Published online 19 June 2012

Nucleic Acids Research, 2012, Vol. 40, No. 18 ***e142***
doi:10.1093/nar/gks549

Customized optimization of metabolic pathways by combinatorial transcriptional engineering

Jing Du[1], Yongbo Yuan[1], Tong Si[1], Jiazhang Lian[1] and Huimin Zhao[1,2,*]

[1]Department of Chemical and Biomolecular Engineering, Institute for Genomic Biology and [2]Departments of Chemistry, Biochemistry and Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Received February 17, 2012; Revised May 3, 2012; Accepted May 15, 2012

ABSTRACT

A major challenge in metabolic engineering and synthetic biology is to balance the flux of an engineered heterologous metabolic pathway to achieve high yield and productivity in a target organism. Here, we report a simple, efficient and programmable approach named 'customized optimization of metabolic pathways by combinatorial transcriptional engineering (COMPACTER)' for rapid tuning of gene expression in a heterologous pathway under distinct metabolic backgrounds. Specifically, a library of mutant pathways is created by *de novo* assembly of promoter mutants of varying strengths for each pathway gene in a target organism followed by high-throughput screening/selection. To demonstrate this approach, a single round of COMPACTER was used to generate both a xylose utilizing pathway with near-highest efficiency and a cellobiose utilizing pathway with highest efficiency that were ever reported in literature for both laboratory and industrial yeast strains. Interestingly, these engineered xylose and cellobiose utilizing pathways were all host-specific. Therefore, COMPACTER represents a powerful approach to tailor-make metabolic pathways for different strain backgrounds, which is difficult if not impossible to achieve by existing pathway engineering methods.

INTRODUCTION

Microbial synthesis of value-added compounds typically depends upon the creation or introduction of heterologous metabolic pathways into the production hosts (1–4). Often, in order to enable commercially viable production of these compounds, the metabolic flux in the heterologous pathways must be optimized to avoid metabolic burden from over-expression of certain genes, redox imbalance from unmatched cofactor specificity, accumulation of unstable or toxic intermediates or other bottlenecks that result in growth inhibition (5–11). Traditional approaches for balancing metabolic flux involve identifying bottlenecks in metabolic pathways and debottlenecking the pathways. The most common debottlenecking strategies include the over-expression of key metabolic genes, deletion of competing pathways and the improvement of certain catalytic enzymes through protein engineering. Traditional approaches usually focus on a certain key step in the pathway instead of the whole dynamic metabolic network (the network consisting of the multi-step heterologous pathway and the overall metabolic background of the host strain). As a result, most traditional approaches have met with limited success in the optimization of multi-step pathways (12).

Recently, a number of innovative approaches have been developed that balance metabolic flux at a global level through either perturbation of global transcription machinery (13), genomic-scale mapping of fitness altering genes (14,15) or multiplex automated genome engineering (16). In addition, new strategies have been developed to balance the metabolic flux within a target pathway by tuning pathway gene expression through engineering of the promoters (13), ribosome binding sites (7) and intergenic regions (5). These new approaches have enabled simultaneous optimization of a metabolic pathway to a certain degree. However, due to the distinct metabolic backgrounds between laboratory strains and industrial strains, pathways optimized in laboratory strains may not be transferable to industrial production strains (17). Therefore, it is highly desirable to simultaneously balance and fine-tune the metabolic flux within a heterologous pathway based on the metabolic background of the host strain.

To this end, we developed the COMPACTER method that is capable of simultaneously optimizing multiple genes in a metabolic pathway and tailoring the target

*To whom correspondence should be addressed. Tel: +1 217 333 2631; Fax: +1 217 333 5052; Email: zhao5@illinois.edu

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

pathway to a production host of interest. In this approach, a metabolic pathway is introduced into a host strain on a single copy vector with each gene under the control of a distinct promoter and terminator pair. The use of a single copy vector will eliminate the potential issue of unstable gene expression levels due to the varying copy number of a multi-copy vector and offer a greater flexibility compared to an integrative vector because the same pathway library assembled on a single copy vector can be freely transferred to different strains. The use of a distinct promoter and terminator pair for each pathway gene is designed to avoid repetitive sequences that may result in undesired recombination events in the target pathway. Nucleotide analog mutagenesis (18) is used to generate a series of promoter mutants of varying strengths. These promoter mutants are assembled with the metabolic genes of interest into a library of mutant pathways with different expression patterns using the DNA assembler method (19). In principle, this library should contain mutant pathways with all possible combinations of expression levels present in the promoter mutants. A high-throughput screening or selection method is then used to identify a mutant pathway with a balanced metabolic flux from the library of mutant pathways in the target host (Figure 1).

MATERIALS AND METHODS

Strains, media and cell cultivation

Saccharomyces cerevisiae strain INVSc1 (*MATa his3Δ1 leu2 trp1-289 ura3-52 MATα his3Δ1 leu2 trp1-289 ura3-52*) was purchased from Invitrogen (Life Technologies, Grand Island, NT, USA). *S. cerevisiae* L2612 (*MATα leu2-3 leu2-112 ura3-52 trp1-298 can1 cyn1 gal*$^{+}$) was a gift from Prof. Yong-su Jin (20). Still Spirits (Classic) Turbo Distiller's Yeast was purchased from Homebrew Heaven (Everett, WA, USA). *Escherichia coli* DH5α (Cell Media Facility, University of Illinois at Urbana-Champaign, Urbana, IL, USA) was used for recombinant DNA manipulation. Yeast strains were cultivated in either synthetic dropout medium (0.17% Difco yeast nitrogen base without amino acids and ammonium sulfate, 0.5% ammonium sulfate and 0.083% amino acid drop out mix) or YPA medium (1% yeast extract, 2% peptone and 0.01% adenine hemisulfate) supplemented with sugar as a carbon source. *E. coli* strains were cultured in Luria broth (LB; Fisher Scientific, Pittsburgh, PA, USA). *S. cerevisiae* strains were cultured at 30°C and 250 rpm for aerobic growth, and 30°C and 100 rpm for oxygen limited conditions. *E. coli* strains were cultured at 37°C and 250 rpm unless specified otherwise. All restriction enzymes and plasmids, including pRS414, pRS415 and pRS416, were purchased from New England Biolabs (Ipswich, MA, USA). All chemicals were purchased from Sigma-Aldrich or Fisher Scientific. A detailed description on fermentation conditions and HPLC analysis may be found in the Supplementary Method S1.

Plasmid and strain construction

Most of the cloning work was done using yeast homologous recombination mediated by the DNA assembler method (19). DNA fragments flanked with regions

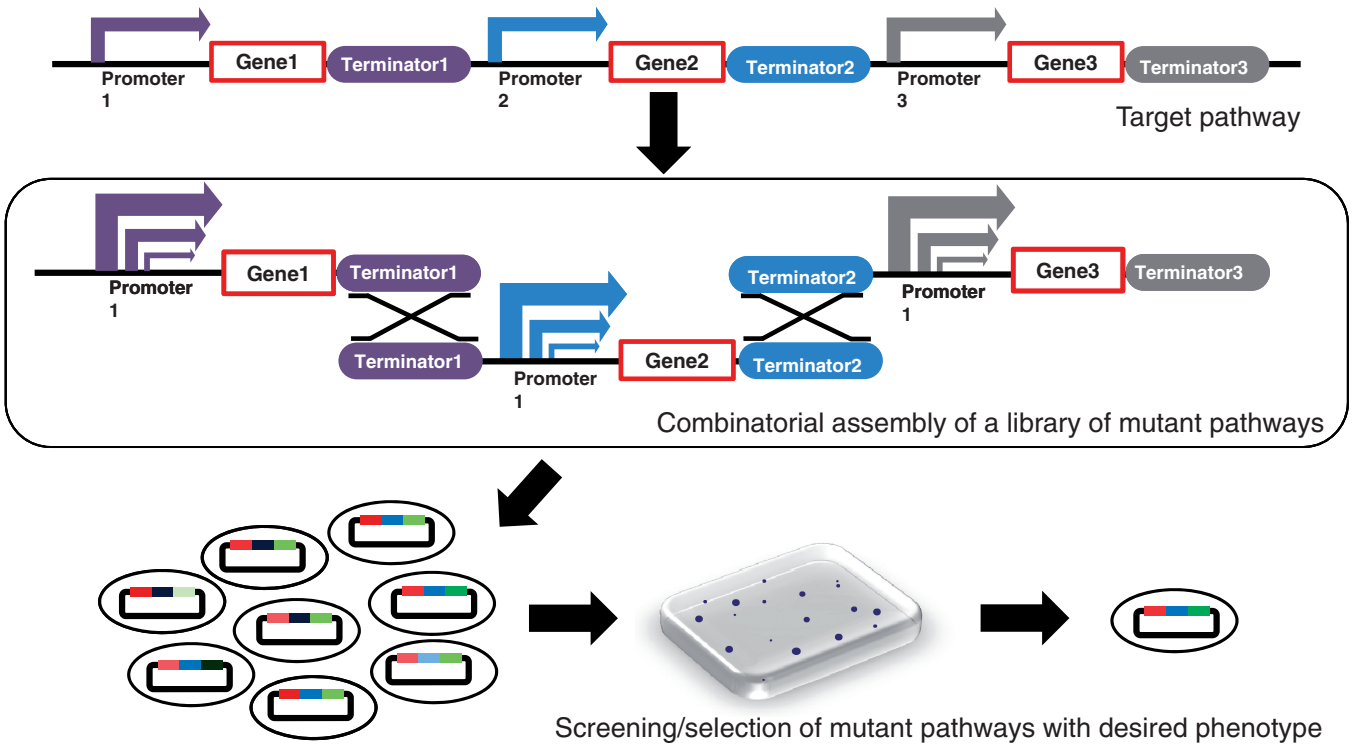

Figure 1. General scheme of the COMPACTER method for combinatorial pathway design. See text for details.

homologous to adjacent DNA fragments were generated with polymerase chain reaction (PCR) and all DNA fragments were purified and co-transformed into *S. cerevisiae* along with the backbone. To confirm the correct clones from transformants, yeast plasmids were isolated using a Zymoprep II yeast plasmid isolation kit (Zymo Research, Irvine, CA, USA) and transferred into *E. coli*. Plasmids from *E. coli* were then isolated and confirmed using diagnostic PCR. A detailed description on creation of promoter mutants may be found in the Supplementary Method S2.

Mutant pathway library creation

To facilitate the construction of pathway libraries, helper plasmids were constructed to contain the DNA sequence homologous to the upstream of the pathway, a unique restriction site, the pathway gene, a terminator and DNA sequence homologous to the downstream of the pathway. The promoter mutant libraries were then cloned into corresponding helper plasmids using the DNA assembler method (19). The resulted constructs were confirmed using diagnostic PCR and the full gene expression cassette consisting of the upstream homologous sequence, the promoter mutant, the gene, the terminator and the downstream homologous sequence were amplified using PCR. The full gene expression cassette were then mixed together and transferred with a linearized single copy shuttle vector to create the pathway library (Supplementary Figures S3 and 8). A detailed description on mutant pathway library creation may be found in the Supplementary Method S3.

Library screening

Libraries of xylose utilizing pathways and cellobiose utilizing pathways were screened using a multi-tiered screening strategy consisting of petri dish plate-based screening, culture tube-based screening and shake-flask based screening. A detailed description of the library screening protocol and subsequent characterization of the positive strains by fermentation and HPLC analysis may be found in the Supplementary Method S4.

Quantitative PCR

Fresh samples of cell cultures harvested at the exponential phase grown under the same condition as mentioned above for xylose or cellobiose fermentation were used to determine the relative expression ratio of metabolic genes via quantitative PCR (qPCR). The total RNA was isolated using the RNeasy mini kit (QIAGEN, Valencia, CA, USA) following the manufacturer's instructions. The RNA samples were then reversed transcribed into cDNA using the Transcriptor First Strand cDNA Synthesis kit using oligo-dT primer following the manufacturer's instructions (Roche, Indianapolis, IN, USA). The qPCR experiments were carried out using the LightCycler® 480 system (Roche, Indianapolis, IN, USA) using the SYBRGreen-based method following the manufacturer's instructions.

RESULTS

Creating a highly efficient xylose utilizing pathway in both laboratory and industrial *Saccharomyces cerevisiae* strains

To demonstrate the utility of the COMPACTER method, we first sought to balance the metabolic flux in an engineered xylose utilizing pathway in *Saccharomyces cerevisiae* for lignocellulosic biofuel production. Xylose is the most abundant pentose sugar in hemicellulose hydrolysates, but cannot be directly utilized by the most commonly used ethanol producer, *S. cerevisiae* (21). Two types of pathways were used to enable xylose utilization in *S. cerevisiae*, the isomerase-based bacterial pathway and the reductase–dehydrogenase-based fungal pathway (22). The fungal xylose utilizing pathway consists of a xylose reductase, xylitol dehydrogenase and xylulokinase (Figure 2a). The fungal xylose utilizing pathway was used to demonstrate the utility of the COMPATER method, because the balancing of enzyme activities and cofactor usage is crucial for the performance of the pathway (22). Despite three decades of effort in optimizing the xylose utilizing pathway in *S. cerevisiae* for biofuels production (21–24), no existing recombinant xylose assimilating strains are efficient enough for commercial application. This may be due to the low expression level and activity of heterologous proteins, redox imbalance resulting from different cofactor preferences for oxidation and reduction reactions and/or suboptimal metabolic flux through the different catalytic steps (25).

Based on the results from an enzyme screening program (Supplementary Figure S1), the xylose reductase from *Candida shehatae* (csXR), the xylitol dehydrogenase from *Candida tropicalis* (ctXDH) and the xylulokinase from *Pichia pastoris* (ppXKS) were selected to construct the xylose utilizing pathway in both laboratory and industrial yeast strains. In the enzyme screening program, enzyme activities of 20 xylose reductases, 22 xylitol dehydrogenases and 19 xylulokinases were assayed using cell lysates. The csXR, ctXDH and ppXKS were chosen because they have the highest activity and closest matched cofactor specificity among all tested enzyme homologs. In the engineered xylose utilizing pathway, the expressions of csXR, ctXDH and ppXKS were driven by PDC1 promoter mutants, TEF1 promoter mutants and ENO2 promoter mutants, respectively. A library of mutant xylose utilizing pathways was generated using pre-selected promoters of varying strengths. Briefly, each of the three constitutive yeast promoters—PDC1, ENO2 and TEF1—in the engineered xylose utilizing pathway was subjected to nucleotide analog mutagenesis (26) to create a pool of promoter mutants. Of note, nucleotide analog mutagenesis was used instead of conventional error-prone PCR because it can generate a very high mutation rate upon introduction of nucleotide analogs in the PCR reaction, which leads to rapid creation of promoter mutants with varying strengths. Approximately 10 promoter mutants of varying strengths were isolated for each promoter using green fluorescent protein (GFP) as a reporter (Supplementary Figure S2). The isolated promoter mutants were then used in the construction of the pathway library (Supplementary Figure S3).

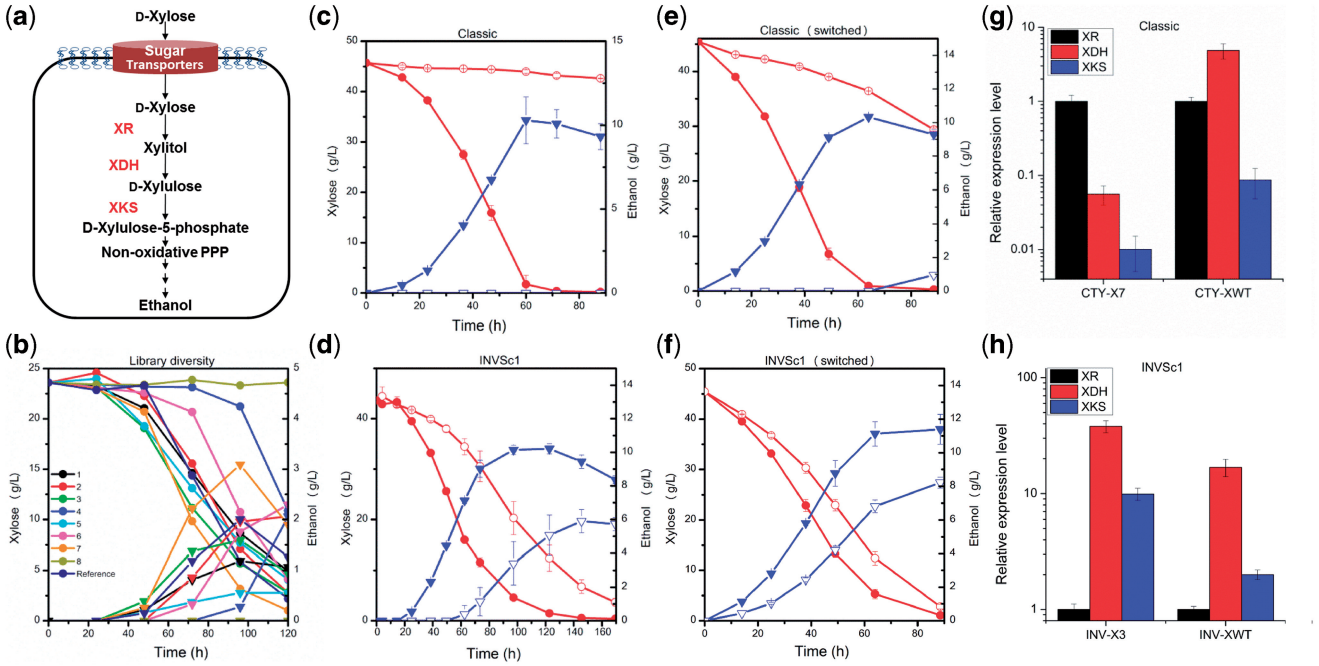

Figure 2. Creation of a highly efficient xylose utilizing pathway in both laboratory and industrial *S. cerevisiae* strains via COMPACTER. Red circle: xylose; blue triangle: ethanol. (**a**) Scheme of the engineered fungal xylose utilizing pathway. (**b**) Xylose fermentation of eight randomly picked clones from the COMPACTER library. (**c**) Comparison of xylose consumption and ethanol production in the industrial strain. Open symbol: the reference industrial strain CTY-XWT; solid symbol: the mutant industrial strain CTY-X7. (**d**) Comparison of xylose consumption and ethanol production in the laboratory strain. Open symbol: the reference laboratory strain INV-XWT; solid symbol: the mutant laboratory strain INV-X3. (**e**) Pathway switching in the industrial strain. Open symbol: the industrial strain harboring the pathway from the mutant laboratory strain CTY-X3; solid symbol: the mutant industrial strain CTY-X7. (**f**) Pathway switching in the laboratory strain. Open symbol: the laboratory strain harboring the pathway from the mutant industrial strain INV-X7; solid symbol: the mutant laboratory strain INV-X3. (**g**) Relative expression levels of XR, XDH and XKS in the mutant industrial strain CTY-X7 and the reference industrial strain CTY-XWT measured using qPCR. The expression levels were normalized by making the expression level of XR as 1. (**h**) Relative expression levels of XR, XDH and XKS in the mutant laboratory strain INV-X3 and the reference laboratory strain INV-XWT measured using qPCR. The expression levels were normalized by making the expression level of XR as 1.

To confirm the diversity of the library, eight colonies from the laboratory yeast strain INVSc1 were randomly picked for analysis. For these colonies, xylose consumption rate, ethanol productivity and ethanol yield were all different, indicating that the resultant pathway library exhibited excellent diversity (Figure 2b). The library was then screened using a colony size-based screening method (Supplementary Figures S4–7) and fast xylose utilizing mutant pathways were identified for both the laboratory and industrial strains (Table 1, Figures 2c and d). For the xylose utilizing pathway optimized in the laboratory strain INVSc1, the optimized mutant laboratory strain INV-X3 improved the ethanol yield by more than 60% (0.25 g/g xylose for INV-X3 versus 0.16 g/g xylose for the reference laboratory strain INV-XWT that harbors the same xylose utilizing pathway under the control of wild-type promoters). Additionally, the INV-X3 laboratory strain consumed xylose 70% faster and produced ethanol 1.5 times faster than the reference laboratory strain INV-XWT (0.4 g/L/h xylose consumption rate and 0.1 g/L/h ethanol production rate; Table 2). More impressively, after only one round of COMPACTER in the industrial strain 'Classic Turbo Yeast', the optimized mutant industrial strain CTY-X7 exhibited a xylose consumption rate of 0.92 g/L/h with an ethanol yield of 0.26 g/g xylose, which is close to the fastest xylose utilizing strain reported in literature (27) (Table 2, Figures 2c). In contrast, the reference industrial strain harboring the same xylose utilizing pathway under wild-type promoters (CTY-XWT) consumed less than 9% of the total xylose and produced no ethanol within 88 h.

The top 10 optimized xylose utilizing laboratory or industrial strains were then isolated and the expression levels of XR, XDH and XKS were determined using quantitative PCR (qPCR). The mutant laboratory strain INV-X3 exhibited a higher expression ratio of XDH/XR and XKS/XR when compared to the reference laboratory strain INV-XWT (more than 1-fold higher XDH/XR ratio and almost 4-fold higher XKS/XR ratio; Figure 2h). Moreover, the top 10 mutant laboratory strains all exhibited higher relative expression ratio of XKS over XR ($4.1 < XKS/XR < 19.3$) when compared with the reference laboratory strain INV-XWT ($XKS/XR \sim 2$), while nine out of the top 10 mutant laboratory strains also exhibited a higher relative expression ratio of XR over XDH (Supplementary Figure S14a). Similarly, the top 10 optimized industrial strains also exhibited very distinct expression levels of XR, XDH and XKS compared to the reference industrial strain CTY-XWT. The top 10 mutants all exhibited a much lower relative expression

Table 1. List of strains and relevant xylose and cellobiose utilizing pathways used in this study

Strain	Host	Pathway	Note
INV-XWT	INVSc1	XWT: xylose pathway under wild type promoters (PDC1p for *csXR*, TEF1p for *ctXDH*, ENO2p for *ppXKS*)	Reference laboratory strain, xylose pathway
INV-X3	INVSc1	X3: Mutant xylose pathway optimized in the INVSc1 strain	Laboratory strain with an optimized xylose pathway
INV-X7	INVSc1	X3: Mutant xylose pathway optimized in the Classic Turbo Yeast strain	Laboratory strain with a switched xylose pathway
CTY-XWT	Classic Turbo Yeast	XWT: xylose pathway under wild type promoters	Reference industrial strain, xylose pathway
CTY-X3	Classic Turbo Yeast	X3: Mutant xylose pathway optimized in the INVSc1 strain	Industrial strain with a switched xylose pathway
CTY-X7	Classic Turbo Yeast	X3: Mutant xylose pathway optimized in the Classic Turbo Yeast strain	Industrial strain with an optimized xylose pathway
INV-CWT	INVSc1	XWT: cellobiose pathway under wild type promoters (PDC1p for *gh1-1*, ENO2p for *cdt-1*)	Reference laboratory strain, cellobiose pathway
INV-C3	INVSc1	X3: Mutant cellobiose pathway optimized in the INVSc1 strain	Laboratory strain with an optimized cellobiose pathway
INV-C59	INVSc1	X3: Mutant cellobiose pathway optimized in the Classic Turbo Yeast strain	Laboratory strain with switched cellobiose pathway
CTY-CWT	Classic Turbo Yeast	XWT: cellobiose pathway under wild type promoters (PDC1p for *gh1-1*, ENO2p for *cdt-1*)	Reference industrial strain, cellobiose pathway
CTY-C3	Classic Turbo Yeast	X3: Mutant cellobiose pathway optimized in the INVSc1 strain	Industrial strain with a switched cellobiose pathway
CTY-C59	Classic Turbo Yeast	X3: Mutant cellobiose pathway optimized in the Classic Turbo Yeast strain	Industrial strain with an optimized cellobiose pathway

DNA sequences of promoter mutants in the optimized pathways are listed in Supplementary Information S1.

Table 2. Xylose fermentation profiles of the optimized and reference strains compared with top xylose fermenting strains in literature

Host strain	Laboratory strain (INVSc1)		Industrial strain (Classic Turbo Yeast)					D452-2	IR-2	IR-2
Seed culture	INV-XWT SCD[a]	INV-X3 SCD[a]	CTY-XWT YPD[b]	CTY-XWT YPX[c]	CTY-X7 YPD[b]	CTY-X7 YPX[c]	CTY-X7 YPX[c]	DA24-16 (27) YPD[b]	MA-R4 (31) YPD[b]	MA-R5 (31) YPD[b]
Initial OD	1	1	10	2	10	2	10	1	2.8 g DCW/L	2.8 g DCW/L
Xylose consumption rate (g xylose/L/h)	0.24	0.40	0.06	0.03	0.74	0.73	0.92	1.06[d]	1.07	1.29[d]
Ethanol productivity (g ethanol/L/h)	0.04	0.10	0	0	0.17	0.17	0.24	0.35[d]	0.36	0.50[d]
Yield (g ethanol/g xylose)	0.15	0.25	0	0	0.24	0.23	0.26	0.31 ~ 0.33[d]	0.34	0.37[d]

The reference strains are csXR, ctXDH and ppXKS driven by wild-type PDC1, TEF1 and ENO2 promoters in corresponding strains.
[a]SC-Ura supplemented with 2% glucose.
[b]YP medium supplemented with 2% glucose.
[c]YP medium supplemented with 2% xylose.
[d]Calculated approximately according to figure shown in corresponding manuscript.

ratio of XDH/XR and XKS/XR when compared to the reference industrial strain CTY-XWT (Supplementary Figure S14b). Specifically, the CTY-X7 industrial strain exhibited a relative expression ratio of XDH/XR almost two orders of magnitude lower when compared with the reference industrial strain CTY-XWT and an expression ratio of XKS/XR more than eight times lower (Figure 2g).

Creating a highly efficient cellobiose utilizing pathway in both laboratory and industrial *Saccharomyces cerevisiae* strains

To demonstrate the generality of the COMPACTER approach for pathway engineering, we sought to optimize the cellobiose utilizing pathway that we recently constructed (28). Cellobiose, a disaccharide with two glucose molecules linked in a $\beta(1 \rightarrow 4)$ bond, is an intermediate product from the enzymatic hydrolysis of cellulose (28). The cellobiose utilizing pathway consists of a cellobiose transporter and an intracellular β-glucosidase for the uptake and complete hydrolysis of cellobiose (Figure 3a). Fermentation of cellobiose rather than glucose represents a novel strategy for sugar utilization that has been shown to alleviate glucose repression and improve the co-fermentation of mixed sugars (27,28).

To optimize the cellobiose utilizing pathway, the cellobiose transporter (CDT) gene (*cdt-1*) and the β-glucosidase (BGL) gene from *Neurospora crassa* (*gh1-1*)

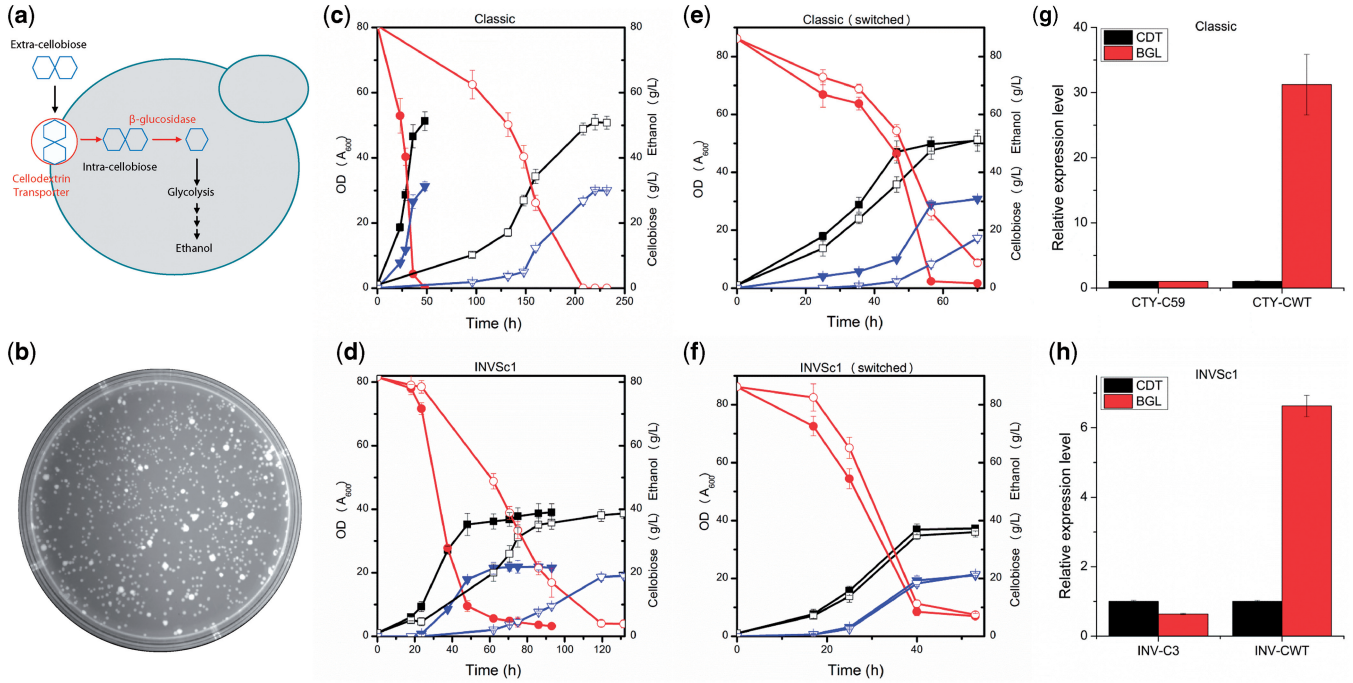

Figure 3. Creation of a highly efficient cellobiose utilizing pathway in both laboratory and industrial *S. cerevisiae* strains via COMPACTER. Red circle: cellobiose; black square: OD (A_{600}); blue triangle: ethanol. (**a**) Scheme of the engineered cellobiose utilizing pathway. (**b**) Library screening on an YPAC agar plate. (**c**) Comparison of cellobiose consumption and ethanol production in 250 ml flask fermentations in the industrial strain. Open symbol: the reference industrial strain CTY-CWT; solid symbol: the mutant industrial strain CTY-C59. (**d**) Comparison of cellobiose consumption and ethanol production in 250 ml flask fermentations in the laboratory strain. Open symbol: the reference laboratory strain INV-CWT; solid symbol: the mutant laboratory strain INV-C3. (**e**) Pathway switching in the industrial strain. Open symbol: the industrial strain harboring the pathway from the mutant laboratory strain CTY-C3; solid symbol: the mutant industrial strain CTY-C59. (**f**) Pathway switching in the laboratory strain. Open symbol: the laboratory strain harboring the pathway from the mutant industrial strain INV-C59; solid symbol: the mutant laboratory strain INV-C3. (**g**) Relative expression levels of CDT and BGL in the mutant industrial strain CTY-C59 and the reference industrial strain CTY-CWT measured using qPCR. The expression levels were normalized by making the expression level of CDT as 1. (**h**) Relative expression levels of CDT and BGL in the mutant laboratory strain INV-C3 and the reference laboratory strain INV-CWT measured using qPCR. The expression levels were normalized by making the expression level of CDT as 1.

were assembled into a single copy expression vector under the control of the ENO2 and PDC1 promoters, respectively. A library of cellobiose utilizing pathways derived from combinations of 10 ENO2 promoter mutants and 11 PDC1 promoter mutants were assembled in the industrial and laboratory *S. cerevisiae* strains. Similar to the above-mentioned xylose utilizing strains, the resultant mutant strains exhibited drastically different cellobiose fermentation abilities due to dissimilar expression levels of the sugar transporter and β-glucosidase (Supplementary Figure S9). The strains harboring the pathway library were screened using a colony size-based screening method (Supplementary Figures S10–13) and fast cellobiose utilizing mutant strains were identified for both laboratory and industrial strains (Table 1 and Figures 3c and d). For the 'Classic Turbo Yeast' industrial strain, the optimized mutant industrial strain CTY-C59 exhibited a 5.4-fold higher cellobiose consumption rate and a 5.3-fold higher ethanol productivity of 0.74 g/L/h, when compared to the reference industrial strain CTY-CWT, which harbored the same cellobiose utilizing pathway under the control of the wild-type promoters (0.39 g/L/h cellobiose consumption rate and 0.14 g/L/h ethanol production rate). Similarly, for the INVSc1 laboratory strain, the optimized mutant laboratory strain INV-C3 exhibited a 2.1-fold higher cellobiose consumption rate (1.50 g/L/h) and a 2.3-fold higher ethanol productivity (0.37 g/L/h) when compared to the reference laboratory strain INV-CWT that harbored the same cellobiose utilizing pathway under the control of the wild-type promoters (Table 3).

After analyzing the promoter mutants present in all optimized strains, it was observed that all of the five cellobiose utilizing mutant pathways in the INVSc1 strain are identical, consisting of the same ENO2 and PDC1 promoter mutants. Eight of the 10 cellobiose utilizing mutant pathways in the 'Classic Turbo Yeast' strain contained the same ENO promoter mutant. The expression levels of *cdt-1* and *gh1-1* in the top 10 optimized pathways of the industrial strain and the top five optimized pathways in the laboratory strain were also determined and compared to the reference pathways. The optimized mutant laboratory strain INV-C3 exhibited a relative expression ratio of BGL/CDT one order of magnitude lower than the reference laboratory strain (BGL/CDT ~ 0.6 for INV-C3 and BGL/CDT ~ 6.6 for INV-CWT; Figure 3h). Similarly, the optimized mutant industrial strains all exhibited a relative BGL/CDT expression ratio much lower than the reference industrial strain CTY-CWT (optimized strains $0.5 < \text{BGL/CDT} < 1.8$, reference strain BGL/CDT ~ 31; Supplementary Figure S15).

Table 3. Summary of cellobiose fermentation profiles

	(Classic)-125		(Classic)-250		(INVSc1)-125		(INVSc1)-250		D452-2
	CTY-CWT	CTY-C59	CTY-CWT	CTY-C59	INV-CWT	INV-C3	INV-CWT	INV-C3	Ha *et al.* (27)
Cellobiose consumption rate (g cellobiose/L/h)	0.36	1.60	0.39	2.18	0.60	1.54	0.7	1.5	1.67
Ethanol productivity (g ethanol/L/h)	0.14	0.65	0.14	0.74	0.16	0.51	0.16	0.37	0.7
Yield (g ethanol/g cellobiose)	0.42	0.44	0.37	0.39	0.32	0.37	0.23	0.27	0.42

Two different shake-flasks, 125 and 250 ml, were used in fermentations. The fermentation behavior of the mutant strains constructed in this study was compared to the fastest cellobiose fermenting strain reported in the literature.

Specifically, the CTY-C59 strain exhibited a relative expression ratio of BGL/CDT more than 30 times lower when compared to the reference industrial strain CTY-CWT (Figure 3g).

Engineered xylose or cellobiose utilizing pathways are host-specific

During the sequence analysis of optimized xylose or cellobiose utilizing pathways, we noticed that the optimized expression patterns of pathways with identical metabolic genes in different strain backgrounds differed drastically. To investigate whether the engineered xylose or cellobiose utilizing pathways are dependent on the strain backgrounds, the most highly optimized mutant pathways found in the laboratory and industrial strains were exchanged. The resulting pathways showed different fermentation performances in a new strain background, indicating that the optimized pathways are strain-specific (Figures 2e and f and 3e and f). Furthermore, the relative expression levels of genes involved in the target pathways measured using qPCR also confirmed the optimal expression profile of the same heterologous pathway was indeed different in distinct strain backgrounds (Figures 2g and h and 3g and h).

As indicated by the pathway exchange experiment and the qPCR measurement, the pathways generated using COMPACTER were truly strain-specific. Surprisingly, the optimal expression pattern for a balanced metabolic flux through the xylose utilizing pathway differed drastically in laboratory and industrial strains. In the laboratory strain, a higher expression level of XDH than XR was required for efficient xylose assimilation. In the industrial strain, however, a higher expression level of XR than XDH was optimal. The distinct expression patterns of optimized pathways indicated the context-dependency of the heterologous pathways in different metabolic backgrounds.

DISCUSSION

Engineering microorganisms to produce chemicals and fuels by introducing heterologous metabolic pathways is a conceptually simple task (1–4), yet the challenge of balancing metabolic flux through these pathways in order to maximize product titers and productivities remains overwhelming (6–11). Previously, two groups attempted to optimize a heterologous metabolic pathway using a combinatorial approach (29,30). However, both studies were limited to a very small library of metabolic pathways, one in a prokaryotic system (29) and the other in yeast using an inefficient *in vitro* cloning method (30). In contrast, the high efficiency of the COMPACTER method enables the creation of a large library of multi-step metabolic pathways through a single yeast transformation, which significantly increases the chance of obtaining an optimized pathway with balanced metabolic flux. In addition, unlike the previous study, in which a single promoter library was used to study the biological effect of varying expression levels of a target gene in a pathway (26), the COMPACTER method creates a library of metabolic pathways using a pool of promoters for each pathway gene so that the metabolic flux of the entire pathway can be rapidly optimized.

Herein, we have demonstrated that the COMPACTER method is an efficient approach to tailor-make pathways for biofuels production from lignocellulosic biomass independent of knowledge concerning previous relevant metabolic engineering studies. In a single round, a recombinant xylose utilizing industrial strain was constructed that displayed 69% of the xylose consumption rate of the fastest reported xylose utilizing strain (making it ranked 4th fastest overall; Table 2). Similarly, after a single round of COMPACTER, a recombinant cellobiose utilizing industrial strain was constructed that exhibited the highest cellobiose consumption rate and ethanol productivity ever reported in literature (Table 3).

The use of the newly developed DNA assembler method enabled the rapid creation of a mutant pathway library with varying gene expression levels through a single step of yeast transformation (19). The efficiency and diversity of the mutant pathway library achieved by DNA assembler was previously found to be excellent by DNA-sequencing analysis of 40 randomly selected individual constructs (32). In this study, the diversity of the pathway library was further investigated by checking the fermentation performance of randomly selected pathway mutants. The distinct fermentation behaviors of the mutant pathways from both the xylose and cellobiose utilizing pathway libraries indicated that the overall library diversity was also excellent (Figure 2b and Supplementary Figure S9).

Both the xylose and cellobiose utilizing pathway libraries were screened using a colony size-based screening method to identify positive mutants with desired phenotype. Since both of the target phenotypes can be linked

with cell viability on certain type of carbon sources, the colony size on an agar plate with a specific sugar as a sole carbon source gives a very good indication of the efficiency of the mutant pathway (Figure 3b). A set of large colonies were isolated and a culture tube-based screening was then performed to identify mutant pathways with improved fermentation ability. During the screening of 50 large colonies isolated from the xylose utilizing mutant pathways in the 'Classic Turbo Yeast' strain, we confirmed that the specific growth rates of the strains correlated very well with the xylose consumption rates and ethanol production (Supplementary Figure S5).

After isolation of the positive mutant pathways, the top mutant pathways for each target pathway were analyzed to determine the optimized expression pattern. For the xylose utilizing pathway, the fact that not all the pathway steps in our target pathways were equally sensitive to the change of promoters indicated that there were multiple points of optimality (Supplementary Figure S14). In addition, we noticed from our study that the relative expression levels (the ratio of expression levels) affected the performance of the pathway more than the absolute expression levels. To confirm that the isolated mutant pathway was truly optimal, the PDC1 promoter in the INV-X3 mutant was replaced by all other nine PDC1 promoter mutants. As we expected, the nine altered constructs all showed lower specific growth rates and xylose consumption rates compared to the INV-X3 strain, which indicated that the optimal mutant pathway identified was truly optimal (Supplementary Figure S19). For the cellobiose utilizing pathway, all the top five mutant pathways isolated from the INVSc1 strain contained the same ENO2 and PDC1 promoter mutants. The fact that the highly diversified pathway library converged into a single solution after screening indicated that the selected pathway was highly optimized. For the 'Classic Turbo Yeast' strain, eight of the 10 final selected strains contained the same ENO2 promoter mutant, whereas the top 10 strains performed similarly even though different PDC promoters were present in the mutant pathways. This also indicated there were multiple points of optimality. However, when the top three mutant pathways from the initial screening were analyzed more carefully in shake flasks, the mutant pathway CTY-C59 outperformed the other two mutant pathways. This clearly showed that switching one promoter in the optimized pathway will yield a sub-optimal pathway, which supports the conclusion that the selected pathway was indeed optimized (Supplementary Figure S11).

The highly mobile pathway libraries constructed using the DNA assembler method enabled rapid pathway construction and optimization in different strain backgrounds. As a result, optimization from the same pathway libraries in the laboratory and industrial yeast strains yielded different optimized pathways for both the xylose and cellobiose utilizing pathways. The distinct optimized pathways for the laboratory and industrial strains were confirmed by exchanging the plasmids harboring the optimal pathways and qPCR. As expected, the optimized strains all outperformed the switched strains, indicating that the optimized pathways were highly host specific. The effect of pathway switching was especially drastic for the xylose utilization pathway: the CTY-X3 strain exhibited a very poor xylose fermentation ability compared to the CTY-X7 strain (Table 1 and Figure 2e). qPCR analysis of these strains indicated the optimal XR to XDH expression ratios were very different in the laboratory and industrial strains (Figures 2g and h). Although the context-dependent nature of heterologous pathways in different host strains has been reported (17), this is the first comprehensive and quantitative report of such context-dependency using two important heterologous pathways for lignocellulosic biofuel production. The dependence of pathway optimization on a particular strain may result from differences in expression levels of endogenous genes involved in the pathway, different behavior of the same promoter under different strain backgrounds, the overall availability of cofactors or stress responses. Since the promoter mutants were characterized only in the laboratory strain, qPCR was performed to check the expression levels of same pathways in different strain backgrounds. The xylose utilizing mutant pathway optimized in the 'Classic Turbo Yeast' strain (X7) exhibited similar expression ratios in both the INVSc1 (INV-X7) and 'Classic Turbo Yeast' strains (CTY-X7). The expression ratios of the INV-X7 and CTY-X7 strains were completely different from the CTY-XWT and CTY-X3 strains. The fact that CTY-X7 and INV-X7 exhibited a similar expression ratio while CTY-WT, CTY-X3 and CTY-X7 exhibited very different expression ratios indicated that the expression ratio difference was more affected by the promoter mutants in the pathway than the background of the host strain (Table 1 and Supplementary Figure S16). The enzyme activity levels in the different xylose utilizing strains were also analyzed to determine the origin of the host-specificity of the optimized pathways. The optimized pathways showed very different enzyme activities compared to the control pathways consisting of wild-type promoters. In Supplementary Figure S14, the optimized pathway in the INVSc1 strain (INV-X3) would require a higher expression of XDH while the optimized pathway in 'Classic Turbo Yeast' strain (CTY-X7) would require a higher expression of XR. The enzyme activity results were consistent with the expression level: the CTY-X7 strain did exhibit a higher XR activity and lower XDH activity compared to the INV-X3 strain, which would result in a higher XR to XDH ratio (Supplementary Figure S17). The enzyme activity results indicated that the optimization of the pathway not only happened at the expression level, but it also affected the enzyme activities in the resulted strains. In other words, the optimized pathway that performed better in a specific strain background was due to the interaction of optimized promoters and the metabolic background of the host strain. The introduction of promoter mutants simultaneously at multiple steps in the pathway altered the overall expression levels, which consequently changed the translation level and protein folding of the heterologous enzymes. As a result, the simultaneous optimization made the heterologous pathway fit better in the specific strain background and generated recombinant strains with

better performance. The fact that the optimized strains differed in so many different aspects further demonstrated the power of the strategy described in this study— the power of rapid generation and examination of a large number of mutant pathways through a single step of *in vivo* pathway optimization (Supplementary Figures S17 and 18). Within this context, it may be very difficult to reverse engineer an industrial *S. cerevisiae* production strain using knowledge gained from the engineering of the simplified model strains used in laboratories. As a corollary, the ability of COMPACTER to tailor-make metabolic pathways rapidly in different strain backgrounds is truly advantageous in pathway engineering.

Finally, due to its simplicity and versatility, the COMPACTER method can be used for simultaneous optimization of gene expression levels in any metabolic pathway. It can rapidly optimize microbial cell factories for the production of many useful compounds such as fuels, platform chemicals and pharmaceuticals. In addition, it can be extremely effective at generating metabolic pathway libraries, gene circuits, gene switches, synthetic regulatory networks and even artificial chromosomes with varying expression patterns for applications in metabolic engineering and synthetic biology.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–19, Supplementary Methods 1–4 and Supplementary Information 1.

ACKNOWLEDGEMENTS

J. Du acknowledges both the Chia-chen Chu graduate fellowship from the School of Chemical Sciences and the Henry Drickamer Fellowship support from the Department of Chemical and Biomolecular Engineering at the University of Illinois.

FUNDING

Funding for open access charge: Energy Biosciences Institute.

Conflict of interest statement. None declared.

REFERENCES

1. Du,J., Shao,Z. and Zhao,H. (2011) Engineering microbial factories for synthesis of value-added products. *J. Ind. Microbiol. Biotechnol.*, **38**, 873–890.
2. Liao,J.C., Atsumi,S. and Hanai,T. (2008) Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature*, **451**, 86–89.
3. Pfeifer,B., Ajikumar,P.K., Xiao,W.H., Tyo,K.E.J., Wang,Y., Simeon,F., Leonard,E., Mucha,O., Phon,T.H. and Stephanopoulos,G. (2010) Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. *Science*, **330**, 70–74.
4. Keasling,J.D., Ro,D.K., Paradise,E.M., Ouellet,M., Fisher,K.J., Newman,K.L., Ndungu,J.M., Ho,K.A., Eachus,R.A., Ham,T.S. *et al*. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, **440**, 940–943.
5. Pfleger,B.F., Pitera,D.J., Smolke,C.D. and Keasling,J.D. (2006) Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nat. Biotechnol.*, **24**, 1027–1032.
6. Dueber,J.E., Wu,G.C., Malmirchegini,G.R., Moon,T.S., Petzold,C.J., Ullal,A.V., Prather,K.L. and Keasling,J.D. (2009) Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.*, **27**, 753–759.
7. Salis,H.M., Mirsky,E.A. and Voigt,C.A. (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.*, **27**, 946–950.
8. Anthony,J.R., Anthony,L.C., Nowroozi,F., Kwon,G., Newman,J.D. and Keasling,J.D. (2009) Optimization of the mevalonate-based isoprenoid biosynthetic pathway in *Escherichia coli* for production of the anti-malarial drug precursor amorpha-4,11-diene. *Metab. Eng.*, **11**, 13–19.
9. Stephanopoulos,G. and Luetke-Eversloh,T. (2008) Combinatorial pathway analysis for improved L-tyrosine production in *Escherichia coli*: identification of enzymatic bottlenecks by systematic gene overexpression. *Metab. Eng.*, **10**, 69–77.
10. Bond-Watts,B.B., Bellerose,R.J. and Chang,M.C.Y. (2011) Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nat. Chem. Biol.*, **7**, 222–227.
11. Liao,J.C., Shen,C.R., Lan,E.I., Dekishima,Y., Baez,A. and Cho,K.M. (2011) Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. *Appl. Environ. Microbiol.*, **77**, 2905–2915.
12. Bujara,M. and Panke,S. (2010) Engineering in complex systems. *Curr. Opin. Biotechnol.*, **21**, 586–591.
13. Alper,H. and Stephanopoulos,G. (2007) Global transcription machinery engineering: a new approach for improving cellular phenotype. *Metab. Eng.*, **9**, 258–267.
14. Warnecke,T.E., Lynch,M.D., Karimpour-Fard,A., Lipscomb,M.L., Handke,P., Mills,T., Ramey,C.J., Hoang,T. and Gill,R.T. (2010) Rapid dissection of a complex phenotype through genomic-scale mapping of fitness altering genes. *Metab. Eng.*, **12**, 241–250.
15. Warner,J.R., Reeder,P.J., Karimpour-Fard,A., Woodruff,L.B.A. and Gill,R.T. (2010) Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. *Nat. Biotechnol.*, **28**, 856–862.
16. Wang,H.H., Isaacs,F.J., Carr,P.A., Sun,Z.Z., Xu,G., Forest,C.R. and Church,G.M. (2009) Programming cells by multiplex genome engineering and accelerated evolution. *Nature*, **460**, 894–898.
17. Matsushika,A., Inoue,H., Murakami,K., Takimura,O. and Sawayama,S. (2009) Bioethanol production performance of five recombinant strains of laboratory and industrial xylose-fermenting *Saccharomyces cerevisiae*. *Bioresour. Technol.*, **100**, 2392–2398.
18. Gherardi,E. and Zaccolo,M. (1999) The effect of high-frequency random mutagenesis on in vitro protein evolution: a study on TEM-1 beta-lactamase. *J. Mol. Biol.*, **285**, 775–783.
19. Shao,Z., Zhao,H. and Zhao,H. (2009) DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res.*, **37**, e16.
20. Jin,Y.S., Ni,H.Y., Laplaza,J.M. and Jeffries,T.W. (2003) Optimal growth and ethanol production from xylose by recombinant *Saccharomyces cerevisiae* require moderate D-xylulokinase activity. *Appl. Environ. Microbiol.*, **69**, 495–503.
21. Van Vleet,J.H. and Jeffries,T.W. (2009) Yeast metabolic engineering for hemicellulosic ethanol production. *Curr. Opin. Biotechnol.*, **20**, 300–306.
22. Hahn-Hagerdal,B., Karhumaa,K., Fonseca,C., Spencer-Martins,I. and Gorwa-Grauslund,M.F. (2007) Towards industrial pentose-fermenting yeast strains. *Appl. Microbiol. Biotechnol.*, **74**, 937–953.
23. Karhumaa,K., Hahn-Hagerdal,B. and Gorwa-Grauslund,M.F. (2005) Investigation of limiting metabolic steps in the utilization of xylose by recombinant *Saccharomyces cerevisiae* using metabolic engineering. *Yeast*, **22**, 359–368.
24. Sedlak,M., Bera,A.K., Ho,N.W.Y. and Khan,A. (2011) A genetic overhaul of *Saccharomyces cerevisiae* 424 A(LNH-ST) to improve xylose fermentation. *J. Ind. Microbiol. Biotechnol*, **38**, 617–626.
25. Hector,R.E., Qureshi,N., Hughes,S.R. and Cotta,M.A. (2008) Expression of a heterologous xylose transporter in a *Saccharomyces cerevisiae* strain engineered to utilize xylose improves aerobic xylose consumption. *Appl. Microbiol. Biotechnol.*, **80**, 675–684.

26. Alper,H., Fischer,C., Nevoigt,E. and Stephanopoulos,G. (2005) Tuning genetic control through promoter engineering. *Proc. Natl Acad. Sci. USA*, **102**, 12678–12683.
27. Ha,S.J., Galazka,J.M., Kim,S.R., Choi,J.H., Yang,X.M., Seo,J.H., Glass,N.L., Cate,J.H.D. and Jin,Y.S. (2011) Engineered *Saccharomyces cerevisiae* capable of simultaneous cellobiose and xylose fermentation. *Proc. Natl Acad. Sci. USA*, **108**, 504–509.
28. Li,S.J., Du,J., Sun,J., Galazka,J.M., Glass,N.L., Cate,J.H.D., Yang,X.M. and Zhao,H.M. (2010) Overcoming glucose repression in mixed sugar fermentation by co-expressing a cellobiose transporter and a beta-glucosidase in *Saccharomyces cerevisiae*. *Mol. BioSyst.*, **6**, 2129–2132.
29. Ramon,A. and Smith,H.O. (2011) Single-step linker-based combinatorial assembly of promoter and gene cassettes for pathway engineering. *Biotechnol. Lett.*, **33**, 549–555.
30. Lu,C. and Jeffries,T. (2007) Shuffling of promoters for multiple genes to optimize xylose fermentation in an engineered *Saccharomyces cerevisiae* strain. *Appl. Environ. Microbiol.*, **73**, 6072–6077.
31. Matsushika,A., Inoue,H., Watanabe,S., Kodaki,T., Makino,K. and Sawayama,S. (2009) Efficient bioethanol production by a recombinant flocculent *Saccharomyces cerevisiae* strain with a genome-integrated $NADP^+$-dependent xylitol dehydrogenase Gene. *Appl. Environ. Microbiol.*, **75**, 3818–3822.
32. Du,J. (2011) Metabolic engineering of *Saccharomyces cerevisiae* for efficient ethanol production from pentose sugars, *Ph.D. Thesis*. University of Illinois, Urbana, IL.