Appl Microbiol Biotechnol
DOI 10.1007/s00253-011-3492-z

BIOENERGY AND BIOFUELS

Mutant selection and phenotypic and genetic characterization of ethanol-tolerant strains of *Clostridium thermocellum*

Xiongjun Shao · Babu Raman · Mingjun Zhu · Jonathan R. Mielenz · Steven D. Brown · Adam M. Guss · Lee R. Lynd

Received: 3 June 2011 / Revised: 6 July 2011 / Accepted: 15 July 2011

Abstract *Clostridium thermocellum* is a model microorganism for converting cellulosic biomass into fuels and chemicals via consolidated bioprocessing. One of the challenges for industrial application of this organism is its low ethanol tolerance, typically 1–2% (*w*/*v*) in wild-type strains. In this study, we report the development and characterization of mutant *C. thermocellum* strains that can grow in the presence of high ethanol concentrations. Starting from a single colony, wild-type *C. thermocellum* ATCC 27405 was sub-cultured and adapted for growth in up to 50 g/L ethanol using either cellobiose or crystalline cellulose as the growth substrate. Both the adapted strains retained their ability to grow on either substrate and displayed a higher growth rate and biomass yield than the wild-type strain in the absence of ethanol. With added ethanol in the media, the mutant strains displayed an inverse correlation between ethanol concentration and growth rate or biomass yield. Genome sequencing revealed six common mutations in the two ethanol-tolerant strains including an alcohol dehydrogenase gene and genes involved in arginine/pyrimidine biosynthetic pathway. The potential role of these mutations in ethanol tolerance phenotype is discussed.

Keywords *Clostridium thermocellum* · Ethanol tolerance · Genome sequencing · Strain adaptation · Mutations

Introduction

Plant biomass is the only foreseeable sustainable source of organic fuels, chemicals, and materials available to humanity (Lynd et al. 1999). Cellulosic biomass is particularly attractive

Electronic supplementary material The online version of this article (doi:10.1007/s00253-011-3492-z) contains supplementary material, which is available to authorized users.

X. Shao · L. R. Lynd (✉)
Thayer School of Engineering, Dartmouth College,
8000 Cummings Hall,
Hanover, NH 03755, USA
e-mail: lee.lynd@dartmouth.edu

B. Raman · J. R. Mielenz · S. D. Brown · A. M. Guss
Biosciences Division, Oak Ridge National Laboratory,
One Bethel Valley Road,
Oak Ridge, TN 37831, USA

X. Shao · B. Raman · J. R. Mielenz · S. D. Brown · A. M. Guss · L. R. Lynd
BioEnergy Science Center (BESC),
Oak Ridge National Laboratory,
One Bethel Valley Road,
Oak Ridge, TN 37831, USA

M. Zhu
School of Bioscience & Bioengineering,
South China University of Technology,
University Town,
Guangzhou 510006, China

L. R. Lynd
Mascoma Corporation,
67 Etna Road,
Lebanon, NH 03766, USA

Present Address:
B. Raman
Bioprocess R&D, Dow AgroSciences,
9330 Zionsville Road,
Indianapolis, IN 46268, USA

Published online: 27 August 2011

in this context because of its widespread availability and low cost (Lynd et al. 2003). Various biological processing technologies have been proposed for converting cellulosic biomass into biofuel such as ethanol (Lynd et al. 2002). By combining the cellulase production, biomass hydrolysis, and sugar fermentation into one step, consolidated bioprocessing (CBP) offers a great potential for cost reduction by eliminating costly cellulase addition and consolidating capital equipment.

Clostridium thermocellum is a candidate microorganism for CBP since it can rapidly hydrolyze cellulosic material, with the aid of a complexed cellulase system termed the cellulosome, and ferment the hydrolysis products to ethanol and organic acids. Genetic tools have been recently developed and successfully applied to eliminate lactic acid and acetic acid production in *C. thermocellum* (Ladisch et al. 2009; Tripathi et al. 2010) in attempts to increase the ethanol yield. Commercial ethanol production requires ethanol titers higher than 40 g/L for economical product recovery and hence the need to develop strains that have better ethanol tolerance phenotype. In *C. thermocellum*, as well as other thermophilic bacteria (Lovitt et al. 1988), ethanol tolerance is an inducible rather than constitutive phenotype. Thus, while wild-type (WT) strains can only tolerate up to 2% (*w*/*v*) ethanol, sequential transfer in increasing ethanol concentration, with or without mutagenesis, has resulted in strains that are able to withstand as high as 50–55 g/L ethanol [(Williams et al. 2007; Timmons et al. 2009; Wang et al. 1983; Sudha Rani and Seenayya 1999); see Supplemental Table S1 for a summary of previous studies].

The most common phenotype observed for ethanol tolerant strains is higher ethanol yield than the wild type and reduced optimal growth temperature (Herrero and Gomez 1980). Ethanol-tolerant isolates of *C. thermocellum* were also found to better tolerate other organic solvents, suggesting perhaps a similar effect of solvents on cell membrane and a general cellular response mechanism (Sudha Rani and Seenayya 1999). Studies on ethanol tolerant mutants showed alterations in membrane fatty acid composition, increase in membrane rigidity for counteracting the fluidizing effect of ethanol (Herrero et al. 1982; Timmons et al. 2009), and profound changes in membrane protein profile (Williams et al. 2007). Recently, genome sequencing of an ethanol-tolerant strain adapted on cellobiose revealed more than 400 mutations, as compared to the WT strain (Brown et al. 2010).

To date, there remains a discrepancy between the tolerance of *C. thermocellum* to exogenously added ethanol and the maximum levels of ethanol produced by this organism. While strains tolerant to exogenously added ethanol have been isolated, a concomitant increase in ethanol titer has not occurred in these strains. This discrepancy has been attributed to inhibition of cellular metabolism by other metabolic by-products such as salts of organic acids for *Thermoanaerobacter thermosaccharolyticum* (Lynd et al. 2001) but, in general, is not well understood. Motivated by a desire to assist future systems biology studies and genetic engineering aimed at understanding the mechanisms of increased ethanol tolerance in *C. thermocellum*, we report here a well-documented selection approach as well as characterization of resultant strains at the level of both growth studies and genome sequencing.

Materials and methods

Strain and culture conditions

C. thermocellum ATCC 27405 was obtained from the American Type Culture Collection (Manassas, VA, USA). A single colony was isolated and denoted as WT. Chemically defined media for thermophilic clostridia (MTC) medium was prepared according to the concentrations listed in Supplemental Table S2. All chemicals were reagent grade and obtained from Sigma (St. Louis, MO, USA), unless indicated otherwise. Solution A contained either Avicel PH105 (FMC Biopolymer, Philadelphia, PA, USA) or cellobiose supplemented with appropriate amounts of DI water (Milli-Q). Solutions B, C, D, E, and F were injected aseptically into solution A using a syringe. Prior to combining all the solutions, they were purged with N_2 (Airgas Northeast, White River Junction, VT, USA) and sterilized by autoclaving at 121°C for 45 min except for solution A with cellobiose, which was autoclaved for 25 min.

Adaptation of *C. thermocellum* in ethanol

Adaptation of *C. thermocellum* was performed by duplicate serial transfers in crimp-sealed 25 mL Balch tubes. The tubes were sealed empty and purged with N_2 and sterilized by autoclaving at 121°C. The tubes were then injected with 9 ml MTC media containing Avicel or cellobiose. Ethanol, also purged with N_2, was added to each tube using a 1-mL syringe to have a final concentration of 0–50 g/L with an increment of 5 g/L. Each inoculation/transfer was 10% volume (1 mL). Cultures were grown in an incubator (New Brunswick Scientific, Innova 4080) with temperature controlled at 55°C and rotation speed set at 200 rpm.

Serial transfer to obtain ethanol tolerant mutants involved inoculation into medium with elevated ethanol concentrations alternated with medium without added ethanol. The parameter *R* was defined as the ratio of final OD over initial OD for cellobiose and the ratio of final pellet nitrogen over initial pellet nitrogen for Avicel within 72 h incubation. The criteria for transfer were (a) transfer to

higher ethanol concentration if $R \geq 4$, (b) maintain current ethanol concentration if $2 \leq R < 4$, and (c) transfer to previous ethanol concentration if $R < 2$. Single colonies of ethanol tolerant strains were isolated from the final cultures and denoted as E50A for adaptation using Avicel and E50C for adaptation using cellobiose.

Isolation of single colonies

Agar (Fisher Scientific, Pittsburgh, PA, USA) solution (18.75 g/L) was prepared, and 40 mL was distributed into each of eight 125-mL serum bottles (Wheaton, Millville, NJ, USA). The bottles were crimp-sealed, purged with N_2, and sterilized by autoclaving at 121°C for 25 min. The sterilized bottles were stored in a 60°C oven to prevent solidification of agar. Sterile anaerobic solutions B, C, D, and E of MTC media, pre-heated to 60°C, were injected into the bottles as per the medium recipe (Supplemental Table S2). A mixture consisting of yeast extract, cellobiose, and MOPS was purged with N_2, filter sterilized, and injected into the agar-containing bottles, giving a final concentrations of 5 g/L yeast extract, 2 g/L cellobiose, and 10 g/L MOPS. The agar-containing bottles were then transferred into an anaerobic chamber (Coy Laboratory Products, Grass Lake, MI, USA). Final adaptation culture (0.5 mL) was inoculated into the first agar-containing bottle followed by serial transfers into the other bottles as follows: 0.5, 0.5, 5, 5, 5, 5, and 5 mL. The contents of each of the last five bottles were poured into two Petri dishes (BD Biosciences, Bedford, MA, USA). The dishes were allowed to sit for 30 min to solidify the agar and then incubated at 55°C. Colonies were picked using a needle after 32–48 h incubation. A picked colony was transferred into a microcentrifuge tube with 1 mL sterilized DI water, which was mixed and injected into a crimp-sealed 125-mL serum bottle with 50 ml Avicel or cellobiose MTC media. The bottle was incubated at 55°C and 200 rpm. After about 24 h, stock culture was prepared with 33% glycerol and stored at −80°C.

Characterization of the ethanol tolerant strains

To determine growth rate without added ethanol, *C. thermocellum* strains WT, E50A, and E50C were cultured in MTC media with Avicel or cellobiose. Crimp-sealed 125 mL serum vials with 35 mL DI water and 0.25 g Avicel or cellobiose were purged with N_2 and sterilized by autoclaving at 121°C. After autoclaving, sterile and anaerobic solution B, C, D, E, and F were injected. The vials were then incubated at 55°C and followed by inoculation of 10% by volume with the inoculum prepared in MTC media with 5 g/L Avicel or cellobiose from stock culture. After inoculation, the bottles were incubated at 55°C and 200 rpm. Samples were taken at various times for analysis of pellet nitrogen and product concentrations.

To characterize ethanol tolerance of the two adapted strains of *C. thermocellum*, E50A and E50C were cultured in Avicel or cellobiose MTC media with various ethanol concentrations. The preparation of media and inoculum was the same as the above except that ethanol, purged with N_2 and supplemented with 2% volume solution D, were injected into each bottle 5 h after the inoculation. Samples were taken using syringe at various time points after inoculation.

Analytical methods

The optical density (OD) of cultures grown in 25-mL Balch tubes (Bellco Glass, Vineland, NJ, USA) was measured directly (without sampling) using a Thermo Spectronic Genesys 10VIS spectrophotometer (Rochester, NY, USA) at 600 nm. Pellet nitrogen, used as a proxy for cell growth for insoluble substrate, was measured using a Shimadzu TOC/TON analyzer equipped with an automatic sampler. Pellet samples were collected by centrifugation of 1 mL sample at 21,130×*g* for 5 min, followed by three washes that involved resuspension of the pellet in 1 mL deionized water, centrifugation as above, and removal of the supernatant. The washed pellet samples were either analyzed directly or stored at −20°C until analysis. Fermentation product concentrations were obtained using a Waters HPLC system with an Aminex HPX-87H column operated at 60°C.

Genome sequencing and analysis

Genomic DNA for E50A and E50C was extracted using Genomic-tip 500/G (Qiagen, Valencia, CA, USA). The DNA samples were shipped on dry ice to the DOE Joint Genome Institute (JGI, Walnut Creek, CA, USA), and samples were sequenced using JGI's whole-genome shotgun sequencing method to produce a high-quality draft sequence. Sequencing was initiated with creation of 3-, 8-, and 40-kb DNA libraries, performed from both sides of the library insert, producing paired ends typically resulting in approximately 8–9× depth. Sequenced reads were aligned using MAQ (Li et al. 2008). A report with single nucleotide polymorphisms (SNPs) and statistical analysis was returned. The genes were annotated according to the *C. thermocellum* ATCC 27405 analysis file on the website of Oak Ridge National Laboratory (http://genome.ornl.gov/microbial/cthe/).

Results

Adaptation

A culture of *C. thermocellum* originating from a single colony isolate was sequentially transferred in growth medium containing progressively increasing ethanol con-

centrations, with every other culture grown in the absence of ethanol, using either crystalline cellulose (Avicel) or cellobiose as the substrate (Fig. 1). Optical density (OD) was used for growth measurement for the adaptation using soluble cellobiose, while cell pellet nitrogen was used for insoluble Avicel. The wild-type strain has an ethanol tolerance of ~15 g/L. To attain tolerance to 50 g/L ethanol, adaptation using cellobiose took 40 transfers, while adaptation using Avicel took 48 transfers. However, to reach 45 g/L ethanol, it took only 30 transfers for adaptation using Avicel as compared to 36 transfers for adaptation using cellobiose. For these adapted cultures, single pure cultures were isolated for further characterization: E50C isolated from cellobiose-grown cultures and E50A isolated from Avicel-grown cultures.

Effect of adaptation

Strain E50C, although adapted to tolerate high ethanol concentrations exclusively with cellobiose as the growth substrate, retained its capability to solubilize and grow on crystalline cellulose. A comparison of growth for the wild-type and selected strains using either Avicel or cellobiose is given in Fig. 2. Interestingly, both E50A and E50C strains grow faster than the wild-type strain when no ethanol is added. Among the three strains, strain E50C has the fastest growth rate using either cellobiose or Avicel, while the wild-type strain has the slowest growth rate (Table 1).

Product profiles for the three strains during the course of growth on cellobiose and Avicel are shown in Fig. 3. Acetic acid is the major product for all three strains on both

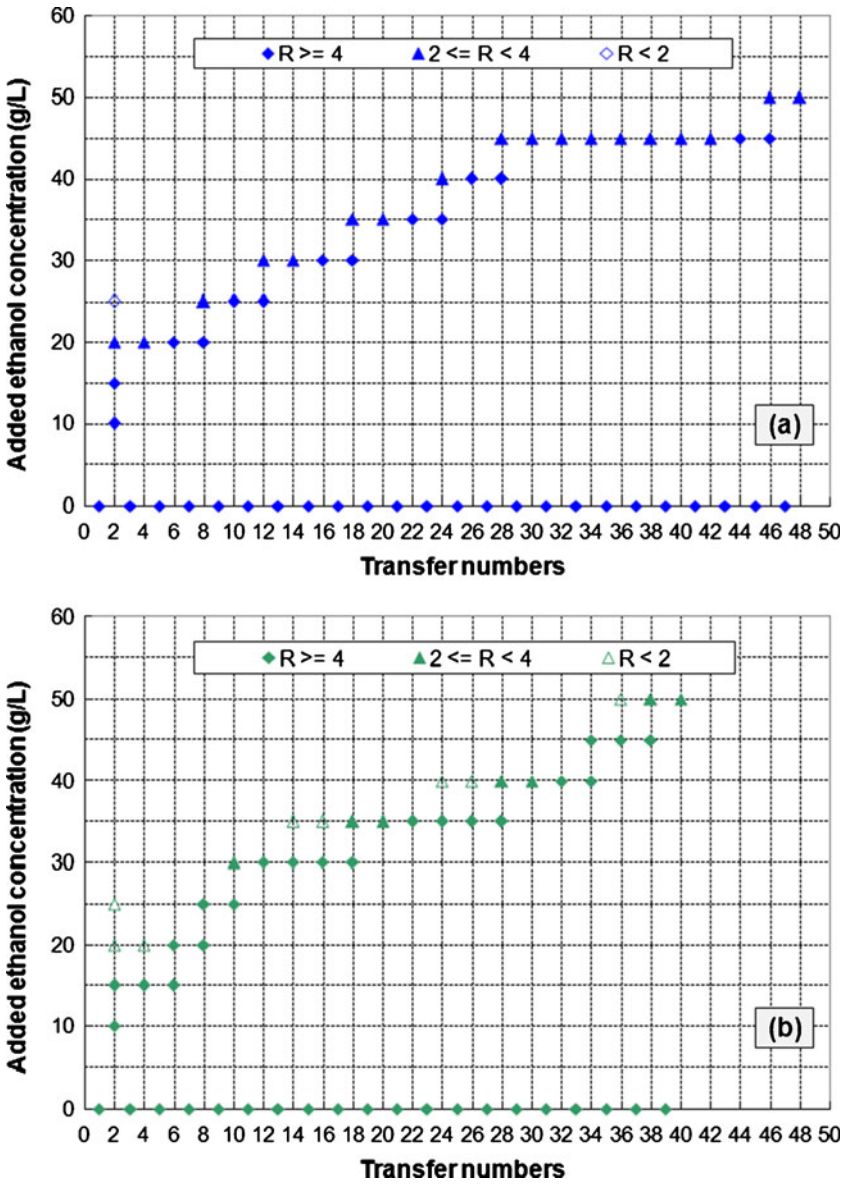

Fig. 1 Adaptation map for higher ethanol tolerance **a** on Avicel, *R* is the ratio of final pellet nitrogen over initial pellet nitrogen within 72 h, **b** on cellobiose, *R* is the ratio of final OD over initial OD with 72 h

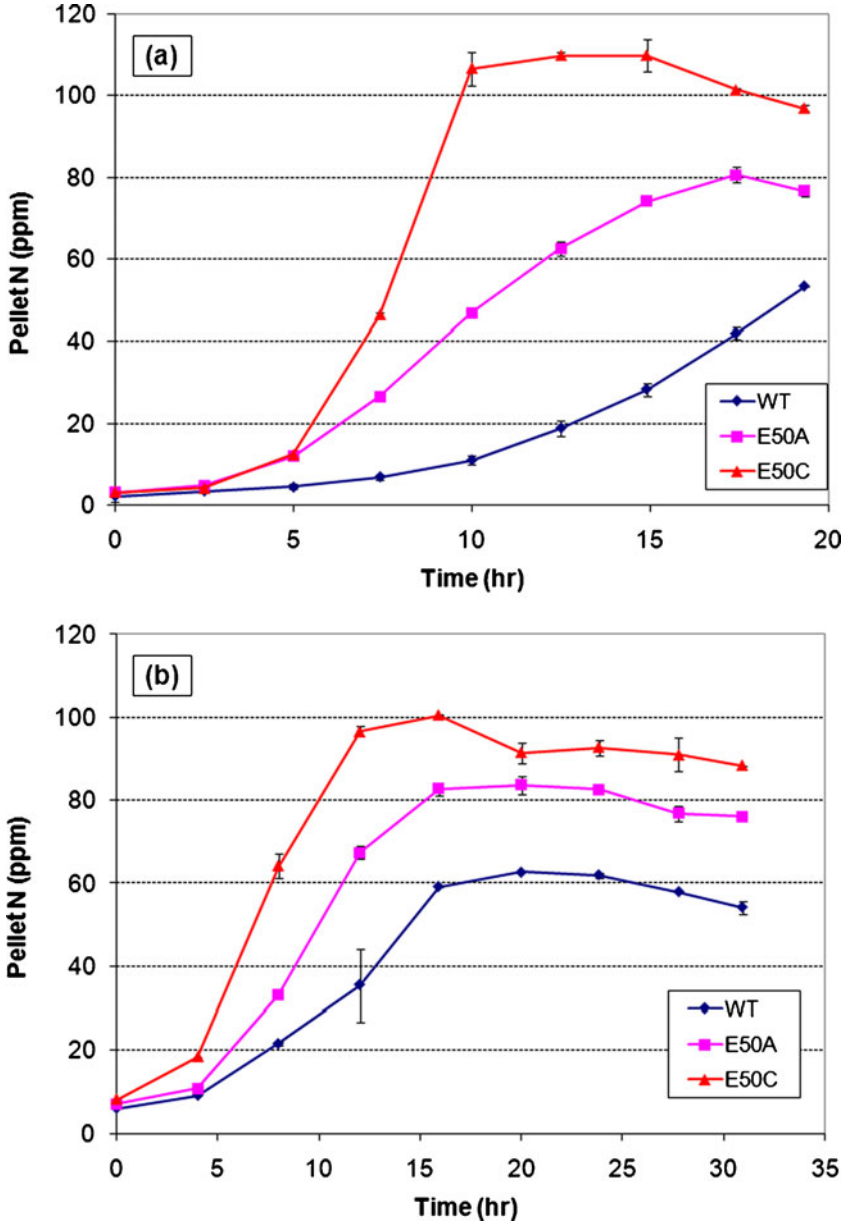

Fig. 2 Comparison of growth (*pellet nitrogen curve*) among E50A, E50C, and wild type (WT) **a** using cellobiose and **b** using Avicel

substrates, followed by ethanol and lactic acid, with E50A producing much more lactic acid than either E50C or WT. Cellobiose was consumed in 10 h for E50C and in 17.5 h for E50A, while there was still about 1.5 g/L cellobiose left after 20 h for the WT strain (Fig. 3). Among the three strains, E50C has the lowest mass ratio of organic acids (acetate plus lactate) to ethanol (Supplemental Table S3). During growth on cellobiose, significant glucose accumulation is observed for E50C and to a lesser extent for E50A, but glucose is consumed when cellobiose is exhausted in both cultures. During growth on Avicel, although substrate consumption was not readily followable, E50C completed product formation most quickly among the three strains and also had the lowest ratio of organic acids to ethanol in the fermentation broth. The wild type had a final product titer of about 2.4 g/L compared to 2.1 and 2.0 g/L for E50A and E50C respectively.

Table 1 Specific growth rates for E50A, E50C, and wild-type strains using cellobiose or Avicel as substrate

	Specific growth rate (h^{-1})	
	Cellobiose	Avicel[a]
Wild-type	0.141±0.017	0.158±0.053
E50A	0.224±0.004*	0.228±0.090
E50C	0.334±0.010	0.280±0.031

[a] Used only early exponential phase data from Figs. 2 and 4

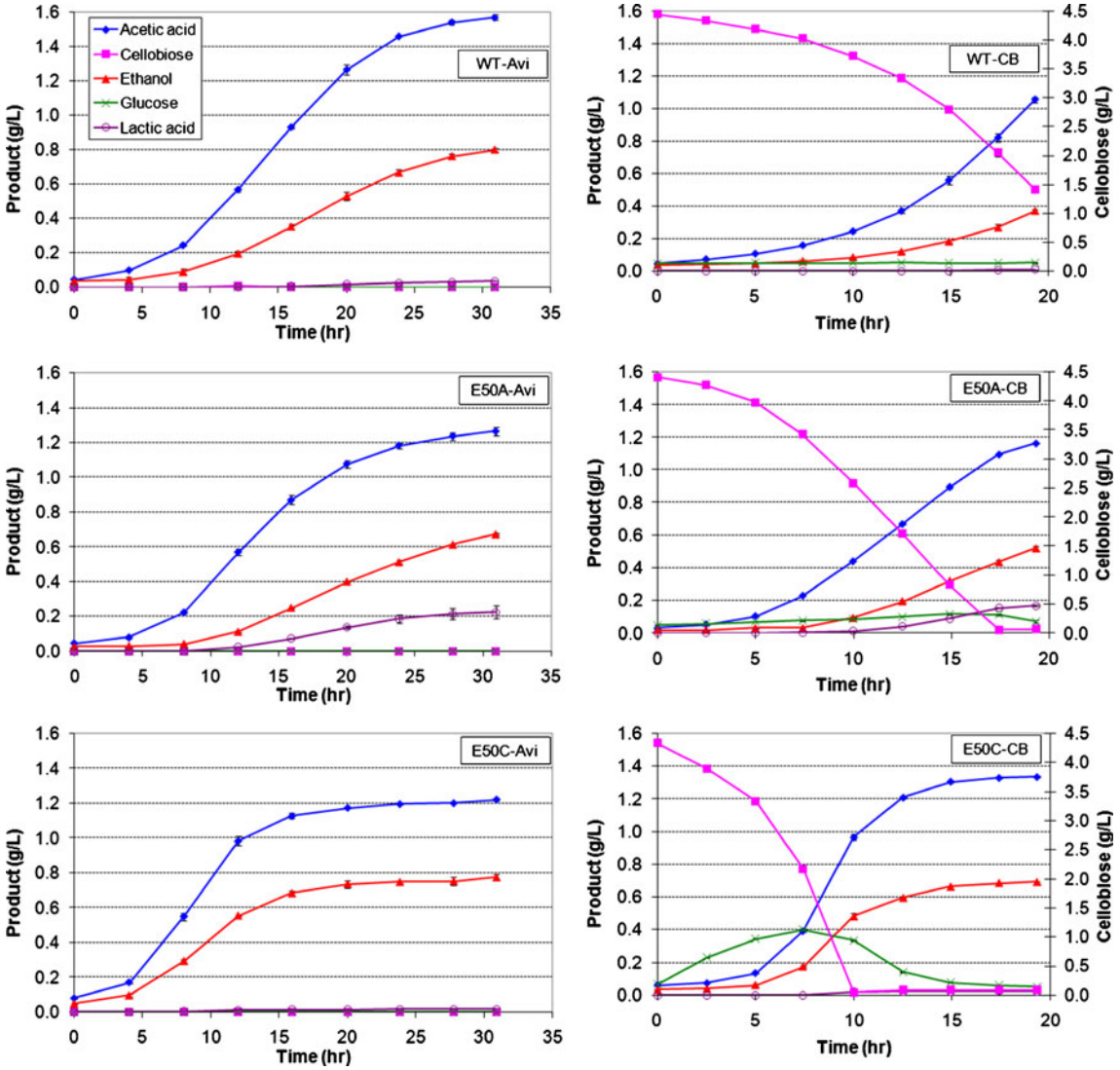

Fig. 3 Comparison of products among wild type (WT), E50A, and E50C using cellobiose (CB) and Avicel (Avi)

Effect of ethanol on growth

The ethanol tolerant strains were evaluated for growth characteristics in response to various ethanol concentrations up to 50 g/L (Fig. 4). Similar to growth patterns observed in the absence of ethanol (Fig. 2), E50C grows more rapidly on both cellobiose and Avicel, as compared to E50A, under all conditions tested. We observed that ethanol concentrations higher than 30 g/L caused significant growth inhibition for both tolerant strains, and this effect is more pronounced during growth on Avicel. This indicates that both strains reach a critical sensitivity threshold around 30–40 g/L ethanol, more clearly seen when examining the maximum cell concentration (as measured by total nitrogen content in the cell pellet) and growth rate data (Fig. 4). For both E50A and E50C, during growth on either Avicel or cellobiose, the maximum cell concentration decreases with increasing ethanol concentration, with the largest change occurring between 30 and 40 g/L added ethanol.

Genetic changes in the ethanol tolerant strains

In an effort to unravel the genetic changes associated with the ethanol tolerance phenotype in *C. thermocellum*, the genomes of the mutant strains were sequenced to identify SNPs or other alterations in their genomic sequences. Genome sequencing revealed 10 and 39 nonsynonymous SNPs in the E50A and the E50C strain, respectively (Table 2). In addition, there were six synonymous SNPs in E50C; non-coding regions in E50A and E50C strains contained five and seven SNPs, respectively (Supplemental Table S4). Six mutated genes were shared by both strains

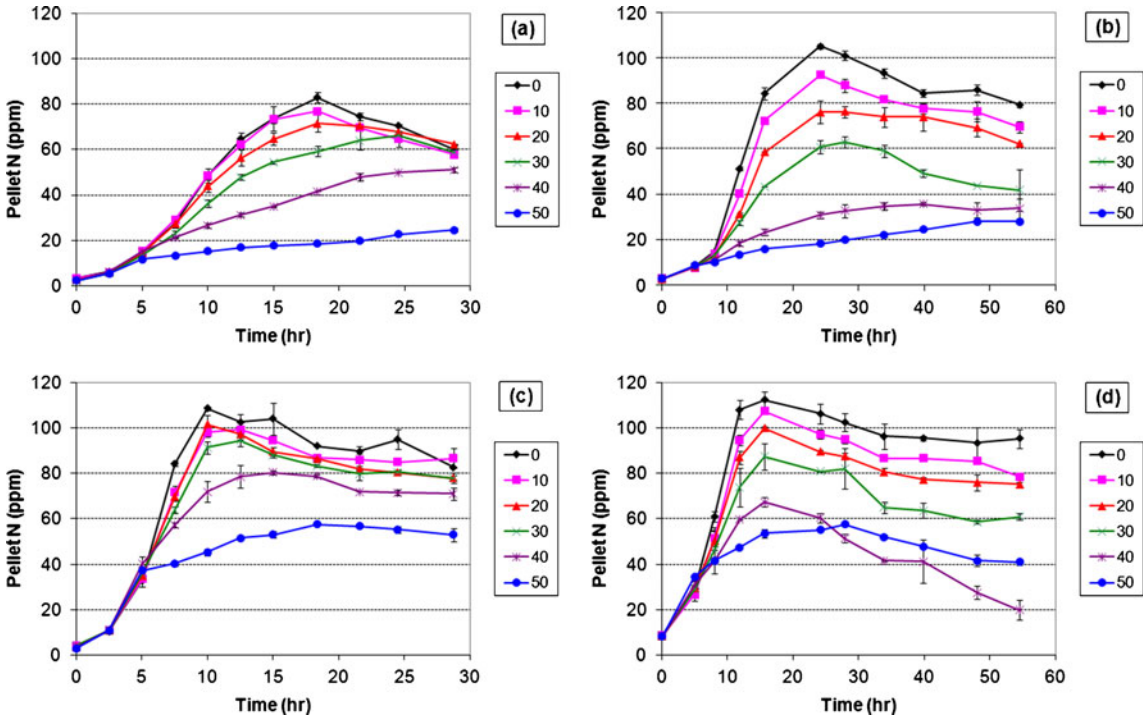

Fig. 4 Growth of selected strains in various ethanol concentrations (0–50 g/L): **a** E50A on cellobiose, **b** E50A on Avicel, **c** E50C on cellobiose, and **d** E50C on Avicel. Ethanol, purged with N_2 and supplemented with 2% volume solution D, were injected 5 h after the inoculation

including four with identical genetic changes. Specifically genes Cthe0390 (putative glucokinase), Cthe1866 (*argD*, acetylornithine aminotransferase), Cthe2699 (putative transcriptional regulator), and Cthe2870 (protein of unknown function) had identical changes in both strains. Cthe0423 (*adhE*, bi-functional aldehyde/alcohol dehydrogenase, involved in ethanol production from acetyl-CoA) and Cthe0953 (*pyrB*, aspartate carbamoyltransferase, involved in pyrimidine biosynthesis from carbamoyl phosphate), on the other hand, were independently mutated in these two strains. Since these independent mutations suggest a functional role in ethanol tolerance, homology-based structural modeling [using I-TASSER (Roy et al. 2010)] was used to identify the location of altered amino acid residues within the protein structures to gain insight into the effect of the mutations. The two independent mutations in the AdhE protein in strains E50A and E50C both lie within the nicotinamide co-factor binding site of the ADH domain in the proteins (Supplemental Figure S1). These mutations in the protein active site could have potential direct implications in enzymatic catalysis including, but not limited to, possible alterations in co-factor specificity, catalytic efficiency, and reaction kinetics.

In addition, E50C strain has mutations in another arginine biosynthetic pathway gene (Cthe1868, *carB*), another gene involved in pyrimidine biosynthesis (Cthe1923, *pyrG*), and genes encoding ABC-type transporter proteins involved in polar amino acid (Cthe1456) and sugar (Cthe0392) transport. Mutations were also found in regulatory genes including a family 24 sigma factor (Cthe2992) and a *relA/spoT* homologue (Cthe1344). RelA/SpoT is involved in the regulation of synthesis and degradation of cellular alarmone, ppGpp, which has known functions in starvation stress survival and heat shock response.

Interestingly, one third of the mutated genes (12 genes) in E50C and two of the mutated genes in E50A encode membrane proteins, potentially to alleviate the well known negative effects of ethanol on plasma membrane and associated proteins (Timmons et al. 2009; Ingram 1990; Ding et al. 2009; Jeffries and Jin 2000; Burdette et al. 2002). In agreement with this, two membrane proteins related to phospholipid (Cthe1000) and choline (Cthe1396) biosynthesis were mutated in E50C. The former protein, phosphotidate cytidylyl transferase was found to be twofold less abundant in the ethanol adapted strain in an earlier study (Williams et al. 2007). Other mutated membrane proteins in strain E50C include two anti-sigma factors with CBM (Cthe0316) and GH domains (Cthe1471), proteins that were recently suggested to be involved in carbohydrate sensing and regulation of carbohydrate active enzyme genes in *C. thermocellum* (Kahel-Raifer et al. 2010; Nataf et al. 2010; Bahari et al. 2011). Two cellulosomal genes, Cthe1890 in E50A and Cthe0270 in E50C, latter encoding a glycoside hydrolase family 18 chitinase (*chiA*), were also mutated in the E50C strain.

Table 2 List of non-synonymous single nucleotide polymorphisms in coding sequences of ethanol-tolerant mutant strains, E50A and E50C, of *Clostridium thermocellum*

Loci	Gene	Annotation	SNP Type	Position	Strand	Ref-Nucl	Ref-aa: codon	E50A-aa: codon	E50C-aa: codon
YP_001036523| Cthe_0089	*mreC*	Rod shape-determining protein MreC	CDS, non-syn	123124	+	C	P:CCC		H:CAC
YP_001036702| Cthe_0270	*chiA*	Glycoside hydrolase, family 18	CDS, non-syn	334469	+	C	A:GCC		V:GTC
YP_001036731| Cthe_0300		Allergen V5/Tpx-1 related	CDS, non-syn	376403	−	C	M:ATG		I:ATA
YP_001036747| Cthe_0316		PA14	CDS, non-syn	399731	+	G	G:GGC		V:GTC
YP_001036821| Cthe_0390		**ROK domain containing protein**	CDS, non-syn	484879	+	C	A:GCG	V:GTG	V:GTG
YP_001036823| Cthe_0392		ABC transporter, cellular inner-membrane translocator, BPD_transp_2	CDS, non-syn	488060	+	A	Q:CAG		L:CTG
YP_001036854| Cthe_0423	***adhE***	**Iron-containing alcohol Dehydrogenase**	CDS, non-syn	533007	+	G	G:GGA	R:AGA	
YP_001036854| Cthe_0423	***adhE***	**Iron-containing alcohol dehydrogenase**	CDS, non-syn	532831	+	A	D:GAT		G:GGT
YP_001037054| Cthe_0626		Hypothetical protein	CDS, non-syn	773016	+	T	D:GAT		E:GAA
YP_001037248| Cthe_0820		Hypothetical protein	CDS, non-syn	995331	−	C	G:GGC		D:GAC
YP_001037259| Cthe_0831		Polyprenyl synthetase	CDS, non-syn	1013252	−	G	L:CTT		F:TTT
YP_001037323| Cthe_0896	*dnaG*	DNA primase	CDS, non-syn	1072967	−	G	A:GCG	V:GTG	
YP_001037380| Cthe_0953	***atc/pyrB***	**Aspartate carbamoyltransferase**	CDS, non-syn	1142635	−	G	P:CCG	T:ACG	
YP_001037380| Cthe_0953	***atc/pyrB***	**Aspartate carbamoyltransferase**	CDS, non-syn	1142613	−	A	V:GTT		G:GGT
YP_001037421| Cthe_0996	*polC*	DNA polymerase III, alpha subunit	CDS, non-syn	1191220	−	G	A:GCC		V:GTC
YP_001037425| Cthe_1000		Phosphatidate cytidylyltransferase	CDS, non-syn	1198439	−	A	I:ATT		S:AGT
YP_001037769| Cthe_1344	*spoT*	RelA/SpoT family protein	CDS, non-syn	1635804	−	G	A:GCC		G:GGC
YP_001037777| Cthe_1352		UDP-glucose 6-dehydrogenase	CDS, non-syn	1644916	−	C	A:GCC		T:ACC
YP_001037820| Cthe_1396		Phospholipase D/ Transphosphatidylase	CDS, non-syn	1705801	−	A	C:TGT		W:TGG
YP_001037878| Cthe_1456		ABC-type polar amino acid transport system, ATPase subunit	CDS, non-syn	1776014	−	T	T:ACC		P:CCC
YP_001037893| Cthe_1471		Glycoside hydrolase, family 5	CDS, non-syn	1787955	+	A	T:ACT		P:CCT
YP_001037982| Cthe_1563		ABC-type antimicrobial peptide transport system ATPase subunit	CDS, non-syn	1891877	+	G	R:CGT		L:CTT
YP_001038187| Cthe_1773		Peptidase S16, lon-like	CDS, non-syn	2095223	−	A	L:TTG		M:ATG
YP_001038280| Cthe_1866	***argD***	**Acetylornithine and succinylornithine aminotransferases**	CDS, non-syn	2212180	+	A	E:GAA	G:GGA	G:GGA
YP_001038282| Cthe_1868	*cpsL/carB*	Carbamoyl-phosphate synthase, large subunit	CDS, non-syn	2216549	+	G	R:AGG		S:AGT
YP_001038304| Cthe_1890		Cellulosome enzyme, dockerin type I	CDS, non-syn	2245282	−	C	S:AGC	N:AAC	
YP_001038335| Cthe_1923	*pyrG*	CTP synthase	CDS, non-syn	2293547	−	C	A:GCC		T:ACC

Table 2 (continued)

Loci	Gene	Annotation	SNP Type	Position	Strand	Ref-Nucl	Ref-aa: codon	E50A-aa: codon	E50C-aa: codon
YP_001038353| Cthe_1942		Hypothetical protein	CDS, non-syn	2318413	−	C	S:AGC		N:AAC
YP_001038379| Cthe_1968		Cell divisionFtsK/ SpoIIIE	CDS, non-syn	2348981	−	C	D:GAT		N:AAT
YP_001038553| Cthe_2157		Hypothetical protein	CDS, non-syn	2565493	−	G	T:ACG		M:ATG
YP_001038717| Cthe_2322		DNA recombinase	CDS, non-syn	2768722	−	G	P:CCA		T:ACA
YP_001038720| Cthe_2325		Recombinase	CDS, non-syn	2772229	−	C	D:GAT		N:AAT
YP_001038848| Cthe_2453		Conserved hypothetical protein	CDS, non-syn	2927348	−	G	L:CTC		F:TTC
YP_001038869| Cthe_2474		Phage terminase, large subunit, PBSX family	CDS, non-syn	2946446	+	T	A:GCT		A:GCA, A:GCT
YP_001038940| Cthe_2545		Conserved hypothetical protein	CDS, non-syn	3013893	−	G	L:CTT		F:TTT
YP_001039091| Cthe_2699		**PemK-like trancriptional modulator of MazE toxin/MazF**	CDS, non-syn	3185426	+	A	H:CAT	P:CCT	P:CCT
YP_001039117| Cthe_2725		DNA-directed RNA polymerase, beta' subunit	CDS, non-syn	3217494	+	G	A:GCA	T:ACA	
YP_001039156| Cthe_2764		TROVE	CDS, non-syn	3263865	+	C	T:ACT		S:AGT
YP_001039261| Cthe_2870		**Protein of unknown function DUF21**	CDS, non-syn	3390930	+	G	A:GCT	T:ACT	T:ACT
YP_001039338| Cthe_2947	*proS2*	Prolyl-tRNA synthetase	CDS, non-syn	3464842	+	C	A:GCC		V:GTC
YP_001039381| Cthe_2992		RNA polymerase, sigma-24 subunit, ECF subfamily	CDS, non-syn	3513403	+	G	A:GCT		T:ACT
YP_001039393| Cthe_3004		Ferredoxin	CDS, non-syn	3527765	−	C	G:GGA		-:TGA
YP_001039417| Cthe_3028		Pyridoxal-dependent decarboxylase	CDS, non-syn	3551206	−	G	H:CAC		Y:TAC
YP_001039432| Cthe_3043		Hypothetical protein	CDS, non-syn	3568985	+	G	D:GAC	N:AAC	
YP_001039438| Cthe_3049		TPR repeat	CDS, non-syn	3576218	−	G	S:TCC		F:TTC

Genes with mutations in both strains are highlighted in bold

SNP Single nucleotide polymorphism; *non-syn* non-synonymous SNP; *Ref-Nucl* mutated nucleotide in the reference sequence; *aa:codon* amino acid: codon; *CDS* coding region

Discussion

In this study, we successfully increased the ethanol tolerance of wild-type strain *C.thermocellum* ATCC 27405 from about 15 to 50 g/L using either Avicel or cellobiose as substrate. Ethanol-tolerant mutants selected in this study on both Avicel and cellobiose grew well on Avicel, which was however not the case for the ethanol tolerant strain selected in an earlier study (Williams et al. 2007). The observation that similar degrees of ethanol tolerance can be developed in a similar number of transfers on both substrates is significant because it implies that functions associated with cellulose hydrolysis, as well as cell growth, do not introduce substantial additional ethanol sensitivity. Selection of ethanol tolerant strains was carried out with exposure to elevated ethanol on every other transfer, rather than on every transfer as used in prior studies (Herrero and Gomez 1980; Tailliez et al. 1989; Williams et al. 2007; Wang et al. 1983). We hypothesize that alternately growing cultures in the presence and absence of elevated ethanol may have allowed some mutants to recover and grow that would not have done so otherwise, although we did not systematically compare strategies for selecting ethanol-tolerant mutants, and it appears that both strategies are effective. Independent of development of ethanol tolerance, strains transferred multiple times as described herein were

able to grow more rapidly on both Avicel and cellobiose than control strains without being transferred for (Fig. 2). More detailed study of this phenomenon has been initiated.

For each mutant selected in this study, the total number of mutations found is an order of magnitude smaller than that reported previously (Brown et al. 2010). Two genes related to carbamoyl-phosphate (carbomoyl-P) metabolism were mutated in both the E50A and E50C strains (Cthe0593 and Cthe1866), whereas another gene directly involved in carbamoyl-P synthesis (Cthe1868) was mutated only in the E50C strain (Fig. 5). *C. thermocellum* has two copies of genes encoding the large and small subunits of the carbamoyl-P synthase enzyme catalyzing its synthesis from ammonia and ATP. One copy (Cthe0949, 0950) is encoded along with the pyrimidine pathway genes (Cthe0947-0953) and another copy (Cthe1867, 1868) along with the arginine pathway genes (Cthe1863-1869). Carbamoyl-P combines with ornithine to initiate the reactions leading to generation of arginine (and urea). However, a homologue for the arginase enzyme, which catalyzes the hydrolysis of arginine to urea and ornithine, has not been identified in *C. thermocellum*, suggesting the potential lack of a functional urea cycle (Fig. 5). Carbamoyl-P is also a central metabolite in the pathway leading to pyrimidine biosynthesis. Mutations in genes related to carbamoyl-P metabolism suggest potential regulation of flux through the pathways leading to arginine and pyrimidine biosynthesis in the ethanol-adapted strains.

Carbamyl compounds, such as the urea cycle intermediates ornithine, citrulline, etc., and also urea are known to react with ethanol and form ethylcarbamate or urethane. Ethyl carbamate has been demonstrated to be toxic to *Bacillus subtilis* (Giovanni-Donnelly et al. 1967) and is known to inhibit the activity of enzymes such as hexokinase (Isenberg et al. 1954). The mutations observed in arginine biosynthetic pathway genes (Cthe1866 and Cthe1868) and also in the putative glucokinase gene (Cthe0390) might be related with minimizing this inhibitory effect by regulating the flux through the pathway, and thus, the steady-state levels of the carbamyl intermediates under the threshold for spontaneous reaction with ethanol. Since the growth medium used in the adaptation experiments contained urea, further research is being pursued to investigate the effects of urea on ethanol tolerance in *C. thermocellum*.

Identification of the mutated residues within co-factor binding sites in the AdhE protein in both E50A and E50C strains suggests potential alterations in co-factor specificity, enzymatic activity, and/or stability, although this remains to be shown. Bacterial systems such as *Thermoanaerobacter ethanolicus* contain primary and secondary alcohol dehydrogenases with differing co-factor specificities. In *C. thermocellum*, conflicting biochemical studies suggest that the alcohol dehydrogenases are either NADH specific (Lamed and Zeikus 1980; Pei et al. 2010) or capable of utilizing NADH or NADPH (Rydzak et al. 2009). Among the four Fe-containing alcohol dehydrogenases in *C. thermocellum*, Cthe0423 (the bi-functional aldehyde/alcohol dehydrogenase; mutated in this study) is the third most abundant transcript in the cell, while the other alcohol

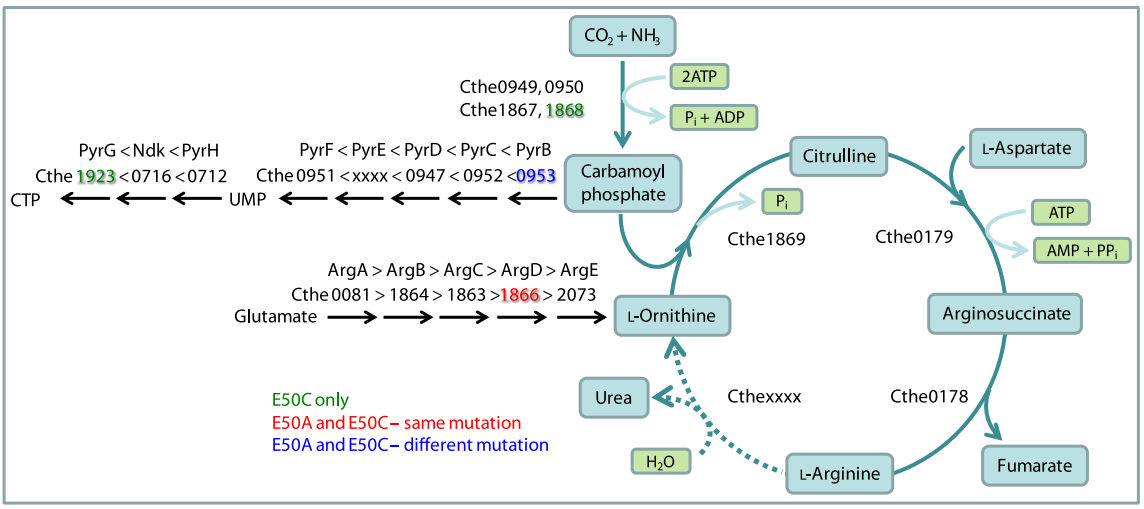

Fig. 5 Figure illustrating the arginine and pyrimidine biosynthetic pathways along with loci numbers of homologous genes in *C. thermocellum* encoding the enzymes catalyzing the various reactions. Carbamoyl-P, a common central intermediate in the two pathways, combines with ornithine to initiate a series of reactions leading to generation of arginine (and urea). The urea cycle, however, may not be fully operational in *C. thermocellum* due to the lack of the homologue for the arginase enzyme which catalyzes the hydrolysis of arginine to ornithine and urea. Genes mutated in the pathway are color-coded depending on the strains displaying the mutations. Genes upstream and downstream of carbamoyl-P and upstream of ornithine were mutated in the ethanol adapted strains suggesting potential regulation of flux through the pyrimidine and arginine biosynthetic pathways

dehydrogenases are transcribed in much lower abundance (Gowen and Fong 2010), suggesting that Cthe0423 is the main ethanol dehydrogenase in *C. thermocellum*.

In prokaryotic systems, there is increasing evidence for the link between alcohol dehydrogenases and maintenance of cellular redox-balance under ethanol stress conditions. For example, ethanol-adapted strain 39EA of *T. ethanolicus* (formerly *Clostridium thermohydrosulfuricum*) was found to lack detectable levels of NAD-linked ADH activity as compared to the wild-type strain (Lovitt et al. 1988). Similarly, *T. ethanolicus* strain 39E H8 adapted to high ethanol levels also lacked activity for the primary alcohol dehydrogenase that is involved in nicotinamide co-factor recycling while increasing the percentage of transmembrane fatty acids (Burdette et al. 2002). More recently, a repressor–operator system was identified in *T. ethanolicus*, whereby the expression of ethanol fermentation pathway enzymes is regulated at the transcriptional level via sensing of a redox signal, such as intracellular NADH concentration (Pei et al. 2011). Therefore, the observed mutations in the AdhE protein in *C. thermocellum* needs to be biochemically characterized for co-factor specificity and altered enzyme properties.

While genome sequencing reveals cellular adaptation and evolution to imposed stress at the genetic level, further investigations on understanding the mechanisms of ethanol tolerance could be pursued using systems biology methods, genetic tools, and biochemical assays.

Acknowledgements The authors are grateful for the support provided by funding grants from the BioEnergy Science Center (BESC), a US Department of Energy (DOE) Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science and Mascoma Corporation. The authors are also grateful for the genome sequencing support provided by the DOE Joint Genome Institute (JGI). Oak Ridge National Laboratory is managed by University of Tennessee UT-Battelle LLC for the Department of Energy under contract no. DE-AC05-00OR22725.

References

Bahari L, Gilad Y, Borovok I, Kahel-Raifer H, Dassa B, Nataf Y, Shoham Y, Lamed R, Bayer EA (2011) Glycoside hydrolases as components of putative carbohydrate biosensor proteins in *Clostridium thermocellum*. J Ind Microbiol Biotechnol 38:825–832. doi:10.1007/s10295-010-0848-9

Brown SD, Yang SY, Guss A, Yang Z, Karpinets T, Klingeman DM, Tschaplinski TJ, Giannone RJ, Hettich RL, Engle NL, Dice L, Rodriguez Jr M, Mielenz J, Cottingham R, Hauser L, Gorin A, Davison BH, Palumbo AV, Lynd L, Keller M (2010) Analysis of the ethanol stress and tolerance mechanisms for *Clostridium thermocellum* through the integration of genome resequencing and systems biology studies. In: 110th ASM General Meeting, San Diego, CA, USA, May 2010

Burdette DS, Jung S-H, Shen G-J, Hollingsworth RI, Zeikus JG (2002) Physiological function of alcohol dehydrogenases and long-chain (C_{30}) fatty acids in alcohol tolerance of *Thermoanaerobacter ethanolicus*. Appl Environ Microbiol 68(4):1914–1918

Ding J, Huang X, Zhang L, Zhao N, Yang D, Zhang K (2009) Tolerance and stress response to ethanol in the yeast *Saccharomyces cerevisiae*. Appl Microbiol Biotechnol 85(2):253–263

Giovanni-Donnelly RD, Kolbye SM, Dipaolo JA (1967) The effect of carbamates on *Bacillus subtilis*. Mutat Res-Fund Mol M 4 (5):543–551

Gowen CM, Fong SS (2010) Genome-scale metabolic model integrated with RNAseq data to identify metabolic states of *Clostridium thermocellum*. Biotechnol J 5(7):759–767

Herrero AA, Gomez RF (1980) Development of ethanol tolerance in *Clostridium thermocellum*: effect of growth temperature. Appl Environ Microbiol 40(3):571–577

Herrero AA, Gomez RF, Roberts MF (1982) Ethanol-induced changes in the membrane lipid composition of *Clostridium thermocellum*. Biochim Biophys Acta 693(1):195–204

Ingram LO (1990) Ethanol tolerance in bacteria. Crit Rev Biotechnol 9(4):305–319

Isenberg HD, Schatz A, Angrist AA, Schatz V, Trelawny GS (1954) Microbial metabolism of carbamates II. Nitrification of urethan by *Streptomyces nitrificans*. J Bacteriol 68(1):5–9

Jeffries TW, Jin Y-S (2000) Ethanol and thermotolerance in the bioconversion of xylose by yeasts. Adv Appl Microbiol 47:221–268

Kahel-Raifer H, Jindou S, Bahari L, Nataf Y, Shoham Y, Bayer EA, Borovok I, Lamed R (2010) The unique set of putative membrane-associated anti-σ factors in *Clostridium thermocellum* suggests a novel extracellular carbohydrate-sensing mechanism involved in gene regulation. FEMS Microbiol Lett 308(1):84–93

Ladisch M, Flatt J, Lynd L, Rajgarhia V, Wenger K, Hogsett DA, Wyman CE, Belcher A, van Rooyen J, Sivasubramanian MS, DiMasi D, Shao X, Draeger J, Kim Y, Ximenes E, Mosier N (2009) Development and deployment of consolidated bioprocessing for production of ethanol. In: 31st symposium on biotechnology for fuels and chemicals, “”San Francisco, CA, USA, 3–6 May 2009

Lamed R, Zeikus JG (1980) Ethanol production by thermophilic bacteria: relationship between fermentation product yields of and catabolic enzyme activities in *Clostridium thermocellum* and *Thermoanaerobium brockii*. J Bacteriol 144(2):569–578

Li H, Ruan J, Durbin R (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res 18(11):1851–1858

Lovitt RW, Shen GJ, Zeikus JG (1988) Ethanol-production by thermophilic bacteria—biochemical basis for ethanol and hydrogen tolerance in *Clostridium thermohydrosulfuricum*. J Bacteriol 170(6):2809–2815

Lynd LR, Wyman CE, Gerngross TU (1999) Biocommodity engineering. Biotechnol Progr 15(5):777–793

Lynd LR, Baskaran S, Casten S (2001) Salt accumulation resulting from base added for pH control, and not ethanol, limits growth of *Thermoanaerobacterium thermosaccharolyticum* HG-8 at elevated feed xylose concentrations in continuous culture. Biotechnol Progr 17(1):118–125

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology (vol 66, pg 506, 2002). Microbiol Mol Biol R 66(4):739–739

Lynd LR, Jin H, Michaels JD, Wyman CE, Dale B (2003) Bioenergy: background, potential, and policy. Center for Strategic and International Studies, Washington

Nataf Y, Bahari L, Kahel-Raifer H, Borovok I, Lamed R, Bayer EA, Sonenshein AL, Shoham Y (2010) *Clostridium thermocellum* cellulosomal genes are regulated by extracytoplasmic polysaccharides via alternative sigma factors. PNAS 107:8646–18651. doi:10.1073/pnas.1012175107

Pei J, Zhou Q, Jiang Y, Le Y, Li H, Shao W (2010) *Thermoanaerobacter* spp. control ethanol pathway via transcriptional regulation and versatility of key enzymes. Metab Eng 12(5):420–428

Pei J, Zhou Q, Jing Q, Li L, Dai C, Li H, Wiegel J, Shao W (2011) The mechanism for regulating ethanol fermentation by redox levels in *Thermoanaerobacter ethanolicus*. Metab Eng 13(2):186–193

Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 5:725–738

Rydzak T, Levin DB, Cicek N, Sparling R (2009) Growth phase-dependant enzyme profile of pyruvate catabolism and end-product formation in *Clostridium thermocellum* ATCC 27405. J Biotechnol 140(3–4):169–175

Sudha Rani K, Seenayya G (1999) High ethanol tolerance of new isolates of *Clostridium thermocellum* strains SS21 and SS22. World J Microbiol Biotechnol 15:173–178

Tailliez P, Girard H, Longin R, Beguin P, Millet J (1989) Cellulose fermentation by an asporogenous mutant and an ethanol-tolerant mutant of *Clostridium thermocellum*. Appl Environ Microbiol 55 (1):203–206

Timmons MD, Knutson BL, Nokes SE, Strobel HJ, Lynn BC (2009) Analysis of composition and structure of *Clostridium thermocellum* membranes from wild-type and ethanol-adapted strains. Appl Microbiol Biotechnol 82:929–939

Tripathi SA, Olson DG, Argyros DA, Miller BB, Barrett TF, Murphy DM, McCool JD, Warner AK, Rajgarhia VB, Lynd LR, Hogsett DA, Caiazza NC (2010) Development of pyrF-based genetic system for targeted gene deletion in *Clostridium thermocellum* and creation of a pta mutant. Appl Environ Microbiol 76:6591–6599. doi:10.1128/AEM.01484-10

Wang DIC, Avgerinos GC, Biocic I, Wang S-D, Fang H-Y, Young FE (1983) Ethanol from cellulosic biomass. Philos Trans R Soc London Ser B 300:323–333

Williams TI, Combs JC, Lynn BC, Strobel HJ (2007) Proteomic profile changes in membranes of ethanol-tolerant *Clostridium thermocellum*. Appl Microbiol Biotechnol 74(2):422–432