Genomics on Pretreatment Inhibitor Tolerance of *Zymomonas mobilis*

Shihui Yang, Martin Keller, and Steven D. Brown

Contents

Abstract The development and use of robust ethanologenic microorganisms resistant to industrially relevant pretreatment inhibitors will be a critical component in the successful generation of biofuel on the industrial scale. Recent progress to understand the genetic basis of pretreatment inhibitor tolerance using genomics and systems biology tools for metabolic engineering for the model ethanologenic bacterium *Zymomonas mobilis* is reviewed in this chapter. The importance of accurate genome annotations and the integration of systems biology data for annotation improvement are highlighted, and case studies that describe the identification and characterization of the *Z. mobilis nhaA*, *hfq*, and *himA* inhibitor tolerance related gene targets are presented.

S. Yang • M. Keller • S.D. Brown (✉)
Biosciences Division and BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
e-mail: brownsd@ornl.gov

Z.L. Liu (ed.), *Microbial Stress Tolerance for Biofuels*, Microbiology Monographs 22,
DOI 10.1007/978-3-642-21467-7_7, © Springer-Verlag Berlin Heidelberg 2012

1 Introduction

A core challenge for next-generation biomass-based cellulosic biofuels is overcoming biomass recalcitrance, or gaining access to its sugars that can then be converted to biofuels (Himmel et al. 2007; Alper and Stephanopoulos 2009). Biomass pretreatment is necessary for optimal release of C-5 and C-6 sugars but can also create a range of inhibitory by-products such as aldehydes, ketones, organic acids, and phenols (Pienkos and Zhang 2010; Palmqvist and Hahn-Hägerdal 2000; Klinke et al. 2004; Liu and Blaschek 2010). Synergistic or additive inhibitory effects are also likely among different hydrolysate inhibitors or metabolic by-products generated during the fermentation such as ethanol, acetate, and lactate (see recent reviews (Mills et al. 2009; Almeida et al. 2007)). An increased lag phase and slower growth increase the biofuel production costs due to reduced production rates and decreased yields (Kadar et al. 2007; Takahashi et al. 1999).

Acetic acid is one major organic acid inhibitor. It is generated by the de-acetylation of hemicelluloses during the pretreatment of biomass. At pH 5.0, about 36% of acetic acid is in the uncharged and undissociated form (HAc). In this form, it is able to pass through the bacterial plasma membrane, leading to uncoupling of the HAc and anion accumulation which causes cytoplasmic acidification (Lawford and Rousseau 1993). Its importance comes from the significant concentrations of acetate that are produced relative to fermentable sugars (McMillan 1994). The produced acetate concentration is also dependent on the feedstock used during the conversion process. An approach to overcoming possible inhibition caused by pretreatment is to remove the inhibitors from the biomass physically or chemically after pretreatment (Pienkos and Zhang 2010). This requires additional equipment and time, thus leading to higher cost. For example, acetate removal processes have been described, but they are energy- or chemical-intensive, and a full-cost analysis has not been reported (McMillan 1994). Applications of inhibitor-tolerant microorganisms appear promising for lower-cost cellulosic biofuel conversion (Almeida et al. 2007; Liu and Blaschek 2010; Liu et al. 2004, 2005, 2008, 2009; Liu and Moon 2009).

Yeast strains are among the current leading industrial biocatalyst microorganisms for fuel production (Hahn-Hagerdal et al. 2006). However, bacteria such as *Escherichia coli, Zymomonas mobilis*, and others are being engineered, developed, and deployed to address commercially important inoculum requirements (Dien et al. 2003; Alper and Stephanopoulos 2009). *Z. mobilis* is a Gram-negative facultative anaerobic bacterium with desirable industrial biocatalyst characteristics, such as high specific productivity, high ethanol yield, and ethanol tolerance (12% v/v) (Dien et al. 2003; Panesar et al. 2006; Rogers et al. 2007). The genome sequence of strain ZM4 has been determined (Seo et al. 2005) and an updated annotation was released recently (Yang et al. 2009a). In addition, the genome sequence and annotation of *Z. mobilis* NCIMB 11163 strain has been reported (Kouvelis et al. 2009) with more strains to be finished or sequenced. Wild-type *Z. mobilis* strains can only utilize a limited range of carbon sources, namely, glucose, fructose, and sucrose. To overcome this limitation, recombinant strains have been engineered to ferment hexose and pentose sugars such

as xylose, arabinose, and other substrates with high yields (Deanda et al. 1996; Zhang et al. 1995), but a low tolerance to acetic acid and a decreased tolerance to ethanol have been reported (Dien et al. 2003; Lawford and Rousseau 1998; Lawford et al. 2001; Ranatunga et al. 1997). In addition, recent achievements to improve transformation efficiency by modifying the DNA restriction-modification systems (Kerr et al. 2010), cellulase expression and secretion (Linger et al. 2010), as well as the genome-scale modeling and *in silico* analysis (Widiastuti et al. 2010), will aid future metabolic engineering and synthetic biology endeavors greatly.

The development and use of robust ethanol-generating microorganisms resistant to industrially relevant inhibitors and with a high-yield ethanol production will be a critical component in the successful generation of fuel ethanol on the industrial scale. However, limited progress has been made in understanding the genetic basis of inhibitor tolerance (Stephanopoulos 2007), and there are few examples of metabolic engineering with systems biology tools for bioprocess development to date (Park et al. 2008). In this chapter, the focus is on genome-based approaches to elucidate molecular mechanisms of inhibitor tolerance for *Z. mobilis*.

2 Genome Annotation of ZM4 Using Systems Biology Studies

Genome sequencing projects provide opportunities for fundamental insights and facilitate strain development (Jeffries 2005). The next generation of new sequencing technologies are delivering fast and relatively inexpensive genome information (see recent reviews (MacLean et al. 2009; Metzker 2010)). Since the first complete microbial genome was published in July 1995 (Fleischmann et al. 1995), the number of finished microbial genomes has grown rapidly. As of August 24, 2010, 1,213 microbial genome sequencing projects have been completed with 3,422 in progress. Detailed information on prokaryotic genome sequencing projects can be accessed at the NCBI Microbial Genomes Resources database: http://www.ncbi.nlm.nih.gov/genomes/MICROBES/microbial_growth.html or the Genomes OnLine Database at: http://www.genomesonline.org/.

The majority of annotation efforts have focused on automatic bioinformatics approaches that are indispensable and based on similarity searches. However, there are issues related to the quality of genome sequencing, and intrinsic annotation errors have also been raised (Devos and Valencia 2001). Inaccurate prediction of open reading frames (ORFs), hypothetical protein descriptions, and discovery of new regulatory elements such as small regulatory are just some of the examples of issues related to genome sequences. On occasions, scientists are faced with different annotation versions generated by different groups for the same genome sequence.

In the case of *Z. mobilis* ZM4, many differences can be seen between the primary annotation and one performed by the J. Craig Venter Institute (JCVI) (http://cmr.jcvi.org/cgi-bin/CMR/GenomePage.cgi?org=ntzm01). Differential gene expression for ORFs predicted by JCVI but absent from the primary annotation has been reported

(Yang et al. 2009b). In addition, the existence of ZM4 plasmids has been reported previously (Yablonsky et al. 1988), but they were not included in the original genome annotation for the strain (Seo et al. 2005). The ZM4 genome annotation has been improved using an updated microbial genome annotation pipeline, the addition of annotated DNA sequences for five plasmids, and data generated from several proteomics studies (Yang et al. 2009a). Almost one-third of the original genome ORF predictions were changed, including important genes such as *nhaA* (ZMO0119) (see detailed descriptions at a later section). The 156 new plasmid gene models represent coding sequences for important genes like an iron-containing alcohol dehydrogenase, hypothetical genes with unknown functions, genes for plasmid maintenance, transport, regulation, metabolism, as well as genes belonging to restriction-modification systems and phage-related genes (Yang et al. 2009a). It is therefore feasible to apply proteomics and next-generation sequencing information for genome annotation improvements, an activity that has received extensive attention recently with several other genome annotations undergoing similar improvements (Armengaud 2009; Baudet et al. 2010; Payne et al. 2010; Wright et al. 2009). The accurate *Z. mobilis* ZM4 genome sequence and annotation are essential components for successful systems biology studies in this and other important ethanologenic microorganisms. In the case of the *Z. mobilis* ZM4 genome update, the improvement was conducted in collaboration with the authors of the primary sequence, which meant the research community was better served by a unified GenBank accession number.

3 Identification of Genes Tolerant to Acetate

Classic strain development that combines random mutagenesis and selection has a long history of success in generation of biocatalysts with industrially designed traits (Parekh et al. 2000; Patnaik 2008). However, the genetic loci contributing to the phenotypic strain changes can be difficult to identify. Systems biology tools and greater access to next-generation sequencing technologies are being increasingly exploited to gain insights into molecular mechanisms that link genotypes to important phenotypes. This section discusses strategies of tolerant gene identifications against acetate in *Z. mobilis*.

3.1 nhaA

An acetate-tolerant *Z. mobilis* mutant (AcR) was created via chemical mutagenesis with *N*-methyl *N*′-nitro *N*-nitrosoguanidine and selection in a continuous culture with a progressively increasing concentration of sodium acetate in the medium feed (Joachimstahl et al. 1998). AcR can efficiently produce ethanol in the presence of 20 g/L NaAc, while the parent organism ZM4 is inhibited above 12 g/L NaAc under

the same conditions (Joachimstahl et al. 1998). Acetic acid was inhibitory to the wild-type-derived strain ZM4(pZB5), which contains the plasmid pZB5 expressing *Escherichia coli* genes for pentose metabolism and xylose assimilation (Zhang et al. 1995) on xylose medium. The major inhibition mechanisms were possibly the intracellular de-energization and acidification (Kim et al. 2000). A recombinant strain was generated by transforming plasmid pZB5 into the AcR background, which can utilize both xylose and glucose with increased acetate resistance and improved fermentation characteristics in the presence of 12 g/L NaAc (Jeon et al. 2002). However, strain AcR was generated while many systems biology tools were being developed or had yet to be conceived, and the molecular mechanism of AcR sodium acetate tolerance was elusive until recently (Yang et al. 2010a).

The mutations in the AcR strain were identified and confirmed through the combination of microarray-based comparative genome sequencing (CGS), next-generation 454-pyroresequencing, and Sanger sequencing (Fig. 1) (Yang et al. 2010a). The CGS results from AcR identified a 1,461-bp (~1.5 kb) region of deleted DNA, which was confirmed using polymerase chain reaction, agarose gel electrophoresis, and Sanger sequencing analysis (Fig. 2). CGS results also identified 38

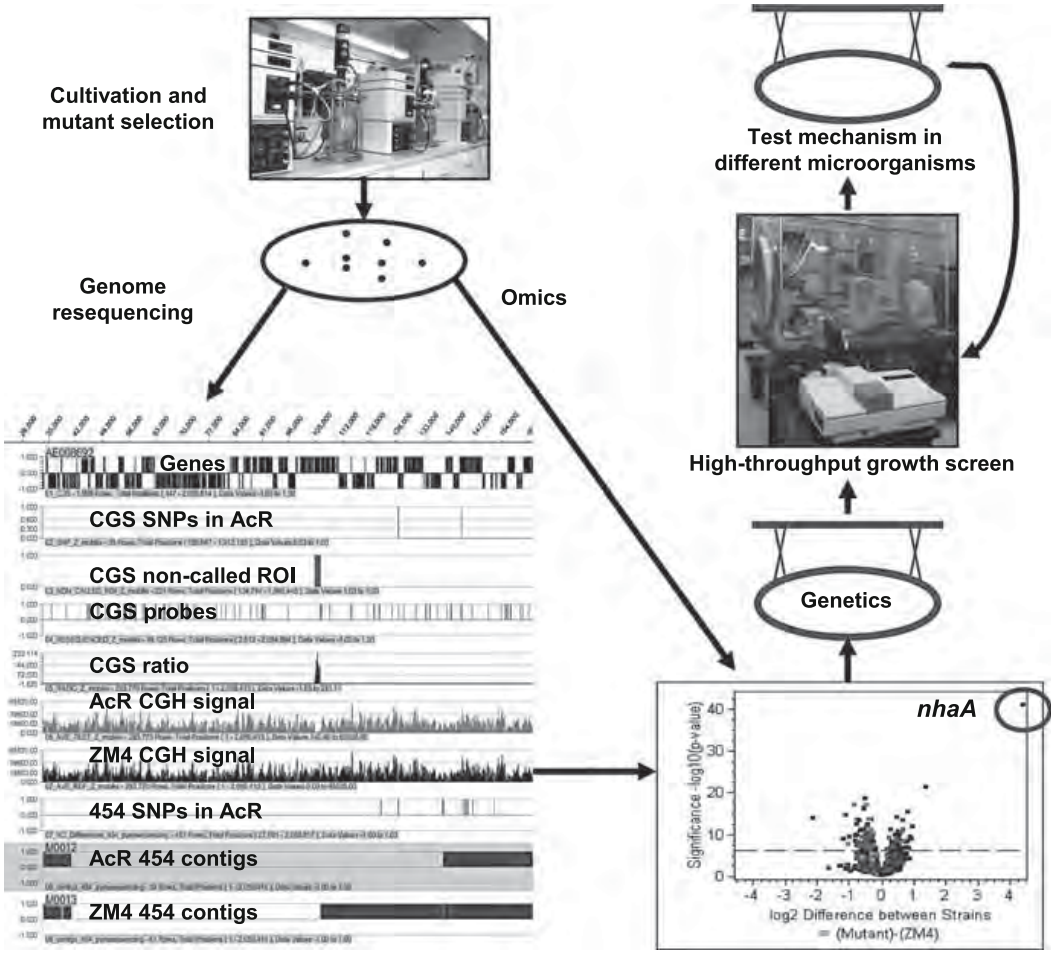

Fig. 1 Overview of the scheme used to identify sodium proton antiporter tolerance mechanisms

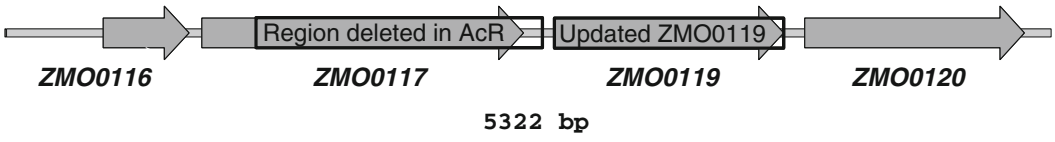

Fig. 2 *Z. mobilis nhaA* (ZMO0119) and its adjacent genes. *ZMO0116*, *ZMO0117*, *ZMO0119*, and *ZMO0120* indicate *Z. mobilis* ZM4 genes. The open box labeled "Region deleted in AcR" is present in ZM4 but deleted in the AcR mutant. The open box labeled "Updated ZMO0119" represents the updated annotation of ZMO0119, which was also used for *nhaA* overexpression plasmid p42-0119 construction

putative AcR single-nucleotide polymorphisms (SNPs), 26 of which were within coding regions and 12 within intergenic regions. From the 454-pyrosequencing shotgun and paired-end sequencing reads generated for ZM4 and AcR, 200 and 219 high confidence differences (HCDiffs) were identified for strains ZM4 and AcR, respectively, compared to the ZM4 reference genome (GenBank accession: AE008692) (Seo et al. 2005). An analysis of the putative mutations shared between ZM4 and AcR identified that most did not contribute to the AcR phenotype and led to improvements in the ZM4 chromosome sequence (Yang et al. 2009a). Only two confirmed SNPs were unique to AcR, with one synonymous SNP (i.e., no change at the amino acid level) found within *ZMO1184* encoding a hypothetical protein and a nonsynonymous SNP in *kup* (*ZMO1209*) encoding a putative potassium transporter. Therefore, the only differences between strains AcR and ZM4 after the ZM4 reference genome annotation update were the 1.5-kb deletion region that truncated *ZMO0117* and DNA upstream of the *nhaA* gene *ZMO0119* (Fig. 2), and two SNPs that affected *ZMO1184* and *ZMO1209*.

To further investigate the correlation between genotypic differences with phenotypic changes, transcriptomics studies were conducted to compare gene expression profiles between wild-type ZM4 and the acetate-tolerant mutant AcR under selective conditions. An analysis of variance (ANOVA) was conducted using JMP Genomics (SAS Institute Inc., Cary, NC) to identify significant differences in exponential and stationary phase transcriptomic profiles for ZM4 and AcR growing either in the presence of NaCl (146 mM or 8.6 g/L NaCl, pH 5.0) or NaAc (146 mM or 12 g/L NaAc, pH 5.0) (Fig. 1). Microarray analysis showed that *nhaA* expression was significantly increased (>16-fold) in strain AcR compared to ZM4 under all conditions tested (Fig. 1). The 1,461-bp deletion of AcR included a 1,363-bp region of *ZMO0117* with only a 275-bp 5′ fragment left and a 160-bp ZM4 *nhaA* upstream region with only 98-bp of the *nhaA* upstream region unchanged in AcR (Fig. 2). A consistently decreased *ZMO0117* signal was detected in each condition in the AcR strain compared to that of ZM4 in transcriptomics studies (Fig. 1). These findings suggested that the deletion in AcR enhanced the *nhaA* expression and likely led to enhanced NaAc tolerance in strain AcR.

To test the hypothesis that the deletion in AcR resulted in higher *nhaA* expression, which augmented NaAc tolerance, a deletion mutant ZM4DM0117 was generated to mimic the AcR 1,461-bp deletion in the wild-type ZM4 strain background by marker exchange (Fig. 2). To test the influence of *ZMO0117* on NaAc tolerance, a *ZMO0117*

insertion mutant strain ZM4IM0117 was constructed (Yang et al. 2010a). The *ZMO0118* gene was combined with *ZMO0119* in the recent update to the ZM4 genome (Yang et al. 2009a, shown in Fig. 2), which demonstrates the importance of working with the best available genome annotation.

To test the correlation between *nhaA* overexpression and NaAc tolerance, a plasmid p42-0119 for *nhaA* overexpression was generated and introduced into wild-type ZM4 background through conjugation and selection (Yang et al. 2010a). The overexpression and mutant strains grew similarly to wild-type ZM4 under anaerobic conditions in RM broth without NaAc supplementation (Fig. 3a). ZM4 wild type and the ZM4IM0117 were unable to grow with the supplementation of 195 mM (or 16 g/L) NaAc at pH 5.0, while the positive control strain AcR grew well (Fig. 3b). The expression of *nhaA* in ZM4 via plasmid p42-0119 restored the growth of ZM4 under these selective conditions, reaching three-fourths of the AcR growth rate. The final cell density (OD_{600nm}) of ZM4 (p42-0119) was only 13% less than that of AcR. ZM4DM0117 was able to grow in the presence of NaAc, achieving more than half of the growth rate and three-fourths of the final cell density of the AcR strain. The similar growth for the insertional mutant ZM4IM0117 as wild-type ZM4 indicated ZM0117 was not responsible for NaAc tolerance. ZM4 NaAc tolerance was augmented substantively by either additional *nhaA* copies provided via plasmid p42-0119 or by recreating the deleted DNA region of AcR in ZM4 wild-type background, which further suggested that the deletion in AcR truncated the *nhaA* promoter region resulted in higher *nhaA* expression, and in turn conferred the tolerance against NaAc.

To investigate the role of *nhaA* with different forms of acetate, ZM4 and AcR strains were grown with the supplementation of the same molar concentrations (195 mM) of sodium chloride (NaCl), NaAc, potassium acetate (KAc), or ammonium acetate (NH_4OAc) (Yang et al. 2010a). Both the sodium and acetate ions had an inhibitory effect on the growth of both *Z. mobilis* wild-type ZM4 and AcR, with decreases in both growth rate and final cell density. The acetate ion was more toxic than the sodium ion. *Z. mobilis* grew more rapidly in the presence of 195 mM NaCl, and the final cell density was higher compared to growth with the supplementation of same molar concentration of NH_4OAc or KAc. At the same molar concentration (195 mM), NaAc was more toxic than KAc or NH_4OAc for ZM4, and the combination of elevated Na^+ and Ac^- ions exerted a synergistic inhibitory effect on ZM4, with its growth totally inhibited.

The AcR strain was selected for sodium acetate tolerance (Joachimstahl et al. 1998). It also has an enhanced tolerance to NaCl, but not NH_4OAc or KAc as compared to the *Z. mobilis* wild-type ZM4 (Yang et al. 2010a). Strain ZM4DM0117 and ZM4 harboring the *nhaA* expression plasmid p42-0119 similarly had enhanced tolerance to NaCl that did not extend to NH_4OAc or KAc. The increased tolerance to NaAc for these strains therefore may be due mostly to an increased tolerance to the sodium ion arising from overexpression of the Na^+/H^+ antiporter gene *nhaA*. The strains were also tested for tolerance to other pretreatment inhibitors such as furfural, HMF, or vanillin, and advantages were not observed. These data again further suggested that NhaA mostly confers enhanced specific tolerance to Na^+ but not to

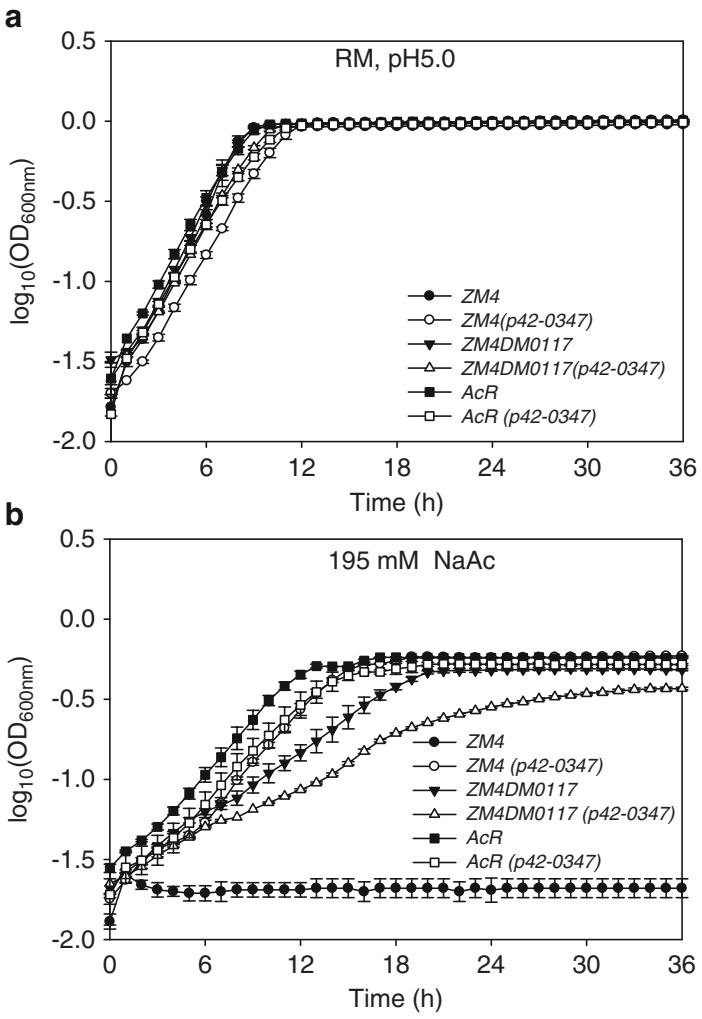

Fig. 3 Higher levels of *Z. mobilis* NaAc tolerance were not achieved through overexpression of both *hfq* and *nhaA*. The impact of Hfq overexpression (via p42-0347) in *Z. mobilis* wild type, acetate-tolerant mutant AcR, and a *Z. mobilis* deletion mutant ZM4DM0117 was assessed at different concentrations of NaAc and compared to that of corresponding parental strains: (**a**) RM broth only without NaAc supplementation as control and (**b**) RM broth with 195 mM NaAc). The growth of *Z. mobilis* strains were monitored by Bioscreen C (Growth Curves USA, NJ) under anaerobic conditions. Strains included in this study were: ZM4 (*Z. mobilis* ZM4 wild-type), AcR (previously described ZM4-derived acetate-tolerant mutant), AcR(p42-0347) (AcR containing a gateway plasmid p42-0347 for *hfq* (ZMO0347) expression), ZM4(p42-0347) (ZM4 containing a gateway plasmid p42-0347 for *hfq* (ZMO0347) expression), ZM4DM0117 (a deletion mutant of ZM4 that mimics the 1.5-Kb deletion in AcR), and ZM4DM0117(p42-0347) (ZM4DM0117 containing a gateway plasmid p42-0347 for *hfq* (ZMO0347) expression). This experiment has been repeated at least three times with similar results. Duplicates were used for each condition

other inhibitors, which reinforces the idea that "you get what you select for." Therefore, inhibitors used in selection regimes need to reflect the real conditions for desired performance where strains are also likely to face a number of different inhibitors.

A similar approach could be used for two other random mutants in the American Type Culture Collection (ATCC), the *Z. mobilis* (ATCC 31822) flocculent mutant

strain ZM401 and the ethanol-tolerant *Z. mobilis* mutant ZM481 (ATCC 31823). The investigation of the genetic differences between the wild-type and mutant strains may provide molecular mechanisms for ethanol tolerance and enhanced flocculation for strain development purposes. The development of high-throughput random mutant generation and selection for improved industrial processing traits, such as those tolerant to high substrate loading and a high concentration of hydrolysate, is needed to further develop industrial microorganisms, and subsequent characterization will allow for more-rapid strain development.

3.2 hfq

Z. mobilis ZM4 gene expression and metabolomic profiles during aerobic and anaerobic conditions were investigated, and it was determined that the ethanol production by *Z. mobilis* decreased with several inhibitory secondary metabolites produced in aerobic conditions (Yang et al. 2009b). This study also revealed that the expression of the putative *hfq* gene ZMO0347 was increased in anaerobic stationary phase compared to that in aerobic conditions (Yang et al. 2009b). Hfq is an RNA chaperone with pleiotropic regulatory roles involved in numerous stress responses (Tsui et al. 1994; Sittka et al. 2008; Zhang et al. 2003; Valentin-Hansen et al. 2004). However, little was known about the role of *Z. mobilis* Hfq in multiple pretreatment inhibitor tolerances until a recent reverse genetics study (Yang et al. 2010b). In this study, an *hfq* insertional mutant was generated in an ZM4 acetate-tolerant strain AcR with the pKnock-Km suicide plasmid system, and plasmid p42-0347 overexpressing *hfq* gene ZMO0347 was introduced into ZM4 wild type, acetate-tolerant mutant AcR, and *hfq* mutant AcRIM0347 by conjugation and selection (Yang et al. 2010b).

An *hfq* mutant (strain AcRIM0347) was unable to grow with the supplementation of 195 mM ammonium acetate or potassium acetate (Yang et al. 2010b). Both the final cell density and the growth rate of the *hfq* mutant were reduced by at least 25% and about 60% in the presence of 195 mM sodium chloride or sodium acetate as compared to that of the parental strain AcR. Consistent with previous reports (Joachimstahl et al. 1998; Yang et al. 2010a), the growth of wild-type ZM4 was completely inhibited in the presence of 195 mM sodium acetate. The introduction of an *hfq*-expressing plasmid (p42-0347) into wild-type ZM4 allowed wild-type ZM4 to obtain a similar growth rate and final cell density to those of acetate-tolerant strain AcR with the supplementation of 195 mM sodium acetate (Yang et al. 2010b). As *hfq* plays a central role in normal *Z. mobilis* physiology, the growth rate of *hfq* mutant AcRIM0347 was reduced to about 20% even without any inhibitor in rich medium (RM) although the final cell density of AcRIM0347 was similar to that of the AcR parental strain. The resistance of AcR to both sodium ion and acetate ion decreased when the *hfq* gene of AcR was inactivated by an insertional mutation. The AcRIM0347 *hfq* mutation was complemented partially by the introduction of an *hfq*-expressing plasmid p42-0347 into the strain.

The reduced inhibitor tolerance of an *hfq* mutant of acetate-tolerant strain AcR and enhanced acetate tolerance of the acetate-sensitive *Z. mobilis* wild-type strain by *hfq* overexpression indicated that *hfq* is important for optimal *Z. mobilis* growth. In addition, the study also showed the possibility to identify inhibitor-tolerant gene targets by top-down systems biology studies followed by reverse genetics approaches.

3.3 nhaA *and* hfq

The *hfq* overexpression plasmid p42-0347 was introduced into acetate-tolerant *Z. mobilis* strains AcR (Joachimstahl et al. 1998) and ZM4 deletion mutant ZM4DM0117 (Yang et al. 2010a), which overexpress the *nhaA* gene to examine whether or not even higher levels of NaAc tolerance could be achieved (Fig. 3). All the strains grew similarly in RM broth, except that those carrying plasmid DNA had slightly reduced growth rates (Fig. 3a), consistent with previous reports (Yang et al. 2010a,2010b). The combined overexpression of *hfq* and *nhaA,* either in an AcR or in a ZM4DM0117 background, did not augment the NaAc tolerance phenotype (Fig. 3b). ZM4 is unable to grow in RM with 195 mM (16 g/L) NaAc, while strain AcR grows well (Joachimstahl et al. 1998; Yang et al. 2010a) (Fig. 3b). The introduction of *hfq*-overexpressing plasmid p42-0347 can improve the NaAc tolerance of wild-type *Z. mobilis* with 195 mM NaAc (Yang et al. 2010b) but neither the deletion mutant ZM4DM0117 nor acetate-tolerant mutant AcR that both already have enhanced NaAc tolerance through *nhaA* overexpression (Yang et al. 2010a) (Fig. 3b). A similar trend for the growth phenotypes was observed under more inhibitory conditions, i.e., 243 mM (20 g/L) or 364 mM (30 g/L) NaAc for AcR strain containing p42-0347 plasmid. The growth rate of ZM4DM0117 (p42-0347) was approximately one quarter less than that of ZM4DM0117 in RM with 195 mM NaAc (Fig. 3b). In addition, the final culture turbidity of ZM4DM0117(p42-0347) in RM with 195 mM NaAc, as measured by OD_{600nm} units, was also reduced by more than one-fifth to 0.37 ± 0.007 compared to the parental strain ZM4DM0117 (Fig. 3b). This indicates that higher levels of NaAc tolerance are not achieved by combining the two independent *hfq* and *nhaA* overexpression mechanisms for *Z. mobilis* NaAc tolerance.

3.4 himA

Another approach to identifying inhibitor-tolerant gene targets and to better understanding microbial physiology uses targeted mutant library construction and characterization. For example, scientists at NREL and DuPont constructed a transposon mutant library of a xylose-utilizing *Z. mobilis* strain and identified a *himA* gene involved in acetate tolerance of *Z. mobilis* (Viitanen et al. 2009). They further engineered a *himA* markerless mutant with reduced *himA* activity and increased

ethanol production compared to parental strains when cultured in a mixed-sugar medium containing xylose, especially in the presence of acetate (Viitanen et al. 2009).

In a similar approach, scientists at the Energy Biosciences Institute (EBI) constructed a "bar-coded" transposon library of *Z. mobilis*. They have established a pooled transposon library containing insertions in 1,695 different genes from 14,009 transposon insertion mutants that includes most non-essential genes in the *Z. mobilis* genome. In addition, a high-throughput 96-well growth screen has been carried out to determine the inhibitory concentration of various inhibitors and potential fuel molecules. The details about this ongoing project can be accessed at the website: http://www.energybiosciencesinstitute.org/index.php?option=com_content&task=view&id=124&Itemid=20. The gene targets identified through this study will hopefully add more inhibitor-tolerant genes for future metabolic engineering or synthetic biology endeavors.

4 Heterologous Expression for Strain Improvement

Heterologous expression of genomic DNA from resistant microorganisms is another strategy that can be employed for strain development purposes. *Deinococcus radiodurans* is an extremely tolerant microorganism isolated in highly radioactive and extreme environments (White et al. 1999). The *D. radiodurans* IrrE protein was identified as a regulator of *recA* expression (Earl et al. 2002), and its heterogeneous expression in *E. coli* promotes DNA repair and protection against oxidative damage (Gao et al. 2003). Although *D. radiodurans* and *E. coli* are quite different organisms, the *irrE* gene protects *E. coli* against multiple stresses, including oxidative, osmotic, and thermal shocks, and confers greater salt tolerance in plants (Pan et al. 2009). Recently, researchers have shown that the *D. radiodurans irrE* gene also confers improved *Z. mobilis* cell viability, abiotic stress tolerance, and ethanol production (Zhang et al. 2010). Numbers of transcripts for key *Z. mobilis* genes (pyruvate decarboxylase and alcohol dehydrogenase) and their enzyme activities were higher in IrrE-expressing *Z. mobilis* as compared to empty vector control strains (Zhang et al. 2010). These studies and others show the potential for heterogeneous expression to expand the genetic pool for strain improvement.

5 Conclusion and Perspectives

In conclusion, recent studies using *Z. mobilis* as a model indicated that accurate genome annotation is crucial for systems biology studies and, in turn, that the data generated from systems biology studies are important for genome annotation improvements. A paradigm for rapid identification and characterization of process-relevant traits created by classical strain development has been proposed through

the integration of systems biology and next-generation sequencing approaches with genetics tools. This affirms the notion that near-term pathway engineering strategies benefit from a combinatorial approach (Alper and Stephanopoulos 2009) as well as the potential to identify the inhibitor-tolerant gene targets by forward genetics (*hfq* case). Gene targets identified from the approaches above can be extended to other industrial biocatalysts by homolog searching and genetics tools (Yang et al. 2010a, b). The phenotypic trait of the acetate-tolerant AcR mutant is largely due to truncation of the *nhaA* promoter region in the AcR, which suggests that future investigations into transcription unit architecture will be a valuable area to pursue through the application RNA-Seq or tiling array technologies. At the same time, integration of information from other omics platforms such as proteomics and metabolomics will provide a more comprehensive profile for metabolic engineering and modeling (Lee et al. 2010). Finally, regulatory networks need to be taken into consideration to better understand and manipulate microbial physiology. (Alper et al. 2006; Alper and Stephanopoulos 2007; Tyo et al. 2007; Cho et al. 2007).

Acknowledgments The authors would like to thank Meghan M. Drake for her careful review and suggestions. The BioEnergy Science Center is a US Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. This manuscript has been authored by UT-Battelle, LLC, under Contract No. DE-AC05-00OR22725 with the US Department of Energy.

References

Almeida JRM, Modig T, Petersson A, Hahn-Hagerdal B, Liden G, Gorwa-Grauslund MF (2007) Increased tolerance and conversion of inhibitors in lignocellulosic hydrolysates by *Saccharomyces cerevisiae*. J Chem Technol Biotechnol 82:340–349

Alper H, Stephanopoulos G (2007) Global transcription machinery engineering: a new approach for improving cellular phenotype. Metab Eng 9:258–267

Alper H, Stephanopoulos G (2009) Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? Nat Rev Micro 7:715–723

Alper H, Moxley J, Nevoigt E, Fink GR, Stephanopoulos G (2006) Engineering yeast transcription machinery for improved ethanol tolerance and production. Science 314:1565–1568

Armengaud J (2009) A perfect genome annotation is within reach with the proteomics and genomics alliance. Curr Opin Microbiol 12:292–300

Baudet M, Ortet P, Gaillard JC, Fernandez B, Guerin P, Enjalbal C, Subra G, de Groot A, Barakat M, Dedieu A, Armengaud J (2010) Proteomics-based refinement of *Deinococcus deserti* genome annotation reveals an unwanted use of non-canonical translation initiation codons. Mol Cell Proteomics 9:415–426

Cho BK, Charusanti P, Herrgard MJ, Palsson BO (2007) Microbial regulatory and metabolic networks. Curr Opin Biotechnol 18:360–364

Deanda K, Zhang M, Eddy C, Picataggio S (1996) Development of an arabinose-fermenting *Zymomonas mobilis* strain by metabolic pathway engineering. Appl Environ Microbiol 62:4465–4470

Devos D, Valencia A (2001) Intrinsic errors in genome annotation. Trends Genet 17:429–431

Dien BS, Cotta MA, Jeffries TW (2003) Bacteria engineered for fuel ethanol production: current status. Applied Microbiol Biotechnol 63:258–266

Earl AM, Mohundro MM, Mian IS, Battista JR (2002) The IrrE protein of *Deinococcus radiodurans* R1 is a novel regulator of *recA* expression. J Bacteriol 184:6216–6224

Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM et al (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269:496–512

Gao GJ, Tian B, Liu LL, Sheng DH, Shen BH, Hua YJ (2003) Expression of *Deinococcus radiodurans* PprI enhances the radioresistance of *Escherichia coli*. DNA Repair 2:1419–1427

Hahn-Hagerdal B, Galbe M, Gorwa-Grauslund MF, Liden G, Zacchi G (2006) Bio-ethanol–the fuel of tomorrow from the residues of today. Trends Biotechnol 24:549–556

Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD (2007) Biomass recalcitrance: Engineering plants and enzymes for biofuels production. Science 315:804–807

Jeffries TW (2005) Ethanol fermentation on the move. Nat Biotechnol 23(1):40–41

Jeon YJ, Svenson CJ, Joachimsthal EL, Rogers PL (2002) Kinetic analysis of ethanol production by an acetate-resistant strain of recombinant *Zymomonas mobilis*. Biotechnol Lett 24:819–824

Joachimstahl E, Haggett KD, Jang JH, Rogers PL (1998) A mutant of *Zymomonas mobilis* ZM4 capable of ethanol production from glucose in the presence of high acetate concentrations. Biotechnol Lett 20:137–142

Kadar Z, Maltha SF, Szengyel Z, Reczey K, De Laat W (2007) Ethanol fermentation of various pretreated and hydrolyzed substrates at low initial pH. Appl Biochem Biotechnol 137:847–858

Kerr AL, Jeon YJ, Svenson CJ, Rogers PL, Neilan BA (2010) DNA restriction-modification systems in the ethanologen, *Zymomonas mobilis* ZM4. Appl Microbiol Biotechnol 89:761–769

Kim IS, Barrow KD, Rogers PL (2000) Nuclear magnetic resonance studies of acetic acid inhibition of *rec Zymomonas mobili*s ZM4(pZB5). Appl Biochem Biotechnol 84–6:357–370

Klinke HB, Thomsen AB, Ahring BK (2004) Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. Appl Microbiol Biotechnol 66:10–26

Kouvelis VN, Saunders E, Brettin TS, Bruce D, Detter C, Han C, Typas MA, Pappas KM (2009) Complete genome sequence of the ethanol producer *Zymomonas mobilis* NCIMB 11163. J Bacteriol 191:7140–7141

Lawford HG, Rousseau JD (1993) The effect of acetic acid on fuel ethanol-Production by *Zymomonas*. Appl Biochem Biotechnol 39:687–699

Lawford HG, Rousseau JD (1998) Improving fermentation performance of recombinant *Zymomonas* in acetic acid-containing media. Appl Biochem Biotechnol 70–2:161–172

Lawford HG, Rousseau JD, Tolan JS (2001) Comparative ethanol productivities of different *Zymomonas* recombinants fermenting oat hull hydrolysate. Appl Biochem Biotechnol 91–3:133–146

Lee K, Park J, Kim T, Yun H, Lee S (2010) The genome-scale metabolic network analysis of *Zymomonas mobilis* ZM4 explains physiological features and suggests ethanol and succinic acid production strategies. Microb Cell Fact 9:94

Linger JG, Adney WS, Darzins A (2010) Heterologous expression and extracellular secretion of cellulolytic enzymes by *Zymomonas mobilis*. Appl Environ Microbiol 76:6360–6369

Liu Z, Blaschek H (2010) Biomass conversion inhibitors and in situ detoxification. In: Vertes A, Qureshi N, Yukawa H, Blaschek H (eds) Biomass to biofuels: strategies for global industries. Wiley, West Sussex, p 27

Liu ZL, Moon J (2009) A novel NADPH-dependent aldehyde reductase gene from *Saccharomyces cerevisiae* NRRL Y-12632 involved in the detoxification of aldehyde inhibitors derived from lignocellulosic biomass conversion. Gene 446:1–10

Liu ZL, Slininger PJ, Dien BS, Berhow MA, Kurtzman CP, Gorsich SW (2004) Adaptive response of yeasts to furfural and 5-hydroxymethylfurfural and new chemical evidence for HMF conversion to 2,5-bis-hydroxymethylfuran. J Ind Microbiol Biotechnol 31:345–352

Liu ZL, Slininger PJ, Gorsich SW (2005) Enhanced biotransformation of furfural and hydroxymethylfurfural by newly developed ethanologenic yeast strains. Appl Biochem Biotechnol 121–124:451–460

Liu ZL, Moon J, Andersh BJ, Slininger PJ, Weber S (2008) Multiple gene-mediated NAD(P) H-dependent aldehyde reduction is a mechanism of in situ detoxification of furfural and 5-hydroxymethylfurfural by *Saccharomyces cerevisiae*. Applied Microbiol Biotechnol 81:743–753

Liu ZL, Ma M, Song M (2009) Evolutionarily engineered ethanologenic yeast detoxifies lignocellulosic biomass conversion inhibitors by reprogrammed pathways. Mol Genet Genomics 282:233–244

MacLean D, Jones JD, Studholme DJ (2009) Application of 'next-generation' sequencing technologies to microbial genetics. Nat Rev 7:287–296

McMillan JD (1994) Conversion of hemicellulose hydrolyzates to ethanol. In: Himmel ME, Baker JO, Overend RP (eds) Enzymatic conversion of biomass for fuels production, vol 566, ACS Symposium Series., pp 411–437

Metzker ML (2010) Sequencing technologies – the next generation. Nat Rev Genet 11:31–46

Mills T, Sandoval N, Gill R (2009) Cellulosic hydrolysate toxicity and tolerance mechanisms in *Escherichia coli*. Biotechnol Biofuels 2:26

Palmqvist E, Hahn-Hägerdal B (2000) Fermentation of lignocellulosic hydrolysates II: inhibitors and mechanisms of inhibition. Biores Technol 74:25–33

Pan J, Wang J, Zhou ZF, Yan YL, Zhang W, Lu W, Ping S, Dai QL, Yuan ML, Feng B, Hou XG, Zhang Y, Ma R, Liu T, Feng L, Wang L, Chen M, Lin M (2009) IrrE, a global regulator of extreme radiation resistance in *Deinococcus radiodurans*, enhances salt tolerance in *Escherichia coli* and *Brassica napus*. PLoS One 4:2

Panesar PS, Marwaha SS, Kennedy JF (2006) *Zymomonas mobilis*: an alternative ethanol producer. J Chem Technol Biotechnol 81:623–635

Parekh S, Vinci VA, Strobel RJ (2000) Improvement of microbial strains and fermentation processes. Appl Microbiol Biotechnol 54:287–301

Park JH, Lee SY, Kim TY, Kim HU (2008) Application of systems biology for bioprocess development. Trends Biotechnol 26:404–412

Patnaik R (2008) Engineering complex phenotypes in industrial strains. Biotechnol Prog 24:38–47

Payne SH, Huang ST, Pieper R (2010) A proteogenomic update to Yersinia: enhancing genome annotation. BMC Genomics 11:460

Pienkos PT, Zhang M (2010) Role of pretreatment and conditioning processes on toxicity of lignocellulosic biomass hydrolysates. Cellulose 16:20

Ranatunga TD, Jervis J, Helm RF, McMillan JD, Hatzis C (1997) Identification of inhibitory components toxic toward *Zymomonas mobilis* CP4(pZB5) xylose fermentation. Appl Biochem Biotechnol 67:185–198

Rogers PL, Jeon YJ, Lee KJ, Lawford HG (2007) *Zymomonas mobilis* for fuel ethanol and higher value products. In: Biofuels, vol 108. Advances in Biochemical Engineering/ Biotechnology. pp 263–288

Seo JS, Chong HY, Park HS, Yoon KO, Jung C, Kim JJ, Hong JH, Kim H, Kim JH, Kil JI, Park CJ, Oh HM, Lee JS, Jin SJ, Um HW, Lee HJ, Oh SJ, Kim JY, Kang HL, Lee SY, Lee KJ, Kang HS (2005) The genome sequence of the ethanologenic bacterium *Zymomonas mobilis* ZM4. Nat Biotechnol 23:63–68

Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, Binnewies TT, Hinton JC, Vogel J (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator. Hfq PLoS Genet 4:8

Stephanopoulos G (2007) Challenges in engineering microbes for biofuels production. Science 315:801–804

Takahashi CM, Takahashi DF, Carvalhal MLC, Alterthum F (1999) Effects of acetate on the growth and fermentation performance of *Escherichia coli* KO11. Appl Biochem Biotechnol 81:193–203

Tsui HC, Leung HC, Winkler ME (1994) Characterization of broadly pleiotropic phenotypes caused by an *hfq* insertion mutation in *Escherichia coli* K-12. Mol Microbiol 13:35–49

Tyo KE, Alper HS, Stephanopoulos GN (2007) Expanding the metabolic engineering toolbox: more options to engineer cells. Trends Biotechnol 25:132–137

Valentin-Hansen P, Eriksen M, Udesen C (2004) The bacterial Sm-like protein Hfq: a key player in RNA transactions. Mol Microbiol 51:1525–1533

Viitanen PV, Tao L, Knoke K, Zhang Y, Caimi PG, Zhang M, Chou Y, Franden M (2009) Process for the production of ethanol from a medium comprising xylose, employing a recombinant *Zymomonas* strain having a reduced *himA* expression. Patent WO/2009/058938

White O, Eisen JA, Heidelberg JF, Hickey EK, Peterson JD, Dodson RJ, Haft DH, Gwinn ML, Nelson WC, Richardson DL, Moffat KS, Qin H, Jiang L, Pamphile W, Crosby M, Shen M, Vamathevan JJ, Lam P, McDonald L, Utterback T, Zalewski C, Makarova KS, Aravind L, Daly MJ, Minton KW, Fleischmann RD, Ketchum KA, Nelson KE, Salzberg S, Smith HO, Venter JC, Fraser CM (1999) Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. Science 286:1571–1577

Widiastuti H, Kim JY, Selvarasu S, Karimi IA, Kim H, Seo JS, Lee DY (2010) Genome-scale modeling and in silico analysis of ethanologenic bacteria *Zymomonas mobilis*. Biotechnol Bioeng 108:655–665

Wright JC, Sugden D, Francis-McIntyre S, Riba-Garcia I, Gaskell SJ, Grigoriev IV, Baker SE, Beynon RJ, Hubbard SJ (2009) Exploiting proteomic data for genome annotation and gene model validation in *Aspergillus niger*. BMC Genomics 10:61

Yablonsky MD, Goodman AE, Stevnsborg N, Delima OG, Demorais JOF, Lawford HG, Rogers PL, Eveleigh DE (1988) *Zymomonas mobilis* CP4: a clarification of strains *via* plasmid profiles. J Biotechnol 9:71–79

Yang S, Pappas KM, Hauser LJ, Land ML, Chen G-L, Hurst GB, Pan C, Kouvelis V, Typas M, Pelletier DA, Klingeman DM, Chang Y-J, Samatova NF, Brown SD (2009a) Improved genome annotation for Zymomonas mobilis. Nat Biotechnol 27:893–894

Yang S, Tschaplinski TJ, Engle NL, Carroll SL, Martin SL, Davison BH, Palumbo AV, Rodriguez M Jr, Brown SD (2009b) Transcriptomic and metabolomic profiling of *Zymomonas mobilis* during aerobic and anaerobic fermentations. BMC Genomics 10:34

Yang S, Land ML, Klingeman DM, Pelletier DA, Lu T-YS, Martin SL, Guo HB, Smith JC, Brown SD (2010a) Paradigm for industrial strain improvement identifies sodium acetate tolerance loci in *Zymomonas mobilis* and *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 107:10395–10400

Yang S, Pelletier DA, Lu TY, Brown SD (2010b) The *Zymomonas mobilis* regulator *hfq* contributes to tolerance against multiple lignocellulosic pretreatment inhibitors. BMC Microbiol 10:135

Zhang M, Eddy C, Deanda K, Finkestein M, Picataggio S (1995) Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis*. Science 267:240–243

Zhang A, Wassarman KM, Rosenow C, Tjaden BC, Storz G, Gottesman S (2003) Global analysis of small RNA and mRNA targets of Hfq. Mol Microbiol 50:1111–1124

Zhang Y, Ma RQ, Zhao ZL, Zhou ZF, Lu W, Zhang W, Chen M (2010) *irrE*, an exogenous gene from *Deinococcus radiodurans*, improves the growth of and ethanol production by a *Zymomonas mobilis* strain under ethanol and acid stresses. J Microbiol Biotechnol 20:1156–1162