGGATCCGGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCT
TCTGCTTAATTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGG
GCCGTTGCTTCGCAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGGAGAG
CGTTCACCGACAAACAACAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTT
CGTTTTATTTGATGCCTGGCTAAGAAACCATTATTATCATGACATTAACCTATAAAA
ATAGGCGTATCACGAGGCCCTTTCGTCTTCACCTCGAGAATTGTGAGCGGATAACAA
TTGACATTGTGAGCGGATAACAAGATACTGAGCACATCAGCAGGACGCACTGACCG
AATTCATTAAAGAGGAGAAAGATATCGTCGACTTTACCGTTCCCAAAAATAAAGAA
GGAGGAACAGCATGTTGTCTAAGCAATCGATCGCCAATTTGACGGAGGCAGACCTC
GCAGGGAAACGGGTTTTTGTCCGGGTAGATTTTAATGTGCCCCTAGATAACGGCAGC
ATCACCGACGACACCAGGATTCGGGCAGCCTTACCCACCATCAAAGACCTGTTGAG
CAAAGGCGCCAAAGTTATTTTGGGCAGTCATTTTGGCCGTCCCAAGGGCAAAGTAGT
GGACAGCATGCGCCTCACCCCCGTTGGCGATCGCCTAGGGGAATTATTGGGCCAGC
CGGTGGTCAAATGCGACGATTGCATTGGTGCTGAAGTTACAGCCAAAATTGCTAGCC
TACCCAATGGTGGTGTGGCCCTGTTGGAAAATCTCCGCTTCCATGCCGGGGAAGAAG
GTAACGATGCTGAATTTGCCAAAGCTTTAGCGGCCAACGCCGACCTCTACGTTAACG
ATGCCTTTGGTACTGCCCACCGGGCCCACGCTTCCACCGAAGGGGTCACCCATTTCC
TCAGCCCCAACGTTGCTGGTTACCTAATCGAAAAGGAATTACAGTTCCTCCAAGGAG
CCATCGAAGCCCCCAAACGTCCCCTAGTGGCGATCGTGGGAGGTTCCAAAGTGTCC
AGTAAAATCGGTGTGATCGAAACCCTATTGGACAAGTGCGATAAGTTGATCATCGG
CGGCGGCATGATTTTCACCTTCTACAAAGCCCAAGGTTTAAACACCGGCAAATCCCT
GGTGGAAGAAGACAAATTGGACTTGGCCAAATCCCTCATGGCTAAAGCCAAAGAAA
AAGGCGTGGAATTTCTCCTGCCCACGGACGTAGTAGTGGCCGACAACTTTGCCCCCG
ATGCCAATGCCCAAACCGTTGGTGTCGATGCAATTCCCGATGGTTGGATGGGTCTAG
ACATTGGTCCCGACTCCGTCAAAACCTTCCAGGATGCCCTCGCTGGTTGTGGCACTG
TCATCTGGAACGGCCCCATGGGGGTATTTGAATTTGACAAATTTGCCGTTGGTACCG
AGGCGATCGCCTGCAGCTTGGCTGAATTGACCGCCAGTGGCACTGTCACCATCATCG
GTGGTGGAGATTCTGTCGCCGCAGTGGAAAAAGTGGGAGTGGCCGAAAAAATGAGC
CATATTTCCACCGGTGGGGGCGCTAGCCTGGAATTGCTAGAAGGTAAAGTTCTGCCC
GGCATTGCCGCTTTAGATGACCGATAACTAGACCCGGGCTCGAGCTAGCAAGCTTGG
CCGGATCCGGCCGGATCCGGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGGAT
GGCCTTCTGCTTAATTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCT
CCGGGCCGTTGCTTCGCAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGG
AGAGCGTTCACCGACAAACAACAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGC
CTTTCGTTTTATTTGATGCCTGGCAGTTCCCTACTCTCGCATGGGGAGACCCCACACT
ACCATCGGCGCTACGGCGTTTCACTTCTGAGTTCGGCATGGGGTCAGGTGGGACCAC
CGCGCTACTGCCGCCAGGCAAATTCTGTTTTATCAGCCGTTACCCCACCTACTAGCT

AATCCCATCTGGGCACATCCGATGGCAAGAGGCCCGAAGGTCCCCCTCTTTGGTCTT
GCGACGTTATGCGGTATTAGCTACCGTTTCCAGTAGTTATCCCCCTCCATCAGGCAG
TTTCCCAGACATTACTCACCCGTCCGCCACTCGTCAGCAAAGAAGCAAGCTTAGATC
GACCTGCAGGGGGGGGGGGGAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACA
TTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACATAAACA
GTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCTTGCTCGAGGCCG
CGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAAT
GTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGA
GTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGT
CAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCG
TACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGGAAAACAGCATTCCA
GGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTT
CCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTA
TTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGAT
TTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAA
GCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAAC
CTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATC
GCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCT
TCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAA
TTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAATTGGTTGT
AACACTGGCAGAGCATTACGCTGACTTGACGGGACGGCGGCTTTGTTGAATAAATC
GAACTTTTGCTGAGTTGAAGGATCAGATCACGCATCTTCCCGACAACGCAGACCGTT
CCGTGGCAAAGCAAAAGTTCAAAATCACCAACTGGTCCACCTACAACAAAGCTCTC
ATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGCGATTCAGGCCTGGTAT
GAGTCAGCAACACCTTCTTCACGAGGCAGACCTCAGCGCCCCCCCCCCCCTGCAGGT
CGATCTGGTAACCCCAGCGCGGTTGCTACCAAGTAGTGACCCGCTTCGTGATGCAAA
ATCCGCTGACGATATTCGGGCGATCGCTGCTGAATGCCATCGAGCAGTAACGTGGCA
CCCCGCCCCTGCCAAGTCACCGCATCCAGACTGAACAGCACCAAGAGGCTAAAACC
CAATCCCGCCGGTAGCAGCGGAGAACTACCCAGCATTGGTCCCACCAAAGCTAATG
CCGTCGTGGTAAAAATCGCGATCGCCGTCAGACTCAAGCCCAGTTCGCTCATGCTTC
CTCATCTAGGTCACAGTCTTCGGCGATCGCATCGATCTGATGCTGCAGCAAGCGTTT
TCCATACCGGCGATCGCGCCGTCGCCCTTTCGCTGCCGTGGCCCGCTTACGAGCTCG
TTTATCGACCACGATCGCATCCAAATCCGCGATCGCTTCCCAGTCCGGCAATTCAGT
CTGGGGCGTCCGTTTCATTAATCCTGATCAGGCACGAAATTGCTGTGCGTAGTATCG
CGCATAGCGGCCAGCCTCTGCCAACAGCGCATCGTGATTGCCTGCCTCAACAATCTG
GCCGCGCTCCATCACCAAGATGCGGCTGGCATTACGAACCGTAGCCAGACGGTGAG
CAATGATAAAGACCGTCCGTCCCTGCATCACCCGTTCTAGGGCCTCTTGCACCAAGG
TTTCGGACTCGGAATCAAGCGCCGAAGTCGCCTCATCCAGAATTAAAATGCGTGGAT
CCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTG
GCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTC
ATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTG
GGCGCCATCTCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAAC
CTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGATCGACC
GATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGAC
TATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCG
GCAGCGCTCTGGGTCATTTTCGGCGAGGACCGCTTTCGCTGGAGCGCGACGATGATC

GGCCTGTCGCTTGCGGTATTCGGAATCTTGCACGCCCTCGCTCAAGCCTTCGTCACT
GGTCCCGCCACCAAACGTTTCGGCGAGAAGCAGGCCATTATCGCCGGCATGGCGGC
CGACGCGCTGGGCTACGTCTTGCTGGCGTTCGCGACGCGAGGCTGGATGGCCTTCCC
CATTATGATTCTTCTCGCTTCCGGCGGCATCGGGATGCCCGCGTTGCAGGCCATGCT
GTCCAGGCAGGTAGATGACGACCATCAGGGACAGCTTCAAGGATCGCTCGCGGCTC
TTACCAGCCTAACTTCGATCACTGGACCGCTGATCGTCACGGCGATTTATGCCGCCT
CGGCGAGCACATGGAACGGGTTGGCATGGATTGTAGGCGCCGCCCTATACCTTGTCT
GCCTCCCCGCGTTGCGTCGCGGTGCATGGAGCCGGGCCACCTCGACCTGAATGGAA
GCCGGCGGCACCTCGCTAACGGATTCACCACTCCAAGAATTGGAGCCAATCAATTCT
TGCGGAGAACTGTGAATGCGCAAACCAACCCTTGGCAGAACATATCCATCGCGTCC
GCCATCTCCAGCAGCCGCACGCGGCGCATCTCGGGCAGCGTTGGGTCCTGGCCACG
GGTGCGCATGATCGTGCTCCTGTCGTTGAGGACCCGGCTAGGCTGGCGGGGTTGCCT
TACTGGTTAGCAGAATGAATCACCGATACGCGAGCGAACGTGAAGCGACTGCTGCT
GCAAAACGTCTGCGACCTGAGCAACAACATGAATGGTCTTCGGTTTCCGTGTTTCGT
AAAGTCTGGAAACGCGGAAGTCAGCGCCCTGCACCATTATGTTCCGGATCTGCATCG
CAGGATGCTGCTGGCTACCCTGTGGAACACCTACATCTGTATTAACGAAGCGCTGGC
ATTGACCCTGAGTGATTTTTCTCTGGTCCCGCCGCATCCATACCGCCAGTTGTTTACC
CTCACAACGTTCCAGTAACCGGGCATGTTCATCATCAGTAACCCGTATCGTGAGCAT
CCTCTCTCGTTTCATCGGTATCATTACCCCCATGAACAGAAATCCCCCTTACACGGA
GGCATCAGTGACCAAACAGGAAAAAACCGCCCTTAACATGGCCCGCTTTATCAGAA
GCCAGACATTAACGCTTCTGGAGAAACTCAACGAGCTGGACGCGGATGAACAGGCA
GACATCTGTGAATCGCTTCACGACCACGCTGATGAGCTTTACCGCAGCTGCCTCGCG
CGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCAC
AGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGG
GTGTTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTG
TATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGC
GGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCC
GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCA
GCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAA
GAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTG
CTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCA
AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGG
AAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCC
TTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTT
CGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCG
ACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACT
TATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGC
GGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGT
ATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTC
TTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCA
GATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTC
TGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAA
AAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAG
TATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTAT
CTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATA
ACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA

CCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCG
AGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCC
GGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTG
CTGCAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTC
CCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTC
CTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTT
ATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA
CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCT
CTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTG
CTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTG
AGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTT
TCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGG
AATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTG
AAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAA
AAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCT
AAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCT
TTCGTCTTCAAGAATT

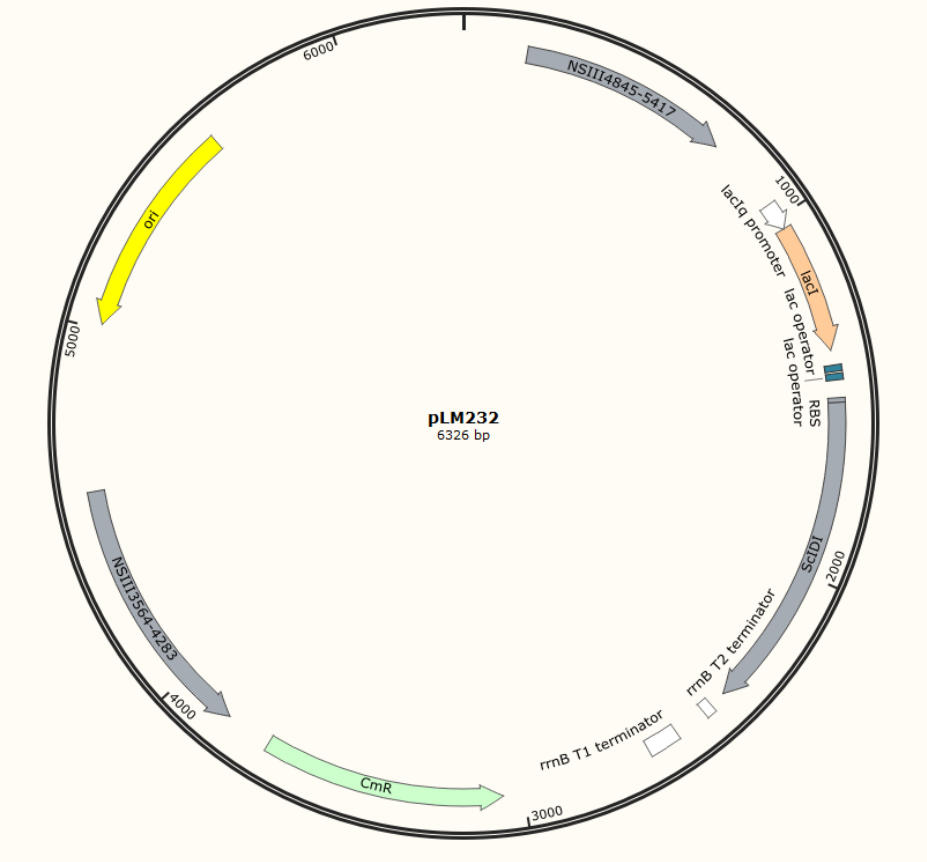

Figure A-7 Plasmid map of pLM232

Sequences:

CGCCGGGGCTGGCAGCTTAGTCCTGCGCAATCTCTACTACATCTGCCAACCCAGTGA
AATTTTGATCTTTGCTGGCAGTAGTCGCCGCAGTAGTGATGGCCGCCGAGTTGGCTA
TCGCTTGGTCAAGGGCGGCAGCAGCCTGCGGGTACCTCTGCTGGAAAAAGCGCTCC
GCGTGTCATTGGTGACAGCTCTGAACGGCTTAGATCTGTTCTTGGTCGGCTTGATGG
GATCAGGGAAAACTACGATCGGGAAGCTGCTGGCGGAGTCCTTGGGCTACACCTAC
GTCGACACCGACAGTTTGATTGAGAACGTCACTGGGCGATCGATTCCAGAGATTTTT
GCCAGCGATGGCGAAGCTGGCTTCCGGCAAATCGAAACGCAGGTCTTAGAAGAAGT
TGCGAGCTATCGCCGCCTAGTAGTGGCAACGGGTGGCGGCATTGTGATTCGGCCTGA
GAACTGGAGCTATTTGCAACAGGGCTTGGTGATTTGGCTAGATGTACCCATCCCCGA
ATTGCTACGCCGTCTAGAAGGTGATCAGAATCGACCGCTTCTACAAACTGAGGCTCC
GGCTACCACGCTACAAGCTTTGTGGGAGCAGCGCCGCGATCGCTATGCCCAAGCCG
ATCTCCGTATTGCGATCGAGGCTTCGGAAGATCCTGAAGTAACTATGCAGCGAATTT
TAGAAGTTATTCCGAGTGTTCTTAAAAACTCTGCTGATAGCTCTCCTGCAGAGATAG
AGACTTAAATCTGGGCCGCGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGC
TGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGAC
CGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAG
GCAGCAGATCAATTCGCGCGCGAAGGCGAAGCGGCATGCATTTACGTTGACACCAT
CGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAAT

TCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGT
GTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAA
ACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCGCGT
GGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCT
GGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACT
GGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAG
CGGCGGTGCACAATCTTCTCGCGCAACGCCGGCGTATCACGAGGCCCTTTCGTCTTC
ACCTCGAGAATTGTGAGCGGATAACAATTGACATTGTGAGCGGATAACAAGATACT
GAGCACATCAGCAGGACGCACTGACCGAATTTCGACAGGCCTAAAGAGGAGAAAAT
GACAGCGGATAACAACTCTATGCCACACGGAGCAGTTAGCAGTTACGCAAAGCTCG
TCCAAAATCAAACTCCAGAGGATATCTTAGAAGAGTTCCCTGAGATTATTCCTCTCC
AACAGCGCCCTAACACCCGGTCCTCGGAAACTTCTAACGACGAGAGCGGCGAGACT
TGCTTCTCTGGTCACGACGAGGAACAAATTAAACTCATGAATGAGAACTGCATTGTA
TTGGATTGGGACGATAACGCAATTGGTGCGGGCACCAAAAAAGTTTGCCATTTGATG
GAAAACATTGAGAAAGGCTTATTGCATCGGGCGTTCAGTGTGTTTATCTTTAATGAG
CAGGGGGAACTGTTGCTCCAACAACGTGCTACAGAGAAGATTACCTTTCCCGATCTG
TGGACGAATACTTGCTGCTCGCATCCCTTATGTATCGATGATGAGCTGGGATTAAAA
GGGAAGTTAGATGATAAGATCAAAGGTGCGATTACAGCCGCGGTACGCAAGTTGGA
TCACGAACTGGGGATTCCCGAGGATGAGACAAAAACTCGCGGTAAATTCCACTTTCT
CAATCGTATCCACTATATGGCCCCATCGAATGAACCCTGGGGTGAGCACGAGATTGA
CTACATTCTGTTCTATAAAATTAACGCAAAGGAAAACTTAACAGTTAATCCCAATGT
TAATGAAGTGCGCGACTTCAAATGGGTTAGCCCGAACGATTTGAAAACGATGTTCGC
GGACCCCTCCTATAAATTTACACCGTGGTTTAAAATCATCTGTGAAAACTACTTATTC
AATTGGTGGGAGCAGTTGGACGACCTCTCCGAGGTTGAAAACGACCGTCAAATTCA
TCGTATGCTGTAGCAAGCTTGGCCGGATCCGGCCGGATCCGGAGTTTGTAGAAACGC
AAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAATTTGATGCCTGGCAGTTTATGG
CGGGCGTCCTGCCCGCCACCCTCCGGGCCGTTGCTTCGCAACGTTCAAATCCGCTCC
CGGCGGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAACAGATAAAACGAA
AGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCAGTTCCCTACTC
TCGCGGCCGCCGATCCTCTAGTATGCTTGTAAACCGTTTTGTGAAAAAATTTTTAAA
ATAAAAAAGGGGACCTCTAGGGTCCCCAATTAATTAGTAATATAATCTATTAAAGGT
CATTCAAAAGGTCATCCACCGGATCAGCTTAGTAAAGCCCTCGCTAGATTTTAATGC
GGATGTTGCGATTACTTCGCCAACTATTGCGATAACAAGAAAAAGCCAGCCTTTCAT
GATATATCTCCCAATTTGTGTAGGGCTTATTATGCACGCTTAAAAATAATAAAAGCA
GACTTGACCTGATAGTTTGGCTGTGAGCAATTATGTGCTTAGTGCATCTAACGCTTG
AGTTAAGCCGCGCCGCGAAGCGGCGTCGGCTTGAACGAATTGTTAGACATTAGCCG
TTGATTGCAGGTGCAGTCAGAGCAACCGGGGTCGCTTCGCCTGCTGCCAAGTCGAGC
GGGAAGTTGTGAGCATTACGCTCGTGCATCACTTCCATGCCCAAGTTGGCACGGTTC
AACACATCTGCCCAGGTGTTGATCACACGGCCTTGGCTATCCAAAACCGACTGGTTG
AAGTTGAAACCATTCAGGTTGAACGCCATGGTGCTGATGCCCAAGGACGTGAACCA
GATGCCCACGACCGGCCATGCAGCCAGGAAGAAGTGCAGCGAACGGCTGTTATTGA
ACGATGCGTATTGGAAGATCAAGCGACCAAAGTAACCGTGGGCTGCCACGATGTTG
TAGGTCTCTTCCTCTTGACCAAATTTGTAGCCGTAGTTTTGGCTCTCGGTCTCGGTCG
TCTCACGCACCAGCGAGCTGGTCACCAACGAACCGTGCATTGCCGAGAACAGCGAA
CCACCGAACACACCAGCCACACCCAGCATGTGGAAGGGGTGCATCAAAATGTTGTG
CTCTGCTTGGAACACGAACATGAAGTTGAAGGTGCCGCTGATACCCAAGGGCATGC

CGTCCGAGAACGAACCTTGGCCGATCGGGTAGATCAGGAACACTGCAGTCGCTGCT
GCAACCGGTGCGCTGTATGCAACACAGATCCAGGGGCGCATACCGAGGCGGTACGA
CAGTTCCCACTCACGGCCCATGTAGCAGAAAACGCCCGGTGATGACGGTGAAAACC
TCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGG
AGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGC
CATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCA
GAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGT
AAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCG
CTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTT
ATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAA
AAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC
CCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG
ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCC
GACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCT
TTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTG
GGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT
CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGT
AACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTG
GCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCC
AGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG
GTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTC
AAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC
GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAA
ATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACA
GTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATC
CATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATC
TGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATC
AGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTAT
CCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAG
TTAATAGTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTCACGCTCGTC
GTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATC
CCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAG
TAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACT
GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGT

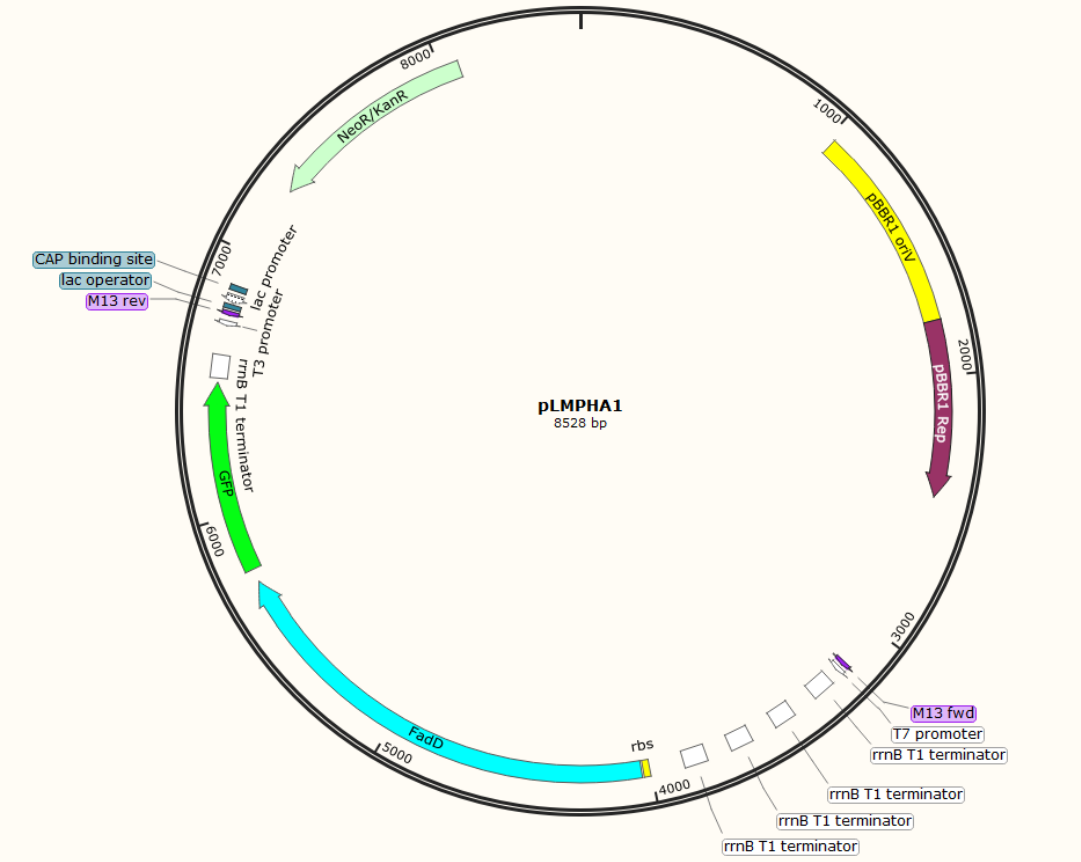

Figure A-8 Plasmid map of pLMPHA1, pLMPHA2-6 were built by replacing the FadD sequence with *PcaF-1, PHAG, TET, TF-RNAp, TF-Cra*

Sequence:

CTCGGGCCGTCTCTTGGGCTTGATCGGCCTTCTTGCGCATCTCACGCGCTCCTGCGGC
GGCCTGTAGGGCAGGCTCATACCCCTGCCGAACCGCTTTTGTCAGCCGGTCGGCCAC
GGCTTCCGGCGTCTCAACGCGCTTTGAGATTCCCAGCTTTTCGGCCAATCCCTGCGG
TGCATAGGCGCGTGGCTCGACCGCTTGCGGGCTGATGGTGACGTGGCCCACTGGTGG
CCGCTCCAGGGCCTCGTAGAACGCCTGAATGCGCGTGTGACGTGCCTTGCTGCCCTC
GATGCCCCGTTGCAGCCCTAGATCGGCCACAGCGGCCGCAAACGTGGTCTGGTCGC
GGGTCATCTGCGCTTTGTTGCCGATGAACTCCTTGGCCGACAGCCTGCCGTCCTGCG
TCAGCGGCACCACGAACGCGGTCATGTGCGGGCTGGTTTCGTCACGGTGGATGCTGG
CCGTCACGATGCGATCCGCCCCGTACTTGTCCGCCAGCCACTTGTGCGCCTTCTCGA
AGAACGCCGCCTGCTGTTCTTGGCTGGCCGACTTCCACCATTCCGGGCTGGCCGTCA
TGACGTACTCGACCGCCAACACAGCGTCCTTGCGCCGCTTCTCTGGCAGCAACTCGC
GCAGTCGGCCCATCGCTTCATCGGTGCTGCTGGCCGCCCAGTGCTCGTTCTCTGGCG
TCCTGCTGGCGTCAGCGTTGGGCGTCTCGCGCTCGCGGTAGGCGTGCTTGAGACTGG
CCGCCACGTTGCCCATTTTCGCCAGCTTCTTGCATCGCATGATCGCGTATGCCGCCAT
GCCTGCCCCTCCCTTTTGGTGTCCAACCGGCTCGACGGGGGCAGCGCAAGGCGGTGC
CTCCGGCGGGCCACTCAATGCTTGAGTATACTCACTAGACTTTGCTTCGCAAAGTCG

TGACCGCCTACGGCGGCTGCGGCGCCCTACGGGCTTGCTCTCCGGGCTTCGCCCTGC
GCGGTCGCTGCGCTCCCTTGCCAGCCCGTGGATATGTGGACGATGGCCGCGAGCGG
CCACCGGCTGGCTCGCTTCGCTCGGCCCGTGGACAACCCTGCTGGACAAGCTGATGG
ACAGGCTGCGCCTGCCCACGAGCTTGACCACAGGGATTGCCCACCGGCTACCCAGC
CTTCGACCACATACCCACCGGCTCCAACTGCGCGGCCTGCGGCCTTGCCCCATCAAT
TTTTTTAATTTTCTCTGGGGAAAAGCCTCCGGCCTGCGGCCTGCGCGCTTCGCTTGCC
GGTTGGACACCAAGTGGAAGGCGGGTCAAGGCTCGCGCAGCGACCGCGCAGCGGCT
TGGCCTTGACGCGCCTGGAACGACCCAAGCCTATGCGAGTGGGGGCAGTCGAAGGC
GAAGCCCGCCCGCCTGCCCCCCGAGCCTCACGGCGGCGAGTGCGGGGGTTCCAAGG
GGGCAGCGCCACCTTGGGCAAGGCCGAAGGCCGCGCAGTCGATCAACAAGCCCCGG
AGGGGCCACTTTTTGCCGGAGGGGGAGCCGCGCCGAAGGCGTGGGGGAACCCCGCA
GGGGTGCCCTTCTTTGGGCACCAAAGAACTAGATATAGGGCGAAATGCGAAAGACT
TAAAAATCAACAACTTAAAAAAGGGGGGTACGCAACAGCTCATTGCGGCACCCCCC
GCAATAGCTCATTGCGTAGGTTAAAGAAAATCTGTAATTGACTGCCACTTTTACGCA
ACGCATAATTGTTGTCGCGCTGCCGAAAAGTTGCAGCTGATTGCGCATGGTGCCGCA
ACCGTGCGGCACCCTACCGCATGGAGATAAGCATGGCCACGCAGTCCAGAGAAATC
GGCATTCAAGCCAAGAACAAGCCCGGTCACTGGGTGCAAACGGAACGCAAAGCGC
ATGAGGCGTGGGCCGGGCTTATTGCGAGGAAACCCACGGCGGCAATGCTGCTGCAT
CACCTCGTGGCGCAGATGGGCCACCAGAACGCCGTGGTGGTCAGCCAGAAGACACT
TTCCAAGCTCATCGGACGTTCTTTGCGGACGGTCCAATACGCAGTCAAGGACTTGGT
GGCCGAGCGCTGGATCTCCGTCGTGAAGCTCAACGGCCCCGGCACCGTGTCGGCCT
ACGTGGTCAATGACCGCGTGGCGTGGGGCCAGCCCCGCGACCAGTTGCGCCTGTCG
GTGTTCAGTGCCGCCGTGGTGGTTGATCACGACGACCAGGACGAATCGCTGTTGGGG
CATGGCGACCTGCGCCGCATCCCGACCCTGTATCCGGGCGAGCAGCAACTACCGAC
CGGCCCCGGCGAGGAGCCGCCCAGCCAGCCCGGCATTCCGGGCATGGAACCAGACC
TGCCAGCCTTGACCGAAACGGAGGAATGGGAACGGCGCGGGCAGCAGCGCCTGCCG
ATGCCCGATGAGCCGTGTTTTCTGGACGATGGCGAGCCGTTGGAGCCGCCGACACG
GGTCACGCTGCCGCGCCGGTAGCACTTGGGTTGCGCAGCAACCCGTAAGTGCGCTGT
TCCAGACTATCGGCTGTAGCCGCCTCGCCGCCCTATACCTTGTCTGCCTCCCCGCGTT
GCGTCGCGGTGCATGGAGCCGGGCCACCTCGACCTGAATGGAAGCCGGCGGCACCT
CGCTAACGGATTCACCGTTTTTATCAGGCTCTGGGAGGCAGAATAAATGATCATATC
GTCAATTATTACCTCCACGGGGAGAGCCTGAGCAAACTGGCCTCAGGCATTTGAGA
AGCACACGGTCACACTGCTTCCGGTAGTCAATAAACCGGTAAACCAGCAATAGACA
TAAGCGGCTATTTAACGACCCTGCCCTGAACCGACGACCGGGTCGAATTTGCTTTCG
AATTTCTGCCATTCATCCGCTTATTATCACTTATTCAGGCGTAGCACCAGGCGTTTAA
GGGCACCAATAACTGCCTTAAAAAAATTACGCCCCGCCCTGCCACTCATCGCAGTCG
GCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAA
ATATTAACGCTTACAATTTCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGC
GATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCA
AGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGAC
GGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTGGGGCCGC
AATTCCCAATTCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCT
TTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGG
AGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGC
CATAAACTGCCAGGAATTAATTCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAA
AGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACA

AATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGC
AGGACGCCCGCCATAAACTGCCAGGAATTAATTCCAGGCATCAAATAAAACGAAAG
GCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCC
TGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGA
GGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGAATTAATTCCAGGCATCAAA
TAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGG
TGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAG
CAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGAATTGGG
GATCGGAAGCTGAATTCGAGCTCAGGCCTTGGTACCCGGGATCCGATTAACTCACAC
AGGAGGGTATCAcctaggTCAGGCGATCTTCTTCAAGCCTTGCTTCTTGAGCTCTTCATC
ACGCAGTTCGCGGCGCAGGATCTTGCCCACGTTGGTGGTCGGCAGCGAATCGCGGA
ACTCGATGTGGCGCGGTACCTTGTAGCCGGTAACGTTGGCACGCATGTGCTCCATCA
CCTGCTCCTTGGTCACGGTCATGCCCGGCTTGACCACGATGAAGACCTTGATCACTT
CACCCGACTTCTCGTCCGGCACACCGATGGCTGCGCACTGCAGCACGCCCGGCAGG
GCCGCCAACACGTCTTCCAGCTCGTTGGGGTATACGTTGAAGCCCGAGACCAGGATC
ATGTCTTTCTTGCGGTCGACGATGCGCATGTAGCCGTCCGCCTGGATCACGGCGATG
TCGCCGGTCTTCAGCCAGCCTTCGCTGTCGAGAATCTCGGCCGTGGCGTCTTCACGC
TGCCAGTAGCCCTTCATCACCTGCGGGCCCTTGACACACAGCTCGCCCACTTCGCCA
AGCGGCAACTCATTGCCGGCGTCGTCGATGACCTTGCACAGGGTCGACGGCACCGG
AATACCGATGGTGCCCACCTGGTTCGCTTCCGAGGGGTTCACCGCCGCCACCGGGCT
GGTTTCGGTCATGCCGTAGCCTTCGCAGATGGCGCAACCGGTAACGGCCTTCCAGCG
CTCGGCCACGCTTAGCTGCAAGGCCATACCGCCCGACAGGGTGATTTTCAGCGCCGA
GAAATCCAGGGCTCGGAACGCCTCGTTGTTGCACAGGGCAACGAACAGGGTGTTGA
GGCCCACAAAGCCGCTGAACTTCCACTTGCCCAGTTCCTTGACCATCGCCGGCAAGT
CACGCGGGTTGCTGATCAGCACGTTGTGGTTGCCGATCAGCATCATCGCCATGCAAT
GGAAGGTAAACGCATAGATGTGGTACAGCGGCAGCGGGGTGATGAGGATTTCGCAG
CCTTCGTGCAGGTTGGAGCCCATCAGTGCACGGCACTGCAGCATGTTGGCCACCAGG
TTGCGGTGGGTCAGCATGGCGCCTTTGGCCACACCGGTGGTACCGCCGGTGTACTGC
AGCACCGCCACGTCGTTGGCCTGCGGGTTGGCTTCGGTCACGGGCTGGCCCTTGCCC
AGCGCCAGGGCGTCATTGAAACGCACGGCCTGCGGCAGGTTGTAGGCCGGCACCAT
CTTCTTCACGTACTTGATCACGCTGTTGATCAGCAGGCGCTTGAGTGGTGGCAGCAG
GTCGGCAACTTCAGTGACGATAACGTGCCTGACCTGGGTTTTGGGCACCACTTTTTC
CGCCAGGTGGGCCATGTTGGCCAGGCACACCAGGGCCTTGGCACCCGAGTCGTTGA
ACTGGTGTTCCATCTCCCGCGCGGTGTACAGCGGGTTGGTGTTGACCACGATCAGCC
CGGCACGCATGGCACCGAAGACCGCGACCGGGTATTGCAGGACATTGGGCAGTTGC
ACGGCAATGCGGTCACCCGGCTTGAGGTCGGTATGCTGCTGCAGCCAGGCGGCAAA
CGCCCCCGACAACGCATACAACTCGCCATAAGTGATAGTCTTGCCCAGGTTGCTAAA
GGCCGGTTTGTCGGCAAAGCGTTGGCAGGATTGCTTGAGTACTGCCTGGATATTGGG
GAATTCGTCAGGATTGATTTCCGCCGTAATCCCGGCTGGGTACTTATCCTTCCAAAA
ATTTTCGATCATAAGCTTAATTCCCTAACTAACTAAAGATTAACTTTATAAGGAGGA
AAAACATATGAGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGA
ATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTG
ATGCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTG
TTCCATGGCCAACACTTGTCACTACTTTGACTTATGGTGTTCAATGCTTTTCAAGATA
CCCAGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGT
ACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAG

TCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTA
AAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAAT
GTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAG
ACACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTC
CAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTG
CCCTTTCGAAAGATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAA
CAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAATAGGAATTAATTCCA
GGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTT
GTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACG
TTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGG
AATTGGGGGCGGCCCCGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCG
CTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATT
CCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGT
GAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCT
GTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTA
TTGGGCGCATGCATAAAAACTGTTGTAATTCATTAAGCATTCTGCCGACATGGAAGC
CATCACAAACGGCATGATGAACCTGAATCGCCAGCGGCATCAGCACCTTGTCGCCTT
GCGTATAATATTTGCCCATGGGGGTGGGCGAAGAACTCCAGCATGAGATCCCCGCG
CTGGAGGATCATCCAGCCGGCGTCCCGGAAAACGATTCCGAAGCCCAACCTTTCAT
AGAAGGCGGCGGTGGAATCGAAATCTCGTGATGGCAGGTTGGGCGTCGCTTGGTCG
GTCATTTCGAACCCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAAGGCGATAGAAG
GCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAG
CCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGAT
AGCGGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGAAAAGCGGCCATTT
TCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTCACGACGAGATCCTCGCC
GTCGGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGAT
GCTCTTCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTC
GCTCGATGCGATGTTTCGCTTGGTGGTCGAATGGGCAGGTAGCCGGATCAAGCGTAT
GCAGCCGCCGCATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGTGA
GATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCT
TCAGTGACAACGTCGAGCACAGCTGCGCAAGGAACGCCCGTCGTGGCCAGCCACGA
TAGCCGCGCTGCCTCGTCCTGCAGTTCATTCAGGGCACCGGACAGGTCGGTCTTGAC
AAAAAGAACCGGGCGCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAG
CCGATTGTCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGGA
GAACCTGCGTGCAATCCATCTTGTTCAATCATGCGAAACGATCCTCATCCTGTCTCTT
GATCAGATCTTGATCCCCTGCGCCATCAGATCCTTGGCGGCAAGAAAGCCATCCAGT
TTACTTTGCAGGGCTTCCCAACCTTACCAGAGGGCGCCCCAGCTGGCAATTCCGGTT
CGCTTGCTGTCCATAAAACCGCCCAGTCTAGCTATCGCCATGTAAGCCCACTGCAAG
CTACCTGCTTTCTCTTTGCGCTTGCGTTTTCCCTTGTCCAGATAGCCCAGTAGCTGAC
ATTCATCCCAGGTGGCACTTTTCGGGGAAATGTGCGCGCCCGCGTTCCTGCTGGCGC
TGGGCCTGTTTCTGGCGCTGGACTTCCCGCTGTTCCGTCAGCAGCTTTTCGCCCACGG
CCTTGATGATCGCGGCGGCCTTGGCCTGCATATCCCGATTCAACGGCCCCAGGGCGT
CCAGAACGGGCTTCAGGCGCTCCCGAAGGT