Published in final edited form as:

Biotechnol Bioeng. 2013 October ; 110(10): 2677–2686. doi:10.1002/bit.24952.

Using the Cre-lox system to randomize target gene expression states and generate diverse phenotypes

Bradley Niesner¹ and Narendra Maheshri^{1,*}

¹Department of Chemical Engineering, Massachusetts Institute of Technology, 25 Ames Street, Cambridge, MA 02139

Abstract

Modifying the expression of multiple genes enables both deeper understanding of their function and the engineering of complex multigenic cellular phenotypes. However, deletion or overexpression of multiple genes can be laborious and involves multiple sequential genetic modifications. Here we describe a strategy to randomize the expression state of multiple genes in *S. cerevisiae* using Cre-loxP recombination. By inserting promoters flanked by inverted loxP sites in front of a gene of interest we can randomly turn it OFF or ON, or between 4 distinct expression states. We show at least 6 genes can be randomized independently and argue that using orthogonal loxP sites should increase this number to at least 15. Finally, we show how combining this strategy with mating allows easy introduction of native regulation as an additional expression state and use this to probe the role of 4 different enzymes involved in base excision repair in tolerate exposure to methyl methane sulfonate (MMS), a genotoxic DNA alkylating agent. The set of vectors developed here can be used to randomize the expression of both heterologous genes and endogenous genes, and could immediately prove useful for metabolic engineering in yeast. Because Cre-loxP recombination works in many organisms, this strategy should be readily extendable.

Keywords

Cre-loxP recombination; gene expression; metabolic and cellular engineering

INTRODUCTION

Identifying epistatic relationships between genes aids in understanding the molecular basis of multigenic phenotypes. Classical epistasis studies focus on combinations of deletions in a limited set of genes to gain insight into pathways leading to a particular phenotype. Global approaches sample pairwise combinations of deletion mutants on a much larger scale, and have been used to understand many pathways, including those involved in chromosomal structure (Jasnos and Korona 2007 ; Tong, et al. 2004), metabolism (Segre, et al. 2005), and growth fitness (Jasnos and Korona 2007). Combining deletion of genes with constitutive gene overexpression can yield additional insights in dosage compensation (Acar, et al. 2010)

*Correspondence to narendra@mit.edu.

and improve metabolic engineering (Alper, et al. 2005b) but systematic combinations have not yet been applied on a global scale.

A major hurdle to systematically exploring the impact of deletion and/or altered expression of combinations of multiple genes even in genetically tractable organisms is that targeted gene modification is generally performed sequentially. Several methods have made progress to overcome this hurdle, though each has its limitations. Titratable promoters allow rapid sampling of different combinations of gene expression using small molecules, but such promoters are limited in number and inducible signals may have pleiotropic effects (Cox, et al. 2007 ; Guo, et al. 2008 ; Maya, et al. 2008). Library approaches can also be employed to fine-tune expression levels, often in the context of optimizing metabolic flux through pathways (Alper, et al. 2005a ; Jensen and Hammer 1998 ; Pfleger, et al. 2006), but they are limited to a single gene or operon. Two techniques that generate diverse libraries by rapidly modifying multiple genes are multiplex automated genome engineering (MAGE) (Wang, et al. 2009) and random ligation of yeast artificial chromosomes (YACS) (Naesby, et al. 2009). In MAGE, a large pool of small DNA oligomers is transformed in *E.coli* in an automated and continuous fashion; targeting the ribosomal binding sites of 24 genes involved in lycopene biosynthesis rapidly increased production (Wang, et al. 2009). However, MAGE requires a high oligo transformation efficiency currently possible only in *E. coli* and may not be easily deployed. An alternative approach has been to randomly ligate multiple gene cassettes into a yeast artificial chromosome (YAC); random ligation of components of the flavonoid biosynthetic pathway followed by screening resulted in the creation of novel strains that could produce a variety of different flavonoids (Naesby, et al. 2009). However while this approach is useful for constructing heterologous pathways, other techniques are needed to explore the impact of modifying endogenous gene expression.

Clearly, new strategies are needed for the rapid and targeted exploration of multiple gene expression states. Here we describe a strategy to randomize the expression levels of both native and heterologous genes in *S. cerevisiae* with a limited number of required genetic modifications by using the Cre/loxP system (Sternberg and Hamilton 1981). We develop and test vectors that include both a 2-state and 4-state promoter design where the downstream gene can adopt either 2 (OFF and ON) or 4 (OFF, LO, INT, HI) constitutive expression states after ‘randomization’ with Cre expression. By introducing one ‘randomizable’ promoter per gene, a range of discrete expression states is generated in a previously isogenic population upon induction of Cre. Up to 5 copies of a 2-state promoter, and 2 copies of a 4-state promoter are reliably randomized without significant loss of the cassette. Additional genes can be added by using two sets of orthogonal loxP sites, which allowed 6 copies of a 2-state promoter and could in principle be extended to up to 15 copies. Furthermore, we demonstrate how combining a simple mating strategy and the 2-state promoters can randomize the expression state of 4 different genes involved in base excision repair between an OFF, ON, and natively regulated state. By screening the resulting library for growth in the presence of methyl methane sulfonate (MMS) – a genotoxic agent that alkylates DNA – we recover a set of genotypes which confer tolerance against this stress.

MATERIALS AND METHODS

Strain and plasmid construction

Yeast strains were constructed using standard techniques (Guthrie and Fink 2004). All yeast strains used here are in the W303 background (Thomas and Rothstein 1989) and detailed in Supplemental Table I. Plasmids and construction notes are listed in Supplemental Table II. The yeast integration vectors used were linearized by digestion with *HindIII*, *PstI*, and *EcoRV* in the *HIS3*, *URA3*, and *TRP1* marker regions respectively (Sikorski and Hieter 1989). The PCR-based promoter replacement cassettes based on those designed by (Longtine, et al. 1998) were amplified using the sequences 5' – TGACGTGCGCAGCTCAGGGG – 3', and 5' – ATCGATGAATTTCGAGCTC – 3' on the 3' end of sense and antisense primers, respectively. Primers contained approximately 50 bp of homology to the 5' ends to the desired genomic locus.

Flow cytometry

Flow cytometry was performed using a LSRII HTS (Becton-Dickinson). Venus, tdTomato, and Cerulean proteins were detected using a 488-nm laser and 530/30-nm filter, a 561-nm laser and 610/20-nm filter, and a 405-nm laser and 450/50-nm filter, respectively. For each sample, ~100,000 cells were measured and visualized on a contour plot of forward versus side scatter. Cells within a contour containing 10% of the population were used for further analysis.

'Floxed' promoter randomization

Cells containing 'floxed' promoter and galactose-inducible Cre recombinase were grown in 2% raffinose liquid media overnight, up to an optical density at 600 nm (OD₆₀₀) of ~1. They were then inoculated into 2% galactose media at an OD₆₀₀ of 0.1-0.2 and allowed to grow for 6 hours for randomization unless otherwise noted in the text. Finally, cells were inoculated in glucose media overnight (>18 hrs) to inhibit Cre expression and allow fluorophores to reach steady-state expression levels. The galactose-inducible Cre recombinase was expressed from a centromeric *LEU2*-marked plasmid (kind gift from Paul Wiggins). The centromeric plasmid was maintained in the population by selecting for leucine autotrophy at each step.

Sporulation and haploid selection

Diploid cells heterozygous at the *CAN1* locus (*CAN1/can1*) were sporulated by overnight growth in rich liquid yeast peptone (YP) media with 1% acetate followed by washing and suspension in 1% potassium acetate + complete amino acids for two days. Haploids were selected by plating the sporulated population on synthetic defined media with 600 mg/L canavanine, and lacking arginine and leucine.

Growth rate measurements

Growth rates in liquid media were measured using 96 deep well plates agitated at 30°C. Three biological replicates per clone were inoculated at 0.003 OD₆₀₀ in 800 µL rich media (yeast peptone with 2% glucose or YPD) or YPD + 0.0035% MMS and grown to early

exponential phase (~0.1 OD₆₀₀), at which point optical density was measured approximately once per hour. OD₆₀₀ data points between 0.1 and 1 were fit well to simple exponential growth, which was used to determine the growth rate. Optical density measurements were taken with a Varioskan Flash (ThermoScientific).

Scoring colony size

We measured colony size in the library as a measure of diversity in growth rate using a method similar to (Klein-Marcuschamer and Stephanopoulos 2008). Briefly, approximately 400 cells were plated on 15 cm plates, allowed to grow 24 hours post-plating, and then imaged using a digital camera. Metamorph (Molecular Devices) software was used to segment and record the size of colonies on the plate.

RESULTS AND DISCUSSION

The orientation of a “floxed” promoter flanked by loxP sites can be randomized in a repeatable and controllable manner

A genetic sequence flanked by loxP sites (“floxed”) can be inverted or excised in the presence of the Cre recombinase depending on the orientation of the loxP sequences (Nagy 2000). We reasoned that the orientation of a constitutive promoter flanked by inverted loxP sites would effectively be randomized after a period of continuous Cre expression. We constructed a series of yeast integration plasmids based on pRS3 series vectors (Sikorski and Hieter 1989) in which the promoter region of the *TEF1*, *ACT1*, *ADH1*, or *MYO2* genes are flanked with loxP sites (“floxed”) and drive expression of a fluorescent reporter (Figure 1A). To test this randomization strategy, we integrated a floxed *TEF1* promoter driving YFP (*P_{TEF1}-YFP*) at the *HIS3* locus. We randomized the promoter by employing a centromeric plasmid containing the galactose-inducible *GAL1* promoter driving Cre (Sauer 1987). Cells pre-grown in 2% raffinose were transferred to 2% galactose to turn on Cre expression, and grown for 6 hours. Then, cells were transferred to 2% glucose and grown for >18 hrs to allow dilution or expression of the fluorescent marker, depending on the promoter state. In Figure 1B, flow cytometric analysis of YFP expression in this randomized population indicates that large fractions of the population are found in both the ON (35%) or OFF (65%) state for the promoter. Expression in the OFF state is indistinguishable from a control strain lacking the fluorescence reporter. We used a PCR strategy to verify promoter inversion had occurred (Figure 1C). Importantly, randomization was dependent on growth in galactose, alleviating any concerns about leaky Cre expression when grown in 2% raffinose, when the galactose pathway is de-repressed.

While Cre is highly efficient, multiple inversion events may result in irreversible loss of the floxed promoter or recombination with other genomic regions leading to a large scale rearrangement. In fact, individual clones chosen after randomization whose promoters were in the ON and OFF states could be randomized again; the resulting phenotypic distribution was identical regardless of the initial state (Supplemental Figure 1). Still, we sought to minimize the time of Cre expression so as to reduce the chance of irreversible rearrangements. Therefore, we repeated the randomization but varied the amount of time the cells spent growing in galactose. Increasing the period of Cre expression to 6 hours

increased the eventual fraction of cells in the OFF state, beyond which there was no change (Supplemental Figure 2).

Interestingly, we never saw a 50%/50% split in ON and OFF cells after randomization; instead, we observed a bias towards OFF cells. This bias was not due to growth differences in cells with the different expression states (data not shown) and depended on the genomic location where the floxed *TEF1* promoter was integrated: the OFF state fraction after 6 hours of galactose induction was 70($\pm$ 3)%, 75($\pm$ 2)%, and 88($\pm$ 2)% at the *URA3*, *HIS3*, and *TRP1* loci respectively (n=3 replicates). While we do not know the origin of this bias, perhaps the forward and backward reaction rates of the inversion depend on either adjacent sequence or the regional genomic environment. Nevertheless, >10% of the population was always found to be ON post-randomization.

Randomizing expression in multiple genes and with multiple expression levels

We next sought to test the feasibility of randomizing multiple floxed promoters present in the same cell, as each additional promoter results in a combinatorial increase in the number of achievable genotypes. To do so, we constructed haploid strains with 2 or 3 copies of the floxed *TEF1* promoter integrated at the *URA3*, *HIS3*, and *TRP1* loci, driving distinct fluorescent proteins. In both cases, we found 4 or 8 possible expression phenotypes within a randomized population present in a manner consistent with independent flipping of each cassette (Supplemental Figure 3). We then mated haploid strains to create diploids with 4, 5, or 6 copies of the promoter. Either 4 or 5 copies of the 2-state promoter strains still randomized independently (Supplemental Figure 4), but curiously repeated attempts of randomizing 6 copies failed, even with varying the time of Cre expression. Different attempts resulted in correlated expression states, with most of the genes either all ON or all OFF (data not shown). One possibility for the failure of the 6 copy 2-state promoter may be due to the greater chance of recombination between loxP sites on different promoters, leading to large scale rearrangements. Therefore, with little or no recombination occurring, all the genes would be in the initial ON state, and when recombination occurs, rearrangements could lead to all genes in the OFF state.

We also extended the recombinase strategy to create a 4-state cassette that leads to four expression states (OFF, LO, MED, or HI) that span a > 100-fold range. The 4-state cassette consists of two pairs of floxed promoters in a divergent orientation (Figure 2A). In the presence of Cre, one pair of promoters may be irreversibly excised, but the remaining promoter pair is still be capable of inversion events (Figure 2A). We first tested the functionality of one copy of the 4-state cassette integrated at the *URA3* locus. Upon randomization, we observed the anticipated range of reporter expression levels in the population. Using primers specific to each promoter, we confirmed that recombination resulted in any of the four promoters driving reporter expression and that one of the promoter pairs is excised during randomization. Individual clones chosen from the population exhibited distinct expression levels indicative of one of the four promoters (data not shown). We then mated two haploid strains with the integrated 4-state cassette driving YFP and RFP, respectively, to generate a two copy 4-state diploid. Upon randomization, we observed a similar range of expression levels with little correlation in the expression state of

each reporter (Figure 2B). We chose 60 individual clones and analyzed their expression profiles by flow cytometry. A subset is shown in Figure 2C. We confirmed the existence of all possible promoter rearrangements in the randomized population with PCR (Figure 2D), and used the same scheme to identify specific genotypes among the chosen clones (Figure 2E). Figure 2B includes expression ranges of each of the 4 promoters (top and right of the plot) taken from the analyzed clones. While loss of one promoter pair is expected for the 4-state cassette in the presence of Cre, we expected to find one intact promoter pair upstream of the fluorescent reporters in each clone. However, in some cells with no reporter expression we were unable to detect the presence of any part of the cassette by PCR (Figure 2E, far right). We suspected in these cases an irreversible recombination event may have deleted the cassette. To test this possibility, we subjected all of the clones to another round of galactose induction for six hours. Approximately 25% of the loci in the clones failed to re-randomize (15/60 for YFP, 17/60 for RFP). For any application that does not require repeated randomization, this irreversibility should still be acceptable because all possible genotypes are realized within the population.

While small numbers of 2- and 4-state promoters recombine and randomize successfully, increasing the number of floxed promoters leads to irreversible recombination events suggestive of deletion. Further addition of floxed promoters would exacerbate this problem. To address this issue, we utilized orthogonal loxP sites (commonly referred to as loxN and lox2772) that are capable of recombining with themselves, but do not interact with loxP sequences or each other (Lee and Saito 1998). We verified the invertability of alternative loxP-flanked promoters (data not shown), and then added 2 copies of a loxN-flanked *TEF1* promoter to make a new 6 copy strain.

To randomize this strain we found it best to express Cre for 12 hours. We then chose and analyzed the expression of 187 clones using flow cytometry. However, solely assaying expression was inadequate in completely characterizing the genotype because we were unable to distinguish clones that have 1 versus 2 promoters in the ON state driving expression of any particular reporter. We complemented the flow cytometry assay with PCR, assaying the promoter orientations in 38 selected clones using a strategy similar to that in Figure 1C. Because CFP and YFP share nearly identical sequence, the PCR strategy was also limited in uniquely determining the genotype. Still, we compared the combined results of the fluorescence and PCR assays in Table I, in which each entry represents either a unique genotype or a small combination of genotypes. Entries corresponding to genotypes that should never be realized are shaded gray. We found the number of clones observed with a particular fluorescence phenotype and PCR-based genotype was largely consistent with what would be expected assuming each allele behaved independently. The one clone found in the gray region lacked only YFP expression which was confirmed multiple times; However, we could not verify the floxed promoter was in the OFF position by PCR; most likely there was an irreversible loss of the floxed promoter upstream of YFP, similar to what was observed for the 2 color, 4-state clones in Figure 2E.

Randomizing expression of DNA repair genes to explore their protective role against DNA damaging stress

To demonstrate the application and utility of this recombinase strategy, we sought to randomize the expression state of four yeast genes involved in base excision repair (BER). BER pathways typically originate when small, non-helix distorting DNA lesions such as alkylation damage are recognized by DNA glycosylases, which remove the damaged nucleotide and leave an abasic (AP) site. In yeast the genes *APN1* and *APN2* encode AP endonucleases that cleave the phosphodiester linkage at the AP site, initiating further events that lead to error-free repair. However, if unrepaired AP sites remain when DNA synthesis occurs, the replication fork can stall. At this point, repair can proceed through an error-free or a mutagenic lesion bypass pathway. The mutagenic pathway requires a “insertase” polymerase to insert a base opposite the AP site, and then the extender polymerase, Pol ζ , a heterodimer of the gene products of *REV3* and *REV7* (Boiteux and Guillet 2004). Thus, stalled replication forks can be restarted, sometimes at the cost of mutations (Figure 3A). Although *APN1* constitutes the major source of AP endonuclease activity, deleting either *APN1* or *APN2* has no phenotype (Johnson, et al. 1998 ; Ramotar, et al. 1991). The *apn1* background does have a mild sensitivity to methyl methane sulfonate (MMS), a DNA alkylating agent which indirectly increases the AP site load, resulting in increased mutation rates. The *apn1 apn2* background exhibits a much more pronounced MMS sensitivity, with a much larger increase in mutation rate. Deletion of either *REV3* or *REV7* in the *apn1 apn2* background reduces the elevated mutation rate seen both in the presence of absence of MMS, confirming the importance of Pol ζ in mutagenic lesion bypass (Johnson, et al. 1998).

We reasoned that modulating the expressions states of *APN1*, *APN2*, *REV3*, and *REV7* would change the repair capacity of AP sites through various pathways (Figure 3A) and ultimately affect the growth rate, especially in the presence of MMS. Previous studies indicate the *apn1 apn2* and *apn1 apn2 rev3* and *apn1 apn2 rev7* backgrounds exhibit reduced growth rates in the presence of MMS (Johnson, et al. 1998). However, there are limited studies looking at the effects of overexpressing some of these genes in the presence of a high AP site load, particularly when other genes are compromised. Both *REV3* and *REV7* have been moderately overexpressed (1.8 and 3-fold) to examine the effect of increased Pol ζ activity in the presence of UV. This lead to a mild (3-fold) increase in the background mutation rate but increased UV sensitivity and a reduction in growth rate (Rajpal, et al. 2000). Therefore overexpression of all four genes may not be obviously beneficial because it may shift repair away from error-free post-replicative repair pathways and/or severely increase the mutational load. We randomized expression these 4 genes between an OFF state similar to a deletion, an ON state where the gene is constitutively expressed, and a NATIVE state representing native regulation. This library could then be subjected to a high AP site load by addition of MMS and genotypes that maintain high growth could be isolated.

To create a library with $(3 \text{ expression states})^4 \text{ genes} = 81 \text{ phenotypes}$, we began by constructing two haploid cells of opposite mating types as follows. In each haploid, promoters of 2 of the 4 genes were replaced with a floxed promoter. We did this using a second set of plasmids we developed to target native genes based on PCR-based disruption

vectors (Longtine, et al. 1998). Primers can be designed to amplify a cassette containing a selectable marker followed by one of two different floxed promoters (*PTEF1* and *P_{ACT1}*) for integration and replacement of the native promoter upstream of any yeast gene (Figure 1A). The initial promoter orientation is ON. The haploids were then mated and sporulated. Because only one haploid has canavanine sensitivity, we could specifically select for haploids after sporulation (Figure 4A). These haploids could then be randomized by growth in galactose. All 4 floxed promoters are present on different chromosomes, minimizing the possibility of interpromoter recombination. We verified the presence of all 3 expressions states for each gene in the bulk mixed population after galactose induction of Cre using PCR (Figure 4B). As an initial step to ascertain any diversity in growth rate, we plated the library on rich media (YPD) and compared colony sizes after 24 hours of growth (Supplemental Figure 5). The two haploids used for mating differed at the *ADE2* locus (*ADE2+* and *ade2-1*), so as controls we plated haploids isogenic to both but without the floxed promoter. While the loss of the *ADE2* auxotrophic marker had a significant effect on decreasing colony size, the mixed populations both before and after randomization exhibited a range of colony sizes that spanned both controls. The marked density toward lower growth rates suggested at least some of the difference was due to differences in DNA repair gene expression and not due to differences in the *ADE2* locus. We then selected for clones that could grow in the presence of high levels of MMS by plating 400 cells (~5x coverage of 81 different genotypes in the library) on rich media (YPD) with 0.035% MMS. Sixteen colonies were observed after growth for 5 days compared to two colonies for a control with a wild type haploid plated. The growth rates of 11 clones in liquid YPD media + 0.0035% MMS as well as the wild type strain was measured (Figure 4C) and each clone was genotyped by PCR (Table II).

Nine of the clones exhibited high growth rates in MMS that were indistinguishable from each other and significantly higher than wild type ($p < 0.025$). All these clones had *APN2* under constitutive expression and *APN1* either constitutively expressed or under native regulation. Yeast express *APN2* at low levels during log phase growth, (410 versus 7300 molecules per cell for *APN1* (Ghaemmaghami, et al. 2003)), and *Apn1p* is responsible for 97% of AP endonuclease activity in lysate (Boiteux and Guillet 2004). Therefore, these results indicate that AP endonuclease activity is probably rate limiting at the higher AP site loads conferred by MMS treatment and overexpression of *APN2* likely increases the cell's capacity for error-free BER. This is further supported by the lower growth rates of clones 10 & 11 which overexpressed *APN1* but not *APN2*.

We found no clear role for Pol ζ because 4 of the fast growing 9 clones had *REV7* in the OFF state. We never found both *REV3* and *REV7* in the ON state, consistent with Pol ζ overexpression reducing growth as seen with UV (Rajpal, et al. 2000). Surprisingly, however, none of the clones were found to be OFF for *REV3* and in every clone which was OFF for *REV7*, *REV3* was ON. However, it remains unclear whether *REV3* expression is required for fast growth in MMS. Because there is no function of Rev3p ascribed outside its role in forming Pol ζ by dimerizing with Rev7p, it is difficult to imagine a requirement for *REV3* expression. There may be a strong bias in the randomization of *REV3* status such that the majority of the library has *REV3* either ON or natively regulated. However, when we

determined the status of *REV3* in 20 unselected clones post-randomization we found a 12/5/3 split for NATIVE/ON/OFF. In the supporting text we estimate that even if the 2-state promoter driving *REV3* expression had a 98%/2% bias for the ON state, there would still be a 50% chance of finding a clone that grows in high MMS with *REV3* OFF. This assumes that native or constitutive *APN1* expression and constitutive *APN2* expression is sufficient for the fast MMS growth phenotype. We cannot definitively rule out whether the *REV3* OFF state correlates with the *APN1* and/or *APN2* OFF state, but with fluorescent reporters we always found 4 floxed promoters switched independently (Figure 2B). Therefore, the library selection suggests a *REV3*-specific role for fast growth in MMS could be tested further.

Summary

By flanking promoters with inverted loxP repeats, we have shown how to rapidly and repeatedly randomize the expression state of multiple promoters in budding yeast, leading to genetically stable phenotypes which can be interrogated for desirable properties. This recombinase-based strategy should extend to the large number of organisms in which the Cre-loxP system has been shown to function (Nagy 2000). It also confers an additional advantage over traditional genetic methods that delete or overexpress genes. Because all target genes are initially expressed at native levels or overexpressed, certain genotypes may be present after randomization that would be difficult to construct with traditional methods. This would be true of genetic suppressors that consist of multiple genes -- for example, a combination of two sets of synthetic lethal genes that are not lethal when both sets are deleted in combination.

As the number of genomically integrated promoters and loxP sites increases, we find an increased probability of irreversibly ending in one expression state. Because we are unable to detect the state by PCR we surmise that recombination, perhaps between different floxed promoters, leads to some large scale genomic rearrangement. We found a critical parameter is the amount of time Cre is expressed. Nevertheless, up to 5 2-state promoters could be randomized within one cell, with the amount of time Cre is expressed as a critical parameter for success. Utilizing the other two loxP variants could allow 15 2-state promoters to be randomized. An alternative strategy to further prevent 'cross-talk' between multiple promoters is to express multiple versions of Cre with altered loxP site preferences (Santoro and Schultz 2002) or even other recombinases (Blakely, et al. 1993 ; Groth and Calos 2004). However, as more genes are chosen to be simultaneously randomized, an important design constraint is that floxed promoters with identical loxP sites should be present on different chromosomes to prevent irreversible inter-cassette recombination events. Placing the floxed promoters on opposite arms of one chromosome may be sufficient although we have not tested this.

Our results suggest no more than 2 copies of a 4-state promoter with each type of loxP sequence can be used without large numbers of irreversible recombination events (Fig. 2). However, use of orthogonal Cre-loxP pairs should extend the numbers of genes capable of being randomized. Even with just the loxP variants tested here, potentially ~6 4-state cassettes could be integrated and randomized successfully. This allows the sampling $4^6 = 4096$ expression states with 2 sets of 3 integration steps followed by a single mating. These

could be useful in cellular engineering contexts where fine-tuning of gene expression is the ultimate goal.

Perhaps the more pressing limitation to constructing a strain with large numbers of floxed promoters is not ‘cross-talk’ between promoters but the technical difficulties in repeated integration of constructs that share homology and limited number of markers (the constructs provided have 5 independent markers, Figure 1). Still, using the current strategy we employed for randomizing BER proteins with 2-state promoters, constructing a strain capable of realizing between $2^{10} = 1024$ and $3^{10} = 59049$ expression states would involve 5 transformation steps in two haploid yeast that could be done in parallel. In general, 2^N (or 3^N) genotypes can be realized with $N/2$ parallel modifications and one mating step, (followed by sporulation). Moreover, genome engineering techniques continue to improve, especially in model organisms like *S. cerevisiae*. These techniques enabled the development of synthetic chromosome arms in which loxPsym variants, which lack orientation, are interspersed to allow random deletion of portions of genetic material (Dymond, et al. 2011). This work supports the feasibility of using the same methods to generate large numbers of floxed promoters allowing not just the deletion, but modulation of many genes’ expression levels and regulation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank B. Paul and S. Finney-Manchester for discussions. This work was funded by an award from the Du Pont / MIT Alliance (to N.M.), GM095733 (to N.M.) and MIT startup funds (to N.M.).

REFERENCES

- Acar M, Pando BF, Arnold FH, Elowitz MB, van Oudenaarden A. A general mechanism for network-dosage compensation in gene circuits. *Science*. 2010; 329:1656–1660. [PubMed: 20929850]
- Alper H, Fischer C, Nevoigt E, Stephanopoulos G. Tuning genetic control through promoter engineering. *Proc Natl Acad Sci U S A*. 2005a; 102:12678–12683. [PubMed: 16123130]
- Alper H, Miyaoku K, Stephanopoulos G. Construction of lycopene-overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets. *Nat Biotechnol*. 2005b; 23:612–616. [PubMed: 15821729]
- Blakely G, May G, McCulloch R, Arciszewska LK, Burke M, Lovett ST, Sherratt DJ. Two related recombinases are required for site-specific recombination at dif and cer in *E. coli* K12. *Cell*. 1993; 75:351–361. [PubMed: 8402918]
- Boiteux S, Guillet M. Abasic sites in DNA: repair and biological consequences in *Saccharomyces cerevisiae*. *DNA Repair (Amst)*. 2004; 3:1–12. [PubMed: 14697754]
- Cox RS 3rd, Surette MG, Elowitz MB. Programming gene expression with combinatorial promoters. *Mol Syst Biol*. 2007; 3:145. [PubMed: 18004278]
- Dymond JS, Richardson SM, Coombes CE, Babatz T, Muller H, Annaluru N, Blake WJ, Schwerzmann JW, Dai J, Lindstrom DL, Boeke AC, Gottschling DE, Chandrasegaran S, Bader JS, Boeke JD. Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature*. 2011; 477:471–476. [PubMed: 21918511]

- Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS. Global analysis of protein expression in yeast. *Nature*. 2003; 425:737–741. [PubMed: 14562106]
- Groth AC, Calos MP. Phage Integrases: Biology and Applications. *Journal of Molecular Biology*. 2004; 335:667–678. [PubMed: 14687564]
- Guo ZS, Li Q, Bartlett DL, Yang JY, Fang B. Gene transfer: the challenge of regulated gene expression. *Trends Mol Med*. 2008; 14:410–418. [PubMed: 18692441]
- Guthrie, C.; Fink, GR. *Guide to yeast genetics and molecular and cell biology*. Part A. Elsevier Academic Press; Amsterdam: 2004.
- Jasnos L, Korona R. Epistatic buffering of fitness loss in yeast double deletion strains. *Nat Genet*. 2007; 39:550–554. [PubMed: 17322879]
- Jensen PR, Hammer K. The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Appl Environ Microbiol*. 1998; 64:82–87. [PubMed: 9435063]
- Johnson RE, Torres-Ramos CA, Izumi T, Mitra S, Prakash S, Prakash L. Identification of APN2, the *Saccharomyces cerevisiae* homolog of the major human AP endonuclease HAP1, and its role in the repair of abasic sites. *Genes Dev*. 1998; 12:3137–3143. [PubMed: 9765213]
- Klein-Marcuschamer D, Stephanopoulos G. Assessing the potential of mutational strategies to elicit new phenotypes in industrial strains. *Proc Natl Acad Sci U S A*. 2008; 105:2319–2324. [PubMed: 18252819]
- Lee G, Saito I. Role of nucleotide sequences of loxP spacer region in Cre-mediated recombination. *Gene*. 1998; 216:55–65. [PubMed: 9714735]
- Longtine MS, McKenzie A 3rd, Demarini DJ, Shah NG, Wach A, Brachet A, Philippsen P, Pringle JR. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast*. 1998; 14:953–961. [PubMed: 9717241]
- Maya D, Quintero MJ, de la Cruz Munoz-Centeno M, Chavez S. Systems for applied gene control in *Saccharomyces cerevisiae*. *Biotechnol Lett*. 2008; 30:979–987. [PubMed: 18239858]
- Naesby M, Nielsen SV, Nielsen CA, Green T, Tange TO, Simon E, Knechtel P, Hansson A, Schwab MS, Titiz O, Folly C, Archila RE, Maver M, van Sint Fiet S, Boussemghoune T, Janes M, Kumar AS, Sonkar SP, Mitra PP, Benjamin VA, Korrapati N, Suman I, Hansen EH, Thybo T, Goldsmith N, Sorensen AS. Yeast artificial chromosomes employed for random assembly of biosynthetic pathways and production of diverse compounds in *Saccharomyces cerevisiae*. *Microb Cell Fact*. 2009; 8:45. [PubMed: 19678954]
- Nagy A. Cre recombinase: the universal reagent for genome tailoring. *Genesis*. 2000; 26:99–109. [PubMed: 10686599]
- Pflegler BF, Pitera DJ, Smolke CD, Keasling JD. Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nat Biotechnol*. 2006; 24:1027–1032. [PubMed: 16845378]
- Rajpal DK, Wu X, Wang Z. Alteration of ultraviolet-induced mutagenesis in yeast through molecular modulation of the REV3 and REV7 gene expression. *Mutat Res*. 2000; 461:133–143. [PubMed: 11018586]
- Ramotar D, Popoff SC, Gralla EB, Demple B. Cellular role of yeast Apn1 apurinic endonuclease/3'-diesterase: repair of oxidative and alkylation DNA damage and control of spontaneous mutation. *Mol Cell Biol*. 1991; 11:4537–4544. [PubMed: 1715020]
- Santoro SW, Schultz PG. Directed evolution of the site specificity of Cre recombinase. *Proc Natl Acad Sci U S A*. 2002; 99:4185–4190. [PubMed: 11904359]
- Sauer B. Functional expression of the cre-lox site-specific recombination system in the yeast *Saccharomyces cerevisiae*. *Mol Cell Biol*. 1987; 7:2087–2096. [PubMed: 3037344]
- Segre D, Deluna A, Church GM, Kishony R. Modular epistasis in yeast metabolism. *Nat Genet*. 2005; 37:77–83. [PubMed: 15592468]
- Sikorski RS, Hieter P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics*. 1989; 122:19–27. [PubMed: 2659436]
- Sternberg N, Hamilton D. Bacteriophage P1 site-specific recombination: I. Recombination between loxP sites. *J Mol Biol*. 1981; 150:467–486. [PubMed: 6276557]

- Thomas BJ, Rothstein R. Elevated recombination rates in transcriptionally active DNA. *Cell*. 1989; 56:619–630. [PubMed: 2645056]
- Tong AH, Lesage G, Bader GD, Ding H, Xu H, Xin X, Young J, Berriz GF, Brost RL, Chang M, Chen Y, Cheng X, Chua G, Friesen H, Goldberg DS, Haynes J, Humphries C, He G, Hussein S, Ke L, Krogan N, Li Z, Levinson JN, Lu H, Menard P, Munyana C, Parsons AB, Ryan O, Tonikian R, Roberts T, Sdicu AM, Shapiro J, Sheikh B, Suter B, Wong SL, Zhang LV, Zhu H, Burd CG, Munro S, Sander C, Rine J, Greenblatt J, Peter M, Bretscher A, Bell G, Roth FP, Brown GW, Andrews B, Bussey H, Boone C. Global mapping of the yeast genetic interaction network. *Science*. 2004; 303:808–813. [PubMed: 14764870]
- Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM. Programming cells by multiplex genome engineering and accelerated evolution. *Nature*. 2009; 460:894–898. [PubMed: 19633652]

Figure 1. Randomizing gene expression using the Cre-lox system

A) A set of (*left*) yeast integration vectors and (*right*) PCR amplifiable disruption vectors with loxP-flanked ('floxed') promoters. Promoters, downstream genes, markers, and loxP sites were varied to generate a set of vectors adaptable to randomize any gene's expression. The floxed promoters and reporter can be excised and swapped using *EcoRI/XhoI* and *EcoRI/NotI* sites, respectively. B) Fluorescence histograms of a floxed *TEF1* promoter driving YFP, *HIS3*:: P_{TEF1} -YFP. All cells (*left*) begin in the ON state before Cre expression (*right*) but randomly settle in an ON or OFF state after Cre induction. Expression is normalized to the mean expression of the *HIS3*:: P_{TEF1} -YFP strain. N = 10,000 cells per plot. C) Genomic PCR of multiple strains using the given primer configuration (bottom) shows the promoters randomize only after Cre induction.

Figure 2. Randomizing gene expression level using four-state promoters

A) Placing two floxed pairs of native promoters in tandem creates a “four-state” promoter capable of randomizing the identity of the promoter proximal to the gene. Labeled arrows show a primer scheme to assay the identity of the promoter driving the reporter after randomization in the presence of Cre. Possible irreversible recombination events and products are shown between loxP sites of similar color, each resulting in one remaining promoter pair which can still reversibly recombine. B) 2D histogram of RFP and YFP expression measured from two different four-state promoters with six hours Cre induction. Fluorescence values of single cells are normalized to the average fluorescence of a non-randomized control strain with the strongest promoter *P_{ACT1}* driving expression (N = 9,449). 1D expression histograms on the edges of the 2D histogram show the tight population variability of each of the four component promoters: green-*P_{ACT1}*, blue-*P_{ADH1}*, red-*P_{MYO2}*, yellow-minimal TATA box. C) 1D expression histograms from clones picked from a randomized population. D) Genomic PCR from a randomized population using the scheme in (A). Larger bands in each column correspond to RFP, smaller ones to YFP. Four genotypes corresponding to each of the two tandem promoter pairs are seen in the population. E) Genotyping individual clones picked from a randomized population using genomic PCR as in D). In each gel, the four lanes proceeding the ladder in each gel correspond to the primer combinations in D). Numbers on each gel label the clone and its corresponding fluorescence profile is also labeled in (C). Two of the clones (far right) are examples of strains in which one reporter cannot be detected through PCR and has an expression profile which is indistinguishable from autofluorescence measured from a strain containing no fluorescent reporter.

Figure 3. Role of selected target genes in Base Excision Repair (BER) of DNA lesions

DNA bases with chemical modifications from damage are first converted to abasic sites. The error-free repair of these sites begins with cleavage of the phosphodiester linkage by AP endonucleases, coded by *APN1* and *APN2* in yeast. Unprocessed abasic sites can lead to replication fork stalling during DNA synthesis. The replication fork is potentially restarted by error-prone trans-lesion polymerases, including the heterodimeric Pol ζ , which is coded by the *REV3* and *REV7* genes in yeast.

Figure 4. Modulating base excision repair pathways to easily tolerate the genotoxic agent methylmethane sulfonate (MMS)

A) A protocol for generating a gene expression library containing combinations of three expression levels at desired loci. At each step, the expression levels which exist in the population for each gene are written to the right of the gene name. In the first two steps, the populations are isogenic, with each cell containing the expression states shown next to each gene. After sporulation, random assortment results in a genetically heterogeneous haploid population. Finally, randomization with Cre expression results in a population where each target locus has one of three possible expression levels. These populations can then be plated with or without selection and phenotypically scored based on colony size. B) PCR using genomic DNA isolated from the sporulated, randomized population indicates all possible genotypes are present at each locus. C) Selected clones were able to grow on plates with YPD + 0.035% MMS. Their growth rates in liquid YPD + 0.0035% MMS media were compared to the wild type strain (W303 background). Mean growth rate is plotted. Error bars represent 95% c.i. (based on a student's *t*-test with three biological replicates).

Table I

Six copy 2-state promoters produce a wide range of genotypes after switching

Mol.State														Total ^b	
Expr State	RFP	ON	ON/OFF	OFF	ON	ON/OFF	OFF	ON	ON/OFF	OFF	ON	ON/OFF	OFF	OFF	OFF
RFP/YFP/CFP	XFP	ON	ON	ON	ON/OFF	ON/OFF	ON/OFF	ON/OFF	ON/OFF	ON/OFF	ON/OFF	ON/OFF	ON/OFF	ON/OFF	ON/OFF
ON/ON/ON		5 (1.8) ^a	3 (1.4)		0 (3.8)	2 (3.1)									10/37
ON/ON/OFF					0 (2.2)	2 (2.8)									2/71
ON/OFF/ON			1 (0)		2 (2.8)	2 (2.2)									5/10
ON/OFF/OFF								4 (3.3)	2 (2.7)						6/15
OFF/ON/ON				2 (0.9)			2 (2.1)								4/10
OFF/ON/OFF							3 (5)								3/12
OFF/OFF/ON							4 (4)								4/11
OFF/OFF/OFF														4 (4)	4/21

^a Value is the number of cells found in the particular molecular state (as determined by PCR). Parenthetical values indicate expected number of cells in each molecular state assuming each allele inverts independently, given the observed fluorescent expression state.

^b Number of single clones with the observed fluorescent expression state (167 total). 38/167 were probed by PCR

Table II

Verified genotypes and growth rates of selected clones

clone	APN1	APN2	REV3	REV7	ADE2	A/α	Growth rate ^a (hr ⁻¹)
1	ON	ON	NAT	NAT	-	A	0.56±0.12
2	NAT	ON	ON	OFF	-	α	0.54±0.11
3	ON	ON	ON	OFF	+	A	0.53±0.1
4	ON	ON	NAT	ON	+	A	0.49±0.06
5	NAT	ON	ON	OFF	-	A	0.48±0.06
6	ON	ON	NAT	ON	-	A	0.47±0.05
7	ON	ON	NAT	ON	+	α	0.47±0.05
8	NAT	ON	ON	OFF	-	α	0.43±0.03
9	ON	ON	NAT	ON	-	A	0.42±0.02
10	ON	OFF	ON	OFF	+	A	0.32±0.02
11	ON	OFF	NAT	ON	-	α	0.23±0.02
W303	NAT	NAT	NAT	NAT	-	A	0.13±0.01
W303	NAT	NAT	NAT	NAT	+	A	0.13±0.02

^aGrowth range is based on the mean and 95% confidence interval from a student's *t*-test based on three biological replicates.