

*Research article*

## **Model-driven *in Silico* *glpC* Gene Knockout Predicts Increased Succinate Production from Glycerol in *Escherichia Coli***

**Bashir Sajo Mienda \* and Mohd Shahir Shamsir**

Bioinformatics Research Group (BIRG), Department of Biosciences and Health Sciences, Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, 81310 Skudai Johor Bahru, Malaysia

\* **Correspondence:** Email: bsmienda@gmail.com; Tel: +60149136476;  
Fax: +607-5531279.

**Abstract:** Metabolic engineered targeting for increased succinate production in *Escherichia coli* using glycerol as a low cost carbon source has attracted global attention in recent years. Succinate production in engineered *E. coli* has progressed significantly using an experimental trial and error approach. The use of a model-guided, targeted metabolic gene knockout prediction for increased succinate production from glycerol under anaerobic conditions in *E. coli* still remains largely underexplored. In this study, we applied a model-driven, targeted *glpC/b2243 in silico* metabolic gene knockout using *E. coli* genome scale model *iJO1366* under the OptFlux software platform with the aim of predicting high succinate flux. The results indicated that the mutant model lacking the *glpC/b2243* gene will demonstrate increased succinate flux that is 30% higher than its wild-type control model. We can hypothesize that an additional NADH molecule was generated following the deletion of the gene and/or the alternatively preferred *GldA-DhaKLM* fermentative route for glycerol metabolism in *E. coli* may have been activated. Although the exact metabolic mechanism involved in increasing the succinate flux still remains obscure; the current study informs other studies that a model-driven, metabolic *glpC/b2243* gene knockout could be applicable in filling our knowledge gap using a comprehensive experimental inquiry in the future; leading to a better understanding of the underlying metabolic function of this gene in relation to succinate production in *E. coli* from glycerol.

**Keywords:** *In silico* gene knockout; glycerol, succinate; *E. coli* genome scale model; anaerobic glycerol-3-phosphate dehydrogenase (*glpC*); OptFlux

---

## 1. Introduction

Much attention has been focused on the use of glycerol as a solitary carbon source in recent years for bio-based, industrial chemical production. This is because glycerol is an abundant, inexpensive carbon source, with a high degree of reducing potential and it is currently being generated as a byproduct of biofuel industry [1]. Glycerol could be converted into a value added succinic acid by using metabolically engineered *Escherichia coli* strains. Succinic acid is a potential platform chemical that could serve as a precursor for the synthesis of other commodity and specialty chemicals with a wide spectrum of applications [2,3].

Metabolic engineering of *Escherichia coli* for increased succinate production from glycerol has progressed significantly using an experimental trial by error approach [1,4–9]. One of the notable examples was reported by Gonzalez and colleagues [1], where they introduced a heterologous pyruvate carboxylase (*pyc*) from *Lactococcus lactis* into *E. coli* for the production of succinate using glycerol as the substrate in minimal salt medium under micro-aerobic conditions. The succinate synthesis was achieved by the blockage of competing metabolic byproducts [1]. In *E. coli* central metabolism, glycerol utilization for succinate production proceeds via three (3) different routes to dihydroxyacetone phosphate (DHAP) which are: (1) the aerobic *GlpK-GlpD* respiratory route, (2) the anaerobic *GlpK-GlpABC* respiratory route (Figure 1) and (3) the *GldA-DhaKLM* fermentative route (Figure 1) [1,6]. We leveraged the knowledge base created in the aforementioned studies to investigate the underlying metabolic role of *glpC* via *in silico* gene knockout, in relation to *E. coli* succinate production and NADH regeneration under anaerobic conditions when glycerol is the only available substrate.

The advances seen in systems metabolic engineering and high throughput technology led to the accumulation of a wealth of biochemical data; where biochemical reaction networks were assembled into genome scale metabolic models (GEM) [9,10,11]. Genome scale metabolic reconstructions of *E. coli* have been developed [12,13] and used widely for predicting growth phenotypes in various environmental and/or genetic perturbations [14–17]. Model-driven metabolic engineering strategies can be performed using a variety of substrates and different computational tools such as OptFlux [18] and COBRA [19] to mention but a few. A more recent study reports a workflow methodology that uses model-driven biological discovery to understand certain underlying metabolic functions of some selected genes in *E. coli* [20].

A number of research efforts involving engineering of *E. coli* host metabolic pathways and succinic acid overproducing *Mannheimia succiniciproducens* have been conducted using genome comparison and *in silico* metabolic gene knockout with MetaFluxNet to increase succinate production in chassis *E. coli* [21]. In a different study [22], elementary mode analysis for the rational design of efficient glycerol conversion to succinate by *E. coli* was conducted using METATOOL 5.1 [23]. The approach investigated the pathway involved for optimal succinate production coupled with the effects of oxygen levels in relation to succinate production in *E. coli* [22]. Taken together, many investigators have applied experimental trial and error genetic engineering approaches to improve succinate production in *E. coli* using different strategies; ranging from addition of foreign genes [24–28] to metabolic evolution of energy conserving pathways using glucose as the solitary carbon source [29]. Other researchers [30] have constructed several strains of *E. coli*, in which genes involved in succinate competing pathways were knocked out without introduction of foreign genes. This strategy combined with several rounds of metabolic evolution, led to *E. coli* strains, designated

as KJ060 (*ΔldhA*, *ΔadhE*, *ΔackA*, *ΔfocA*, *ΔpflB* and KJ073 (*ΔldhA*, *ΔadhE*, *ΔackA*, *ΔfocA*, *ΔpflB*, *ΔmgsA* *ΔpoxB*) with improved succinate production from a glucose substrate. As promising as this approach might be, there are, as yet, no *E. coli* strains or models lacking the *glpC* gene which have been reported to have increased succinate production using glycerol as the solitary carbon source. Here, we report a model-driven *in silico* *glpC*/b2243 metabolic gene knockout strategy for increased succinate production using an accurate genome scale metabolic reconstruction of *E. coli* iJO1366 [13] with the OptFlux software platform [18]. Minimization of Metabolic Adjustment (MOMA) was used as the gene knockout simulation algorithm [31] under the OptFlux software interface. The results obtained herein demonstrate, for the first time, that the mutant model lacking the *glpC* gene is predicted to increase succinate flux, when glycerol is used as the substrate, by about 30% (see Table 1) when compared to the wild-type control model,. It was thus hypothesized that an additional molecule of NADH was generated following the *glpC* gene knockout and/or the glycerol central metabolism was channeled via the preferred *GldA-DhaKLM* fermentative route for increased generation of DHAP, PEP and subsequently, succinate (Figure 1).

## 2. Materials and Method

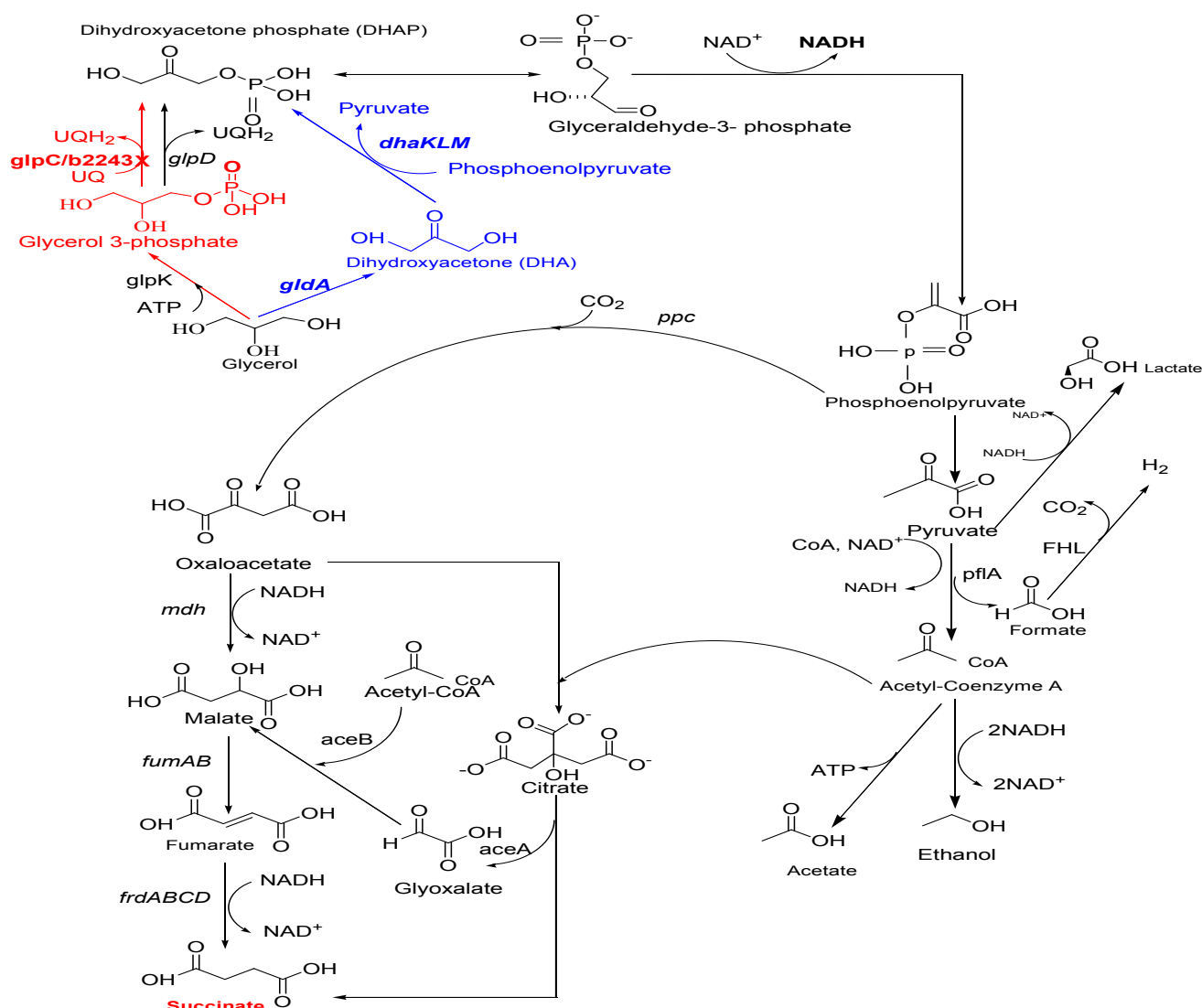
### 2.1. Model and flux balance analysis

The *E. coli* genome scale metabolic model developed by Orth and colleagues [13], designated as *E. coli* iJO1366 was utilized as a basis for metabolic gene knockout simulation, succinate yield, and strain design. This model was shown to be capable of accurate predictions of growth rates, metabolite excretion rates on a variety of substrates, and genetic conditions consistent with experimental data [12,13,15]. For all phenotype simulations, flux balance analysis was conducted. All computations were carried out using the OptFlux software platform version 3.07 (<http://www.optflux.org>), which was implemented using Java programming. Minimization of Metabolic Adjustment (MOMA) algorithm [31] was used for *glpC*/b2243 gene knockout simulation.

Glycerol was chosen as the solitary carbon source under anaerobic conditions. Glycerol uptake rate was constrained to a maximum of 18.5 mmol g DW<sup>-1</sup> h<sup>-1</sup> while its corresponding oxygen uptake rate was set to be zero, as the environmental condition is anaerobic. These values were selected based on closely established experimental observations of aerobic and anaerobic growth in *E. coli* [11,32].

**Table 1. Strain design properties of *E. coli* models designed from glycerol substrate using the OptFlux software platform**

Model Strains of <i>E. coli</i> used	Biomass (h <sup>-1</sup> )	% Biomass	Succinate (mmol g DW <sup>-1</sup> h <sup>-1</sup> )	% Succinate	Ethanol (mmol g DW <sup>-1</sup> h <sup>-1</sup> )	Acetate (mmol g DW <sup>-1</sup> h <sup>-1</sup> )	Formate (mmol g DW <sup>-1</sup> h <sup>-1</sup> )
Orth Model (WT)	0.3959758	100	0.12733	100	6.86885	25.59892	33.99566
GlpC/b2243	0.29609198	74.77	0.16604	130.4	6.86208	25.57664	33.87925



**Figure 1.** Pathways involved in the fermentative utilization of glycerol in *E. coli* under anaerobic conditions partially adopted from ref. [1,2,22]. Generation of Phosphoenolpyruvate (PEP) and pyruvate, which was subsequently carboxylated to intermediates leading to succinate production, is shown. The preferred alternative fermentative GldA-dhaKLM route for glycerol dissimilation in *E. coli* [1] is shown in blue. The shutdown alternative route that functions under anaerobic conditions is shown in red. The knockout gene is called anaerobic glycerol-3-phosphate dehydrogenase (*glpC/b2243*). Abbreviations: *dhaKLM*, dihydroxyacetone kinase; *gldA*, anaerobic glycerol dehydrogenase II; *glpK*, aerobic glycerol kinase; *glpD*, aerobic glycerol-3-phosphate dehydrogenase; *ppc*, phosphoenolpyruvate carboxylase; FHL, formate hydrogen lyase complex; *aceA*, isocitrate lyase; *aceB*, malate synthetase; UQ, Ubiquinone; UQH<sub>2</sub>, denotes where Flavin containing enzymes like *glp* regulon interacts with a quinone to provide or acquire reducing energy.

### 3. Results and Discussions

In this study, an *E. coli* genome scale model was subjected to *glpC*/b2243 metabolic gene knockout for increased succinate production from glycerol under anaerobic conditions. We performed *in silico* deletion of the *glpC*/b2243 gene under anaerobic conditions in order to use model-driven, targeted biological inquiry on the underlying metabolic function of the gene in relation to succinate production in *E. coli* from a glycerol substrate. The results indicated that the mutant model lacking the *glpC* gene was predicted to have increased succinate production that was 30% higher than its wild-type parent model (Table 1). In addition, the growth rate of the mutant model was about 75% of the wild type model (see Table 1). The problem of the impaired growth rate of *E. coli* on glycerol could be addressed by mutational inactivation of *rpoC* and *glpK* as reported recently by Palsson and colleagues [33]. Their method involves subjecting the *E. coli* strain to adaptive laboratory evolution on glycerol, and subsequently identifying point mutations in *rpoC* and *glpK* genes using comparative whole genome sequencing; which ultimately allows the mutant strain to gain a fast growth advantage over its wild-type counterpart on glycerol [33].

In principle, glycerol dissimilation in *E. coli* central metabolic pathways uses aerobic, anaerobic and fermentative routes to produce dihydroxyacetone phosphate (DHAP). Under anaerobic condition the route to DHAP was established to be via the *GlpK- GlpABC* respiratory route or *GldA-DhaKLM* fermentative routes (Figure 1) [1]. It was previously established that for efficient utilization of glycerol for succinate production under anaerobic conditions, *E. coli* preferred the use of the *GldA-DhaKLM* fermentative route; because higher energy NADH is generated when glycerol is consumed through *GldA* rather than through a reduced Flavin *glpD* or *GlpABC* route (Figure 1). We took advantage of this phenomenon and performed *in silico* gene knockout of the *glpC* gene to increase the availability of the NADH required for succinate production. *E. coli* regulates glycerol metabolism based on environmental and genetic conditions. Glycerol-3-phosphate dehydrogenase (*glpD*) is only expressed under aerobic conditions (Figure 1), while, the glycerol-3-phosphate dehydrogenase operon (*glpABC*) is usually expressed under anaerobic conditions. The operon consists of three open reading frames, *glpABC*, encoding polypeptides with varying molecular weights. The *glpC* subunit possesses two cysteine clusters that were characterized by an iron-sulfur binding domain, while the other two polypeptides (*gldAB*) formed a catalytic dimer [34]. The *glpC* subunit was reported to have a membrane anchor binding function for the *glpAB* dimer [34].

The metabolic routes of glycerol dissimilation via *glpK*, glycerol-3-phosphate and *glpABC* to form DHAP (Figure 1) occur under anaerobic condition; but the expression of *glpD* (Figure 1) occurs only under aerobic conditions. This route is disadvantaged for the formation of DHAP, because the simulation conducted in this study was performed under anaerobic conditions, thereby limiting the formation of DHAP, which is required for subsequent formation of phosphoenolpyruvate (PEP). We reasoned that, by knocking out the *glpC* gene, the metabolic route for the production of DHAP via *glpK* will effectively shutdown, favoring glycerol metabolism through the *GldA-DhaKLM* fermentative route to form DHAP under anaerobic conditions. This could be the reason why the deletion of the *glpC* gene in the *E. coli* genome scale model predicted 30% higher succinate production to than the wild-type model.

The *DhaKLM* route has been known to utilize phosphoenolpyruvate (PEP) as a cofactor, channeling the metabolic nodes available for the fermentative glycerol metabolism to succinate (Figure 1) [1]. As a green technology, succinate production from glycerol involved a CO<sub>2</sub> fixing step

onto a 3-carbon intermediate which is subsequently converted to succinate by the TCA reductive branch [35] (Figure 1). The *E. coli* PEP carboxylase (*ppc*) is the primary native carboxylation enzyme for succinate generation, although Gonzalez and colleagues [1] argued that it is not an ideal stable route as the PEP level will be reduced drastically when the fermentative pathway involving *GldA-DhaKLM* is utilized [1]. In a similar study reported by the same group [1], heterologous pyruvate carboxylase (*pyc*) was introduced into *E. coli* from *L. lactis* in order to retain the effective activity of the *GldA-DhaKLM* route, thereby creating an effective node for the subsequent pyruvate conversion to succinate [1]. *E. coli* lacks native pyruvate carboxylase, as such, the introduction of foreign *pyc* gene from *L. lactis* was necessary to achieve higher carbon flux to succinate from pyruvate. On the contrary, the *in silico* metabolic *glpC*/b2243 gene knockout was performed and evaluated without introduction of any foreign gene; but the mutant model achieved a succinate flux increase of about 30% when compared to the wild-type parent model (Table 1). On the basis of this finding, we can now hypothesize that the deletion of the *glpC* gene using glycerol as substrate under anaerobic condition could have generated an additional molecule of NADH and/or increased the generation of DHAP and PEP required for subsequent carboxylation to generate 30% higher succinate flux in the mutant model. The underlying metabolic function of *glpC* under anaerobic condition with a glycerol substrate in relation to the generation of reducing equivalent and succinate production remains obscure.

Although, the exact mechanism involved in increasing the succinate flux by 30% is unknown, it was established that the key to increasing succinate production in *E. coli* was increasing the carboxylation of PEP and pyruvate to a four (4) carbon dicarboxylic acid precursor of succinate [29]. However, glycerol dissimilation in *E. coli* central metabolic pathways under anaerobic conditions proceeds via DHAP to PEP and pyruvate [1,5,29]. The preferred *GldA-dhaKLM* fermentative route may have been activated following the deletion of the anaerobic *glpC* gene. The 30% succinate flux increase after the *glpC* gene knockout could be attributed to a possible increase in the pool of PEP and pyruvate generation. The former is a precursor to succinate production via OAA by CO<sub>2</sub> fixation with the help of *E. coli* native *ppc* (Figure 1), while the latter is required for ATP generation, necessary for cell growth and maintenance. In addition, formate is formed from pyruvate, while additional CO<sub>2</sub> and H<sub>2</sub> are generated through the formate hydrogen lyase (FHL) complex from formate (Figure 1) [5]. Succinate production from glycerol in *E. coli* involved fixing CO<sub>2</sub> onto a 3-carbon intermediate, which is converted to succinate through the TCA cycle reductive branch [1,35]. Therefore, the CO<sub>2</sub> generated from formate might be used to generate more succinate production via PEP to OAA by using *ppc*.

#### 4. Conclusion

The generation of glycerol as a waste from the bioethanol and biodiesel industries makes it an abundant and low cost carbon source for bio-based succinate production in engineered *E. coli* chassis hosts. The potentials of using glycerol for succinate production in *E. coli* under anaerobic conditions coupled with a model-driven, targeted *glpC* *in silico* metabolic gene knockout was evaluated using the *E. coli* genome scale metabolic model *iJO1366*, with the OptFlux software platform in this work. The flux to succinate production was increased following the *glpC* gene deletion under anaerobic conditions with the glycerol substrate. Although, the exact metabolic mechanism of the succinate increase still remains unknown, the underlying metabolic function of this deletion could be unfolded

in the future using model-guided, experimental biological inquiry. The result of this deletion is considered significant for further investigation on the succinate metabolism from glycerol by *E. coli*.

### Conflict of Interest

All authors declare to have no conflict of interests.

### References

1. Blankschien MD, Clomburg JM, Gonzalez R (2010) Metabolic engineering of *Escherichia coli* for the production of succinate from glycerol. *Metab Eng* 12: 409–419.
2. Thakker C, Martinez I, San KY, et al. (2012) Succinate production in *Escherichia coli*. *Biotechnol J* 7: 213–224.
3. McKinlay JB, Vieille C, Zeikus JG (2007) Prospects for a bio-based succinate industry. *Appl Microbiol Biotechnol* 76: 727–740.
4. Clomburg JM, Gonzalez R (2013) Anaerobic fermentation of glycerol: a platform for renewable fuels and chemicals. *Trends Biotechnol* 31: 20–28.
5. Dharmadi Y, Murarka A, Gonzalez R (2006) Anaerobic fermentation of glycerol by *Escherichia coli*: a new platform for metabolic engineering. *Biotechnol Bioeng* 94: 821–829.
6. Durnin G, Clomburg J, Yeates Z, et al. (2009) Understanding and harnessing the microaerobic metabolism of glycerol in *Escherichia coli*. *Biotechnol Bioeng* 103: 148–161.
7. Mattam AJ, Clomburg JM, Gonzalez R, et al. (2013) Fermentation of glycerol and production of valuable chemical and biofuel molecules. *Biotechnol Lett* 35: 831–842.
8. Shams Yazdani S, Gonzalez R (2008) Engineering *Escherichia coli* for the efficient conversion of glycerol to ethanol and co-products. *Metab Eng* 10: 340–351.
9. Yazdani SS, Gonzalez R (2007) Anaerobic fermentation of glycerol: a path to economic viability for the biofuels industry. *Curr Opin Biotechnol* 18: 213–219.
10. Monk J, Palsson BO (2014) Genetics. Predicting microbial growth. *Science* 344: 1448–1449.
11. Varma A, Bernhard BW, Palsson OB (1993) Stoichiometric Interpretation of *Escherichia coli* Glucose Catabolism under Various Oxygenation Rates. *Appl Environ Microbiol* 59: 2465–2473.
12. Feist AM, Henry CS, Reed JL, et al. (2007) A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol* 3: 121.
13. Orth JD, Conrad TM, Na J, et al. (2011) A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism--2011. *Mol Syst Biol* 7: 535.
14. Oberhardt MA, Palsson BO, Papin JA (2009) Applications of genome-scale metabolic reconstructions. *Mol Syst Biol* 5: 320.
15. Feist AM, Zielinski DC, Orth JD, et al. (2010) Model-driven evaluation of the production potential for growth-coupled products of *Escherichia coli*. *Metab Eng* 12: 173–186.
16. Hara KY, Shimodate N, Ito M, et al. (2009) Systematic genome-wide scanning for genes involved in ATP generation in *Escherichia coli*. *Metab Eng* 11: 1–7.
17. McCloskey D, Palsson BO, Feist AM (2013) Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*. *Mol Syst Biol* 9: 661.

18. Rocha I, Maia P, Evangelista P, et al. (2010) OptFlux: an open-source software platform for in silico metabolic engineering. *BMC Syst Biol* 4: 45.
19. Schellenberger J, Que R, Fleming RM, et al. (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat Protoc* 6: 1290–1307.
20. Guzman GI, Utrilla J, Nurk S, et al. (2015) Model-driven discovery of underground metabolic functions in *Escherichia coli*. *Proc Natl Acad Sci U S A* 112: 929–934.
21. Lee SJ, Lee DY, Kim TY, et al. (2005) Metabolic engineering of *Escherichia coli* for enhanced production of succinic acid, based on genome comparison and in silico gene knockout simulation. *Appl Environ Microbiol* 71: 7880–7887.
22. Chen Z, Liu H, Zhang J, et al. (2010) Elementary mode analysis for the rational design of efficient succinate conversion from glycerol by *Escherichia coli*. *J Biomed Biotechnol* 2010: 518743.
23. von Kamp A, Schuster S (2006) Metatool 5.0: fast and flexible elementary modes analysis. *Bioinformatics* 22: 1930–1931.
24. Kim P, Laivenieks M, Vieille C, et al. (2004) Effect of Overexpression of *Actinobacillus succinogenes* Phosphoenolpyruvate Carboxykinase on Succinate Production in *Escherichia coli*. *Appl Environ Microb* 70: 1238–1241.
25. Gokarn RR, Evans JD, Walker JR, et al. (2001) The physiological effects and metabolic alterations caused by the expression of *Rhizobium etli* pyruvate carboxylase in *Escherichia coli*. *Appl Microbiol Biot* 56: 188–195.
26. Sanchez AM, Bennett GN, San KY (2005) Efficient Succinic Acid Production from Glucose through Overexpression of Pyruvate Carboxylase in an *Escherichia coli* Alcohol Dehydrogenase and Lactate Dehydrogenase Mutant. *Biotechnol Prog* 21: 358–365.
27. Wang Q, Wu C, Chen T, et al. (2006) Expression of galactose permease and pyruvate carboxylase in *Escherichia coli* ptsG mutant increases the growth rate and succinate yield under anaerobic conditions. *Biotechnol Lett* 28: 89–93.
28. Vemuri GN, Eiteman MA, Altman E (2002) Effects of Growth Mode and Pyruvate Carboxylase on Succinic Acid Production by Metabolically Engineered Strains of *Escherichia coli*. *Appl Environ Microbiol* 68: 1715–1727.
29. Zhang X, Jantama K, Moore JC, et al. (2009) Metabolic evolution of energy-conserving pathways for succinate production in *Escherichia coli*. *Proc Natl Acad Sci U S A* 106: 20180–20185.
30. Jantama K, Haupt MJ, Svoronos SA, et al. (2008) Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of *Escherichia coli* C that produce succinate and malate. *Biotechnol Bioeng* 99: 1140–1153.
31. Segre D, Vitkup D, Church GM (2002) Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci U S A* 99: 15112–15117.
32. Varma A, Palsson BO (1994) Stoichiometric Flux Balance Models Quantitatively Predict Growth and Metabolic By-Product Secretion in Wild-Type *Escherichia coli* W3110. *Appl Environ Microbiol* 60: 3724–3731.
33. Cheng K-K, Lee B-S, Masuda T, et al. (2014) Global metabolic network reorganization by adaptive mutations allows fast growth of *Escherichia coli* on glycerol. *Nat Commun* 5.



- 
34. Cole ST, Eiglmeier K, Ahmed S, et al. (1988) Nucleotide Sequence and Gene-Polypeptide Relationships of the *glpABC* Operon Encoding the Anaerobic sn-Glycerol-3-Phosphate Dehydrogenase of *Escherichia coli* K-12. *J Bacteriol* 170: 2448–2456.
  35. Sauer U, Eikmanns BJ (2005) The PEP–pyruvate–oxaloacetate node as the switch point for carbon flux distribution in bacteria. *FEMS Microbiol Rev* 29: 765–794.

© 2015, Bashir Sajo Mienda, et al., licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)