De novo creation of MG1655-derived E. coli strains specifically designed for plasmid DNA production

Applied Microbiology and Biotechnology

De novo creation of MG1655-derived E. coli strains specifically designed for plasmid DNA production

--Manuscript Draft--

Manuscript Number:	AMAB-D-12-00981R1
Full Title:	De novo creation of MG1655-derived E. coli strains specifically designed for plasmid DNA production
Article Type:	Original Paper
Section/Category:	Biotechnological products and process engineering
Corresponding Author:	Kristala Prather UNITED STATES
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	
Corresponding Author's Secondary Institution:	
First Author:	Geisa A. L. Goncalves
First Author Secondary Information:	
Order of Authors:	Geisa A. L. Goncalves
	Duarte M. F. Prazeres, Ph.D
	Gabriel A. Monteiro, Ph.D
	Kristala Prather
Order of Authors Secondary Information:	
Abstract:	The interest in plasmid DNA (pDNA) as a biopharmaceutical has been increasing over the last several years, especially after the approval of the first DNA vaccines. New pDNA production strains have been created by rationally mutating genes selected on the basis of E. coli central metabolism and plasmid properties. Nevertheless, the highly mutagenized genetic background of the strains used makes it difficult to ascertain the exact impact of those mutations. To explore the effect of strain genetic background, we investigated single and double knockouts of two genes, pykF and pykA, which were known to enhance pDNA synthesis in two different E. coli strains, MG1655 (wild-type genetic background) and DH5α (highly mutagenized genetic background). The knockouts were only effective in the wild-type strain MG1655, demonstrating the relevance of strain genetic background and the importance of designing new strains specifically for pDNA production. Based on the obtained results, we created a new pDNA production strain starting from MG1655, by knocking out the pgi gene in order to redirect carbon flux to the pentose phosphate pathway, enhance nucleotide synthesis and consequently increase pDNA production. GALG20 (MG1655ΔendAΔrecAΔpgi) produced 25-fold more pDNA (19.1 mg/g DCW) than its parental strain, MG1655ΔendAΔrecA (0.8 mg/g DCW), in glucose. For the first time pgi was identified as an important target for constructing a high-yielding pDNA production strain.
Response to Reviewers:	Dear Editor, Specific responses to each of the reviewers' comments have been provided in a separate document. Thank you, Kris Prather

Massachusetts Institute of Technology
Chemical Engineering Department
Room 66-454
77 Massachusetts Avenue
Cambridge, MA 02139

Kristala L. Jones Prather
Theodore T. Miller Career
Development Associate Professor
Phone: 617-253-1950
Fax: 617-258-5042
E-mail: kljp@mit.edu

July 13, 2012

Dear Editor,

Please receive the revised version of our manuscript entitled "***De novo* creation of MG1655-derived E. coli strains specifically designed for plasmid DNA production**" for consideration for publication in *Applied Microbiology and Biotechnology* (manuscript number AMAB-D-12-00981). Specific responses to each of the reviewers' comments have been provided in a separate document. We have made a sincere effort to address the concerns raised by the reviewers and hope that our responses are found to be satisfactory.

Thank you,

Kristala Jones Prather

Responses to reviewers' comments

Manuscript AMAB-D-12-00981
De novo creation of MG1655-derived *E. coli* strains specifically designed for plasmid DNA production

Original text is referenced according to the page/line numbers of the original manuscript. Revised text is referenced according to the page/line numbers of the revised manuscript.

Reviewer #1: General comments:
The authors conducted the research in an interesting area of developing new E. coli strains for plasmid DNA production based on carbon source metabolic pathway and plasmid DNA yield. The results generated by the authors indicated that mutated wild type strain MG1655 in genes coding pyruvate kinase and phosphoglucose isomerase may result in significant increase in plasmid DNA production yield. The authors presented logical designing of the experiments, reliable data, and reasonable conclusion, which is within the scope of "Applied Microbiology and Biotechnology". While the strategy for experimental designing is reasonably planned and most critical data is presented, however, some data of supporting the final results were not clear or missing (see below specific comments/questions). My recommendation for this manuscript is "acceptable for publication with minor revision".

Specific Comments:

1.Questions to content:

1).Statement in Page 8 Line 16: Nevertheless, pDNA ? was 20g/L (Table 3). There was no GALG20 data either in glucose or glycerol in the Table 3.

First, we thank the reviewer for a careful reading of our manuscript and for your comments and suggestions. The statement on Page 8 Line 16 was only referring to the Pyk mutant strains since these were the only ones constructed at this point. To clarify this, we have changed the sentence to read as follows:

Original text (Page 8, Line 183): "Nevertheless, pDNA yields were significantly lower in glucose than in glycerol for all strains when the initial concentration…"

Revised text (Page 8, Line 184): "Nevertheless, pDNA yields were significantly lower in glucose than in glycerol for all Pyk mutant strains when the initial concentration…".

However, we appreciate the suggestion from the reviewer to include more data in Table 3. Thus we have added additional data on GALG10 and GALG11 and we have included data on GALG20 in Table 3.

Revised Table: Table 3

2). In Table 2, the authors listed cell growth data of the parent strains in this report, which is good. However, why not include GALG10, GALG11, and GALG20 three important strains in comparison of cell growth with the parent strains? If the data is available, the authors should include in the table 2 with a modified table title or present in a separated table.

We do indeed have these data available and have included biomass, acetate, and growth rate data for GALG10, GALG11 and GALG20 in Table 2. The title of Table 2 has been modified to reflect these additions.

Revised Table: Table 2

3). The Table 3 summarized 10 different strains' plasmid DNA production yield in different or different amount of carbon sources. A couple of critical data either were not clear or missed, including: 1). There was no carbon source condition data under 20g/L or 10+10g/L carbon sources for two important mutant strains GALG10 and GALG11, while other strains were tested under those conditions. 2). The best strain in this report is the GALG20, however there was no data showed here in the comparison with all other strains. Question would be: how you pick up this strain in to the Table comparison without initial yield data?

As stated above, we appreciated the reviewer's comments related to data organization especially in Table 3 and have included additional data in this table, including data on GALG10 and GALG11 with 20g/L of glucose. However, these strains were not tested with 10+10g/L of glucose. We first identified glucose inhibition for the wild-type strain MG1655*ΔendAΔrecA* in 20g/L, and conducted tests with different glucose concentrations to understand this effect in the wild-type strain versus the *pyk* double mutant. Through these experiments, we concluded that the best condition was 5+10g/L. After determining the best condition for the wild-type and double mutant strains, we tested the other strains only in the two extremes of 20g/L (inhibition for wild-type) and 5+10g/L (no-inhibition for wild-type strain).

4). Combining information in Table 4 with figure 2, it seems reasonable to conclude that GALG20 has the highest specific productivity with the number of 20mg/g DCW. However, the comparison was poor. There is no systematic comparison of the best strain with others (except MG1655 deltaendAdeltarecA) in cell growth and productivities in different and different amount of carbon sources.

We understand the confusion of the reviewer due to the amount of data presented in this work. However, after all the modifications in Table 3 suggested by the reviewer, the relevant data for all strains constructed is now included. We believe that this now makes Table 4 easier to understand. Table 4 is merely a summary of the best performers and describes the top three strains identified in this study (GALG20, GALG10 and MG1655*ΔendAΔrecA*) related to the best carbon source for each of them and compares results with those obtained with DH5α, a commonly used strain for pDNA production.

5). Figure 2, suggest to insert three corresponding small boxes (white, grey, and solid) inside of the figure, and label the boxes with 5+10g/L of glucose, 5+10g/L of glycerol, and 20g/L of glucose to indicate corresponding bars.

We have inserted the small boxes in Figure 2 as suggested by the reviewer.

Revised Figure: Figure 2

2. Minor editing issues:

1). More than two authors in a reference, et al should be added in the text where the references were cited.

We have followed the Applied Microbiology & Biotechnology journal guidelines which state that we should not use author lists with "et al." As noted:

- All author names should be provided in the references of AMB-manuscripts
- Please do not use an EndNote Style abbreviating long author lists with "et al."

2). On page 2, Line 12, insert "(*E. coli*)" after "*Escherichia coli*".

Revised text: On page 2, Line 12, *E. coli* was inserted after *"Escherichia coli".*

Reviewer #2: In this study the authors have attempted to develop a protocol for enhancing pDNA production by mutating genes in the glycolytic pathway without reducing growth, by studying the influence of carbon source (glucose vs glycerol). In addition, they have tested the effect of mutating pgi gene by shifting carbon flow from pentose phosphate pathway to promoting nucleotide synthesis.
The study has been carefully and "rationally" conducted, and presented well. However, what is novel about this study is unclear. A somewhat minor observation that genetic background may play a role in the effect of mutations could be important.

First, we thank the reviewer for evaluating our manuscript. We do believe that investigating the role of genetic background in the effect of various mutations is novel, and, as the reviewer states, this could be important, especially as researchers attempt to transfer these mutations into other host strains. We also believe that the specific finding of the beneficial effect of the pgi knockout on pDNA production is novel, particularly given the exceptionally high specific yields that resulted.

Several studies have looked looked at the pgi gene and its effect on enhancing production of products (Ahn et al, 2011, FEMS Microbiol Lett, 324:10; enhanced production of shikimate). Also, study of Kabir and Shimizu(2003), J. Biotech 105:11-31, has examined the effect of mutation of 87 genes on production of metabolites particularly in pgi mutants. They have specifically shown that the phb gene has the most significant impact on the pentose phosphate pathway and the glycolytic pathway genes.

We agree with the reviewer that we were not the first to examine the effects of the pgi mutation on enhancing the production of a biochemical product. We have cited 4 papers that describe this mutation as being beneficial. However, no previous work has identified pgi (or rather, its deletion) as a significant factor in pDNA production.

In order to make this study meaningful, the effect of other carbon sources such as fructose, maltose (rather than combinations of glucose, also studied earlier by various investigators) may add significance to this study. Ahn et al, 2011, find that fructose rather than glucose has a significant effect on pDNA production in pgi mutants.

We appreciate the desire of the reviewer to understand the impact of the *pgi* mutation in the context of additional carbon sources. However, in the present study, with the objective of producing pDNA for gene therapy and DNA vaccination, we do not feel that it is necessary to test fructose as carbon source. Our objective was not to determine the effect of the mutation for all carbohydrates, but rather to consider the two most commonly used substrates in industrial fermentations, glucose and glycerol, as described in the Introduction.

Contrary to the reviewer's comment, we do not find any data on the effect of fructose on pDNA production in *pgi* mutants in Ahn et al, 2011. This reference does show that *pgi* mutant strains exhibited significantly reduced cell growth and a decrease of shikimic acid production in glucose as a sole carbon source. The addition of fructose with glucose was used as a strategy to recover cell growth.

Another interesting point that was verified by Ahn et al, 2011, was the concentration of NADPH. In glucose only, the molar yield of NADPH on sugar for the *pgi* mutant was twice that of the parental strain, while in fructose or in the mixture of fructose and glucose, the concentration of NADPH was similar for both strains (*pgi* mutant or parental strain). In our case, we don't have problems with reduction of growth rate in glucose (Table 2), and it was already proposed computationally that increasing NADPH pools would favor pDNA production. Thus, in this case, we consider that glucose alone would be more appropriate for pDNA production. As described in our manuscript: "A large excess of NADPH was detected in Pgi mutant strains (Canonaco 2001; Chin and Cirino 2011; Siedler 2011), which was shown to improve biotransformations in various processes. However, an excess of NADPH can cause a redox imbalance in the cell, imposing stress that may appear as a large reduction in growth rate (Charusanti 2010). In the specific case of pDNA synthesis, a mathematical model has demonstrated that high generation of NADPH would be effective for increasing yields (Cunningham 2009)".

We do appreciate the reviewer highlighting this interesting report on the use of a *pgi* knockout for another biochemical, and we have added this citation to the manuscript.

New text (Page 15, Line 348-356): "Another advantage of GALG20 for pDNA production was that the growth rate was similar to the parental strain in glucose (Table 2), whereas a previous study demonstrated a significant reduction of growth rate in a *pgi* mutant strain grown in the same carbon source (Ahn 2011). The differences in the behavior of these two *pgi* mutants in glucose could be associated with the strain genetic background between the two studies, *E. coli* MG1655*ΔendAΔrecA* and KPM SA1 (*ΔaraLΔaraK*), respectively. In the current work, we have demonstrated that strain background can substantially influence the effect of gene knockouts, such as the differences between the *pykF* and *pykA* mutations in MG1655*ΔendAΔrecA* versus DH5α."

De novo creation of MG1655-derived *E. coli* strains specifically designed for plasmid DNA production

Geisa A. L. Gonçalves[1,2,3], Duarte M. F. Prazeres[1,2,3], Gabriel A. Monteiro[1,2,3], Kristala L. J. Prather[1,4]*

[1]MIT-Portugal Program

[2]Department of Bioengineering, Instituto Superior Técnico (IST), Lisbon, Portugal

[3]IBB-Institute for Biotechnology and Bioengineering, Center for Biological and Chemical Engineering, IST, Lisbon, Portugal

[4]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA

*Corresponding Author

77 Massachusetts Ave., Room 66-454

Cambridge, MA 02139

Phone: (617) 253-1950

Fax: (617) 258-5042

kljp@mit.edu

Keywords: DNA vaccine, plasmid biopharmaceuticals, *Escherichia coli*, strain engineering, metabolic engineering

Abstract

The interest in plasmid DNA (pDNA) as a biopharmaceutical has been increasing over the last several years, especially after the approval of the first DNA vaccines. New pDNA production strains have been created by rationally mutating genes selected on the basis of *E. coli* central metabolism and plasmid properties. Nevertheless, the highly mutagenized genetic background of the strains used makes it difficult to ascertain the exact impact of those mutations. To explore the effect of strain genetic background, we investigated single and double knockouts of two genes, *pykF* and *pykA*, which were known to enhance pDNA synthesis in two different *E. coli* strains, MG1655 (wild-type genetic background) and DH5α (highly mutagenized genetic background). The knockouts were only effective in the wild-type strain MG1655, demonstrating the relevance of strain genetic background and the importance of designing new strains specifically for pDNA production. Based on the obtained results, we created a new pDNA production strain starting from MG1655, by knocking out the *pgi* gene in order to redirect carbon flux to the pentose phosphate pathway, enhance nucleotide synthesis and consequently increase pDNA production. GALG20 (MG1655*ΔendAΔrecAΔpgi*) produced 25-fold more pDNA (19.1 mg/g DCW) than its parental strain, MG1655*ΔendAΔrecA* (0.8 mg/g DCW), in glucose. For the first time *pgi* was identified as an important target for constructing a high-yielding pDNA production strain.

Introduction

Non-viral gene therapy is a promising approach for the treatment of genetic disorders and acquired diseases, and for the prevention of infectious diseases. The discovery that naked plasmid DNA (pDNA) could mediate gene transfer and expression *in vivo* (Wolff 1990) initiated the emergence of a new class of medicinal agents (Prazeres 2011). The efforts to develop such plasmid biopharmaceuticals have increased over the last several years, especially after the approval of the first veterinary DNA vaccines (Han 2009; Williams 2009). The relevance of non-viral gene therapy is underscored by the fact that plasmid DNA accounts for 25% of the vectors used in clinical trials (http://www.wiley.com//legacy/wileychi/genmed/clinical/).

One of the aims of research on plasmid biopharmaceuticals has been the development of *Escherichia coli (E. coli)* host strains capable of producing the large amounts of pDNA required for clinical trials and eventually commercialization (Bower and Prather 2009; Goncalves 2011; Lara and Ramirez 2012). Roughly speaking, volumetric pDNA titers lower than 250 mg/L are typically obtained in laboratory-scale fermentations (Prazeres 2011), whereas titers of the order of 1000-2100 mg/L have been reported in high-yield, preindustrial fermentations (Carnes 2006; Listner 2006; Phue 2008; Luke 2009; Williams 2009). On a dry cell weight (DCW) basis, the plasmid content of *E. coli* cells may be expected to vary anywhere between 0.5 and 5.0 % w/w, depending on the plasmid type, strain, growth conditions and growth phase (Prazeres 2011). *E. coli* host strains of the K-12 and B type with diverse mutagenized genetic backgrounds, such as DH5, DH5α, DH10B, JM108, JM101 and BL21, have all been used for pDNA production (Bower and Prather 2009; Carnes and Williams 2007; Cunningham 2009; Prather 2003; Williams 2009). However, most of these strains were originally developed to facilitate the cloning of heterologous genes and for the production of recombinant proteins, and thus may not be the most appropriate for pDNA

manufacturing. A more rational approach has been pursued recently, whereby genes are selected and mutated on the basis of *E. coli* central carbon metabolism and plasmid properties. This has led to the generation of a plethora of new pDNA production strains, such as DH5α*ΔfruR* (Ow 2007), BL21*ΔendAΔrecA* (Phue 2008), JM101*ΔpykFΔpykA* (Cunningham 2009), DH5α*zwf:rpiA*$^{+}$ (Williams 2009) and W3110*ΔPTSGalP*$^{+}$*ΔpykA*(Pablos 2011).

Reducing acetate and enhancing nucleotide production are some of the strategies that are used to increase pDNA yields. Knockouts of selected genes, such as *pykF, pykA, ackA-pta* and *poxB* have been explored to reduce acetic acid formation in pDNA production strains (Carnes 2011; Cunningham 2009; Pablos 2011). Genes in the pentose phosphate pathway such as *zwf* and *rpiA* have also been overexpressed in order to enhance nucleotide production for pDNA synthesis (Wang 2006; Williams 2009). However, most of these mutations were made in strains with highly mutagenized genetic backgrounds. Thus, it is unclear whether the specific strain background had an effect on the mutations introduced. In addition, these highly mutated strains exhibit a lower growth rate than parental (wild-type) strains, a characteristic that suggests a less healthy organism.

The trade-off between strain engineering and fermentation strategy is still a hurdle for the development of new pDNA production strains. Differences in fermentation strategies (batch or fed-batch) (O'Kennedy *et al.,* 2003; Ow *et al.,* 2007; Ow *et al.,* 2009), medium composition and carbon source (glucose or glycerol) (Oh and Liao 2000) were previously shown to affect *E. coli* metabolism and should also be taken into consideration when producing pDNA. It is well known that glycerol has the advantage of minimizing acetate formation in fermentation processes and it is becoming an inexpensive and attractive carbon source (Carnes 2006; Carvalho 2011). On the other hand, glucose has been used for several years in high-density cultures and for pDNA production and is both accessible and

inexpensive (Luli and Strohl 1990). High acetate formation has been observed with glucose; however, accumulation profiles depend on the strain (Phue 2005). *E. coli* strains have been engineered to optimize glucose uptake and reduce acetate formation (De Anda 2006). For example, the strain VH33 (PTS^-GalP^+), a derivative of W3110 with an inactive phosphotransferase system (PTS) and a strong promoter for *galP*, has shown improved production of pDNA (Pablos 2011; Soto 2011). Interestingly, high yield pDNA production was shown with strain BL21*ΔendAΔrecA* in both glucose and glycerol, while DH5α produced high pDNA yields only in glycerol (Phue 2008).

In this work, we set out to enhance pDNA production by rationally mutating key genes of the glycolytic pathway. We first chose genes which had already been shown to increase pDNA yields, *pykF* and *pykA* (Cunningham 2009; Pablos 2011), and deleted them in two different strains: MG1655*ΔendAΔrecA*, a nearly wild-type genetic background; and DH5α, a commonly used, but highly mutagenized strain. Deletions of *endA* and *recA* were made in MG1655 in order to minimize recombination and non-specific digestion of DNA (Phue 2008; Summers 1998). We also analyzed the influence of carbon source (glucose versus glycerol) among the different strains developed. Finally, we explored the effect of a previously untested mutation with the introduction of a new knockout, *pgi*, in MG1655*ΔendAΔrecA* to redirect carbon flow into the pentose phosphate pathway (PPP) and promote nucleotide synthesis.

Material and Methods

Strains and plasmids

The bacterial strains used in this study are indicated in Table 1. MG1655 was kindly donated by Professor Gregory Stephanopoulos of the Department of Chemical Engineering at the Massachusetts Institute of Technology (Cambridge, MA, USA). Gene deletions of *endA*,

recA, *pykF*, *pykA* and *pgi* in MG1655 and JM101were carried out with P1 transduction using the following strains from the Keio collection (Baba 2006) as donors: JW1666-1 [*ΔpykF751::kan*]; JW1843-1 [*ΔpykA779::kan*]; JW3985-1 [*Δpgi-721::kan*]; JW2912-1 [*ΔendA720::kan*] and JW2669-1 [*ΔrecA774::kan*]. The *recA* gene was always the last to be deleted, since P1 transduction requires an active RecA recombinase. The kanamycin cassette was removed using plasmid pCP20 as described by Datsenko and Wanner (2000), and successfully constructed mutant strains were verified by colony PCR using appropriate primers. The standard protocol for inactivation of chromosomal genes (Datsenko and Wanner 2000) was adapted to knock-out *pykF* and *pykA* in DH5α. Briefly, DH5α was transformed with a *recA* containing plasmid, pKD46*recA*+, and RecA recombinase expression was induced with arabinose. DH5α(pKD46*recA*+) was transformed with the specific PCR fragment generated using appropriate primers, and homologous recombinants were selected with kanamycin. The plasmid pKD46*recA*+ is temperature sensitive and was cured by raising the temperature. The remaining steps of the protocol were performed as described previously by Datsenko and Wanner (2000). All strains were transformed with plasmid pVAX1GFP (3697 bp), derived from Invitrogen's (Carlsbad, CA) pVAX1LacZ as described previously (Azzoni 2007).

Medium and growth conditions

Transformed strains were maintained on agar plates containing 30 µg/mL of kanamycin. A seed bank was prepared from single colonies picked from each plate and inoculated in Luria Bertani (LB) medium supplemented with 30 µg/mL of kanamycin. Cultures were grown to mid-exponential phase at 250 rpm, 37ºC and frozen at -80ºC in 15% (v/v) final concentration glycerol. The inoculum was prepared from frozen seed stocks in complex medium [Bacto peptone, 10 g/L; yeast extract, 10 g/L; $(NH_4)_2SO_4$, 3 g/L; K_2HPO_4, 3.5 g/L; KH_2PO_4, 3.5 g/L; thiamine, 199 mg/L; $MgSO_4$, 1.99 g/L; trace element solution, 1

mL/L (Listner 2006)] supplemented with 30 µg/mL of kanamycin, grown to mid-exponential phase and then used to inoculate batch cultures to an initial optical density at 600 nm (OD_{600}) of approximately 0.1. Cultures were grown at 37ºC for 24 hours in 250 mL shake flasks containing 50 mL of complex medium supplemented with 30 µg/mL of kanamycin, initial pH of 7.1 and aeration by shaking at 250 rpm. Glucose and glycerol were used as the primary carbon sources as indicated in the Results. Cells were sampled at 6, 12 and 24 hours to quantify glucose, glycerol and acetate, and at 12 and 24 hours to quantify pDNA.

Biomass quantification

Samples were taken every 3 hours to determine biomass concentration. OD_{600} was measured in a Beckman Coulter DU 800 spectrophotometer. Dry cell weight (DCW) was determined using a vacuum filtration system. 10 mL of each sample was filtered in previously weighed filter papers and then dried at 42ºC. A linear correlation was determined between cell concentration (g/L) and OD_{600nm}.

Plasmid DNA quantification

Plasmid DNA was quantified from crude lysates prepared from OD_{600nm} = 10 cell pellets using the method described by Listner (2006). The method was modified slightly: cell pellets were harvested by centrifugation at 5000 x *g* for 15 minutes, the 37°C incubation took place with 250 rpm shaking, and 5 µL of RNase A solution (10 mg/mL) was used per mL of lysate. The resulting lysates were analyzed using a Gen-Pak FAX anion-exchange column (Waters Corporation) on an Agilent 1100 Series HPLC system. The HPLC method was run at a constant flow rate of 0.75 mL/min and consisted of a 10-minute linear NaCl gradient from 300 mM to 660 mM, followed by 5.5 minutes at 1 M NaCl, after which the column was flushed with 0.04 M phosphoric acid for 4.5 minutes. Before the next sample injection, the column was equilibrated with 300 mM NaCl for 10 min. Plasmid DNA eluted at

approximately 610 mM NaCl and was detected at 260 nm with a diode array detector (DAD). A highly linear standard curve of pVAX1GFP was prepared using pDNA purified with the Hi-Speed QIAfilter Plasmid Maxi Kit (Qiagen) and quantified using absorbance at 260 nm on an Implen NanoPhotometer.

Metabolite quantification

To determine the concentration of glucose, glycerol and acetic acid, culture samples were centrifuged at 5000 x *g* for 15 minutes, and the aqueous supernatant was used for HPLC analysis on an Agilent 1100 Series HPLC system equipped with an Aminex HPX-87 H anion exchange column (Bio-Rad Laboratories) and refractive index detector (RID). The LC method was run at a constant flow rate of 0.6 mL/min with 5 mM H_2SO_4 as the mobile phase, at 50ºC for 25 min.

Statistical analyses

The effect of the gene knockouts was measured by pDNA yield quantification. T-tests for independent samples (SSPS Statistics) were done to identify significant differences in pDNA yields between parental and mutated strains. Differences were considered statistically significant when the *P* value was < 0.05. At least three independent duplicate experiments were conducted in order to confirm the results. Standard error of mean (SEM) was calculated for all measurements of pDNA yields, biomass, acetate and growth rate.

Results

Effect of carbon source on *pykF* and *pykA* knockout strains

Cuningham (2009) demonstrated the benefits of the double *pykF-pykA* mutations in a JM101 strain grown on glucose (Fig. 1). We chose to investigate these same mutations in the nearly wild-type MG1655*ΔendAΔrecA* and in the commonly used DH5α strain, using either

glycerol or glucose as the primary carbon source. The first experiments were conducted with an initial concentration of glycerol and glucose of 20 g/L and with MG1655*ΔendAΔrecA,* GALG1011 (MG1655*ΔendAΔrecAΔpykFΔpykA*), DH5α and DH5α*ΔpykFΔpykA,* all carrying pVAX1GFP. MG1655*ΔendAΔrecA* and DH5α grew faster than their respective *pykF* and *pykA* double knockout mutant strains in both glucose and glycerol. However, higher biomass content was obtained in glycerol. GALG1011 produced less acetate than MG1655*ΔendAΔrecA* in both carbon sources. In contrast, DH5α and DH5α*ΔpykFΔpykA* produced similar amounts of acetate. In general, less acetate was obtained in glycerol for all strains (Table 2).

The double knockout of *pykF* and *pykA* had a negative impact on pDNA production in both strains when using glycerol as carbon source (Table 3). On the other hand, the double knockout increased pDNA yield 2-fold in GALG1011 *vs.* MG1655*ΔendAΔrecA* when glucose was the main carbon source. In glucose, the phosphotransferase system (PTS) is able to couple pyruvate formation with carbohydrate transport, partially mitigating the effect of *pyk* deletions. However, the PTS system is not utilized for glycerol uptake and hence this additional source of pyruvate is not available (Oh and Liao 2000). Nevertheless, pDNA yields were significantly lower in glucose than in glycerol for all Pyk mutant strains when the initial concentration of both carbon sources was 20 g/L (Table 3). It has been reported that 10 g/L of glucose can inhibit pDNA synthesis in *E. coli* B (Zhi-nan 2005) and that *E. coli* K-12 grows slower and accumulates more acetate than B strains at high glucose concentrations (Phue 2005). Therefore, inhibition of pDNA synthesis could also be expected in K-12 strains with high glucose. Of the strains and conditions examined, MG1655*ΔendAΔrecA* was the best producer of pDNA in glycerol, although it produced the highest quantity of acetate (Tables 2 and 3). It is known that acetate secretion can shuttle carbon away from nucleotide

synthesis which could be disadvantageous for pDNA formation, but the exact effect of acetate on pDNA production is not yet elucidated (Carnes 2011; Wang 2006).

Glucose inhibition in plasmid DNA production

To determine the effect of glucose concentration on pDNA production, MG1655*ΔendAΔrecA* and GALG1011 bearing pVAX1GFP were grown in three different conditions: 5 g/L of glucose initially plus 10 g/L of glucose after 12 hours (5+10 g/L), 10 g/L of glucose initially plus 10 g/L of glucose after 12 hours (10+10 g/L), and 20 g/L with no extra addition of glucose. MG1655*ΔendAΔrecA* and GALG1011 produced 5-fold and 4-fold more pDNA, respectively, in (5+10 g/L) than 20 g/L of glucose (Table 3). However, when the initial glucose concentration was 5 g/L, the difference in pDNA specific yield (mg/g DCW) between these strains was less than 2-fold (Table 3). In (5+10 g/L) glucose, GALG1011 still produced less acetate and grew slower than MG1655*ΔendAΔrecA* (Table 2). Similar biomass was formed for both strains, and the final biomass achieved in (5+10 g/L) glucose was comparable to that obtained in glycerol, though the pDNA yield was still higher in glycerol. Experiments were also performed with glucose addition after 9 hours, but no difference was observed in pDNA yields (data not shown).

After determining the best glucose concentration, (5+10 g/L), for pDNA production in MG1655*ΔendAΔrecA* and GALG1011, we tested DH5α and DH5α*ΔpykFΔpykA,* in order to determine whether the negative performance of the double mutant with 20 g/L glucose (Table 3) was due to inhibition or to strain genetic background. pDNA production increased 2-fold in DH5α when moving from 20 g/L glucose to (5+10 g/L) glucose. However, the negative effect of *pykF* and *pykA* knockouts was maintained. Therefore, it appears that the genetic background of DH5α is suppressing the positive effects of *pykF* and *pykA* gene knockouts observed in GALG1011.

Effect of strain background on *pykF* and *pykA* double knockouts

As stated previously, the beneficial effects of the *pykF-pykA* double knockout for pDNA production were demonstrated in JM101 grown in glucose (Cunningham 2009). There are no reports about the behavior of JM101*ΔpykFΔpykA* bearing plasmid in glycerol. To further examine the effect of strain genetic background, we decided to construct JM101*ΔpykFΔpykA* and to test it in the same rich medium and conditions as GALG1011 and DH5α*ΔpykFΔpykA*. The loss of Pyk is expected to reduce acetate accumulation, and glycerol is frequently used for the same purpose. Thus, the double mutations are not expected to favor pDNA production in glycerol over the unmutated strain; however, because this carbon source was not previously tested, it is unclear if the effect of the mutations would be negative or neutral. Consistent with the results presented by Cunningham (2009), JM101*ΔpykFΔpykA* produced 2-fold more pVAX1GFP (5.3 or 2.5 mg/g DCW) than JM101 (2.6 or 1.3 mg/g DCW) when grown in 5 or 20 g/L glucose, respectively. However, as with the other engineered strains, JM101*ΔpykFΔpykA* (2.2 mg/g DCW) produced less pDNA than JM101 in glycerol-rich medium (8.4 mg/g DCW; Table 3). In general, the double knockout of *pykF* and *pykA* was effective in MG1655*ΔendAΔrecA* and JM101 in glucose, but it was not beneficial in DH5α. The highly mutagenized genetic background of DH5α appears to affect the *pykF* and *pykA* mutation.

Effect of single knockouts of *pykF* or *pykA* gene

The previous results showed that the *pykF* and *pykA* double knockout can be beneficial in glucose and that the mutation had a positive effect in MG1655*ΔendAΔrecA* and JM101 (Table 3). Nevertheless, the *pykF* and *pykA* mutant strain GALG1011 still produced much less plasmid in glucose (6.6 mg/g DCW) than the nearly wild-type strain, MG1655*ΔendAΔrecA,* in glycerol (11.2 mg/g DCW), (Table 3).

In an effort to increase pDNA yields, we decided to investigate *pykF* or *pykA* single knockouts in MG1655*ΔendAΔrecA*. A single knockout could be more effective than the double knockout for pDNA production in glucose because more pyruvate would be formed and higher ATP yields would be obtained through glycolysis. At the same time, carbon flux would be boosted to the pentose phosphate pathway, increasing nucleotide synthesis (Siddiquee 2004). For example, Pablos (2011) showed that a single *pykA* knockout increased pDNA production 3-fold in W3110*ΔPTSGalP*$^{+}$ in glucose.

To determine the effect of a *pykF* or *pykA* single knockout in MG1655*ΔendAΔrecA,* we generated strains GALG10 (MG1655*ΔendAΔrecAΔpykF*) and GALG11 (MG1655*ΔendAΔrecAΔpykA*) and tested them in glucose and glycerol under the same conditions as described previously (Table 3). In both carbon sources, one single mutation was found to be more effective for pDNA production than the double mutation. In glycerol, MG1655*ΔendAΔrecA,* GALG10 and GALG11 produced similar amounts of pDNA (11.2, 8.6 and 10.3 mg/g DCW). In contrast, GALG10 (10.5 mg/g DCW) and GALG11 (13.1 mg/g DCW) produced more pDNA than MG1655*ΔendAΔrecA* (3.6 mg/g DCW) or GALG1011 (6.6 mg/g DCW) in glucose. GALG11 achieved high pDNA yields in glucose (13.1 mg/g DCW), even in comparison with the yields obtained with MG1655*ΔendAΔrecA* in glycerol (11.2 mg/g DCW), (Table3). Nevertheless, both GALG10 and GALG11 produced little pDNA when the initial concentration of glucose was 20 g/L (≤1 mg/g DCW).

Single mutations in *pykF* or *pykA* genes were also constructed in DH5α, generating the strains DH5α*ΔpykF* and DH5α*ΔpykA.* Surprisingly, DH5α*ΔpykF* (5.9 mg/g DCW) produced more pDNA than DH5α (1.8 mg/g DCW) in glucose. On the contrary, DH5α*ΔpykA* (0.9 mg/g DCW) displayed a similar behaviour to DH5α*ΔpykFΔpykA* (0.3 mg/g DCW) and produced much less pDNA. Although the *pykF* single knockout had positive effects in DH5α,

the final pDNA yields obtained with DH5α and DH5α*ΔpykF* bearing pVAX1GFP were lower than the ones obtained with GALG10 and GALG11 (Table 3).

Effect of *pgi* knockout on plasmid DNA production

The main advantages of removing pyruvate kinase (Pyk) in a pDNA production strain through the deletion of both *pykF* and *pykA* genes would be the reduction of acetate and increase of carbon flux into the pentose phosphate pathway (PPP). On the other hand, the total lack of Pyk could reduce energy available for the cells by reducing the ATP content. In this work, we have shown that deletion of both pyruvate kinase isozymes increased pDNA production in the parental strain MG1655*ΔendAΔrecA* in glucose. However, higher amounts of pDNA were obtained when only pyruvate kinase I or II were absent in the parental strain in glucose, a result that suggests the importance of generating some pyruvate and completing glycolysis.

To increase the pentose phosphate pathway flux without compromising the energy obtained from the generation of pyruvate, we tried a different strategy in order to enhance nucleotide production and reduce glycolytic flux at the same time (Fig. 1). It has been reported that strains carrying high copy pDNA require extra synthesis of nucleotides and that the carbon flux into the PPP may not be sufficient to meet cellular needs (Flores 2004). The elimination of phosphoglucose isomerase (*pgi*) aims to redirect the carbon flux into the PPP, enhance synthesis of nucleotides, and also provide high amounts of reducing cofactors (i.e., NADPH). Pgi mutant strains were recently reported to improve the production of xylitol (Chin and Cirino 2011; Chin 2009), chiral compounds for the pharmaceutical market (Siedler 2011), and second generation biofuels (Yao 2011).

GALG20 (MG1655*ΔendAΔrecAΔpgi*) produced 3-fold more pVAXGFP pDNA (11.6 mg/g DCW) than MG1655*ΔendAΔrecA* in (5+10 g/L) glucose (3.6 mg/g DCW). In glycerol,

GALG20 produced similar amounts of pDNA as MG1655*ΔendAΔrecA* (Fig. 2). Surprisingly, this strain had the best performance in 20 g/L of glucose, producing 25-fold more pDNA (19.1 mg/g DCW) than the parental strain (0.8 mg/g DCW), and 46% or 65 % more, respectively, than the previously best identified performers, GALG11 in 5+10 g/L glucose (13.1 mg/g DCW) and MG1655*ΔendAΔrecA* in 20 g/L glycerol (11.5 mg/g DCW).

Discussion

The *de novo* design of *E. coli* strains specifically for pDNA amplification is expected to enhance production yields when compared to strains with highly mutagenized genetic backgrounds. In this study, we performed a series of mutations in the wild type MG1655 strain and in the laboratory strain DH5α in order to improve pDNA yields. The genes coding for pyruvate kinase (*pykA*, *pykF*) and phosphoglucose isomerase (*pgi*) were selected as targets due to their role in the PPP (Fig. 1).

In order to reduce acetate and increase carbon flux into the PPP, a double knockout of *pykF* and *pykA* was first created. Our results show that the strain genetic background can indeed interfere in the outcome of mutations which are otherwise rationally designed. Of the Pyk mutant strains examined, the strains derived from the nearly wild type MG1655*ΔendAΔrecA* (GALG10 and GALG11) were the best producers of pDNA, whereas DH5α*ΔpykFΔpykA* and DH5α*ΔpykA* were the worst producers of pDNA in glucose (Table 3). The negative effect of the *pykF* and *pykA* double knockout on the production of pDNA in DH5α could be ascribed to the interaction between the strain genotype and the new mutations added to the strain, though it is not clear which specific mutations result in the negative phenotype. In contrast, DH5α*ΔpykF* produced more plasmid than DH5α in glucose, a result that points to an essential role of the *pykA* gene in DH5α. Although, DH5α*ΔpykF* had

increased pDNA yields when compared to DH5α, the obtained yields were lower than those obtained with GALG10 and GALG11. Other advantages associated with the use of the nearly wild-type strain MG1655*ΔendAΔrecA* as a starting point for strain engineering instead of DH5α include a more robust growth and a lack of auxotrophies that can benefit pDNA production.

Strains with single mutations of either *pykF* or *pykA* were found to be more efficient in producing pDNA than strains containing the double mutation. The synthesis of pDNA was likely more advantageous when the pyruvate kinase activity was reduced as opposed to abolished because glycolytic flux was down-regulated while still enabling additional ATP production and up-regulation of glucose flux through the PPP. For example, Siddiquee (2004) demonstrated that a *pykF* single knockout in *E. coli* increased glucose flux through the PPP. Moreover, linear programming has predicted different carbon fluxes for *E. coli* lacking pyruvate kinase, resulting in different ATP yields (Phalakornkule 2001). Since cells, as living organisms, are difficult to phenotypically predict, different strains (MG1655 and DH5α) with the same gene knockout (*pykA* or *pykF*) can have diverse carbon fluxes and result in different ATP yields, favoring pDNA synthesis or not.

The abolishment of the *pgi* gene should redirect glucose-6-phosphate preferentially to the PPP, but glycolysis would continue due to the generation of fructose-6-phosphate and glyceraldeyde-3-phosphate (Fig. 1). In general, high amounts of NADPH, nucleotides and ATP would be obtained. Indeed, the inactivation of the *pgi* gene in *E. coli* previously resulted in the redirection of glucose predominantly via the PPP, but a minor fraction of glucose was still catabolized in the Entner-Doudoroff Pathway (EDP) (Canonaco 2001). A large excess of NADPH was detected in Pgi mutant strains (Canonaco 2001; Chin and Cirino 2011; Siedler 2011), which was shown to improve biotransformations in various processes. However, an excess of NADPH can cause a redox imbalance in the cell, imposing stress that may appear

as a large reduction in growth rate (Charusanti 2010). In the specific case of pDNA synthesis, a mathematical model has demonstrated that high generation of NADPH would be effective for increasing yields (Cunningham 2009). Nevertheless, no Pgi mutant strain has been previously reported for the purpose of producing pDNA.

Together the benefits of increasing glucose flux into the PPP, such as enhancement of nucleotides and of NADPH generation, turn a Pgi mutant strain into a high potential pDNA production strain. In fact, GALG20 produced 25-fold more pDNA than its parental strain in 20 g/L of glucose. A previous study has shown that catabolite repression can be relaxed by *pgi* mutants (Yao 2011). Indeed, large amounts of biomass (8.8 g/L) and pDNA (19.1 mg/g DCW) and low amounts of acetic acid (0.02 g/L) were formed by GALG20 in high concentrations of glucose. Another advantage of GALG20 for pDNA production was that the growth rate was similar to the parental strain in glucose (Table 2), whereas a previous study demonstrated a significant reduction of growth rate in a *pgi* mutant strain grown in the same carbon source (Ahn 2011). The differences in the behavior of these two *pgi* mutants in glucose could be associated with the strain genetic background between the two studies, *E. coli* MG1655*ΔendAΔrecA* and KPM SA1 (*ΔaraLΔaraK*), respectively. In the current work, we have demonstrated that strain background can substantially influence the effect of gene knockouts, such as the differences between the *pykF* and *pykA* mutations in MG1655*ΔendAΔrecA* versus DH5α.

In the *pgi* mutant strain, the level of glucose-6-phosphate is significantly higher, which can destabilize the PTS system and in turn, alter the level of cyclic AMP (cAMP) (Morita 2003). On the other hand, *pgi* mutant cells up-regulate the catabolite repressor activator (Cra), due to the decrease in fructose 6-phosphate (Yao 2011). Nevertheless, it is not clear yet how cAMP or Cra mechanisms of gene regulation would correlate with pDNA synthesis. Previous studies suggested that increasing cAMP levels could be associated with

enhancement of pDNA yields (Cunningham 2009), and the elimination of Cra was shown to increase pDNA (Ow 2009). Further work would be necessary to elucidate the role played by cAMP and Cra in plasmid DNA synthesis, and *pgi* mutant cells could be a useful target to study metabolic network behavior associated with pDNA production. The regulation of glucose-6-phosphate-1-dehydrogenase (Zwf) or other enzymes in the PPP could limit pDNA synthesis. Based on previous evidence (Wang 2006), ribose phosphate isomerase (Rpi) could also limit nucleotide production because the relative amount of ribose 5-phosphate (R5P) and xylose 5-phosphate (Xu5P) depends on cellular needs. Overexpression of Zwf and Rpi would be potential alternatives to increase pDNA yields in *pgi* mutant cells.

Among the strains generated and analyzed we selected the top three pDNA producer strains and compared their performance with the commonly used lab strain, DH5α (Table 4). GALG20 and GALG11 performed better in glucose while MG1655*ΔendAΔrecA* and DH5α performed better in glycerol. The *pgi* mutation in the MG1655*ΔendAΔrecA* wild-type strain resulted in the best volumetric (140.8 mg/L) and specific (19.1 mg/g DCW) pDNA yields. To our knowledge, this is the highest reported value for specific plasmid yield for a constant temperature batch process. This figure is only comparable to the 19.2 mg/g DCW obtained by Ow (2009) in a fed-batch system with a DH5αΔFruR strain at 37ºC. Further experiments will be carried out and different fermentation strategies will be explored in order to investigate the true potential of these high pDNA producer strains.

Acknowledgments

This work was supported by the MIT-Portugal Program and Fundação para a Ciência e a Tecnologia (PhD grant SFRH/BD/33786/2009 to Geisa A. L. Gonçalves). We also acknowledge Kevin Solomon (MIT) for providing the plasmid pKD46recA+, Sang-Hwal Yoon (MIT) for constructing MG1655*ΔendAΔrecA*, and Diana Bower (MIT) for the

development of HPLC methods. Special acknowledgment to all members of the Prather Research Group for providing helpful insights that contributed to this work.

The authors have declared no conflict of interest.

References

Ahn J, Chung BKS, Lee DY, Park M, Karimi IA, Jung JK, Lee H (2011) NADPH-dependent *pgi*-gene knockout *Eschericha coli* metabolism producing shikimate on different carbon sources. FEMS Microbiol Lett 324, 10-16.

Azzoni AR, Ribeiro SC, Monteiro GA, Prazeres DMF (2007) The impact of polyadenylation signals on plasmid nuclease-resistance and transgene expression. J Gene Med 9:392 - 402.

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2:2006 - 2008.

Bower DM, Prather KLJ (2009) Engineering of bacterial strains and vectors for the production of plasmid DNA. Appl Microbiol Biotechnol 82:805 - 813.

Canonaco F, Hess TA, Heri S, Wang T, Szyperski T, Sauer U (2001) Metabolic flux response to phosphoglucose isomerase knock-out in *Escherichia coli* and impact of overexpression of the soluble transhydrogenase UdhA. FEMS Microbiol Lett 204(2):247 - 252.

Carnes AE, Hodgson CP, A.Williams J (2006) Inducible *Escherichia coli* fermentation for increased plasmid DNA production. Biotecnhol Appl Biochem 45:155 - 166.

Carnes AE, Luke JM, Vincent JM, Shukar A, Anderson S, Hodgson CP, Williams JA (2011) Plasmid DNA fermentation strain and process-specific effects on vector yield, quality, and transgene expression. Biotechnol Bioeng 108(2):354-363.

Carnes AE, Williams JA (2007) Plasmid DNA manufacturing technology. Recent Pat Biotechnol 1(2):1-16.

Carvalho RJ, Cabrera-Crespo J, Tanizaki MM, Goncalves VM (2011) Development of production and purification processes of recombinant fragment of pneumococcal surface protein A in *Escherichia coli* using different carbon sources and chromatography sequences. Appl Microbiol Biotechnol.

Charusanti P, Conrad TM, Knight EM, Venkataraman K, Fong NL, Xie B, Gao Y, Palsson BO (2010) Genetic basis of growth adaptation of *Escherichia coli* after deletion of *pgi*, a major metabolic gene. PLoS Genet 6(11):e1001186.

Chin JW, Cirino PC (2011) Improved NADPH supply for xylitol production by engineered *Escherichia coli* with glycolytic mutations. Biotechnol Prog 27(2):333 - 341.

Chin JW, Khankal R, Monroe CA, Maranas CD, Cirino PC (2009) Analysis of NADPH supply during xylitol production by engineered *Escherichia coli*. Biotechnol Bioeng 102(1):209 - 220.

Cunningham DS, Koepsel RR, Ataai MM, Domach MM (2009) Factors affecting plasmid production in *Escherichia coli* from a resource allocation standpoint. Microb Cell Fact 8:1 - 27.

Cunningham DS, Liu Z, Domagalski N, Koespsel RR, Ataai MM, Domach MM (2009) Pyruvate kinase-deficient *Echerichia coli* exhibits increased plasmid copy number and cyclic AMP levels. J Bacteriol 191(9):3041 - 3049.

Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A 97(12):6640 - 6645.

De Anda R, Lara AR, Hernandez V, Hernandez-Montalvo V, Gosset G, Bolivar F, Ramirez OT (2006) Replacement of the glucose phosphotransferase transport system by

galactose permease reduces acetate accumulation and improves process performance of *Escherichia coli* for recombinant protein production without impairment of growth rate. Metab Eng 8(3):281 - 290.

Flores S, Anda-Herrera R, Gosset G, Bolívar FG (2004) Growth-rate recovery of *Escherichia coli* cultures carrying a multicopy plasmid, by engineering of the pentose-phosphate pathway Biotechnol Bioeng 87(4):485 - 494.

Goncalves GA, Bower DM, Prazeres DM, Monteiro GA, Prather KL (2011) Rational engineering of *Escherichia coli* strains for plasmid biopharmaceutical manufacturing. Biotechnol J. DOI: 10.1002/biot.201100062

Han Y, Liu S, Ho J (2009) Using DNA as a drug—Bioprocessing and delivery strategies. Chem Eng Res Des 87:343 - 348.

Lara AR, Knabben I, Regestein L, Sassi J, Caspeta L, Ramırez OT, Buchs J (2011) Comparison of oxygen enriched air vs. pressure cultivations to increase oxygen transfer and to scale-up plasmid DNA production fermentations. Eng Life Sci 11(4):382 - 386.

Lara AR, Ramirez OT (2012) Plasmid DNA production for therapeutic applications. Methods Mol Biol 824:271-303.

Listner K, Bentley LK, Chartrain M (2006) A simple method for the production of plasmid DNA in bioreactors. Methods Mol Med 127:295 - 309.

Luke J, Carnes, AE, Hodgson, CP, Williams, JA (2009) Improved antibiotic-free DNA vaccine vectors utilizing a novel RNA based plasmid selection system. Vaccine 27:6454 - 6459.

Luli GW, Strohl WR (1990) Comparison of growth, acetate production and acetate inhibtion of *Escherichia Coli* strains in batch and fed-batch fermentations. Appl Environ Microbiol 56:1004 - 1011.

Morita T, El-Kazzaz W, Tanaka Y, Inada T, Aiba H (2003) Accumulation of glucose 6-phosphate or fructose 6-phosphate is responsible for destabilization of glucose transporter mRNA in *Escherichia coli*. The journal of biological chemestry 278(18):15608 - 15614.

Murarka A, Dharmadi Y, Yazdani SS, Gonzalez R (2008) Fermentative utilization of glycerol by *Escherichia coli* and its implications for the production of fuels and chemicals. Appl Environ Microbiol 74(4):1124 - 1135.

O'Kennedy RD, Ward JM, Keshavarz-Moore E (2003) Effects of fermentation strategy on the characteristics of plasmid DNA production. Biotechnol Appl Biochem 37(Pt 1):83-90.

Oh M, Liao JC (2000) Gene expression profiling by DNA microarrays and metabolic fluxes in *Escherichia coli*. Biotechnol Prog 16:278 - 286.

Ow DS-W, Lee RM-Y, Nissom PM, Philp R (2007) Inactivating FruR global regulator in plasmid-bearing *Escherichia coli* alters metabolic gene expression and improves growth rate. J Biotechnol 131:261 - 269.

Ow DS-W, Yap MG-S, Oh SK-W (2009) Enhancement of plasmid DNA yields during fed-batch culture of a fruR- knockout *Escherichia coli* strain Biotechnol Appl Biochem 52:53 - 59.

Pablos TE, Soto R, Mora EM, Le Borgne S, Ramirez OT, Gosset G, Lara AR (2011) Enhanced production of plasmid DNA by engineered *Escherichia coli* strains. J Biotechnol 158:201 – 212.

Phalakornkule C, Lee S, Zhu T, Koepsel R, Ataai MM, Grossmann IE, Domach MM (2001) A MILP-based flux alternative generation and NMR experimental design strategy for metabolic engineering. Metab Eng 3(2):124 - 137.

Phue J-N, Lee SJ, Trinh L, Shiloach J (2008) Modified *Escherichia coli* B (BL21), a superior producer of plasmid DNA compared with *Escherichia coli* K (DH5a). Biotechnol and Bioeng 101(4):831 - 836.

Phue J, Noronha SB, Hattacharyya R, Wolfe AJ, Shiloach J (2005) Glucose metabolism at high density growth of *E. coli* B and *E. coli* K: differences in metabolic pathways are responsible for efficient glucose utilization in *E. coli* B as determined by microarrays and nothern blot analyses. Biotechnol and Bioeng 90(7):805 - 820.

Prather KJ, Sagar S, Murphy J, Chartain M (2003) Industrial scale production of plasmid DNA for vaccine and gene therapy: plasmid design, production, and purification. Enzyme Microb Technol 33:865 - 883.

Prazeres, DMF (2011) Plasmid Biopharmaceuticals: Basics, Applications and Manufacturing. John Eiley & Sons, Inc., Hoboken, NJ.

Siddiquee KA, Arauzo-Bravo MJ, Shimizu K (2004) Effect of a pyruvate kinase (*pykF*-gene) knockout mutation on the control of gene expression and metabolic fluxes in *Escherichia coli.* FEMS Microbiol Lett 235(1):25 - 33.

Siddiquee KAZ, Arauzo-Bravo MJ, Shimizu K (2004) Metabolic flux analysis of *pykF* gene knockout *Escherichia coli* based on ^{13}C-labeling experiments together with measurements of enzyme activities and intracellular metabolite concentrations. Appl Microbiol Biotechnol 63:407 - 417.

Siedler S, Bringer S, Bott M (2011) Increased NADPH availability in *Escherichia coli*: improvement of the product per glucose ratio in reductive whole-cell biotransformation. Appl Microbiol Biotechnol 92(5):929 - 937.

Soto R, Caspeta L, Barrón B, Gosset G, Ramírez OT, Lara AR (2011) High cell-density cultivation in batch mode for plasmid DNA production by a metabolically engineered *E. coli* strain with minimized overflow metabolism. Biochem Eng J 56:165-171.

Summers D (1998) Timing, self-control and a sense of direction are the secrets of multicopy plasmid stability. Mol Microbiol 29:1137 - 1141.

Wang Z, Xiang L, Shao J, Wegrzyn A, Wegrzyn G (2006) Effects of the presence of ColEI plasmid DNA in *Escherichia coli* on the host cell metabolism. Microb Cell Fact 5:34.

Williams JA, Carnes AE, Hodgson CP (2009) Plasmid DNA vaccine vector design: impact on efficacy, safety and upstream production. Biotechnol Adv 27(4):353 - 370.

Williams JA, Luke J, Langtry S, Anderson S, Hodgson CP, Carnes AE (2009) Generic plasmid DNA production platform incorporating low metabolic burden seed-stock and fed-batch fermentation processes. Biotechnol Bioeng 103(6):1129 - 1143.

Wolff AJ, Malone RW, Williams P, Chong W, Acsadi G, Jani A (1990) Direct gene transfer into mouse muscle *in vivo*. Science 247:1465 - 1468.

Yao R, Hirose Y, Sarkar D, Nakahigashi K, Ye Q, Shimizu K (2011) Catabolic regulation analysis of *Escherichia coli* and its *crp, mlc, mgsA, pgi* and *ptsG* mutants. Microb Cell Fact 10:67.

Zhi-nan X, Wen-he S, Hao C, Pei-lin C (2005) Effects of medium composition on the production of plasmid DNA vector potentially for human gene therapy. J Zhejiang Univ Sci B(5):396 – 400.

Table 1 Bacterial strains used in this study.

Strain	Genotype	Reference
MG1655*ΔendAΔrecA*	*F⁻ λ ⁻ilvG rfb-50 rph1ΔendAΔrecA*	Prather Lab
GALG10	MG1655*ΔendAΔrecAΔpykF*	This study
GALG11	MG1655*ΔendAΔrecAΔpykA*	This study
GALG1011	MG1655*ΔendAΔrecAΔpykF ΔpykA*	This study
DH5α	F- ö80*lac*ZÄM15 Ä(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 ë⁻	Invitrogen
DH5α*ΔpykA*	DH5α *ΔpykA*	This study
DH5α*ΔpykF*	DH5α *ΔpykF*	This study
DH5α*ΔpykFΔpykA*	DH5α *ΔpykFΔpykA*	This study
JM101	F´ *traD36 proA⁺B⁺ lacI^q* Δ*(lacZ)M15/* Δ*(lac-proAB) glnV thi*	ATCC33876
JM101*ΔpykFΔpykA*	JM101*ΔpykFΔpykA*	This study
GALG20	MG1655*ΔendAΔrecAΔpgi*	This study

Table 2 Biomass, acetate and growth rate in glucose versus glycerol for parental strains versus Pyk and Pgi mutant strains.

Strain	**Carbon Source**	**Biomass** (g/L)	**Acetate** (g/L)	**Growth rate** (h^{-1})
MG1655*ΔendAΔrecA*	Glycerol 20 g/L	7.7 ± 0.8	3.86 ± 0.01	0.60 ± 0.01
GALG1011		8.6 ± 0.9	0.08 ± 0.01	0.49 ± 0.01
GALG20		8.5 ± 0.5	0.61 ± 0.04*	0.60 ± 0.05
DH5α		11.3 ± 0.9	0.05 ± 0.01	0.55 ± 0.01
DH5α*ΔpykFΔpykA*		9.0 ± 0.6	0.04 ± 0.01	0.31 ± 0.01
MG1655*ΔendAΔrecA*	Glucose 20 g/L	3.5 ± 0.3	5.74 ± 0.04	0.65 ± 0.01
GALG1011		4.3 ± 0.2	4.81 ± 0.03	0.53 ± 0.01
GALG20		10.9 ± 0.2	1.55 ± 0.02*	0.77 ±0.05
DH5α		3.8 ± 0.2	4.72 ± 0.04	0.49 ± 0.08
DH5α*ΔpykFΔpykA*		3.0 ± 0.4	4.21 ± 0.02	0.28 ±0.01
MG1655*ΔendAΔrecA*	Glucose 5+10 g/L	11.8 ± 0.3	1.98 ± 0.05*	0.55 ± 0.08
GALG1011		12.2 ± 0.2	0.64 ± 0.05*	0.40 ± 0.10
GALG10		9.8 ± 0.4	0.86 ± 0.05*	0.30 ± 0.1
GALG11		12.0 ± 0.4	1.23 ± 0.05*	0.48 ± 0.05
GALG20		9.6 ±0.2	0.68 ± 0.05*	0.67 ± 0.05

Strains were grown in shake flasks at 37ºC. Average value ± standard error of mean (SEM) is shown.

*Values corresponding to hour 8, since at hour 24 no acetate was detectable.

Table 3 Effect of glucose and glycerol in different concentrations on plasmid DNA production for different strains.

Strain	Carbon Source (g/L)		pDNA yield (mg/gDCW)	pDNA yield (mg/L)
MG1655*ΔendAΔrecA*	Glucose	20	0.8 ± 0.1	1.5 ± 0.1
		10+10	0.8 ± 0.1	1.9 ± 1.2
		5+10	3.6 ± 0.7	27.4 ± 1.9
	Glycerol	20	11.5 ± 0.8	79.3 ± 1.4
		5+10	11.2 ± 0.5	75.1 ± 3.9
GALG1011	Glucose	20	1.6 ± 0.3	3.2 ± 0.4
		10+10	1.2 ± 0.5	4.3 ± 1.7
		5+10	6.6 ± 0.4	42.1 ± 3.5
	Glycerol	20	7.6 ± 1.1	50.5 ± 1.3
		5+10	2.7 ± 0.2	20.9 ± 0.2
GALG10	Glucose	20	0.99 ± 0.1	3.45 ±0.3
		5+10	10.5 ± 1.3	81.5 ± 5.7
	Glycerol	5+10	8.6 ± 0.6	69.2 ± 3.9
GALG11	Glucose	20	0.65 ± 0.1	2.4 ± 0.3
		5+10	13.1 ± 0.2	94.1 ± 2.7
	Glycerol	5+10	10.3 ± 0.6	79.4 ± 7.0
GALG20	Glucose	20	19.1 ± 1.5	140.8 ± 0.8
		5+10	11.6 ± 1.1	88.9 ± 0.7
	Glycerol	5+10	10.1 ± 0.1	65.5 ± 1.4
DH5α	Glucose	20	0.8 ± 0.1	1.3 ± 0.1
		5+10	1.8 ± 0.7	9.6 ± 0.5
	Glycerol	20	4.4 ± 0.3	34.7 ± 0.6
DH5α*ΔpykFΔpykA*	Glucose	20	0.4 ± 0.1	2.6 ± 0.2
		5+10	0.3 ± 0.3	0.9 ± 0.2
	Glycerol	20	1.5 ± 0.3	5.9 ± 1.1
DH5α*ΔpykF*	Glucose	5+10	5.9 ± 0.1	36.9 ± 0.4
DH5α*ΔpykA*	Glucose	5+10	0.9 ± 0.1	1.5 ± 0.1
JM101	Glucose	20	1.3 ± 0.2	4.5 ± 0.8
		5	2.5 ± 0.4	35.5 ± 6.8
	Glycerol	20	8.4 ± 1.4	54.3 ± 2.8
JM101 *ΔpykFΔpykA*	Glucose	20	2.6 ± 0.3	12.5 ± 2.8
		5	5.3 ± 1.7	28.5 ± 4.3
	Glycerol	20	2.2 ± 1.3	12.5 ± 0.5

Strains were grown in shake flasks at 37ºC. Average value ± standard error of mean (SEM) is shown.

Table 4 Top three high-yield pDNA production strains identified in this study versus a common strain for pDNA production, DH5α.

Strain	Carbon source	pDNA yield	
		volumetric (mg/L)	Specific (mg/g DCW)
GALG20	Glucose	140.8 ± 0.8	19.1 ± 1.5
GALG11	Glucose	94.1 ± 2.7	13.1 ± 0.2
MG1655*ΔendAΔrecA*	Glycerol	79.3 ± 1.4	11.5 ± 0.8
DH5α	Glycerol	34.7 ± 0.6	4.4 ± 0.3

Strains were grown in shake flasks at 37ºC. Average value ± standard error of mean (SEM) is shown.

Figure legends

Fig. 1 Gene knockout strategies to improve plasmid DNA production in *E. coli.* The knockout of the phosphoglucose isomerase (*pgi*) gene redirects glycolytic flux, increasing fluxes in the pentose phosphate pathway and enhancing nucleotide synthesis and NADPH production. Glycolysis is down-regulated, but proceeds via the formation of fructose 6-phosphate and glyceraldehyde 3-phosphate. The knockouts of pyruvate kinase genes *pykF* and *pykA* reduce acetate formation and increase fluxes in the pentose phosphate pathway.This figure represents a simplified version of *E. coli* central metabolism. Dark arrows represent high carbon flow in the pentose phosphate pathway, light arrows represent downregulation of glycolysis, and blank arrow represent null carbon flow. Abbreviations: G6P, glucose 6-phosphate; F6P, fructose 6-phosphate; FDP, fructose 1,6-diphosphate; G3P, glyceraldehyde 3-phosphate; PEP, phosphoenolpyruvate; 6GPC, 6-phosphogluconate; RU5P, ribulose 5-phosphate; XU5P, xylulose 5-phosphate; R5P, ribose 5-phosphate; S7P, sedoheptulose 7-phosphate; E4P, erythrose 4-phosphate; PTS, phosphotransferase system.

Fig. 2 Effect of *pgi* knockout on pDNA specific yield (mg/g DCW) using MG1655*ΔendAΔrecA* grown in glycerol and glucose. Strains were grown for 24 h in shake flasks (37ºC, 250 rpm) with rich medium supplemented with 5+10 g/L of glucose (white bars), 5+10 g/L glycerol (gray bars) and 20 g/L of glucose (solid bars). Differences between MG1655*ΔendAΔrecA* and GALG20 were statistically significant ($p<0.05$) in different concentrations of glucose, but they were not statistically significant ($p>0.05$) in glycerol. The standard error of the mean (SEM) was calculated is represented with error bars.

Figure 1
Click here to download high resolution image

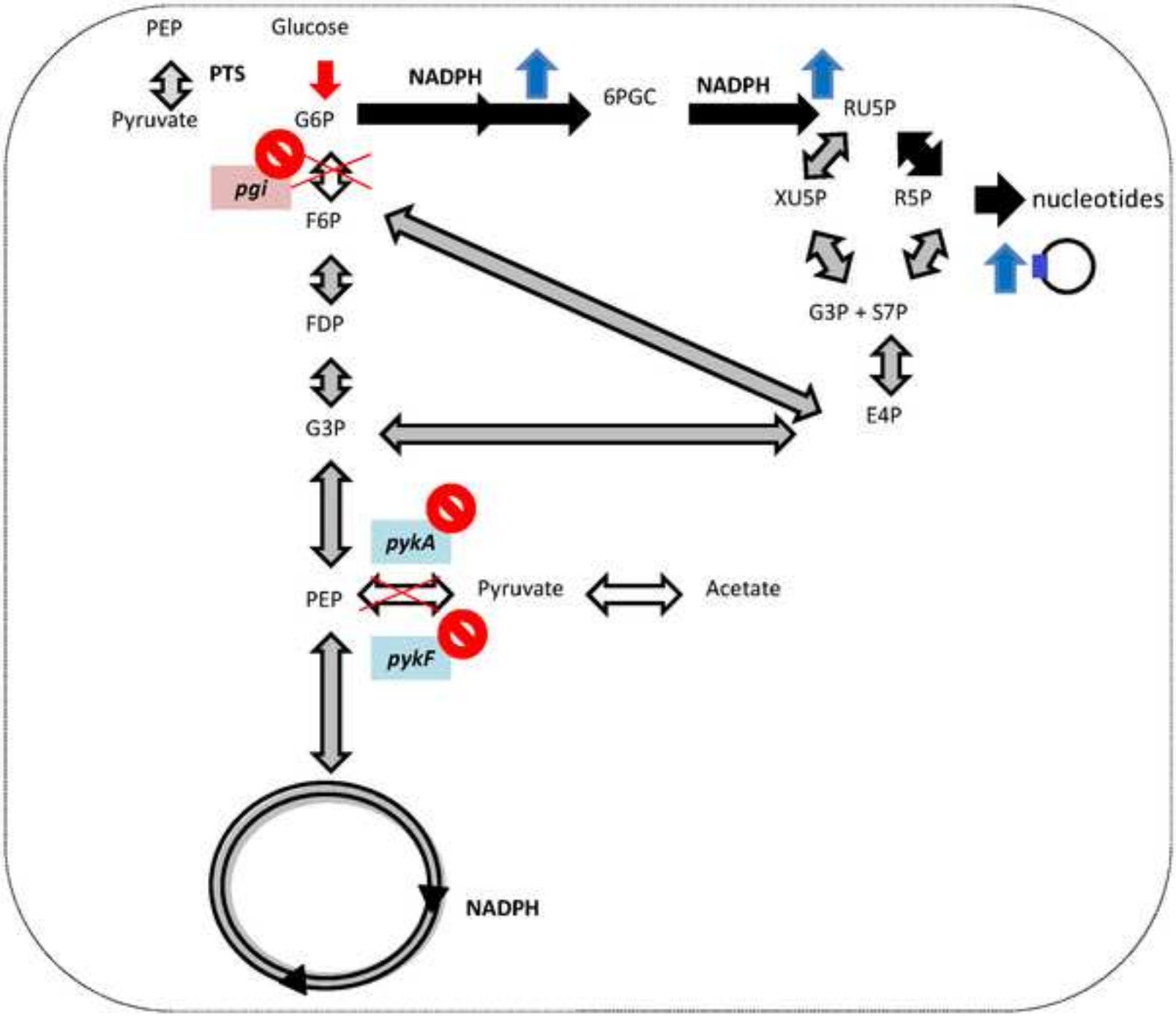

Figure 2
Click here to download high resolution image