UC Berkeley UC Berkeley Previously Published Works

Title

High-Resolution Scanning of Optimal Biosensor Reporter Promoters in Yeast

Permalink

https://escholarship.org/uc/item/4d10205b

Journal

ACS Synthetic Biology, 9(2)

ISSN

2161-5063

Authors

Ambri, Francesca D'Ambrosio, Vasil Di Blasi, Roberto <u>et al.</u>

Publication Date

2020-02-21

DOI

10.1021/acssynbio.9b00333

Peer reviewed

High-resolution scanning of optimal biosensor reporter promoters in yeast

Francesca Ambri^{1#}, Vasil D'Ambrosio^{1#}, Roberto Di Blasi¹, Jerome Maury¹, Simo Abdessamad Baallal Jacobsen ¹, Douglas McCloskey¹, Michael K. Jensen^{1*} & Jay. D Keasling¹⁻⁵

¹ Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark
 ² Joint BioEnergy Institute, Emeryville, CA, USA
 ³ Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
 ⁴ Department of Chemical and Biomolecular Engineering & Department of Bioengineering, University of California, Berkeley, CA, USA

⁵ Center for Synthetic Biochemistry, Institute for Synthetic Biology, Shenzhen Institutes of Advanced Technologies, Shenzhen, China

* To whom correspondence should be addressed. Michael K. Jensen: Email: mije@biosustain.dtu.dk, Tel: +45 6128
4850

These authors contributed equally to this study

Abstract

Small-molecule binding allosteric transcription factors (aTFs) derived from bacteria enable real-time monitoring of metabolite abundances, high-throughput screening of genetic designs, and dynamic control of metabolism. Yet, engineering of reporter promoter designs of prokaryotic aTF biosensors in eukaryotic cells is complex. Here we investigate the impact of aTF binding site positions at single-nucleotide resolution in >300 reporter promoter designs in *Saccharomyces cerevisiae*. From this we identify biosensor output landscapes with transient and distinct aTF binding site position-effects for aTF repressors and activators, respectively. Next, we present positions for tunable reporter promoter outputs enabling metabolite-responsive designs for a total

of four repressor-type and three activator-type aTF biosensors with dynamic output ranges up to 8- and 26-fold, respectively. This study highlights aTF binding site positions in reporter promoters as key for successful biosensor engineering, and that repressor-type aTF biosensors allows for more flexibility in terms of choice of binding site positioning compared to activator-type aTF biosensors.

Keywords biosensor, transcriptional regulators, operator position, yeast,

Introduction

The upsurge of genetic editing capacity in the last decades has allowed microbial cells to reach the spotlight as key players in establishing bio-sustainable alternatives to the production of fine chemicals and fuels ¹. At the heart of robust successful microbial cell factory engineering is the equilibrium between cell growth and the production of valuable metabolites. Finding equilibrium between native growth-supporting metabolism and burdening expression of metabolic pathway genes is often investigated by targeted perturbations of gene regulation using regulatory parts libraries with diversity in their genetic components, expression level, and copy number ^{2,3}.

Following successful design and construction of cell factory libraries, screening isoclonal cell factory designs for optimal product titers is often inferred from chromatography techniques in a low- to semi-throughput manner, ultimately impacting both cost and time required for optimal cell factory development ⁴. As a means to mitigate high costs and long development time-spans for cell factory engineering, development of small-molecule binding biosensors based on allosteric transcription factors (aTFs) from prokaryotes has emerged as a promising research strategy ⁵. Indeed, engineering aTFs to couple the abundance of a target metabolite or chemical of interest,

and report a screenable output which can be analysed in high-throughput (e.g. fluorescence or growth), has enabled screening and selection of cell factory libraries in multiplex, as well as dynamic control of metabolism, and production stability ^{6–9}.

However, when engineering aTF-based biosensors in eukaryotes the general strategy needs to include means to express both the prokaryotic candidate aTF(s), and a reporter gene controlled by a promoter engineered with one or more prokaryotic aTF binding sites (ie. operators)¹⁰. While no generally applicable design exists, a few guiding principles should be adopted to increase the chances of successfully onboarding prokaryotic aTFs in yeast and other eukaryotes. First, when employing repressor-type aTFs, a strong promoter driving the expression of both the aTF and the reporter gene has been successfully reported ^{6,11–14}. In principle, this allows both a high level of repression in the absence of the ligand (OFF state), and a high dynamic range of the aTF-based biosensor upon ligand-induced derepression of reporter gene expression when the aTF detach from the DNA (ON state). In the case of transcriptional activators, employing a stronger promoter to drive the transcription of the reporter gene has been shown to cause leaky expression, and also to reduce the dynamic range, while optimal choice of promoter for controlling the expression of the allosteric transcriptional activator should be determined by testing a few native promoters covering a large expression range ¹⁵. While aTF expression is easily regulated by the use of native promoters, synthetic reporter promoters driving the conditional expression of biosensor outputs often rely on a number of engineering parameters, including dampened native regulation by removal of upstream-regulatory sequences (URSs)^{15,16}, in combination with searches for optimal operator positioning within the eukaryotic promoter ^{6,13,15}. Indeed, the evolutionary encoding of prokaryotic aTF designs is often aTF-specific with respect to combination of operator length, sequence motif, and position relative to the core promoter, impacted by both aTF oligomerization

and mode-of-action of the specific aTF¹⁷. Based on this, the optimization of reporter promoters for aTF-based biosensor designs in eukaryotes has proceeded by trial and error, most often by investigation of a few random insertion sites for single or multi-copy operator positioning around TATA boxes and/or the transcription start site (TSS) of the chosen promoter ¹⁰. As such, in order to quantitatively assess the impact of operator positioning, and from this elucidate designs for optimal biosensor outputs, there is a need for systematic scanning of operator positions in eukaryotic reporter promoter designs.

Here we report the deep-scanning of operator positioning in weak and strong yeast promoters. As initial benchmarks, we investigated the repressor-type vanillic acid-responsive biosensor VanR from the GntR aTF family ^{18,19}, and the activator-type *cis*, *cis*-muconic acid (CCM) and protocatechuic acid biosensors BenM and PcaQ, respectively, from the Lys-R type transcriptional regulator (LTTR) family ^{15,20}. By systematic evaluation of more than three-hundred reporter promoter designs we here lay out design principles that enable successful small-molecule biosensing of a total of seven new and previously engineered biosensors in yeast.

Results and Discussion

To identify optimal designs for aTF-binding sites in yeast promoters, we choose biosensor systems with different modes of action, oligomerization, and operator lengths for an introductory deep-scanning of optimal operator positioning ^{19,21,22}. Initially, VanR from *Caulobacter crescentus* was chosen as a candidate repressor-type biosensor aTF ²³, whereas the LTTR BenM from *Acinetobacter baylyi* ADP1 was chosen as a candidate for an activator-type biosensor ¹⁵. VanR enables de-repression of gene expression when dislocated from promoter binding upon binding to

vanillate ^{24,25}, whereas BenM serves as a transcriptional activator upon binding of its cognate ligand CCM ²⁶. In addition to VanR and BenM, we also included the LTTR activator protocatechuic acid biosensor PcaQ from *Sinorhizobium meliloti*. PcaQ has previously been reported to be non-functional in eukaryotes, yet protocatechuic acid is an important precursor chemical for a broad range of value-added chemicals for which a biosensor could be a promising metabolic engineering tool ^{27–29}. Here, we sought to further investigate if deep-scanning operator site positions would enable successful biosensing of protocatechuic acid in yeast using PcaQ.

In terms of reporter promoter design for the three aTF-based biosensors, we analysed the VanO operator library in the context of the strong constitutive TEF1 promoter (pTEF1)(Fig. 1A), while the weak truncated CYC1 (209bp_pCYC1) promoter was used as a chassis to screen for optimal operator positions for the activator-type biosensors BenM or PcaQ since the only transcriptional activators reported so far in *S. cerevisiae* rely on this promoter to drive the expression of the reporter gene (Fig. 1A)¹⁵. The strong TEF1 promoter was chosen because transcriptional repressors have previously been demonstrated to rely on strong reporter promoters to enable simple validation of successful expression of aTF repressors in non-native hosts driving the expression of the reporter gene (i.e. low OFF state when aTF repressor is expressed)^{6,11,12,14,30}. In all three instances the libraries were constructed to enable control of expression of the reporter gene fluorescence protein (yeGFP).

Taking into account the presence of TATA boxes, TATA-like elements, and transcription start site (TSS) in the two chosen yeast promoters ^{15,31}, we constructed a total of 320 synthetic reporter promoters for the three candidate biosensors (Fig. 1A). For the vanillic acid biosensor VanR, 91 synthetic TEF1 promoters carrying two VanR-operator sites separated by six nucleotides (2xVanO), spanning all single nucleotide positions between the TATA-like element and the TSS

of the TEF1 promoter ³¹, were constructed (Fig. 1A, Supplementary Table S1). This double operator design, spaced by the Eco47III restriction site was used since it was previously reported to have the highest dynamic range among tested VanO design in mammalian cells.¹⁸ For the CCM and protocatechuic acid biosensors BenM and PcaQ, libraries covering a total of 106 and 133 single nucleotide positions of either BenO or PcaO operators in 209bp_pCYC1, were constructed (Fig. 1A, Supplementary Table S1). Because of the relatively large size of the operator sequences compared to the length of the truncated CYC1-based promoter, we originally included only one BenO operator site when constructing the first successful activator-type biosensor for BenM¹⁵, and this design was therefore also adopted for the BenO and PcaO operator libraries in this study. Each design controlling the expression of yeGFP was genomically integrated in strains expressing the corresponding aTF (Fig. 1A). Next, for VanR and BenM, fluorescence values collected in the presence (ON state) and in the absence (OFF state) of inducers, vanillic and cis, cis-muconic acid, respectively, supplemented to the medium (Fig. 1B and Fig. 1C). The PcaO reporter promoter library was integrated into a protocatechuic acid (PCA)-producing strain expressing the gene encoding 3-dehydroshikimate dehydratase (3DSD) from dung mold Podospora pauciseta known to catalyze the conversion of 3-dehydroshikimic acid to PCA in S. cerevisiae^{28,32}. Because of the low OFF state provided by the truncated 209bp_pCYC1 promoter, we initially collected only ON state data assuming that any significant increase in fluorescence could be a direct result of PcaQmediated activation of GFP expression. All the following analysis performed on selected PcaO positions will be then compared to strains without the inducer. (Fig. 1A and Fig. 1D). Yeast strains expressing the aTFs and yeGFP under the control of strong pTEF1 and weak 209bp_pCYC1 promoters were used as controls throughout the study.

For the deep-scanning of optimal VanO positions, we tested 81/91 (89%) pTEF1-2xVanO designs resulting in significant changes (*t*-test, p < 0.05) in ON state compared to OFF state reporter outputs in all positions tested, with the majority of the library (80%) showing ligand-induced changes in yeGFP expression > 2-fold (Fig. 1B). Of particular interest in terms of biosensor dynamic output range, the region covering VanO operator positions 95-78 enabled gradual increases and decreases in vanillic acid-induced reporter outputs spanning fold-changes of 4.2-8.6, while maintaining low OFF state expression (Fig. 1B). Contrastingly, operator library designs extending from position 71 to TSS had higher OFF states, corroborating earlier findings that expanding the distance between TATA box and operator can lead to increases in basal transcription levels of engineered promoters³³, and thus lowering the overall fold-change, as observed here in response to vanillic-acid (Fig. 1B). We speculate that increasing the distance from the TATA box can compromise the ability of the repressor to obstruct RNA polymerase progression or assembly of the transcriptional machinery, and highlight operator positioning within 40-50 bases from the TATA-box to enable the strongest repression.

For the 209bp_pCYC1_BenO library screen, 96 out of the 106 constructed and characterized designs had < 2-fold changes in yeGFP readout upon CCM feeding (Fig. 1C). However, among the 10 designs with ligand-induced fold-changes higher than 2-fold, we observed a distinct hotspot for BenM activation within the minimal design space covering positions 181-183, enabling CCM-induced fold-changes from 9.3-10.6, while BenO at position 197 enabled a fold-change of 4.3 (Fig 1C). The identification of position 183 for functional prokaryotic activator-type biosensor designs in yeast corroborates our earlier studies ¹⁵. Furthermore, as for the VanO library screen, CCM feeding did not lead to changes in yeGFP expression of the control reporter promoter 209bp_pCYC1 (Fig. 1C).

Finally, for the PcaO reporter promoter library, we characterized yeGFP expression for all possible 133 209bp_pCYC1_PcaO promoter designs in yeast strains expressing *3DSD*. Following 16 hours of cultivation, yeGFP expression was measured and compared to control expression from the 209bp_pCYC1 promoter without PcaO (Fig. 1A). From this analysis, we observed that while the majority (85/133) of operator positions enabled > 2-fold higher expression compared to the 209bp_pCYC1 control promoter, distinct activation of yeGFP expression was observed for positions 163, 144, 145, and 108, enabling changes in yeGFP expression from 8- to 30.3-fold compared to a "no-operator" control strain (Fig. 1D).

While there is limited mechanistic evidence on how prokaryotic transcriptional activators function in eukaryotes ³⁴, aTF members of the LTTR subfamily exert transcriptional regulation in prokaryotes by binding as tetramers to two operators, the regulatory binding site (RBS) and activating binding site (ABS), simultaneously ²². In the OFF state this causes DNA bending, thereby preventing access of the RNA polymerase to specific sequences of the target promoter called UP, whereas in the ON state the aTF is displaced from the ABS which subsequently causes relaxation of DNA bending and closer proximity of the RNA polymerase to UP, and ultimately promotes transcription ^{35,36}. Following this mechanism, we hypothesize that BenM and PcaQ, and potentially a multitude of other aTF members of the LTTR subfamily, only function as aTFs in eukaryotic transcriptional regulation in a narrow window of operator positions to ensure native transcriptional machinery can discrimination between OFF and ON states, and that this likely depends on both the size of the aTF operator and the aTF itself. Indeed, the operator sequence of BenM is composed of three sub-binding sites, of which BenM is constitutively bound to site 1 and shifting binding between site 2 or 3, depending on ligand presence ²², while only one binding site

has been experimentally determined for PcaQ so far ³⁷, further suggesting that PcaQ interacts with the reporter promoter differently compared to BenM ^{37,38}.

In summary, deep-scanning operator positions in > 300 reporter promoter identified an 18bp (positions 78-95) region in the native TEF1 promoter that supported both low OFF state expression and up to 8.6-fold ligand-induced de-repression for the repressor-type biosensor VanR, while for activator-type biosensors BenM and PcaQ, optimal operator positioning was defined at single-nucleotide resolution with BenO positions 197, 183, 182 and PcaO positions 163, 145 and 108, yielding up 10.6- and 30.3-fold ligand-induced activation of reporter promoters, respectively.

To more broadly investigate the regional bias of operator designs for repressor-type biosensors and the functional dichotomy of BenO-PcaO operator designs, we next decided to test selected operator positions in other biosensor systems.

First, from the VanO operator library, positions 105, 94 and 83 were selected as low-, middle- and high-response operator sites, with ligand-induced fold-changes of 2.5-, 5-, and 8.4-fold, respectively (Fig. 1B, Fig. 2A). Next, the operators of three reporter promoter designs were replaced by a single copy operator for the new batch of repressor-type aTF biosensors; TetR from *E. coli*, XylR from *Caulobacter crescentus*, and FapR from *Bacillus subtilis* (Fig. 2B). The choice of candidate repressor-type aTF biosensors were guided by the interest i) to cover archetypical variants of these three families of transcriptional regulators, and ii) to span variable operator sequence lengths, from a minimum of 17bp for TetR to 34bp in the case of FapR ^{6,13,39}. Moreover, TetR, XylR, and FapR have previously been reported to work as functional biosensors for anhydrotetracycline, xylose, and malonyl-CoA, respectively, in yeast, thus enabling a relevant benchmark to previous studies ^{6,13,39}. While TetR and XylR biosensor responses to non-native metabolites tetracycline and xylose are readily detected from feeding the ligands to the cultivation

medium ^{13,39}, the characterization of the FapR biosensor output for detection of endogenous malonyl-CoA pools can be assessed by feeding of cerulenin, a fatty acid synthase (FAS) inhibitor that allows the increase of intracellular malonyl-CoA pools ⁶. Similar to the VanR, for each of the three new biosensor systems we included a control strain without an aTF operator in the reporter promoter, and screened all strain designs in the OFF and ON states (Fig. 2B-C). From this screen we observed that position 83, the best performing position selected from the VanO library screen, only supported 1.9- and 1.4-fold induction in fluorescence for XylR and FapR, respectively, while for the TetR biosensor no change in fluorescence was observed (Fig. 2C). However, for XylR, XylO positions 105 and 94 enabled fluorescence induction by 3.4- and 5.2-fold, respectively, which is comparable to the fold-change observed for the XylR biosensor system in yeast previously published, albeit using another promoter chassis ¹⁴. Similarly, for the TetR-TetO biosensor system, both positions 105 and 94 enabled ligand-induced changes in fluorescence of 5.7- and 2-fold, respectively, while the FapR-FapO biosensor induced yeGFP expression by approx. 2-fold for each of these two operator positions (Fig. 2C). As observed from the VanO library, the difference in ON/OFF fold-change is largely affected by the reporter promoter activity in the OFF state (Fig. 1B, Supplementary Figure S1). For instance for the FapO-FapR designs without supplemented ligand (OFF state), positions 105 and 94 support 84% and 80% repression respectively, compared to modest 53% repression at position 83 (Supplementary Figure S1). Importantly, the OFF state repression for FapR observed in this study using a single operator site at position 105 is comparable to a design obtained with three operator sites, located \pm 50bp from TSS, previously published ⁶. Likewise, it should be noted that FapR together with VanR were the only biosensors that did not completely de-repress under the tested ON conditions (13.5 µM

cerulenin and 4 mM vanillic acid) as inferred from the highest ON state at position 83 compared to the control strains without FapR or VanR expressed (Supplementary Figure S1).

Next, for the activator-type biosensors, again based on designs supporting tunable outputs, we selected three designs from both the BenO (positions 197, 183, 182) and the PcaO (positions 163, 145 and 108) library screens (Fig. 1B-C, Fig. 3A), and tested their relevance for biosensing based on other activator-type biosensors from the large LTTR family of aTFs (Fig. 3B). The new activator-type biosensor candidates comprised the BenM homolog and CCM-responsive CatM from Acinetobacter baylyi ADP1 40, and the flavonoid-responsive nodulation genes regulator NodD from *Rhizobium leguminosarum*⁴¹, neither of them previously successfully engineered in eukaryotes. Additionally, we tested two previously studied activator-type biosensors, namely the malonate-responsive MdcR from Klebsiella pneumoniae and arginine-responsive ArgP from E. coli, respectively, of which ArgP has not previously been tested in yeast (Fig. 3B)^{15,42,43}. As strong promoters were successfully used for repressor-type aTF biosensors (Fig. 2B), by default we also chose the strong constitutive TDH3 promoter (pTDH3) for expression of all activator-type aTFs, except NodD and BenM¹⁵, which we expressed from the weak REV1 promoter (pREV1)(Fig. 3B). Interestingly, for NodD, we never succeeded in obtaining any viable transformants when we attempted to transform pTDH3::NodD into yeast (data not shown).

As for the repressor screen, we tested all cognate aTF and operator designs in the presence (ON) and absence (OFF) of inducer ligands, and also included the 'no operator' 209bp-pCYC1 reporter promoter as a control (Fig. 3B-C). From the analysis, beyond validation of the CCM and PCA biosensors (Fig. 1C-D, Fig. 3C), we observed that the malonate biosensor MdcR enabled yeGFP induction ranging from 1.6- to 2.1-fold for all six tested operator positions, with the highest induction reached for reporter promoter with MdcO inserted at position 163, although this position

also giving rise to high OFF state expression (Fig. 3C). MdcR was the only newly tested biosensor that enabled a > 2 fold-change induction, whereas ArgP only sustained a negligible 1.2-fold induction of yeGFP expression for the best operator position 183. In contrast, the CCM and flavonoid biosensors CatM and NodD supported modest 1.3- and 1.6-fold yeGFP induction for design 108 (Fig. 3C). With respect to NodD, cell viability was exclusive observed in the presence of the NodD operator, indicating that this system may be working properly; further characterization and engineering is needed to understand possible detrimental "off-target" events of NodD when no operator is present in the genome. Furthermore, despite LTTRs being structurally and functionally conserved ²⁰, BenM, PcaQ, and MdcR were the only activator-type aTFs enabling > 2-fold reporter gene induction in this study (Fig 3C). Nevertheless, for BenM and PcaQ, the biosensors displayed drastic opposite readouts for their optimal operator-reporter designs, namely 183-182 and 108, respectively (Fig. 3B), implying deeper layers of complexity in the mechanism of induction. Indeed, more detailed knowledge of the specific promoter sequences is required for predictive aTF placement for efficient and conditional RNA polymerase recruitment 27.40.

In summary, from the regional operator screen for functionalization of other repressor-type biosensors, we observed that all four tested aTFs, including VanR, employing operator positions 105 and 94 gave rise to functional biosensor systems (Fig. 2C). Furthermore, we observed that the leakiness of repressor-type biosensor systems increased when the operator site was located more than 40 nucleotides from the TATA-like element (Fig. 1B). Regarding the activator-type biosensors, although no universal design was established, we were able to develop a new activator-based biosensor in yeast. Moreover, our results suggest that onboarding prokaryotic transcriptional activators in *S. cerevisiae* may require deep-scanning of operator positions, and furthermore

substantiate our previous finding ¹⁵, demonstrating truncated minimal 209bp_pCYC1 promoter as a platform promoter chassis for engineering reporter promoters for transcriptional activator-type biosensors in yeast. LIkewise, given its exceptional >20-fold increase in yeGFP expression in a PCA production strain without any further engineering of the aTF, we envision the PCA biosensor PcaQ could be useful for screening and selection of metabolic routes to the shikimate-based products like vanillin glucoside, catechol, and *cis-cis*-muconic acid ^{28,44}.

Lastly, it should be mentioned that beyond testing of operator positioning, we envision that libraries alike the one functionally characterized in this study, would be interesting to mine biochemically by investigating aTF:DNA binding affinity, and also through further functional characterization using combinatorial approaches of aTF biosensor design parameter (e.g. aTF expression and translation initiation). Ultimately, such multivariate parameter testing should enable more predictive engineering of new aTF biosensors for biotechnological and medical applications.

Methods

Strains, chemicals and media

Escherichia coli strain DH5 α was used as a host for cloning and plasmid propagation. Cells were grown at 37°C in 2xYT medium supplemented with 100 µg/mL ampicillin (900 mL deionized water, 16 g bacto tryptone, 10 g bacto yeast extract, 5 g NaCl, adjust the pH to 7, deionized water up to 1 L, autoclave). *Saccharomyces cerevisiae* strains CEN.PK102-5B (MAT*a ura3-52 his3* Δ *1 leu2-3*/*112 MAL2-8*^c*SUC2*) and CEN.PK113-5A (MAT*a*, *trp1 his3* Δ *1 leu2-3*/*112 MAL2-8*^c*SUC2*) were used as basic strains for construction of biosensor systems. Strains were grown at 30°C in synthetic complete (SC) medium (6.7 g yeast nitrogen base without amino acids, appropriate amount of drop-out medium supplement (Sigma-Aldrich), deionized water up to 900 mL, adjust pH to 5.6, (20 g agar in case of plates), autoclave, 100 mL 20% (w/v) glucose); yeast extract peptone dextrose (YPD) complete medium (10 g BactoYeast extract, 20 g BactoPeptone, 20 g Dextrose, (20 g agar in case of plates), deionized water up to 1 L, autoclave); and in Delft minimal medium (per L: 7.5 g/L (NH₄)₂SO₄, 14.4 g/L KH₂PO₄, 0.5 g/L MgSO₄*7H₂O, 2mL/L trace metals, 100mL/L 20% glucose, 1mL vitamins) at pH 4, 4.5 or 6. Trace metal and vitamins solution were prepared according to Ambri et al. ⁴⁵. All reagents were purchased from Sigma-Aldrich unless otherwise specified.

Yeast strains carrying the different biosensor designs were inoculated in Delft minimal medium and grown over-night. Next, S. cerevisiae cells were diluted 1:50, with the exception of XylR and FapR strains which were diluted 1:25, into fresh Delft minimal medium in the presence or absence of inducer. Fluorescence data were collected following defined periods of cultivation depending on the chemical characteristics of the ligand (Supplementary Table S2), and according to previous studies ^{6,14,15,39} More specifically, for strains expressing repressor-type biosensor, TetR from E. coli and VanR from Caulobacter crescentus were grown in the presence of 2.3 µM anhydrotetracycline and 4 mM vanillic acid, respectively, for 16 hours prior to fluorescence measurements. Strains carrying XylR from C. crescentus were cultivated in the presence of 13.3 mM D-(+)-xylose and cultivated for 6 hours before measuring the fluorescence, while strains carrying FapR from Bacillus subtilis were grown for 6 hours before addition of 13.5 µM cerulenin, following which fluorescence was measured 2h later. All strains expressing repressors were grown at pH 6, except for strains expressing VanR cultivated at pH 4. For strains expressing activatortype biosensor, BenM and CatM from Acinetobacter baylyi ADP1 were grown at pH 4.5 in the presence of 1.4 and 3 mM of cis-cis-muconic acid, respectively; strains carrying ArgP from *Escherichia coli* were cultivated 4.5 at pH in the presence of 10 mM of L-Arginine; strains expressing MdcR from *Klebsiella pneumoniae* were grown at pH 4.5 in the presence of 10 mM of malonate; strains carrying NodD from *Rhizobium leguminosarum* were grown at pH 6 in the presence of 0.2 mM of naringenin. All fluorescence data were collected after 22 hours (Skoedjt 2016). Finally, strains expressing PcaQ from *Sinorhizobium meliloti* in combination with genomic integration of the *Podospora pauciseta 3DSD* gene were grown at pH 6, and fluorescence data were collected after 16 hours.

Library design

The VanO operator library was designed by inserting 2x VanO sequences linked by a 6 nucleotide spacer, agcgct¹⁸, in front of every nucleotide between the TATA-like box and the TSS of the native TEF1 promoter. The constructed and characterized library ultimately comprised 81 out of all 91 possible designs. The BenO and PcaO operator libraries were designed inserting BenO sequence (73 bp) or PcaO sequence (80 bp) in front of every nucleotide in the truncated 209bp-CYC1 promoter, with the exception of the positions encoding the two TATA boxes and the TSS. The constructed and characterized BenO library comprised 106 out of all 133 possible designs, while the PcaO operator library spanned all 133 operator positions. For all reporter promoter designs, a customized R script was used for the systematic insertion of operator sequences in TEF1 and 209bp_pCYC1 promoters prior to DNA synthesis.

Synthetic genes and plasmid construction

Codon-optimized genes, promoter libraries, and oligonucleotides were commercially synthesized (Twist Bioscience; Integrated DNA Technologies, Inc). All genetic components were amplified with Phusion U Hot Start DNA Polymerase Master Mix (Thermo Fisher Scientific, Inc.). Plasmids were constructed, propagated, digested and transformed into yeast according to Jensen et al. ⁴⁶(Supplementary Tables S3-6). All constructed plasmids are sequence-verified by Sanger sequencing (Eurofins). Agarose 1% was used for analysis of nucleic acids by gel electrophoresis.

The plasmids carrying the single-copy operator for FapR, XylR and TetR in positions 105, 94 and 83 were constructed by PCR, then USER cloned into plasmids, and subsequently propagated, sequence-verified, digested, and transformed into yeast according to Jensen et al. ⁴⁶. The same approach was applied when constructing plasmids carrying single-copy operators for CatM, MdcR, ArgP and NodR. All reporter promoter sequences can be found in (Supplementary Table S7).

Flow cytometry measurements and data analysis

Prior to flow cytometric acquisition, strains were pre-inoculated and sub-cultivated in 0.5 mL Delft minimal medium in regular 96-format polypropylene deep-well plates as specified for the cultivation conditions for each biosensor. Next, cells were diluted 1:20 in 0.2 ml phosphate-buffered saline (PBS) in regular flat-bottom 96-well plates in order to arrest cell growth prior to flow cytometry acquisition using a Becton Dickinson LSR FORTESSA with a blue 488 nm laser set at FCS 150 V, SSC 250 V and FITC-A 510V or 450V, where the lower voltage was applied for strains carrying the reporter promoter without the operator because of detection limits. For each measurement 10,000 single-cell events were recorded using BD FACSDIVATM software. Raw data were analyzed using FlowLogic software (Inivai Technologies). Mean fluorescence intensity (MFI) of gated populations were determined, and fold changes calculated as the ratio between induced (ON state) and non-induced (OFF state). For all fluorescence measurements, data

represent average of three biological replicates and error bars correspond to standard deviation between measurements.

Acknowledgements

This work was supported by the Novo Nordisk Foundation, and by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie action PAcMEN (grant agreement No. 722287)

Author contributions

F.A., M.K.J. and J.D.K. conceived the study; F.A. and V.D. designed the experiments, F.A., V.D., J.M., S.A.B.J., D.M., and R.D.B. performed the experiments; F.A. and V.D. collected and analyzed the flow cytometric data with the support of R.D.B.; F.A., V.D. and M.K.J. wrote the manuscript.

Conflict of interest

Authors declare no conflict of interest.

Figure legends

Figure 1. Single-nucleotide position scanning for optimal aTF biosensor operators in yeast reporter promoters. (**A**) Schematic outline of the experimental design used to scan for optimal operator positions for prokaryotic aTFs VanR, BenM, and PcaQ in yeast promoters pTEF1 and 209bp_pCYC1. Red marks depict VanO, BenO and PcaO operator positioning in pTEF1 and 209bp_pCYC1 promoters. Black marks labelled T1 and T2 represent TATA boxes 1 and 2. (B) Screening yeast strains expressing yeGFP controlled by 81 different reporter promoters with VanO operator positioned in 81 out of 91 possible single-nucleotide positions between the TATA-like element and transcription start site (TSS) of pTEF1 promoter. Mean fluorescence intensity (MFI) values from flow cytometry measurements of yeGFP intensities in the presence or absence of 4 mM vanillic acid following 16 hours of incubation. (C) Screening yeast strains expressing yeGFP controlled by 106 different reporter promoters with BenO operator positioned between the 5'-end and TSS of the 209bp_pCYC1 promoter, with numbers labelled relative to the 5'-end of the truncated promoter. MFI values from flow cytometry measurements of yeGFP intensities in the presence or absence of 1.4 mM CCM following 22 hours of incubation. (D) Screening yeast strains expressing yeGFP controlled by 133 different reporter promoters with the PcaO operator positioned between the 5'-end and TSS of the 209bp_pCYC1 promoter. MFI values from flow cytometry measurements of yeGFP intensities in strains expressing the gene encoding the 3dihydroshikimate dehydratase (3DSD) from dung mold Podospora pauciseta converting 3dehydroshikimic acid to PCA following 16 hours of incubation. (B-D) MFI values are shown as mean \pm s.d. from at least three (n = 3) biological replicate experiments. AU, arbitrary units. Horizontal dashed lines indicates 2-fold cut-off in MFI between yeGFP expression from cells cultivated in the presence or absence of inducer (B-C), or compared to control strains expressing yeGFP without any operator in 209bp_pCYC1 (D). Vertical dashed lines indicate TATA box positions in the 209bp_pCYC1 promoter. Operator designs and genotypes of all strains are listed in Supplementary Tables S1 and S6.

Figure 2. Testing of selected operator-promoter designs in aTF repressors. (A) Overlay of fluorescence intensity profiles of the operator-promoter designs controlling the expression of yeGFP selected from the single nucleotide deep-scanning of the VanO library (Fig. 1B). (B) Schematic outline of the experimental design used to test for optimal operator positions for prokaryotic repressor aTFs VanR, XylR, FapR and TetR. Red marks depict the selected operator positioning in pTEF1. Black mark labelled T1 represents the TATA box. (C) Mean fluorescence intensity (MFI) values from flow cytometry measurements of yeGFP intensities for the selected operator-promoter designs in the presence (ON state) and absence (OFF state) of 4 mM vanillic acid following 16 hours of incubation (VanR + VanO); 13.3mM D-(+)-xylose following 6 hours of incubation (XylR + XylO); 13.5 μ M cerulenin following 2 hours of incubation (FapR + FapO); 2.3 µM anhydrotetracycline following 16 hours of incubation (TetR + TetO). MFI values are shown as mean \pm s.d. from three (n = 3) biological replicate experiments. AU, arbitrary units. Dashed line indicates 2-fold cut-off in MFI between yeGFP expression from cells cultivated in the presence or absence of inducer. Operator designs and genotypes of all strains are listed in Supplementary Tables S1 and S6.

Figure 3. Testing of selected operator-promoter designs in LysR-Type Transcriptional Regulators (LTTRs). (A) Overlay of fluorescence intensity profiles of the operator-promoter designs controlling expression of yeGFP selected from the single nucleotide deep-scanning of PcaO and BenO libraries (Fig. 1C-D). (B) Schematic outline of the experimental design used to test for optimal operator positions in LTTR members BenM, PcaQ, CatM, MdcR, NodD and ArgP. Red marks depict the selected operator positioning in the minimal 209bp_pCYC1 promoter. Black marks labelled T1 and T2 represent TATA boxes 1 and 2. (C) Mean fluorescence intensity (MFI) values from flow cytometry measurements of yeGFP intensities for the selected operator-promoter designs were collected following 22 hours of incubation in the presence (ON state) and absence (OFF state) of 1.4 mM *cis-cis* muconic acid (BenM + BenO); 10 mM malonic acid (MdcR + MdcO); 3 mM *cis-cis* muconic acid (CatM + CatO); 0.2 mM naringenin (NodD + NodO); and 10 mM L-arginine (ArgP + ArgO). PcaQ + PcaO MFI values were collected following 16 hours cultivation of strains expressing (ON state) or not expressing (OFF state) the gene encoding *Podospora pauciseta* 3-dehydroshikimate dehydratase (*3DSD*). MFI values are shown as mean \pm s.d. from three (n = 3) biological replicate experiments. AU, arbitrary units. Dashed line indicates 2-fold cut-off in MFI between yeGFP expression from cells cultivated in the presence or absence of inducer. Operator designs and genotypes of all strains are listed in Supplementary Tables S1 and S6. * denotes strain not viable in ON state.

Supplementary information

Supplementary Figure S1. Activities of reporter promoter designs for various operators and positions (A) Mean fluorescence intensity (MFI) values from flow cytometry measurements of yeGFP intensities for the selected operator-promoter designs. MFI values are shown as mean \pm s.d. from three (n = 3) biological replicate experiments. AU, arbitrary units.

Table S1. Operator designs in reporter promoters. Here illustrated are VanO design 105 in pTEF1, BenO design 183 and PcaO design 108, both in 209bp_pCYC1; TATA box, TATA-like element and transcription start site are in bold; operator sequences are in italics.

Table S2. Sub-cultivation conditions.

Table S3. Synthetic gene sequences.

Table S4. Oligonucleotides sequences.

Table S5. Plasmids used in this study.

Table S6. Strains used in this study.

Table S7. Reporter promoters used in this study.

References

- (1) Walker, R. S. K.; Pretorius, I. S. Applications of Yeast Synthetic Biology Geared towards the Production of Biopharmaceuticals. *Genes* **2018**, *9* (7).
- (2) Lee, S. Y.; Kim, H. U. Systems Strategies for Developing Industrial Microbial Strains. *Nat. Biotechnol.* **2015**, *33* (10), 1061–1072.
- (3) Nielsen, J. Yeast Systems Biology: Model Organism and Cell Factory. *Biotechnol. J.* **2019**, e1800421.
- (4) Nielsen, J.; Keasling, J. D. Engineering Cellular Metabolism. *Cell* **2016**, *164* (6), 1185–1197.
- (5) Libis, V.; Delépine, B.; Faulon, J.-L. Sensing New Chemicals with Bacterial Transcription Factors. *Curr. Opin. Microbiol.* **2016**, *33*, 105–112.
- (6) David, F.; Nielsen, J.; Siewers, V. Flux Control at the Malonyl-CoA Node through Hierarchical Dynamic Pathway Regulation in Saccharomyces Cerevisiae. *ACS Synth. Biol.* 2016, 5 (3), 224–233.
- (7) Snoek, T.; Romero-Suarez, D.; Zhang, J.; Ambri, F.; Skjoedt, M. L.; Sudarsan, S.; Jensen, M. K.; Keasling, J. D. An Orthogonal and pH-Tunable Sensor-Selector for Muconic Acid Biosynthesis in Yeast. ACS Synth. Biol. 2018, 7 (4), 995–1003.
- (8) Rugbjerg, P.; Sarup-Lytzen, K.; Nagy, M.; Sommer, M. O. A. Synthetic Addiction Extends the Productive Life Time of Engineered Escherichia Coli Populations. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (10), 2347–2352.
- (9) Xiao, Y.; Bowen, C. H.; Liu, D.; Zhang, F. Exploiting Nongenetic Cell-to-Cell Variation for Enhanced Biosynthesis. *Nat. Chem. Biol.* **2016**, *12* (5), 339–344.
- (10) D'Ambrosio, V.; Jensen, M. K. Lighting up Yeast Cell Factories by Transcription Factor-Based Biosensors. *FEMS Yeast Res.* **2017**, *17* (7).
- (11) Teo, W. S.; Hee, K. S.; Chang, M. W. Bacterial FadR and Synthetic Promoters Function as Modular Fatty Acid Sensor- Regulators in Saccharomyces Cerevisiae. *Eng. Life Sci.* 2013, *13* (5), 456–463.
- (12) Li, S.; Si, T.; Wang, M.; Zhao, H. Development of a Synthetic Malonyl-CoA Sensor in Saccharomyces Cerevisiae for Intracellular Metabolite Monitoring and Genetic Screening. *ACS Synth. Biol.* **2015**, *4* (12), 1308–1315.
- (13) Hector, R. E.; Mertens, J. A. A Synthetic Hybrid Promoter for Xylose-Regulated Control of

Gene Expression in Saccharomyces Yeasts. Mol. Biotechnol. 2017, 59 (1), 24–33.

- (14) Teo, W. S.; Chang, M. W. Bacterial XylRs and Synthetic Promoters Function as Genetically Encoded Xylose Biosensors in Saccharomyces Cerevisiae. *Biotechnol. J.* 2015, 10 (2), 315–322.
- (15) Skjoedt, M. L.; Snoek, T.; Kildegaard, K. R.; Arsovska, D.; Eichenberger, M.; Goedecke, T. J.; Rajkumar, A. S.; Zhang, J.; Kristensen, M.; Lehka, B. J.; et al. Engineering Prokaryotic Transcriptional Activators as Metabolite Biosensors in Yeast. *Nat. Chem. Biol.* 2016, *12* (11), 951–958.
- (16) McIsaac, R. S.; Oakes, B. L.; Wang, X.; Dummit, K. A.; Botstein, D.; Noyes, M. B. Synthetic Gene Expression Perturbation Systems with Rapid, Tunable, Single-Gene Specificity in Yeast. *Nucleic Acids Res.* 2013, 41 (4), e57–e57.
- (17) De Paepe, B.; Peters, G.; Coussement, P.; Maertens, J.; De Mey, M. Tailor-Made Transcriptional Biosensors for Optimizing Microbial Cell Factories. *J. Ind. Microbiol. Biotechnol.* 2017, 44 (4-5), 623–645.
- (18) Gitzinger, M.; Kemmer, C.; Fluri, D. A.; El-Baba, M. D.; Weber, W.; Fussenegger, M. The Food Additive Vanillic Acid Controls Transgene Expression in Mammalian Cells and Mice. *Nucleic Acids Res.* **2012**, *40* (5), e37.
- (19) Zheng, M.; Cooper, D. R.; Grossoehme, N. E.; Yu, M.; Hung, L. W.; Cieslik, M.; Derewenda, U.; Lesley, S. A.; Wilson, I. A.; Giedroc, D. P.; et al. Structure of Thermotoga Maritima TM0439: Implications for the Mechanism of Bacterial GntR Transcription Regulators with Zn2+-Binding FCD Domains. *Acta Crystallogr. D Biol. Crystallogr.* 2009, 65 (Pt 4), 356–365.
- (20) Maddocks, S. E.; Oyston, P. C. F. Structure and Function of the LysR-Type Transcriptional Regulator (LTTR) Family Proteins. *Microbiology* 2008, 154 (Pt 12), 3609–3623.
- (21) Ezezika, O. C.; Haddad, S.; Neidle, E. L.; Momany, C. Oligomerization of BenM, a LysR-Type Transcriptional Regulator: Structural Basis for the Aggregation of Proteins in This Family. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2007**, *63* (5), 361–368.
- (22) Bundy, B. M.; Collier, L. S.; Hoover, T. R.; Neidle, E. L. Synergistic Transcriptional Activation by One Regulatory Protein in Response to Two Metabolites. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (11), 7693–7698.
- (23) Thanbichler, M.; Iniesta, A. A.; Shapiro, L. A Comprehensive Set of Plasmids for Vanillateand Xylose-Inducible Gene Expression in Caulobacter Crescentus. *Nucleic Acids Res.* 2007, *35* (20), e137.
- (24) Morawski, B.; Segura, A.; Ornston, L. N. Repression of Acinetobacter Vanillate Demethylase Synthesis by VanR, a Member of the GntR Family of Transcriptional Regulators. *FEMS Microbiol. Lett.* **2000**, *187* (1), 65–68.
- (25) Gerischer, U. Specific and Global Regulation of Genes Associated with the Degradation of Aromatic Compounds in Bacteria. *J. Mol. Microbiol. Biotechnol.* **2002**, *4* (2), 111–121.
- (26) Collier, L. S.; Gaines, G. L.; Neidle, E. L.; Neidle, E. L. Regulation of Benzoate Degradation in Acinetobacter Sp. Strain ADP1 by BenM, a LysR-Type Transcriptional Activator. J. Bacteriol. 1998, 180 (9), 2493–2501.
- (27) MacLean, A. M.; Anstey, M. I.; Finan, T. M. Binding Site Determinants for the LysR-Type Transcriptional Regulator PcaQ in the Legume Endosymbiont Sinorhizobium Meliloti. *J. Bacteriol.* **2008**, *190* (4), 1237–1246.
- (28) Hansen, E. H.; Møller, B. L.; Kock, G. R.; Bünner, C. M.; Kristensen, C.; Jensen, O. R.; Okkels, F. T.; Olsen, C. E.; Motawia, M. S.; Hansen, J. De Novo Biosynthesis of Vanillin in

Fission Yeast (Schizosaccharomyces Pombe) and Baker's Yeast (Saccharomyces cerevisiae). *Appl. Environ. Microbiol.* **2009**, *75* (9), 2765–2774.

- (29) Liu, Q.; Yu, T.; Li, X.; Chen, Y.; Campbell, K.; Nielsen, J.; Chen, Y. Rewiring Carbon Metabolism in Yeast for High Level Production of Aromatic Chemicals. *Nat. Commun.* 2019, 10 (1), 4976.
- (30) Wang, M.; Li, S.; Zhao, H. Design and Engineering of Intracellular-Metabolite-Sensing/regulation Gene Circuits in Saccharomyces Cerevisiae. *Biotechnol. Bioeng.* 2016, *113* (1), 206–215.
- (31) Rhee, H. S.; Pugh, B. F. Genome-Wide Structure and Organization of Eukaryotic Pre-Initiation Complexes. *Nature* **2012**, *483* (7389), 295–301.
- (32) Strucko, T.; Magdenoska, O.; Mortensen, U. H. Benchmarking Two Commonly Used Saccharomyces Cerevisiae Strains for Heterologous Vanillin-β-Glucoside Production. *Metabolic Engineering Communications* 2015, *2*, 99–108.
- (33) Murphy, K. F.; Balázsi, G.; Collins, J. J. Combinatorial Promoter Design for Engineering Noisy Gene Expression. *Proc. Natl. Acad. Sci. U. S. A.* 2007, 104 (31), 12726–12731. https://doi.org/10.1073/pnas.0608451104.
- (34) Snoek, T.; Chaberski, E. K.; Ambri, F.; Kol, S.; Bjørn, S. P.; Pang, B.; Barajas, J. F.; Welner, D. H.; Jensen, M. K.; Keasling, J. D. Evolution-Guided Engineering of Small-Molecule Biosensors. *Nucleic Acids Res.* 2019.
- (35) Tropel, D.; van der Meer, J. R. Bacterial Transcriptional Regulators for Degradation Pathways of Aromatic Compounds. *Microbiol. Mol. Biol. Rev.* **2004**, *68* (3), 474–500, table of contents.
- (36) Devesse, L.; Smirnova, I.; Lönneborg, R.; Kapp, U.; Brzezinski, P.; Leonard, G. A.; Dian, C. Crystal Structures of DntR Inducer Binding Domains in Complex with Salicylate Offer Insights into the Activation of LysR-Type Transcriptional Regulators. *Mol. Microbiol.* 2011, *81* (2), 354–367.
- (37) Maclean, A. M.; Haerty, W.; Golding, G. B.; Finan, T. M. The LysR-Type PcaQ Protein Regulates Expression of a Protocatechuate-Inducible ABC-Type Transport System in Sinorhizobium Meliloti. *Microbiology* **2011**, *157* (Pt 9), 2522–2533.
- (38) Goethals, K.; Van Montagu, M.; Holsters, M. Conserved Motifs in a Divergent Nod Box of Azorhizobium Caulinodans ORS571 Reveal a Common Structure in Promoters Regulated by LysR-Type Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (5), 1646–1650.
- (39) Garí, E.; Piedrafita, L.; Aldea, M.; Herrero, E. A Set of Vectors with a Tetracycline-Regulatable Promoter System for Modulated Gene Expression in Saccharomyces Cerevisiae. *Yeast* 1997, *13* (9), 837–848.
- (40) Ezezika, O. C.; Collier-Hyams, L. S.; Dale, H. A.; Burk, A. C.; Neidle, E. L. CatM Regulation of the benABCDE Operon: Functional Divergence of Two LysR-Type Paralogs in Acinetobacter Baylyi ADP1. *Appl. Environ. Microbiol.* **2006**, 72 (3), 1749–1758.
- (41) Feng, J.; Li, Q.; Hu, H.-L.; Chen, X.-C.; Hong, G.-F. Inactivation of the Nod Box Distal Half-Site Allows Tetrameric NodD to Activate nodA Transcription in an Inducer-Independent Manner. *Nucleic Acids Res.* **2003**, *31* (12), 3143–3156.
- (42) Peng, H. L.; Shiou, S. R.; Chang, H. Y. Characterization of mdcR, a Regulatory Gene of the Malonate Catabolic System in Klebsiella Pneumoniae. J. Bacteriol. 1999, 181 (7), 2302– 2306.
- (43) Laishram, R. S.; Gowrishankar, J. Environmental Regulation Operating at the Promoter Clearance Step of Bacterial Transcription. *Genes Dev.* **2007**, *21* (10), 1258–1272.

- (44) Weber, C.; Brückner, C.; Weinreb, S.; Lehr, C.; Essl, C.; Boles, E. Biosynthesis of Cis,cis-Muconic Acid and Its Aromatic Precursors, Catechol and Protocatechuic Acid, from Renewable Feedstocks by Saccharomyces Cerevisiae. *Appl. Environ. Microbiol.* 2012, 78 (23), 8421–8430.
- (45) Ambri, F.; Snoek, T.; Skjoedt, M. L.; Jensen, M. K.; Keasling, J. D. Design, Engineering, and Characterization of Prokaryotic Ligand-Binding Transcriptional Activators as Biosensors in Yeast. *Methods Mol. Biol.* **2018**, *1671*, 269–290.
- (46) Jensen, N. B.; Strucko, T.; Kildegaard, K. R.; David, F.; Maury, J.; Mortensen, U. H.; Forster, J.; Nielsen, J.; Borodina, I. EasyClone: Method for Iterative Chromosomal Integration of Multiple Genes Saccharomyces Cerevisiae. *FEMS Yeast Res.* 2014, 14 (2), 238–248.