Supporting information

Synthetic core promoters and as universal parts for fine-tuning expression in different yeast species

Rui M. C. Portela[1], Thomas Vogl[2], Claudia Kniely[2], Jasmin E. Fischer[2], Rui Oliveira[1] and Anton Glieder[2]

[1] REQUIMTE/LAQV, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal
[2] Institute for Molecular Biotechnology, NAWI Graz University of Technology, Petersgasse 14/2, 8010 Graz, Austria

Supplementary Material:

S1 – Summary of literature references on *S. cerevisiae* and *P. pastoris* core promoters.

In this section we will succinctly describe the main studies developed to clarify the mechanisms of yeast promoters, focusing on *S. cerevisiae* and *P. pastoris*. The references of S1 and S2 follow the numbering of the main text.

S. cerevisiae

Sugihara *et al*. identified CRMs in TATA-less and TFIID-dependent core promoters, namely *RPS5* [32]. Park *et al.* developed a method for improved transcription start site (TSS) mapping, inferring relationships between core promoter *cis*-regulatory elements, chromatin features and TSS location [33]. A study on the effect of 5'UTR features in gene expression was presented by Dvir *et al.* [34]. It has been shown that yeast core promoters show a high level of conservation, maintaining their functionality even in distantly related species [27]. As illustrative example, the *S. cerevisiae LEU2* core promoter has been shown to remain functional when inserted in *P. pastoris* [8].

Different approaches have been followed to design synthetic promoters for protein expression fine-tuning. Some are focused on CRMs [37,38], while others target both CRMs and core promoters [7,9,18,19,39–42]. For CRMs design, the interaction between transcription factors (TFs) and respective binding sites (TFBSs) have been used to model transcription, either by creating a large library of promoters based on combinatorial arrangement of different TFBS upstream of the natural core promoter [38] or by generating orthogonal synthetic zinc fingers used to wire new synthetic transcriptional cascades [37]. Also, the inclusion of regulatory sequences next to the core promoter can be used to fine-tune transcription. Other design approaches consisted in adding random mutations to a natural promoter (*TEF*) [7,9], randomizing two specific areas of the *PFY1* promoter and changing its expression profile afterwards by adding Tn10 Tet operator sites [41] or generating a minimal promoter based on a large scale screening of random sequences for minimal length, robustness and modulatory [19] has been used for the same purpose.

P. pastoris

Promoter libraries have been generated mostly based on deletions of CRMs sections to research the P_{AOX1} regulatory mechanisms and to determine TFBSs (*e. g.* [24,26]). Promoter libraries have also been created to control gene expression by modifying the core promoter [16,25,43] and 5'UTR [44] sequeces. Berg *et al.* studied random mutagenesis of P_{AOX1} [16], located in both core promoter and CRM regions. They observed expression profile modifications (derepression) when mutating some specific nucleotides in the P_{AOX1} CRM, and modifications in the expression rate when mutating the core promoter sequence. Following a different approach, *P. pastoris* synthetic core promoters have also been designed based on four natural *P. pastoris* core promoters consensus sequence through the addition of some natural TFBSs [25].

S2 – Detailed computational design of synthetic core promoters

Given that the core promoter design method was based on features from a genome wide list of *S. cerevisiae* natural core promoter sequences [28] it can be divided into three parts:

a) Computation of features from the *S. cerevisiae* data set to be used for the synthetic core promoter design;

b) Generation of core promoters sequences based on calculated data;

c) Design space reduction – selection of core promoter sequences to be tested *in vivo*.

As mentioned in the main text, several features were simultaneously incorporated in the design given that they were found to be correlated with maximal promoter activity [28]. The features included in the design process were: *i*) nucleotide occurrence along the sequence of 140 strong natural *S. cerevisiae* core promoters (as reported by [28]), *ii*) the presence and position of the TATA box, *iii*) the position and number other motifs (other than TATA box, as defined by [28]) and *iv*) nucleosome occupancy profiles [28,45].

As described below, all the referred information was calculated in the first design part (*a*). However, the nucleotide occurrence, TATA box position and motif position and frequency was included in the core promoter design (*b*), while the nucleosome occupancy profile was included indirectly in the design process as a selection step (*c*).

It should be highlighted that only some of the motifs described by Lubliner *et al.* were added in this design process. The selection criteria were: strong reported correlation with maximal promoter activity and motif position within the desired core promoter region (from start codon to 150 bp upstream of it given that our core promoters' target length was 150 bp). The list of selected motifs and respective location is provided in Supplementary Tables 7-10.

a) Computation of features from the S. cerevisiae data set to be used for the core promoter design

Firstly, from the whole 729 native *S. cerevisiae* promoters' data set we focused on the 140 strong core promoters and respective 5'UTR. The sequences were trimmed to have a final length of 150 bp (corresponding to 50 bp downstream and 100 bp upstream the transcriptional start site (TSS)). From this subset we computed the:

1. Nucleotide probability distribution along the core promoter sequence – The frequency of each nucleotide was computed separately for consecutive promoter regions, in a sliding windows manner (windows size of 20 bp and windows step of 10 bp). The probability was calculated for each nucleotide and promoter region (frequency of each nucleotide for each promoter region was divided by the windows size). This resulted in a matrix of $n \times w$, with n the number of nucleotides (4) and w the number of windows (14). The sum of these probabilities, column wise, was 1;

2. TATA box position distribution along the sequence – Considering the TATA box consensus sequence (TATAWAWR), all the occurrences location of this motif were annotated. A

Gaussian distribution model was inferred from this set of TATA box locations using the respective average (μ^T) and standard deviation (σ^T);

3. **Position and frequency distribution of motifs along the sequence** – For each of the selected motifs listed on Supplementary Tables 7-10 a similar approach, as compared to the previous step, was used: annotation of number and positions of motif occurrences (respectively, f^{Mi} and p^{Mi}, with i=1,2,..,7). For each set of frequency and positions a Gaussian distribution model was inferred (described by the respective average (μ^{fMi} and μ^{pMi}) and standard deviation (σ^{fMi} and σ^{pMi}));

4. **Average nucleosome occupancy along the promoter sequence** – The last step pre-design computation was the natural nucleosome occupancy average profile. For this step a software package by [45] was used. With it, for each 140 natural core promoters a nucleosome profile were calculated. To avoid sequence edge related error, a 1000bp sequence (derived from the original cloning plasmid) was added to each side of the promoter sequences. The average nucleosome profile (μ^{Nj}) was calculated using the obtained 140 occupancy profiles (each profile consisted of 150 occupancy scores related with each nucleotide, j=1, 2, …, 150).

b) Generation of core promoters sequences based on calculated data

As mentioned in the main text, 4 different groups (named P, T, M, A) were designed useing the previously calculated information. They differ in the presence or absence of a TATA box and/or selected motifs (group P: without TATA box nor motifs; group T: with TATA box and without motifs; group M: with motifs and without TATA box; group A: with TATA box and motifs).

The sequence generation was computed as follows:

1. **Random sequence generation** – 400 sequences, of 150 bp each, were generated with the MATLAB function *randseq*. This function had as input the vector of nucleotide probability w^l (l=1, 2, …, 14) and the sequence length (equal to the window size – 20). Thus, the *randseq* function was used 14 times to generate each sequence;

2. **Removal of randomly occurring motifs (TATA box and selected motifs)** – TATA boxes and any of the selected motifs were searched and replaced by a newly generated sequence. This procedure was repeated until no motif or TATA-box were found in the generated sequences;

3. **Removal of randomly occurring start codons** – Following the previous step approach Start codons upstream of the protein codon region were also removed to avoid frame shift mutations or different N-termini of the reporter protein;

4. **Add Kozak sequence upstream of start codon** – Due to the known relevance of the nucleotides adjacent to the start codon [34], this region was replaced by the P_{AOX1} Kozak sequence (CGAAACG) in the generated sequences;

5. **Separation of sequences in 4 groups** – The 400 sequences were divided in 4 groups of 100 sequences each. The group P had no further modifications as it is characterized by not having TATA box or any other motifs;

6. **Addition of a TATA box to groups T and A** – For each sequence belonging to these groups, a TATA box position was generated (*randn* MATLAB function). The TATA box Gaussian

distribution was taken into account by multiplying the generated number with σ^T and summing μ^T. One TATA box was inserted per core promoter sequence. The sequence originally located in this region was replaced by the TATA box. The sequences in group T had no further modifications as this group is characterized by having a TATA box and not having any other motifs;

7. Addition of motifs to groups M and A – In a similar way as in the previous step, the number of motifs and respective position was generated with the *randn* MATLAB function together with the respective average (μ^{fMi} and μ^{pMi}) and standard deviation (σ^{fMi} and σ^{pMi}). Thus, the frequency of each motif in each sequence also followed a Gaussian distribution model inferred from the natural sequences, meaning that some motifs might be present more than once while others might be absent in a given sequence.

c) Design space reduction – selection of core promoter sequences to be tested in vivo

From the 100 sequences in each group, 28 were selected for experimental screening. For each of the 100 designed sequences a nucleosome occupancy profile was calculated (as described in a-4). Using the calculated profiles, the objective function that was used to select the 28 sequences was:

$$min\left(\sum_{j=1}^{150}\left(\mu^{Nl}-\mu^{sj}\right)^2\right) \qquad \text{Eq. 1}$$

Where μ^{Nj} is the average nucleosome occupancy profile for natural core promoter sequences, μ^{sj} is the nucleosome occupancy profile for each *s* (*s*=1, 2, …, 100) synthetic core promoter sequence along its *j* nucleotide position (*j*). The 28 sequences with a lower sum of squared errors were selected. With it we aimed to select for screening the designed sequences that were more similar to the natural promoters concerning the predicted nucleosome average occupancy.

Supplementary Table 1 – List of primers used to clone the positive and negative controls

Name	Sequence
C-WO-CRM1	TATTGTGAAATAGACGCAGATCGGGAACACTGAAAAATACACAGTTATTATTCATTTAAATGACAGCAATATATAAACAGAAGGAAGCTG CCCTGTCTTA
eGFP-pAOX1-3prime	AAAAGTTCTTCTCCTTTGCTAGCCATCGTTTCGAATAATTAGTTGTTTTTTGATCTTCTC
pAOX1_Syn_dBamHI_SwaI-forward	GATCGGGAACACTGAAAAATACACAGTTATTATTCATTTAAATAGATCTAACATCCAAAGACGAAAGGTTGAATGAAAC
C-WO-Core1	GTTCTTCTCCTTTGCTAGCCATAAGTAGGGGTTAGAACAGTTAAATTTTGATCATG
C-W-HHF2+10	GTTCTTCTCCTTTGCTAGCCATATTTATTGATTATTTGTTTATGGGTGAGTCTAGAAAAGGACGCACTCGTCTTGTATTTATAGATGAAAA GAAAGTAGGGGTTAGAACAGTTAAATTTTGATC
R1	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGAGTCTGGGGCGGTCTTATGCTAATCGTTGGTGTCCTGTCCTAGAGAGTTGCGAACAGGC AAGGCGCCGGAGATAGAGTATCGTGAAGGAATTTATAGAGGGTCTATCGCCGCAGTCGTCGACGGGCAAACAAGCACCGGGCAAGTA GGGGTTAGAACAGTTAAATTTTG
R2	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAACACGTATGGGCCAAGTGTTTGAAATTCAACCTAGTTTTGAGGGTTACGATACGGACCTC CCCTACGCTGCAGTCCTATCAGCGAGGCGCAGCAAGCGTTAGGCGTGCGGCTTCCAGACTCAAGGAGACCTTGGCGATATGAAGTAG GGGTTAGAACAGTTAAATTTTG
R3	GTTCTTCTCCTTTGCTAGCCATCGTTTCGACGCTTTCGAAGAGGATGGCTACTCCCGGCTAGTAAGTTGGCTTGTTTGCTTGACGGATG ACTTAGGGTTTTTAATGAAGGCCGTTGTGTCTACGGACACGACACGGGTGTTCTACTCGCGCCTCTTGGGAGCACTTTAATAAAGTAGG GGTTAGAACAGTTAAATTTTG
R4	GTTCTTCTCCTTTGCTAGCCATCGTTTCGACGTTAGAGATTGGCTTCCTAAGATCGAGTAGTGATTCCGTATATAGGCTCGCTTACCCAG ATCCCCTAGACTGTTCGCTTTGCTATGAGTCTCAACTAACCCTAAATGTCCGTGCCCCGTTTCTATCGTATATTAGGGGACAAGTAGGGG TTAGAACAGTTAAATTTTG
R5	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAAATGTAGCTTTGTCCGTAGAATACCCGGTATAAGACAAGAGCAGTCTAGTAGGAGAGCTC CTTTGACCTGCCGGTTTCTGGGAAGGGCCCAGGAAACGGGTTACGGTTCTTACGACACGAATGCGTGTTTCGTGACTGTTTAAGTAGG GGTTAGAACAGTTAAATTTTG
R6	GTTCTTCTCCTTTGCTAGCCATCGTTTCGATCAACCGGCAATTCAGAAATGCGGGATTTAGTACACCTATTTCTTTACCAACTCCCGCCG CGTTTTAGCTTAATGATGAGCGGTGGGCGTGTTTTGAAAAAAGATGTTAGAGATATATTCTAGTCAGAGGGTTCGACTAGAAAGTAGGG GTTAGAACAGTTAAATTTTG
R7	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTGCTTACCGGTGGATTACGTACGGGGTTGGATGAGAAGTGGGATGGCCCAACACCAATTG ATTGTCAACTCCGACCTGAAGGCTTCACAAGAGTGGAGGGTACAGCTAGTACTGAGTGTGTTAATTGGAGTACGGTCGGCTTAAGTAGG GGTTAGAACAGTTAAATTTTG

Supplementary Table 2 – List of primers used to clone the synthetic promoters of group P

Name	Sequence
P1	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCGAGCACTTCTAGTTTTCGAAGATACAGTTCCACAATAGGTTCTTTCTATCGTCAGTATTCTCGTTC GCAAGCAAATAATTTCGGGATTATAGCGGAGTTTACAAACTAAACAAGATGTGCCTACTAGGCTATTCTAGACTAAAAGTAGGGGTTAGAACAG TTAAATTTTG
P2	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTCGTGAAATTACTAAGGGTTTAACTGCACTCGCGTATTTCAGAAGACTACTCTCGTAGGTTAATGC AACTCTACAAGTGATGACTTTGCTATGAACTACTTGTACTACTTACAATCTGTCTAGCAAACTGCGTCCGAAGTCCTAAGTAGGGGTTAGAACAG TTAAATTTTG
P3	GTTCTTCTCCTTTGCTAGCCATCGTTTCGATAATTTCAACCCTTCGTTCCGAAAACTAGAACGTGAGTTTCTGTTTGGCTATTCCGCGTATATCG TTTCTGAATTATTGAATAGAGCACAATATCAAAATACTAAAAATCGAGTTATTGGGATCGTACCAATACGTGGTTTAAAGTAGGGGTTAGAACAG TTAAATTTTG
P4	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGTTTTCGCTTTAATCTGTTAGGCAATATCCCTATTACGCAAACAGAGACTAACACCAACGACCAAG ACTTATATTTTGTTTGCGCACTACTAGCTAGGAACGTAGATATCAGTTACAAATATAATTCACTACGAGTTATCCGGAAGTAGGGGTTAGAACAG TTAAATTTTG
P5	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAAGACTCAGTTACGACGTAGAGAGGATGGATAGGCCGTTCGTCAACAGAGCGATCTAATTGTTTC GTTACTGATGGAATTGTTGGATAGTGTAAATCTAATAATGGAATTATATTTCAATTACTTGTTTGGGTAAAACGCCTGAAGTAGGGGTTAGAACA GTTAAATTTTG
P6	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCTAGAACTAGGTTCTTCTCCTACTGTCTAAATATCTCTATATTTTAAAGTGAAATTTGGAGTGGTCG TTATAATTACTCGTTTTAGTGCACCCTAGTCGGGGGTCTAAAATTACAGTATACAAGTAAAGTTGTGATGACTCCAAAGTAGGGGTTAGAACAGT TAAATTTTG
P7	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCTAAACTGTTCCTATCTTCTTATGTTTAGTAATGTAGCGTGAAATAATGTCAGACGATTATCTACTA CGAGACTACACTACGATACGTACTAACGAGGAGTGACTTGGGGGTACCGTATAGTTGTAATCTACCTACTTCCGCCAAGTAGGGGTTAGAACA GTTAAATTTTG
P8	GTTCTTCTCCTTTGCTAGCCATCGTTTCGATATGATCGTATGGGCAAACTTCACTCTGTTTCTAATTATAATTTAGCTTCGGATCGTATGAGGGT GGACACCTCGGTTGACTTGACTACGGTTCTAATGAACTTTTAATAAATCGTACCCACCTAATTAGAGAAGTATATAGAAGTAGGGGTTAGAACA GTTAAATTTTG

Supplementary Table 2 (cont.) – List of primers used to clone the synthetic promoters of group P

Name	Sequence
P9	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGATTCTAATGATATCTTCCACGACTGTAGAGCAACGGTTAGCAAACTACTATGTAGATGTTTTAGA TTGTGATTTAGATGCAAACTATGTTCCTTATTTTAACAACAATAGTGCAACTATATTGGAACTACCTGCAGAAAGCAAAGTAGGGGTTAGAACAG TTAAATTTTG
P10	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAAAGAACACGACGTCTATGAACTAAGGTTCAGTCTAAATACTAAATAAATGAACTTGTATCTATTTT TTTGCCTGATATAAGTTGCGTTGGGAAGACTAATTATGAAGATGTTCAAGATAAGATGAATTGAATAACTAAAAAAAAGTAGGGGTTAGAACAGT TAAATTTTG
P11	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTTCTGTACTTAACAGTAGCACTACAATTTCTTAAGTACTAATTTACTTTTATTCCTACTACTAAGTGG TTTTCAGCTATGGTGGTTCAAATATTAGGTAGCCTAGTATCCCACGTACGTAATGAGACAAAAACTAATAATGCAAAGTAGGGGTTAGAACAGTT AAATTTTG
P12	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAGTTTTATTAAAGAATTTAGCTCTAGCTACAATCGATTTATGCAATCGTGCTAGACTGGAATAACTT GTAGCTACGCGTATGGCTTCGTATTGGGGAGAAGTAGTTAACACACGACTATGGATATTATGGTAAATAGTCAATAAAGTAGGGGTTAGAACAG TTAAATTTTG
P13	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTTAATCACGGGTGAAAATTAAGAGAACTTTTAACTTAATGAGACTAGGGGAATAAACTTTGAATTG GTTCTCGTACGTATGCGGTAACTCGTGTATTTGCCCTATGAGTAATAGGTAGAATCAAGAATGTACTACTAATATGGAAGTAGGGGTTAGAACA GTTAAATTTTG
P14	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTATAACTTTAATTTCTTTCAATTTTAAAAAATATAAAACGGTAACTAAAGGTATTTTTCGCGTTAACC AATAACTGATTTTAAGTATATCTGCAGAGTAAGGGTTGATGAAGCAGGTAGCTATTTGAGTAGAATCGTACAATGAAGTAGGGGTTAGAACAGT TAAATTTTG
P15	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGTATTATCTATAACCACTATACTTTTCTATTACTAATTTTCAAGCAATACGTTCGTGTTTGCGTAAGT ATACCCTGGTATCTGTGACTCAAGCTAAGTAACGAGATTATTGACCTACCTTTGGGTGTATCAAGTCTAACAAAGAAGTAGGGGTTAGAACAGT TAAATTTTG
P16	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGTATTAAGAATAAAGTTCTCAGAAGGGTTTAAATGCGAGCTTAATTTGGGATGCTTAGATGTTATC CTTATCTAAATTCACTAATACACAGATAGTTCAGTTAATGAGCAGAATTTTAGTGTGACAGAATCTGTGATGTGTCAAAGTAGGGGTTAGAACAG TTAAATTTTG

Supplementary Table 2 (cont.) – List of primers used to clone the synthetic promoters of group P

Name	Sequence
P17	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAACACTCTCTAATGATACGCGAGCTAATTAGAAGAGAAGCTAGCTTATGAAGTTTATCGGTTGCTC CCACTATCAAACTATAAATAAGTGGAAAAAACTTCTCGTGTCGTTGTTGAACAACTAAAGTCCTATTATCGTGTCCAAAGTAGGGGTTAGAACAG TTAAATTTTG
P18	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAATAGATACAAAATCGTTAACTATAGGTAGTGCTAGTACTTTCAGATACCTTTTTGGTAGGCTAATT TATCTATCAATATATTAAAGTACGGTCTCCCTCGTTGAATGATAACTCAGTACCTAACTAACACTTTAATTAAGGAAAGTAGGGGTTAGAACAGTT AAATTTTG
P19	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAAGCTCGGTAATTCTAGAGTTTTAATTCGTTCTGTTAATTACAATGAAATCTACGTTCTAAATTTTTT CAATAAGTTTCTAATCAATCACGGGCAATTACAAGGATTGAAGTAGTCTACCTTTGTGTTCTAGTGTGGACAGTCAAGTAGGGGTTAGAACAGTT AAATTTTG
P20	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAAATCGACTTCTATTACTAATAAAGTCTAAGACTACGTTATTCAGATGCTACCTTTGAGAGTTTATA TCTACTACTATAACTGTAGTCACACGAACCTAGAATTCAGTTTCCACGGTTTATAATTACTCTATCTAAAGGGATTAAGTAGGGGTTAGAACAGT TAAATTTTG
P21	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTTCAAGGGTCACCTACAATAAGTTCTTTAGAGTATGTATATTAAACTCTATGTTGCAAGATATGGGT AACTATCAGATCGGTAAATCGTCGGTTCTAATACTATGAACTAAAAAGGTCTAACCTATGATCCTATATTCTCCTAAAGTAGGGGTTAGAACAGT TAAATTTTG
P22	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCTTAATCTCGATAATATACTTTAATTACTAAAGGAGAACGATTCACAAACTAAGCAACTACGATATT GTCAACTATGCAAATTTTGATGGGTAGAGTAGTCTAGATTGTTATATTCAACTGCGAAAAAGATAGCGGAAACTTGAAGTAGGGGTTAGAACAG TTAAATTTTG
P23	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTCAGTTCTTCTAGAGGTATCTGGTTTTCTAGATTATTCGTTTTTTATTCTACGGAATTTGGAATTTG GTGTTTGAAGGTGTTACCTCTGCGACAGATTTTGTTATTCTAAAATAGTACCTCCCGAGTGAGCAAACAATTGACGAAGTAGGGGTTAGAACAG TTAAATTTTG
P24	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTCAAAATTCAACTAATTACAATTTAAGATATTAAAAAATCTTTAGCGCTATGGGGTGAGCTAGTTAA TGGCAAGACAAGATAGCTAAATTTAAGTTCCGTATATGTCTACTAGTGTTTCGTGGACTCAAATTAAGAAAGATGAAAGTAGGGGTTAGAACAGT TAAATTTTG

Supplementary Table 2 (cont.) – List of primers used to clone the synthetic promoters of group P

Name	Sequence
P25	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAATTAATTATTTACTTCTAGTCAGCACTAATGAGGTTTCAGCTTTTATTGAACTTCTGTCAGATTAGT ATACTAGCTACAACTATAAATCTGCGAAGCTAAGTTTAGGAAATAATATCTGTATTTATGCTCCCGATTCGAGTCAAGTAGGGGTTAGAACAGTT AAATTTTG
P26	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTGATTTATGCGTTATGGATATGTTAAGGCAACGATAATTATTAAAGGAATAGTCGTTAGAGCCGTG TGAATTCTACTACGTATTATTCTAAAGAGTCACTACTGATGTCCTTATCTATTAGTGATATATTTCGACCAGACGAAAAGTAGGGGTTAGAACAGT TAAATTTTG
P27	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGTCTAGCTAAACCTTCTACAATGTGAATTATTCAAACGTAGATCGTAGGATTCTAAGGTTCGTGGA CAGTAGTTGTTTATAGGGGGGCTCTAGAGAGTTTGATTAGCGATATTTAGAGACCAAATTTCACCTGATAGCCTAGCAAGTAGGGGTTAGAACA GTTAAATTTTG
P28	GTTCTTCTCCTTTGCTAGCCATCGTTTCGACTAAGTTCACACGGGCTAATATTTAAGATAAACTATAGGATGCTAATATTTACGTATAATGGAAG GTCCTGTAATATCCTCCAGATGGATTTGTTAACTATTTTATACTGATCGTAAAGTGATTAATGTTGAATCCCCTAGCAAGTAGGGGTTAGAACAG TTAAATTTTG

Supplementary Table 3 – List of primers used to clone the synthetic promoters of group M

Name	Sequence
M1	GTTCTTCTCCTTTGCTAGCCATCGTTTCGATCTAATATTCAGAACTATATCGTTTAGCGGGATGGGCAAGTGCCGCCCCTATTTTTAAAATGAAT AACTACTAGATTTCACACGGGTTTGTTTTGTTGATATGTTATTACAATCTAGCTCAATGATTATACTTGATCTCTTTAAGTAGGGGTTAGAACAGT TAAATTTTG
M2	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGTAATATTCAAGGTGGGTTGATGCACCCTAAGGTGGCTATTCTTTTTGCCTGGTGTTTACTGCTAT TTCTGTGGCACTTTAGAACAGTTCCTTCGACAGTCTTGTTTTCCGCTAAGGAGAGTAATGGACGGATTACCGAAGCGAAGTAGGGGTTAGAACA GTTAAATTTTG
M3	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGACTCACTTGGAGCTTGTTAAGTAATCAGGGCGTCGGTTGTTGAAAGTTTGATTGGTTACGGTTCT TTCGATTTTCGGCCTTTAGTCTACTACTAGTCTCCTTTAACCGCTAGTCTGCACTAGTACACCCAATCTCTAATTCCAAGTAGGGGTTAGAACAG TTAAATTTTG

Supplementary Table 3 (cont.) – List of primers used to clone the synthetic promoters of group M

Name	Sequence
M4	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTTTGTCTATAAAGAGATTTCGAGGACACTACGCTAGCAGATTGTGAGATTAATCGTTTTGAGCAAG TTATCAAAGAAATTTCACTGCTGGGCTTTTCTTGGCCACTCTCACTATCTTTACTGATCTCGTACTACTAGGTAAACAAGTAGGGGTTAGAACAG TTAAATTTTG
M5	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTGTATTGTGGTCCTTATACAGGACGTTGTATCACCTGAGGTTCTTTTTCTAGCTACACCAAAAGAT TATTAATACCTAAACTTAGTGAGATAAGTTATGATGTTATAACTAGTTATGTCAAGACGGGCTAACTCCAATAGACAAAGTAGGGGTTAGAACAG TTAAATTTTG
M6	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCTATTAACAATGCGCGTAGATGATTAGATGCTTAACCTTATACTAAGAGTTGATGACGCGGCCGTT GCTTTTTCAAGATCTTAAGTTTTTCAGATCTTTGCTTCAAATCGCTAACTATTAATAATACGCCCTAGAAAAATCGTAAGTAGGGGTTAGAACAGT TAAATTTTG
M7	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTAGTTTCGATATGGGTAAAAGGGCTATATGAGAGGGTACTCAGTGTCCTGGAAAAATTTTTGTTGG TAAGTTCGAATCTATAGTGTTAAGCTAGGCTGTTCTATTGCTAATAGTCCGTCTTTGCGTCTTCAACGTATTTGGTCAAGTAGGGGTTAGAACAG TTAAATTTTG
M8	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCGTTAGGGTCTATTTTTATGAGGACAACAGCGCTCTAGTATACCTTTCGTAGGGCCGGAAACTAT CTAAGTGCCGCTATCGACTAGAAGCTTATTATCCCCAAGATCAAATATATTGTTGAAAAGGATTATCTCAACGGCTGCAAGTAGGGGTTAGAAC AGTTAAATTTTG
M9	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTTAACTTAACAAAATACGTAGATGATTCCTCTTTTAACTAAAGGAATGACTATTAACTCGATAGCTC CTAGAGAACGTACCAGATTTTGGTGGTTTTCTTTTTTTGTCCTACTTTCTGATTACTACTATACAAGAAGTTTAAGAAGTAGGGGTTAGAACAGTT AAATTTTG
M10	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAGTAAAATTATCTTCCTTCAGTTGGCAAGGTCCTCCACACGGATACTTTATCCTAATAGAGTTGCG ACAACTACTATGAACTATCCTTAGGTAAAAGCGGCCCCAAAAATAAACTGTACCTTGTACAACTTACTAAAGTACGTAAGTAGGGGTTAGAACAG TTAAATTTTG
M11	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGACTGAGGTCCTTCGTATTTAGGTGTATGCTAATGTAACCTCTATCCTATTCGAGTCACAGTGGCT CCCAAGTAAATGTCCACACTTAATGAACGCTACTCAATTATAACCACTATGTTAGCCTTAAAATGGCTACTCAAATAAAGTAGGGGTTAGAACAG TTAAATTTTG

Supplementary Table 3 (cont.) – List of primers used to clone the synthetic promoters of group M

Name	Sequence
M12	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAACTGGGGTCACTCAAAATTGAGACCCTAGACTTTAAGTATCCTACAGCTAGTACCTGCACCGCT ACCTCAAAACTTTGAACGTTGAAATCGATTGCAACGAACTTGTAACGATCCTGTTAGGAAGCTAAGTGTATAGTGAATAAGTAGGGGTTAGAAC AGTTAAATTTTG
M13	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCCGATCAAAGATTTTTCTAGATAGTTATTCACTACGGTTTTATTCTATACGAACTAAGCTTCACGGT AGTGTACTATTCGCTATGTGGCTACTTTGACTCCGCTGGATTCTATCGCTTTTAGCAATATATAATGAAGTTATTTAAGTAGGGGTTAGAACAGT TAAATTTTG
M14	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGGAACCGAGCTCTATTGTGAATACTGCTGCCGGTACGCTCGGTCTGATTAGAATCTCTATAGATA TGGAGCCGACTCGCGTGTTTGGACGATGTATATCTATTTGAACCCCAAAATTTAATCGCTACAATCCTCGAAAAAAATAAGTAGGGGTTAGAAC AGTTAAATTTTG
M15	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAATTTCAATACTTGAAATAGTGAGACTTATCTAACTATTAGTAAGTATGGAAGAGAACTGAAGAAAG AGACTATTATAAGAATTCAATACCTCTTATTTTTGAAAGCTAGATCTGAGTTACTGGACTTTCCTCGACACTACAGAAGTAGGGGTTAGAACAGT TAAATTTTG
M16	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAATAACCTGTGATATTACACTCTCTTATGATCTTTCTATCAGGTTCCTTATGCTTCTGGCTATTCGA TAATTCCTAGTCTGCTCACAAAACCAGATTCTATCTTTCCGTACTATCTTCTTGTTAAGAAATCCCTGTCTAGAGAAAGTAGGGGTTAGAACAGT TAAATTTTG
M17	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTCCTGTTTTGGTTGCGAGATTACCCCTTTCTAAGTTTTCTCTACGTTATGGGGCTCCAACTAGCTA ATCGGTTGTTCGTAACTGGTCTGTTCGGCGGTTAGTGAGATACCAGGTGTATTTGGGTTATTGTACGAACAAATATTAAGTAGGGGTTAGAACA GTTAAATTTTG
M18	GTTCTTCTCCTTTGCTAGCCATCGTTTCGATAAGGTTGACTTTGGACTCGCGTATTCTTGCTATCGAAGTTTGATAGAGTGATCGTCTCTATCTA TTTTTGGGAGTACTGTGCAACTGAGTATTGCAAAATAACTTGATTGTTTATGAGATATGCTAGGTGTATGAAACCCTAAGTAGGGGTTAGAACAG TTAAATTTTG
M19	GTTCTTCTCCTTTGCTAGCCATCGTTTCGATAAAATTTTGCCGGGCTGTACGTCTTATCTAGATGTCGTTAAAACCTCAGGCCAAGCTCTATATA CTGCAACAAACCGCTAGCAAAGAAATTTAATATACTACACTCTTAAAAATGTATAGAGTTATTTTACTCAATTAACTAAGTAGGGGTTAGAACAGT TAAATTTTG

Supplementary Table 3 (cont.) – List of primers used to clone the synthetic promoters of group M

Name	Sequence
M20	GTTCTTCTCCTTTGCTAGCCATCGTTTCGATTGTATCTACTTTTCCCACTATTCAGCGTAAACAGACTAATGCTCCTATCTCTAGAGCTTGGTAGA TTAGTACGTGAAGGTATTAAAAATCTTTTTGTTTCCAATTGAAGGAGAGTTTACCTAACCTTACTGTAAGAGTGTCAAGTAGGGGTTAGAACAGT TAAATTTTG
M21	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCAAGGCAACTCGAGTATAAGTCTATTTGTATGCCGTTGATTCAAGAGAGTTCTGTGCCGTTAAAAA TTAGAATGTAATTAGAAGTAGCAATTCAGATACGATTGAATGGCCAATATCTGAAATTTAAGGTAGGGACTAACAACAAGTAGGGGTTAGAACA GTTAAATTTTG
M22	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTTACTTTCTTAATCTAAACAAGGATTCTATTCTTTGCGCCTAAAGGTTTAGGAAGTTTTCTGTTTGC CGCGTAGTTGATTTCACAAGAACAGTGAAGTATGGCTCGATCTACTAAAATTGAAAGCTAAACGTGGGATAGGTAAAAGTAGGGGTTAGAACAG TTAAATTTTG
M23	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAAAAGTCAATATATTGGAAAATTGCTCAATTAAATTGATGTACCCTGTAAGTGGTAAACTTTGTAAA CGTAGTTCACTCAGAATATATCGAGCCACGACTTTAGAAAATCCTTATCTACTTACGAAACTTAGAGTTCTCGAATAAGTAGGGGTTAGAACAGT TAAATTTTG
M24	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTGCCTTTTTTTTGAGACTGTTATTTAGTGATACACTAGTGTTAGGTATAATCGTGATTCTTAGATTG TTATTTAGATATTTCTATAGACGGCTAACTTTTTACCAACTTAAACTTACAGTATATACGCTCTCTATAATTGGCTAAGTAGGGGTTAGAACAGTT AAATTTTG
M25	GTTCTTCTCCTTTGCTAGCCATCGTTTCGATCTAGTAATTTATAAAACTACCTTGAGTGAGTGATCGATATTGTGATTCTGATCACGATGATCTCA CCTCGATATTGGATAGATGACTCCTTAAGTGCCAAGAACCTAACTAAACTTGATTGGTTAATTATTACGGGTGATAAAGTAGGGGTTAGAACAGT TAAATTTTG
M26	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAATATGTGTCCCTATAAGTGGTCCTACTCACGAAGGATGAATCAAATGAGTATGGATTCAAGGAAG GTAACTATGGATCTACTTGTGTGAAGGAACTATTTATCTTTTATGTGAGATAGGTGCCTCTAGTCAAATATTAACCTAAGTAGGGGTTAGAACAG TTAAATTTTG
M27	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGACTAAAGACCTTTATAAGGGACTGCGTGGTAAATAATGCTATTAAACTCTTTTTGATATTAGGTAT AGTACCTTATCGCTCAACAGTTTTACTACTTCTGCCGCCTCTGATTACTGATCGACAATATTATAGTATTATAAATAAGTAGGGGTTAGAACAGTT AAATTTTG
M28	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTATAAATAATGTAATAACTTTAAGATATGGATTGCACTTGTAACTTAAGCAATTGAAGTTTGCTAAC TAGTTATGTATCTTTTTCTACCTTGGCCGTCCTATTAAATATCTACAAGAGAGATCACAATTGACCGAAAGATTACAAGTAGGGGTTAGAACAGT TAAATTTTG

Supplementary Table 4 – List of primers used to clone the synthetic promoters of group T

Name	Sequence
T1	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCCTATATACGTAGCTACGAGTTAATGACTACAAATATCGCTCTACTCGTAGATGGAGGATAAGGAACACAAAGGAGCTCGGTTCTATACAAATTCGTTTAGTATTGATTTTATTTATATCAATGTCGTAACGTCGTGTGAAATAAGTAGGGGTTAGAACAGTTAAATTTTG
T2	GTTCTTCTCCTTTGCTAGCCATCGTTTCGATTACTAATCGAAATTACGTCACGTAGTAAGTCAATGTATTAATTACAGAGTATCCTTAATTATAAGATTGCGTCAGCGAATCGTGTATCTATTCCTATTTACTGTGGATAACGTGAACTTATATATATCTGGTTAAAGCGCCAAGTAGGGGTTAGAACAGTTAAATTTTG
T3	GTTCTTCTCCTTTGCTAGCCATCGTTTCGACTATACGATCCACTTCTACTTTCTAGGTATGAAGATGTATGTTAGATCTCGTTTTGTTAGTCGTTAGCCGTGCAATACGTTACTTGACCCTGATACTTATATAGACTATTATACTTTTACTTGTGGTGTTTCTATCAATTTAAAGTAGGGGTTAGAACAGTTAAATTTTG
T4	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGAGGCTACGCGATCGATAGGTTATTATAACCACACACTCCAATATATTTATACTGTCCTCGTCTAACTTTAGTCTCTAAACGTTCTCGTGTGATACTTATCCTAGTTAATATGGTTACTACGTAAAACTTGCAGTATCCCCCGAAGTAGGGGTTAGAACAGTTAAATTTTG
T5	GTTCTTCTCCTTTGCTAGCCATCGTTTCGACTAGCTATTATGGGTAAATCTGAACAACTGTAGTAACTGGAAAATCGTTGTTTTTATAAACTACACTGCTAGAGCTATCCCTCTGAGTTAAGAATTCGTTGAAGTCTATCGTCCTACTACACGTAATTCGTTCTATTCACTAAAGTAGGGGTTAGAACAGTTAAATTTTG
T6	GTTCTTCTCCTTTGCTAGCCATCGTTTCGACTAATTAGGGGAAGCGTTTCTTACAGCTACTGAATCTAGTGCGTTACCTCTATACGTATAAGTACTGTGAAACCAATGCTATCTACTATATTCGTTAACTTTTTATATATTTAATGTTTTTTTTAATAACACTGGCTATTTAAAGTAGGGGTTAGAACAGTTAAATTTTG
T7	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGGGTCTAGGAAGGAGCGAGTTCTACAAACACGATGGGGGGATGAAGCTCCTATTATTTAATGTAGATCTAGAGTAATCGTATTTATATTTTAATTATGGGGGTGATCAAGGATACTAGACTTACAATCCTCTAATCTGACTGAAAGTAGGGGTTAGAACAGTTAAATTTTG
T8	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAGACTAATCAGATAATGAGTACGGGAGAAGAGAAAATCTAAGTAGACTACTCGTAGCGAACGTTACCCTCGTTCTACTATGCTTTTATACTATCGAACAAATAGACTATTCCTATAAAAAAGTATTCCTTGGGATTCGTCCCTAAGTAGGGGTTAGAACAGTTAAATTTTG

Supplementary Table 4 (cont.) – List of primers used to clone the synthetic promoters of group T

Name	Sequence
T9	GTTCTTCTCCTTTGCTAGCCATCGTTTCGATAGGATTATGGTGTAGGACAGCACTCACCTATCAAATTCGTTTTTCCTACTCGTACCCTTCTAATA TTCCAATTCTAAATAGGACTATTATATCGGATGTTACACTGCGCCTTATATACTTGGCTATCAAATCTCTAAAGTTAAGTAGGGGTTAGAACAGTT AAATTTTG
T10	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGTAAAAGACTACTTCGGCTATCTAAAAGTTTAACAGTTCAGGGTTTATCGCAAGATTACTCGAGTA ACTGATTTCCTTAAGATAAGTTTTCGCGAACTCACCTAGGAGGTACCTTTTATAATGGATGTCCTTGGTAAGAGAGAAAGTAGGGGTTAGAACA GTTAAATTTTG
T11	GTTCTTCTCCTTTGCTAGCCATCGTTTCGATCCTTAGTAAGAATCTTCCTAGAAAAGGCAAAATAGATATCTCTCTCTAGTTGCTATGGTTGGTTT CCAGGCTACCTTGACCAAATTTTCTATGTCTAGCTTCAGACGATCTATTTTTATATAAAAGTAAATGTTTATAACAAAGTAGGGGTTAGAACAGTT AAATTTTG
T12	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCGGCTCTCCTTCTGTTTATCGATATTCTAATTTATCGATGGTTATGTTGCGACGTGTTCTAACCTAT ATTCTGGGAGATTATTTATAACTGATAAACTGTCTATTATAGTTTTCTAACTGTGGCTATAGATTTCAGGGTGATTAAGTAGGGGTTAGAACAGTT AAATTTTG
T13	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGGCTTTTCTCAAAACGATTTCTTATCGGACCGATATTAGTGTTAACTATATAATATTCGCCGGAAC GGAGACTAGATATATTAAACCTTATATAGACGTAGTTCGTCTAGAATTCAAGTCGTATAGTGAGAAGTTTAGCTAAGAAGTAGGGGTTAGAACAG TTAAATTTTG
T14	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTATTCCTAAGACTAGAATACTATCACTAATCTAGAACGGGAGATGTAGGCTCGAGAGATCCAGTCT GCGTAAATATAGCGCCACCAAACGTACTGATTTTTATATGTTTACCAAAAGTACGACCCCTATTGTTTCCTACCAATAAGTAGGGGTTAGAACAG TTAAATTTTG
T15	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGATAACTATAATAGTATATAGGGAGCTATAATAATGCAAATTCCTCGGTGGTTGGAAATTATTACTT TAAATATTATAATTTATATACTAGTCGCCCGTTCGTATGCCCAAGATATATTCGTTTTGCCACTATCGTTTCGTTTAAGTAGGGGTTAGAACAGTT AAATTTTG
T16	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTATTTTTCTAACCAGCGTAGGATGGTAAATGAAATGTAGATCTATGGGTTGCGTTACTACTTTCCA GATTCGTTTTAATTTACAAATATAAGTATAGCTTATATACAGTAGATATATATGATTGGACGATCAACAAGACGTATAAGTAGGGGTTAGAACAGT TAAATTTTG

Supplementary Table 4 (cont.) – List of primers used to clone the synthetic promoters of group T

Name	Sequence
T17	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTCAACGTATTAGGCAAGAATATTTTTATGTAAAAAGCTTCTTAAATTCGGACTATAGTAGTTCCCTC CCCTTCCCACTACGCCGCAAAGTTTGTTTTGATCGTTTATATAGATTCGGTTCTGTACCTTGATCCCAAGCAAAAAAAGTAGGGGTTAGAACAGT TAAATTTTG
T18	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCAACTATTATGGGCTCTGTTATCAGTAGTATTCGATGTTCAGTAGTGTGTCTATGTTTATCGAAGAT CTCTGTAGTTACCAATGGTGCGTATTTTTCTACCGTTTTTTATACTATCCAGACAGTTAATCTATCGGCCTCCTAGAAGTAGGGGTTAGAACAGT TAAATTTTG
T19	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTATGATATATAGCTATAAACTTACACGTCAGGAATATGTTCACTCTCGCAACAAATCACACACTAAC CCCTATAAATTCGTTTAGCTTTCGTTCTAACAGTGTCTGGTAGAGTACTTATATAAATTTATAACAAGATCAATTAAAGTAGGGGTTAGAACAGTT AAATTTTG
T20	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCTTATAAACTTATGCAATGATCAACTAAATCTACGAATTAGATAATATTTGCACGCGAGGTTACCTT CTATACGATAACCTGATATTTATAAACTGAGAACCCCAAAGAAAGGTTTTAAGTATCTAACACTATCGTGCAGAACAAGTAGGGGTTAGAACAGT TAAATTTTG
T21	GTTCTTCTCCTTTGCTAGCCATCGTTTCGATATGTATCAGTACCCTAGGAATATCGTATAAATGTTCTATTTTCTATGGAATATGGGTGTCAGCTT TATATATTGTGCACCTGATAATTGTAGGACCAAATTTGGGTCGTAATCTCTACTAATCAAGATTGGCAAGAATTCTAAGTAGGGGTTAGAACAGT TAAATTTTG
T22	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAATAAGAGTTTTATGTACTTAACGTACTACCTGGAAAACGGACAAACTAAGCTACTAAGTAAGCTA TGTTAGTACGCGTCTATATCTATTGAATGTCTAATTAAGAGATTATTTATAACTACCTAAAATTAATACCGTCACGAAAGTAGGGGTTAGAACAGT TAAATTTTG
T23	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAGCTTCTACTATTTCAACTAGAGTTCTAAATCTTAAAAGTAAATTCGTATCTCAGATTGGGGATTCG TCGGTGGAAGGATACAACTTTTATAAGAAATTACTACGAATAGACTTTTAATGATACAATCAATGTGGGTTAAACGAAGTAGGGGTTAGAACAGT TAAATTTTG
T24	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTACTGAGCTATAGCGTGGTTTAGCTAAACAATCGAAAAATCTCTGGGGTAACTAGAGCGCTTAATA GGTACTTGTGCTAAATTTGTACTGATAATTTACGAGCGAGACAATATTTATAGACTGCCCTAAGTAAGTTCGAAATTAAGTAGGGGTTAGAACAG TTAAATTTTG

Supplementary Table 4 (cont.) – List of primers used to clone the synthetic promoters of group T

Name	Sequence
T25	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGTGATTAAATATACCTATTCTCGTGGAACGAACGGGGTCAGTTTTGGGTCTGAAGATCGTCAATCT CCTGAATTTATACGATATCAGATAATGTTTTTGATGAGGTGAGATTGACTTGCAGGACCACCTTTTATAGTCTATATAAGTAGGGGTTAGAACAGT TAAATTTTG
T26	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTACCTTCCTATACTACTCTAGAATGTAACCTATAATCTTAATAACTTCTATCTGCTTTTTATCGATTC GAATCTAGATGTTTGATTATGCTTTAAATGAAAAATTCTAACGACTTCAGCTTTTATAGATAAACCTAGAGGAGGAAGTAGGGGTTAGAACAGTTA AATTTTG
T27	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAACAGAAAAAGAATGTGATGTTAAAAGCTTTTTACGTGCAAAATACCCTCACTTCAGATATCGGAAC AACAAGGGTATGCTATACAAACGAAAGGTTACTGAGGGACTTGACCTACGTCTGTGACGGTTTATATACTAGTAAAAAGTAGGGGTTAGAACAG TTAAATTTTG
T28	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAATAAAATAAGGTTTCTTTTAACTAGTTATAACAGTTTTACTAGACTTATACTTCTATAGGTTATTCTA AGTTATACAATAATCTAGACAGCAATAAGCTATATTTATAGAAAACCGAGCCTTATCTATCGTGAAAACGAGCTAAGTAGGGGTTAGAACAGTTA AATTTTG

Supplementary Table 5 – List of primers used to clone the synthetic promoters of group A

Name	Sequence
A1	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTTTTGTTGTTGTTTGCGAATATATACTAATTCAGCTTAATGAACACCTAACTATACTCTTTTGAAATA CAAGTGGGGGTGACTATTTCACCTACAGGTGTCAAGTTCTAAAGGTATTTAAATCCCTTTTATAAGGTAATTTATAAGTAGGGGTTAGAACAGTT AAATTTTG
A2	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTCTTAAAATGTTGAGCCGCACCTGGGCACAAAGAAAGTATTGCGTTGTTATACGACCTTGCGGCC GACGAAATGCTATTATGATAATGCTTGATGTAGTGGGTACAAGATTATTGTGTAAAATATTTATACAACTAAGAATATAAGTAGGGGTTAGAACA GTTAAATTTTG
A3	GTTCTTCTCCTTTGCTAGCCATCGTTTCGATAACGCTACTGAGGATTCTTTAACTGCGTATGATCTAGCTCTAGAATTATTTTATTCTCGTTAGTT ACTCCTTATATACCTAAGATAGTTTACGCCGAACTCGCTCAATTGATCAACGTTATCTAACTTGAGCTTTAAAGTGAAGTAGGGGTTAGAACAGT TAAATTTTG

Supplementary Table 5 (cont.) – List of primers used to clone the synthetic promoters of group A

Name	Sequence
A4	GTTCTTCTCCTTTGCTAGCCATCGTTTCGACCACTTTACTGACAGCTACCCAGAAGGGGACTAAGATATCGGATTTGTGGGAGTTAAAGAAGTT AATAACAACGTGTATAGATAAAAATCTTTTTATAATGAACAAGCTTACTGTATTCGCACAACGAAGTCGCTTTGCTTAAAGTAGGGGTTAGAACA GTTAAATTTTG
A5	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAGCACAATTTTGAGCTAATTGATGCTATGCTAGGAAACTTTCTGTTGATCGTTACTATGTTTTTAGT TATAGAAGGCTATCAATTCTCGATTGTTGATAAACGGGGATCTAACTTTTATACTCAAAAACAACGGGAAATTAGGAAGTAGGGGTTAGAACAGT TAAATTTTG
A6	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGTGTGCTTGACAATCCTCCTAGCACTGGCTTTGTTCTTTGACCTACTGAATTCCTCTAGTAAATAC TACTGTGTAATATTTATATACGGGAGTCGAAAGCCCAGCAAACTTTCCTCTATCTACAACAAAATGTACTACTAACCAAGTAGGGGTTAGAACAG TTAAATTTTG
A7	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAGATACTAAATTTCCTGTGGGATTTGATTGATCTACGCTACCACGTAGAACAATATTAAACCAATAC TATCTGGCACTTATGCGAATTCTTGAGATATTGTTCAAGATTCCAATATAAGCTTTTATATCTGTGTTCTTCAATAAAGTAGGGGTTAGAACAGTT AAATTTTG
A8	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGATGGATTAATGGGTAGCTGTATGACGATAGTATAACTTTTGATGACTTTAATTTAACCTCGACTCA AGTAAACTTTATGTCGCTAATTCTTGATCTAATTTCTTATATATAACTAGACAATCAGGACCCTAAACTATGATTAAAGTAGGGGTTAGAACAGTT AAATTTTG
A9	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGGATCCTATGAGACGTATTTCTTCTATTGGATTGTTTCAATCTACTATGAAGGTGATCGTCTACTAC AGCCACCCCTACTCCAACCTAATACTTTTATATGGGTTGGTTTCTTTGCCCCGATTCCGCTAATAATACAAATTTTAAGTAGGGGTTAGAACAGT TAAATTTTG
A10	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCTTACGGCTAGTAAGTACTTGATTCGGACCCCTATATACAGGATATTTGTATTGGGAACTATCGTA TAAGCTTACGTGCTCTATCTTAGGTAGGGAAAAACACCACTATCTACAGAACCTATCGTTTTTATTTATATATTAATAAGTAGGGGTTAGAACAGT TAAATTTTG
A11	GTTCTTCTCCTTTGCTAGCCATCGTTTCGGGTATCTCTAGACTATTGCGACTTTCTACTACAACCCACTACTAATTGATTGCGTCTATAGATCTTC GAACACTATATGGATTTAAAAGTACTTTTATAATGCTATTCTAAGAATTACTTTAAGAATTGGATGTTAATTCAATAAGTAGGGGTTAGAACAGTTA AATTTTG

Supplementary Table 5 (cont.) – List of primers used to clone the synthetic promoters of group A

Name	Sequence
A12	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTTGTAACTACTAAAAGTTAATATTAAGGTATTACTATGGATCGCTCTAATTTCAATCTTATACGTTTG TCTAACTCGCTCAAAATATTTATACGAAATTCTTATATTGGTTAAGGTTGAAAACTCAATTTAGGTTCACACGATAAGTAGGGGTTAGAACAGTTA AATTTTG
A13	GTTCTTCTCCTTTGCTAGCCATCGTTTCGACAAATTGCTATTTTATTGTCTTAATAGTATCTGTAGATTATTGATATATGATATAGAGGTAACTCTT TTCCAGAATAGAAAGGTCACTAGTCTTTCTGTACAACCTAAACCGTAATTTTTATATTGACCAAAATAGAAAGAGAAGTAGGGGTTAGAACAGTT AAATTTTG
A14	GTTCTTCTCCTTTGCTAGCCATCGTTTCGACTATAGGTTTCTGTTGGTAATTACGTTTACTCTTATCGTGGTGGTACAAAAAATACTAAAGTATCA AATCTTGAATTTATAGAACTAAACTAAGTGGATACCTATTACTACCTAGTTACTATTATAATTATATATATGAAAAAAGTAGGGGTTAGAACAGTTA AATTTTG
A15	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTTCTACCCTTTCCAATATATCGGGATATTGATTAATAATTACTGCTAAGCGCGAGGCTTATATAAGG GTATGTAACGAGAGGAATATCTTTCAGGGTCAACGTTGGATTACCCTATTTTACCTATTATATAAGGGTTGCGAGTAAGTAGGGGTTAGAACAG TTAAATTTTG
A16	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCAGCTTGCGCTAACCTATTCTAAATGTAAATTGCACTATCTACTTACTTATATCTTTCTAGCAACGT AAGTAGTTGGAACTATTACTCAAAAATTCAATTTCTTTTATACCAGATCAATTACCAGTCTATGTGCGACTTATGTAAGTAGGGGTTAGAACAGTT AAATTTTG
A17	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCCTACTAATCTATGGGATAAAGAAGTGAAATTGATGCGATTGCGTGACAAAGAGAGCTATACTTAC TCTGTCTTATTAGTTTTACTGTGGCTTTTTATATATTTCTACTTTAAGACTAGATTGACCTATTTAGAATATATAATAAGTAGGGGTTAGAACAGTT AAATTTTG
A18	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTAGTGGCTGTATGTAGTCTCAATTCTAATATTCTTATCAATTGATTATTTTGGTTGAATGTAGGTAC TAATCTTATAAATTAAGAAGGTTAATAAATCTTAACTCAATGAGAGAATAGTTTACAATTTTTTATATCTAAACAGAAGTAGGGGTTAGAACAGTTA AATTTTG
A19	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCCTATAATTGGGAGAAACGAGGATAGTTGAATTCTAGCTCCTATAGTATGAACGGTTATGTAAAGG AATGAGATGCTATCTTATATACGAACTTGAGTTTAAAAGGTTATTTTTCTTAAATACTCCACCCAATGGAATTTAGTAAGTAGGGGTTAGAACAGT TAAATTTTG

Supplementary Table 5 (cont.) – List of primers used to clone the synthetic promoters of group A

Name	Sequence
A20	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTAATATATTTATTATTGTTCACCTCTCTGCTATGGAGTTACGACTCGTCGACGGGATATGTTTACTA GGAGGTCTTACAGGGTAGATCTTTCGAAACTAGTGATTTATATATATCTTTTTCTATGACCTCCTAGTACGTTATCAAGTAGGGGTTAGAACAGT TAAATTTTG
A21	GTTCTTCTCCTTTGCTAGCCATCGTTTCGCTTTAACTACGTCGAGGGACCTCTATTTAGTCTGGCCGAGTTGCGTAGCTAAAGTTTAAGATCTAC TAAAGTTTAGTATGTGTTCGATAATGATTTTCTCTCACGTTCGCTTTTCGAAGTTTTTTCGCCTTTTATAATCGTTAAAGTAGGGGTTAGAACAGT TAAATTTTG
A22	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTGTCTACAACTTACTAATCAACTCACTACGACTTTAATAGTCAACAAGAGAACGAAAGAGATTCTC CAAATACAATCTTTGTTTATATATGTGATGTCGATTAAGTCTTCCTATCTACTAATGTACACTGCAGTGAGGATGTCAAGTAGGGGTTAGAACAG TTAAATTTTG
A23	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAAGTGATCGTAAGAGAATTTGAATCTTGACTAAGCTAGGCTTAGGTGTTGAAGTAAGCGCAAATGA ACTTAATCTTGTTTGCTCTTTAGTAAAATTCAGTGATTGGAGATCACTTTTATAGAGAACAAGTATTTCGTCAGCAAAAGTAGGGGTTAGAACAGT TAAATTTTG
A24	GTTCTTCTCCTTTGCTAGCCATCGTTTCGTATTTGGCTGCCCGTTACGTAACCCAAAAATAGTGAATTTCCTGGCCGAATAGTTCTGTTCAGTGG GCCTTCTACTATGTATGGGAGCAAAAACCAACACCCTTATATAAGTATTCTACAAATTAATCTTGGCACCCAAAAGAAAGTAGGGGTTAGAACAG TTAAATTTTG
A25	GTTCTTCTCCTTTGCTAGCCATCGTTTCGATTAACAATACGCTTACCTCGCAATTTGGTATCTAGATTGCTTAAGCAGGCCAATCCTTAATGTGC TTTCTTCTCGTTGTCAATTAATCTGATTGACGCTTTCTATTATTTATACCCGTCTTGGTGAATTCAACTATATAGAAAAGTAGGGGTTAGAACAGT TAAATTTTG
A26	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAATGATATAGAAATGTCCTATATGTAATTCTAAGACAGTCTAATTTTCTATATAAAATTGCAACTACT ATCTAGGAAACTGCAATGTCACACAGTAAAAACCTGGGCCTTATCTTATATATGTTAGAGCCTATCCTTAGTGACAAGTAGGGGTTAGAACAGTT AAATTTTG
A27	GTTCTTCTCCTTTGCTAGCCATCGTTTCGAATAGGAGTGAAGGTATAATTCTGACTCTGAGCTTTTTTGTAGTAAATGTTTCGCTAGATAGATATC TATGATTAGTACTAGTAAAATAGTCCCTGAATATTATAATGTATTAGTGTCCCAGATTTTTTTAGTTTATATACTGAAGTAGGGGTTAGAACAGTTA AATTTTG
A28	GTTCTTCTCCTTTGCTAGCCATCGTTTCGACCCTTCCGATTATTAATGTTACGGTACCTTCTTTGTTGTAAATGTTCGAAGAATTGTTTCTTCCGT TTAAACCAAGAGTTGCTTGTACTGCGTTGGTATATTTGGTATACAACCTCCTACTTTATATATTTCTGCCCTATCGAAGTAGGGGTTAGAACAGT TAAATTTTG

Supplementary Table 6 – List of primers used test the ten best synthetic core promoters fused to different CRMs in *P. pastoris* and *S. cerevisiae*

Name	Sequence
CAT-core	AAAGTTCTTCTCCTTTGCTAGCCATCGTTTCGTTTAATTGTAAGTCTTGACTAGAGCAAG
CAT-CRM-forw	GCTGGCCTTTTGCTCACATGTATTTAAATTAATCGAACTCCGAATGCGGTTC
DAS-core	GAAAAGTTCTTCTCCTTTGCTAGCCATCGTTTCGTTTGTTCGATTATTCTCCAGATAAAATCAACAATAGTTG
DAS-CRM-forw	GCTGGCCTTTTGCTCACATGTATTTAAATAGCAATGATATAAACAACAATTGAGTGACAGG
GAP-core	GAAAAGTTCTTCTCCTTTGCTAGCCATCGTTTCGTGTGTTTTGATAGTTGTTCAATTGATTGAAATAGG
GAP-CRM-forw	GCTGGCCTTTTGCTCACATGTATTTAAATTTTTTGTAGAAATGTCTTGGTGTCCTCG
CAT-CRM-rev	TCCAACAAAGAGGCAACAGAGGTCGGCGCGCCACTGGGTGCTACTGATGAGCAACAGAGGCTATCAC
DAS-CRM-rev	TCCAACAAAGAGGCAACAGAGGTCGGCGCGCCACTGGGTGCTATGCTTTAGTTCTTTTTTGAACCCAAAGGCTATCTGATGAAAAG
GAP-CRM-rev	TCCAACAAAGAGGCAACAGAGGTCGGCGCGCCACTGGGTGCTAGTGGTTTCCAATAATCTCATGACATGCG
seqTomato19-41rev	CGCATGAACTCCTTGATAACTTC
ADH-CRM-rev	TCCAACAAAGAGGCAACAGAGGTCGGCGCGCCACTGGGTGCTAAGACAGCAAACTTTTTTTTATTTCAAATTCAAGTAAC
GAL-CRM-rev	TCCAACAAAGAGGCAACAGAGGTCGGCGCGCCACTGGGTGCTAGATCAAAAATCATCGCTTCGCTGA
GPD-CRM-rev	TCCAACAAAGAGGCAACAGAGGTCGGCGCGCCACTGGGTGCTAAAGTAGGGGAATAATTTCAGGGAACTG
A28-GFP-rev	AAAAGTTCTTCTCCTTTGCTAGCCATCGTTTCGACCCTTCCGATTATTAATGTTACGGTA
T28-GFP-rev	GAAAAGTTCTTCTCCTTTGCTAGCCATCGTTTCGAATAAAATAAGGTTTCTTTTAACTAGTTATAACAGTTTTACTAG
A27-GFP-rev	GAAAAGTTCTTCTCCTTTGCTAGCCATCGTTTCGAATAGGAGTGAAGGTATAATTCTGACTCTGAG
T27-GFP-rev	GAAAAGTTCTTCTCCTTTGCTAGCCATCGTTTCGAACAGAAAAAGAATGTGATGTTAAAAGCTTTTTACG
T26-GFP-rev	AAAGTTCTTCTCCTTTGCTAGCCATCGTTTCGTACCTTGGTATACTACTCTAGAATGTAA
T25-GFP-rev	GAAAAGTTCTTCTCCTTTGCTAGCCATCGTTTCGGTGATTAAATATACCTATTCTCGTGGAACGAAC
M28-GFP-rev	GAAAAGTTCTTCTCCTTTGCTAGCCATCGTTTCGTATAAATAATGTAATAACTTTAAGATATGGATTGCACTTG
T24-GFP-rev	GAAAAGTTCTTCTCCTTTGCTAGCCATCGTTTCGTACTGAGCTATAGCGTGGTTTAGC
T23-GFP-rev	GAAAAGTTCTTCTCCTTTGCTAGCCATCGTTTCGAGCTTCTACTATTTCAACTAGAGTTCTAAATCTTAAAAG
T22-GFP-rev	GAAAAGTTCTTCTCCTTTGCTAGCCATCGTTTCGAATAAGAGTTTTATGTACTTAACGTACTACCTGG

Supplementary Table 6 (cont.) – List of primers used test the ten synthetic core promoters with different CRMs in *P. pastoris* and *S. cerevisiae*

Name	Sequence
A28-CAT-for	GTGATAGCCTCTGTTGCTCATCAGCGATAGGGCAGAAATATATAAAGTAGGAGG
T28-CAT-for	GTGATAGCCTCTGTTGCTCATCAGAGCTCGTTTTCACGATAGATAAGGCTC
A27-CAT-for	GTGATAGCCTCTGTTGCTCATCAGCAGTATATAAACTAAAAAAATCTGGGACACTAATAC
T27-CAT-for	GTGATAGCCTCTGTTGCTCATCAGTTTACTAGTATATAAACCGTCACAGACGTAGG
T26-CAT-for	GTGATAGCCTCTGTTGCTCATCAGCCTCCTCTAGGTTTATCTATAAAAGCTGAAG
T25-CAT-for	GTGATAGCCTCTGTTGCTCATCAGATATAGACTATAAAAGGTGGTCCTGCAAG
M28-CAT-for	GTGATAGCCTCTGTTGCTCATCAGGTAATCTTTCGGTCAATTGTGATCTCTC
T24-CAT-for	GTGATAGCCTCTGTTGCTCATCAGAATTTCGAACTTACTTAGGGCAGTC
T23-CAT-for	GTGATAGCCTCTGTTGCTCATCAGCGTTTAACCCACATTGATTGTATCATTAAAAG
T22-CAT-for	GTGATAGCCTCTGTTGCTCATCAGTCGTGACGGTATTAATTTTAGGTAG
A28-DAS-for	GCCTTTGGGTTCAAAAAAGAACTAAAGCACGATAGGGCAGAAATATATAAAGTAGGAGG
T28-DAS-for	GCCTTTGGGTTCAAAAAAGAACTAAAGCAAGCTCGTTTTCACGATAGATAAGGCTC
A27-DAS-for	GCCTTTGGGTTCAAAAAAGAACTAAAGCACAGTATATAAACTAAAAAAATCTGGGACACTAATAC
T27-DAS-for	CCTTTGGGTTCAAAAAAGAACTAAAGCATTTACTAGTATATAAACCGTCACAGACGTAGG
T26-DAS-for	GCCTTTGGGTTCAAAAAAGAACTAAAGCACCTCCTCTAGGTTTATCTATAAAAGCTGAAG
T25-DAS-for	GCCTTTGGGTTCAAAAAAGAACTAAAGCAATATAGACTATAAAAGGTGGTCCTGCAAG
M28-DAS-for	GCCTTTGGGTTCAAAAAAGAACTAAAGCAGTAATCTTTCGGTCAATTGTGATCTCTC
T24-DAS-for	GCCTTTGGGTTCAAAAAAGAACTAAAGCAAATTTCGAACTTACTTAGGGCAGTC
T23-DAS-for	CCTTTGGGTTCAAAAAAGAACTAAAGCACGTTTAACCCACATTGATTGTATCATTAAAAG
T22-DAS-for	GCCTTTGGGTTCAAAAAAGAACTAAAGCATCGTGACGGTATTAATTTTAGGTAG

Supplementary Table 6 (cont.) – List of primers used test the ten synthetic core promoters with different CRMs in *P. pastoris* and *S. cerevisiae*

Name	Sequence
A28-GAP-for	CGCATGTCATGAGATTATTGGAAACCACCGATAGGGCAGAAATATATAAAGTAGGAGG
T28-GAP-for	CGCATGTCATGAGATTATTGGAAACCACAGCTCGTTTTCACGATAGATAAGGCTC
A27-GAP-for	CATGTCATGAGATTATTGGAAACCACCAGTATATAAACTAAAAAAATCTGGGACACTAATAC
T27-GAP-for	CGCATGTCATGAGATTATTGGAAACCACTTTACTAGTATATAAACCGTCACAGACGTAGG
T26-GAP-for	CGCATGTCATGAGATTATTGGAAACCACCCTCCTCTAGGTTTATCTATAAAAGCTGAAG
T25-GAP-for	CGCATGTCATGAGATTATTGGAAACCACATATAGACTATAAAAGGTGGTCCTGCAAG
M28-GAP-for	CGCATGTCATGAGATTATTGGAAACCACGTAATCTTTCGGTCAATTGTGATCTCTC
T24-GAP-for	CGCATGTCATGAGATTATTGGAAACCACAATTTCGAACTTACTTAGGGCAGTC
T23-GAP-for	CGCATGTCATGAGATTATTGGAAACCACCGTTTAACCCACATTGATTGTATCATTAAAAG
T22-GAP-for	CGCATGTCATGAGATTATTGGAAACCACTCGTGACGGTATTAATTTTAGGTAG
A28-ADH-for	GTTACTTGAATTTGAAATAAAAAAAAGTTTGCTGTCTCGATAGGGCAGAAATATATAAAGTAGGAGG
T28-ADH-for	TACTTGAATTTGAAATAAAAAAAAGTTTGCTGTCTAGCTCGTTTTCACGATAGATAAGGC
A27-ADH-for	GTTACTTGAATTTGAAATAAAAAAAAGTTTGCTGTCTCAGTATATAAACTAAAAAAATCTGGGACACTAATAC
T27-ADH-for	GTTACTTGAATTTGAAATAAAAAAAAGTTTGCTGTCTTTTACTAGTATATAAACCGTCACAGACGTAGG
T26-ADH-for	GTTACTTGAATTTGAAATAAAAAAAAGTTTGCTGTCTCCTCCTCTAGGTTTATCTATAAAAGCTGAAG
T25-ADH-for	GTTACTTGAATTTGAAATAAAAAAAAGTTTGCTGTCTATATAGACTATAAAAGGTGGTCCTGCAAG
M28-ADH-for	ACTTGAATTTGAAATAAAAAAAAGTTTGCTGTCTGTAATCTTTCGGTCAATTGTGATCTC
T24-ADH-for	GTTACTTGAATTTGAAATAAAAAAAAGTTTGCTGTCTAATTTCGAACTTACTTAGGGCAG
T23-ADH-for	GTTACTTGAATTTGAAATAAAAAAAAGTTTGCTGTCTCGTTTAACCCACATTGATTGTATCATTAAAAG
T22-ADH-for	TTACTTGAATTTGAAATAAAAAAAAGTTTGCTGTCTTCGTGACGGTATTAATTTTAGGTA

Supplementary Table 6 (cont.) – List of primers used test the ten synthetic core promoters with different CRMs in *P. pastoris* and *S. cerevisiae*

Name	Sequence
A28-GAL-for	TCAGCGAAGCGATGATTTTTGATCCGATAGGGCAGAAATATATAAAGTAGGAGG
T28-GAL-for	TCAGCGAAGCGATGATTTTTGATCAGCTCGTTTTCACGATAGATAAGGCTC
A27-GAL-for	TCAGCGAAGCGATGATTTTTGATCCAGTATATAAACTAAAAAAATCTGGGACACTAATAC
T27-GAL-for	TCAGCGAAGCGATGATTTTTGATCTTTACTAGTATATAAACCGTCACAGACGTAGG
T26-GAL-for	TCAGCGAAGCGATGATTTTTGATCCCTCCTCTAGGTTTATCTATAAAAGCTGAAG
T25-GAL-for	TCAGCGAAGCGATGATTTTTGATCATATAGACTATAAAAGGTGGTCCTGCAAG
M28-GAL-for	TCAGCGAAGCGATGATTTTTGATCGTAATCTTTCGGTCAATTGTGATCTCTC
T24-GAL-for	TCAGCGAAGCGATGATTTTTGATCAATTTCGAACTTACTTAGGGCAGTC
T23-GAL-for	TCAGCGAAGCGATGATTTTTGATCCGTTTAACCCACATTGATTGTATCATTAAAAG
T22-GAL-for	TCAGCGAAGCGATGATTTTTGATCTCGTGACGGTATTAATTTTAGGTAG
A28-GPD-for	CAGTTCCCTGAAATTATTCCCCTACTTCGATAGGGCAGAAATATATAAAGTAGGAGG
T28-GPD-for	CAGTTCCCTGAAATTATTCCCCTACTTAGCTCGTTTTCACGATAGATAAGGCTC
A27-GPD-for	GTTCCCTGAAATTATTCCCCTACTTCAGTATATAAACTAAAAAAATCTGGGACACTAATAC
T27-GPD-for	CAGTTCCCTGAAATTATTCCCCTACTTTTTACTAGTATATAAACCGTCACAGACGTAGG
T26-GPD-for	CAGTTCCCTGAAATTATTCCCCTACTTCCTCCTCTAGGTTTATCTATAAAAGCTGAAG
T25-GPD-for	CAGTTCCCTGAAATTATTCCCCTACTTATATAGACTATAAAAGGTGGTCCTGCAAG
M28-GPD-for	CAGTTCCCTGAAATTATTCCCCTACTTGTAATCTTTCGGTCAATTGTGATCTCTC
T24-GPD-for	CAGTTCCCTGAAATTATTCCCCTACTTAATTTCGAACTTACTTAGGGCAGTC
T23-GPD-for	CAGTTCCCTGAAATTATTCCCCTACTTCGTTTAACCCACATTGATTGTATCATTAAAAG
T22-GPD-for	CAGTTCCCTGAAATTATTCCCCTACTTTCGTGACGGTATTAATTTTAGGTAG

Supplementary Table 7 – Group P synthetic core promoters TATA box position and positive and negative motifs frequency and position

Promoter name	Postive motifs frequency and position										Negative motifs frequency and position				TATA box position
	TTTT	Position	TTCT	Position	CAA	Position	ATCA	Position	CAAT	Position	AAGA	Position	AGCG	Position	
P1	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P2	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P3	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P4	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P5	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P6	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P7	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P8	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P9	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P10	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P11	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P12	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P13	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P14	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P15	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P16	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P17	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P18	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P19	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P20	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P21	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P22	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P23	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P24	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-

Supplementary Table 7 (cont.) – Group P synthetic core promoters TATA box position and positive and negative motifs frequency and position

Promoter name	Postive motifs frequency and position										Negative motifs frequency and position				TATA box position
	TTTT	Position	TTCT	Position	CAA	Position	ATCA	Position	CAAT	Position	AAGA	Position	AGCG	Position	
P25	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P26	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P27	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-
P28	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-

Supplementary Table 8 – Group M synthetic core promoters TATA box position and positive and negative motifs frequency and position

Promoter name	Postive motifs frequency and position										Negative motifs frequency and position				TATA box position
	TTTT	Position	TTCT	Position	CAA	Position	ATCA	Position	CAAT	Position	AAGA	Position	AGCG	Position	
M1	1	-67	0	-	0	-	0	-	0	-	1	-148	0	-	-
M2	0	-	1	-90	1	-26	0	-	0	-	1	-109	1	-119	-
M3	0	-	0	-	4	-49; -46; -23; -16	1	-59	0	-	1	-73	1	-117	-
M4	0	-	1	-83	2	-49; -10	0	-	1	-49	2	-123; -105	0	-	-
M5	1	-69	0	-	2	-34; -13	0	-	1	-13	0	-	0	-	-
M6	0	-	0	-	0	-	2	-60; -30	0	-	2	-104; -88	1	-118	-
M7	1	-62	0	-	0	-	0	-	0	-	2	-135; -127	0	-	-
M8	0	-	1	-94	0	-	0	-	0	-	0	-	1	-83	-
M9	0	-	1	-81	0	-	1	-32	0	-	1	-106	0	-	-
M10	0	-	0	-	1	-30	0	-	0	-	0	-	0	-	-

Supplementary Table 8 (cont.) – Group M synthetic core promoters TATA box position and positive and negative motifs frequency and position

Promoter name	Postive motifs frequency and position										Negative motifs frequency and position				TATA box position
	TTTT	Position	TTCT	Position	CAA	Position	ATCA	Position	CAAT	Position	AAGA	Position	AGCG	Position	
M11	0	-	0	-	0	-	0	-	0	-	0	-	1	-103	-
M12	1	-80	0	-	1	-27	0	-	1	-27	0	-	0	-	-
M13	0	-	0	-	0	-	0	-	0	-	0	-	3	-123; -110; -88	-
M14	0	-	1	-59	1	-23	1	-54	1	-23	0	-	0	-	-
M15	0	-	3	-88; -71; -63	1	-20	0	-	0	-	1	-101	0	-	-
M16	1	-96	0	-	0	-	1	-37	0	-	3	-126; -123; -110	0	-	-
M17	0	-	0	-	2	-20; -16	0	-	0	-	0	-	0	-	-
M18	0	-	0	-	4	-50; -37; -21; -15	2	-60; -52	0	-	0	-	0	-	-
M19	0	-	1	-96	1	-17	0	-	0	-	1	-116	1	-87	-
M20	1	-96	0	-	1	-10	0	-	1	-10	1	-100	0	-	-
M21	1	-72	2	-89; -78	2	-44; -34	1	-46	0	-	0	-	0	-	-
M22	0	-	1	-95	1	-42	0	-	0	-	0	-	0	-	-
M23	0	-	1	-89	3	-43; -30; -22	1	-45	3	-43; -30; -22	0	-	0	-	-
M24	0	-	0	-	1	-19	2	-63; -37	0	-	1	-67	0	-	-
M25	0	-	0	-	2	-50; -32	4	-69; -60; -54; -42	1	-50	0	-	0	-	-
M26	0	-	0	-	0	-	0	-	0	-	1	-112	0	-	-
M27	0	-	0	-	0	-	1	-64	0	-	1	-58	1	-87	-
M28	0	-	0	-	2	-46; -40	0	-	1	-40	1	-87	0	-	-

Supplementary Table 9 – Group T synthetic core promoters TATA box position and positive and negative motifs frequency and position

Promoter name	Postive motifs frequency and position										Negative motifs frequency and position				TATA box position
	TTTT	Position	TTCT	Position	CAA	Position	ATCA	Position	CAAT	Position	AAGA	Position	AGCG	Position	
T1	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-124
T2	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-134
T3	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-110
T4	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-58
T5	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-65
T6	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-116
T7	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-93
T8	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-95
T9	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-125
T10	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-126
T11	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-130
T12	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-93
T13	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-100
T14	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-110
T15	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-93
T16	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-111
T17	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-116
T18	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-117
T19	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-128
T20	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-97
T21	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-79
T22	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-123
T23	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-98
T24	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-124

Supplementary Table 9 (cont.) – Group T synthetic core promoters TATA box position and positive and negative motifs frequency and position

Promoter name	Postive motifs frequency and position										Negative motifs frequency and position				TATA box position
	TTTT	Position	TTCT	Position	CAA	Position	ATCA	Position	CAAT	Position	AAGA	Position	AGCG	Position	
T25	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-141
T26	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-132
T27	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-141
T28	0	-	0	-	0	-	0	-	0	-	0	-	0	-	-115

Supplementary Table 10 – Group A synthetic core promoters TATA box position and positive and negative motifs frequency and position

Promoter name	Postive motifs frequency and position										Negative motifs frequency and position				TATA box position
	TTTT	Position	TTCT	Position	CAA	Position	ATCA	Position	CAAT	Position	AAGA	Position	AGCG	Position	
A1	0	-	0	-	4	-21; -17; -14; -11	0	-	0	-	1	-66	0	-	-138
A2	0	-	0	-	2	-48; -19	0	-	1	-48	0	-	0	-	-136
A3	0	-	1	-53	0	-	1	-41	0	-	0	-	1	-112	-85
A4	2	-95; -94	1	-68	0	-	0	-	0	-	1	-99	0	-	-105
A5	0	-	1	-80	2	-27; -18	1	-55	1	-27	0	-	0	-	-126
A6	0	-	0	-	3	-46; -39; -15	0	-	0	-	0	-	0	-	-92
A7	0	-	1	-56	2	-36; -32	2	-38; -34	1	-36	2	-144; -97	0	-	-133
A8	0	-	0	-	1	-48	1	-50	0	-	2	-111; -98	1	-91	-118
A9	0	-	0	-	2	-40; -35	1	-63	2	-40; -35	1	-119	0	-	-105
A10	0	-	0	-	1	-27	0	-	0	-	1	-93	0	-	-144
A11	1	-94	0	-	1	-24	1	-56	1	-24	1	-72	0	-	-105
A12	1	-90	0	-	1	-9	0	-	0	-	2	-108; -65	1	-85	-98

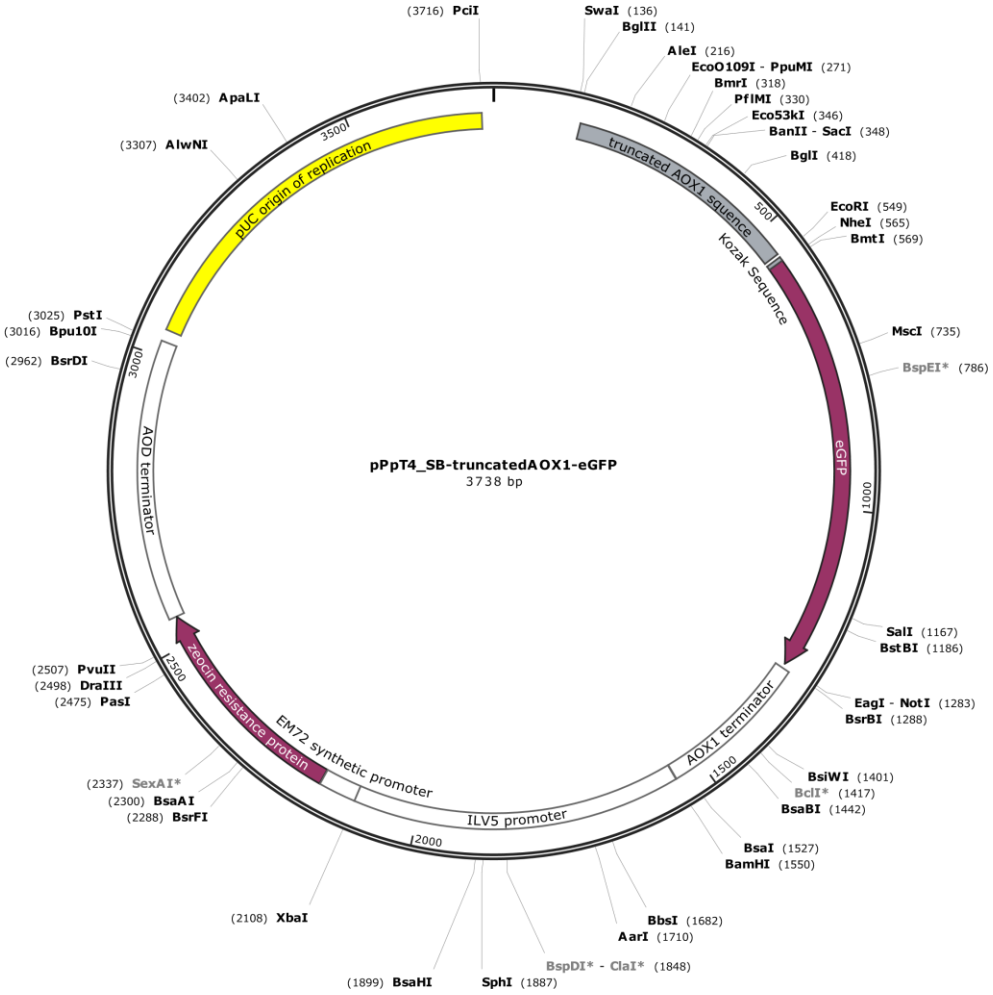
Supplementary Table 10 (cont.) – Group A synthetic core promoters TATA box position and positive and negative motifs frequency and position

Promoter name	Postive motifs frequency and position										Negative motifs frequency and position				TATA box position
	TTTT	Position	TTCT	Position	CAA	Position	ATCA	Position	CAAT	Position	AAGA	Position	AGCG	Position	
A13	0	-	2	-87; -82	3	-50; -25; -14	2	-59; -52	3	-50; -25; -14	2	-101; -74	0	-	-130
A14	3	-60; -59; -58	1	-91	1	-23	0	-	0	-	1	-79	0	-	-143
A15	0	-	0	-	1	-36	1	-38	1	-36	1	-96	0	-	-70
A16	2	-99; -98	0	-	2	-39; -13	0	-	1	-39	2	-110; -62	0	-	-115
A17	0	-	0	-	2	-48; -39	1	-41	2	-48; -39	1	-80	0	-	-107
A18	0	-	1	-92	1	-49	1	-51	1	-49	2	-106; -80	0	-	-140
A19	0	-	0	-	2	-35; -16	0	-	1	-16	2	-123; -88	0	-	-93
A20	0	-	0	-	1	-23	0	-	1	-23	3	-123; -96; -82	0	-	-119
A21	0	-	0	-	1	-49	0	-	0	-	0	-	1	-117	-142
A22	0	-	1	-58	0	-	0	-	0	-	2	-115; -85	0	-	-95
A23	1	-100	0	-	2	-34; -27	0	-	0	-	2	-92; -82	0	-	-126
A24	2	-99; -98	0	-	1	-12	0	-	0	-	1	-136	0	-	-115
A25	0	-	0	-	2	-46; -34	0	-	1	-46	2	-128; -78	1	-106	-120
A26	1	-63	0	-	0	-	0	-	0	-	1	-121	0	-	-126
A27	1	-93	0	-	1	-46	0	-	0	-	0	-	0	-	-146
A28	0	-	1	-60	2	-45; -42	0	-	0	-	1	-69	0	-	-135

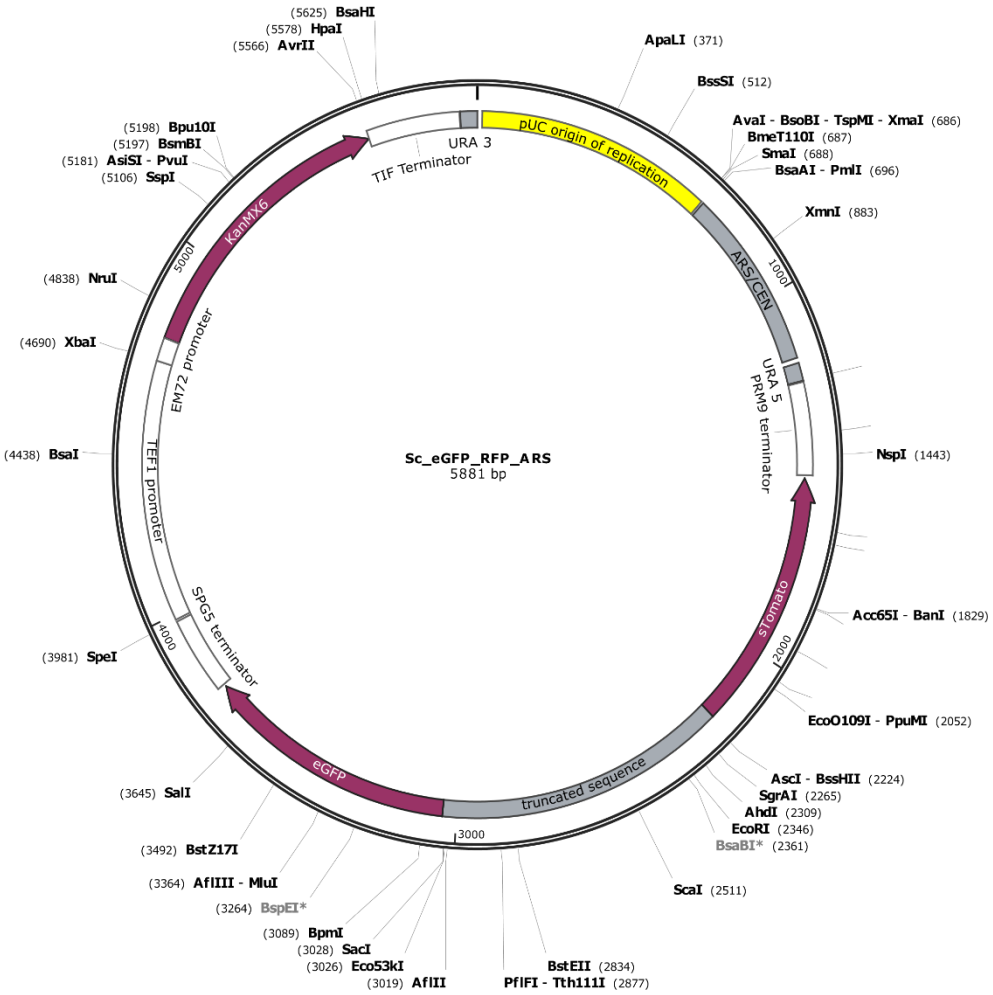
Supplementary Table 11 – Blast result for the 10 synthetic promoters with highest activity against the *P. pastoris* CBS 7435 genome. The results consist of the minimum *E*-value for each case, the presence of gaps, match localization (in the *P. pastoris* chromosome and genome and in the synthetic promoter sequence) and brief description of the sequence function in the *P. pastoris* genome.

Seq. Name	Min. *E* value	Gaps	Chromosome	Genome location (bp)	Syn. Seq. Location	Syn. Seq. location (from start codon)	Match description
A28	0.083	0	2	497292 to 497314	76 to 98	-74 to -52	Coding sequence of catalytic subunit of (1,3)-β-D-glucan synthase
T28	0.29	1	3	2049828 to 2049792	42 to 77	-108 to -73	130bp upstream of protein coding sequence with putative serine active lipase domain (possible promoter region)
	0.29	0	1	2266300 to 2266316	59 to 75	-91 to -75	Coding sequence of putative protein with unknown function
A27	1	0	3	943316 to 943288	125 to 149	-25 to -1	Coding sequence of hypothetical protein
	1	0	1	212956 to 212931	43 to 68	-107 to -82	Coding region of essential component of the Rix1 complex
T27	0.083	0	4	1369772 to 1369755	123 to 140	-27 to -10	10bp upstream of nucleolar protein coding sequence (possible promoter region)
T26	0.083	0	1	1928476 to 1928498	139 to 117	-33 to -11	Inter gene sequence (between a nucleolar protein and a transcription factor)
T25	0.29	0	4	565027 to 565053	20 to 46	-130 to -104	Coding sequence of hypothetical protein
	0.29	0	2	1165582 to 1165601	33 to 52	-117 to -98	Coding sequence of hypothetical protein
	0.29	0	1	1209655 to 1209634	61 to 82	-89 to -68	Coding sequence of hypothetical protein
M28	0.083	0	4	482161 to 482129	29 to 61	-121 to -89	Coding sequence of subunit of TFIIH and nucleotide excision repair factor 3 complexes
	0.083	1	1	450941 to 450977	149 to 112	-38 to -1	Coding sequence of Flavin adenine dinucleotide (FAD) synthetase
T24	1	0	1	1442424 to 1442406	47 to 65	-103 to -85	Coding sequence of phosphatidylserine decarboxylase of the mitochondrial inner membrane
T23	0.29	0	1	1877072 to 1877098	28 to 54	-122 to -96	Coding sequence of hypothetical protein
T22	1	0	3	619259 to 619274	80 to 95	-70 to -55	Coding sequence of Component of the ESCRT-II complex

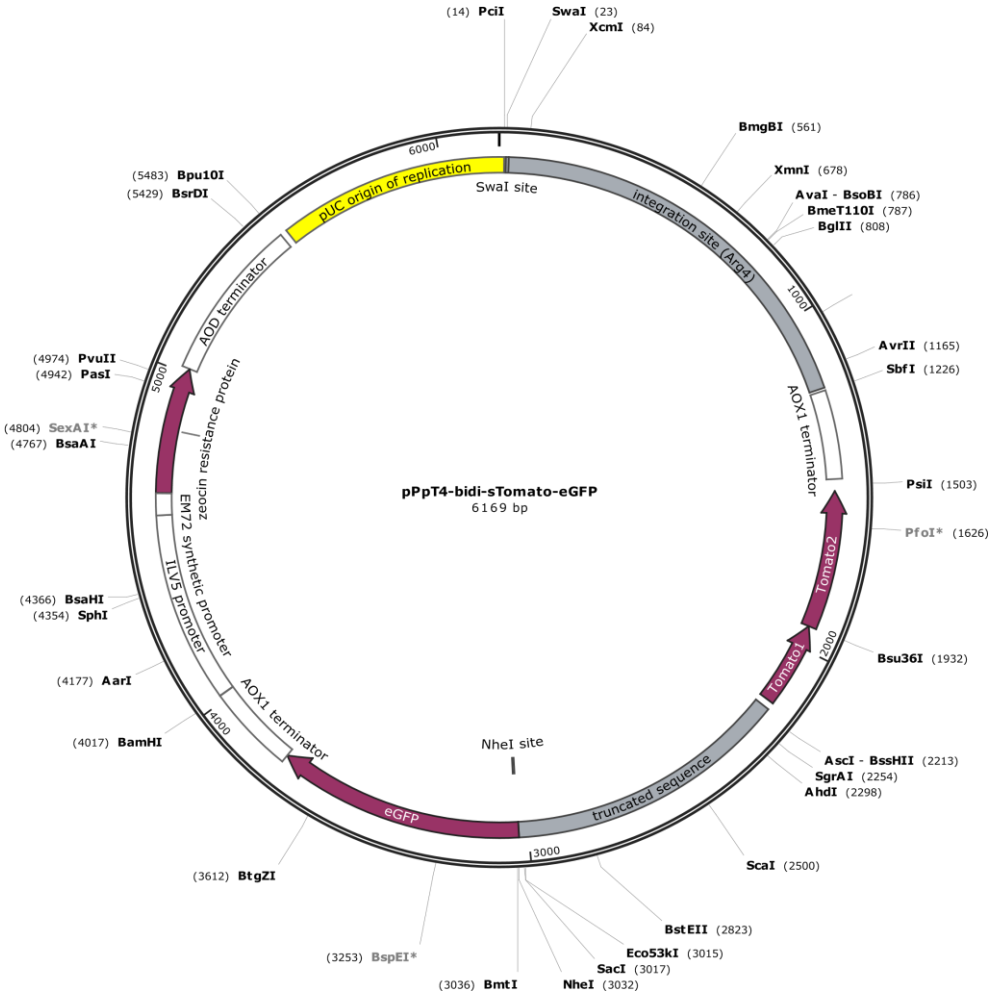
Supplementary Figure 1 – Map of *P. pastoris/E. coli* shuttle vector pPpT4_SB-truncatedAOX1-eGFP with main features highlighted: Restriction enzymes, eGFP, zeocin resistance marker, promoters and terminators and origin of replication.



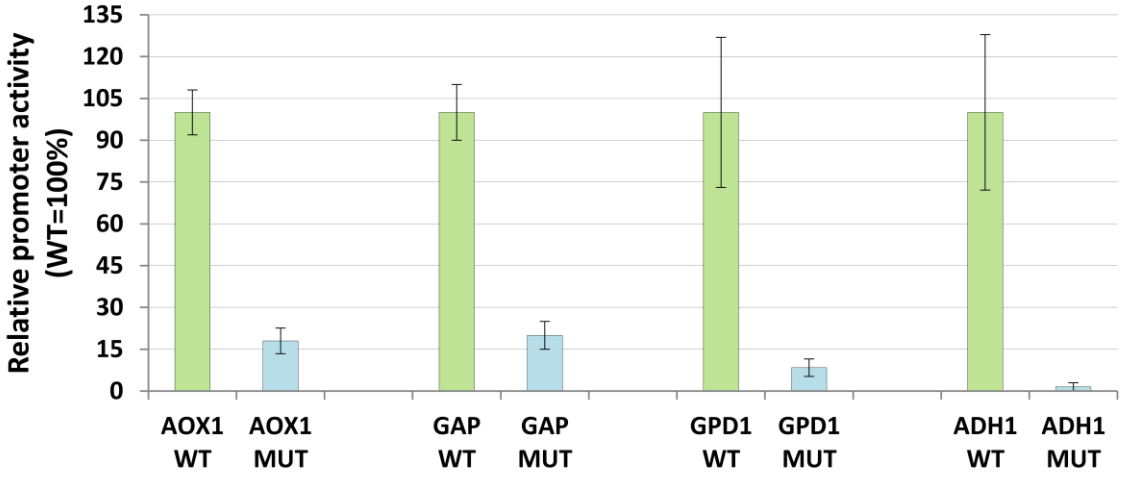
Supplementary Figure 2 – Map of Sc_eGFP_RFP_ARS with main features highlighted: Restriction enzymes, eGFP, sTomato (RFP), promoters and terminators, kanamycin resistance marker and autonomous replicating sequence (ARS).



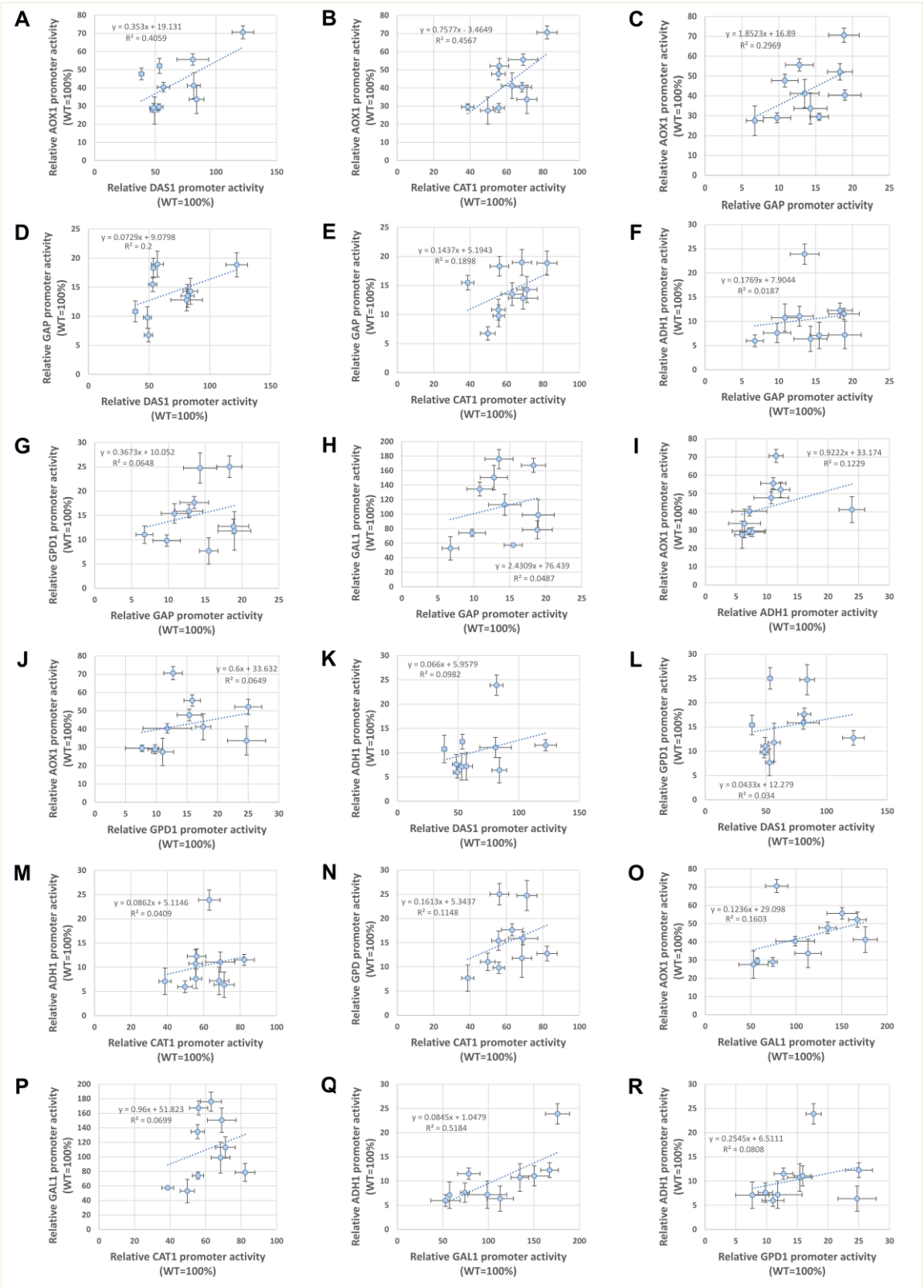
Supplementary Figure 3 – Map of *P. pastoris/E. coli* shuttle vector pPpT4-bidi-sTomato-eGFP with main features highlighted: Restriction enzymes, eGFP, sTomato (RFP), zeocin resistance marker, promoters and terminators and origin of replication.



Supplementary Figure 4 – Expression of the P_{AOX1}, P_{GAP}, P_{ScGPD1} and P_{ScADH1} promoters depends on the TATA box motif. The TATA box motif in the natural promoter sequence was mutated by replaying three nucleotides of this motif by cytosine. The reporter protein fluorescence of the mutated (MUT) promoters is compared to the unmodified wildtype (WT) promoter.

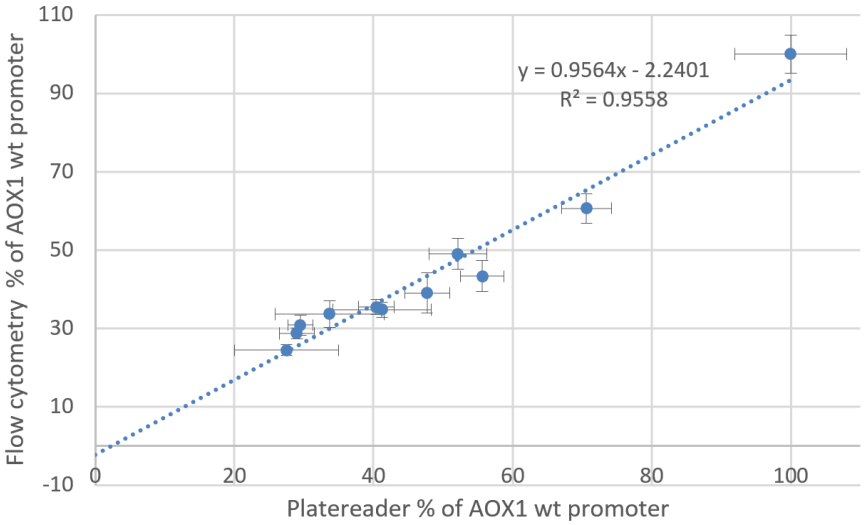


Supplementary Figure 5 – Additional correlation diagrams for comparisons shown in main Figure 5A. The heatmap in main Figure 5A was generated from 21 correlation diagrams. Three representative diagrams are shown as panels B-D in Figure 5 of the main manuscript, the remaining 18 are shown here.

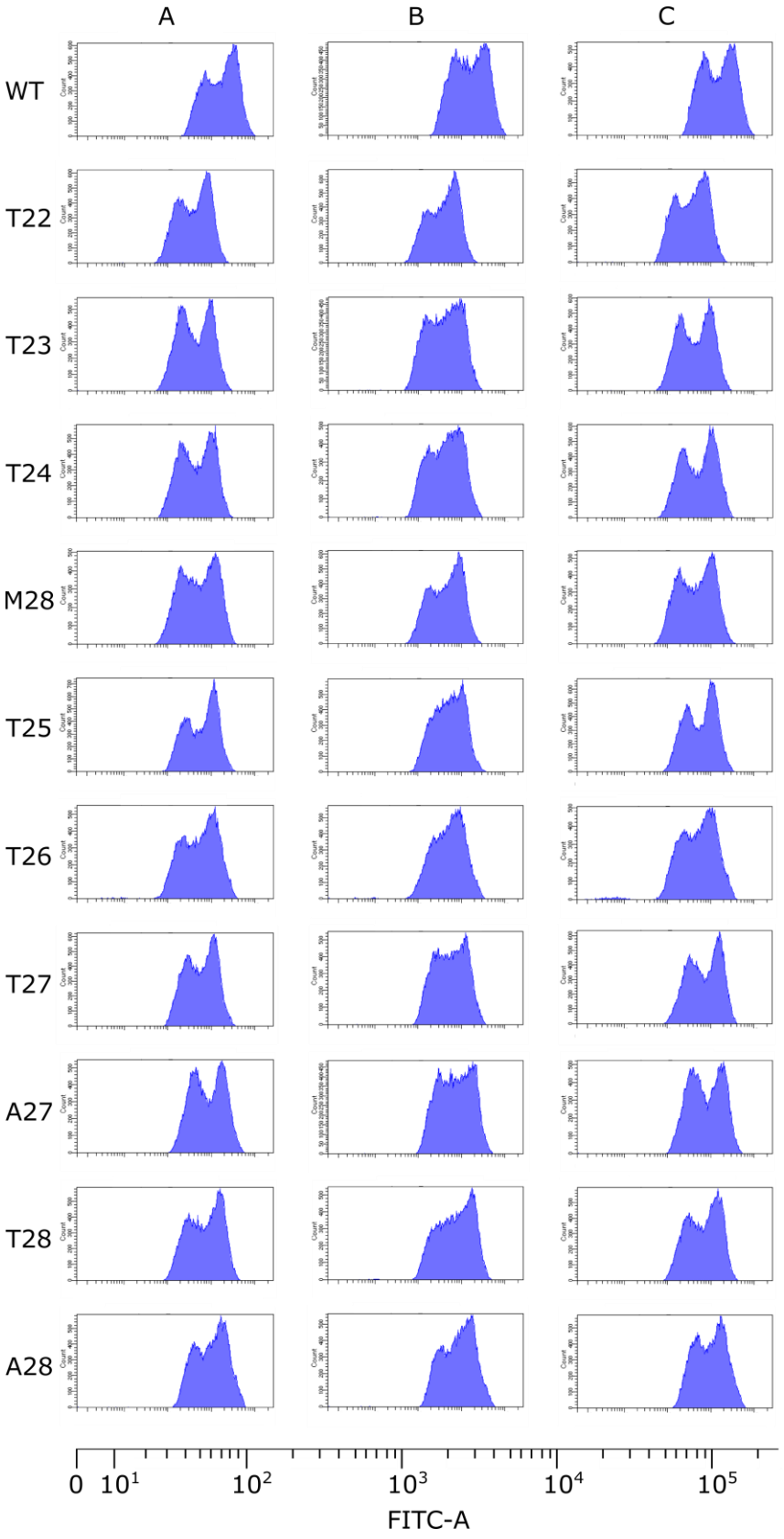


Supplementary Figure 6– The plate reader based fluorescence measurements are in excellent agreement with flow cytometry measurements. A: the *P. pastoris* strains bearing the top ten synthetic core promoters fused to the *AOX1* CRM and the *AOX1* wild type core promoter were regrown in deep well plates and measured by flow cytometry (see materials and methods section). The subset of synthetic core promoters fused to the *AOX1* CRM was selected, as this inducible promoter is tightly regulated, representing worst case conditions regarding on/off behavior. The plate reader measurements shown here are identical to Figure 4B in the main manuscript. A linear regression line was calculated, the formula and R^2 are shown in the figure. **B**: histograms of cell counts and respective measured fluorescence (log scale) of three representative biological replicates measurements (30000 events each) for each synthetic core promoter strain and the *AOX1* wild type promoter are shown. Note the different scaling of the y-axis for each plot. The histograms show, that the cell populations are highly similar when comparing synthetic core promoters to each other and also to the natural *AOX1* core promoter. Notably, all strains measured showed two separate fluorescence histogram peaks, indicating distinct cell population. These populations may be caused by the methanol inducible nature of our system: cells are at first grown on glucose and then induced with methanol. The different cell populations may be attributable to 'older' cells, having been grown on glucose and subsequently induced, and 'new' cells emerging from cell divisions after methanol induction and hence only grown under these conditions. These two peaks occurred for all core promoters tested (the synthetic ones and the native control) and hence these differences appear to be an inherent trait of the methanol induced yeast cells. The flow cytometry data was also used to calculate noise levels. For this purpose a squared coefficient of variance was calculated with the eight biological replicates of each strain. No clear difference was found between the synthetic core promoters and the wild type core promoter (data not shown).

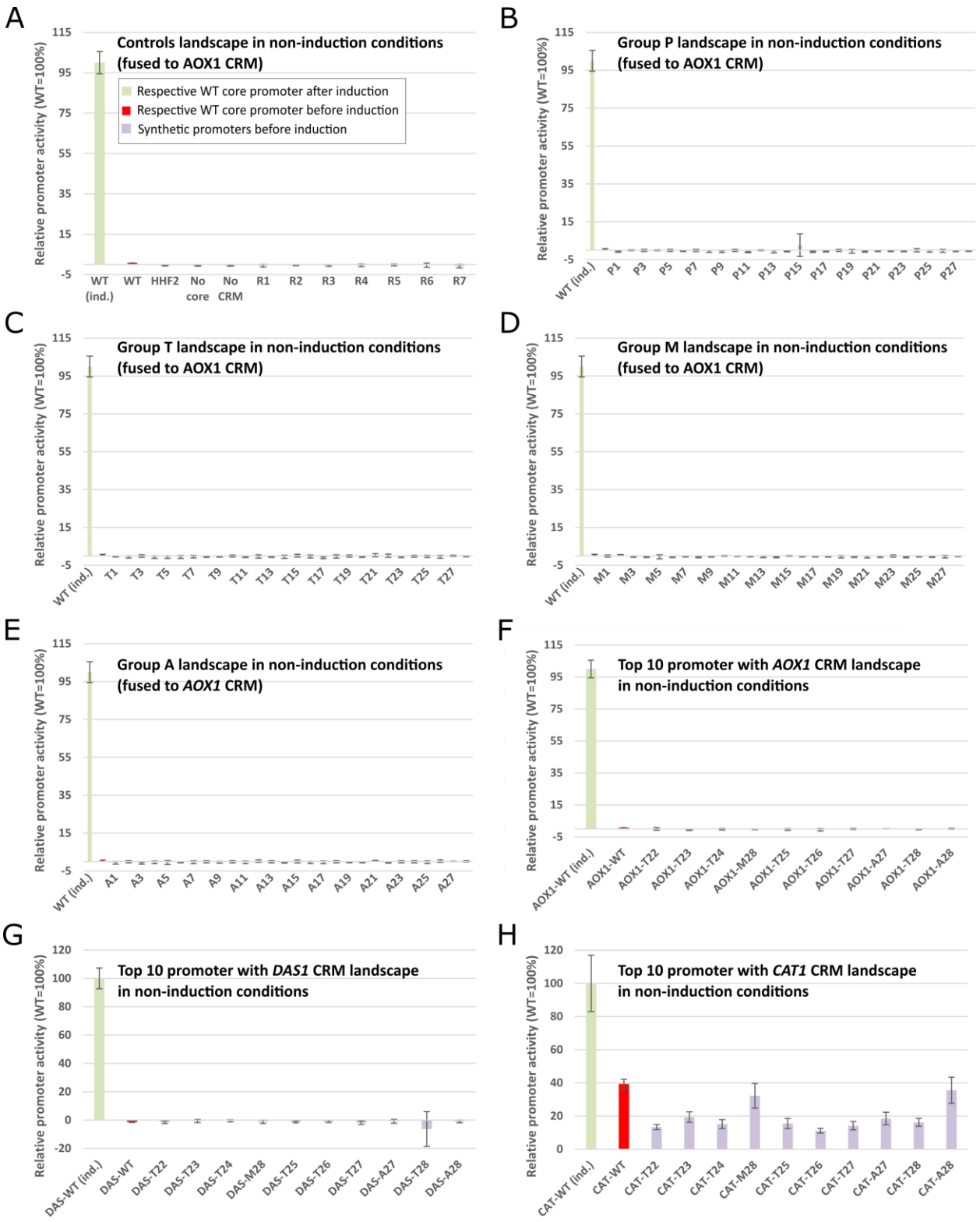
A



B


A
B
C
WT
T22
T23
T24
M28
T25
T26
T27
A27
T28
A28
0
10^1
10^2
10^3
10^4
10^5
FITC-A

Supplementary Figure 7 – **Reporter protein fluorescence measurements under non-inducing conditions yield highly similar results as the respective natural core promoters.** (A) controls fused to the AOX1 CRM, (B-E) four synthetic core promoters' groups fused to AOX1 CRM and (F-H) top 10 synthetic core promoters fused to other inducible CRMs (AOX1, DAS1 and CAT1). The panels shown in this figure are complementary to Figure 2 (panels A and C-F) and Figure 4 (panels B-D) in the main manuscript. In contrast to the main manuscript, where reporter protein fluorescence under methanol-induced conditions is shown, reporter protein fluorescence shown in this figure was measured 60 h after inoculation on BMD media (see methods section). The synthetic core promoters are shown in violet, and the respective wild type core promoter as control is shown in red (before induction) and green (after induction). All values represent single measurements of at least three independent cultivations in separate 96-well deep-well plates. The average and standard deviation of at least three biological replicates are shown. The AOX1 and DAS1 CRMs are tightly repressed under non-inducing conditions [24] and all fusions of synthetic core promoters to the CRMs maintained the same tight repression as the respective wild type core promoters. The CAT1 promoter CRM shows naturally a strong depression profile (expression starts when glucose is depleted [24]). Also the fusions to the synthetic core promoters to the CAT1 CRM maintained this regulatory profile. The relative strengths under depressed (shown here) and induced conditions (shown in main manuscript Figure 4 panel D) vary to a similar degree. These measurements show that the synthetic core promoters control expression strength under different conditions, while leaving the regulatory mode unaffected.


A
Controls landscape in non-induction conditions
(fused to AOX1 CRM)
Respective WT core promoter after induction
Respective WT core promoter before induction
Synthetic promoters before induction
Relative promoter activity (WT=100%)
WT (ind.)
WT
HHF2
No core
No CRM
R1
R2
R3
R4
R5
R6
R7
B
Group P landscape in non-induction conditions
(fused to AOX1 CRM)
C
Group T landscape in non-induction conditions
(fused to AOX1 CRM)
D
Group M landscape in non-induction conditions
(fused to AOX1 CRM)
E
Group A landscape in non-induction conditions
(fused to AOX1 CRM)
F
Top 10 promoter with AOX1 CRM landscape
in non-induction conditions
G
Top 10 promoter with DAS1 CRM landscape
in non-induction conditions
H
Top 10 promoter with CAT1 CRM landscape
in non-induction conditions