J Ind Microbiol Biotechnol
DOI 10.1007/s10295-016-1818-7

BIOENERGY/BIOFUELS/BIOCHEMICALS - ORIGINAL PAPER

Improved glycerol to ethanol conversion by *E. coli* using a metagenomic fragment isolated from an anaerobic reactor

Inés Loaces[1] · Cecilia Rodríguez[1] · Vanesa Amarelle[1] · Elena Fabiano[1] · Francisco Noya[1]

Received: 25 February 2016 / Accepted: 30 July 2016

Abstract Crude glycerol obtained as a by-product of biodiesel production is a reliable feedstock with the potential to be converted into reduced chemicals with high yields. It has been previously shown that ethanol is the primary product of glycerol fermentation by *Escherichia coli.* However, few efforts were made to enhance this conversion by means of the expression of heterologous genes with the potential to improve glycerol transport or metabolism. In this study, a fosmid-based metagenomic library constructed from an anaerobic reactor purge sludge was screened for genetic elements that promote the use and fermentation of crude glycerol by *E. coli*. One clone was selected based on its improved growth rate on this feedstock. The corresponding fosmid, named G1, was fully sequenced (41 kbp long) and the gene responsible for the observed phenotype was pinpointed by in vitro insertion mutagenesis. Ethanol production from both pure and crude glycerol was evaluated using the parental G1 clone harboring the ethanologenic plasmid pLOI297 or the industrial strain LY180 complemented with G1. In mineral salts media containing 50 % (v/v) pure glycerol, ethanol concentrations increased two-fold on average when G1 was present in the cells reaching up to 20 g/L after 24 h fermentation. Similar fermentation experiments were done using crude instead of pure glycerol. With an initial OD_{620} of 8.0, final ethanol concentrations after 24 h were much higher reaching 67 and 75 g/L with LY180 cells carrying the control fosmid or the G1 fosmid, respectively. This translates into a specific ethanol production rate of 0.39 g h^{-1} OD^{-1} L^{-1}.

Keywords Crude glycerol · Biodiesel · Ethanol · Functional metagenomics

Introduction

Biodiesel is one of the best alternative liquid fuels. It is an energy efficient renewable fuel made from vegetable oil and animal fats [24]. It is also environmentally friendly, biodegradable and the only alternative fuel with the potential to completely displace its petroleum counterpart. Conventional diesel engines can use 100 % biodiesel with only minor modifications to fuel system and little impact on performance [22]. Biodiesel is produced by transesterification or alcoholysis of triacylglycerol with methanol or ethanol in a process that generates about 10 % (w/w) glycerol as a by-product. As a result of its worldwide production, 2 million tons of glycerol per year consistently reached the market for the past 10 years [8] driving down glycerol market prices and increasing disposal costs [55]. Finding ways to extract more value from this by-product is essential for balancing the economic equation of the biodiesel industry and increasing the competitiveness of this biofuel against its fossil counterpart [53].

Many commercially important compounds are currently produced from sugars, such as glucose and xylose, obtained from starch, sucrose or lignocellulose. When compared to these sugars, glycerol is a more reduced substrate, better suited to achieve higher yields in those situations where production of the final product is limited by the availability of reducing equivalents [9, 12]. Several previous works have evaluated the utilization of glycerol to produce diverse

✉ Francisco Noya
fnoya@iibce.edu.uy

[1] Departamento de Bioquímica y Genómica Microbianas, Instituto de Investigaciones Biológicas Clemente Estable, Av. Italia 3318, 11600 Montevideo, Uruguay

Published online: 13 August 2016

value-added chemicals such as 1,3-propanediol, hydrogen or amino acids [30, 39, 45]. The finding that ethanol is the main product of glycerol fermentation by *E. coli* [32], together with its reported ability to grow on glycerol as the sole carbon source [28], opened up the possibility to use this industrial workhorse to also produce ethanol from biodiesel production main by-product.

Many improvements and modifications were done in glycerol fermentation conditions and *E. coli* metabolic pathways, looking for better ethanol yields [10, 11, 19, 29, 36, 37, 49, 51, 54]. In one study, mutations in many seemingly unrelated pathways (amino acids, pentose phosphate, TCA cycle, gluconeogenesis, amino sugar and nucleotides) showed positive effects in glycerol to ethanol conversion [50]. However, none of the work done so far aimed at identifying novel heterologous genes that could enhance this process in a strain already optimized for ethanol production.

Metagenomics is the analysis of all the genomes that compose a microbial community [18]. It involves the extraction of the genomic DNA directly from the environmental sample without cultivation. This DNA can be either sequenced, or fragmented, cloned into a vector and introduced into a host to construct a metagenomic library. Functional metagenomics implies the screening of this library for phenotypes or functions of interest that are newly expressed in the host cells [1]. One major advantage of this approach is that no previous information on the target genes is required. Thus, this strategy has allowed the identification of entirely new classes of genes [4, 14, 44].

In this study, a functional metagenomic approach was used to identify environmental DNA fragments that were able to improve both the growth of *E.coli* with crude glycerol as the sole carbon source, as well as ethanol production by an ethanologenic strain.

Materials and methods

Metagenomic library construction

Metagenomic DNA was isolated as previously described [43] from the purge sludge of an anaerobic reactor fed with ruminal digesta [total solids (TS) 19.3 % (w/w); volatile solids (VS) 17.3 % (w/w)] and tannery fats (TS 25.1 %, VS 16.6 %), inoculated with an activated purge sludge of an oil-wastewater treatment plant (TS 25.6 %, y VS 21.7 %). DNA fragments ranging from 30 to 50 kbp were selected and cloned into fosmid pCC1FOS and transduced into EPI300-T1^{R} cells (Epicentre Biotechnologies, Madison, WI, US) as previously described [26].

Table 1 Composition of crude glycerol

Component	% (wt/wt)
Glycerol	78.6
Methanol	4.0
Ash	4.7
Water	2.8
OMNG[a]	13.9
pH	10–12

[a] *OMNG* organic matter non glycerol

Screening of metagenomic library

Crude glycerol was obtained as a by-product of the biodiesel industry (ALUR, Montevideo, Uruguay) and its composition is detailed in Table 1. Agar-plates of minimal media M9 (2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 50 mM $Na_2HPO_4{\cdot}7H_2O$, 20 mM KH_2PO_4, 0.05 % NaCl, 0.1 % NH_4Cl, thymine 50 μg/mL, leucine 200 μM) containing 4.0 % (v/v) crude glycerol as the only carbon source were used for the screening of the metagenomic library. The library was grown under aerobic conditions overnight on M9 agar-plates with 1 % (w/v) glucose and then replica plated onto screening media. Positive clones were selected based on visually distinguishable growth after 16 h of incubation at 30 °C. For growth assays in liquid media, crude glycerol was previously neutralized. For this purpose, a 50 % (v/v) crude glycerol dilution was adjusted to pH 7.0 with 1 % HCl, filtered through filter paper and the resulting liquid fraction was sterilized by autoclaving. The growth of selected clones was assessed in liquid M9 media with 4.0 % (v/v) neutralized crude glycerol as the sole carbon source by measuring the optical density at 620 nm (OD_{620}) during 24 h in 96 well plates using Varioskan Flash Multimode Reader (Thermo Scientific, MA, US).

G1 fosmid sequencing and in silico analysis

Fosmid G1 was sequenced in an Ion Torrent PMG platform, using Ion PGM TM Sequencing Kit (Life Technologies, Invitrogen, CA, US) in Ion 314 ™ Chip (Life Technologies). The sequencing library was constructed using Ion Xpress™ Plus commercial Kit (fragmentation of 200 bp) (Life Technologies) and Ion PGM™ Template OT2 200 Kit (Life Technologies) was used for the template preparation. 114,683 reads were obtained, with a mean read length of 206 bp. *E. coli* genomic DNA contamination was filtered out by the Mirabait software using the genome of K12 as template. The reads obtained were assembled using the software Mira 4.0 rc5, with a coverage of 40X (7500 reads of 235 bp on average were randomly selected). Manual

curation was made with Gap5 program. Open reading frames (ORFs) were predicted from the assembled contigs using MetaGene [38] hosted on the WebMGA online server [52]. The translated protein sequences of the predicted genes were subjected to Blastp comparison [59] with the non-redundant protein database (NR) and the CDD domain database (including Clusters of Orthologous Groups classification, COG). The PhyloPythiaS server [40] was used to infer the phylogenetic origin of the environmental DNA sequence.

ORF G1_23 sequence was compared to the Pfam database [41] at http://pfam.xfam.org/. Complementary domain analysis was made with HMMER [15] using profile Hidden Markov Model (HMMs). The prediction of transmembrane motifs was performed by TMHMM 2.0 server [23], and protein subcellular location prediction was determined by PSORTb v.3.0 [33]. The prediction of signal peptides was performed using SignalP 4.1 [35].

Closest homologs to ORF G1_23 were selected and multiple sequences alignments were computed with MEGA 6 [48] by Muscle method. MEGA 6 was also used for the construction of phylogenetic trees using the Maximum likelihood method.

The DNA sequence of G1 is available in the GenBank database under the accession number KU952095.

Gene identification by mutagenesis

G1 fosmid was in vitro mutagenized using EZ-Tn5 Insertion Kit (Epicentre Biotechnologies) according to the manufacturer's instructions. The mutagenized mixture was used to transform fresh electrocompetent EPI300-T1^R cells and Tn5-mutant fosmids were selected in LB plates supplemented with chloramphenicol 12.5 μg/ml and kanamycin 50 μg/ml. Transformants were replica plated onto M9 plates containing 4 % (v/v) crude glycerol and incubated at 30 °C for 16 h. Colonies displaying growth impairment regarding the parental strain EPI300 (G1) were selected. Single Tn5 insertion was confirmed by restriction analysis. DNA flanking the Tn5 was sequenced by conventional Sanger method (Macrogen, Seoul, Korea) using primer KAN-2 FP-1 (5′-ACCTACAACAAAGCTCTCATCAACC-3′).

Ethanol production assays

For ethanol production assays two strains were used: (1) EPI300-T1^R complemented with both fosmid G1 and the ethanol production enhancer plasmid pLOI297 (carrying *adhB* and *pdc* from *Zymomonas mobilis* CP4; ATCC 68239) [2]; (2) the ethanologenic strain LY180 [Δ*frd*BC::(Zm frg *cel*YEc) Δ*ldh*A::(Zm frg *cas*A BKo) *adh*E::(Zm frg *est*ZPp FRT) Δ*ack*A::FRT rrlE::(*pdc adh*A *adh*B FRT) Δ*mgs*A::FRT], a derivative of LY168 [21, 31], complemented with G1 fosmid. All the strains were conserved at −80 °C.

Batch fermentations were carried out in Fleakers minifermenters (Corning, New York, USA) containing 200 mL of M9 minimal media for EPI300-T1^R cells, or NBS media [58] for LY180 cells, supplemented with different concentrations of neutralized glycerol and incubated without aeration at 37 °C, and 150 rpm. pH was adjusted to 7.0 before autoclaving and remained un-controlled during the fermentations. Fleakers were inoculated with an initial OD_{620} of 1 or 8, using the cellular pellet of overnight cultures grown in LB media supplemented with 20 g/L of glucose. To provide vitamins and additional nutrients in the experiments using pure glycerol, tryptone and yeast extract at 1 % (w/v) final concentration were added. 1 mM HCl-betaine was included when glycerol content was above 20 %. LY180 cells are unable to growth on pure glycerol so 1 mM selenite was added to the media.

For ethanol production kinetics assays, fermentations were carried out in 1.2 L fermenters, instead of Fleakers. Just after inoculation, oxygen was displaced from the headspace with nitrogen and incubation proceeded without aeration.

Quantitative determinations

Ethanol and oxygen were determined by gas phase detectors BCP-EtOH and BCP-O2EC (BlueSens, Herten, Germany). H_2 and CO_2 were determined by gas chromatography using a 310 C gas chromatograph (SRI, CA, USA) with argon as gas carrier.

Glycerol determinations were done as previously described [6]. 0.5 mL of each sample was transferred into a 10 mL test tube containing 1.5 mL of 50 % ethanol and 1.2 mL of 10 mM sodium periodate. The mixture was shaken for 30 s. Then, 1.2 mL of 0.2 M acetylacetone was added and the reaction was carried on at 70 °C for 1 min. Glycerol concentration was determined by reading the absorbance at 410 nm against a titration curve.

Volatile fatty acids (VFA) were determined by HPLC. Samples were centrifuged at 15.000 rpm and 5 °C for 15 min, and 100 μL of the supernatant was injected in a Waters modular HPLC system (Waters Associates, Milford, MA, USA). The HPLC system consisted in a binary HPLC pump (Waters 1525), an autosampler (Waters 717 plus), an ion-exclusion column Rezex ROA-Organic Acid H+ (8 %) (300*7.8 mm) (Phenomenex, Torrance, CA, USA), and a photodiode array detector (Waters 2998) linked to Empower 2 (Waters) chromatography data software. The temperature of the column was set at 30 °C. Mobile phase was H_2SO_4 0.015 M at 0.3 mL/min isocratic elution. Quantifications were performed using a standard solution containing 5 mg/mL of each VFA.

All assays were done in triplicates. Significance analysis were made using a Tukey's test (p value < 0.05) using the Infostat software.

Results and discussion

Metagenomic library screening

Escherichia coli is naturally able to grow aerobically on glycerol as the sole carbon source [28]. However, crude glycerol obtained as a by-product of biodiesel production has a very alkaline pH and is contaminated with organic and inorganic compounds which hinders cell growth. The impurities of the industrial feedstock were mainly composed of ash, methanol, non-glycerol organic materials and water (Table 1). The glycerol content was between 70 and 78 % (wt/wt). Therefore, we set to further improve the growth of *E.coli* on crude glycerol by its complementation with random environmental DNA. For this end, we used a previously constructed metagenomic library to search for clones eliciting an improved growth in crude glycerol when compared with the parental strain EPI300-T1^{R}.

The 100,000-clones metagenomic library was screened on agar-plates of M9 minimal media with 4 % crude glycerol as the sole carbon source. After 16 h of incubation at 30 °C under aerobic conditions the clones that grew faster than the parental strain were selected. This advantage was transient since after 48 h of incubation all colonies developed similar sizes.

Selected clones from the primary screen were confirmed by growth assays in liquid media. Since crude glycerol contains residual fatty acids that precipitate when mixed with buffered M9 media and interfere optical measurements, liquid assays were done with crude glycerol previously neutralized and cleared by filtration. Liquid M9 media was supplemented with 4 % neutralized crude glycerol and inoculated with the clones selected in the primary screen. Growth rate was measured by OD_{620} and the assay was repeated at least three times for each selected clone (data not shown). One of these clones, named G1, consistently showed improved rates compared with the parental strain and was selected for further characterization.

G1 fosmid sequence analysis

The G1 fosmid (41,312 bp long, including the pCC1FOS backbone) was fully sequenced and assembled. Its nucleotide sequence did not show homology to any reference genome in the NCBI database (refseq_genomic by October 2015). The PhyloPythiaS server [40] inferred that G1 was related to *Alistipes*. Thirty-eight putative ORFs were predicted and their analysis showed homology to proteins from Phyla Bacteroidetes (*Bacteroides*, *Hymenobacter*, *Elizabethkingia*, *Alistipes sp*. and *Prevotella sp*.) and Firmicutes (*Desulfitobacterium*) (Table 2). From these results, no ORF could be inferred responsible for the growth phenotype conferred by G1 fosmid.

To identify the gene responsible for the improved growth rate of the G1 clone in crude glycerol, fosmid G1 was subjected to in vitro Tn5-mediated mutagenesis. After Tn5 insertion, over 200 km-resistant colonies were identified on M9-crude glycerol plates. Only the smallest of these were selected and the single insertion of the transposon was verified by DNA restriction analysis. Three colonies were identified in this way and were sequenced to map Tn5 insertion inside fosmid G1 sequence (Fig. 1). In two of them, mt9 and mt34, Tn5 interrupted the sequence of ORF G1_23, while in mt40, it was inserted 5' of this ORF. Loss of function as a result of Tn5 insertion in the three mutants was confirmed by growth assays (Fig. 2). As negative control, EPI300 carrying a randomly selected fosmid was used. The mutant whose growth rate was most affected was mt9. This phenotype is consistent with the position of the Tn5 insertion inside the coding region but close to the first codon of the ORF. It is likely that no gene product is expressed in this mutant. The mt34 mutant displayed an intermediate growth rate, slightly better than the negative control. In this mutant Tn5 is inserted at about three quarters of the coding region and a partially functional product could be expressed from it. The least affected was the mt40 mutant. In this case, the mutation is located upstream of the ORF most likely inside the regulatory region of the gene. The gene product is not affected and is likely that the observed loss of function was due to an untimed or reduced gene expression.

ORF G1_23 showed only modest amino acid sequence homology with hypothetical proteins from *Bacteroidetes* and *Prevotella* (Fig. 3, Table 3). No Pfam or COG conserved domains could be identified. A putative signal peptide encompassing the first 23 residues of ORF G1_23 was identified by SignalP [5] and PSORTb V3.0 [33]. The cleavage site was predicted between amino acids 23 and 24 (VSC-QS $D = 0.715$ D-cutoff $= 0.570$, data not shown). A transmembrane domain was predicted on the first 25 aminoacids of ORF G1_23 with 40 % probability and the remainder of the protein was located extracellularly with high probability (100 %). Since G1_23 is only homologous to hypothetical proteins, no function could be assigned to it. Taking into account that the fosmid G1 was identified as a growth enhancer in crude glycerol, one may speculate that G1_23 could function as a detoxifier of one or more growth inhibitors present in the crude glycerol (Table 1). For example, it has been shown an improvement in tolerance to crude glycerol when overexpressing trehalose biosynthetic genes in *E. coli* [34]. It is also possible

Table 2 ORFs assigned to G1 sequence

#ORF[a]	Start	End	Strand	Length (aa)	Closest homolog using BLASTP[b]			
					Description [organism][c,d]	*E*-value[c]	Identity (%)[c]	Gene bank ID[c]
G1_1	41040	255	−	175	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase [Alistipes sp. CAG:831]	1.00E−40	75	WP_021928565.1
G1_2	249	932	−	227	Outer membrane protein beta-barrel domain protein [Alistipes sp. CAG:831]	6.00E−38	39	WP_021928564.1
G1_3	929	1345	−	138	Lipoprotein [Bacteroidales bacterium CF50]	2.00E−31	46	AGY53676.1
G1_4	1464	2723	−	419	Putative uncharacterized protein [Bacteroides sp. CAG:770]	2.00E−114	45	WP_022285662.1
G1_5	2720	3589	−	289	GTP pyrophosphokinase [Bacteroidales bacterium CF50]	8.00E−78	47	AGY53675.1
G1_6	3568	4956	−	462	relA/SpoT family protein [Bacteroides sp. CAG:545]	0	66	WP_022017717.1
G1_7	4966	5211	−	81	Ribonuclease BN [Alistipes sp. CAG:514]	4.00E−11	44	WP_021839834.1
G1_8	5232	5954	−	240	Ribonuclease BN [Bacteroides sp. CAG:545]	2.00E−45	37	WP_022017716.1
G1_9	5959	6954	−	331	NAD-dependent DNA ligase LigA [Alistipes sp. CAG:831]	3.00E−144	65	WP_021928558.1
G1_10	6869	7987	−	372	NAD-dependent DNA ligase LigA [Alistipes sp. CAG:831	2.00E−153	64	WP_021928558.1
G1_11	8079	9311	+	410	6-phosphofructokinase [Alistipes timonensis]	0	67	WP_010260254.1
G1_12	9376	9615	+	79	Alpha-amylase [Hymenobacter norwichensis]	0.007	34	WP_022825614.1
G1_13	9882	10928	+	348	Alpha-amylase [Bacteroides barnesiae]	9.00E−97	48	WP_018709950.1
G1_14	11163	11564	+	133	Multispecies: RNA polymerase subunit sigma-24 [Desulfitobacterium]	7.00E−04	40	WP_014794121.1
G1_15	11561	11779	+	72	RNA polymerase sigma factor, sigma-70 family [Gordonibacter pamelaeae]	0.002	35	WP_015538795.1
G1_16	11837	12070	−	77	ADP-ribosylglycohydrolase family protein [Prevotella ruminicola]	1.00E−04	41	WP_013063668.1
G1_17	12034	12333	−	99	Hypothetical protein [prevotella sp. fd3004]	1.00E−19	57	WP_036911212.1
G1_22	20459	20749	+	96	Hypothetical protein brdcf_p399 [bacteroidales bacterium cf50]	1.00E−22	58	AGY53026.1
G1_23	**21400**	**22734**	**−**	**444**	**Hypothetical protein [Prevotella ruminicola]**	**8.00E−28**	**33**	**WP_013064512.1**
G1_24	23124	24230	+	368	Transcriptional regulator [Prevotella fusca]	1.00E−44	30	WP_025077699.1
G1_25	24227	24547	−	106	Phosphonate metabolism protein PhnP [uncultured bacterium]	2.00E−34	56	EKD31793.1
G1_26	24493	24999	−	168	Beta-lactamase domain-containing protein [Bacteroides sp. CAG:545]	2.00E−62	66	WP_022015881.1
G1_28	26416	26763	−	115	Dihydroneopterin aldolase [Alistipes sp. cag:435]	2.00E−32	46	WP_022324543.1
G1_29	26872	27588	−	238	Uroporphyrin-iii c/tetrapyrrole methyltransferase [Alistipes sp. cag:435]	3.00E−93	57	WP_022324546.1
G1_30	27530	29476	−	648	Hypothetical protein BRDCF_p1264 [Bacteroidales bacterium CF50]	9.00E−179	44	AGY53891.1
G1_31	29632	31986	−	784	Hypothetical protein BRDCF_p1264 [Bacteroidales bacterium CF50]	4.00E−55	25	AGY53891.1
G1_32	32050	33078	+	342	tRNA threonylcarbamoyladenosine modification protein TsaD [Alistipes]	0	73	WP_004328979.1
G1_33	33176	34621	+	481	Oxidoreductase [Bacteroides sp. CAG:770]	0	62	WP_022285915.1
G1_34	35079	36149	+	356	Putative PD-(D/E)XK nuclease family protein [Elizabethkingia anophelis]	4.00E−24	28	CDN76750.1
G1_35	36306	36833	−	175	Peptidyl-prolyl cis–trans isomerase [Alistipes sp. cag:831]	4.00E−72	70	WP_021927072.1
G1_36	36830	37420	−	196	Uncharacterized protein [Alistipes sp. cag:831]	1.00E−66	54	WP_021927077.1

Table 2 continued

#ORF[a]	Start	End	Strand	Length (aa)	Closest homolog using BLASTP[b]			
					Description [organism][c,d]	*E*-value[c]	Identity (%)[c]	Gene bank ID[c]
G1_37	37503	37754	−	83	50S ribosomal protein L31 [Alistipes sp. CAG:831]	4.00E−47	87	WP_021927126.1
G1_38	37823	38455	−	210	Hypothetical protein [Bacteroides fragilis]	4.00E−28	36	WP_032568461.1
G1_39	38961	39902	+	313	Phosphomannomutase [Bacteroides sp. CAG:770]	5.00E−137	63	AGY53846.1
G1_40	39937	40383	+	148	Hypothetical protein BRDCF_p1220 [Bacteroidales bacterium CF50]	1.00E−43	52	AGY53847.1
G1_41	40395	41036	+	213	Hypothetical protein BRDCF_p1220 [Bacteroidales bacterium CF50]	1.00E−71	47	AGY53847.1

[a] ORFs identified by MetaGene

[b] Assigned by RPS Blast

[c] Closest homolog identified by BlastP 2.2.30+ using non-redundant protein sequences

[d] ORF interrupted by Tn5 is shown in bold. ORFs belonging to pCC1FOS backbone are omitted

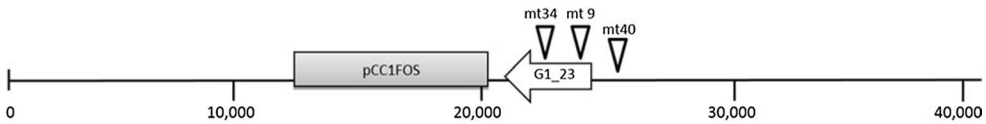

Fig. 1 Relative positions of transposon Tn5 insertion sites on G1 fosmid (41,312 bp). The *arrowheads* indicate the Tn5 insertion sites on each G1 mutant (mt34 at position 22,285; mt9 at 22,332; mt40 at 22,826). The location of ORF G1_23 (21,400–22,734) is indicated by an *arrow*, and the pCC1FOS backbone, between positions 12,510 and 20,519, by a *grey box*

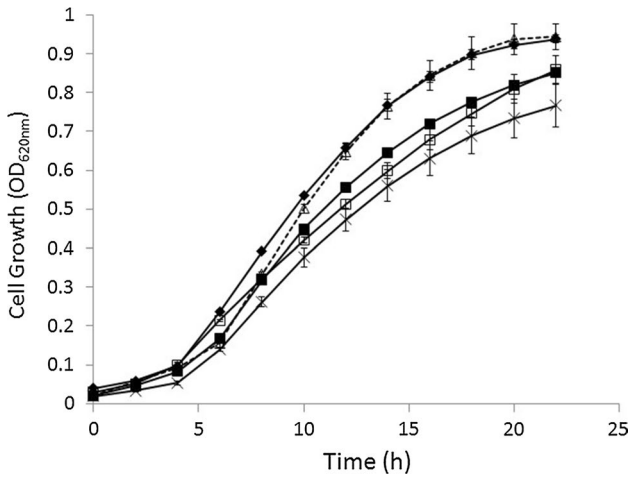

Fig. 2 Growth (as OD_{620}) in M9 4 % (v/v) crude glycerol of EPI300 (G1) (*filled diamond*), selected Tn5 mutants: mt9 (*multiplication sign*), mt 34 (*filled square*), mt40 (*open triangle*) and EPI300 carrying a randomly picked fosmid as negative control (*open square*). Incubation at 37 °C during 24 h. Experiments were conducted in triplicates, and averages were plotted. *Error bars* indicate standard deviations

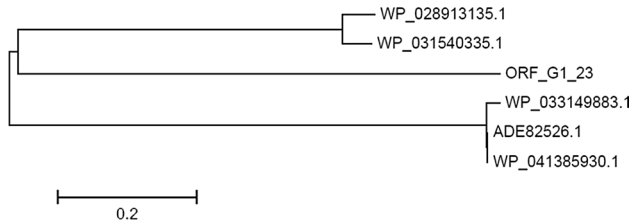

Fig. 3 Phylogenetic relationship of ORF G1_23 and its closest homologs identified using NCBI BlastP. Multiple sequences alignments were computed with MEGA 6 using the Muscle method. Phylogenetic tree was constructed using the Maximum likelihood method also implemented in MEGA 6

that G1_23 function is related to glycerol transport and/or metabolism. For example, the expression of glycerol dehydrogenase (Gcy) and dihydroxyacetone kinase (Dak) in *Saccharomyces cerevisiae* improved growth on glycerol [57].

Ethanol production using EPI300 strains

To determine whether fosmid G1 improves ethanol production from glycerol by *E.coli*, EPI300 (G1) was complemented with the ethanologenic plasmid pLOI297. This plasmid carries the PET operon [2] [alcohol dehydrogenase (*adh*) and pyruvate decarboxylase (*pdc*) genes from *Zymomonas mobilis*]. Due to the very low K_M of *pdc* for pyruvate (0.4 mM) this operon confers *E. coli* the ability

Table 3 Closest homologs of ORF G1_23 identified using NCBI BlastP

Related proteins (accession number)	Origin	Identity (%)	Max score	Query cover (%)	*E*-value	Reference
Hypothetical protein (WP_031540335.1)	Bacteroidetes	38	126	62	1.00E−28	
Hypothetical protein (WP_041385930.1)	Prevotella ruminicola	32	125	90	4.00E−28	
Hypothetical protein PRU_1480 (ADE82526.1)	Prevotella ruminicola 23	33	124	87	9.00E−28	[42]
Hypothetical protein (WP_033149883.1)	Prevotella sp. RM4	31	120	90	4.00E−26	
Hypothetical protein (WP_028913135.1)	Prevotella sp. MA2016	37	100	52	1.00E−19	

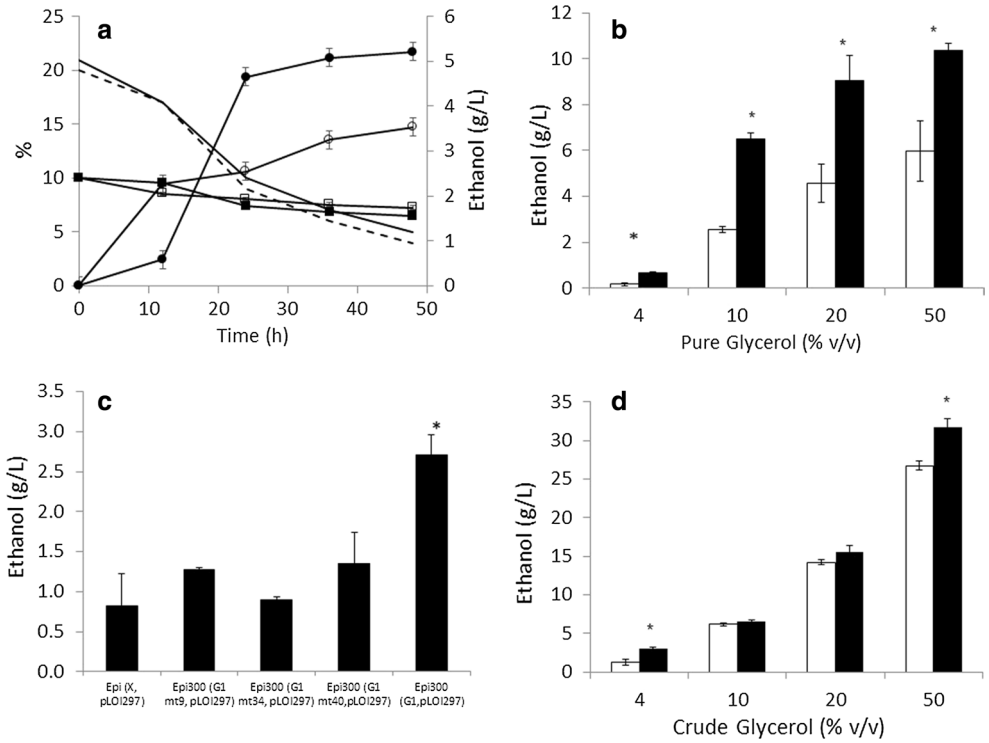

Fig. 4 Ethanol production by EPI300 (pLOI297) strains, in M9 minimum media with pure or crude glycerol as carbon source. **a** Kinetics of glycerol (*squares*) and oxygen consumptions (*lines*), and ethanol production (*circles*) from 10 % (v/v) pure glycerol during 48 h at 37 °C and 150 rpm by EPI300 (X, pLOI297) (*open symbols* and *dashed line*), EPI300 (G1, pLOI297) (*solid symbols* and *continuous line*); **b** ethanol production with different concentrations of pure glycerol added to media using EPI300 (X, pLOI297) (*empty bars*) or EPI300 (G1, pLOI297) (*full bars*) after 24 h of incubation at 37 °C, 150 rpm; **c** ethanol production from 10 % (v/v) pure glycerol using EPI300 (X, pLOI297), EPI300 (G1, pLOI297) and selected Tn5 mutants; **d** ethanol production with different concentrations of crude glycerol added to media (% v/v) using EPI300 (X, pLOI297) (*empty bars*) or EPI300 (G1, pLOI297) (*full bars*) after 24 h of incubation at 37 °C, 150 rpm. Experiments were done in triplicates, and averages were plotted. *Error bars* indicate standard deviations, (*asterisk)* is present when statistical significant differences were found between strains (*t* test, 95 % confidence level)

to produce ethanol as the principal fermentative product [47] by outcompeting other acidogenic pathways. Cells were grown under microaerophilic conditions in M9 media supplemented with 10 % pure glycerol, 1 % tryptone and 1 % yeast extract since *E. coli* requires a rich-media to effectively ferment glycerol [12]. Under these conditions oxygen can act as the final electron acceptor for reducing equivalents not eliminated by the *pdc/adh* pathway. Ethanol and glycerol were measured over time. As negative control, EPI300 carrying pLOI297 and a randomly selected fosmid (X) was used. Ethanol rose sharply during 12 h for EPI300 (X, pLOI297) and during 24 h for EPI300 (G1, pLOI297), being the maximum ethanol concentrations 3.5 g/L and 5.2 g/L by 48 h, respectively (Fig. 4a). Glycerol consumption after 48 h was 2.7 and 3.5 %, respectively, which represents an ethanol yield of 0.13 and 0.15 g per g of glycerol, respectively. These data underline the fact that the presence of G1 increased ethanol production rate (g/h) by 56 % and glycerol consumption by 30 % but did not improve ethanol yield much which was still far from its theoretical maximum of 0.50 g/g. Residual oxygen was 5 % by 48 h. One can conclude that, in these conditions, the majority of the

glycerol consumed is not being converted into ethanol (see below).

Cell growth was not affected significantly by the presence of G1. The initial OD_{620} was 1.0 and, by 48 h, OD_{620} was 1.2 for the control and 1.3 for the cells harboring G1 (data not shown). Although, G1 was selected by the better growth rate on glycerol containing plates, under microaerobiosis, better growth is no longer a dominant characteristic. A plausible explanation might be that under these conditions energy metabolism funnels the flux of carbon away from anabolism and growth. From a practical point-of-view, this is a desirable outcome.

As expected, increasing glycerol concentrations, increased ethanol production. However, in the presence of G1, ethanol concentrations were at least double of those obtained with the control fosmid (X) (Fig. 4b).

To determine whether the ORF G1_23 implicated on growth improvement on crude glycerol was also involved in the better ethanol yields obtained from the fermentation of pure glycerol, we repeated the fermentation experiments with the Tn5-induced mutants described above. The G1 fosmid mutants mt9, mt34 and mt40, introduced into EPI300 cells and complemented with pLOI297, were evaluated for ethanol production in 10 % of pure glycerol. All the mutants in ORF G1_23 showed maximum levels of ethanol production similar to the negative control strain (Fig. 4c) and significantly lower ($p < 0.05$) than the strain carrying the parental G1 fosmid. These results showed that ORF G1_23 is involved not only in improving growth on crude glycerol but also in better ethanol production from pure glycerol. Since G1_23 did improve ethanol production from pure glycerol it is less probable that its function is related to detoxification of inhibitors present in crude glycerol. Instead, it is more likely related to the transport or metabolism of glycerol itself.

To assess the ability of G1 fosmid to improve the fermentation of crude glycerol obtained from the industrial production of biodiesel, neutralized crude glycerol instead of pure glycerol was used in fermentation assays. Increasing initial concentrations of crude glycerol were added to minimal media, and ethanol production was evaluated. As with pure glycerol, G1 enhanced ethanol production albeit the improvements were not as large (Fig. 4d). In absolute terms, maximum ethanol concentrations were on average two- to three-fold higher with neutralized crude glycerol than with pure glycerol (Fig. 4b, d).

Furthermore, in both pure and crude glycerol, production of acetic, succinic, and propionic acids was detected (Table 4). This is not surprising since, in *E. coli,* ethanol, acetate and succinate are produced when glycerol is metabolized through the glycerol dehydrogenase and dihydroxyacetone kinase pathway [17, 58]. Interestingly, fermentations with crude glycerol showed a marked reduction in acetic, propionic and succinic accumulation and, at the same time, much better ethanol yields suggesting that the other components present in crude glycerol helped shift the flux of carbon from pyruvate into the *pdc/adh* pathway encoded by pLOI297 and away from the native routes. This is similar to previous reports with pLOI297 in which ethanol conversion from xylose was over the theoretical yield reflecting that pyruvate was also generated from the catabolism of other complex nutrients [2].

Table 4 Volatile acids co-produced during pure or crude glycerol fermentation by EPI300 (X, pLOI297) or EPI300 (G1, pLOI297) strains

Strain	Volatile acid concentration (mg/mL)[a]			
	Epi (X, pLOI297)		Epi (G1, pLOI297)	
Glycerin	Pure	Crude	Pure	Crude
Acetic	2.572	0.072	4.627	0.110
Propionic	3.158	0.172	3.688	0.105
Succinic	1.657	ND	1.891	ND

[a] Concentrations of acids detected on supernatants of fermentation assays of 50 % of corresponding carbon source

[b] *ND* not detected

Since it has been shown that only the synthesis of ethanol is required when using glycerol [12, 32], to obtain a better yield it is viable to eliminate the by-products that take carbon away from ethanol.

Ethanol production using LY180 strain

LY180 strain was genetically modified to produce ethanol as the main fermentation product and, at the same time, to tolerate high ethanol and furfural concentrations [2, 31, 56]. Both acetate kinase (*ack*) and fumarate reductase (*frd*) were interrupted to prevent acetate, succinic and propionic production. LY180 converts pyruvate into ethanol mainly by the expression of pyruvate decarboxylase (*pdc*) and alcohol dehydrogenase (*adh*) from *Z. mobilis.* At low oxygen levels, any redox imbalance resulting from the use of the *pdc/adh* pathway is eliminated by respiration [16]. Also, LY180 has a rebuilt pyruvate formate-lyase (*pfl*) pathway that can be used to metabolize glycerol producing ethanol and formate. Formate is later converted into CO_2 and hydrogen by the enzymatic formate hydrogenlyase complex FHL [46].

To pursue better ethanol yields from glycerol by eliminating the undesired by-products, we introduced fosmid G1 and a control fosmid into LY180 strain. We measured ethanol production by LY180 after 24 h under microaerophilic conditions with increasing concentrations of pure glycerol (Fig. 5a). As expected, LY180 led much better ethanol yields than EPI300 (pLOI297). The effect of G1 was also very significant increasing ethanol production

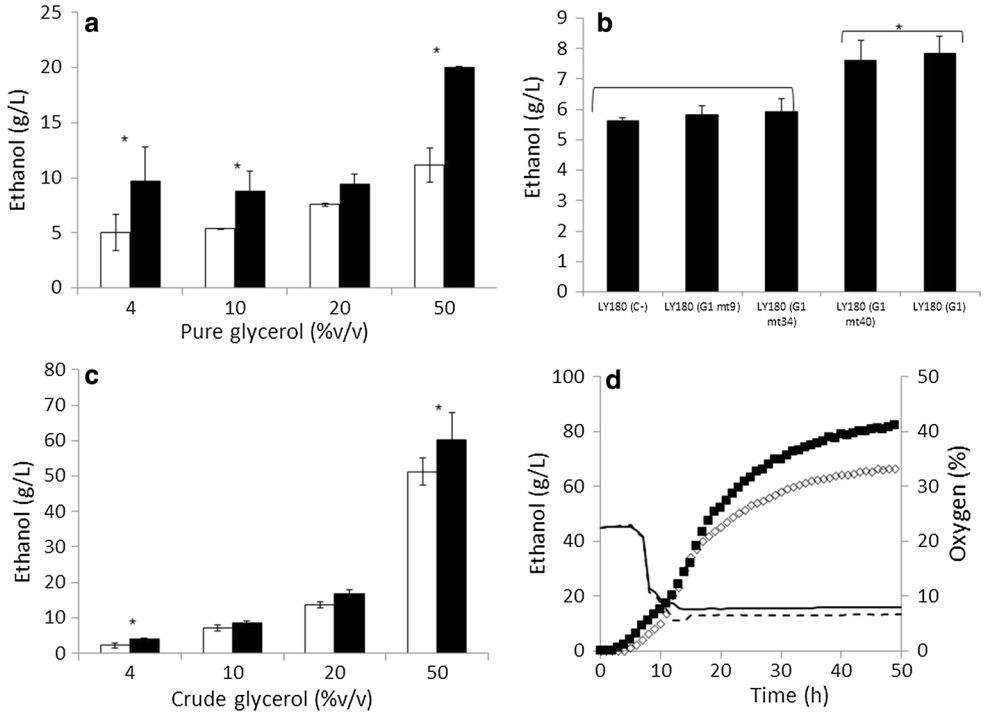

Fig. 5 Ethanol production by LY180 strains in NBS minimum media with pure or crude glycerol as carbon sources. **a** Ethanol production with different concentrations of pure glycerol added to media by LY180 (X) (*empty bars*) or LY180 (G1) (*full bars*); **b** ethanol production from 10 % (v/v) crude glycerol after 24 h incubation at 37 °C, 150 rpm using LY180 (X), LY180 (G1) and selected Tn5 mutants; **c** ethanol production on NBS media supplemented with crude glycerol (% v/v) by LY180 (X) (*empty bars*) or LY180 (G1) (*full bars*); **d** kinetics of oxygen consumption (*lines*) and ethanol production (*symbols*) during 50 h from 50 % (v/v) crude glycerol by LY180 (X) (*dashed line*, *diamond*) or LY180 (G1) (*solid line*, *filled square*). Experiments were done in triplicates, and averages were plotted. *Error bars* indicate standard deviations, *asterisk* is present when statistical significant differences were found between strains (*t* test, 95 % confidence level)

up to 94 % at 4 % glycerol. At 50 % glycerol, 20 g/L of ethanol were obtained after 24 h incubation. To relate ORF G1_23 with the observed improvement in ethanol production, we evaluated the effect of the Tn5-induced mutants in LY180 cells. We measured ethanol concentration after 24 h fermentation on minimal media containing 4 % pure glycerol and LY180 cells harboring the G1 fosmid, its mutants or a randomly picked control fosmid (Fig. 5b). Both mt9 and mt34 performed similar to the negative control, as also seen in EPI300 (pLOI297) cells. However, mt40 behaved similar to fosmid G1 wild type suggesting that the mutation on mt40 was complemented or suppressed by the LY180 host genetic background. Tn5 insertion in mt40 mapped upstream of the coding region of ORF G1_23 which could indicate that it affected the promoter or other regulation elements of the gene. LY180 derives from *E.coli* W [21, 31] which is quite different from K-12, the parental linage of most laboratory cloning strains. It has a larger genome size, encodes more genes, and it grows on a much broader range of carbon substrates [3]. Then, it is plausible to speculate that there are also differences that affect regulation of gene expression which silence the mutation of mt40.

Since we wanted to test the conversion of crude glycerol into ethanol, we performed the fermentations assays with LY180 cells harboring fosmid G1 or a control fosmid, and increasing concentrations of neutralized crude glycerol (Fig. 5c). At concentrations of 10, 20 and 50 % of the carbon source, the yields of ethanol production obtained with LY180 in crude glycerol (Fig. 5c) were higher than those obtained with pure glycerol (Fig. 5a). At 50 % crude glycerol, LY180 cells yielded 51 g/L of ethanol almost five-fold more ethanol than when 50 % pure glycerol was used. These results are in accordance with those obtained with EPI300 cells confirming that neutralized crude glycerol is a very interesting substrate for ethanol production. Similarly to the observations in EPI300, G1 had a positive effect on ethanol production. At 50 % crude glycerol, the presence of fosmid G1 yielded an additional 10 g/L of ethanol (Fig. 5c). As expected, in this case, no by-products such as acetate, succinate, or propionate were produced (data not shown).

To further improve ethanol yields, we increased the initial cell density eight times (from OD_{620} = 1.0 to an OD_{620} = 8.0). After 24 h, final ethanol concentration from

50 % crude glycerol rose from 51 to 67 g/L with LY180 cells carrying the control fosmid and from 60 to 75 g/L for LY180 harboring fosmid G1. This represents a specific ethanol production rate of 0.39 g h^{-1} OD^{-1} L^{-1}.

Ethanol production and oxygen consumption were continuously measured during LY180 fermentation of 50 % crude glycerol (Fig. 5d). After a lag phase of 7 h, oxygen was rapidly consumed and leveled off at 6–8 %. Low oxygen favored the production rate of ethanol and reached 67 and 82 g/L by 50 h of fermentation with LY180 (X) and LY180 (G1), respectively. This ethanol concentration is close to the limit that these cells can tolerate without compromising their integrity [20]. Still, this concentration of ethanol is high enough to make economically feasible its distillation for commercial purposes [27]. Trace amounts of CO_2 and H_2 were detected in the headspace consistently to the usage of the pfl/FHL pathway [32] (data not shown).

The fermentative conversion of glycerol to ethanol has previously been reported by metabolically engineered *E. coli* strains. Durnin et al. [13] reported the production of more than 20 g/L of ethanol accompanied by hydrogen or formate production from 60 g/L of crude glycerol by an *E. coli* strain with dihydroxyacetone kinase (*dhaKLM*) and glycerol dehydrogenase (*gldA*) genes under control of $P_{LtetO-1}$. Similar results were obtained from 50 g/L of crude glycerol after 96 h of fermentation by *E. coli* MG1655 (pZSKLMgldA) strain [7]. An *E. coli* strain carrying a plasmid that overexpressed glycerol uptake genes *dhaKLM*, *gldA*, and *glpK*, as well as the ethanol pathway gene *adhE* produced up to 41.9 g/L of ethanol by 72 h from 110 g/L glycerol [29].

Taken together the results obtained with both *E. coli* strains confirm that neutralized crude glycerol is an excellent carbon source for ethanol production and that the presence of fosmid G1 enhances this production even further. Crude glycerol contains about 80 % of glycerol. The remaining 20 % is composed of contaminants (Table 1), including micronutrients, fatty acids and proteins, that might enhance cell metabolism and ethanol production [25, 51] explaining the significant larger yield observed.

Conclusions

The crude glycerol resulting from biodiesel production has recently became an abundant low cost feedstock. In this work, we intended to convert it to ethanol by looking for novel genes that can improve microorganisms with genetic backgrounds that have been designed for maximum ethanol production. With the results presented in this work we conclude that neutralized crude glycerol, after just a 50 % dilution, is an excellent carbon source for obtaining distillable concentrations of ethanol after only 24 h of fermentation. We also conclude that the metagenomic fosmid G1 has a positive effect both in cell growth and ethanol production in microaerophilic conditions. One ORF, named G1_23, is the one most likely responsible for these phenotypes.

The positive effect of G1 was more evident with pure glycerol than with crude glycerol, which rules out the possibility that G1 is involved in tolerance mechanisms against contaminants present in the crude glycerol. Then, G1 is either involved in glycerol transport or glycerol metabolism. Although it was not fully characterized, sequence analyses suggested that G1_23 encodes a transmembrane or extracellular protein. This fact favors towards a role of G1 in glycerol transport.

The fermentation of pure and crude glycerol by EPI300 carrying the ethanologenic plasmid pLOI297 and the fosmid G1 achieved final ethanol concentrations of 10.4 and 31.7 g/L of ethanol, respectively. Not surprisingly, even better yields were obtained when the EPI300 strain was replaced with LY180. This strain not only overexpresses the *Z. mobilis* homoethanol pathway but also has deletions that eliminate by-products. When strain LY180 was transformed with the G1 fosmid ethanol production increased over 20 % from 50 % crude glycerol. LY180 carrying the G1 fosmid was able to produce up to 75 g/L of ethanol after 24 h fermentation with an initial OD_{620} of 8.0. This represents a specific ethanol production rate of 0.39 g h^{-1} OD^{-1} L^{-1} which is suitable for the commercial production of this biofuel.

Acknowledgments We are very grateful to Dr. Lonnie Ingram (University of Florida) for providing pLOI297 plasmid and LY180 strain, as well as to Dr. Liliana Borzacconi (Departamento de Ingeniería de Reactores, Facultad de Ingeniería, Universidad de la República) for providing purge sludges. Special thanks to Nikolai Guchin (ANCAP, Uruguay) and Darío Rodríguez (ALUR, Uruguay) for providing crude glycerol. This work was supported by Grants #FSE_2011_6383 from Agencia Nacional de Investigación e Innovación (ANII, Uruguay), by the Programa de Desarrollo de las Ciencias Básicas (PEDECIBA, Uruguay) and by Sistema Nacional de Investigadores (SNI, Uruguay).

References

1. Aakvik T, Lale R, Liles M, Valla S (2011) Metagenomic libraries for functional screening. In: de Bruijn F (ed) Handbook of molecular microbial ecology I, 1st edn. Wiley-Blackwell, Hoboken, New Jersey, pp 171–181. doi:10.1002/9781118010518.ch22
2. Alterthum F, Ingram LO (1989) Efficient ethanol production from glucose, lactose, and xylose by recombinant *Escherichia coli*. Appl Environ Microbiol 55:1943–1948
3. Archer CT, Kim JF, Jeong H, Park JH, Vickers CE, Lee SY, Nielsen LK (2011) The genome sequence of *E. coli* W (ATCC 9637): comparative genome analysis and an improved genome-scale reconstruction of *E. coli*. BMC Genom 12:1–20. doi:10.1186/1471-2164-12-9
4. Beloqui A (2006) Novel polyphenol oxidase mined from a metagenome expression library of bovine rumen: biochemical properties, structural analysis, and phylogenetic relationships. J Biol Chem 281:22933–22942. doi:10.1074/jbc.M600577200

5. Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: signalP 3.0. J Mol Biol 340:783–795. doi:10.1016/j.jmb.2004.05.028
6. Bondioli P, Della Bella L (2005) An alternative spectrophotometric method for the determination of free glycerol in biodiesel. Eur J Lipid Sci Technol 107:153–157
7. Cintolesi A, Clomburg JM, Rigou V, Zygourakis K, Gonzalez R (2012) Quantitative analysis of the fermentative metabolism of glycerol in *Escherichia coli*. Biotechnol Bioeng 109:187–198. doi:10.1002/bit.23309
8. Ciriminna R, Pina CD, Rossi M, Pagliaro M (2014) Understanding the glycerol market. Eur J Lipid Sci Technol 116:1432–1439
9. Clomburg JM, Gonzalez R (2013) Anaerobic fermentation of glycerol: a platform for renewable fuels and chemicals. Trends Biotechnol 31:20–28. doi:10.1016/j.tibtech.2012.10.006
10. Cofré O, Ramírez M, Gómez JM, Cantero D (2012) Optimization of culture media for ethanol production from glycerol by *Escherichia coli*. Biomass Bioenergy 37:275–281. doi:10.1016/j.biombioe.2011.12.002
11. Chaudhary N, Ngadi MO, Simpson BK, Kassama LS (2011) Biosynthesis of ethanol and hydrogen by glycerol fermentation using *Escherichia coli*. Adv Chem Eng Sci 1:83–89. doi:10.4236/aces.2011.13014
12. 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. doi:10.1002/bit.21025
13. Durnin G, Clomburg J, Yeates Z, Alvarez PJ, Zygourakis K, Campbell P, Gonzalez R (2009) Understanding and harnessing the microaerobic metabolism of glycerol in *Escherichia coli*. Biotechnol Bioeng 103:148–161. doi:10.1002/bit.22246
14. Ferrer M, Golyshina OV, Chernikova TN, Khachane AN, Reyes-Duarte D, Santos VA, Strompl C, Elborough K, Jarvis G, Neef A, Yakimov MM, Timmis KN, Golyshin PN (2005) Novel hydrolase diversity retrieved from a metagenome library of bovine rumen microflora. Environ Microbiol 7:1996–2010. doi:10.1111/j.1462-2920.2005.00920.x
15. Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. Nucl Acids Res 39(suppl 2):W29–W37. doi:10.1093/nar/gkr367
16. Geddes R, Shanmugam KT, Ingram LO (2015) Combining treatments to improve the fermentation of sugarcane bagasse hydrolysates by ethanologenic *Escherichia coli* LY180. Bioresour Technol 189:15–22. doi:10.1016/j.biortech.2015.03.141
17. Gonzalez R, Murarka A, Dharmadi Y, Yazdani SS (2008) A new model for the anaerobic fermentation of glycerol in enteric bacteria: trunk and auxiliary pathways in *Escherichia coli*. Metab Eng 10:234–245. doi:10.1016/j.ymben.2008.05.001
18. Handelsman J (2004) Metagenomics: application of genomics to uncultured microorganisms. Microbiol Mol Biol Rev 68:669–685. doi:10.1128/mmbr.68.4.669-685.2004
19. Hu H, Wood TK (2010) An evolved *Escherichia coli* strain for producing hydrogen and ethanol from glycerol. Biochem Biophys Res Commun 391:1033–1038. doi:10.1016/j.bbrc.2009.12.013
20. Jarboe L, Grabar T, Yomano L, Shanmugan K, Ingram L (2007) Development of ethanologenic bacteria. Adv Biochem Engin Biotechnol 108:237–261. doi:10.1007/10_2007_068
21. Jarboe LR, Grabar TB, Yomano LP, Shanmugan KT, Ingram LO (2007) Development of ethanologenic bacteria. In: Olsson L (ed) Biofuels. Springer, Berlin, Heidelberg, pp 237–261. doi:10.1007/10_2007_068
22. Johnson DT, Taconi KA (2007) The glycerin glut: options for the value-added conversion of crude glycerol resulting from biodiesel production. Environ Prog 26:338–348. doi:10.1002/ep.10225
23. Krogh A, Larsson B, Von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567–580
24. Lai E (2014) Biodiesel: environmental friendly alternative to petrodiesel. J Pet Environ Biotechnol 5:e122
25. Lee SJ, Kim SB, Kang SW, Han SO, Park C, Kim SW (2012) Effect of crude glycerol-derived inhibitors on ethanol production by Enterobacter aerogenes. Bioprocess Biosyst Eng 35:85–92. doi:10.1007/s00449-011-0607-y
26. Loaces I, Amarelle V, Muñoz-Gutierrez I, Fabiano E, Martinez A, Noya F (2015) Improved ethanol production from biomass by a rumen metagenomic DNA fragment expressed in *Escherichia coli* MS04 during fermentation. Appl Microbiol Biotechnol 99:9049–9060. doi:10.1007/s00253-015-6801-0
27. Macrelli S, Mogensen J, Zacchi G (2012) Techno-economic evaluation of 2nd generation bioethanol production from sugar cane bagasse and leaves integrated with the sugar-based ethanol process. Biotechnol Biofuels 5:1–18. doi:10.1186/1754-6834-5-22
28. Martínez-Gómez K, Flores N, Castañeda HM, Martínez-Batallar G, Hernández-Chávez G, Ramírez OT, Gosset G, Encarnación S, Bolivar F (2012) New insights into *Escherichia coli* metabolism: carbon scavenging, acetate metabolism and carbon recycling responses during growth on glycerol. Microb Cell Fact 11:46
29. Matthew S, Wong ML, Black Ryan W, Le Thao Q, Puthli Sharon, Campbell Paul, Monticello Daniel J (2014) Microaerobic conversion of glycerol to ethanol in *Escherichia coli*. Appl Environ Microbiol 80:3276–3282
30. Meiswinkel TM, Rittmann D, Lindner SN, Wendisch VF (2013) Crude glycerol-based production of amino acids and putrescine by Corynebacterium glutamicum. Bioresour Technol 145:254–258. doi:10.1016/j.biortech.2013.02.053
31. Miller EN, Jarboe LR, Yomano LP, York SW, Shanmugam KT, Ingram LO (2009) Silencing of NADPH-dependent oxidoreductase genes (yqhD and dkgA) in furfural-resistant ethanologenic *Escherichia coli*. Appl Environ Microbiol 75:4315–4323. doi:10.1128/AEM.00567-09
32. 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 74:1124–1135. doi:10.1128/AEM.02192-07
33. Nancy YY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ (2010) PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics 26:1608–1615
34. Nguyen ADQ, Kim YG, Kim SB, Kim CJ (2013) Improved tolerance of recombinant *Escherichia coli* to the toxicity of crude glycerol by overexpressing trehalose biosynthetic genes (otsBA) for the production of β-carotene. Bioresour Technol 143:531–537. doi:10.1016/j.biortech.2013.06.034
35. Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10:1–6
36. Nikel PI, Giordano AM, de Almeida A, Godoy MS, Pettinari MJ (2010) Elimination of d-Lactate synthesis increases poly(3-Hydroxybutyrate) and ethanol synthesis from glycerol and affects cofactor distribution in recombinant *Escherichia coli*. Appl Environ Microbiol 76:7400–7406. doi:10.1128/aem.02067-10
37. Nikel PI, Ramirez MC, Pettinari MJ, Méndez BS, Galvagno MA (2010) Ethanol synthesis from glycerol by *Escherichia coli* redox mutants expressing adhE from Leuconostoc mesenteroides. J Appl Microbiol 109:492–504. doi:10.1111/j.1365-2672.2010.04668.x
38. Noguchi H, Park J, Takagi T (2006) MetaGene: prokaryotic gene finding from environmental genome shotgun sequences. Nucleic Acids Res 34:5623–5630. doi:10.1093/nar/gkl723
39. Papanikolaou S, Fick M, Aggelis G (2004) The effect of raw glycerol concentration on the production of 1,3-propanediol by Clostridium butyricum. J Chem Technol Biotechnol 79:1189–1196. doi:10.1002/jctb.1103

40. Patil KR, Roune L, McHardy AC (2012) The phylopythias web server for taxonomic assignment of metagenome sequences. PLoS One 7:e38581
41. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, Finn RD (2012) The Pfam protein families database. Nucleic Acids Res 40:D290–D301. doi:10.1093/nar/gkr1065
42. Purushe J, Fouts DE, Morrison M, White BA, Mackie RI, Coutinho PM, Henrissat B, Nelson KE (2010) Comparative genome analysis of Prevotella ruminicola and Prevotella bryantii: insights into their environmental niche. Microb Ecol 60:721–729. doi:10.1007/s00248-010-9692-8
43. Rodríguez MC, Loaces I, Amarelle V, Senatore D, Iriarte A, Fabiano E, Noya F (2015) Est10: a novel alkaline esterase isolated from bovine rumen belonging to the new family XV of lipolytic enzymes. PLoS One 10:e0126651. doi:10.1371/journal.pone.0126651
44. Santero E, Floriano B, Govantes F (2016) Harnessing the power of microbial metabolism. Curr Opin Microbiol 31:63–69. doi:10.1016/j.mib.2016.03.003
45. Sarma SJ, Brar SK, Sydney EB, Le Bihan Y, Buelna G, Soccol CR (2012) Microbial hydrogen production by bioconversion of crude glycerol: a review. Int J Hydrog Energy 37:6473–6490. doi:10.1016/j.ijhydene.2012.01.050
46. Sawers R, Blokesch M, Böck A (2004) Anaerobic formate and hydrogen metabolism. EcoSal Plus 2004. doi:10.1128/ecosalplus.3.5.4
47. Senthilkumar V, Gunasekaran P (2005) Bioethanol production from cellulosic substrates: engineered bacteria and process integration challenges. J Sci Ind Res 64:845–853
48. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. doi:10.1093/molbev/mst197
49. Trinh CT, Srienc F (2009) Metabolic engineering of *Escherichia coli* for efficient conversion of glycerol to ethanol. Appl Environ Microbiol 75:6696–6705. doi:10.1128/aem.00670-09
50. Valle A, Cabrera G, Cantero D, Bolivar J (2015) Identification of enhanced hydrogen and ethanol *Escherichia coli* producer strains in a glycerol-based medium by screening in single-knock out mutant collections. Microb Cell Factories 14:93
51. Valle A, Cabrera G, Muhamadali H, Trivedi DK, Ratray NJW, Goodacre R, Cantero D, Bolivar J (2015) A systematic analysis of TCA *Escherichia coli* mutants reveals suitable genetic backgrounds for enhanced hydrogen and ethanol production using glycerol as main carbon source. Biotechnol J 10:1750–1761. doi:10.1002/biot.201500005
52. Wu S, Zhu Z, Fu L, Niu B, Li W (2011) WebMGA: a customizable web server for fast metagenomic sequence analysis. BMC Genom 12:444
53. Yang F, Hanna MA, Sun R (2012) Value-added uses for crude glycerol–a byproduct of biodiesel production. Biotechnol Biofuels 5:13. doi:10.1186/1754-6834-5-13
54. Yazdani S, Gonzalez R (2008) Engineering *Escherichia coli* for the efficient conversion of glycerol to ethanol and co-products. Metabolic Eng 10:340–351. doi:10.1016/j.ymben.2008.08.005
55. Yazdani SS, Gonzalez R (2007) Anaerobic fermentation of glycerol: a path to economic viability for the biofuels industry. Curr Opini Biotechnol 18:213–219. doi:10.1016/j.copbio.2007.05.002
56. Yomano L, York S, Ingram L (1998) Isolation and characterization of ethanol-tolerant mutants of *Escherichia coli* KO11 for fuel ethanol production. J Ind Microbiol Biotechnol 20:132–138
57. Yu KO, Kim SW, Han SO (2010) Engineering of glycerol utilization pathway for ethanol production by *Saccharomyces cerevisiae*. Bioresour Technol 101:4157–4161. doi:10.1016/j.biortech.2010.01.066
58. Zhang X, Shanmugam KT, Ingram LO (2010) Fermentation of glycerol to succinate by metabolically engineered strains of *Escherichia coli*. Appl Environ Microbiol 76:2397–2401. doi:10.1128/aem.02902-09
59. Zhang Z, Schwartz S, Wagner L, Miller W (2000) A greedy algorithm for aligning DNA sequences. J Comput Biol 7:203–214. doi:10.1089/10665270050081478