Title

Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast

Journal

Nature Chemical Biology, 12(11)

ISSN

1552-4450

Authors

Skjoedt, Mette L
Snoek, Tim
Kildegaard, Kanchana R
et al.

Publication Date

2016-11-01

DOI

10.1038/nchembio.2177

Peer reviewed

nature chemical biology

ARTICLE

PUBLISHED ONLINE: 19 SEPTEMBER 2016 | DOI: 10.1038/NCHEMBIO.2177

Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast

Mette L Skjoedt[1,10], Tim Snoek[1,10], Kanchana R Kildegaard[1], Dushica Arsovska[1], Michael Eichenberger[2,3], Tobias J Goedecke[1], Arun S Rajkumar[1], Jie Zhang[1], Mette Kristensen[1], Beata J Lehka[4,5], Solvej Siedler[1], Irina Borodina[1], Michael K Jensen[1]* & Jay D Keasling[1,6–9]

Whole-cell biocatalysts have proven a tractable path toward sustainable production of bulk and fine chemicals. Yet the screening of libraries of cellular designs to identify best-performing biocatalysts is most often a low-throughput endeavor. For this reason, the development of biosensors enabling real-time monitoring of production has attracted attention. Here we applied systematic engineering of multiple parameters to search for a general biosensor design in the budding yeast *Saccharomyces cerevisiae* based on small-molecule binding transcriptional activators from the prokaryote superfamily of LysR-type transcriptional regulators (LTTRs). We identified a design supporting LTTR-dependent activation of reporter gene expression in the presence of cognate small-molecule inducers. As proof of principle, we applied the biosensors for *in vivo* screening of cells producing naringenin or *cis,cis*-muconic acid at different levels, and found that reporter gene output correlated with production. The transplantation of prokaryotic transcriptional activators into the eukaryotic chassis illustrates the potential of a hitherto untapped biosensor resource useful for biotechnological applications.

Bio-based production of chemicals and fuels is an attractive avenue to reduce dependence on petroleum. For bio-based production, biocatalysts must often be genetically modified to increase production. However, the current efficiency of genome-engineering methods and parts prospecting allows for unprecedented genotype diversity that vastly outstrips our ability to screen for best cell performance[1,2].

To meet current demand, bioengineers have started to develop genetically encoded devices and systems that enable screening and selection of better-performing biocatalysts in higher throughput. Genetic devices including oscillators, amplifiers and recorders, which have been developed based on fine-tuned relationships between input and output signals, are promising tools for programming and controlling gene expression in living cells[3–5]. These devices sense extracellular or intracellular perturbations and actuate cellular decision-making processes akin to logic gates in electrical circuits. Hence, from a diverse set of inputs, molecular gating components such as RNA aptamers and allosterically regulated transcription factors have been engineered to control outputs for applications such as high-throughput screening, actuation on cellular metabolism and evolution-based selection of optimal cell performance[6–8].

A key component in many of the reported devices is a ligand-inducible transcriptional regulator. Transcriptional regulators are straightforward and powerful components, with many uses in genetic designs. Owing to their modular structure, transcriptional regulators have proven to be versatile platforms for genetically encoded Boolean logic functions[9,10]. In particular, gene switches based on ligand-binding transcriptional repressors bind to genomic targets in the absence of their cognate ligand and thereby repress gene expression of the downstream gene(s), whereas binding between ligand and repressor causes the release of the repressor from the DNA and thereby a derepression[11]. In such 'NOT' gates, simple steric hindrance of RNA polymerase progression, as in the case of the tetracycline-responsive gene switch TetR, have for decades been used for conditional control of gene expression in both prokaryotic and eukaryotic chassis[12,13]. Transcriptional repressors and other artificial transcriptional regulators can be further engineered, for example, via the addition of nuclear localization signals, destabilization domains and transcriptional activation regions, to repurpose conditional repressors into activators[13–15]. Though conceptually intriguing and practically relevant, the repurposing of logic gates can suffer from the inherent need for extensive engineering[9,16,17].

Though most ligand-inducible genetic devices adopted for eukaryotes historically have been founded on transcriptional repressors, a hitherto untapped resource for use in genetic designs is ligand-inducible transcriptional activators. Bacterial genomes encode a multitude of ligand-inducible activators amenable for integration into synthetic genetic devices[18,19]. In bacteria, transcriptional activation can take place through a transcriptional activator binding to an operator site in a promoter, thereby improving its ability to guide RNA polymerase to initiate transcription, or transcriptional activation can rely on interactions with the RNA polymerase itself such as when a housekeeping σ factor is replaced by another σ factor[20]. Examples of prokaryotic transcriptional activators used for genetic designs in other nonnative prokaryotic chassis include arabinose-inducible AraC and quorum-sensing LuxR[7,21]. However, so far no direct transplantation of prokaryotic ligand-inducible transcriptional activators has been reported in eukaryotes.

Here we report the direct transplantation of a prokaryotic transcriptional activator as a biosensor for *cis,cis*-muconic acid (CCM) in the budding yeast *Saccharomyces cerevisiae*. Based on a multiparametric engineering strategy, we identified a functional design

[1]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark. [2]Evolva SA, Reinach, Switzerland. [3]Department of Biology, Technical University Darmstadt, Darmstadt, Germany. [4]Evolva Biotech A/S, Copenhagen, Denmark. [5]Department of Science and Environment, Roskilde University, Roskilde, Denmark. [6]Joint BioEnergy Institute, Emeryville, California, USA. [7]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA. [8]Department of Chemical and Biomolecular Engineering, University of California, Berkeley, Berkeley, California, USA. [9]Department of Bioengineering, University of California, Berkeley, Berkeley, California, USA. [10]These authors contributed equally to this work.*e-mail: mije@biosustain.dtu.dk

for the biosensor. The design is applicable to a range of other biosensors founded on small-molecule-induced transcriptional activators from the LTTR family. As proof of principle, we applied two of these biosensors for real-time monitoring of bulk and fine chemical product accumulation in yeast cells engineered to produce CCM and naringenin, respectively. This constitutes to our knowledge the first successful direct transfer of prokaryotic transcriptional activators into a eukaryotic chassis to activate gene expression without reconfiguring any motifs and domains.

RESULTS

A prokaryote transcription activator in yeast

To investigate the potential to build orthogonal biosensors using prokaryotic transcriptional activators in a eukaryotic chassis, we initially selected BenM from *Acinetobacter* sp. ADP1 for several reasons. First, it belongs to the LTTR family, which is one of the most abundant families of transcriptional regulators found in a diverse range of prokaryotes[22]. Second, in *Acinetobacter* sp. ADP1, BenM serves as a native CCM-inducible transcriptional activator[23] (**Supplementary Results**, **Supplementary Fig. 1**). CCM is an intermediate from aromatic compound catabolism and an important precursor for bioplastics. Moreover, CCM biosynthesis recently has been refactored in yeast, yet without any high-throughput screening option available[24,25]. Third, BenM has a well-characterized DNA-binding site (here termed *BenO*) and mode of action (**Fig. 1a** and **Supplementary Fig. 1**). Finally, this protein does not require any binding to regulatory subunits apart from its cognate inducers, which should ensure its orthogonality in nonnative chassis[26].

Engineering transcriptional repressors from prokaryotes into eukaryote chassis has emphasized the importance of operator positioning in synthetic eukaryote promoters in relation to transcriptional output[17,27]. Hence, we first sought to identify optimal positioning of *BenO* when introduced into a eukaryote promoter. As a first expression cassette, we first used the full-length (491 base pair (bp)) *CYC1* promoter (here referred to as CYC1p) to control the expression of green fluorescence protein (GFP)[28]. CYC1p recently has been reported as a suitable promoter for introduction of other nonnative TF binding sites in yeast[29], and throughout this study we based all engineered reporter gene promoters on chromosomally integrated full-length or truncated versions of that promoter. Initially, we introduced *BenO* into the 491-bp CYC1p immediately upstream of one of the two TATA boxes: TATA-1β (designated 491bp_CYC1p_ *BenO_T1*) or TATA-2α (designated 491bp_CYC1p_*BenO_T2*)[30]; or upstream of both TATA boxes (designated 491bp_CYC1p_*BenO_T1/T2*) (**Supplementary Fig. 2a**). Outputs from these engineered promoters were compared by flow cytometry to expression from the native CYC1p (491bp_CYC1p) using GFP as the reporter (**Fig. 1a** and **Supplementary Fig. 2a**). In general, introducing *BenO* negatively impacted CYC1p activity (**Fig. 1b**). However, when we simultaneously expressed BenM from the *TEF1* promoter, we observed 20-fold and fivefold induction of expression from 491bp_CYC1p_*BenO_T1* and 491bp_CYC1p_*BenO_T1/T2* compared to the promoter activities without expression of BenM. For 491bp_CYC1p_*BenO_T2*, we observed a modest 30% reduction in expression. BenM did not increase expression of native CYC1p without *BenO* (**Fig. 1b**). Taken together, these data show that BenM can function as a transcriptional activator in yeast.

Protonated CCM is directly taken up by yeast at pH 4.5 without any growth defects (**Supplementary Fig. 3a**,**b**). This enables testing of CCM-inducibility of the genetic devices by simple supplementation of the medium with 200 mg/L CCM at pH 4.5. Following 24 h of cultivation, we measured GFP output using flow cytometry. We observed modest increases (1.3–2.2-fold; **Fig. 1b**) in reporter output from all versions of CYC1p that harbored *BenO*, whereas no change was observed from the native CYC1p (**Fig. 1b**). Also, all engineered promoters showed substantial transcriptional activities in the control medium (no CCM) compared to background autofluorescence (**Fig. 1b**).

To lower the basal activity of the engineered promoters, we removed upstream activating sequences (UAS1 and UAS2) and introduced *BenO* into truncated versions of the CYC1p (designated

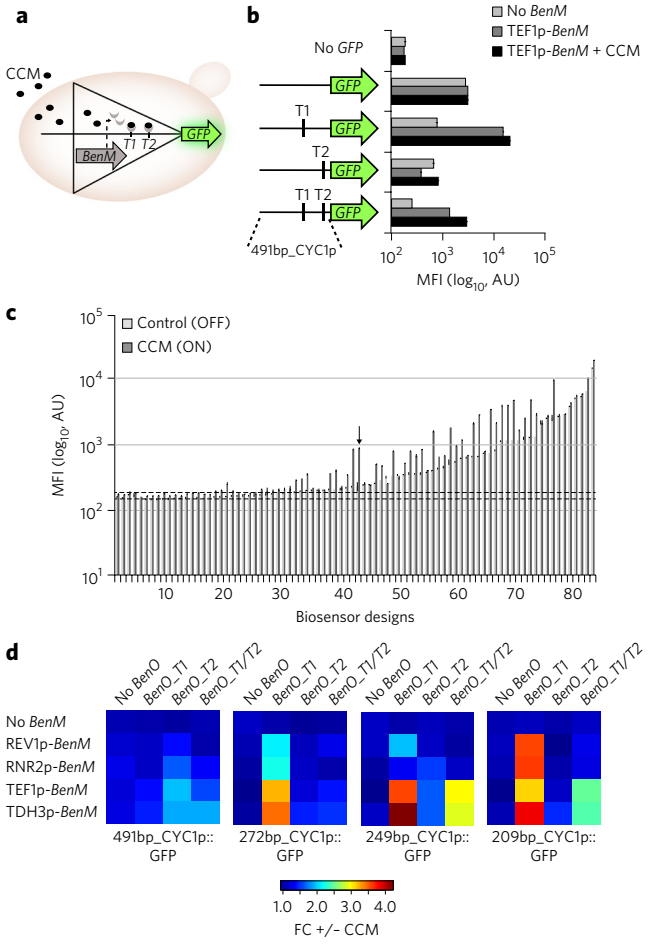

Figure 1 | Engineering the CCM-responsive prokaryotic transcriptional activator BenM in yeast. (**a**) Schematic outline of native and synthetic full-length (491 bp) CYC1 promoter variants with different *BenO* positioning and number (*T1* and/or *T2*). The transcriptional activator BenM from *Acinetobacter* sp. ADP1 controls expression of GFP from the synthetic CYC1 promoter with *BenM* operator (*BenO*) integrated at position *T1* and/or *T2*. CCM further induced BenM-dependent expression of GFP. (**b**) Mean fluorescence intensity (MFI) values from flow cytometry measurements of GFP intensities in the presence or absence of BenM expressed from the constitutively active *TEF1* promoter, and following 24 h of incubation in the presence or absence of 1.4 mM CCM. (**c**) Screening 84 yeast strains expressing all possible combinations of BenM expression levels (TDH3p, TEFp, RNR2p and REV1p) individually or in combination with native or engineered CYC1p reporter promoters of different lengths (491 bp, 272 bp, 249 bp and 209 bp), *BenO* positioning, and number (*T1* and/or *T2*) by flow cytometry. Outputs are ordered according to GFP fluorescence intensity in control medium. Dashed lines indicate background fluorescence as inferred from BenM-expressing strains without GFP, and arrow indicates the best-performing biosensor. Genotypes and GFP expression levels of all 84 strains are listed in **Supplementary Tables 1** and **2**, respectively. (**d**) Heatmaps showing fold change in GFP expression (FC) between CCM-induced and control cultures of the 80 strains shown in **c**. For **b** and **c**, MFI values are shown as mean ± s.d. from three (n = 3) biological replicate experiments. AU, arbitrary units.

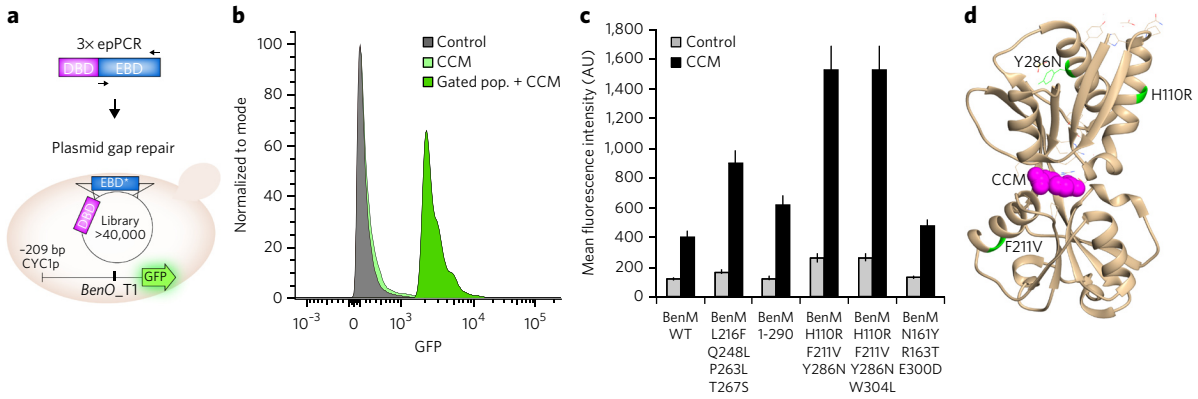

Figure 2 | High-throughput engineering and screening of BenM variants with improved CCM inducibility. (**a**) Purified products from three rounds of error-prone PCR (epPCR), using the sequence encoding the EBD of BenM as template, were co-transformed into yeast together with a linearized centromeric plasmid, to allow for *in vivo* library reconstitution by gap repair and expression of wild-type BenM DBD fused to ~40,000 variants of the EBD. Transformed yeast contained a chromosomal integration of 209bp_CYC1p_*BenO_T1* controlling the expression of GFP to allow for FACS-based screening of BenM variants with improved CCM inducibility. (**b**) Representative flow cytometry histograms of fluorescence intensities obtained from a population of yeast cells expressing CCM sensor variants in control and CCM inducer (CCM) medium. Control, CCM-induced and sorted (Gated pop. + CCM) cell populations were normalized to mode for comparison. Proportions of cells in each histogram were calculated by FlowJo software as described in Online Methods. (**c**) Isolated BenM variants were grown as clonal cultures, validated by flow cytometry and the sequences encoding EBDs of variants with significantly higher GFP expression under CCM-induced cultivation were sequenced (*t*-test, $P < 0.05$). Mean fluorescence intensities are shown as mean ± s.d. from three ($n = 3$) biological replicate experiments. AU, arbitrary units. (**d**) Ribbon representations of the EBD of BenM (PDB 2F7A) with the residue changes identified in BenM H110R,F211V,Y286N highlighted in green. Bound CCM is highlighted as a magenta Van der Waals sphere.

272bp_CYC1p, 249bp_CYC1p and 209bp_CYC1p, **Supplementary Fig. 2a**)[28,31]. Also, to improve the dynamic range of the genetic device, we tuned the production of BenM by placing the *benM* gene under the transcriptional control of three other native yeast promoters: *TDH3* promoter, *RNR2* promoter and *REV1* promoter (here referred to as TDH3p, RNR2p and REV1p, respectively). Together with TEF1p, this system allows for an expression range covering almost three orders of magnitude[32]. By combining and chromosomally integrating all possible BenM expression cassettes with all CYC1p-derived reporter constructs, we generated a total of 84 yeast strains, including control strains (**Fig. 1c**, **Supplementary Table 1** and **Supplementary Fig. 4**). Analyzing basal and CCM-induced GFP expression for all strains by flow cytometry, we observed reporter outputs that spanned more than two orders of magnitude from the lowest to the highest GFP levels, with most of the high outputs resulting from GFP-encoding reporter genes expressed from full-length CYC1p backbones in strains expressing BenM as well (**Fig. 1c**, **Supplementary Fig. 4** and **Supplementary Table 2**). Low-expressing strains mostly comprised truncated CYC1p reporter variants without *BenO* or *BenM*. These data showed that the *BenO_T1* positioning allowed CCM-inducibility of all truncated variants of CYC1p, with the highest dynamic range observed for the minimal promoters 249bp_CYC1p_*BenO_T1* and 209bp_CYC1p_*BenO_T1* (3.2–4.7-fold) (**Fig. 1d** and **Supplementary Table 2**). Among the genetic devices tested, strain MeLS0049 with 209bp_CYC1p_*BenO*_T1 controlled by BenM expressed from REV1p showed both low basal activity and high CCM-inducibility (3.8-fold), and therefore, we regarded it as most suitable for application as a CCM biosensor.

High-throughput prototyping of biosensor variants

The dynamic range of a biosensor output is an important parameter when evaluating applicability of a biosensor for screening and selection. For this reason, we applied a high-throughput engineering strategy to identify BenM variants with higher dynamic ranges when expressed from the weak *REV1* promoter. Previous mutagenesis studies identified residues important for ligand binding in LTTR effector binding domains (EBDs)[22]. To screen and select BenM variants with improved dynamic range, we first performed PCR-based mutagenesis of the sequence encoding BenM EBD (residues 90–304) (**Fig. 2a**). Following mutagenesis, we harnessed yeast's homologous recombination machinery for plasmid gap repair of sequences encoding variant EBDs with the BenM DNA-binding domain (DBD) (**Fig. 2a**). We analyzed a population derived from ~40,000 transformants by fluorescence-activated cell sorting (FACS) using a two-step approach, in which we first removed the variants showing increased basal activity. Next, we compared fluorescence output from the population of transformants in control and CCM medium, and sorted all cells showing higher fluorescence in CCM medium than in control medium (**Fig. 2b**). We cultivated the sorted cells as clones, and validated them by flow cytometry (**Fig. 2c**). We identified five BenM variants with higher dynamic ranges than wild-type BenM (**Fig. 2c**). Sequencing of the clones encoding BenM variants identified a triple mutant with point mutations encoding substitutions H110R, F211V and Y286N in the BenM EBD (**Fig. 2c**). Plasmid-based expression of BenM H110R,F211V,Y286N showed doubled GFP output upon CCM induction (sixfold), compared to induction for the plasmid-based expression of wild-type BenM (**Fig. 2c**). Substitutions in BenM H110R,F211V,Y286N were not positioned in the immediate vicinity of the CCM binding site (**Fig. 2d**). Similar to all other genetic devices engineered in this study, we integrated the sequence encoding BenM H110R,F211V,Y286N into the genome for stable expression.

LTTR-based biosensor specificity and orthogonality

To assess the potential application of the LTTR-based biosensor for CCM in yeast, we next investigated the specificity of BenM, as well as its potential impact on the host transcriptome. First, by testing a range of diacids supplied to the growth medium at pH 4.5 with identical molar concentrations to CCM (1.4 mM), we observed that among the diacids tested both BenM and BenM H110R,F211V,Y286N induced GFP expression specifically in response to CCM (**Fig. 3a**). Second, to test for transcriptional orthogonality of BenM H110R,F211V,Y286N in yeast, we used RNA sequencing (RNA-seq) to quantify and compare the transcriptomes of cells with (MeLS0284)

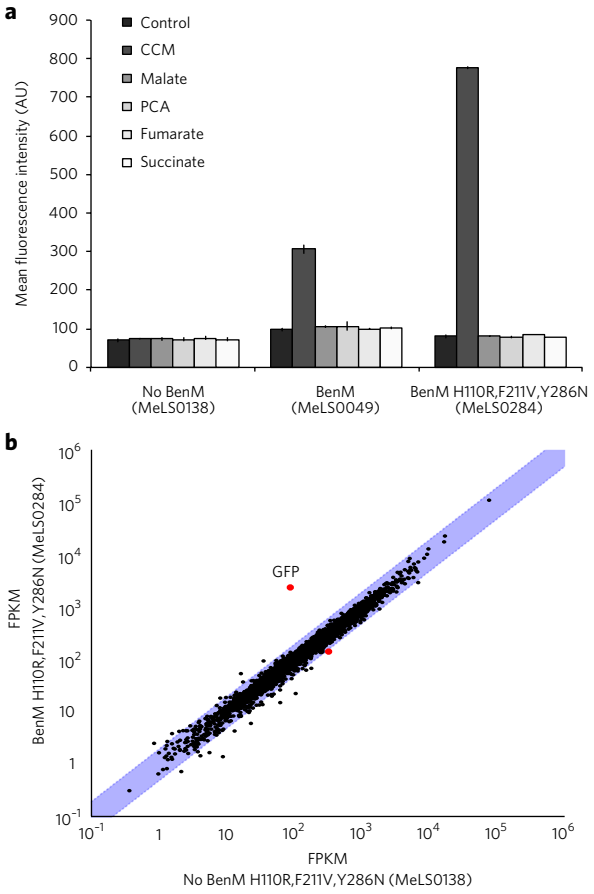

Figure 3 | Biosensor specificity and transcriptional orthogonality. (**a**) Specificity of the CCM biosensor was tested by addition of various dicarboxylic acids (1.4 mM) to the growth medium. GFP expression was measured by flow cytometry following 24 h of cultivation. (**b**) RNA sequencing FPKM (fragments per kilobase per million) are plotted for yeast cells stably expressing 209bp_CYC1p_*BenO_T1*::GFP reporter construct and BenM H110R,F211V,Y286N versus cells only expressing the reporter construct following 24 h of cultivation in medium supplemented with CCM. Purple area indicates twofold cut-off and red dots significantly differentially regulated genes as inferred from cuffdiff[48] (>2-fold, $P < 0.05$) (see **Supplementary Dataset 1** and **Supplementary Fig. 5**). All data points are averaged from three ($n = 3$) biological replicates.

or without (MeLS0138) expression of BenM H110R,F211V,Y286N. As the genetic device had low basal activity (**Supplementary Fig. 4** and **Supplementary Table 2**), we analyzed yeast transcriptomes following 24 h of cultivation in the presence of CCM. Here we observed that the average GFP transcript abundance from strain MeLS0284 was ~27-fold higher compared to strain MeLS0138 (**Fig. 3b**, **Supplementary Fig. 5** and **Supplementary Dataset 1**). Apart from genes encoding GFP and BenM, only one other gene, *TCS3*, encoding the Golgi-associated retrograde protein complex component, passed our stringent cutoff (*t*-test, $P < 0.05$, FDR < 5%, greater than twofold changes) showing a modest decrease (2.3×) in expression level when BenM H110R,F211V,Y286N was expressed (**Fig. 3b**). We found no match to *BenO* in this gene's promoter suggesting that the minor transcriptome perturbations could be due to noise in RNA-seq measurements or indirect effects.

A design for engineering LTTR-based biosensors in yeast

The genetic device developed in this study represents to our knowledge the first example of transplanting a prokaryotic transcriptional activator into a eukaryotic chassis and its use to activate gene expression without the need to modify the protein beyond codon optimization[14,16]. Acknowledging the vast numbers of transcriptional activators found among LTTR members[22], the optimal reporter promoter design (209bp_CYC1p_*BenO_T1*) could prove valid for other metabolic engineering and biotechnological applications. To test the generality of the biosensor design for engineering other small-molecule-binding transcriptional activators as biosensors in yeast, we selected four other candidates from the LTTR family; FdeR from *Herbaspirillum seropedicae,* PcaQ from *Sinorhizobium meliloti*, ArgP from *Escherichia coli* and MdcR from *Klebsiella pneumonia*, with co-inducers naringenin, protocatechuic acid (PCA), L-arginine and malonic acid, respectively[21,33–36]. In this proof-of-principle study, we selected the four candidates based on a minimal set of information, including knowledge about operator sequences, experimental evidence for ligand-inducible control of target operons, and their mode of action in the native chassis (i.e., activation; **Supplementary Fig. 1**). Furthermore, all of the metabolites mentioned above can passively diffuse across the yeast plasma membrane, with the exception of malonic acid, which requires the expression of the dicarboxylic acid transporter MAE1 from *Schizosaccharomyces pombe*[37]. For this purpose, we integrated the gene encoding MAE1 into cells expressing MdcR (**Supplementary Table 3**). Based on knowledge about the operator sequences, the ligand-inducible control of target promoters and the mode of action (transcriptional activation), we directly replaced *BenO* located upstream of the TATA box 1β (*T1*) in the 209-bp_CYC1p with operator sequences for each of these LTTRs (**Fig. 4a**, **Supplementary Fig. 2a**,**b**, and **Supplementary Table 4**). We first tested whether expression of GFP could be activated upon low and high expression of individual LTTRs. All LTTRs activated GFP expression from the 209bp_CYC1p_*T1* when we expressed the LTTR from the strong *TDH3* promoter compared to yeast cells without expression of an LTTR (1.4×–8.1×), with BenM showing the strongest activation (8.1×) (**Fig. 4a**). Similarly, GFP expression was induced by ArgP when the weak REV1p controlled expression of the LTTR (2.2×). This proves the broad applicability of the reporter promoter design and that biosensor output is tunable depending on the expression level of the LTTR. Next, we tested whether each LTTR could further induce GFP expression when its cognate inducer is supplied to the growth medium (**Fig. 4b**). For this purpose we prepared media supplemented with one of the following ligands: 1.4 mM CCM, 0.2 mM naringenin, 30 mM L-arginine, 1.4 mM PCA or 10 mM malonic acid (these concentrations have been previously reported to be relevant concentrations in terms of bio-based production and microbial physiology[21,25,38–40]). Here, in addition to BenM, ArgP was the only LTTR enabling a significant ligand-inducible increase in GFP expression when LTTR expression was controlled by REV1p (*t*-test, $P < 0.05$; **Fig. 4b**). However, when expressing LTTRs from the *TDH3* promoter, all LTTRs except PcaQ significantly increased GFP expression (1.4×–4.1×) when their cognate ligand was present in the cultivation medium (*t*-test, $P < 0.05$; **Fig. 4b**). All tested LTTRs activated expression of GFP when their operators were placed in the *T1* position of the 209bp_CYC1p scaffold (**Supplementary Table 4**). Furthermore, just as for BenM, yeast expressing FdeR, ArgP and MdcR from the strong *TDH3* promoter further induced GFP expression upon addition of their cognate inducers (**Fig. 4b**).

Many of the characterized LTTRs regulate operons by binding prototypic LTTR box patterns 5′-T-N11-A-3′ and 5′-TTA-N7/8-GAA-3′[22]. In addition to transcriptional orthogonality (**Fig. 3b**), we therefore further tested whether individual LTTRs would cross-react with operators for another LTTR. For this purpose, we expressed LTTRs ArgP and MdcR together with the 209bp_CYC1p_*T1* with operators for MdcR (here called *MdcO*) or ArgP (here called *ArgO*) driving GFP expression. As controls we tested GFP expression from 209bp_CYC1p_*T1* with *MdcO* or *ArgO* without expression of

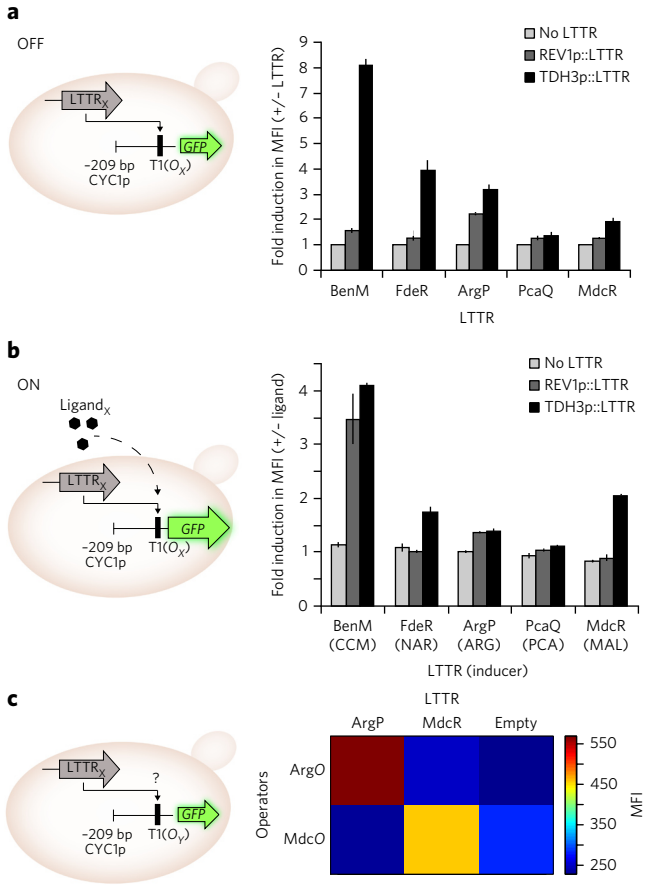

Figure 4 | Application of transcriptional activators from the LTTR family as biosensors in yeast. (**a**) Illustration of LTTR-mediated activation of GFP expression by binding to the cognate operator in position *T1* of 209bp_CYC1p (left). The 209bp_CYC1p_*T1* reporter promoter design supports GFP expression when controlled by individual LTTR transcriptional activators expressed from either a weak (REV1p) or a strong (TDH3p) promoter (right). The *y* axis shows fold induction in mean fluorescence intensity (MFI) in cells expressing individual LTTRs relative to cells not expressing the LTTRs. (**b**) Illustration on external application of individual ligands for induction of LTTR-mediated activation of GFP expression (left). External application of individual ligands can induce LTTR-mediated activation of GFP expression (right). The *y* axis shows fold induction in MFI for cells grown for 24 h in medium containing either CCM, naringenin (NAR), L-arginine (ARG), protocatechuic acid (PCA) or malonic acid (MAL) compared to cells growing in control medium. (**c**) Heatmap showing orthogonality of MdcR- and ArgP-mediated transcriptional regulation of GFP expression controlled by either reporter promoter 209bp_CYC1p_MdcO_*T1* or 209bp_CYC1p_ArgO_*T1* (**Supplementary Table 4**). Color key shows MFI from three (n = 3) biological replicate experiments. For **a** and **b**, MFI values and their error bars are shown as mean ± s.d. from three (n = 3) biological replicates.

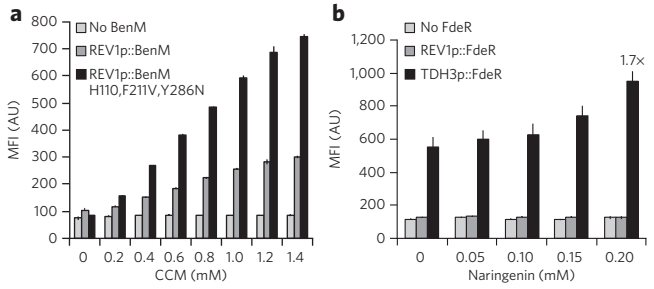

Figure 5 | Biosensor sensitivity and operational range. (**a**) Response functions of wild-type BenM and engineered BenM H110R,F211V,Y286N expressed in yeast from REV1p as measured by flow cytometry using various concentrations of CCM (24 h) and the 209bp_CYC1p_*BenO_T1* promoter controlling the expression of GFP. A yeast strain without *BenM* expressed was used as a control for background GFP fluorescence from the 209bp_CYC1p_*BenO_T1* promoter. (**b**) Response function measurement for the naringenin biosensor when FdeR is expressed from a weak (REV1p) or a strong (TDH3p) promoter using various concentrations of naringenin (24 h) and the 209bp_CYC1p_*FdeO_T1* reporter promoter controlling the expression of GFP. A yeast strain without *FdeR* expressed was used as a control for background GFP fluorescence from the 209bp_CYC1p_*FdeO_T1* promoter. For **a** and **b**, mean fluorescence intensity (MFI) values and their error bars are shown as mean ± s.d. from three (n = 3) biological replicate experiments.

LTTRs. Flow cytometry analysis showed specificity between LTTR transcriptional activators and their inferred operator (**Fig. 4c**). This is in agreement with another study on cross-reactivity between promoter and transcriptional regulators of the TetR family, and the fact that LTTR residues in both the conserved N-terminal DBDs and the divergent EBDs are important for DNA binding[12,22].

In vivo application of LTTR-based biosensors in yeast

Based on our engineering efforts and characterization of prokaryote LTTR-based biosensors imported into yeast, we next addressed whether such biosensors would support real-time monitoring of product accumulation *in vivo* and thereby potentially provide high-throughput screening assays of biocatalysts. To test this, we selected CCM and naringenin, for which highest titers in shake-flask-cultivated haploid yeast of ~1 mM (141 mg/L) and 0.2 mM (54 mg/L), respectively, have recently been reported[25,41]. Also, these two products are of general interest to biotechnology, because CCM is a platform chemical for the production of several valuable consumer bioplastics[24], and naringenin belongs to a class of secondary metabolites called flavonoids with nutritional and agricultural value[42].

Before applying the biosensors for *in vivo* detection of these metabolites, we first tested their operational range and induction kinetics. For BenM and BenM H110R,F211V,Y286N, we observed a weakly sigmoidal input-output relationship between CCM concentration and GFP output following 24 h of cultivation. For chromosomally integrated constructs encoding BenM H110R,F211V,Y286N and BenM, a maximum of 10-fold and 3.5-fold induction was reached in the presence of the highest soluble CCM concentrations (1.4 mM, 200 mg/L), respectively (**Fig. 5a**). Induction kinetics of BenM and BenM H110R,F211V,Y286N were similar. This is in line with BenM substitutions likely not being involved in direct binding of CCM (**Fig. 2d**) but rather altering BenM binding to DNA to support increased GFP expression.

Similarly, for FdeR we first tested naringenin sensitivity and operational range of the sensor. As for CCM, we only tested the operational range for concentrations of naringenin at which it is soluble in growth medium (i.e., < 0.2 mM). Expression of FdeR controlled by the weak *REV1* promoter did not support induction of GFP expression at any of the tested concentrations (**Fig. 5b**), yet when expression of FdeR was controlled by the strong *TDH3* promoter, we observed a maximum 1.7-fold increase in GFP expression following 24 h cultivation in the presence of 0.2 mM naringenin (**Fig. 5b**). The operational ranges of BenM and FdeR were within the ranges of reported CCM and naringenin production titers in yeast, and which could make these biosensors applicable for screening such biocatalysts.

Next, we transformed the CCM biosensor (209bp_CYC1p_*BenO_T1*::GFP and REV1p::BenM^{H110R,F211V,Y286N}) into a small library of six yeast strains engineered to produce CCM. CCM

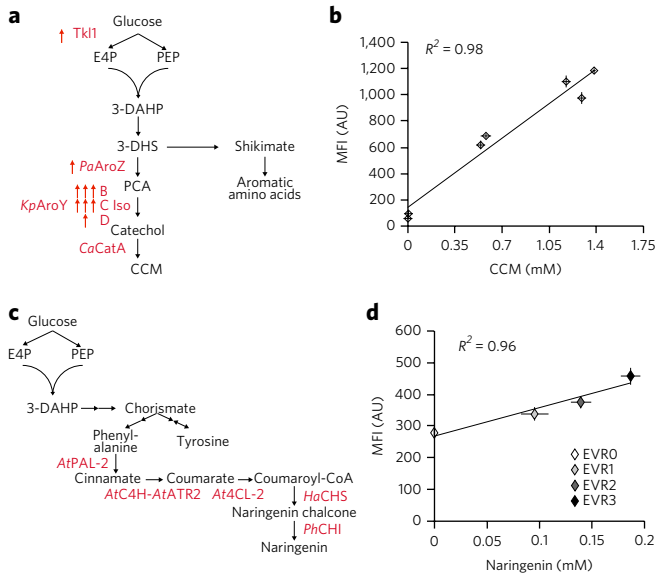

Figure 6 | *In vivo* application of CCM and naringenin biosensors in yeast. (**a**) Schematic representation of the heterologous three-step CCM production pathway[25] for testing BenM as a biosensor for *in vivo* CCM production in yeast. Overexpression of Tkl1 was included together with balancing of the heterologous three-step pathway (PaAroZ, KpAroY and CaCatA) using single-locus or multi-loci integration of sequences encoding AroY subunits B and C (Iso, isoform) (Online Methods and **Supplementary Table 1**). (**b**) Following 24 h of cultivation, CCM titers and mean fluorescence intensities (MFIs) were quantified and plotted for each strain. (**c**) Schematic representation of a heterologous five-step naringenin production pathway adopted from ref. 43. For the hydroxylation of cinnamate to coumarate, a fusion protein of AtC4H and AtATR2 was used. For testing FdeR as a biosensor for *in vivo* naringenin production in yeast, mean fluorescence intensity (MFI) in three different strains engineered with one copy of the five-step naringenin production pathway (EVR1) or with one (EVR2) or two (EVR3) additional integrations of bottleneck enzymes were compared to a control strain (EVR0, ctrl) without the production pathway. Following 48 h of cultivation, naringenin titers and MFI values were quantified and plotted. For **b** and **d**, data are average of three biological replicates. MFI values and metabolite quantifications are presented as means ± s.d. from three (n = 3) biological replicate experiments.

production with a final titer of 149 mg/L has been recently reported in haploid yeast using a three-step heterologous pathway consisting of an *AroZ* homolog from *Podospora anserina* encoding dehydroshikimate dehydratase (PaAroZ), the *AroY* gene from *Klebsiella pneumonia* encoding the multisubunit protocatechuic acid decarboxylase (PCA-DC) and the *CatA* gene encoding catechol 1,2-dioxygenase from *Candida albicans* (CaCatA) (**Fig. 6a**)[25]. From that study it was clear that PCA-DC was a rate-limiting step for flux through the upper part of the shikimate pathway toward CCM. It had been suggested that an increased supply of precursor toward erythrose-4-phosphate (E4P) could improve CCM production. For this reason we introduced single or multiple copies of genes encoding different PCA-DC subunits from *K. pneumonia* and introduced no or one additional copy of transketolase (*Tkl1*) from *S. cerevisiae* (**Fig. 6a**). We cultured the six-membered CCM production strain library and a wild-type CCM null background strain individually. After 24 h of cultivation, we analyzed the medium for CCM concentration using high-pressure liquid chromatography (HPLC) and analyzed the cells by flow cytometry for GFP intensity measurements. We observed a strong correlation (r = 0.98) between GFP output and CCM production titers, spanning a range of 0.00016 mM to 1.39 mM (0.023–197.6 mg/L) (**Fig. 6b**). We obtained the highest titers in strain ST4245-2 with multiple TY integrations of genes encoding AroY subunits B and C, and Tkl1 (**Fig. 6a**,**b**). To further examine the performance of the CCM biosensor, we monitored GFP output and CCM production titers following 72 h of cultivation. GFP outputs were saturated at titers >1.41 mM (200 mg/L) (**Supplementary Fig. 6a**,**b**). However, the strain that produced the most CCM after 72 h (3.03 mM, 430.8 mg/L) also produced the most CCM and had the highest fluorescence after 24 h, emphasizing the applicability of the CCM biosensor for screening high-producing strains during early stages of cultivation.

Finally, we transformed 209bp_CYC1p_*FdeO_T1*::GFP and TDH3p::FdeR into yeast strains with a five-step heterologous naringenin pathway[43]. For building a small library of naringenin-producing strains, we chromosomally introduced either a single copy of the pathway (*EVR1*), or with one and two additional integrations of genes encoding bottleneck enzymes (AtPAL-2 and HaCHS for EVR2; AtPAL-2, HaCHS, and AtC4H:L5:AtATR2 for EVR3) (**Fig. 6c** and **Supplementary Table 1**). Following 48 h of cultivation, we analyzed the medium for naringenin concentration using ultra-performance liquid chromatography (UPLC) and analyzed the cells by flow cytometry for GFP intensity measurements. As observed for the CCM biosensor, the naringenin biosensor also had a strong correlation (r = 0.96) between GFP output and naringenin titers, spanning a range of 0.094 mM to 0.184 mM (25.61–50.18 mg/L) (**Fig. 6d**), with the highest titer obtained in strain EVR3 containing two additional integrations of genes encoding bottleneck enzymes on top of the full copy of the five-step naringenin pathway. For the naringenin sensor, we observed a poorer correlation between biosensor output and titers at 24 h (r = 0.87) compared to our 48 h (r = 0.96) measurements (**Supplementary Fig. 6c**,**d**). However, just as for the CCM biosensor, the strain that produced the most naringenin at 48 h (0.184 mM, 50.18 mg/L) also produced the most naringenin (0.045 mM, 12.25 mg/L) and had the highest fluorescence at 24 h.

Taken together, the two applications of the LTTR-based biosensors suggest that simple expression of the LTTR and an engineered reporter promoter (209bp_CYC1p_*T1*::GFP) with an operator site in position T1 allows for direct transplantation of prokaryotic transcriptional activators as biosensors to screen for the best-performing biocatalysts. Though some of the transcriptional activators used in this study derived from prokaryotes with growth optima at higher temperatures compared to yeast, BenM showed a higher dynamic range in output at 30 °C compared to 37 °C (**Supplementary Fig. 7**), illustrating robustness of LTTR performance.

DISCUSSION

Systematic engineering and meticulous characterization have for decades pushed forward the sequence-function understanding of genetic parts and interactions thereof. This has allowed the rational engineering of parts and genetic circuits useful for a range of applications in biotechnology. Although most of the genetic devices originate from prokaryotes, transplantation into eukaryotes has been reported for bioswitches, used to construct orthogonal genetic devices to control a cellular response to a defined input[14,44,45]. Specifically, genetic devices enabling the manipulation of transcription through the transplantation of prokaryote transcriptional repressors have inspired researchers, in their quest for tools to screen, select and actuate cellular responses[17,46]. In this study we showed that ligand-inducible transcriptional activators from the largest family of transcriptional regulators found in prokaryotes can be ported to the eukaryotic chassis and used to measure the level of a small molecule inside the cell and activate transcription. The LTTR-based transcriptional activators function as is in yeast without any further engineering and without co-expression of other molecular components (i.e., σ factors). In fact, through a systematic engineering approach we provide a framework from which new ligand-binding

transcriptional activators from the LTTR family can be designed through the simple swapping of a candidate LTTR operator sequence into the 209bp_CYC1p scaffold promoter at a defined position (*T1*) (**Supplementary Fig. 2a,b** and **Supplementary Tables 4** and **5**). Also, we provided two proof-of-principle examples for such biosensors to screen *in vivo* for the best-performing biocatalysts.

Compared to many of the studies using transcriptional repressors as biosensors in eukaryotes[12,13], the biosensor outputs based on ligand-inducible transcriptional activators presented in this study have lower dynamic ranges falling within one order of magnitude. This is in agreement with the observation from using BenM and FdeR as biosensors in *E. coli*[21,39]. This can pose a challenge for their applicability in genetic designs where a larger dynamic range is needed. However, we demonstrated in this study how these biosensors could be subjected to biosensor-based FACS for identification of biosensor designs with improved characteristics (i.e., dynamic range), which may expand their applicability for metabolic engineering. We envision that this can be exploited for high-throughput screening of libraries of genetic designs for metabolites for which there exists no high-throughput screening assay or biosensor.

Apart from dynamic range, another key performance measure for biosensors is their operational range. In our study, we demonstrated how biosensors could be used in laboratory strains with limited engineering to improve titers, which at their best still were far from being commercially relevant. Indeed, in diploid yeast, production of 559.3 mg/L CCM has been recently reported[24], whereas an *E. coli-E. coli* co-cultivation study has reported the production of 2 g/L CCM[47]. Though tolerance to low-pH fermentations should make yeast an economically feasible chassis for bio-based production of dicarboxylic acids such as CCM, the CCM biosensor design based on BenM may need to be adjusted or evolved as production hosts become better and the product titers become higher. Additionally, the biosensor will need to be matched to the production kinetics of the individual strain or library of biocatalysts.

Nevertheless, the LTTR-based ligand-inducible transcriptional activators reported here are much-needed tools for optimizing biocatalysts that produce chemicals and fuels for which there exist no high-throughput screen or selection. This should spur interest in developing many other orthogonal logic gates based on LTTR members, which could serve as a vast and valuable reservoir for developing new ligand-inducible genetic circuits enabling high-throughput screening, reprogramming and growth-coupled strain selection for bio-based production of chemicals. Furthermore, as the mode of action of transcriptional activators (YES) differs from that of repressors (NOT), a future possibility for higher-order designs in cellular reprogramming can now be exploited in greater diversity.

Received 19 August 2015; accepted 30 June 2016;
published online 19 September 2016

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. RNA-seq data are available in the ArrayExpress: E-MTAB-4836.

References

1. Jakočiūnas, T., Jensen, M.K. & Keasling, J.D. CRISPR/Cas9 advances engineering of microbial cell factories. *Metab. Eng.* **34**, 44–59 (2016).
2. Esvelt, K.M. & Wang, H.H. Genome-scale engineering for systems and synthetic biology. *Mol. Syst. Biol.* **9**, 641 (2013).
3. Elowitz, M.B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).
4. Wang, B., Barahona, M. & Buck, M. Amplification of small molecule-inducible gene expression via tuning of intracellular receptor densities. *Nucleic Acids Res.* **43**, 1955–1964 (2015).
5. Farzadfard, F. & Lu, T.K. Genomically encoded analog memory with precise in vivo DNA writing in living cell populations. *Science* **346**, 1256272 (2014).
6. Michener, J.K. & Smolke, C.D. High-throughput enzyme evolution in *Saccharomyces cerevisiae* using a synthetic RNA switch. *Metab. Eng.* **14**, 306–316 (2012).
7. Raman, S., Rogers, J.K., Taylor, N.D. & Church, G.M. Evolution-guided optimization of biosynthetic pathways. *Proc. Natl. Acad. Sci. USA* **111**, 17803–17808 (2014).
8. Choi, J.H. & Ostermeier, M. Rational design of a fusion protein to exhibit disulfide-mediated logic gate behavior. *ACS Synth. Biol.* **4**, 400–406 (2015).
9. Ausländer, S., Ausländer, D., Müller, M., Wieland, M. & Fussenegger, M. Programmable single-cell mammalian biocomputers. *Nature* **487**, 123–127 (2012).
10. Khalil, A.S. *et al.* A synthetic biology framework for programming eukaryotic transcription functions. *Cell* **150**, 647–658 (2012).
11. Folcher, M., Xie, M., Spinnler, A. & Fussenegger, M. Synthetic mammalian trigger-controlled bipartite transcription factors. *Nucleic Acids Res.* **41**, e134 (2013).
12. Stanton, B.C. *et al.* Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat. Chem. Biol.* **10**, 99–105 (2014).
13. Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. USA* **89**, 5547–5551 (1992).
14. Stanton, B.C. *et al.* Systematic transfer of prokaryotic sensors and circuits to mammalian cells. *ACS Synth. Biol.* **3**, 880–891 (2014).
15. Gilbert, L.A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013).
16. Gossen, M. *et al.* Transcriptional activation by tetracyclines in mammalian cells. *Science* **268**, 1766–1769 (1995).
17. Teo, W.S. & Chang, M.W. Bacterial XylRs and synthetic promoters function as genetically encoded xylose biosensors in *Saccharomyces cerevisiae*. *Biotechnol. J.* **10**, 315–322 (2015).
18. Lee, N., Francklyn, C. & Hamilton, E.P. Arabinose-induced binding of AraC protein to araI2 activates the araBAD operon promoter. *Proc. Natl. Acad. Sci. USA* **84**, 8814–8818 (1987).
19. Shadel, G.S. & Baldwin, T.O. The Vibrio fischeri LuxR protein is capable of bidirectional stimulation of transcription and both positive and negative regulation of the luxR gene. *J. Bacteriol.* **173**, 568–574 (1991).
20. Lee, D.J., Minchin, S.D. & Busby, S.J.W. Activating transcription in bacteria. *Annu. Rev. Microbiol.* **66**, 125–152 (2012).
21. Siedler, S., Stahlhut, S.G., Malla, S., Maury, J. & Neves, A.R. Novel biosensors based on flavonoid-responsive transcriptional regulators introduced into *Escherichia coli*. *Metab. Eng.* **21**, 2–8 (2014).
22. Maddocks, S.E. & Oyston, P.C.F. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* **154**, 3609–3623 (2008).
23. Collier, L.S., Gaines, G.L. III & Neidle, E.L. Regulation of benzoate degradation in Acinetobacter sp. strain ADP1 by BenM, a LysR-type transcriptional activator. *J. Bacteriol.* **180**, 2493–2501 (1998).
24. Suastegui, M. *et al.* Combining Metabolic Engineering and Electrocatalysis: Application to the Production of Polyamides from Sugar. *Angew. Chem.* **128**, 2414–2419 (2016).
25. Curran, K.A., Leavitt, J.M., Karim, A.S. & Alper, H.S. Metabolic engineering of muconic acid production in *Saccharomyces cerevisiae*. *Metab. Eng.* **15**, 55–66 (2013).
26. 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. USA* **99**, 7693–7698 (2002).
27. Wang, M., Li, S. & Zhao, H. Design and engineering of intracellular-metabolite-sensing/regulation gene circuits in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **113**, 206–215 (2016).
28. Olesen, J., Hahn, S. & Guarente, L. Yeast HAP2 and HAP3 activators both bind to the CYC1 upstream activation site, UAS2, in an interdependent manner. *Cell* **51**, 953–961 (1987).
29. McIsaac, R.S., Gibney, P.A., Chandran, S.S., Benjamin, K.R. & Botstein, D. Synthetic biology tools for programming gene expression without nutritional perturbations in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **42**, e48 (2014).
30. Li, W.Z. & Sherman, F. Two types of TATA elements for the CYC1 gene of the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **11**, 666–676 (1991).
31. Pfeifer, K., Arcangioli, B. & Guarente, L. Yeast HAP1 activator competes with the factor RC2 for binding to the upstream activation site UAS1 of the CYC1 gene. *Cell* **49**, 9–18 (1987).
32. Lee, M.E., Aswani, A., Han, A.S., Tomlin, C.J. & Dueber, J.E. Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. *Nucleic Acids Res.* **41**, 10668–10678 (2013).
33. 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.* **181**, 2302–2306 (1999).

34. MacLean, A.M., MacPherson, G., Aneja, P. & Finan, T.M. Characterization of the beta-ketoadipate pathway in Sinorhizobium meliloti. *Appl. Environ. Microbiol.* **72**, 5403–5413 (2006).
35. Laishram, R.S. & Gowrishankar, J. Environmental regulation operating at the promoter clearance step of bacterial transcription. *Genes Dev.* **21**, 1258–1272 (2007).
36. 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* **157**, 2522–2533 (2011).
37. Chen, W.N. & Tan, K.Y. "Malonate uptake and metabolism in *Saccharomyces cerevisiae*". *Appl. Biochem. Biotechnol.* **171**, 44–62 (2013).
38. Opekarová, M. & Kubín, J. On the unidirectionality of arginine uptake in the yeast *Saccharomyces cerevisiae*. *FEMS Microbiol. Lett.* **152**, 261–267 (1997).
39. Rogers, J.K. & Church, G.M. Genetically encoded sensors enable real-time observation of metabolite production. *Proc. Natl. Acad. Sci. USA* **113**, 2388–2393 (2016).
40. Rikhvanov, E.G., Varakina, N.N., Rusaleva, T.M., Rachenko, E.I. & Voĭnikov, V.K. [The effect of sodium malonate on yeast thermotolerance]. *Mikrobiologiia* **72**, 616–620 (2003).
41. Koopman, F. *et al.* De novo production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **11**, 155 (2012).
42. Winkel-Shirley, B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol.* **126**, 485–493 (2001).
43. Naesby, M. *et al.* Yeast artificial chromosomes employed for random assembly of biosynthetic pathways and production of diverse compounds in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **8**, 45 (2009).
44. Gupta, R.K., Patterson, S.S., Ripp, S., Simpson, M.L. & Sayler, G.S. Expression of the Photorhabdus luminescens lux genes (luxA, B, C, D, and E) in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **4**, 305–313 (2003).
45. Galloway, K.E., Franco, E. & Smolke, C.D. Dynamically reshaping signaling networks to program cell fate via genetic controllers. *Science* **341**, 1235005 (2013).
46. Kim, T., Folcher, M., Doaud-El Baba, M. & Fussenegger, M. A synthetic erectile optogenetic stimulator enabling blue-light-inducible penile erection. *Angew. Chem. Int. Edn Engl.* **54**, 5933–5938 (2015).
47. Zhang, H., Li, Z., Pereira, B. & Stephanopoulos, G. Engineering *E. coli-E. coli* cocultures for production of muconic acid from glycerol. *Microb. Cell Fact.* **14**, 134 (2015).
48. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).

Acknowledgments

This work was supported by the Novo Nordisk Foundation and by the European Union Seventh Framework Programme (FP7-KBBE-2013-7-single-stage) under grant agreement no. 613745, Promys (M.E. & S.S.). We acknowledge A. Koza and E. Özdemir for technical assistance.

Author contributions

M.L.S., T.S., J.D.K. and M.K.J. conceived this project. M.L.S., T.S. and M.K.J. designed all of the experiments. M.L.S., T.S. and D.A. performed all flow cytometry analyses. M.L.S., T.S., D.A., B.J.L., J.Z., K.R.K., S.S., T.J.G. and M.E. constructed all strains and plasmids. M.K. and K.R.K. performed all analytical measurements, and M.K.J. performed the RNA-seq experiment. M.L.S., T.S., M.K.J., I.B., A.S.R. and K.R.K. analyzed the data. M.K.J. wrote the paper.

Competing financial interests

The authors declare no competing financial interests.

Additional information

Any supplementary information, chemical compound information and source data are available in the online version of the paper. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Correspondence and requests for materials should be addressed to M.K.J.

ONLINE METHODS

Strains, chemicals and media. *Saccharomyces cerevisiae* CEN.PK102-5B (MAT*a ura3-52 his3Δ1 leu2-3/112 MAL2-8*^{c}*SUC2*), CEN.PK113-5A (MAT*a, trp1 his3Δ1 leu2-3/112 MAL2-8*^{c}*SUC2*) and CEN.PK113-7D (wild type, *MAT*a *MAL2-8*^{c}*SUC2)* strains were obtained from P. Kötter (Johann Wolfgang Goethe University, Frankfurt, Germany). EasyClone plasmids used in this work are described ref. 49. *E.* strain DH5α was used as a host for cloning and plasmid propagation. The chemicals and Pfu TURBO DNA polymerase were obtained from Sigma-Aldrich and Agilent Technologies Inc., respectively. All acids used were >97% in purity. *S. cerevisiae* cells were grown at 30 °C in synthetic complete medium as well as drop-out medium, and agar plates were prepared using pre-mixed drop-out powders (Sigma-Aldrich). Mineral medium was freshly prepared as described previously[49]. For all media containing diacids, 1.4 mM of the individual diacids were dissolved in mineral medium and pH adjusted to 4.5 before sterile filtration. For CCM several dilutions were made to examine the performance of the CCM biosensor. For naringenin, mineral medium was supplemented with 0, 0.05, 0.10 or 0.20 mM naringenin, dissolved in ethanol; the final ethanol concentration for each medium was 2% (v/v), and the final pH of the medium was adjusted to 6.0. *E. coli* cells were grown at 37 °C in Luria-Bertani (LB) medium supplemented with 100 µg/mL ampicillin.

Synthetic genes and oligonucleotides. Oligonucleotides and synthetic genes were commercially synthesized (Integrated DNA Technologies, Inc. and Thermo Fisher Scientific Inc., respectively). Sequences of synthetic genes and oligonucleotides can be found in **Supplementary Tables 4** and **5**, respectively.

Plasmids, strains and library construction. Except *A. thaliana At4CL-2* (NM_113019.3) and *S. cerevisiae ScTKL1* (NM_001184171.1), all genes encoding *K. pneumoniae AroY.B* (AAY57854.1), *AroY.Ciso* (BAH20873.1), *AroY.D* (AAY57856.1), *C. albicans CatA* (XP_722784.1), *P. anserina AroZ* (XP_001905369.1), *Acinetobacter* sp. ADP *BenM* (AAC46441.1), *A. thaliana AtC4H* (NM_128601.2), *A. thaliana* AtATR2 (NM_179141.2), *A. thaliana AtPAL2* (NM_115186.3), *Petunia hybrida PhCHI* (X14589), *Hypericum androsaemum HaCHS* (AF315345), *Schizosaccharomyces pombe MAE1* (NM_001020205.2), *Sinorhizobium meliloti PcaQ* (NC_003078.1), *E. coli ArgP* (NC_000913.3), *K. pneumonia MdcR* (U14004), and *Herbaspirillum seropedicae* SmR1 *FdeR* (Hsero_1002, UniProtKB-D8J0W4_HERSS) were codon-optimized for expression in yeast (see **Supplementary Table 4** for full sequences). All gene fragments and correct overhangs for USER cloning were amplified by PCR using oligonucleotides listed and described in **Supplementary Table 5**. Unless otherwise stated, the amplified products were USER-cloned into EasyClone integrative plasmids[49], and confirmed by sequencing.

The list of the constructed plasmids can be found in **Supplementary Table 3**. Transformation of yeast cells was carried out by the lithium acetate method[50], and strains selected on synthetic drop-out medium (Sigma-Aldrich), selecting for appropriate markers. For selection of strain carrying KanMX and HypMX, the media were supplemented with 200 µg/mL G418 sulfate and 200 µg/mL hygromycin B, respectively. Transformants were genotyped using oligonucleotides described in **Supplementary Table 5**. The resulting strains are listed in **Supplementary Table 1**.

To establish the CCM producing strains, we expressed the dehydroshikimate DHS dehydratase from *P. anserina* (PaAroZ), the PCA decarboxylase genes from *K. pneumoniae* (*KpAroY.B, KpAroY.Ciso* and *KpAroY.D*), and the catechol 1, 2 dioxygenase CDO from *C. albicans* (*CaCatA*) in *S. cerevisiae*. It has been reported that the conversion of PCA to catechol by PCA decarboxylase is a limiting step[25]. For this reason we expressed the *KpAroY.B* and *KpAroY.Ciso* genes in either single or multiple genomic integrations to create a small library of CCM production strains. In addition, Tkl1 was overexpressed to improve the precursor supply.

To establish a naringenin-producing strain, we integrated the full pathway containing the phenylalanine ammonium lyase gene from *A. thaliana* (*AtPAL-2*), the fusion of cinnamate 4-hydroxylase from *A. thaliana* and NADPH-cytochrome P450 reductase from *A. thaliana* (*AtC4H:L5:AtATR2*), the 4-coumarate-CoA ligase 2 from *Arabidopsis thaliana* (*At4CL-2*), the naringenin-chalcone synthase from *Hypericum androsaemum* (*HaCHS*), and the chalcone isomerase from *P. hybrida* (*PhCHI*) to make strain EVR1 (**Supplementary Table 1**). Strains EVR2 and EVR3 contained one and two additional integrations of bottleneck enzyme genes (*AtPAL-2* and *HaCHS* for EVR2; *AtPAL-2, HaCHS*, and *AtC4H:L5:AtATR2* for EVR3) (**Supplementary Table 1**).

Mutagenesis of BenM and library preparation. For optimization of the CCM inducibility of BenM, purified products from three consecutive rounds of error-prone PCR (epPCR) of the sequences encoding the EBD (residues 90–304) of BenM, were co-transformed into yeast together with linearized centromeric plasmid according to ref. 51, to allow for *in vivo* gap repair and library reconstitution of wild-type BenM DBD fused to EBD variants expressed from REV1p. For epPCR we used the GeneMorph II kit according to manufacturer's description (Agilent Technologies). Transformed yeast contained a chromosomal integration of the 209bp_CYC1p_*BenO_T1* promoter controlling the expressing of GFP at EasyClone site 4 on chromosome XII (ref. 52), to allow for FACS-based screening of improved CCM-inducible BenM variants.

Metabolite quantification using HPLC and UPLC-MS. The CCM production strains were cultivated in 24-deep-well plate with air-penetrable lids (EnzyScreen) to test for the production of CCM. Colonies from the individual strain were inoculated in 1 mL synthetic drop-out medium (Sigma-Aldrich), selecting for *URA*, *HIS* and *LEU* markers, and grown at 30 °C with 250 r.p.m. agitation at 5 cm orbit cast for 24 h. 300 µL of the overnight cultures were used to inoculate 3 mL mineral medium (pH 4.5) in 24-deep well plate and incubated for 24–72 h at the same conditions as above. Experiments were performed in triplicate. The culture broth was centrifuged 3,500 r.p.m. and the supernatant analyzed for CCM concentration using HPLC. For this purpose, samples were analyzed for 45 min using Aminex HPX-87H ion exclusion column with a 1 mM H_2SO_4 flow of 0.6 mL/min. The temperature of the column was 60 °C. Refractive index and UV detectors (Dionex) were used for detection of CCM (250 nm). CCM concentrations were quantified by comparison with the spectrum of the standards. For the naringenin production strains, 300 µl culture broth was extracted with 300 µl MeOH in a 10-min incubation (300 r.p.m., 5 cm amplitude, 30 °C) in a 96-square-deep-well microtiter plate (Greiner Masterblock, 96-well, 2 ml, P, V-bottom) and subsequently clarified by centrifugation at 4,000*g* for 5 min. Clarified broth extract was subsequently diluted four times with 50% MeOH, and 2 µl was injected on a Acquity UPLC system (Waters) coupled to a Acquity TQ mass detector (Waters). Separation of the compounds was achieved on a Acquity UPLC BEH C18 column (Waters, 1.7 µm, 2.1 mm × 50 mm), kept at 55 °C. Mobile phases A and B were water containing 0.1% formic acid and acetonitrile containing 0.1% formic acid, respectively. A flow of 0.6 mL/min was used. The gradient profile was as follow: 0.3 min constant at 10% B, a linear gradient from 10% B to 25% B in 3.7 min, a second linear gradient from 25% B to 100% B in 1 min, a wash for 1 min at 100% B and back to the initial condition of 10% B for 0.6 min. The mass analyzer was equipped with an electrospray (ESI) source and operated in negative mode. Capillary voltage was 3.0 kV; the source was kept at 150 °C and the desolvation temperature was 350 °C; desolvation and cone gas flow were 500 L/h and 50 L/h respectively. $[M\text{-}H]^-$ ions of naringenin (271 *m/z*) was tracked in SIR mode. Naringenin was quantified using a quadratic calibration curve with authentic standards ranging from 0.039 mg/L to 20 mg/L using QuanLynx software (Waters).

Transport assays. Overnight grown CEN.PK113-5A cells were diluted (OD_{600} = 0.1) into SC medium lacking His and Leu, with or without 1.4 mM (200 mg/L) CCM. Media samples were taken at both 0 h and 24 h, while samples for measuring cellular lysates (10^8 cells) were harvested at 24 h. For quantification of CCM by LC-MS, cultures were harvested by centrifugation. For extracellular CCM quantification the supernatant was centrifuged twice and filtered (0.2 µm) before analysis. For intracellular CCM quantification harvested pellets were washed twice in ice-cold isotonic saline solution (0.9% NaCl) and centrifuged at 5,000*g* before cells were extracted in an aqueous 0.1% formic acid solution and sonicated for 15 min. Following this, samples were centrifuged at 13,000*g* and supernatants were filtered (0.2 µm) before analysis. LC-MS data were collected on EVOQ EliteTriple Quadrupole Mass Spectrometer system coupled with an Advance UHPLC (Bruker). Samples were held at 4 °C during the analysis. A 1-µL sample was injected onto a ACQUITY HSS T3 C18 UHPLC column

(Waters), with a 1.8-µm particle size and 2.1 × 100 mm dimensioning. The column was held at a temperature of 30 °C. The solvent system used was 0.1% formic acid (mobile phase A) and acetonitrile with 0.1% formic (mobile phase B). The flow rate was 0.400 mL/min with an initial solvent composition of 100% mobile phase A held until 0.50 min, then changed until it reached %A = 5.0 and %B = 95.0 at 1.00 min. This was held until 1.79 min when the solvent was returned to the initial conditions and the column was re-equilibrated until 4.00 min. The eluent was sprayed into the heated ESI probe of the mass spectrophotometer, which was held at 250 °C and 2,500 V. Sheath/nebulizer/cone gas flow rate of 50/50/20 units and cone temperature was 350 °C. Two transitions were chosen in negative multiple reaction monitoring (MRM) mode for quantification of CCM: *m/z* 141.70–96.80 (quantification transition) and *m/z* 141.70–53.1 (confirmation transition). Triplicate measurements were made for all samples.

Fluorescence-activated cell sorting. A two-step method was used to sort for BenM variants that specifically induce in the presence of CCM. Cells (10× library size, ~400,000 cells) were inoculated in mineral medium without inducer and incubated for 24 h at 30 °C, diluted into PBS, and then GFP intensity of individual cells was measured using a BD Biosciences Aria (Becton Dickinson) with a blue laser (488 nm) by applying tight gates on the FSC and SSC channels. Only cells displaying autofluorescence intensities were sorted to limit autoactivating BenM variants. Sorted cells were recovered in mineral medium, followed by subculturing (1:100) into mineral medium containing 1.4 mM (200 mg/L) *cis,cis*-muconic acid. The cells were grown for 24 h at 30 °C, washed and subjected to a second round of FACS. Cells exhibiting high levels of GFP (top 1%) were sorted, recovered in mineral medium and plated for single colonies on SC medium lacking His and Leu. Individual clones were subsequently validated using flow cytometry.

Flow cytometry measurements and data analysis. Cells grown for 24 h in control (mineral medium, pH 4.5) or inducing medium (mineral medium pH 4.5 + 1.4 mM CCM, 1.4 mM protocatechuic acid, 10 mM malonic acid, 0.2 mM naringenin, or 50 mM L-arginine) were diluted into PBS to arrest cell growth. Cells were then analyzed by flow cytometry using a Fortessa flow cytometer (Becton Dickinson) with a blue laser (488 nm), for validation of single strains. For each strain 10,000 single-cell events were recorded. Events were analyzed using FlowJo software (TreeStar Inc.). The fluorescence arithmetic mean of the gated cell population was calculated, and the fold-change determined by dividing the mean fluorescence of the induced (ON) state with the mean fluorescence of the control (OFF) state. For flow cytometry for CCM and naringenin producing cells we tested mean fluorescence intensities from 10,000 cells pr. strain following 72 and 48 h, respectively. The data represent the average of three ($n = 3$) biological replicates and error bars correspond to the s.d. between these measurements.

Transcriptome analysis. To study the impact of ligand-induced BenM on genome-wide gene expression, triplicate cultures of strains MeLS0138 and MeLS0284 were grown for 24 h at 30 °C in 50 ml mineral medium pH 4.5 with 1.4 mM CCM. Following this, total RNA was extracted essentially as previously described[53]. Briefly, 15 ml samples of the six cultures were harvested into a pre-chilled 50 ml tube with crushed ice and then immediately centrifuged at 4 °C, 4,000 r.p.m. for 5 min. Subsequently, the pellets were resuspended in 2 ml of RNAlater Solution (Ambion, Life Technologies) and incubated on ice for 1 h. Next, cells were pelleted by centrifugation (12,000 r.p.m. for 10 s) and transferred to liquid nitrogen, and stored at −80 °C until further analysis. Total RNA extraction was performed using RNeasy Mini Kit (QIAGEN). For this purpose, samples were thawed on ice, and 600 µl of buffer RLT containing 1% (v/v) β-mercaptoethanol was added directly to the cells, before being transferred into a 2 ml extraction tube containing 500 µl glass beads and disrupted using the PRECELLYS24 (Bertin Technologies) for 2 × 20 s at 6,500 r.p.m. The cell mixture was pelleted, and supernatant was transferred to a new tube. Total RNA was purified according to the manufacturer's protocol, and genomic DNA removed using Turbo DNA-free Kit (Ambion). The quantity and quality of the RNA samples were measured using Qubit 2.0 Fluorometer using the Qubit RNA BR Assay (Thermo Fisher Scientific) and Agilent 2100 Bioanalyzer using the RNA 6000 Nano Kit (Agilent Technologies), respectively. For sequencing, we used 3 µg of total RNA as input for TruSeq Stranded mRNA Sample Preparation kit before sequencing on the MiSeq System using MiSeq Reagent Kit v3 150 cycles at a 2 × 75 bp read length configuration (Illumina) obtaining 38 M reads.

Bioinformatic resources. Two-dimensional heatmap plots were generated using the plot3D package and the R GUI. For ribbon structure representation of CCM-binding domain of BenM H110,F211V,Y286N the UCSF Chimera software was used[54]. For RNA-seq data analysis, TopHat (2.0.13) and Cufflinks (2.2.1) suite were used as previously described[48]. Expression levels (fragments per Kilobase of exon per million fragments mapped; FPKM) from three ($n = 3$) biological replicates of the conditions tested were processed with cuffdiff to obtain fold change differences and to perform statistical testing. A *q*-value cutoff of < 0.05 was used to identify genes that have significant differential expression. Additionally, a >2-fold cut-off selection criterion was applied. Reference genome, and annotations for CEN.PK113-7D strain were retrieved from Saccharomyces Genome Database (SGD; http://www.yeastgenome.org/). Genes with FPKM = 0 for any replicate were removed from consideration.

Database for RNA-seq data. RNA-seq data are available in the ArrayExpress database (E-MTAB-4836); see also **Supplementary Dataset 1**.

49. Jensen, N.B. *et al.* EasyClone: method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae. *FEMS Yeast Res.* **14**, 238–248 (2014).
50. Gietz, R.D. & Schiestl, R.H. Large-scale high-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2**, 38–41 (2007).
51. Eckert-Boulet, N., Pedersen, M.L., Krogh, B.O. & Lisby, M. Optimization of ordered plasmid assembly by gap repair in Saccharomyces cerevisiae. *Yeast* **29**, 323–334 (2012).
52. Mikkelsen, M.D. *et al.* Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metab. Eng.* **14**, 104–111 (2012).
53. Kildegaard, K.R. *et al.* Evolution reveals a glutathione-dependent mechanism of 3-hydroxypropionic acid tolerance. *Metab. Eng.* **26**, 57–66 (2014).
54. Pettersen, E.F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

doi:10.1038/nchembio.2177