Metabolic Engineering 10 (2008) 154–165

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/ymben

A genomics approach to improve the analysis and design of strain selections

T.E. Warnecke [a], M.D. Lynch [b], A. Karimpour-Fard [a], N. Sandoval [a], R.T. Gill [a,*]

[a] Department of Chemical and Biological Engineering, University of Colorado, ucB424/ECCH120, Boulder, CO, USA
[b] OpX Biotechnologies, Boulder, CO, USA

ARTICLE INFO

Article history:
Received 28 January 2008
Received in revised form
17 March 2008
Accepted 3 April 2008
Available online 9 May 2008

Keywords:
3-Hydroxypropionic acid
Tolerance
Selection
Strain engineering
Multi-SCale Analysis of Library Enrichments (SCALEs)

ABSTRACT

Strain engineering has been traditionally centered on the use of mutation, selection, and screening to develop improved strains. Although mutational and screening methods are well-characterized, selection remains poorly understood. We hypothesized that we could use a genome-wide method for assessing laboratory selections to design selections with enhanced sensitivity (true positives) and specificity (true negatives) towards a single desired phenotype. To test this hypothesis, we first applied multi-SCale Analysis of Library Enrichments (SCALEs) to identify genes conferring increased fitness in continuous flow selections with increasing levels of 3-hydroxypropionic acid (3-HP). We found that this selection not only enriched for 3-HP tolerance phenotypes but also for wall adherence phenotypes (41% false positives). Using this genome-wide data, we designed a serial-batch selection with a decreasing 3-HP gradient. Further examination by ROC analysis confirmed that the serial-batch approach resulted in significantly increased sensitivity (46%) and specificity (10%) for our desired phenotype (3-HP tolerance).

1. Introduction

Selection is a powerful yet poorly understood strain-engineering tool. This fact has limited the effective use of selection in a broad range of industrial biotechnology applications. Laboratory selections are based on enrichment for the fittest members of a population and dilution of the less fit members of the population. The challenge here lies in defining fitness in a manner that enables the design of selections directed at phenotypes of interest. That is, multiple phenotypes (fast growth, high tolerance, etc.) can simultaneously contribute to overall fitness in a particular environment. Once the environment changes, the relative weight of each phenotype in defining fitness can also change (Vasi et al., 1994; Elena and Lenski, 2003; Gall et al., 2007). While the complicated nature of laboratory selection is well-established, understanding of the relationships among overall fitness, selectable phenotypes, and fitness altering genes has remained poor (Gottschal, 1993; Dykhuizen and Dean, 1994; Finkel and Kolter, 1999; Cooper and Lenski, 2000; Elowitz et al., 2002). This lack of understanding has delayed progress in strain-engineering efforts, and thus the development of efficient, robust, and well-characterized industrial strains.

New genomics tools can be used to address these issues (Crameri et al., 1994; Shoemaker et al., 1996; Giaever et al., 1999; Winzeler et al., 1999; Gill et al., 2002; Kacmar et al., 2006; Alper and Stephanopoulos, 2007). One such tool, multi-SCale Analysis of Library Enrichments (SCALEs), is capable of quantifying enrichment patterns corresponding to $>10^6$ individual clones comprising genomic libraries with defined insert sizes (Lynch et al., 2007). In *Escherichia coli*, this method has enabled the identification of genetic mechanisms underlying biofilm formation, as well as mechanisms underlying the growth inhibition due to several antimicrobial compounds, including aspartic acid anti-metabolites and 1-napthol (Bonomo et al., 2007; Gall et al., 2007; Lynch et al., 2007). In one case, selections involving three different aspartic acid anti-metabolites resulted in the identification of distinct mechanisms for alleviating tolerance that would not have been readily distinguished via traditional library screening methods. In another case, selections performed in the presence of 1-napthol via single or serial-batch methods resulted in the enrichment of genetic elements that were specific to each selection strategy. It was shown that genotypes selected in single batch conferred rapid growth phenotypes while those selected via serial transfer exhibited general improvements in multiple phenotypes, including growth rate, lag time, and maximum 1-napthol tolerance. This result led to a general approach for relating genome-scale fitness measurements to the multiple phenotypes that contribute to overall fitness in any selection (Gall et al., 2007). Here, we describe the extension of this approach to enable the design of more efficient strain selections. Specifically, we have used this strategy to analyze and then re-design strain selections with improved specificity towards a desired phenotype (3-hydroxypropionic acid (3-HP) tolerance in *E. coli*).

* Corresponding author.
E-mail address: rtg@colorado.edu (R.T. Gill).

1096-7176/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymben.2008.04.004

Engineering of 3-HP tolerance in *E. coli* is a relevant model system for our studies for several reasons. First, biological production of this platform chemical has been proposed as an alternative to the current petrochemical route to the production of several commodity chemicals including 1,3-propanediol, acrylic acid, methyl acrylate, and acrylamide (Werpy and Petersen, 2004). The combined market for these chemicals is well over $1 billion/year (Paster et al., 2003). Second, while biological production of 3-HP has been reported, the yield remains well below what is required for commercialization (Suthers and Cameron, 2001; Vollenweider and Lacroix, 2004). Third, severe growth inhibition has been observed to 3-HP concentrations as low as 10 g/L in minimal media (pH 7.0), which limits the economic feasibility of 3-HP production (data not shown). Importantly, while the exact mechanisms underlying 3-HP toxicity are not known, it is known that such mechanisms involve specific anion effects on metabolism and that addition of extracellular amino acids to the culture medium partially alleviates 3-HP toxicity (data not shown). Overall, 3-HP tolerance represents a trait that is a fundamental barrier to commercialization, yet remains poorly understood and thus not well-addressed by rational metabolic engineering approaches.

Here, we have used SCALEs to improve the design of selections for the identification of 3-HP tolerant phenotypes. Selection design involves decision making about a range of criteria, each of which modulates selective pressure, including (i) growth strategy (continuous flow, single batch, repeated batch, planktonic, agar plates), (ii) 3-HP concentration in the environment (initial concentration (C_0), rate of change in concentration (dC/dt), and final concentration (C_f)), and (iii) media conditions (rich, minimal, supplemented minimal), among others. For our first selection, we chose to use a continuous flow reactor fed with minimal media and increasing amounts of 3-HP. We employed the SCALEs method to track and analyze changes in our genomic library population ($>10^6$ clones). On the basis of the functions of genes identified to confer increased fitness, we determined that this selection strategy enriched for wall adherent or biofilm phenotypes at the expense of true 3-HP tolerance phenotypes, the result of which was an increased false positive rate for this selection. We then used this information to redesign our selections in a way to reduce the selective advantage for biofilm formation and increase the enrichment of true 3-HP tolerant phenotypes. This second selection utilized repeated batch cultures with decreasing concentrations of 3-HP. We determined that this second selection enriched for a distinct collection of clones when compared with the first selection. We then confirmed that a collection of such clones contained genetic elements conferring increased tolerance to 3-HP. Finally, we performed statistical analyses to conclude that the second selection indeed displayed a higher sensitivity (true positive rate) and specificity (true negative rate) for 3-HP tolerant phenotypes.

2. Methods

2.1. Bacteria, plasmids, and media

Wild-type *E. coli* K12 (ATCC #29425) was used for the preparation of genomic DNA. Genomic libraries were constructed using the pSMART-LCKAN (Lucigen, Middleton, WI). Cultures for library construction were cultivated in Luria–Bertani (LB) medium at 37 °C. Selections were fed with MOPS minimal medium (Neidhardt, 1974). Antibiotic concentrations were as follows: 20 μg kanamycin/mL and 100 μg chloramphenicol/mL.

2.2. Genomic library construction

Cultures of the *E. coli* K12 were cultivated overnight in 500 mL of LB at 37 °C to an optical density of 1.0 measured by absorbance at 600 nm (OD_{600}). DNA was extracted using a Genomic DNA Purification kit (Qiagen) according to manufacturer's instructions. Five samples containing 50 μg of purified genomic DNA were digested using two blunt-cutter restriction enzymes: AluI and RsaI (Invitrogen). Both enzymes have four base pair recognition sequences and are used in tandem to ensure a random digestion of the genomic DNA. Digestion reactions were carried out with a total volume of 50 μL. The reactions contained 1 unit of Rsa1, 1 unit Alu1, 50 mM Tris–HCl (pH 8.0), and 10 mM $MgCl_2$ and were incubated at 37 °C for 1, 2, 5, 10, and 15 min, respectively. The partially digested DNA was immediately mixed and separated based on size using agarose gel electrophoresis. DNA fragments of 0.5, 1, 2, 4, and greater than 8 kb were excised from the gel and purified with a Gel Extraction Kit (Qiagen).

The purity of the DNA fragments was quantified using UV absorbance, each with an A260/A280 absorbance ratio of >1.7. Ligation of the purified, fragmented DNA with the pSMART-LCKan vectors was performed with the CloneSmart Kit (Lucigen) according to manufacturer's instructions. The ligation product was then electroporated into *E. Cloni* 10 GF' Elite Electrocompetent Cells (Lucigen), plated on LB+kanamycin, and incubated at 37 °C for 24 h. Dilution cultures with 1/1000 of the original transformation volume were plated on LB+kanamycin in triplicate to determine accurate transformation efficiency and to confirm greater than 10^5 transformants per library corresponding to greater than 99% probability of complete library coverage.

Colonies were harvested by gently scraping the plates into TB media. The cultures were immediately resuspended by vortexing, and aliquoted into 15–1 mL freezerstock cultures with a final glycerol concentration of 15% (v/v) (Sambrook, 2001). The remainder of the culture was pelleted by centrifugation for 15 min at 3000 rpm. Plasmid DNA was extracted according to the manufacturer's instructions using a HiSpeed Plasmid Midi Kit (Qiagen). To confirm insert sizes and positive transformant numbers, plasmids were isolated from random clones for each library size using a Qiaprep Spin MiniPrep Kit (Qiagen). Purified plasmids were then analyzed by either PCR or restriction digestion. PCR using the SL1 (5′-CAG TCC AGT TAC GCT GGA GTC-3′) and SR2 (5′-GGT CAG GTA TGA TTT AA A TGG TCA GT) primers was performed on eight clones from the 0.5, 1, and 2 kbp insert libraries. Restriction digestions with the enzyme EcorV were carried out for eight clones from the 2, 4, and 8 kbp insert libraries. Inspection by electrophoresis showed that the required number of colonies contained an insert of the expected size for complete genomic representation (greater than 10^5 clones) and that chimeras were not present.

2.3. Transformation of library DNA

Purified plasmid DNA from each library was introduced into MACH1™-T1® (Invitrogen) by electroporation. MACH1™-T1® cultures were made electrocomptent by a standard glycerol wash procedure on ice to a final concentration of 10^{11} cells/mL (Sanbrook et al.). 1/1000 volume of the original transformations was plated on LB+kanamycin in triplicate to determine transformation efficiency and adequate transformant numbers ($>10^6$). The original cultures were combined and diluted to 100 ml with MOPS minimal media+kanamycin and incubated at 37 °C for 6 h or until reaching an OD_{600} of 0.20.

2.4. Continuous flow reactor selections

The transformed library was introduced into a CFR with working volume of 50 mL. MOPS minimal media+kanamycin was introduced at a controlled volumetric flow rate by use of a