FDH Knockout and TsFDH Transformation Led to Enhanced Growth Rate of *Escherichia Coli*

Roya Razavipour
Islamic Azad University Science and Research Branch
Saman Hosseini Ashtiani (saman.hosseini-ashtiani@dbb.su.se)
Stockholm University: Stockholms Universitet https://orcid.org/0000-0003-2381-3410
Abbas Akhavan Sepahy
Islamic Azad University Tehran North Branch
Mohammad Hossein Modarressi
Tehran University of Medical Sciences
Bijan Bambai
NIGEB: National Institute for Genetic Engineering and Biotechnology

Research

Keywords: Formate Dehydrogenase, Metabolic CO2 Leak, Glycerol minimal medium, Spearman's rank correlation coefficient, RNA-seq, linear regression, Principal Component Analysis (PCA)

Posted Date: December 14th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1146279/v1

Abstract

Background:

Increased Atmospheric CO_2 to over 400 ppm has prompted global climate irregularities. Reducing the released CO_2 from biotechnological processes could remediate these phenomena. In this study, we sought to find a solution to reduce the amount of CO_2 in the process of growth and reproduction by preventing the conversion of formic acid into CO_2.

Results:

The (bio)chemical conversion of formic acid to CO_2 is a key reaction. Therefore, we compared the growth of BL21, being a subfamily of K12, alongside two strains in which two different genes related to the formate metabolism were deleted, in complex and simple media. Experimental results were entirely consistent with metabolic predictions. Subsequently, the knockout bacteria grew more efficiently than BL21. Interestingly, TsFDH, a formate dehydrogenase with the tendency of converting CO_2 to formate, increased the growth of all strains compared with cells without the TsFDH. Most mutants grew in a simple medium containing glycerol, which showed that glycerol is the preferred carbon source compared to glucose for the growth of *E. coli*.

Conclusion:

These results explain the reasons for the inconsistency of predictions in previous metabolic models that declared glycerol as a suitable carbon source for the growth of *E. coli* but failed to achieve it in practice. To conduct a more mechanistic evaluation of our observations, RNA sequencing data analysis was conducted on an *E. coli* RNA-seq dataset. The gene expression correlation outcome revealed the increased expression levels of several genes related to protein biosynthesis and glycerol degradation as a possible explanation of our observations.

Background

CO_2 is easily formed by the oxidation of organic molecules during respiration in living organisms or combustion in regular mechanical engines. This molecule is thermodynamically stable with a low chemical activity. Today, the atmospheric CO_2 concentration is approaching alarming levels (from 300 ppm to 417 *ppm* in about 50 years), this in turn has led to elevated frequencies of extreme climate conditions like drought, flooding, wild fire and tropical storms in different regions of the world (1). The development of innovative methods for reducing the released CO_2 into the atmosphere and/or the assimilation of CO_2 into organic matter is in demand more than ever (2). Respiration is a more economical process for extracting chemical energy from organic matter in comparison with anaerobic fermentation. There is still an undesired side effect of respiration, that is, the release of CO_2. Engineering bacterial strains with the aim of reducing carbon dioxide release into the atmosphere or even fixing the

atmospheric CO_2 has environmental and economic advantages in biotechnological processes. There have been a number of efforts to reduce the carbon dioxide release during biomass production by metabolic engineering (3). The central pathways and cycles of metabolism are the first targets for manipulating enzymes responsible for critical biochemical reactions, or regulatory proteins controlling the expression of some enzymes to reduce CO_2 release (4). One of the interesting candidates for reducing the amount of released CO_2 is formate dehydrogenase (FDH). Theoretically, FDHs are enzymes capable of reversible conversion of CO_2 to formate, which is the simplest organic acid (5). However, the major drawback of the biotechnological application of FDHs is the fact that the majority of these enzymes favor the oxidation of formate to produce CO_2 under physiological conditions (6). There are three known FDHs in *E. coli* genome, namely, FdhH, FdhN, and FdhO. The newly identified pressure induced FDH (FHL) is another identified FDH in E. coli. In search of an “ideal” FDH we chose to express FDH from *Thiobacillus sp*. KNK65MA (TsFDH) as an enzyme with favorable kinetics for formate formation (7). The crystal structure of active enzyme (PDB: 3WR5) shows a homo-tetramer of 406 amino acid-long polypeptide. There are 5 extra residues at N-terminal of the recombinant protein, compared with the sequence in UniProt (8) (accession code: Q76EB7). We sought to observe the growth of *E. coli* fdhD and fdhF knockouts as well as BL21 cells during heterologously expressed acidophilic TsFDH. Our observations demonstrate a clear growth rate advantage in cells expressing the recombinant enzymes in BL21 compared with BL21s without TsFDH. Both FDH knockout strains, namely JW3866 and JW4040, transformed with TsFDH plasmid demonstrated meaningful growth advantage in LB as well as M9 + glycerol media. In order to perform an *in silico* evaluation of our observations, we opted for examining the transcriptomic correlations between the target (knockout) genes and the rest of the genes in *E. coli*. The subsequent correlation analysis based on the RNA-seq data revealed possible transcriptomic level evidence behind the observed increased growth of the cells expressing the recombinant TsFDH.

Results

Experimental results

Growth kinetics of bacterial cells with or without plasmid on LB and M9+Glycerol is presented using Microsoft Excel (2019) in Fig. 1 and 2, respectively. Samples were taken within 24 hours at different time intervals. The preferential growth dynamics of FDH knockout strain with or without pET+TsFDH over standard BL21 strain were recognizable more clearly after eight hours post inoculation when LB was used as the growth medium. In samples grown on M9+glycerol, we observed a pronounced shift in growth divergence towards earlier time intervals compared with LB medium. This demonstrates the critical role of FDH in growth efficiency of *E. coli* on glycerol as the carbon source. SDS-PAGE analysis of the strains confirms the expression of TsFDH under the experimental conditions (Additional file 1: Fig. S1).

In silico analysis results

correlation analysis:

All correlations with each of the knockout genes were calculated (Additional file 2 and 3: Table S1 and S2) and the top ones with p-values and FDRs less than 0.01 were chosen (Additional file 4 and 5: Table S3 and S4). Table 1a lists the top 20 anti-correlated genes with fdhF and Table 1b lists those anti-correlated with fdhD.

Table 1
a Top 20 genes anti-correlated with fdhF

Gene ID	Gene name	R	P-value	FDR
ECB_t00058	leuU	-0.4571561	3.21E-09	4.33E-09
ECB_t00033	asnT	-0.4556892	3.65E-09	4.92E-09
ECB_t00036	asnV	-0.4546738	4.00E-09	5.37E-09
ECB_t00035	asnU	-0.4542874	4.14E-09	5.55E-09
ECB_t00046	argQ	-0.4519145	5.09E-09	6.81E-09
ECB_t00034	asnW	-0.4516529	5.21E-09	6.96E-09
ECB_t00049	argV	-0.4504972	5.76E-09	7.68E-09
ECB_t00047	argZ	-0.4490646	6.52E-09	8.69E-09
ECB_t00043	valY	-0.445059	9.20E-09	1.22E-08
ECB_t00048	argY	-0.44344	1.06E-08	1.40E-08
ECB_t00013	leuW	-0.443264	1.07E-08	1.42E-08
ECB_01180	ychH	-0.4423339	1.16E-08	1.53E-08
ECB_t00016	valT	-0.4382053	1.64E-08	2.15E-08
ECB_t00042	valX	-0.4360204	1.97E-08	2.57E-08
ECB_t00018	valZ	-0.4343724	2.26E-08	2.95E-08
ECB_t00041	valU	-0.4335911	2.41E-08	3.14E-08
ECB_t00086	leuQ	-0.4305087	3.10E-08	4.02E-08
ECB_t00068	hisR	-0.4252534	4.73E-08	6.13E-08
ECB_t00084	leuV	-0.4227489	5.78E-08	7.46E-08
ECB_t00076	glyT	-0.418321	8.19E-08	1.05E-07

Table 1
b Top 20 genes anti-correlated with fdhD

Gene ID	Gene name	R	P-value	FDR
ECB_t00047	argZ	-0.50717	2.60E-11	3.56E-11
ECB_t00046	argQ	-0.5056082	3.06E-11	4.18E-11
ECB_t00049	argV	-0.5041045	3.57E-11	4.88E-11
ECB_t00048	argY	-0.5031612	3.94E-11	5.38E-11
ECB_t00058	leuU	-0.4951881	8.85E-11	1.20E-10
ECB_t00076	glyT	-0.4899046	1.50E-10	2.02E-10
ECB_t00055	pheV	-0.4894842	1.56E-10	2.10E-10
ECB_t00079	pheU	-0.4879634	1.81E-10	2.44E-10
ECB_t00029	leuZ	-0.4832916	2.85E-10	3.83E-10
ECB_t00056	ileX	-0.452855	4.69E-09	6.13E-09
ECB_t00083	leuX	-0.4528277	4.70E-09	6.14E-09
ECB_t00051	metZ	-0.4467	7.99E-09	1.04E-08
ECB_t00067	argX	-0.4460439	8.46E-09	1.10E-08
ECB_t00077	thrT	-0.4400051	1.41E-08	1.82E-08
ECB_t00053	metV	-0.4324591	2.64E-08	3.39E-08
ECB_t00052	metW	-0.4299643	3.24E-08	4.15E-08
ECB_t00034	asnW	-0.4247936	4.91E-08	6.26E-08
ECB_t00035	asnU	-0.4168034	9.22E-08	1.17E-07
ECB_t00036	asnV	-0.4163967	9.52E-08	1.21E-07
ECB_t00033	asnT	-0.4157713	9.99E-08	1.27E-07

PCA analysis:

According to the PCA results (Additional file 1: Fig. S2), it is postulated that the top anti-correlated genes for both knockouts are closely associated with one another. As a comparison, the PCA plot was also generated using all the genes, indicating that the other genes show more expression divergence. Moreover, the second PCA (Additional file 1: Fig. S2b) is a confirmation that the dataset, being based on different growth media, reflects a wide range of expression levels for different genes in each data point, which is crucial for demonstrating the correlations between the fluctuating gene expression levels.

Linear regression analysis

The knockout genes exhibit a negative slope of the fitted linear regression model against the top anti-correlated genes, while the slope of the fitted model is positive among all the top anti-correlated genes themselves (Fig. 3).

[IMAGE-C:\Workspace\ACDC\ImageHandler\f3

Metabolic pathways

Using BioCyc database of microbial genomes and metabolic pathways, all the significantly anti-correlated genes (with p-values and FDRs less than 0.01 and "Spearman's ρ" < -0.2) were shown to be involved in tRNA-charging pathway, t-RNA processing pathway (PWY0-1479), glycerol and glycerophosphodiester degradation (PWY0-381), glycerol degradation I (PWY-4261), glycerophosphodiester degradation (PWY-6952) and tetrapyrrole biosynthesis I (PWY-5188).

Discussion

Increasing the growth efficacy of industrially important microorganisms is a novel goal in biotechnological applications. One of the strategies to boost the growth rate is to reduce the organic carbon leak, i.e., the release of CO_2 as one of the main end products in respiration process. Different *E. coli*, strains are the workhorse for the production of some well-known biopharmaceuticals, like G-CSF, Romiplostim and Asparaginase. Therefore, *E. coli* is a suitable model microorganism and developing a strain of *E. coli* with higher growth rate is in demand. Previously, other researchers have approached this challenge by defining fermentation conditions and controlling aeration rates (9) or with genetically overexpressing ArcA transcription factor (10). Here, we introduce a new approach by targeting one of the main enzymes responsible for converting organic formate into inorganic wasteful CO_2, i.e., formate dehydrogenase. There are three known FDHs in *E. coli*, namely, respiratory FDH, anaerobically expressed FDH and newly identified pressure induced FDH (FHL). FDHF is the cytosolic form, while FDHN and FDHO are membrane bound, with FDHN responsible for nitrogen cycle and FDHO active in sulfur metabolism (11). All these enzymes prefer the oxidation of formate into CO_2 under physiological conditions. Scanning BRENDA for FDHs with tendency towards the production of formate from CO_2 revealed that there are few candidate FDHs with formate production (CO_2 reduction) preference. Among these few candidates TsFDH has one of the highest k_{cat} and acceptable oxygen tolerance in microbial world. Increasing the efficacy of the bacterial growth, especially strains used in industry, is one of the major goals of biotechnology. A comprehensive study of *E. coli's* metabolism using the Regulon DB.ccg.unam.mx software showed that the format dehydrogenase enzymes were successfully expressed in recombinant form in *E. coli* (12). The increase in the growth of recombinant *E. coli* strains harboring TsFDH clearly demonstrate the feasibility of this approach for developing bacterial strains for biotechnological applications with higher biomass to carbon source ratio caused by the prevention of carbon dioxide leakage. This solution has so far been remained out of sight and to the best of our knowledge not tried

yet. In order to prove our hypothesis, we expressed the recombinant fdhD gene from *Thiobacillus* in *Escherichia coli* strain BL21. In order to better compare the role of TsFDH in growth efficacy, we also transformed two FDH knockout strains from K12, namely JW3688 and JW4040 with pET+TsFDH. Different growth rates were observed between the original knockout strains and the recombinant strains containing the plasmids in ΔFDHF (W4040) and ΔFDHD (W3866). Conversion of CO_2 to fuels and value-added chemicals is part of an aspiration to solve the energy and environmental issues. Since the Kelvin cycle is the mainstream of CO_2 fixation pathway in plants, algae and cyanobacteria, most engineering efforts are directed towards the Kelvin cycle for converting CO_2 into valuable materials. Heterotrophic microorganisms generally do not assimilate CO_2 through the central metabolism (13). Over the past five years, there has been great success in the production of CO_2 derivatives, which have the potential to be used as fuel and valuable chemicals by autotrophic germs. In this research, we showed that the removal of the main chain of the wild type formate dehydrogenase gene from *E. coli* and its replacement with the formate dehydrogenase gene from *Thiobacillus* Sp. KNK65MA strongly increased the growth rate of *E. coli* cells. In a previous study by Palsson *et al*., the whole-cell *in silico* model of *E. coli* metabolic network predicted that glycerol should be a preferred carbon source over glucose. However, the experimental findings were not consistent with the mentioned predictions. They indicated the adaptive evolution phenomenon for the bacteria to go from sub-optimal to the predicted optimal growth rate on glycerol (14–16). Our outcomes may explain these discrepancies between the *in silico* predictions and the experimental results from a different angle of view by elaborating on the role of *E. coli's* FDHs. Moreover, Palsson et al. demonstrated the role of glycerol metabolic enzymes in increasing the efficiency of *E. coli* growth on glycerol (14). Our findings may also reveal the presence of an alternative mechanism underlying CO_2 release during glycerol metabolism via *E. coli*'s FDH. Moreover, the replacement of the native FDH with TsFDH might lead to a potential path of a feedforward loop leading to glycerol efficiency as a carbon source. FDH decreases the growth rate of *E. coli* cells by converting valuable organic carbon resources into wasteful CO_2. Our *in silico* analysis showed that the omission of FDH leads to a significant increase in the expression of a number of important genes playing key roles in highly relevant metabolic pathways, particularly protein biosynthesis pathways and glycerol degradation pathways. These results were deciphered on the basis of the negative correlations between the two target genes (knockouts) and the rest of bacterial genes in the RNA-seq dataset. Our *in silico* achievements are in compliance with the experimental results and could be conceived as the mechanistic justifications of our experimental observations. Subsequently, it could be hypothesized that the reduced CO_2 leak may lead to a feedback loop boosting the production of organic carbon resources necessary for the higher growth rate of the bacteria. According to a study (13), *E. coli*'s FDHs have a strong tendency for regenerating CO_2 from formate. Among the studied formate dehydrogenases, TsFDH has potential advantages as a biocatalyst in the field of CO_2 reduction (7). TsFDH shows a favorable conversion of CO_2 to formate with up to a 21.2-fold higher turnover number than the conversion rate of well-known FDH enzymes such as *Candida boidini's* FDH.

Conclusions

The main problems with studying most FDHs published so far are protein instability, sensitivity to oxygen, the low conversion rate. The FDH of this study (TsFDH) shows some biochemical advantages over previously studied FDHs such as Candida bolidini's FDH including higher turnover number and insensitivity to the environmental oxygen (17–19). All in all, we showed that the FDH of *Thiobacillus* sp. KNK65MA will increase the growth rate of *E. coli* strain BL21 as well as K12 knockout strains (W4040 and W3688) with relatively high values owing to the prevention of CO_2 leakage from metabolic pathways. We showed that in the initial incubation periods, the growth rate was approximately equal between the expressing strain BL21 and *E. coli* 4040 (F chain with code 4509) and *E. coli* 3866 (D chain with code 4598), while at 24h the growth rates of the transformed and knockout bacteria were much higher than that of the BL21 control, which indicates the effect of formate dehydrogenase gene on the *E. coli* metabolism. In other words, there's a much higher growth rate by eliminating the mentioned chains in knockout strains. Additionally, the bacterial cells containing the recombinant TsFDH plasmid showed a significant growth rate compared to the cells lacking the recombinant plasmid. For example, between the original BL21 and BL21 cells harboring TsFDH plasmid, a growth enhancement of more than 200% was observed on M9+glycerol. By promoting *E. coli*'s growth rate through reducing CO_2 release we achieved two novel goals, namely, increasing the "biomass/carbon source" ratio and reducing atmospheric CO_2 during fermentation processes. We are currently performing *in silico* modeling and molecular dynamics simulations to identify target residue(s) in TsFDH for increasing the catalytic efficiency during the formate production reaction.

Materials And Methods

Escherichia coli Strains, Plasmids and Media

All *E. coli* strains and media used in this study are presented in Table 2. *Escherichia coli* BL21(DE3) was used for the expression of the recombinant FDH from *Thiobacillus sp.* KNK65MA (TsFDH). Two FDH knockout strains, JW3866 and JW4040 were purchased from Keio Collection.

Table 2
Strains and plasmids

Strains and plasmids	Related characteristics	Source
Strains	[*lon*] *omp*T *gal* (λ DE3) [*dcm*] Δ*hsd*S	Invitrogen
E. coli Bl21(DE3)	K12 Δ*fdhD*	Dharmacon
E. coli JW3866	K12 Δ*fdhF*	Dharmacon
E. coli JW4040	Ap^R, *T7* promoter, lac operator	Novagen
Plasmids	pET-21α, containing TsFDH gene from *Thiobacilusis sp* KNK65MA	This research
pET-21α		
pET-21α-TsFDH		

M9 medium with glycerol as carbon source and LB medium as a complex medium all containing 30 µg kanamycin were used for measuring bacterial growth. For BL21 with pET+TsFDH, the same media with ampicillin were used. In all samples containing pET+TsFDH, IPTG (0.5 mM final concentration) was added to the medium. The metabolic reactions consuming or producing formate (map01200 and C00058) were obtained from KEGG (20) (https://www.genome.jp/pathway/map01200+C00058). Using the results from KEGG pathway search for all the carbon fixation reactions, the contributing FDHs were identified. The kinetic parameters including K_{cat} and K_m of FDHs (EC: 1.17.1.9) for efficient formate formation were obtained from Brenda enzyme data bank (21) and the published articles were reviewed and compared in different bacteria (Online Resource 1). This approach revealed some interesting FDHs with relatively better kinetic parameters. Although, the results obtained by TsFDH might be quite satisfactory, we assume there are still some FDHs that deserve the attention for replacing the indigenous FDHs of *E. coli* for improving the growth efficiency. Our mentioned assumption is based on the ambiguity of assay conditions for some of the reported FDHs and lack of a gold standard for the kinetics comparisons. Scanning the kinetic parameters for a desired FDH suggested the *Thiobacillus sp*. KNK65MA (7).

Amino acid and nucleotide sequences of *Thiobacilusis sp* KNK65MA formate dehydrogenase were obtained from UniProt (accession # Q76EB7). cDNA of TsFDH was synthetized in pET21a by ZistEghtesadMad based on reference sequence (Q76EB7). Two knockout strains of K12 *Escherichia coli*, JW 4040 and JW 3866, with the deletion of fdhF and fdhD genes, respectively were purchased from Dharmacon. The stocks of the Knockout *E. coli* strains were cultured on LB broth and M9+Glycerol media followed by incubation at $37°^C$ for 24 hours (22).

The expression strain BL21 was also used as control to compare the growth rates. All strains were cultured at the same time under the same conditions on LB broth media at 37°C and 200 rpm. Competent cells of the BL21, *E. coli* 4040, *and E. coli* 3866 were prepared as previously mentioned (22). pET21, a plasmid containing a fusion gene to express format dehydrogenase of *Thiobacillus sp.* KNK65MA

(pET+TsFDH), was transformed in competent BL21cells, *E. coli* JW4040 and *E. coli* JW3866 on LB Agar with Amp (100 mg / ml) followed by incubation overnight at 37°C. The colonies containing plasmid were selected and cultured on a 10 ml LB broth with Amp as a primary culture and incubated at 37°C, 200 rpm for 24 hours. Then the culture was carried out in 200 ml of the LB broth containing Amp (100 mg / ml) and they were incubated at 37°C and at 200 rpm for 24 hours. Also, the bacteria BL21, *E. coli* JW 4040 and *E. coli* JW3866 lacking plasmid were simultaneously cultivated and incubated on LB broth or M9-Glycerol + 50 µg/ml Kanamycin under identical conditions with plasmid-containing samples.

Media and culture conditions

M9 medium + glycerol containing 30 µg kanamycin was used for measuring bacterial growth for BL21 with pET+TsFDH. The same media with ampicillin were also used. In all samples containing pET+TsFDH, IPTG (0.5 mM final concentration) was added to the medium.

Growth measurements

Bacteria were grown in batch cultures at 37°C in shaker incubator in 50 mL flasks. 1000 µL samples were taken in triplicate at indicated time intervals and the absorbance was measured at 600nm. During incubation, plasmid-free bacteria and plasmid-containing ones were sampled at different times, namely 2h, 4h, 6h, 8h, 10h, 12h and 24hr. To determine the growth rate of bacteria at the above time intervals, using a spectrophotometer at 600 nm wavelength, optical absorption, cell growth and growth rates were measured.

In silico analysis

Data preparation

In order to achieve a deeper insight into our observations, we applied correlation analysis, PCA and linear model analysis. We looked for an *E. coli* RNA dataset which could reflect the maximum possible transcriptional variations so that we would be able to calculate significant correlations between the genes. Moreover, the number of genes involved in the gene expression profile was important to calculate as many correlations as possible. With this aspiration, we fetched an *E. coli* RNA-seq dataset comprising 152 RNA-seq count samples under 34 different growth conditions (GEO accession GSE94117). These samples were taken from both exponential and stationary phases. One unique aspect of this highly pertinent dataset is the fact that it is sampled under 34 different growth conditions leading to a wider range of differentially expressed genes because of the different metabolic needs (23).

Data preprocessing

Using Python version 3.6.1, 152 samples of RNA-seq count files were merge. The counts were converted into count per million (CPM) and were log2 transformed. The resulting data were z-score transformed per gene across all samples. Quality control was performed as sample-level box plots before and after data preprocessing (Additional file 1: Fig. S3 and S4).

Correlation analysis

The Spearman rank-order correlation coefficient, being a nonparametric measure, examines the monotonic relationship between the ordinal values of the variables. Contrary to the Pearson correlation, the Spearman's rank correlation is not based on the assumption that the variables are normally distributed. Spearman correlation coefficient spans between -1 and +1 with 0 indicating no correlation. Correlation coefficients of -1 or +1 imply perfect monotonic relationship. Using the spearman function from the sub-package scipy.stats (24) the correlations between each of the two gene knockouts and the rest of the genes were calculated. The negatively correlated genes were chosen for fdhD and fdhF, all of which with p-values and FDRs less than 0.01 and "Spearman's ρ" < -0.2. Among these anti-correlated genes, the top 20 ones were chosen for further analysis. Since the top anti-correlated genes were already of p-values and FDRs below the specified threshold, there was no need for fishing out the differentially expressed genes in advance. In other words, the significant gene expression co-variations are reflected in the significant Spearman correlation coefficients.

PCA analysis

PCA as a dimensionality reduction technique was used to compare the gene expression profile dispersion of the bacteria based on the diversity of the expression levels of their genes. To this aim, pca function from mixOmics R package was used (25).

Linear regression analysis

Scatter plots of the top 40 anti-correlated genes were generated against each of the knoked out genes. Linear regression models were fit to the data points to show the overall trendlines (26).

Pathway analysis

BioCyc (27) database of microbial genomes and metabolic pathways was used to find the pathways each of the anti-correlated genes are attributed to.

Abbreviations

PCA
Principal Component Analysis
FDH
Formate Dehydrogenase
TsFDH
Thiobacillus sp. KNK65MA
DE3
Escherichia coli BL21
CPM
Count Per Million

Declarations

Acknowledgements

We thank Payam Emami, Rui Benfeitas and Paulo Czarnewski at the National Bioinformatics Infrastructure Sweden (NBIS) at SciLifeLab for their fruitful discussions and help with the RNA-seq data analyses. We also thank Roghaieh Ghaderi Ternik for her help and support. Saman Hosseini Ashtiani thanks his supervisor Professor Arne Elofsson for his valuable discussions and support in addition to acknowledging the fund by EU-ITN project ProteinFactory (MSCA-ITN-2014-ETN-642836) and the Swedish Research Council (Grant 2016-03798).

Authors' contributions

RR and BB conducted experiments. AAS and MHM helped with the design of experiments. SHA conceived and executed the bioinformatics analysis sections. BB, RR and SHA wrote the manuscript. All authors read and approved the manuscript.

Funding

- We had no specific funding for this project.

Availability of data and materials

- All the data supporting the conclusions of this paper are included in the context of the paper and the additional files. All the codes for the preprocessing and the analysis of the RNA-seq dataset are available at https://github.com/SamanAshtiani/ecoli_fdh.git.

Declarations

Ethics approval and consent to participate

- Not applicable.

Consent for publication

- Not applicable.

Competing interests

-Not applicable.

References

1. Solomon S, Plattner G-K, Knutti R, Friedlingstein P. Irreversible climate change due to carbon dioxide emissions. Proc Natl Acad Sci U S A [Internet]. 2009 Feb 10 [cited 2021 Jul 7];106(6):1704–9.

Available from: http://www.ncbi.nlm.nih.gov/pubmed/19179281

2. Gong F, Zhu H, Zhang Y, Li Y. Biological carbon fixation: From natural to synthetic. J CO2 Util. 2018 Dec 1;28:221–7.
3. Cotton CA, Edlich-Muth C, Bar-Even A. Reinforcing carbon fixation: CO2 reduction replacing and supporting carboxylation. Curr Opin Biotechnol. 2018 Feb 1;49:49–56.
4. Hädicke O, Kamp A von, Aydogan T, Klamt S. OptMDFpathway: Identification of metabolic pathways with maximal thermodynamic driving force and its application for analyzing the endogenous CO2 fixation potential of Escherichia coli. PLOS Comput Biol [Internet]. 2018 Sep 1 [cited 2021 Jul 7];14(9):e1006492. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1006492
5. Moon M, Park GW, Lee JP, Lee JS, Min K. Recent progress in formate dehydrogenase (FDH) as a non-photosynthetic CO2 utilizing enzyme: A short review. J CO2 Util. 2020 Dec 1;42:101353.
6. Hoelsch K, Sührer I, Heusel M, Weuster-Botz D. Engineering of formate dehydrogenase: synergistic effect of mutations affecting cofactor specificity and chemical stability. Appl Microbiol Biotechnol 2012 976 [Internet]. 2012 May 17 [cited 2021 Jul 7];97(6):2473–81. Available from: https://link.springer.com/article/10.1007/s00253-012-4142-9
7. Choe H, Joo JC, Cho DH, Kim MH, Lee SH, Jung KD, et al. Efficient CO2-reducing activity of NAD-dependent formate dehydrogenase from Thiobacillus sp. KNK65MA for formate production from CO 2 gas. PLoS One. 2014;9(7):e103111.
8. Consortium TU, Bateman A, Martin M-J, Orchard S, Magrane M, Agivetova R, et al. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res [Internet]. 2021 Jan 8 [cited 2021 Sep 18];49(D1):D480–9. Available from: https://academic.oup.com/nar/article/49/D1/D480/6006196
9. Riesenberg D, Schulz V, Knorre WA, Pohl HD, Korz D, Sanders EA, et al. High cell density cultivation of Escherichia coli at controlled specific growth rate. J Biotechnol. 1991;20(1):17–27.
10. Basan M, Hui S, Williamson JR. ArcA overexpression induces fermentation and results in enhanced growth rates of E. coli. Sci Rep. 2017;7(1):1–7.
11. Iwadate Y, Funabasama N, Kato JI. Involvement of formate dehydrogenases in stationary phase oxidative stress tolerance in Escherichia coli. FEMS Microbiol Lett. 2017;364(20):fnx193.
12. Santos-Zavaleta A, Salgado H, Gama-Castro S, Sánchez-Pérez M, Gómez-Romero L, Ledezma-Tejeida D, et al. RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in E. coli K-12. Nucleic Acids Res [Internet]. 2019 Jan 8 [cited 2021 Sep 19];47(D1):D212–20. Available from: https://academic.oup.com/nar/article/47/D1/D212/5160972
13. Gong F, Liu G, Zhai X, Zhou J, Cai Z, Li Y. Quantitative analysis of an engineered CO2-fixing Escherichia coli reveals great potential of heterotrophic CO2 fixation. Biotechnol Biofuels. 2015;8(1):86.
14. Ibarra RU, Edwards JS, Palsson BO. Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. Nat 2002 4206912 [Internet]. 2002 Nov 14 [cited 2021 Sep 20];420(6912):186–9. Available from: https://www.nature.com/articles/nature01149

15. Cheng K-K, Lee B-S, Masuda T, Ito T, Ikeda K, Hirayama A, et al. Global metabolic network reorganization by adaptive mutations allows fast growth of Escherichia coli on glycerol. Nat Commun 2014 51 [Internet]. 2014 Jan 31 [cited 2021 Sep 20];5(1):1–9. Available from: https://www.nature.com/articles/ncomms4233
16. K M-G, N F, HM C, G M-B, G H-C, OT R, et al. New insights into Escherichia coli metabolism: carbon scavenging, acetate metabolism and carbon recycling responses during growth on glycerol. Microb Cell Fact [Internet]. 2012 Apr 18 [cited 2021 Sep 23];11. Available from: https://pubmed.ncbi.nlm.nih.gov/22513097/
17. Yamamoto H, Mitsuhashi K, Kimoto N, Kobayashi Y, Esaki N %J A microbiology, biotechnology. Robust NADH-regenerator: Improved α-haloketone-resistant formate dehydrogenase. 2005;67(1):33–9.
18. Slusarczyk H, Felber S, Kula MR, Pohl M. Stabilization of NAD-dependent formate dehydrogenase from Candida boidinii by site-directed mutagenesis of cysteine residues. Eur J Biochem. 2000;267(5):1280–9.
19. Hartmann T, Leimkühler S. The oxygen-tolerant and NAD+-dependent formate dehydrogenase from Rhodobacter capsulatus is able to catalyze the reduction of CO2 to formate. FEBS J. 2013;280(23):6083–96.
20. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res [Internet]. 2017 Jan 4 [cited 2021 Aug 18];45(D1):D353–61. Available from: https://academic.oup.com/nar/article/45/D1/D353/2605697
21. Chang A, Jeske L, Ulbrich S, Hofmann J, Koblitz J, Schomburg I, et al. BRENDA, the ELIXIR core data resource in 2021: new developments and updates. Nucleic Acids Res [Internet]. 2021 Jan 8 [cited 2021 Sep 18];49(D1):D498–508. Available from: https://academic.oup.com/nar/article/49/D1/D498/5992283
22. Sambrook J. Molecular cloning: a laboratory manual [Internet]. Third edition. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press, [2001] ©2001; 2012. Available from: https://search.library.wisc.edu/catalog/999897924602121
23. Caglar MU, Houser JR, Barnhart CS, Boutz DR, Carroll SM, Dasgupta A, et al. The E. coli molecular phenotype under different growth conditions. Sci Rep. 2017;7(October 2016):1–15.
24. Kokoska S, Zwillinger D. CRC Standard Probability and Statistics Tables and Formulae, Student Edition [Internet]. CRC Standard Probability and Statistics Tables and Formulae, Student Edition. CRC Press; 2000 [cited 2021 Jul 7]. Available from: https://www.taylorfrancis.com/books/mono/10.1201/b16923/crc-standard-probability-statistics-tables-formulae-student-edition-stephen-kokoska-daniel-zwillinger
25. Rohart F, Gautier B, Singh A, Cao K-AL. mixOmics: An R package for 'omics feature selection and multiple data integration. PLOS Comput Biol [Internet]. 2017 Nov 1 [cited 2021 Jul 7];13(11):e1005752. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005752

26. Davison ML, Kim S-K, Close C. Factor Analytic Modeling of Within Person Variation in Score Profiles. http://dx.doi.org/101080/00273170903187665 [Internet]. 2009 [cited 2021 Jul 7];44(5):668–87. Available from: https://www.tandfonline.com/doi/abs/10.1080/00273170903187665
27. Karp PD, Billington R, Caspi R, Fulcher CA, Latendresse M, Kothari A, et al. The BioCyc collection of microbial genomes and metabolic pathways. Brief Bioinform [Internet]. 2018 Mar 27 [cited 2021 Jun 3];20(4):1085–93. Available from: https://academic.oup.com/bib/article/20/4/1085/4084231

Supplementary

Supplementary Information is not available with this version.

Figures

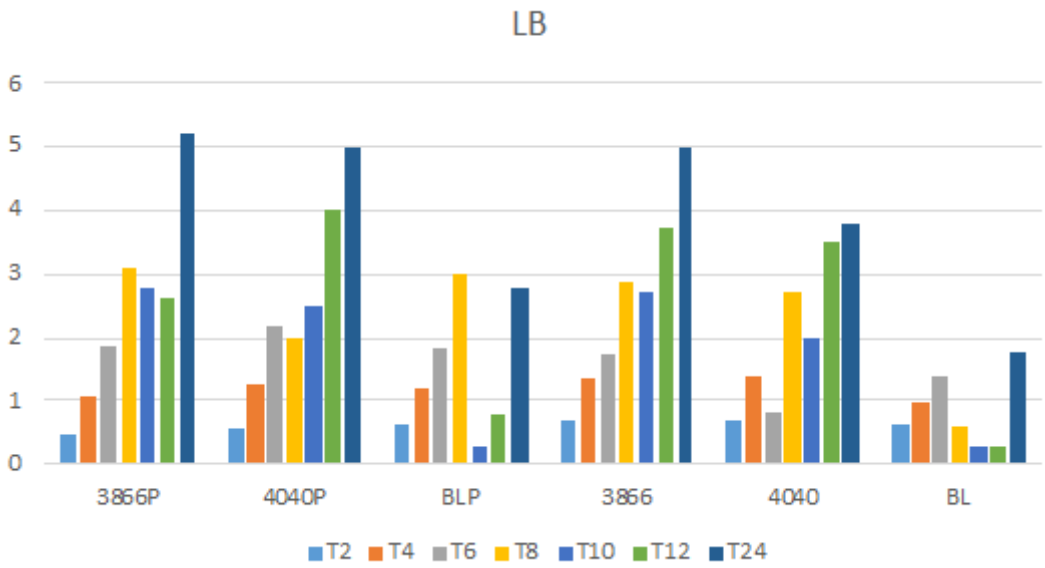

Figure 1

Growth rate comparison among bacteria on LB medium at OD 600 nm at different time points

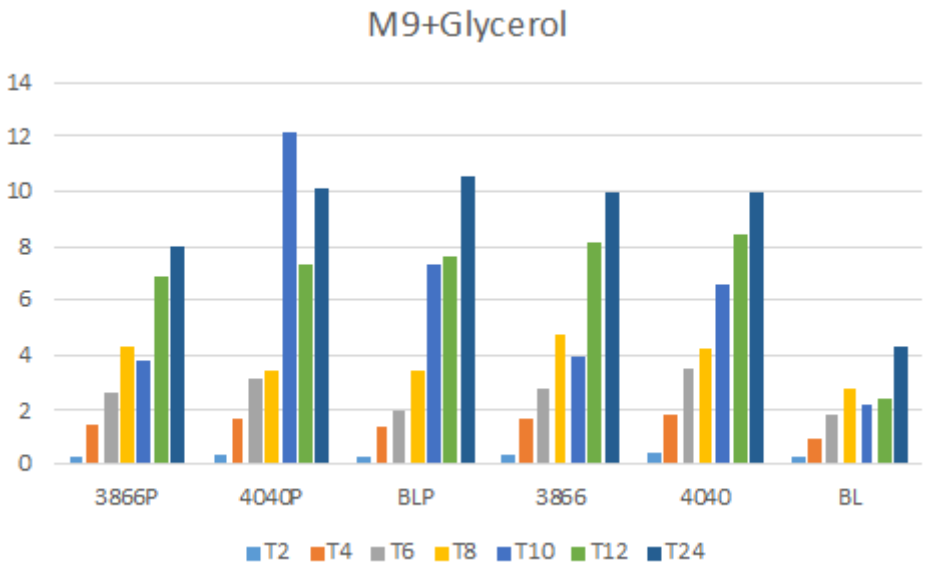

Figure 2

Growth rate comparison among bacteria on M9+Glycerol medium at OD 600 nm at different time points

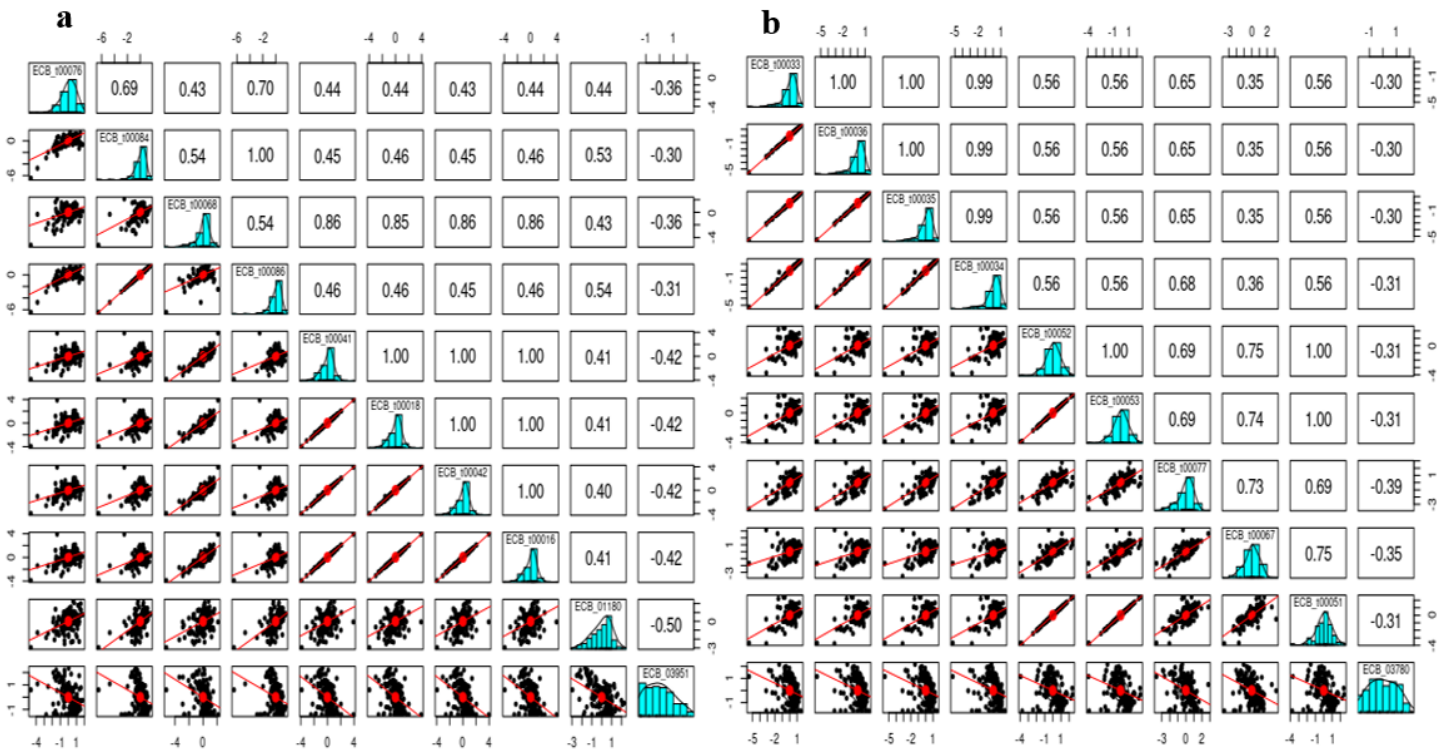

Figure 3

Top anti-correlated genes exhibit a negative slope of the fitted linear model against the knock-out target genes i.e., fdhF (a) and fdhD (b). For better resolution only top nine anti-correlated genes are depicted in the plot