



RESEARCH

Open Access



A toolset of constitutive promoters for metabolic engineering of *Rhodospiridium toruloides*

Luísa Czamanski Nora^{1,2,3} , Maren Wehrs^{2,3,4}, Joonhoon Kim^{2,5}, Jan-Fang Cheng^{6,8}, Angela Tarver^{6,8}, Blake A. Simmons^{2,3}, Jon Magnuson^{2,5}, Miranda Harmon-Smith⁶, Rafael Silva-Rocha¹, John M. Gladden^{2,3,7}, Aindrila Mukhopadhyay^{2,3,8} , Jeffrey M. Skerker^{8,9*} and James Kirby^{2,3,7*}

Abstract

Background: *Rhodospiridium toruloides* is a promising host for the production of bioproducts from lignocellulosic biomass. A key prerequisite for efficient pathway engineering is the availability of robust genetic tools and resources. However, there is a lack of characterized promoters to drive expression of heterologous genes for strain engineering in *R. toruloides*.

Results: This data describes a set of native *R. toruloides* promoters, characterized over time in four different media commonly used for cultivation of this yeast. The promoter sequences were selected using transcriptional analysis and several of them were found to drive expression bidirectionally. Promoter expression strength was determined by measurement of GFP and mRuby2 reporters by flow cytometry. A total of 20 constitutive promoters (12 mono-directional and 8 bidirectional) were found, and are expected to be of potential value for genetic engineering of *R. toruloides*.

Conclusions: A set of robust and constitutive promoters to facilitate genetic engineering of *R. toruloides* is presented here, ranging from a promoter previously used for this purpose (P7, glyceraldehyde 3-phosphate dehydrogenase, GAPDH) to stronger monodirectional (e.g., P15, mitochondrial adenine nucleotide translocator, ANT) and bidirectional (e.g., P9 and P9R, histones H3 and H4, respectively) promoters. We also identified promoters that may be useful for specific applications such as late-stage expression (e.g., P3, voltage-dependent anion channel protein 2, VDAC2). This set of characterized promoters significantly expands the range of engineering tools available for this yeast and can be applied in future metabolic engineering studies.

Keywords: *Rhodotorula*, Bidirectional promoters, Metabolic engineering, Synthetic biology, Transformation, Fluorescent reporters, Promoter characterization

Background

With the aim of replacing many petrochemical-derived fuels and chemicals with renewable alternatives, the field of metabolic engineering is moving towards the development of microbial hosts that use lignocellulosic biomass as feedstocks for bioproduction. The oleaginous,

carotenogenic yeast *Rhodospiridium toruloides* (also known as *Rhodotorula toruloides*) is a promising host for sustainable bioproduction due to its innate capacity to efficiently utilize both hexoses and pentoses present in lignocellulosic biomass, as well as its capacity to grow on lignin derived intermediates and aromatics, such as *p*-coumaric acid and ferulic acid [1]. So far, much of the work on metabolic engineering or bioproduct formation in *R. toruloides* has focused on lipid accumulation [2, 3]. However, the high lipid content indicates a suitability for applications to any acetyl-CoA-derived product, from

*Correspondence: jmskerker@lbl.gov; jkirby@lbl.gov

² DOE Joint BioEnergy Institute, 5885 Hollis St, Emeryville, CA 94608, USA

⁸ Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720, USA
Full list of author information is available at the end of the article



fatty alcohols, terpenes and methyl ketones to TCA-cycle derived products, with many applications such as fuels, chemicals, and pharmaceuticals [1, 4–7].

To improve prospects for more advanced metabolic engineering in this fungal host, the development of robust genetic tools and resources is an immediate priority [8]. A major need on the tools front is for efficient and stable promoters to express genes of interest and selection markers. A growing branch of synthetic biology involves the pursuit and improvement of standard parts that are reliable, orthogonal and robust for non-conventional hosts [9, 10]. A small selection of promoters has already been characterized to modulate expression of heterologous genes in *R. toruloides*, the majority of them being metabolite-responsive promoters rather than constitutive [8, 11–14]. While promoters that are responsive to certain metabolites are valuable tools, a toolset of well characterized constitutive promoters remains necessary to explore the full potential of strain engineering. In this context, functional bidirectional promoters would be a valuable addition to the toolset, since they can be convenient for optimization of multi-gene pathways.

The most common method used for transformation of *R. toruloides* is *Agrobacterium tumefaciens*-mediated transformation (ATMT), a useful strategy for genomic integration of heterologous genes but is a laborious and time-consuming process [15, 16]. For this work, a high-efficiency chemical transformation method was adopted from a protocol developed for *Saccharomyces cerevisiae* [17] to streamline the engineering workflow for *R. toruloides*. RNA sequencing analysis was performed to select promoters from *R. toruloides* that are likely to result in high- and medium-level constitutive expression. Several of these promoters were predicted to be bidirectional. The aim of this study was to characterize native *R. toruloides* promoters by monitoring expression of two fluorescent reporters in four different media over a cultivation period of 7 days. Of the set of promoters tested, several appear to be good candidates for driving heterologous expression in this host, thus expanding the toolset available for genome editing.

Results

Construction and characterization of promoter library

RNA-sequencing data derived from *R. toruloides* cultivated on several different carbon sources was used to identify constitutive promoters driving high and medium expression levels of native genes (predicted expression levels are classified as described in “Methods” section). All selected promoters are considered to be constitutive, based on variance in transcript abundance across six different growth conditions. Of the 29 promoter regions selected, 13 comprise a complete intergenic sequence

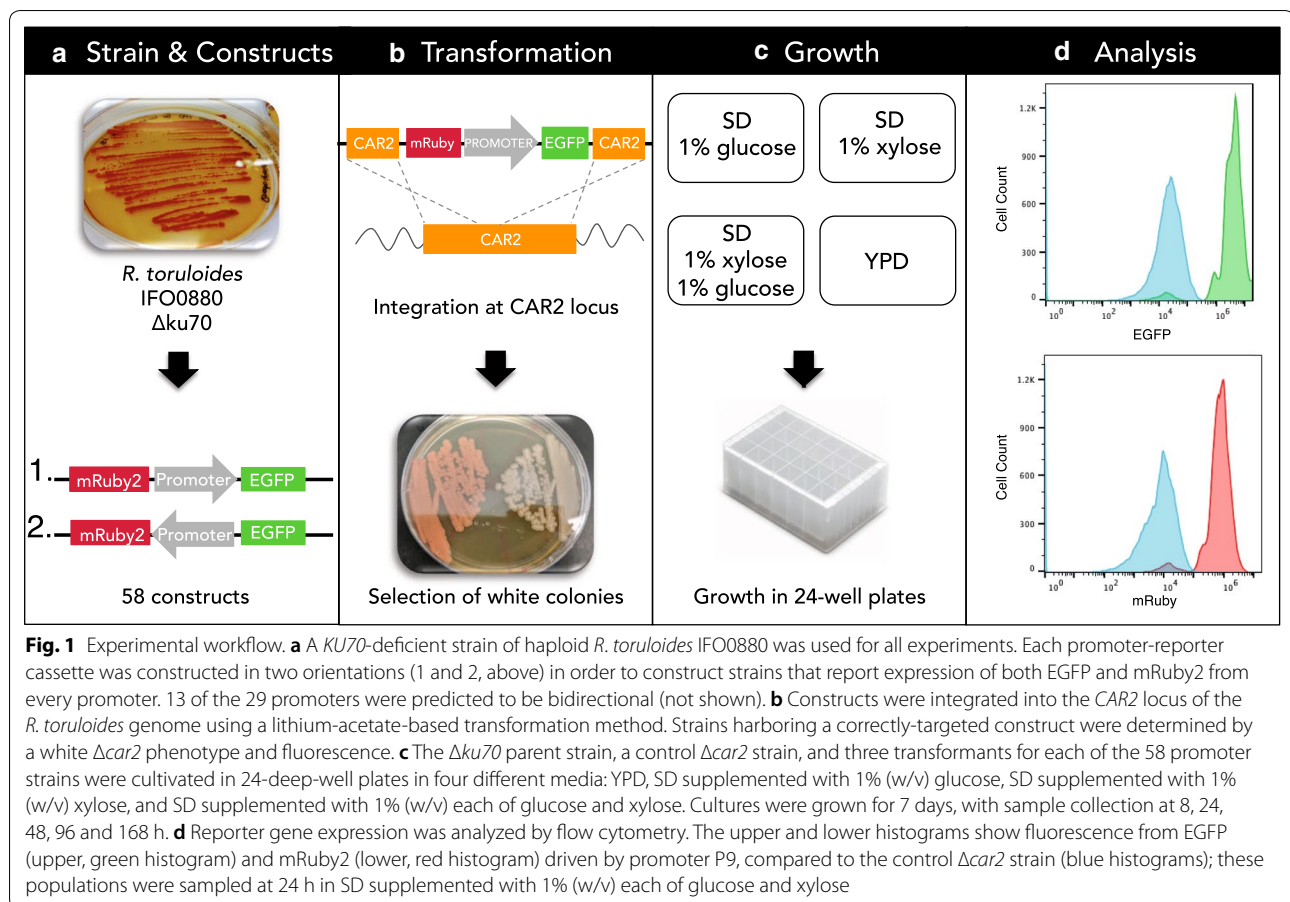
(i.e., between start codons of two neighboring divergent genes) and were therefore identified as putative bidirectional promoters. Each promoter was cloned into a reporter gene construct consisting of mRUBY2 and EGFP in two orientations, designated as orientation 1 and 2 (Fig. 1a and Additional file 1: Figure S1). A strain containing the reporter gene cassette lacking a promoter was used as the negative control, referred to here as $\Delta car2$. In total, 59 constructs were transformed and integrated into the *R. toruloides* genome, 29 constructs in orientation 1 and 29 in orientation 2 (Fig. 1a), and one control with no promoter.

The workflow from construction of expression constructs to promoter analysis by flow cytometry is shown in Fig. 1. Each construct was flanked by 1 kb sequences homologous to the *R. toruloides* *CAR2* gene, which encodes a bifunctional phytoene/lycopene synthase [11, 12]. Disruption of the *CAR2* gene through homologous recombination results in white colonies due to lack of carotenoid production (Fig. 1b). Three confirmed transformants for each construct were inoculated into 24-deep-well plates containing LB medium, grown overnight, and then sub-cultured into the various test media at a starting OD₆₀₀ of 0.05. One rich medium (YPD) and three defined media (SD supplemented with 1% glucose, 1% xylose, or 1% of each sugar) were selected to represent typical media compositions [1] to culture *R. toruloides* (Fig. 1c).

To select appropriate time points for reporter fluorescence measurements, we monitored optical density at 600 nm (OD₆₀₀) and measured extracellular sugars by high performance liquid chromatography (HPLC) in all four media. Representative data, for the parental strain, *R. toruloides* IFO0880 $\Delta ku70$ and a strain harboring reporter construct P14, can be found in Additional file 2: Figure S2 and Additional file 3: Figure S3, respectively. Samples were collected at 8, 24, 48, 96 and 168 h for measurement of reporter fluorescence by flow cytometry. For constructs containing bidirectional promoters, fluorescence of both reporters was measured simultaneously.

Monodirectional promoters

Of the 16 putative monodirectional promoters selected for investigation, 12 resulted in medium to high fluorescence when driving EGFP expression in *R. toruloides*. Hierarchical clustering of a heatmap generated from flow cytometry data (Fig. 2) shows that the strongest monodirectional promoters under most conditions are P14 (translation elongation factor 1, TEF1), P17 (hypothetical protein, RTO4_15825), and P15 (mitochondrial adenine nucleotide translocator, ANT)—see Additional file 5: Table S1 for more information on promoter sequences. Overall (in decreasing order of



strength) promoters P14, P17, P15, P4, P27, P18, and P10 are classified as strong promoters, while promoters P7, P1, P5, P3 and P6 are clustered as medium-strong promoters. It is worth noting that promoter P7, which natively drives expression of GAPDH and is commonly viewed as a benchmark strong promoter, is classified as medium-strong within this promoter set.

In order to evaluate consistency of expression of different coding sequences, promoters were also used to drive expression of mRuby2 (Fig. 3). Again, P14, P15 and P17 were clustered as the strongest promoters while P6, P3 and P1 were classified as medium-strong promoters.

Bidirectional promoters

Each of the 13 promoters that was predicted to be bidirectional (i.e., selected by transcriptomics and comprising a complete intergenic sequence in *R. toruloides*) was positioned in between the divergent coding sequences for EGFP and mRuby2, allowing expression in both directions to be monitored simultaneously. Of these, eight resulted in medium-to-high reporter fluorescence in both directions in at least one condition. Hierarchical

clustering of a heatmap generated from flow cytometry data shows that the strongest bidirectional promoter pairs are P9 and P9R (histones H3 and H4, respectively), P12 and P12R (small subunit ribosomal proteins S28e and S5e, respectively), and P19 and P19R (ribosomal proteins S15e and LP2, respectively) (Fig. 4). The forward (EGFP) and reverse (mRuby2) promoter strengths for these three bidirectional pairs are well-balanced across a range of conditions. Under most conditions, promoters P9 and P9R, constituting the strongest bidirectional promoter pair, are similar in strength to the strongest of the unidirectional promoters. Promoter pairs P13/R, P28/R, P22/R, P29/R and P8/R were clustered as medium-strength promoters in at least one direction. Of these, P29 and P29R are the most evenly-balanced, while P8 and P8R are the most divergent in terms of promoter strength. P8R (ribosomal protein, L37e) is significantly stronger than P8 (ubiquitin/40S ribosomal protein S27a fusion), across most of the conditions tested.

Fluorescence from a second reporter was measured for each of these bidirectional promoters by cloning them in orientation 2 (Fig. 5 and Additional file 1: Figure S1). Promoters P9/R, P19/R and P12/R were again classified

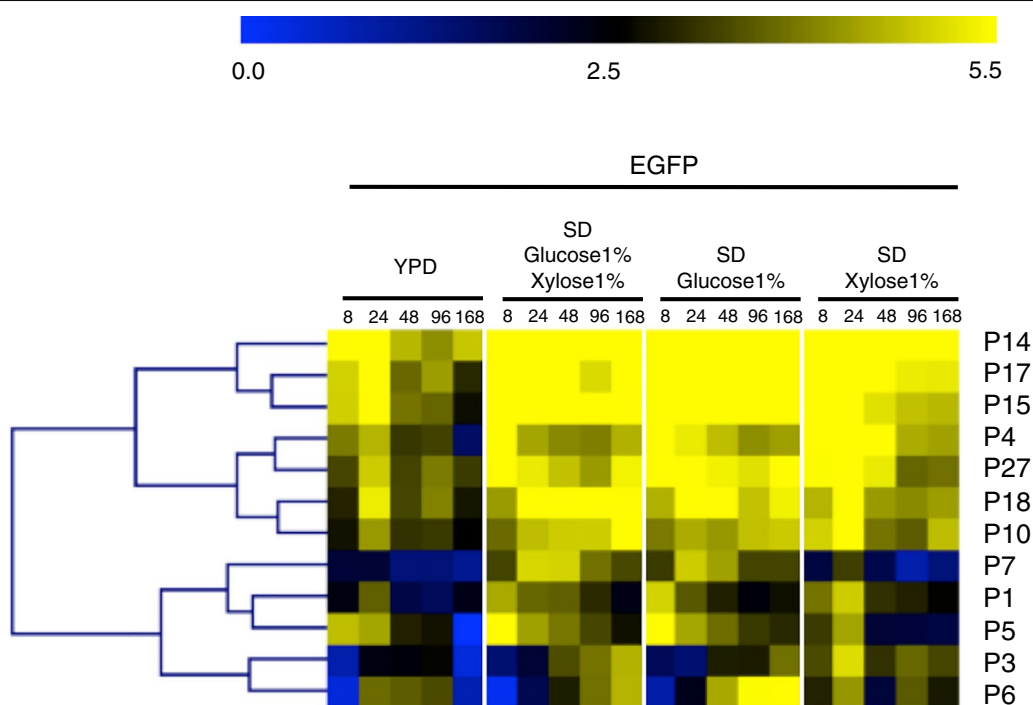


Fig. 2 Expression of EGFP from monodirectional promoters. Fluorescence was measured by flow cytometry of cell populations sampled at five time points from four different media for strains harboring promoter constructs in orientation 1. The heatmap was generated using MultiExperiment Viewer (MeV), and promoters were clustered by hierarchical clustering using Euclidean distance. Fluorescence expression values are on a Log2 scale (color scale is shown in the upper bar), calculated as described in "Methods". Each column represents a time point, shown in hours. Each row represents a promoter, named on the right side of the heatmap

as the strongest bidirectional promoter pairs while P8/R and P29/R were overall the weakest promoter pairs. Promoter P22 was classified as medium bidirectional promoter when assayed in both orientations.

Correlation of mRuby2 and EGFP expression

To evaluate promoter reliability and robustness, the expression level of the two reporter genes (EGFP and mRuby2) driven by the same promoter were compared. Linear regression and correlation analysis of fluorescence from the two reporters was performed for all the promoters comparing EGFP from orientation 1 and mRuby2 from orientation 2 (EGFP expression for these promoters is shown in Figs. 2 and 4 and mRuby2 expression is shown in Figs. 3 and 5) in the four different media at 48 h (Fig. 6).

Promoter P1, P12, and P29 expressed EGFP at significantly higher levels than mRuby2. This could suggest a composability effect, in which contextual effects such as the mRNA secondary structure surrounding the ATG is less favorable for EGFP than for mRuby2 [18–20]. The benchmark promoter P7 (GAPDH) expressed mRuby2 at higher levels than EGFP in strains grown in YPD medium or SD with 1% xylose, indicating that expression levels

from this promoter may be less predictable in certain media. Most of the strongest promoters, including P4, P9, P10, P14, P15, P17, P18 and P27, produce similar levels of EGFP and mRuby2 fluorescence, and are recommended for metabolic engineering needs.

Discussion

The lack of standard parts that can be reliably used for genetic manipulation is often a bottleneck in metabolic engineering studies involving non-model microbial hosts. The aim of this study was to find and characterize a collection of promoters to support reliable gene expression in the fungal host *R. toruloides*. We selected 29 *R. toruloides* promoters, based on transcriptomics studies, and examined them in four different media over 7 days using a dual-reporter system.

P_{GAPDH} is one of the most widely used promoters for engineering a variety of yeasts, including *R. toruloides*, where it was employed for expression of heterologous terpene synthases [1]. However, 13 of the promoters investigated in this study were found to be stronger than P_{GAPDH} under the majority of conditions (Figs. 2 and 4). Several of these promoters are currently in use for engineering production of various bioproducts in *R.*

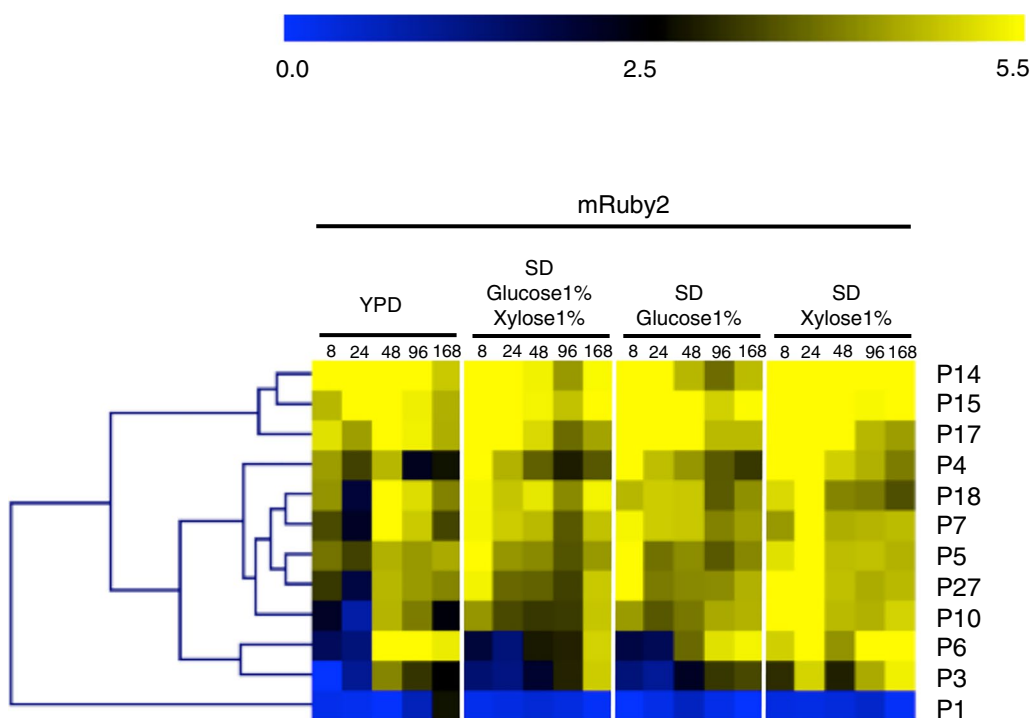


Fig. 3 Expression of mRuby2 from monidirectional promoters in orientation 2. Fluorescence was measured by flow cytometry of cell populations sampled at five time points from four different media. The heatmap was generated using MultiExperiment Viewer (MeV), and promoters were clustered by hierarchical clustering using Euclidean distance. Fluorescence expression values are on a Log₂ scale (color scale is shown in the upper bar), calculated as described in "Methods". Each column represents a time point, shown in hours. Each row represents a promoter, named on the right side of the heatmap

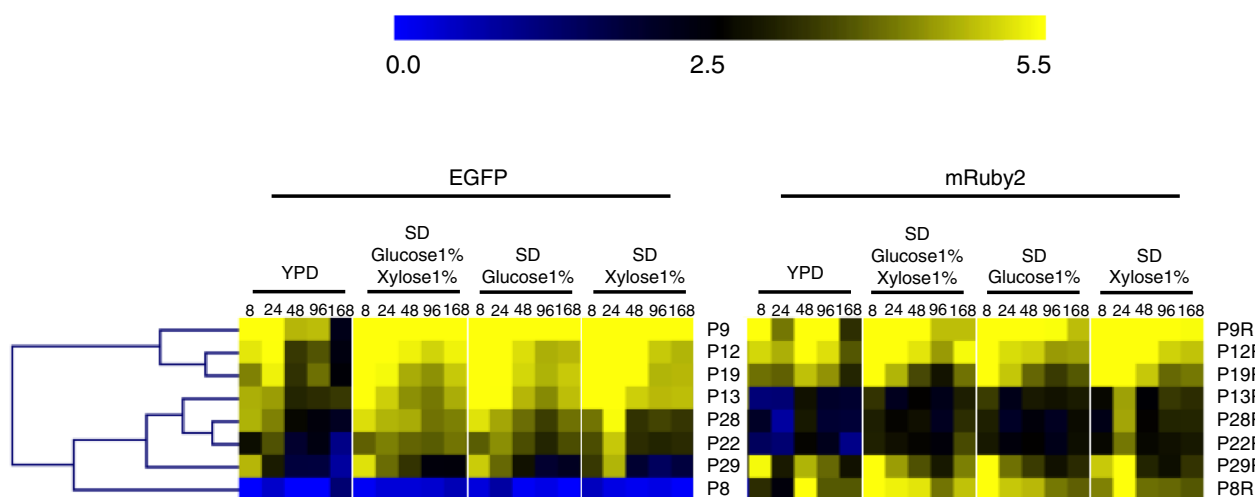
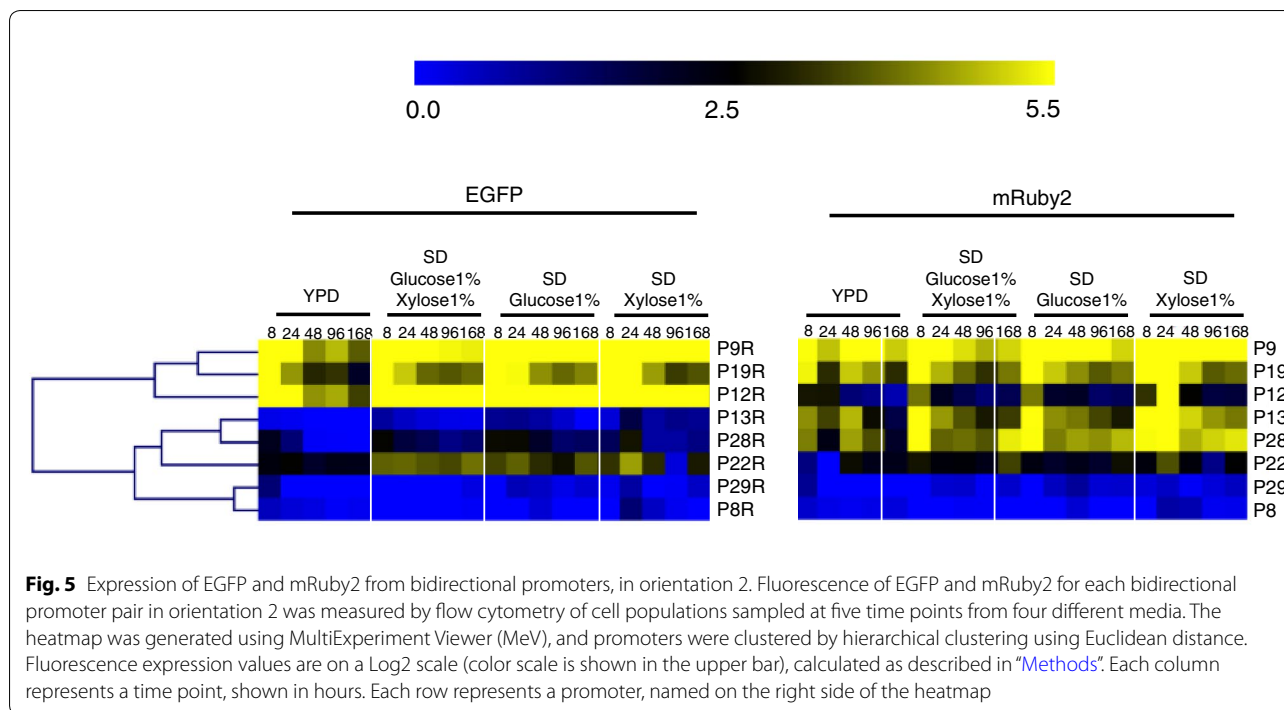


Fig. 4 Expression of EGFP and mRuby2 from bidirectional promoters. Fluorescence of EGFP and mRuby2 for each bidirectional promoter pair was measured in strains harboring promoter constructs in orientation 1. The heatmap was generated using MultiExperiment Viewer (MeV), and promoters were clustered by hierarchical clustering using Euclidean distance. Fluorescence expression values are on a Log₂ scale (color scale is shown in the upper bar), calculated as described in "Methods". Each column represents a time point, shown in hours. Each row represents a promoter, named on the right side of the heatmap. All time points for all media are shown for both EGFP and mRuby2



toruloides, and reports on their utility are anticipated in the near future. The first of these reports on the use of promoters P14 and P15 to produce the non-ribosomal peptide, indigoidine, at titers approaching 3 g/L from a sorghum lignocellulosic hydrolysate [7].

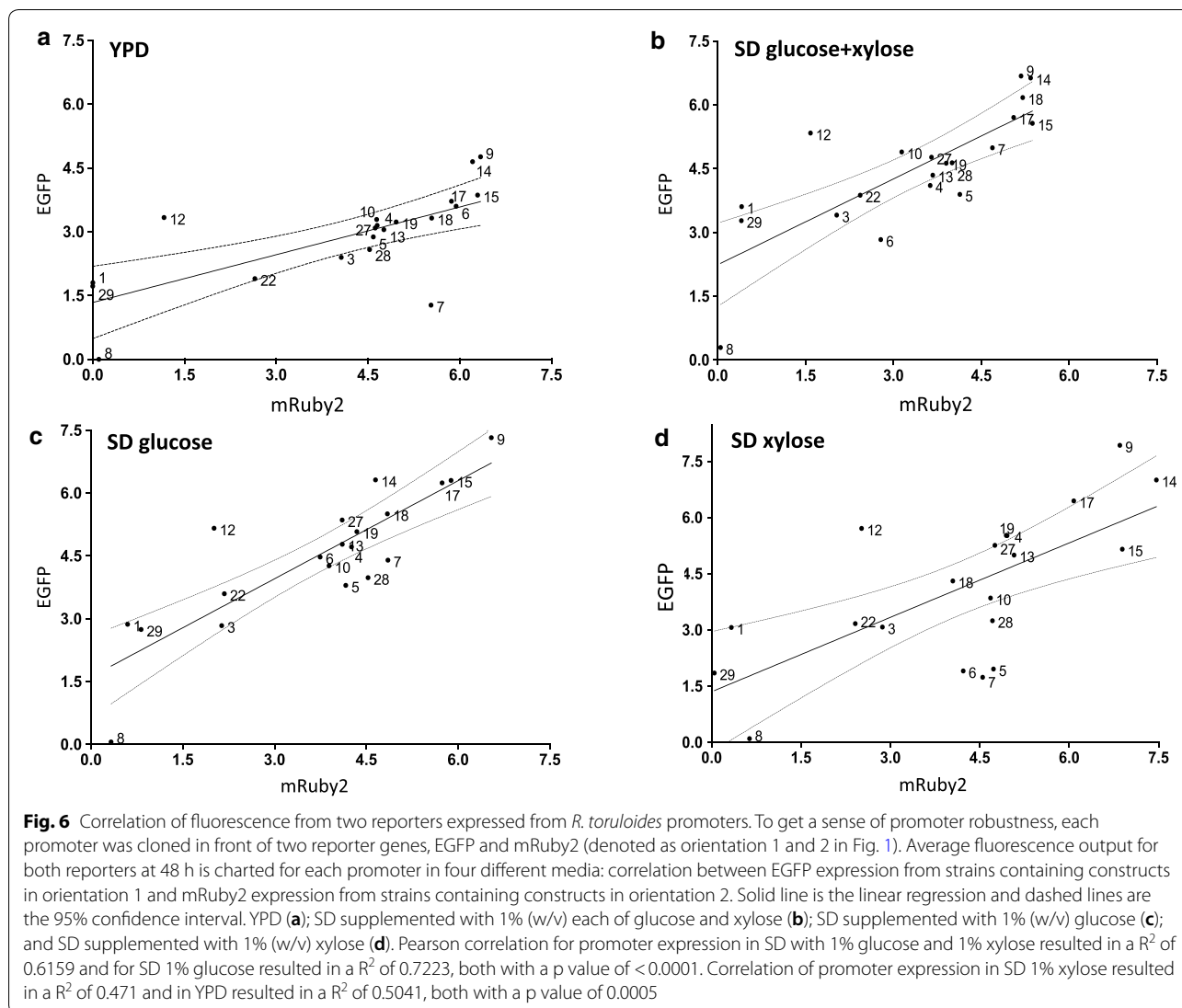
Temporal changes in promoter strength can be an important factor in guiding their selection for pathway engineering. Under control of several of the promoters, reporter fluorescence reaches its maximum early (at 8 or 24 h) and then diminishes over time or drops to a lower but stable level (Figs. 2 and 4). Interestingly, this occurs for two of the strongest promoters, P15 (*ANT*) and P9 (histone H3) when strains are cultured in YPD medium, but not in the SD media. This is likely due to the cells reaching stationary phase significantly faster in YPD compared to SD media (Additional file 2: Figure S2 and Additional file 3: Figure S3). In contrast, expression from promoters P3 (voltage-dependent anion channel protein 2, *VDAC2*) and P6 (unknown function) increases over time under some conditions. These two promoters may be useful for applications where late expression of a protein is desired (e.g., generation of metabolic products that are inhibitory to growth or consumed during log phase).

Correlation of fluorescence from EGFP and mRuby2 at the 48 h suggests that most of the promoters express both reporter genes at similar rates, suggesting a reasonable level of composability, with a few exceptions (Fig. 6). A similar analysis of fluorescence from the two reporters driven by promoter P9, which includes every time point

and growth medium, indicates that the expression level of the two reporters is similar under all conditions for this promoter (Additional file 4: Figure S4).

Promoters from *R. toruloides* that were posited to have medium or strong transcriptional activity in both directions were also investigated. Bidirectional promoters can reduce design and assembly complexity and also help keep DNA construct sizes within reasonable limits, a useful feature when engineering multigene pathways [21–23]. Of the 14 promoters that were predicted to be bidirectional according to RNA-sequencing data (P1, P2, P8, P9, P11, P12, P13, P19, P21, P22, P24, P25, P28 and P29) fluorescence of both reporters was detected for 8 of them, indicating bidirectional transcription (Fig. 4). Expression in the forward and reverse directions from the strongest bidirectional promoter pairs (P9/P9R, P12/P12R and P19/P19R) is balanced (Figs. 4 and 5), a feature that should prove valuable for metabolic engineering applications where equal expression of two genes is sought. Each of these three promoters drives expression of two genes that are closely related in function. Promoters P9 and P9R, the bidirectional pair natively responsible for expression of histones 3 and 4, respectively (Additional file 5: Table S1), was shown to be the strongest promoter of the entire collection, under most conditions.

Promoters P6, P9, P10, P12, P14 and P15 contain introns in their predicted 5' UTRs, a feature often associated with increased promoter strength. Introns can increase transcript levels by affecting the rate of



transcription, nuclear export, transcript stability and even mRNA translation [24]. In *S. cerevisiae* it was shown that intronic sequences had a positive impact on expression of ribosomal proteins, both transcriptionally and post-transcriptionally [25].

So far, there have been three reports of relatively constitutive promoters for use in *R. toruloides*, totaling a set of eight promoters: *GAPDH* (also called *GPD1*), *FBA1*, *PGK1*, *PGL1*, *TPI1*, *LDPI*, *ACC1* and *FAS1* [11, 12, 14, 15]. Since normalization of the reporter genes used to characterize these promoters differs between these studies, it is challenging to compare strength of these promoters with the collection presented here.

Synthetic biology studies often rely heavily on the fine-tuning of gene expression levels, to achieve a metabolic balance and a high product titer, rate and yield [5, 10, 26, 27]. Despite the on-going expansion of characterized

promoters for use in synthetic biology, the construction of non-native production pathways with predictable outcome remains challenging and requires further characterization. Yet, the majority of engineering work to date performed to date in *R. toruloides* has relied on random-integration strategies such as ATMT [1, 2, 15, 16], that offer little control over integration locus or copy number. The strong promoters characterized in this work will complement the recent development of CRISPR technology for *R. toruloides*, allowing for site-specific integration while maintaining high expression rates [28].

Conclusions

RNA sequencing data was found to be a useful starting point for identification of *R. toruloides* promoters that can be used for heterologous expression. The collection of 12 monodirectional and 8 bidirectional native

promoters presented in this work is the largest promoter set published for *R. toruloides*, including the first bidirectional promoters reported for this organism. Among these were 13 promoters that are stronger than the benchmark GAPDH promoter, the most commonly used promoter for yeasts. This characterized promoter set expands the *R. toruloides* genetic toolbox and is likely to be valuable for future metabolic engineering efforts in this promising host.

Methods

Media and growth conditions

Synthetic defined (SD) medium was made using Difco yeast nitrogen base (YNB) with ammonium sulfate and without amino acids (Becton, Dickinson & Co., Sparks, MD) and complete supplemental mixture (CSM; Sunrise Science Products, San Diego, CA), following manufacturer's directions. Lysogeny broth (LB) was made using Difco LB (Miller) mixture (Becton, Dickinson & Co.) and yeast peptone dextrose (YPD) medium was made as normal, including 2% peptone, 1% yeast extract (Becton, Dickinson & Co.) and 2% glucose. Xylose and glucose were from Sigma-Aldrich (St. Louis, MO).

For the promoter studies, all 60 *R. toruloides* IFO0880 $\Delta ku70$ strains (the parent strain, 58 strains containing promoter constructs, a $\Delta car2$ control strain) were inoculated into 24-deep-well plates containing 2 mL of LB (Miller) and grown overnight at 30 °C with shaking at 200 rpm. These cultures were then used to inoculate (at a 1:100 dilution) four different media in 24-deep-well plates: SD 1% xylose, SD 1% glucose, SD 1% xylose plus 1% glucose, and YPD. These cultures were grown to exponential phase overnight at 30 °C with shaking at 200 rpm and were then used to inoculate 3 mL of identical media in 24-deep-well plates at a starting OD₆₀₀ of 0.05. These cultures were grown for 7 days at 30 °C in a shaker at 200 rpm with samples taken every 8, 24, 48, 96 and 168 h.

Strains and plasmid construction

The parent strain for all work described here is *R. toruloides* IFO0880 harboring a deletion for the non-homologous end joining (NHEJ) gene, *KU70*, to facilitate a higher proportion of site-specific integration events in each transformation (Fig. 1a) [2]. *R. toruloides* strain IFO0880 $\Delta ku70$, a haploid strain, was used for transformations to favor homologous recombination. Strains and plasmids used in this study are available upon request through the Joint BioEnergy Institute Strain Registry (<https://public-registry.jbei.org/>). Gene synthesis and plasmid construction were performed by Joint Genome Institute (JGI) and two constructs were designed for each promoter so that it was positioned between EGFP and

mRuby2 coding sequences in both directions. Constructs also contain a Nourseothricin acetyltransferase (*nat1*) cassette, conferring resistance to Nourseothricin, and homology regions corresponding to the *CAR2* locus. A schematic representation of the expression cassette built for integration into *R. toruloides* genome can be seen in Fig. 1 and selected promoters and their combined annotations can be seen in Additional file 5: Table S1.

RNA-sequencing

RNA-sequencing was performed at the Joint Genome Institute and used to identify a set of promoters that drive high or medium expression in different media including SD medium with 2% glucose (NCBI Accession SRP183741, SRP183740, SRP183739), and M9 minimal medium with added trace elements and 2% glucose (NCBI Accession SRP164139, SRP164140, SRP164141) or 10 mM *p*-coumaric acid (NCBI Accession SRP164142, SRP164143, SRP164144). Another dataset was generated at the University of California, Berkeley and included in the analysis (NCBI Accession GSE128360), which consists of samples grown in YNB minimal medium with 3% glucose and different carbon to nitrogen ratios (7 mM ammonium chloride or 168 mM ammonium chloride), and a library made of pooled samples grown in several media including YNB medium with different carbon sources (glucose, xylose, glycerol, or glutamate) and YPD medium with different conditions (growth phase, osmolarity, temperature, and oxygen availability). Sequenced reads were trimmed and filtered using BBTools (<https://jgi.doe.gov/data-and-tools/bbtools/>), mapped to IFO0880 genome sequence using HISAT2 [29], mapped reads were assigned to genes using featureCounts [30], and read counts were used to calculate Fragments Per Kilobase of transcript per Million mapped reads (FPKM). High expression was defined as log₂ of the mean FPKM values above 10, and medium expression was defined as log₂ of the mean FPKM values between 8 and 10. In order to select promoters among these that resulted in relatively constitutive expression, 20 genes with the lowest variance across different growth conditions were selected. Subsequently, promoter regions of high and medium expression genes were selected for testing, spanning 1 kb of upstream sequence or a shorter distance in cases where an upstream CDS was identified. In some cases, the predicted promoter region of a selected gene was cut short by an upstream CDS on the opposite strand and was annotated as a putative bidirectional promoter.

R. toruloides transformation

Plasmids were purified from *E. coli* DH10B strains using a QIAprep plasmid miniprep kit (Qiagen, Redwood City, CA, USA) and digested with *Pvu*II (FastDigest,

Thermo Scientific, Waltham, MD, USA) to release each construct from the vector backbone (with exception of constructs 5, 14, 18, 34, 43 and 47 that were amplified by PCR using Phusion High Fidelity DNA Polymerase (Thermo Scientific, Waltham, MD, USA) since they had *PvuII* sites inside their promoter sequences). The transformation method was adapted from Gietz [17]. *R. toruloides* IFO0880 $\Delta ku70$ strain was inoculated into 10 mL of YPD and incubated overnight at 30 °C with shaking at 200 rpm. Then, the culture was diluted to an OD₆₀₀ of 0.2 in fresh YPD and incubated at 30 °C on a shaker at 200 rpm until it reached an OD₆₀₀ of 0.8. The cultures were harvested, centrifuged at 3000×g for 10 min, washed once with 5 mL of water and then resuspended in 1 mL 100 mM lithium acetate (pH 7.5) (LiAc) in 1.7 mL tubes. The cells were centrifuged again and the LiAc was removed. Next, the transformation mixture was added, consisting of 240 µL PEG 4000 (50% w/v), 36 µL 1.0 M LiAc (pH 7.5), 10 µL single-stranded salmon sperm DNA (10 mg/mL) and 5 µg of the transforming DNA in 74 µL of water. After mixing by vortexing, cells were incubated at 30 °C for 30 min. Subsequently, 34 µL DMSO was added and cells were heat shocked at 42 °C for 15 min. Cells were then centrifuged at 7200 rpm in a microfuge and the supernatant was removed by pipetting. Cells were resuspended in 1.5 mL YPD, transferred to culture tubes, and allowed to recover overnight at 30 °C at 200 rpm. The cells were plated in YPD agar plates containing 100 µg/mL of Nourseothricin Sulfate (G-Biosciences, St. Louis, MO) and incubated at 30 °C for 2 days. The plates were then stored at 4 °C for 2 days to differentiate the white from orange colonies. For each construct, at least three independent transformants were grown overnight in YPD and tested for fluorescence by flow cytometry (Fig. 1b).

Flow cytometry

High-throughput flow cytometry experiments were performed using the Accuri C6 flow cytometer equipped with an autosampler (BD Accuri™ RUO Special Order System, Becton, Dickinson and Company, New Jersey, USA). All fluorescence measurements were performed on biological triplicates. A total of 30,000 events were recorded at a flow rate of 35 µL/min and a core size of 16 µm for each sample. Both fluorescent proteins were excited at 488 nm, EGFP emission was detected at 530 nm and mRuby2 emission was detected at 675 nm. Data acquisition was performed as described in the Accuri C6 Sampler User's Guide. The acquired data was analyzed in the FlowJo® software (Becton, Dickinson & Company), where populations were gated and the median of all fluorescence curves were obtained.

Statistics

For the fluorescence values, the median value of each fluorescence curve of each replicate was extracted using FlowJo®. The mean of the triplicates was calculated from the median values, as well as the standard deviation (STD) and the coefficient of variation (CV). The CVs of all promoters were normalized to a scale of 0 to 1. The resulting means were then divided by the mean fluorescence of the negative control $\Delta car2$ to normalize all samples in relation to background fluorescence on the EGFP and mRuby2 channels. The resulting normalized values were then transformed to log₂ values represented in the figures as Relative Fluorescence Units (RFU). Heatmaps were assembled using Multiple Experiment Viewer (MeV), where promoters were clustered by hierarchical clustering using Euclidean distance as the distance metric selection through average linkage clustering. The remaining graphs were created using GraphPad Prism® (GraphPad Software, San Diego, CA), also used to calculate linear regression as needed.

Growth rate experiments

For growth rate experiments, strains were grown in the same conditions as the fluorescence measurement experiments. Samples were taken at 0, 8, 24, 48, 96 and 168 h after inoculation. OD₆₀₀ was monitored using a SpectraMax Plus 384 spectrophotometer (Molecular Devices, San Jose, CA, USA).

High performance liquid chromatography (HPLC) analysis

Sugars were quantified on an Agilent Technologies 1200 series HPLC (Agilent Technologies, Santa Clara, CA, USA) equipped with an Aminex HPX-87H column (Bio-Rad, Hercules, CA) as described previously [31]. Samples were filtered through 0.45 µm filters (VWR, Visalia, CA, USA) before injection of 5 µL of each sample onto the column. Sugars were monitored by a refractive index detector, and concentrations were calculated by integration of peak areas and comparison to standard curves for the compounds of interest.

Additional files

Additional file 1: Figure S1. Schematic representation of the expression cassettes to be inserted into the *R. toruloides* genome through homologous recombination. Representative design of the constructs in orientation 1 (top), showing promoter P1 driving EGFP expression and P1R driving mRuby2 and orientation 2 (lower), where the promoter fragment has been reversed. Promoter-reporter constructs, paired with a *NAT* cassette (conferring resistance to nourseothricin), are flanked by *CAR2* gene fragments to facilitate integration of the cassette into the *CAR2* locus of the *R. toruloides* chromosome.

Additional file 2: Figure S2. Growth of the parental *R. toruloides* $\Delta ku70$ strain and a strain harboring promoter construct P14. Growth was monitored by OD₆₀₀ measurements, taken at 8, 24, 48, 96 and 168 h from

strains grown in: YPD (A); SD supplemented with 1% (w/v) each of glucose and xylose (B); SD supplemented with 1% (w/v) glucose (C); and SD supplemented with 1% (w/v) xylose (D).

Additional file 3: Figure S3. Glucose and xylose consumption in the four media used for this study. Two strains are shown: the parental *R. toruloides* $\Delta ku70$ strain and strain harboring promoter construct P14. Sugars were quantified by HPLC from samples taken at 8, 24, 48, 96 and 168 h from strains grown in: YPD (A); SD supplemented with 1% (w/v) each of glucose and xylose (B); SD supplemented with 1% (w/v) glucose (C); and SD supplemented with 1% (w/v) xylose (D).

Additional file 4: Figure S4. Correlation between EGFP and mRuby2 expression from promoter P9 in all 4 media and all 5 time points. Fluorescence expression values are on a Log2 scale. $R^2 = 0.5537$ for a P value of 0.0002. Solid line is the linear regression and dashed lines are 95% confidence interval.

Additional file 5: Table S1. Promoters investigated and strains generated in this study.

Acknowledgements

The authors would like to thank Samuel Coradetti for sharing RNA sequencing data, Di Liu and Gina Geiselman for the helpful comments and Sophie Comyn for assistance with confocal microscopy of *R. toruloides*. Strains and plasmids are available in The Joint BioEnergy Institute Inventory of Composable Elements (JBEI-ICE) at <https://public-registry.jbei.org/folders/437>.

Authors' contributions

JMG and JMS conceived the project. MW, JaK, LCN, and AM designed the experiments. LCN performed the experiments, data analysis and wrote the manuscript. JoK performed the RNA sequencing analysis. JFC, AT, and MHS constructed all plasmids. JaK, JMS, JoK, MW, AM, RSR, JMG, and BAS critically revised the manuscript. All authors read and approved the final manuscript.

Funding

LCN was funded by São Paulo Research Foundation (FAPESP) Research Internship Abroad (BEPE) Scholarship number 2018/02227-6 and RSR was funded by São Paulo Research Foundation (FAPESP) Young Investigator Award 2012/22921-8. This work conducted by the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under contract DE-AC02-05CH11231 with Lawrence Berkeley National Laboratory. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>). The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231 with Lawrence Berkeley National Laboratory.

Availability of data and materials

Strains generated and analyzed here are available in The Joint BioEnergy Institute Inventory of Composable Elements (JBEI-ICE) at <https://public-registry.jbei.org/folders/437>.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Ribeirão Preto Medical School, University of São Paulo, Av. Bandeirantes 3900, Ribeirão Preto, São Paulo CEP 14049-900, Brazil. ² DOE Joint BioEnergy Institute, 5885 Hollis St, Emeryville, CA 94608, USA. ³ Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720, USA. ⁴ Institut für Genetik, Technische Universität Braunschweig, 38106 Brunswick, Germany. ⁵ Chemical and Biological Process Development Group, Pacific Northwest National Laboratory, 902 Battelle Blvd, Richland, WA 99354, USA. ⁶ Department of Energy, Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA. ⁷ Department of Biomass Science and Conversion Technology, Sandia National Laboratories, 7011 East Ave, Livermore, CA 94550, USA. ⁸ Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720, USA. ⁹ QB3-Berkeley, University of California, Berkeley, CA 94720, USA.

Received: 1 April 2019 Accepted: 24 June 2019

Published online: 29 June 2019

References

1. Yaegashi J, Kirby J, Ito M, Sun J, Dutta T, Mirsiaghi M, et al. *Rhodospiridium toruloides*: a new platform organism for conversion of lignocellulose into terpene biofuels and bioproducts. *Biotechnol Biofuels*. 2017;10:1–13.
2. Zhang S, Ito M, Skerker JM, Arkin AP, Rao CV. Metabolic engineering of the oleaginous yeast *Rhodospiridium toruloides* IFO0880 for lipid overproduction during high-density fermentation. *Appl Microbiol Biotechnol*. 2016;100:9393–405.
3. Wang Y, Zhang S, Zhu Z, Shen H, Lin X, Jin X, et al. Systems analysis of phosphate-limitation-induced lipid accumulation by the oleaginous yeast *Rhodospiridium toruloides*. *Biotechnol Biofuels*. 2018. <https://doi.org/10.1186/s13068-018-1134-8>.
4. Yuan J, Ching CB. Mitochondrial acetyl-CoA utilization pathway for terpene production. *Metab Eng*. 2016;38:303–9. <https://doi.org/10.1016/j.ymben.2016.07.008>.
5. Campbell K, Xia J, Nielsen J. The impact of systems biology on bioprocessing. *Trends Biotechnol*. 2017;35:1156–68. <https://doi.org/10.1016/j.tibtech.2017.08.011>.
6. Park YK, Nicaud JM, Ledesma-Amaro R. The engineering potential of *Rhodospiridium toruloides* as a workhorse for biotechnological applications. *Trends Biotechnol*. 2017;36:304–17. <https://doi.org/10.1016/j.tibtech.2017.10.013>.
7. Wehrs M, Gladden JM, Liu Y, Platz L, Pahl J-P, Moon J, et al. Sustainable bioproduction of the blue pigment indigoidine: expanding the range of heterologous products in *R. toruloides* to include non-ribosomal peptides. *Green Chem*. 2019;21:3394–406.
8. Johns AMB, Love J, Aves SJ. Four inducible promoters for controlled gene expression in the oleaginous yeast *Rhodotorula toruloides*. *Front Microbiol*. 2016;7:1–12.
9. Nora LC, Westmann CA, Martins-Santana L, de Alves LF, Monteiro LMO, Guazzaroni ME, et al. The art of vector engineering: towards the construction of next-generation genetic tools. *Microb Biotechnol*. 2018;12(1):125–47.
10. Martins-Santana L, Nora LC, Sanches-Medeiros A, Lovate GL, Cassiano MHA, Silva-Rocha R. Systems and synthetic biology approaches to engineer fungi for fine chemical production. *Front Bioeng Biotechnol*. 2018;6:1–18. <https://doi.org/10.3389/fbioe.2018.00117/full>.
11. Liu Y, Mei C, Koh J, Ngoh ST, Ji L. Engineering an efficient and tight D-amino acid-inducible gene expression system in *Rhodospiridium/Rhodotorula* species. *Microb Cell Fact*. 2015;14:1–16.
12. Liu Y, Koh CMJ, Ngoh S Te, Ji L. Erratum: Engineering an efficient and tight D-amino acid-inducible gene expression system in *Rhodospiridium/Rhodotorula* species. [*Microb Cell Fact*. 2015; 14:170] <https://doi.org/10.1186/s12934-015-0357-7>. *Microb Cell Fact*. 2017;16:1–16.
13. Wang Y, Lin X, Zhang S, Sun W, Ma S, Zhao ZK. Cloning and evaluation of different constitutive promoters in the oleaginous yeast *Rhodospiridium toruloides*. *Yeast*. 2016;33:99–106.

14. Liu Y, Yap SA, Koh CMJ, Ji L. Developing a set of strong intronic promoters for robust metabolic engineering in oleaginous *Rhodotorula* (*Rhodospiridium*) yeast species. *Microb Cell Fact*. 2016;15:1–9.
15. Liu Y, Koh CMJ, Sun L, Hlaing MM, Du M, Peng N, et al. Characterization of glyceraldehyde-3-phosphate dehydrogenase gene *RtGPD1* and development of genetic transformation method by dominant selection in oleaginous yeast *Rhodospiridium toruloides*. *Appl Microbiol Biotechnol*. 2013;97:719–29.
16. Lin X, Wang Y, Zhang S, Zhu Z, Zhou YJ, Yang F, et al. Functional integration of multiple genes into the genome of the oleaginous yeast *Rhodospiridium toruloides*. *FEMS Yeast Res*. 2014;14:547–55.
17. Gietz RD. Yeast transformation by the LiAc/SS carrier DNA/PEG method. *Methods Mol Biol*. 2014;1163:33–44.
18. Mutalik VK, Guimaraes JC, Cambray G, Lam C, Christoffersen MJ, Mai Q, et al. Precise and reliable gene expression via standard transcription and translation initiation elements. *Nature Methods*. 2013;10(4):354–60.
19. Kosuri S, Goodman DB, Cambray G, Mutalik VK, Gao Y. Composability of regulatory sequences controlling transcription and translation in *Escherichia coli*. *Proc Natl Acad Sci USA*. 2013;110(34):14024–29.
20. Goodman DB, Church GM, Kosuri S. Causes and effects of N-terminal codon bias in bacterial genes. *Science*. 2013;342:475–9.
21. Vogl T, Kickenweiz T, Pitzer J, Sturmberger L, Weninger A, Biggs BW, et al. Engineered bidirectional promoters enable rapid multi-gene co-expression optimization. *Nat Commun*. 2018;9:3589. <https://doi.org/10.1038/s41467-018-05915-w>.
22. Xu Z, Camblong J, Wei W, Gagneur J, Perocchi F, Clauder-Münster S, et al. Bidirectional promoters generate pervasive transcription in yeast. *Nature*. 2009;457:1033.
23. Neil H, Malabat C, Aubenton-carafa Y, Xu Z, Steinmetz LM, Jacquier A. Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature*. 2009;457:1038–42. <https://doi.org/10.1038/nature07747>.
24. Shaul O. How introns enhance gene expression. *Int J Biochem Cell Biol*. 2017;91:145–55.
25. Hoshida H, Kondo M, Kobayashi T, Yarimizu T, Akada R. 5'-UTR introns enhance protein expression in the yeast *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol*. 2017;101(1):241–51.
26. Hwang HJ, Lee SY, Lee PC. Engineering and application of synthetic nar promoter for fine-tuning the expression of metabolic pathway genes in *Escherichia coli*. *Biotechnol Biofuels*. 2018;11:1–13. <https://doi.org/10.1186/s13068-018-1104-1>.
27. Venayak N, Anesiadis N, Cluett WR, Mahadevan R. Engineering metabolism through dynamic control. *Curr Opin Biotechnol*. 2015;34:142–52. <https://doi.org/10.1016/j.copbio.2014.12.022>.
28. Otoupal PB, Ito M, Arkin AP, Magnuson JK. Multiplexed CRISPR-Cas9 based genome editing of *Rhodospiridium toruloides*. *mSphere*. 2019;4(2):e00099-19.
29. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12:357. <https://doi.org/10.1038/nmeth.3317>.
30. Liao Y, Smyth GK, Shi W. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30:923–30.
31. Reider Apel A, Mukhopadhyay A, Keasling JD, Ouellet M, Szmids-Middleton H. Evolved hexose transporter enhances xylose uptake and glucose/xylose co-utilization in *Saccharomyces cerevisiae*. *Sci Rep*. 2016;6:1–10. <https://doi.org/10.1038/srep19512>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

