

# Modulating Pathway Performance by Perturbing Local Genetic Context

Carmen Lopez, Yuxin Zhao, Rick Masonbrink, and Zengyi Shao\*

Cite This: *ACS Synth. Biol.* 2020, 9, 706–717

Read Online

ACCESS |



Metrics &amp; More



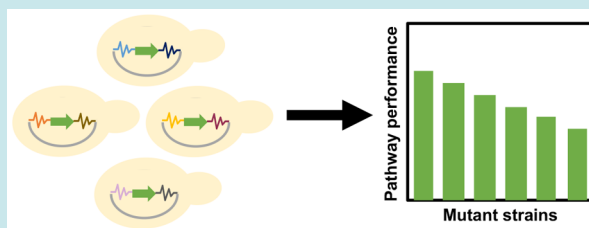
Article Recommendations



Supporting Information

**ABSTRACT:** Combinatorial engineering is a preferred strategy for attaining optimal pathway performance. Previous endeavors have been concentrated on regulatory elements (e.g., promoters, terminators, and ribosomal binding sites) and/or open reading frames. Accumulating evidence indicates that noncoding DNA sequences flanking a transcriptional unit on the genome strongly impact gene expression. Here, we sought to mimic the effect imposed on expression cassettes by the genome. We created variants of the model yeast *Saccharomyces cerevisiae* with significantly improved fluorescence or cellobiose consumption rate by randomizing the sequences adjacent to the GFP expression cassette or the cellobiose-utilization pathway, respectively. Interestingly, nucleotide specificity was observed at certain positions and showed to be essential for achieving optimal cellobiose assimilation. Further characterization suggested that the modulation effects of the short sequences flanking the expression cassettes could be potentially mediated by remodeling DNA packaging and/or recruiting transcription factors. Collectively, these results indicate that the often-overlooked contiguous DNA sequences can be exploited to rapidly achieve balanced pathway expression, and the corresponding approach could be easily stacked with other combinatorial engineering strategies.

**KEYWORDS:** combinatorial pathway engineering, balanced pathway expression, genetic context, modulation of gene expression, yeasts



The use of microorganisms to synthesize valuable compounds ranging from commodity chemicals, such as biofuels and food-additives, to nutraceuticals and pharmaceuticals, is regarded as a green and sustainable strategy. Initial biosynthesis can be established by the expression of a complex pathway that usually involves multiple endogenous and heterologous genes in a selected microbial host; however, production optimization requires a delicate balance of the expression level of all genes in the pathway. High-level production is often accompanied by overconsumption of intracellular resources, thus impairing cell growth. A strategy that can identify the optimum point of interest between vital cellular processes and flux of the recombinant pathway is therefore highly sought-after. Traditional approaches for modulating gene expression are based on the identification of bottlenecks and targeted alleviation in a sequential manner. However, sequential manipulation is time-consuming and provides only narrow coverage of expression combinations, whereas combinatorial pathway engineering enables simultaneous modulation of all the genes at play, thus yielding optimized phenotypes.<sup>1</sup>

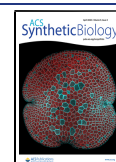
Conventional combinatorial approaches for pathway optimization mainly target regulatory elements (e.g., promoters,<sup>2</sup> terminators,<sup>3</sup> and ribosomal binding sites<sup>4</sup>) and/or open reading frames (ORF).<sup>5</sup> The numerous examples found in the literature suggest that targeting regulatory elements to modulate pathway expression is the prevailing choice, attributed to easy-to-execute protocols.<sup>6</sup> The genetic informa-

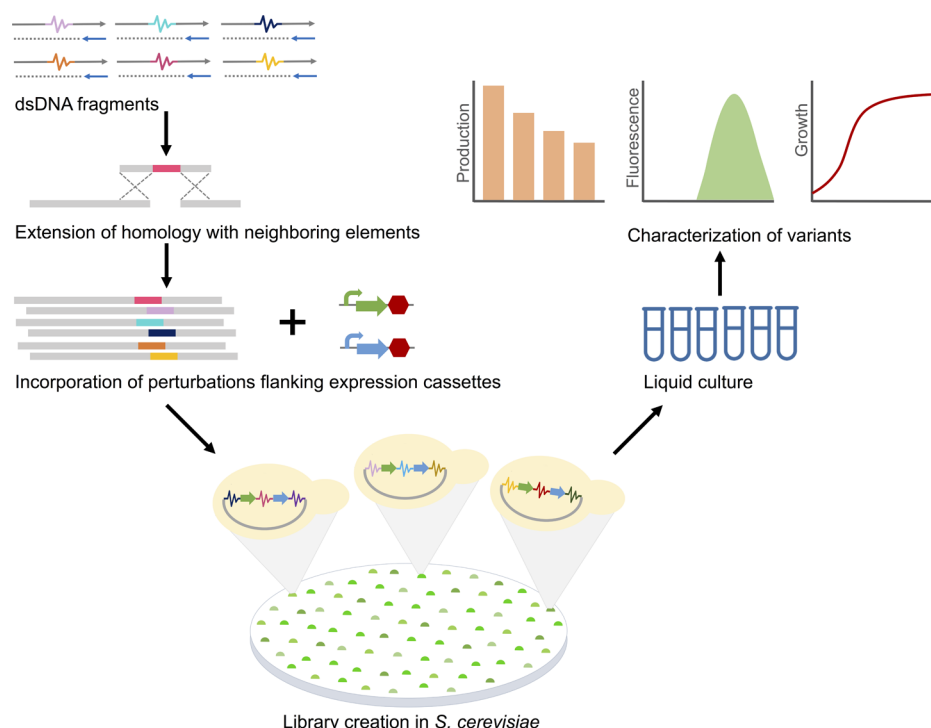
tion on eukaryotic organisms is highly packaged; in order to fit meters-long genomes in cells at the microscale, negatively charged strands of DNA are wrapped around histones in the form of nucleosomes, which are often depicted as “beads on a string”.<sup>7</sup> Genome-wide studies have demonstrated that nucleosome positioning strongly influences gene expression<sup>8,9</sup> and that the positioning pattern is largely dependent on DNA sequences.<sup>10</sup> For example, homopolymeric runs of polyA are elements found in promoters that serve as barriers to the formation of nucleosomes, thus stimulating gene expression.<sup>11–13</sup> In addition, increasing evidence has pointed to new regulator binding sites (e.g., upstream activation sequences<sup>14</sup> or downstream cleavage sites<sup>15</sup>) existing outside previously defined transcriptional units. These observations suggest the potential application of noncoding sequences flanking transcriptional units as new engineering targets.

The purpose of this work was to develop a strategy to rapidly modulate the expression level of multiple genes in a target pathway. Using libraries of sequences flanking the gene expression cassettes, we explored the impact of perturbing the

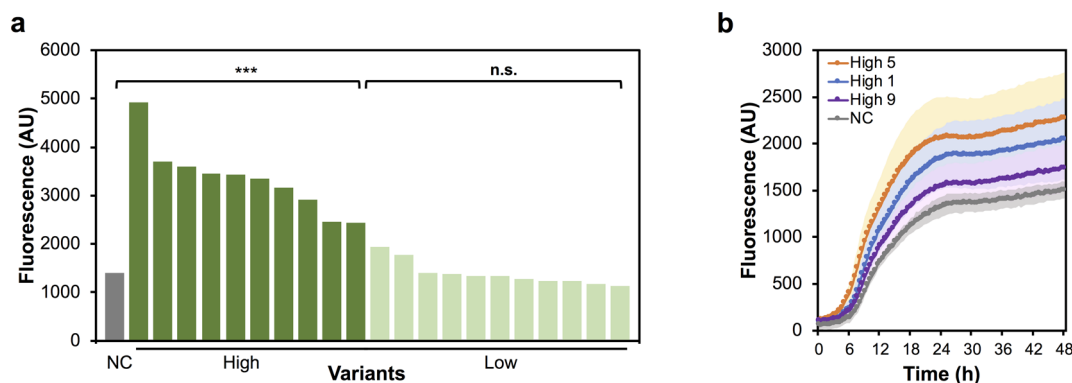
Received: November 2, 2019

Published: March 24, 2020





**Figure 1.** Schematic illustration of introducing perturbation sequences to the local genetic context of a gene or a pathway and identifying the variants with favorable expression dynamics. 90-nt primers were designed to include 30-nt random sequences in the middle, which can be easily converted to double-stranded DNA using a single 20-nt primer annealing at the 3' end. The regions flanking the 30-bp perturbation were lengthened by overlap extension PCR and created longer homologous overlaps with the neighboring fragments. The library of perturbations, gene expression cassettes, and the linearized plasmid backbone were transformed to yeast to form plasmids. Colonies were subjected to the corresponding high-throughput screening. The selected variants were assessed by downstream analyses, e.g., flow cytometry, growth assay, and productivity evaluation.



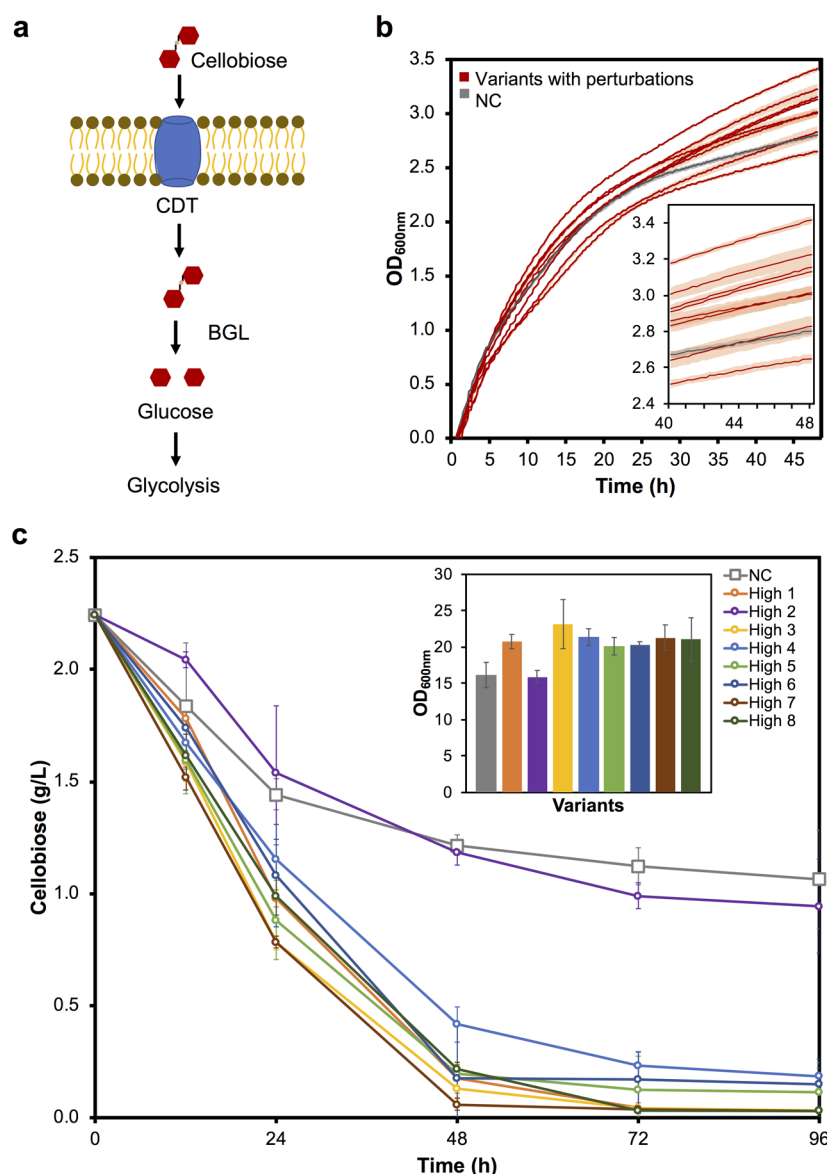
**Figure 2.** Fluorescence profile for strains bearing perturbations flanking the *gfp* expression cassette. (a) The selected strains were analyzed by flow cytometry after 44-h cultivation. A two-sample *t* test was conducted to determine the effect of perturbations on GFP expression. Significant difference (\*\*\*) was observed between the high-expression group and the negative control without a perturbation (NC);  $p < 0.001$ . The mark "n.s." indicates "not significant" as determined by the statistical test. (b) Fluorescence for the verified clones with higher GFP expression was monitored continuously over 48 h. The shaded area on each curve represents the standard deviation of three biological replicates.

local genetic context of the transcriptional units and generated improved phenotypes in *Saccharomyces cerevisiae*. Our results suggest that the modulation effect is highly sequence-specific to both the pathway and the expression carrier and a single mutation could completely abolish the improved phenotype. Nucleosome occupancy analyzed experimentally and by *in silico* modeling suggested that 30-bp is sufficient to remodel chromatin packaging in the promoter region. Such a strong context dependency reinforces the need to incorporate this often-overlooked element (i.e., noncoding DNA) as another engineering target into the design space. The perturbation of

local genetic context therefore serves as a stackable strategy to fulfill the essential task of diversifying the designable elements in the synthetic biology toolbox.

## RESULTS AND DISCUSSION

The perturbation of local genetic context was developed as a novel strategy based on the use of short library-generated primers containing 30 randomized nucleotides (nts) and screening to assess differences in expression compared to a nonperturbed reference. The inspiration behind this method arose from the observation that integration of a gene or a



**Figure 3.** Enhancing cellobiose utilization by perturbing the local genetic context of the pathway. (a) The cellobiose utilizing pathway includes a cellobiose transporter (CDT) and a  $\beta$ -glucosidase (BGL). (b) Growth curves for the perturbed variants selected by colony size according to Fiji. (c) Cellobiose consumption profiles of the improved variants during 96-h fermentation. The inset shows the OD<sub>600</sub> values at 48 h. A two-sample  $t$  test with  $p < 0.05$  showed significant differences between the perturbed strains (except High 2) and the strain bearing the cellobiose utilization pathway without perturbation (NC, negative control). Error bars represent the standard deviation of three biological replicates.

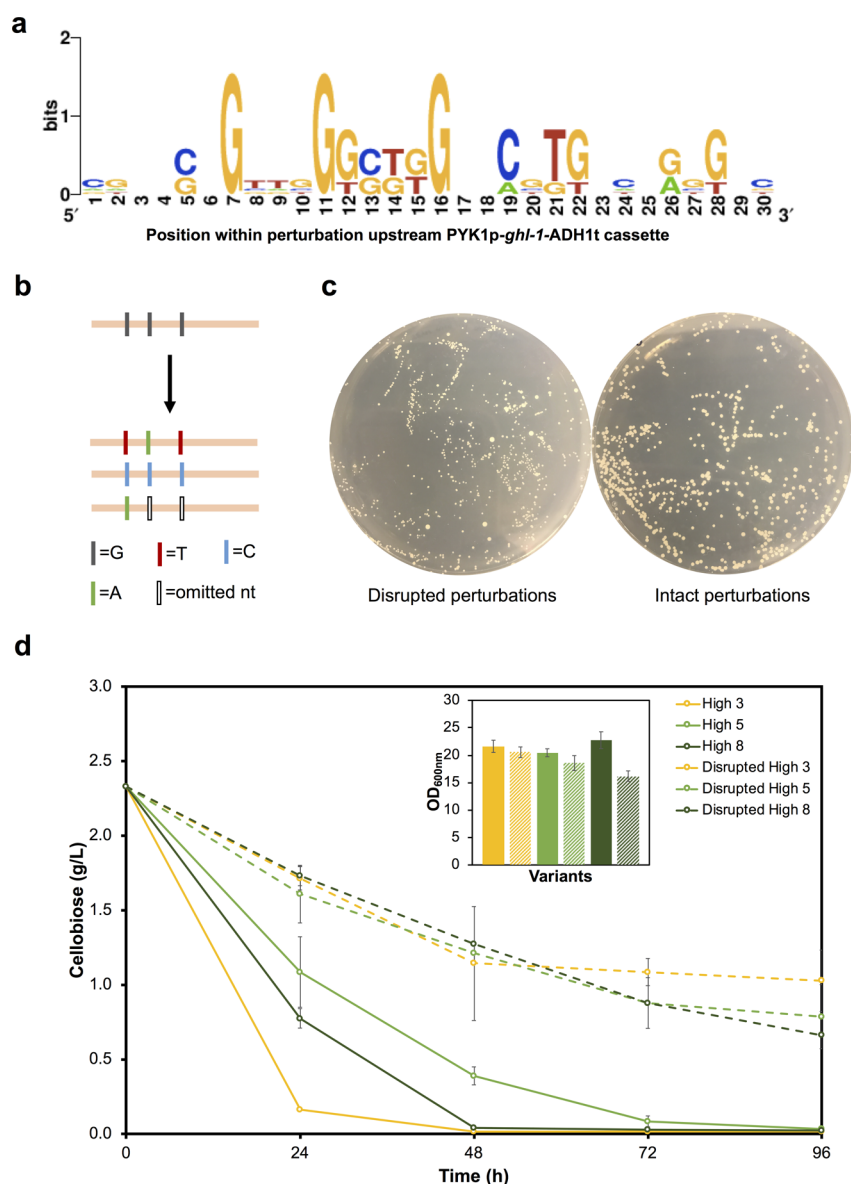
biosynthetic pathway into a genome led to varying levels of expression depending on the integration locus.<sup>16,17</sup> Hence, we hypothesized that random sequences flanking expression cassettes could emulate what happens in the genome, yielding a useful collection of sequences that promote variable gene expression.

Our strategy first utilizes 90-nt oligonucleotides consisting of 30 randomized nts flanked by 30-nt overlaps homologous to the neighboring sequences (Figure 1). These single-stranded oligonucleotides are subsequently converted to double-stranded fragments using regular primers that anneal to the 3' ends. The homology regions flanking the perturbations are lengthened by overlap extension PCR (OE-PCR). The 30-bp random sequences constituting the perturbation, are therefore incorporated at the flanks of transcriptional units, *i.e.*, upstream of promoters and downstream of terminators, and assembled into an episomal plasmid using the DNA assembler approach

that relies on homologous recombination in the model yeast *S. cerevisiae* to efficiently assemble large pathways.<sup>18,19</sup> High-throughput screening of the generated library reveals the variants of interest. Following plasmid recovery, the perturbation sequences are retrieved for further analysis.

#### GFP Expression Impacted by Local Genetic Context.

As a proof of concept, we implemented an expression cassette containing green fluorescent protein (GFP) under the control of a medium-strength yeast promoter and terminator (TEF2p and TEF2t<sup>20</sup>). We hypothesized that using a constitutive promoter would attenuate the dynamic expression variation associated with different growth stages. In order to estimate the variation range of GFP expression yielded using this method, screening was facilitated by the image software FIJI. The variants at the extreme of the GFP spectrum (high and low mean fluorescence) were collected and evaluated using flow cytometry. GFP expression was profiled for the variants of

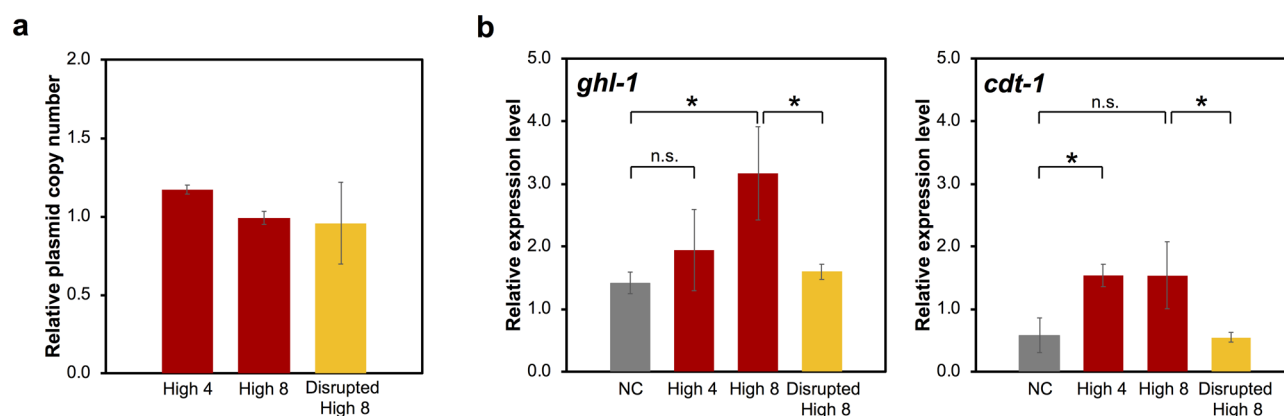


**Figure 4.** Inspection of the perturbation sequences. (a) Sequence logo showcasing the perturbations upstream of the *ghl-1* expression cassette. Positions 7, 11, and 16 show conserved nucleotides in the perturbation sequences analyzed from seven improved variants. (b) Schematic illustrating the point mutations employed to disrupt the original perturbation sequences. (c) The exemplary variants with an intact perturbation and with a disrupted perturbation had different growth rates on the plates. (d) Cellobiose consumption profile during 96-h fermentation. Inset shows the OD<sub>600</sub> values at 48 h. Disruption of perturbations significantly reduced the sugar consumption rates of the three selected variants (two-sample *t* test;  $p < 0.05$ ). Error bars represent the standard deviation of three biological replicates.

interest in Figure 2a. In line with our prediction, library members showed distinct levels of gene expression. Variants classified as high, *i.e.*, the top 10 variants with high mean fluorescence, showed up to a 3.5-fold increase ( $p < 0.001$ ) compared to the nonperturbed control (NC). In contrast, variants with slightly decreased fluorescence (labeled as “low”) showed no significant differences. As often occurs in library screening approaches, promising variants arise as a combined consequence of targeted manipulation, unpredicted background alteration, and the common variation associated with growth at the first site of screening. After retransformation of the isolated plasmids, the GFP expression levels displayed by the variants in the high group generally decreased; however, the enhanced performance compared with the nonperturbed control was largely retained (up to 2-fold increase; Figure S1). To obtain a more complete measurement for expression

performance, the total fluorescence was continuously monitored using a fluorescence spectrometer over 48 h, instead of an end-point assay using flow cytometry analysis that measured the mean fluorescence intensity of biological triplicates (at 44 h). Time-course profiles for the selected variants that showed consistently higher GFP intensity than the control over several rounds of transformation are presented in Figure 2b. The results suggest that the perturbations in these variants caused significant increases in GFP expression regardless of the growth stage, highlighting the ease with which this method is able to modulate the expression of a simple expression cassette. The perturbation sequences were then retrieved by sequencing (summarized in Table S1).

**Effects of Perturbing the Local Genetic Context of the Cellobiose Utilization Pathway.** Predicated on the initial proof of concept, we sought to investigate the utility of



**Figure 5.** Analysis of the plasmid copy numbers and the transcriptional levels in the selected strains. (a) Relative plasmid copy number was determined with respect to *LEU2* in the genome at 24 h. (b) Relative expression level with respect to *ALG9* was determined for the genes comprising the cellobiose utilization pathway at 24 h. The selected perturbations increased the expression levels of both genes. Significant difference (\*) was determined by two-sample test ( $p < 0.05$ ); n.s. stands for “not significant”. Error bars represent the standard deviation of three biological replicates.

fine-tuning the genetic context to balance pathway expression. The cellobiose utilization pathway is comprised of disaccharide transporter encoded by *cdt-1* and a  $\beta$ -glucosidase encoded by *ghl-1*<sup>1,21</sup> (Figure 3a). Cellobiose fermentation has become an attractive strategy for producing value-added chemicals from cellulosic biomass in *S. cerevisiae*, mainly because cellobiose can be co-utilized with xylose without causing carbon catabolite repression.<sup>22</sup> Following manipulations analogous to those used to construct the GFP<sup>+</sup> variants, we heterologously expressed and perturbed the two gene expression cassettes required for cellobiose uptake and assimilation in *S. cerevisiae*. The lack of additional carbon sources in the medium forced the yeast to power its growth *via* cellobiose consumption. Therefore, the effect of the perturbations could be inferred from cellular growth. Promising variants were screened by colony size from the transformation plates and their growth in small batch fermentation was monitored for 48 h. Notably, most of the perturbed CDT<sup>+</sup>BGL<sup>+</sup> variants selected on their large colony size exhibited improved growth profiles compared to the nonperturbed counterpart. As shown in Figure 3b, seven out of the eight perturbed variants obtained after retransforming the isolated DNA accumulated up to 49% more biomass, suggesting that the corresponding 30-bp sequences flanking the two gene expression cassettes enabled more efficient cellobiose utilization. To verify that cellobiose consumption was indeed enhanced by the presence of the certain perturbations, we measured the cellobiose concentration in spent media over the course of 96-h fermentation (Figure 3c). The seven perturbed variants had <20% of the available cellobiose present in the medium at 2 days of culture and cellobiose was nearly depleted at the end of fermentation, which was in stark contrast to the nonperturbed pathway control.

Intrigued by the changes observed in cellobiose utilization, we examined the identity of the 30-bp perturbations. Alignment of the retrieved perturbation sequences revealed a noteworthy trend: a conserved guanine nucleotide was found at positions 7, 11, and 16 in the perturbations located upstream of the *YK1p-ghl-1-ADH1t* cassette (Figure 4a and Table S1). To exclude the possibility that these conserved nucleotides were random, we sought to determine the essentiality of guanine at these specific positions. Thus, we mutated the guanines at positions 7, 11, and 16 in clones High 3, High 5,

and High 8 and assembled new versions of plasmids pRS415-P3-CB, pRS415-P5-CB, and pRS415-P8-CB using new primers (Figure 4b and Table S2). The resulting variants with disrupted perturbations showed drastic changes on plates (Figure 4c); moreover, cellobiose utilization characterized in liquid medium was significantly impaired compared with that of the intact-perturbation counterparts (Figure 4c,d, Table S3). These results provide evidence that the actual sequence, particularly of a few key positions in the perturbations, is crucial for effectively balancing gene expression in a pathway.

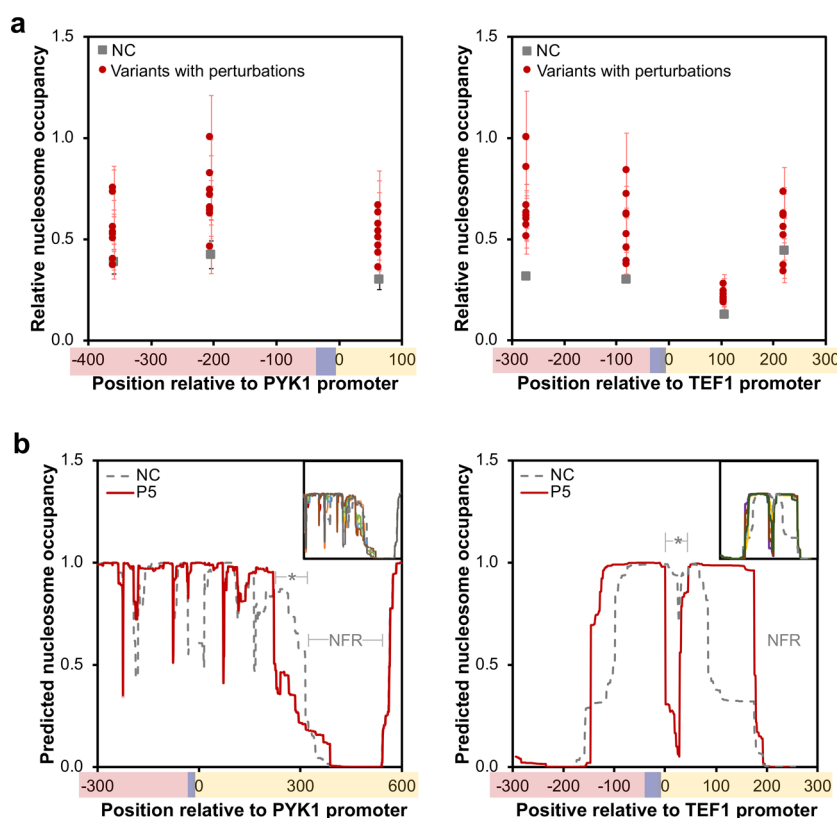
**The Potential Mechanisms for Noncoding Sequences to Exert Modulation Roles.** Understanding how perturbation sequences exert modulation roles to improve strain performance requires biochemical characterization in depth. As of today, the mechanistic studies are scattered and scarce from the perspective of synthetic biology. Although the execution of this perturbation strategy does not require any knowledge on relevant mechanisms, it is important to draw community's attention to the key players and the interesting phenomena that were not well remarked previously. The underlying mechanisms are likely multifaceted considering the diverse characteristics of the perturbation sequences.

**1. Modulating Transcriptional Expression Levels.** The first potential mechanism was revealed by a collection that includes strains High 8, Disrupted High 8, and High 4, the latter of which does not carry the conserved guanines but still has a rapid cellulose consumption rate. A plasmid copy number assay as described by Moriya and colleagues<sup>23</sup> was carried out initially to rule out a gene dosage effect. As shown in Figure 5a, these three strains carry comparable copy numbers of plasmids (1–1.2 copies per cell). We proceeded to analyze the transcriptional levels of *ghl-1* and *cdt-1* and found that strain High 8 had significantly higher *ghl-1* expression and strain High 4 had significantly higher *cdt-1* expression compared to the strain without perturbation (NC) (Figure 5b). Interestingly, for both genes, disrupting the conserved guanines (strain Disrupted High 8) removed the enhanced performance, which was not only reflected by the poor cell growth on cellobiose but also supported by the transcriptional analysis. These results provide direct evidence that perturbations can induce transcriptional modulation; in some perturbation sequences, the beneficial modulation requires specific motifs.

**Table 1.** Prediction of the Transcription Factor-Binding Motifs Based on the Perturbation Sequences Applied to the Cellobiose Pathway<sup>a</sup>

|              | query perturbation | standard name | systematic name | maximum score (%) | DNA binding domain |
|--------------|--------------------|---------------|-----------------|-------------------|--------------------|
| <i>ghl-1</i> | P3                 | <u>RAP1</u>   | YNL216W         | 91.5              | Myb/SANT           |
|              | P5                 | MATALPHA2     | YCR039C         | 100               | Homeodomain        |
|              | P8                 | <u>SUT1</u>   | YGL162W         | 100               | Zinc cluster       |
| <i>cdt-1</i> | P3                 | <u>NSI1</u>   | YDR026C         | 100               | Myb/SANT           |
|              | P5                 | None          | N/A             | N/A               | N/A                |
|              | P8                 | UPC2          | YDR213W         | 89.3              | Zinc cluster       |

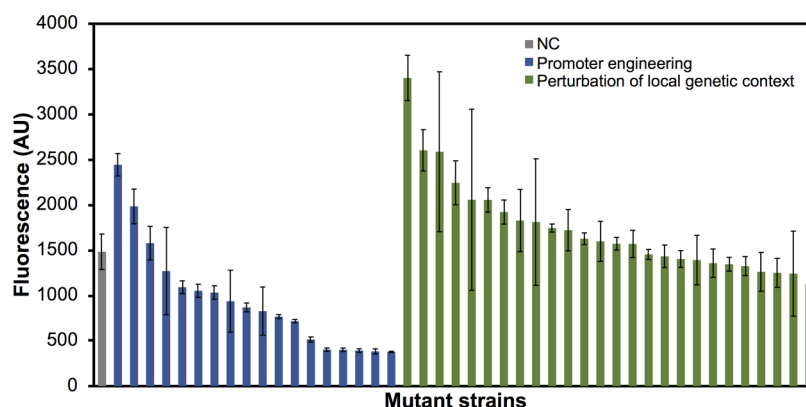
<sup>a</sup>Perturbation sequences listed in Table S1 were scanned using YeTFaSCo. Briefly, sequences were input individually into the scan tool available at <http://yefasco.cabr.utoronto.ca/index.php> where they were analyzed using the *expert curated- no dubious* as motif set. From the resulting table, the transcriptional factor motif with the highest score was reported for each perturbation sequence. Maximum score indicates the likelihood of optimal binding sites that are more like the motif than the background distribution. RAP1, SUT1, and NSI1 are three transcription factors involved in chromatin packaging. None: there was no hit with any transcription factor.



**Figure 6.** The nucleosome occupancy profiles for the 5' upstream sequences of the *PYK1p-ghl-1* and the *TEF1p-cdt-1* expression cassettes. (a) Relative nucleosome occupancy determined experimentally by qPCR and reported with respect to *REC102*, a heterochromatic control. The region relevant to *TEF1p* showed occupancies significantly higher in five perturbed strains in the first two positions, in four perturbed strains in the third position and in one perturbed strain in the fourth position. The region relevant to *PYK1p* showed significantly higher occupancies in four perturbed strains in the first position, in two perturbed strains in the second position, and in one perturbed strain in the third position (two sample-test,  $p < 0.05$ ). (b) Nucleosome occupancy predicted by NuPop, a Hidden Markov Model.<sup>35,36</sup> An exemplary strain bearing a perturbation is compared against a negative control (without any perturbation). Inset shows the profiles of all the strains that were characterized. Individual profiles compared to the negative control are listed in Figure S2 and S3. Genetic context is represented in the axis of abscissas: the yellow shade indicates the promoter region, the purple rectangle indicates the 30-bp perturbation sequence, and the red shade indicates the upstream region to the promoter. NFR stands for the nucleosome-free region and asterisk (\*) shows outstanding changes in occupancy between P5 and NC. Error bars represent error propagation.

**2. Recruiting Transcriptional Regulators.** On the basis of the observation that the conserved guanines influence cellobiose consumption and gene expression by comparison of the strains High 8 and Disrupted High 8, we were intrigued by the possibility that a more active transcription was caused by recruiting specific transcriptional regulators (TFs). We therefore scanned the 30-bp perturbation sequences for motifs where potential TFs could bind using the online server

YeTFaSCo, a database collecting all the available TF specificities for *S. cerevisiae*.<sup>24</sup> The strains High 3, High 5, and High 8 were found to carry the binding sites for five transcription factors, including RAP1, SUT1, NSI1, MAT-ALPHA2, and UPC2 (Table 1). The potential binding sites for RAP1 and SUT1 carry the conserved guanine at the 11th position; its disruption severely impaired cellobiose consumption (Figure 4c, d) and gene expression (Figure 5b),



**Figure 7.** Assessing the performance of the perturbation strategy against promoter engineering for modulating gene expression in *S. cerevisiae*. A set of library members from both methods were compared to investigate their diversification power. GFP fluorescence was measured for both groups by flow cytometry after 44-h fermentation. There was a significant difference in fluorescence between the variants generated by promoter engineering compared to the variants generated by the perturbation strategy;  $p < 0.05$ . NC indicates negative control. Error bars represent the standard deviation of three biological replicates.

which signaled an essential role borne by these TFs. Interestingly, RAP1, SUT1, and NSI1 were previously reported to be related to chromatin remodeling, suggesting a more complex mechanism involving multiavenues. RAP1 has been validated as an effective chromatin opener in the *HIS* promoter,<sup>25</sup> whereas SUT1 and NSI1 were reported to elicit nucleosome displacement.<sup>26</sup> A mechanism was proposed in a recent study for the capability of RAP1 for invading a compact chromatin, locally inducing an open state, and facilitating the establishment of a stable conformation by coordination with the RSC remodeling complex.<sup>27</sup>

**3. Inducing Changes in Local Nucleosome Occupancy.** It has been long known of the antagonistic relationship between TFs and nucleosomes (known as the accessibility model).<sup>28</sup> Studies focused on general TFs imply their roles in the formation of nucleosome-free regions<sup>29</sup> likely by promoting a transient uncoiling of DNA off the nucleosome.<sup>30</sup> To investigate whether the perturbation sequences indeed caused a measurable chromatin remodeling effect for the high-performance strains, we carried out the micrococcal nuclease digestion followed by quantitative PCR (qPCR) with tiling array.<sup>31,32</sup> This approach allows a preferential cleavage of the internucleosomal regions leaving a nucleosome-protected DNA fraction from which the segments with a mononucleosome length are recovered to serve as templates for qPCR. Figure 6a shows the nucleosome occupancy generated for multiple positions upstream to the two promoters (*i.e.*, *PYK1p* and *TEF1p*) used in the cellobiose utilization pathway. The result indicated that incorporation of perturbation sequences generally increased the occupancy of the two promoters at their 5' ends. At a quick glance, this result did not support the transcription-level assay and the finding of the potential TF-binding sites, based on which one would be allured to expect lower nucleosome occupancy. However, it is essential to point out that nucleosome remodeling is a dynamic process, and relying on a few noncontiguous locations at one sample collection time to draw a conclusion might be overstretching at this stage. Indeed, there are reports that an inverse correlation between transcription rate and nucleosome occupancy is not always observed.<sup>33,34</sup> Nevertheless, if comparison were only made between the plasmids before and after the perturbation, the nucleosome occupancy assay consistently suggested that the insertion of the 30-bp perturbation sequences altered the

transcriptional environment for the two genes, which created an opportunity to screen for improved variants with desired expression dynamics. Whether the impact was mediated directly by nucleosome positioning or recruitment of TFs or in combination requires high-resolution characterization of these interplaying factors at both sequence and time scales.

#### 4. Inducing Chromatin Remodeling at a Distance.

Considering that Hidden Markov Model is often used to predict nucleosome occupancy,<sup>32,35</sup> we adopted this computational tool to examine chromatin remodeling for an expanded region. Interesting trends appeared in the promoter regions of both expression cassettes. Both promoters showed clear nucleosome-free regions (NFR). In the *PYK1* promoter, six of the perturbed strains showed decreased levels of occupancy in the ~100 bp upstream to the NFR (marked by \* in Figure 6b and Figure S2). In the *TEF1* promoter, one can find a distinct high occupancy region with a length of approximately 180 bp in the nonperturbed control strain. In comparison, the perturbed strains show a profile where the high occupancy region has expanded to ~300 bp and resolved into two peaks. This finding is supported by all the high-performance strains (Figure S3) although only High 5 is presented in Figure 6b for clarity. More interestingly, the region of ~30 bp at the beginning of the promoter (right after the perturbation sequence) shows a significantly reduced level of occupancy across all the high-performance strains (Figure 6b and Figure S3). Consistent with the mechanism above showing that the local nucleosome occupancy was altered by the perturbation, Hidden Markov Model also qualitatively supported the altered chromatin-packaging patterns after the perturbation sequences were introduced. The full mechanism should be multifaceted considering the diverse characteristics of the perturbation sequences. In our view, meaningful conclusion should be drawn based on the comparison between the perturbed strains and the nonperturbed one within the same type of measure at this early stage.

In summary, it is likely that the beneficial impact was mediated by a combined effect of several mechanisms. These four preliminary mechanistic studies collectively pointed to one scenario where the perturbation sequences bearing potential TF-binding sites recruit nucleosomes in a different manner from the one used by the nonperturbed plasmid. If certain plasmid architecture leads to a more favorable

expression dynamics of the pathway, the corresponding variant will stand out from the subsequent screening.

**Compatibility of the Perturbation Strategy with Other Pathway Engineering Approaches.** We were interested in assessing the practicality of our method. Its implementation is simple as it only requires slight modifications to primers and does not halt the plasmid construction workflow (Figure S4). In order to benchmark this method with other available strategies for combinatorial pathway engineering, we chose to compare the perturbation strategy with the promoter engineering strategy,<sup>2</sup> as they both rely on PCR to quickly generate variations in gene expression. An experiment to measure GFP fluorescence under the control of the TEF2 promoter including either perturbations using a 30-bp oligonucleotide library or a mutagenized promoter library was performed. The results indicated that our approach is particularly useful for creating variants with up-regulated expression, whereas promoter engineering usually generates variants with lower expression than the control (the two groups are significantly different;  $p < 0.05$ ) (Figure 7). Contingent to particular applications, both methods offer room for optimization. For example, it has been reported that changes in parameters such as the PCR cycle number and the concentration of the nucleotide analogues improve the rate of mutagenesis for the promoter engineering strategy.<sup>2,37</sup> In our strategy, the 30-bp flanking sequences could be extended to introduce larger-scale spatial variations. In addition, the use of characterized perturbations in tandem may also aid to expand the dynamic range that could be attained, in an analogous manner to the introduction of multiple upstream activating sequences in the promoter region.<sup>38</sup> Although local genetic context serves a regulator-like function, it does not exclude the use of other existing fine-tuning strategies. Therefore, the combination of these complementary tools is most likely to lead to an enhanced diversification space.

Orthogonality is regarded as a highly desirable feature when designing biological parts in synthetic biology. We also assessed the capacity of perturbations to exert their modulation impact on a different biosynthetic pathway. Namely we transferred the perturbation sequences from P3, P5, and P8 to the upstream regions of the expression cassettes comprising the betaxanthin biosynthetic pathway (Figure S5a). Drawing from the observation that perturbations retrieved from GFP did not share similarities with those identified in the cellobiose case, our expectation was that P3, P5, and P8 would not be able to elicit significant changes in a totally different genetic context. This hypothesis was validated experimentally as shown in Figure S5b. Secondary evidence also arose from the integration of the perturbed cellobiose pathway to the genome, which showcased an undeniable influence of the genetic context on function. Despite gene dosage was comparable between our plasmid system and the integration design, growth on cellobiose was severely impaired in the integration strains and cellobiose was utilized with much lower efficiency by the integration strains (Figure S6). Altogether, these findings imply that perturbation of local genetic context requires a user to go through the screening step for each assessed phenotype, as changes in the neighboring sequences to the expression cassettes will likely abolish the remodeling effect induced by the perturbations.

Microbial cell factories are regarded as promising platforms for producing high-value chemicals from cheap carbon sources. Intensive efforts in the form of up-regulation or down-

regulation of key enzymatic steps and disruption of byproduct-yielding reactions are often required to divert resources from pathways supporting cell growth to metabolic pathways synthesizing the product of interest. In this regard, the use of perturbing local genetic context to balance gene expression works as a straightforward diversification strategy to create libraries containing combinatorial modulations of the pathway expression. In the current study, we demonstrated that the DNA sequences contiguous to the expression cassette greatly impacted the measurable trait (*i.e.*, cellobiose consumption). Our results show that incorporation of 30-bp upstream sequences to an expression cassette is sufficient to induce a favorable change in expression dynamics likely *via* recruiting TFs and/or altering nucleosome occupancy. Accumulating evidence demonstrates that the spatial arrangement of nucleosomes in eukaryotes is far from random. Using yeast as example, we demonstrated that local genetic context plays a major role in balancing gene expression in complex metabolic pathways. By creating a method that couples library construction with high-throughput screening, we were able to rapidly obtain improved phenotypes that exhibited more favorable gene expression and chromatin packaging dynamics. This study not only sheds light on the previously overlooked DNA elements, but also facilitates genetic manipulation *via* chromatin remodeling. We envision this method to be particularly useful in unexploited microorganisms that harbor great metabolic potentials but still lack detailed characterization of biological parts.

## METHODS

**Strains and Culture Media.** *Saccharomyces cerevisiae* YSG50 (MAT $\alpha$ , *ade2*-1, *ade3* $\Delta$ 22, *ura3*-1, *his3*-11, 15, *trp1*-1, *leu2*-3, 112, *can1*-100) was used as the host for plasmid assembly in this study. YPAD consists of 1% (corresponding to 10 g/L) yeast extract, 2% peptone, 0.01% adenine hemisulfate, and 2% dextrose. Selective media were prepared using 0.083% complete supplement mixture lacking the auxotrophic nutrient of interest (uracil, leucine, or tryptophan), 0.16% yeast nitrogen base without amino acids and ammonium sulfate, 0.5% ammonium sulfate, and 2% dextrose or cellobiose. *Escherichia coli* BW25141 (*lacI*<sup>q</sup> *rrnB*<sub>T14</sub>  $\Delta$ *lacZ*<sub>WJ16</sub>  $\Delta$ *phoBR*580 *hsdR*514  $\Delta$ *araBAD*<sub>AH33</sub>  $\Delta$ *rhaBAD*<sub>LD78</sub> *galU*95 *endA*<sub>BT333</sub> *uidA*( $\Delta$ *MluI*):*pir*<sup>+</sup> *recA1*), used for plasmid enrichment, was cultured in Luria–Bertani medium supplemented with 100  $\mu$ g/mL ampicillin. The strains and corresponding genotypes are summarized in Table S4.

**Plasmid Construction and Transformation.** All plasmids described in this work were constructed using DNA assembler<sup>18</sup> and are listed in Table S5. For *S. cerevisiae* transformation, 90-nt single-stranded DNA fragments carrying a 30-nt random sequence flanked by 30-nt regions homologous to neighboring elements were synthesized (IDT, Coralville, IA). A list of primers is provided in Table S2. For each design, the single-stranded oligonucleotide was converted to a double-stranded DNA fragment using the primer annealing to the 3'-end. OE-PCR was performed to link the upstream fragment (~1 kb), the 90-bp fragment carrying the perturbation sequence, and the downstream fragment (~500 bp) to increase the length of the homology regions flanking the perturbation. Promoters (TEF2p, PYK1p, TEF1p, PGK1p, TDH3p, and CCW12p), genes (*gfp*, *ghl-1*, *cdt-1*, *ARO4*, *TYRH*, and *DOD*), and terminators (TEF2t, ADH1t, PGK1t, TDH1t, PGK1t, and ADH1t) were amplified from pRS426-TEF2p-

TEF2t,<sup>20</sup> pRS425-NCU00801-NCU00130<sup>1</sup> (a gift from Dr. Huimin Zhao at the University of Illinois, Urbana–Champaign), pWCD2240 and pWCD2249<sup>39</sup> (a gift from Dr. Dueber at University of California, Berkeley, Berkeley) using Q5 high-fidelity DNA polymerase (New England BioLabs, Ipswich, MA, USA). pRS416 was digested with *Xho*I and *Sac*I to generate a plasmid backbone for constructing plasmid pRS416-PX-TEF2p-GFP-TEF2t; pRS415 was digested with *Xba*I and *Kpn*2I to construct plasmid pRS415-PX-CB or *Bcu*I and *Hind*III to construct plasmid pRS415-PX-BX (Table S5). Restriction enzymes were purchased from New England BioLabs. For assembly, the digested backbone and PCR-amplified fragments were cotransformed *via* electroporation into *S. cerevisiae* following the reported DNA Assembler method.<sup>18</sup> Transformants were plated on selective plates lacking the amino acid corresponding to the auxotrophic selection marker (uracil, leucine, or tryptophan). Colonies appeared after 2–3 days of incubation at 30 °C. Yeast plasmids were isolated using the Zymoprep Yeast Plasmid Miniprep II Kit (Zymo Research, Irvine, CA, USA) and transformed into *E. coli* BW25141 for enrichment. The *E. coli* plasmids were isolated using the QIAprep Spin Plasmid Mini-prep Kit (Qiagen, Valencia, CA, USA). The constructs were verified by restriction digestion and Sanger sequencing (ISU DNA facility, Ames, IA, USA). Additional protocol details regarding pathway assembly can be found in previous studies.<sup>18,19</sup>

**Library Screening for GFP<sup>+</sup> Strains and Fluorescence Analysis.** Colonies transformed with the plasmid pRS416-PX-TEF2p-GFP-TEF2t perturbation library were subjected to fluorescence intensity analysis using a DR46B transilluminator (Clare Chemical Research, Dolores, CO, USA). The collected images were analyzed using FIJI<sup>40</sup> to identify ten colonies with the highest fluorescence and ten with the lowest expression. These colonies were cultured in 2 mL SC-URA medium at 30 °C and 250 rpm for 40–48 h. Cells were diluted in 10 mM phosphate-buffered saline (pH 7.4) to obtain an optical density at 600 nm (OD<sub>600</sub>) of 0.1–0.2. A FACSCanto flow cytometer was used to analyze the fluorescence intensity of the cells at 488 nm. The distribution and percentage of GFP<sup>+</sup> clones were calculated using BD FACSCanto Clinical Software. For rigorous comparison, the verified improved variants were replated and single colonies were inoculated into 2 mL of SC-URA medium and cultured for 48 h to serve as seed culture. The cells were then diluted to an OD<sub>600</sub> of 0.1 and cultured in 96-well plates (Fisher Scientific, Pittsburgh, PA, USA) containing 200  $\mu$ L of medium in each well. Plates were incubated at 30 °C with continuous shaking. GFP intensity measurements were performed every 30 min using a Synergy HTX multimode reader (BioTek, Winooski, VT, USA) for a period of 44 h. The excitation and the emission wavelengths were 485 and 516 nm, respectively.

For promoter engineering, TEF2p was mutagenized by error prone PCR following the manufacturer's recommendations and a previously described protocol.<sup>2</sup> Nucleotide analogues 8-oxo-2'-deoxyguanosine and 6-(2-deoxy- $\beta$ -D-ribofuranosyl)-3,4-dihydro-8H-pyrimido-[4,5-c][1,2]oxazin-7-one were purchased from TriLink Biotechnologies (San Diego, CA, USA). The wild type TEF2p in plasmid pRS416-PX-TEF2p-GFP-TEF2t was replaced with the mutagenized promoter to yield plasmid pRS416-MX-TEF2p-GFP-TEF2t. Following analysis, the yeast plasmids were isolated and enriched in *E. coli* to yield perturbation sequences and promoter sequences carrying mutations.

**Quantification of Betaxanthin Fluorescence.** Colonies transformed with any of the six versions of the plasmid pRS415-PX-BX were incubated at 30 °C for approximately 3–4 days to allow for color development. Because cells were uniformly colored, colonies were randomly selected for growing in 2-ml SC-LEU medium at 30 °C until the cultures reached saturation. Cells were subsequently diluted to an OD<sub>600</sub> of approximately 0.1 and cultured in 96-well plates (Fisher Scientific, Pittsburgh, PA, USA) containing 200  $\mu$ L of medium in each well. Plates were incubated at 30 °C with continuous shaking. Fluorescence measurement was performed every 30 min using a Synergy HTX multimode reader (BioTek, Winooski, VT, USA) for a period of 48 h. The excitation and the emission wavelengths were 485 and 516 nm, respectively.

**Growth Characterization of *S. cerevisiae*.** Sizes of colonies transformed with the plasmid pRS415-PX-CB perturbation library were analyzed using FIJI<sup>40</sup> to identify colonies with the largest surface area. Top eight performers were cultured in SC-LEU medium at 30 °C and 250 rpm overnight to serve as seed cultures. Cells from the seed cultures were diluted to an OD<sub>600</sub> of 0.1 and cultured in 96-well plates containing 200  $\mu$ L of medium in each well at 30 °C with continuous shaking at 205 rpm. OD<sub>600</sub> measurements were performed every 15 min using the Synergy Eon Microplate Spectrophotometer (BioTek) for a period of 48 h. The specific growth rates were calculated as the slopes obtained from semi logarithmic plots (logOD<sub>600</sub> vs time). Following characterization, the yeast plasmids were isolated and enriched in *E. coli* to retrieve the perturbation sequences. To assess the perturbation sequences, an alignment of the retrieved sequences was performed and a sequence logo was prepared using WebLogo.<sup>41</sup> Colonies transformed with plasmids pRS415-dPX-CB were similarly characterized. Growth profiles were verified in two independent transformation events.

**Estimation of Cellobiose Consumption by *S. cerevisiae*.** To seek a direct relationship between growth and cellobiose consumption, larger-volume fermentation was set up. A single colony carrying a specific perturbation sequence was inoculated into 3 mL of SC-LEU to serve as the seed culture and grown overnight. Seed cultures were transferred to 250-ml baffled flasks containing 50 mL of SC-LEU medium and cultivated for 48 h at 250 rpm. Following fermentation, cells were washed twice with water and transferred into 12 mL of SC-LEU medium containing 2 g/L cellobiose to reach an initial OD<sub>600</sub> of 10. Cells were cultivated at 250 rpm and 30 °C for 96 h. Samples were collected at 12, 24, 48, 72, and 96 h. Cellobiose quantification was performed using the DNS reducing sugar assay<sup>42</sup> and a modified version of a previously described method.<sup>43</sup> Briefly, 15  $\mu$ L of the supernatant and 30  $\mu$ L of the DNS reagent were mixed in 0.2 mL PCR tubes. The mixtures were placed in a C1000 touch thermocycler (BioRad, Hercules, CA, USA) and heated to 98 °C for 5 min. The reaction was stopped by cooling down to 4 °C for 1 min. Samples were diluted with water at a ratio of 1:20. Absorbance was measured in 96-well polystyrene plates containing 130  $\mu$ L of the diluted samples at 540 nm using a Synergy Eon Microplate Spectrophotometer. SC-LEU medium supplemented with 0 to 3 g/L cellobiose was used to generate a standard curve (Figure S7). Cellobiose consumption profiles were verified by two independent transformation events.

**qPCR for Assessing the Transcriptional Levels.** To measure the transcriptional levels of *ghl-1* and *cdt-1* after the perturbation sequences were incorporated, qPCR was

performed. Cells were grown in identical conditions as described for the cellobiose consumption study. Cell pellets with an OD<sub>600</sub> of at least 12 were collected at 24 h and stored at −80 °C. RNA was extracted using RNeasy mini kit (Qiagen, Valencia, CA). The genomic DNA was removed using Turbo DNA-free kit (Thermo Fisher Scientific, Waltham, MA) and cDNA libraries were reverse transcribed using RevertAid first strand cDNA synthesis kit (Thermo Fisher Scientific, Waltham, MA) according to manufacturers' instruction. Primers for qPCR were designed using PrimerQuest Tool from IDT and are listed in Table S2. Twenty-five μL reactions were carried out in a StepOnePlus Real-Time PCR system (Thermo Fisher Scientific, Waltham, MA) using Maxima SYBR Green qPCR Master Mix (Thermo Fisher Scientific, Waltham, MA) according to manufacturer's instruction. One μL of cDNA was applied to each sample. Relative quantification of *ghl-1* and *cdt-1* was performed using the corresponding standard curve for each target and then normalized to the expression of *ALG9*, a housekeeping gene that is commonly used in qPCR analysis. Assay was performed in triplicates.

**Plasmid Copy Number Assay.** Copy number was determined for the exemplary plasmids using qPCR following a previously reported protocol.<sup>23</sup> Briefly, cells were grown in selective medium for 24 h and biomass equivalent to approximately 2 OD<sub>600</sub> per ml were collected. Following a series of boiling and freezing steps, total DNA was extracted and stored at −80 °C. Using PrimerQuest Tool from IDT, primer pairs were designed for two genes: *LEU2* was used to quantify pRS415-PX-CB series of plasmid and *ALG9* was used as an internal reference. Primer sequences are listed in Table S2. qPCR reactions were prepared as above, except that 11.5 μL of isolated DNA template was used. A standard curve was constructed for each target using either pure plasmid or genomic DNA sample. Assay was performed in triplicates.

**Nucleosome Maps.** Nucleosome occupancy was mapped for the perturbation vicinity of the *ghl-1* and *cdt-1* expression cassettes. To this end, 200 mL of cultures at approximately OD<sub>600</sub> of 0.8 were prepared from a starting OD<sub>600</sub> of 0.2. Cells were collected by centrifugation and resuspended in 600 μL of Y-Lysis Buffer (Zymo Research, Irvine, CA, USA). Spheroplasts were prepared by an addition of 187.5 units of zymolase and incubation at 30 °C for 1 h with gentle shaking. Cells were washed with 1.2 M sorbitol and resuspended in 500 μL of Nuclei Prep Buffer (Zymo Research, Irvine, CA, USA). An aliquot of 100 μL was used to perform micrococcal nuclease digestion with 0.5 units of the enzyme. The reaction was incubated for 10 min at 37 °C. Nucleosomal DNA was purified using EZ Nucleosomal DNA Prep Kit (Zymo Research, Irvine, CA, USA). Mononucleosomes were isolated by loading the purified DNA on a 2% agarose gel. The bands located between 100 and 200 bp were excised and gel-purified using Zymoclean Gel DNA Recovery kit (Zymo Research, Irvine, CA, USA) to serve as the templates for qPCR.

Nucleosome maps were constructed according to the protocol for qPCR with tiling arrays<sup>31,32</sup> with either three primer pairs for *ghl-1* or four primer pairs for *cdt-1*, which were designed by PrimerQuest Tool from IDT. Amplicons of ~100 bp were placed at an average distance of 83 bp. qPCR reactions were prepared as above, except that 5 ng of nucleosomal DNA was used per 20 μL of reaction. Relative nucleosome occupancy for *ghl-1* or *cdt-1* was calculated using the corresponding standard curve for each amplicon designed for

the target gene. The result was calculated as the ratio with respect to the locus REC102, which was regarded as a common heterochromatic control<sup>44</sup> to facilitate the comparison among plates. Maps shown in Figure 6 were generated by a second normalization to the tallest PCR signal for each target. Assays were performed in duplicates.

*In silico* prediction of nucleosome occupancy was performed using the software tool NuPop R package based on Hidden Markov Model.<sup>35,36</sup> Both the first and fourth order Markov chain was tested using *S. cerevisiae* as the species.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acssynbio.9b00445>.

Figure S1: Fluorescence profile for the perturbed variants after plasmid retransformation; Figure S2: Predicted nucleosome occupancy for the upstream region to *ghl-1*; Figure S3: Predicted nucleosome occupancy for the upstream region to *cdt-1*; Figure S4: Maps for the plasmids constructed in this study; Figure S5: Assessment of orthogonality of the introduced perturbations by a secondary exemplary pathway; Figure S6: Assessment of the cellobiose utilization pathway with the perturbation sequences in the genomic context; Figure S7: Standard curve for cellobiose estimation; Table S1: Perturbation sequences retrieved in this work; Table S2: Primers used in this study; Table S3: Specific growth rates for the strains bearing the cellobiose pathway with the original perturbations and the disrupted perturbations; Table S4: Strains used in this study; Table S5: Plasmids constructed in this work (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

Zengyi Shao – Interdepartmental Microbiology Program, Department of Chemical and Biological Engineering, and NSF Engineering Research Center for Biorenewable Chemicals, Iowa State University, Ames, Iowa 50011, United States; Ames Laboratory, Ames, Iowa 50011, United States; [orcid.org/0000-0001-6817-8006](https://orcid.org/0000-0001-6817-8006); Phone: 515-294-1132; Email: [zyshao@iastate.edu](mailto:zyshao@iastate.edu)

### Authors

Carmen Lopez – Interdepartmental Microbiology Program, Iowa State University, Ames, Iowa 50011, United States  
Yuxin Zhao – The Key Laboratory of Molecular Microbiology and Technology, Ministry of Education, College of Life Sciences, Nankai University, Tianjin 300071, China; Department of Chemical and Biological Engineering, Iowa State University, Ames, Iowa 50011, United States  
Rick Masonbrink – Office of Biotechnology, Iowa State University, Ames, Iowa 50011, United States

Complete contact information is available at:  
<https://pubs.acs.org/doi/10.1021/acssynbio.9b00445>

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We would like to thank Dr. Shawn M. Rigby for performing flow cytometry analysis. This work was supported by the National Science Foundation Grants (1716837 and 1749782). We would also like to acknowledge the scholarship provided by CONACyT to CL.

## ■ REFERENCES

- (1) Du, J., Yuan, Y., Si, T., Lian, J., and Zhao, H. (2012) Customized optimization of metabolic pathways by combinatorial transcriptional engineering. *Nucleic Acids Res.* 40, No. e142.
- (2) Alper, H., Fischer, C., Nevoigt, E., and Stephanopoulos, G. (2005) Tuning genetic control through promoter engineering. *Proc. Natl. Acad. Sci. U. S. A.* 103, 12678–12683.
- (3) Curran, K. A., Morse, N. J., Markham, K. A., Wagman, A. M., Gupta, A., and Alper, H. S. (2015) Short synthetic terminators for improved heterologous gene expression in yeast. *ACS Synth. Biol.* 4, 824–832.
- (4) 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.
- (5) Yuan, Y., and Zhao, H. (2013) Directed evolution of a highly efficient cellobiose utilizing pathway in an industrial *Saccharomyces cerevisiae* strain. *Biotechnol. Bioeng.* 110, 2874–2881.
- (6) Jeschek, M., Gerngross, D., and Panke, S. (2017) Combinatorial pathway optimization for streamlined metabolic engineering. *Curr. Opin. Biotechnol.* 47, 142–151.
- (7) Kornberg, R. D. (1974) Chromatin Structure: a repeating unit of histones and DNA. *Science* 184, 868–871.
- (8) Yuan, G. C. (2005) Genome-Scale Identification of Nucleosome Positions in *S. cerevisiae*. *Science* 309, 626–630.
- (9) Lee, W., Tillo, D., Bray, N., Morse, R. H., Davis, R. W., Hughes, T. R., and Nislow, C. (2007) A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Genet.* 39, 1235–1244.
- (10) Struhl, K., and Segal, E. (2013) Determinants of nucleosome positioning. *Nat. Struct. Mol. Biol.* 20, 267–273.
- (11) Jansen, A., Van Der Zande, E., Meert, W., Fink, G. R., and Verstrepen, K. J. (2012) Distal chromatin structure influences local nucleosome positions and gene expression. *Nucleic Acids Res.* 40, 3870–3885.
- (12) Raveh-Sadka, T., Levo, M., Shabi, U., Shany, B., Keren, L., Lotan-Pompan, M., Zeevi, D., Sharon, E., Weinberger, A., and Segal, E. (2012) Manipulating nucleosome disfavoring sequences allows fine-tune regulation of gene expression in yeast. *Nat. Genet.* 44, 743–750.
- (13) Iyer, V., and Struhl, K. (1995) Poly(dA:dT), a ubiquitous promoter element that stimulates transcription *via* its intrinsic DNA structure. *EMBO J.* 14, 2570–2579.
- (14) Blazeck, J., Garg, R., Reed, B., and Alper, H. S. (2012) Controlling promoter strength and regulation in *Saccharomyces cerevisiae* using synthetic hybrid promoters. *Biotechnol. Bioeng.* 109, 2884–2895.
- (15) Mandel, C. R., Bai, Y., and Tong, L. (2008) Protein factors in pre-mRNA 3'-end processing. *Cell. Mol. Life Sci.* 65, 1099–1122.
- (16) Flagfeldt, D. B., Siewers, V., Huang, L., and Nielsen, J. (2009) Characterization of chromosomal integration sites for heterologous gene expression in *Saccharomyces cerevisiae*. *Yeast* 26, 545–551.
- (17) Wu, X. L., Li, B. Z., Zhang, W. Z., Song, K., Qi, H., Dai, J. B., and Yuan, Y.-J. (2017) Genome-wide landscape of position effects on heterogeneous gene expression in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* 10, 189.
- (18) 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, 1–10.
- (19) Shao, Z., and Zhao, H. (2013) Construction and engineering of large biochemical pathways *via* DNA assembler. *Methods Mol. Biol.* 1073, 85–106.
- (20) Sun, J., Shao, Z., Zhao, H., Nair, N., Wen, F., Xu, J. H., and Zhao, H. (2012) Cloning and characterization of a panel of constitutive promoters for applications in pathway engineering in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 109, 2082–2092.
- (21) Galazka, J. M., Tian, C., Beeson, W. T., Martinez, B., Glass, N. L., and Cate, J. H. D. (2010) Cellodextrin transport in yeast for improved biofuel production. *Science* 330, 84–86.
- (22) Ha, S. J., Galazka, J. M., Kim, S. R., Choi, J. H., Yang, X., 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. U. S. A.* 108, 504–509.
- (23) Moriya, H., Shimizu-Yoshida, Y., and Kitano, H. (2006) *In Vivo* robustness analysis of cell division cycle genes in *Saccharomyces cerevisiae*. *PLoS Genet.* 2, No. e111.
- (24) de Boer, C. G., and Hughes, T. R. (2012) YeTFaSCo: a database of evaluated yeast transcription factor sequence specificities. *Nucleic Acids Res.* 40, D169–D179.
- (25) Yu, L., and Morse, R. H. (1999) Chromatin opening and transactivator potentiation by RAP1 in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 19, 5279–88.
- (26) Yan, C., Chen, H., and Bai, L. (2018) Systematic study of nucleosome-displacing factors in budding yeast. *Mol. Cell* 71, 294–305.
- (27) Mivelaz, M., Cao, A. M., Kubik, S., Zencir, S., Hovius, R., Boichenko, I., Stachowicz, A. M., Kurat, C. F., Shore, D., and Fierz, B. (2020) Chromatin fiber invasion and nucleosome displacement by the Rap1 Transcription Factor. *Mol. Cell* 77, 488.
- (28) Klemm, S. L., Shipony, Z., and Greenleaf, W. J. (2019) Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* 20, 207–220.
- (29) Kaplan, N., Moore, I. K., Fondufe-Mittendorf, Y., Gossett, A. J., Tillo, D., Field, Y., LeProust, E. M., Hughes, T. R., Lieb, J. D., Widom, J., and Segal, E. (2009) The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* 458, 362–366.
- (30) Anderson, J. D., Thäström, A., and Widom, J. (2002) Spontaneous access of proteins to buried nucleosomal DNA target sites occurs *via* a mechanism that is distinct from nucleosome translocation. *Mol. Cell. Biol.* 22, 7147–57.
- (31) Lam, F. H., Steger, D. J., and O'Shea, E. K. (2008) Chromatin decouples promoter threshold from dynamic range. *Nature* 453, 246–250.
- (32) Curran, K. A., Crook, N. C., Karim, A. S., Gupta, A., Wagman, A. M., and Alper, H. S. (2014) Design of synthetic yeast promoters *via* tuning of nucleosome architecture. *Nat. Commun.* 5, 1–8.
- (33) Oberbeckmann, E., Wolff, M., Krietenstein, N., Heron, M., Ellins, J. L., Schmid, A., Krebs, S., Blum, H., Gerland, U., and Korber, P. (2019) Absolute nucleosome occupancy map for the *Saccharomyces cerevisiae* genome. *Genome Res.* 29, 1996–2009.
- (34) Bryant, G. O., Prabhu, V., Floer, M., Wang, X., Spagna, D., Schreiber, D., and Ptashne, M. (2008) Activator control of nucleosome occupancy in activation and repression of transcription. *PLoS Biol.* 6, No. e317.
- (35) Xi, L., Fondufe-Mittendorf, Y., Xia, L., Flatow, J., Widom, J., and Wang, J.-P. (2010) Predicting nucleosome positioning using a duration Hidden Markov Model. *BMC Bioinf.* 11, 346.
- (36) Wang, J.-P., Fondufe-Mittendorf, Y., Xi, L., Tsai, G.-F., Segal, E., and Widom, J. (2008) Preferentially quantized linker DNA lengths in *Saccharomyces cerevisiae*. *PLoS Comput. Biol.* 4, No. e1000175.
- (37) Zaccolo, M., and Gherardi, E. (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.
- (38) Shabbir Hussain, M., Gambill, L., Smith, S., and Blenner, M. A. (2016) Engineering promoter architecture in oleaginous yeast *Yarrowia lipolytica*. *ACS Synth. Biol.* 5, 213–223.
- (39) Deloache, W. C., Russ, Z. N., Narcross, L., Gonzales, A. M., Martin, V. J. J., and Dueber, J. E. (2015) An enzyme-coupled biosensor enables (S)-reticulin production in yeast from glucose. *Nat. Chem. Biol.* 11, 465–471.

(40) Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. (2012) Fiji: An open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682.

(41) Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–90.

(42) Miller, G. L. (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* 31, 426–428.

(43) King, B., Donnelly, M. K., Bergstrom, G. C., Walker, L. P., and Gibson, D. M. (2009) An optimized microplate assay system for quantitative evaluation of plant cell wall-degrading enzyme activity of fungal culture extracts. *Biotechnol. Bioeng.* 102, 1033–1044.

(44) Kee, K., Protacio, R. U., Arora, C., and Keeney, S. (2004) Spatial organization and dynamics of the association of Rec102 and Rec104 with meiotic chromosomes. *EMBO J.* 23, 1815–1824.