Promoter Architecture and Promoter Engineering in *Saccharomyces cerevisiae*

Hongting Tang [1,†]**, Yanling Wu** [1,†]**, Jiliang Deng** [1,†]**, Nanzhu Chen** [1,†]**, Zhaohui Zheng** [1]**, Yongjun Wei** [2] **, Xiaozhou Luo** [1,*] **and Jay D. Keasling** [1,3,4,5,6,*]

1 Center for Synthetic Biochemistry, Shenzhen Institutes for Advanced Technologies, Chinese Academy of Sciences, Shenzhen 518055, China; ht.tang@siat.ac.cn (H.T.); yl.wu@siat.ac.cn (Y.W.); jl.deng@siat.ac.cn (J.D.); nz.chen@siat.ac.cn (N.C.); zh.zheng@siat.ac.cn (Z.Z.)
2 School of Pharmaceutical Sciences, Key Laboratory of Advanced Drug Preparation Technologies, Ministry of Education, Zhengzhou University, Zhengzhou 450001, China; yongjunwei@zzu.edu.cn
3 Joint BioEnergy Institute, Emeryville, CA 94608, USA
4 Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
5 Department of Chemical and Biomolecular Engineering & Department of Bioengineering, University of California, Berkeley, CA 94720, USA
6 Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Kongens Lyngby, Denmark
* Correspondence: xz.luo@siat.ac.cn (X.L.); jdkeasling@lbl.gov (J.D.K.)
† These authors contributed equally to this work.

Abstract: Promoters play an essential role in the regulation of gene expression for fine-tuning genetic circuits and metabolic pathways in *Saccharomyces cerevisiae* (*S. cerevisiae*). However, native promoters in *S. cerevisiae* have several limitations which hinder their applications in metabolic engineering. These limitations include an inadequate number of well-characterized promoters, poor dynamic range, and insufficient orthogonality to endogenous regulations. Therefore, it is necessary to perform promoter engineering to create synthetic promoters with better properties. Here, we review recent advances related to promoter architecture, promoter engineering and synthetic promoter applications in *S. cerevisiae*. We also provide a perspective of future directions in this field with an emphasis on the recent advances of machine learning based promoter designs.

Keywords: promoter architecture; promoter engineering; synthetic promoter; synthetic biology; machine learning; *Saccharomyces cerevisiae*

1. Introduction

S. cerevisiae, as a eukaryotic model organism, has been widely used as a host in synthetic biology because of its clear genetic background, rapid growth, easy cultivation and safety. Currently, yeast cell factories have been developed for the production of recombinant proteins, biofuels, natural and unnatural products, and bulk and fine chemicals [1]. The precise regulation of protein expression is important to balance the intricate metabolic pathways and to ensure a high yield of the desired products. Promoters are the basic transcriptional regulatory elements controlling the quantitative and temporal regulation of protein expression and have been extensively applied to fine-tune the expression of genes in pathway engineering in *S. cerevisiae* [2,3].

In *S. cerevisiae*, endogenous promoters can be classified into two types: constitutive promoters and inducible promoters. Constitutive promoters maintain relatively stable transcription levels with little influence from the intracellular or extracellular stimuli. In order to find strong constitutive promoters

for genetic engineering, many studies have been performed to characterize endogenous promoters according to their expression strength under different growth conditions. Keren et al. [4] measured the activities of 867 native promoters which cover about one-sixth of the *S. cerevisiae* genome under different growth conditions, including various carbon sources. The promoter activities on glucose were comparable to the transcriptome data using single-molecule sequencing [5]. Commonly used constitutive promoters—including P_{TEF1}, P_{TDH3}, P_{PGK1}, P_{TPI1}, P_{CCW12}, and P_{ENO2}—often show slightly different strengths in various studies, which may due to an inconsistent sampling time and experimental conditions [6–8].

Inducible promoters can initiate a dramatic change in the transcriptional level of their corresponding genes upon the presence or absence of specific stimuli. In *S. cerevisiae*, these stimuli, or inducers, range from carbon sources, including glucose, galactose, sucrose, maltose, glycerol, acetate and ethanol, to environmental factors such as the temperature, pH, stress and light, to others, such as metabolites, amino acids, metal ions and hormones [9–11]. Galactose-inducible promoters, including P_{GAL1}, P_{GAL2}, P_{GAL7} and P_{GAL10}, are considered to be strong promoters and have been used to drive a higher enzyme expression level for the efficient production of desired products [8]. By using these promoters, Ro et al. [12] constructed a heterologous pathway for the production of the antimalarial precursor artemisinic acid and Luo et al. [13] achieved the complete biosynthesis of cannabinoids and their unnatural analogues.

The construction of metabolic pathways often involves the expression of multiple genes whose expression levels span several orders of magnitude. The fine regulation of these genes makes it challenging to select the proper promoters. It is usually difficult for endogenous promoters to meet the requirements for the rational design and optimization of metabolic flux. This is mainly because the number of well-characterized promoters is inadequate, their dynamic ranges are poor, and they are often not orthogonal to endogenous regulations. For instance, P_{GAL1} is often used repeatedly for the expression of different enzymes in the construction of metabolic pathways [13], so the gene copy number has to be increased to compensate for the insufficient promoter strength [8], and the utilization of too many galactose-inducible promoters may interfere with the metabolism of galactose due to the depletion of the transcription activator Gal4p [14]. Therefore, to increase the yield of the target products and to reduce the interference from host fitness loss, it is often necessary to develop synthetic promoters for the accurate regulation of multiple genes along the foreign metabolic network. Recently, a wide range of synthetic promoters has been developed to confer constitutive, spatial or temporal gene expressions [15–17]. An in-depth understanding of the constitutive and regulatory mechanisms of promoters is critical for the construction of sophisticated promoters for desired gene expression patterns. In this review, we provide an introduction to the general promoter architecture and the function of its individual elements, and we summarize the design principles and metabolic engineering applications of the synthetic promoters in yeast.

2. Promoter Architecture and Function

S. cerevisiae promoters have multiple essential elements for the accurate transcriptional regulation of genes, including a core promoter region, an upstream activator sequence (UAS), an upstream repressor sequence (URS) and nucleosome-disfavoring sequences, such as poly(dA:dT) sequences, as shown in Figure 1. Each of these elements plays a vital role in tuning promoter activities, thus, it is important to understand the functions and combinatorial regulatory mechanisms of these elements to predict their functions. In order to obtain a comprehensive understanding of the promoter regulatory mechanisms, several tools have been developed, such as YEASTRACT (Yeast search for transcriptional regulators and consensus tracking) and SCPD (*Saccharomyces cerevisiae* promoter database) [18–21].
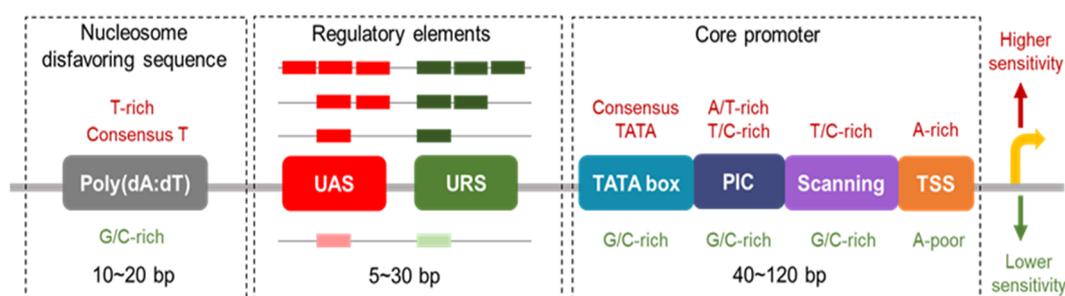
Figure 1. Promoter architecture in *S. cerevisiae*. The crimson rectangle represents the upstream activator sequence (UAS) with a higher activating activity and the pink rectangle is the UAS with a lower activity. Dark green represents the upstream repressor sequence (URS) with a higher repressing activity and light green represents a lower activity. The length of each element is labeled and 40–120 bp is the sequence length between the TATA box and the transcriptional start site (TSS), 5–30 bp is the UAS or URS length, and 10–20 bp is the length of the poly(dA:dT) tracts.

2.1. Core Promoters

The core promoter is the nucleotide sequence that directly interacts with RNA polymerase II (pol-II) and other general transcription factors to form the pre-initiation complex (PIC) to initiate transcription. In metazoans, several conserved functional elements have been identified, including the TATA box, transcriptional start site (TSS), motif 10 element (MTE), downstream promoter element (DPE), and TFIIB recognition element (BRE), while only the TATA box and the TSS region have been identified in *S. cerevisiae*.

TATA boxes are the recognition sites of the TATA-binding protein (TBP), a general transcription factor, with a consensus sequence of TATA(A/T)A(A/T)(A/G) [22]. However, only approximately 19% of all promoters in *S. cerevisiae* contain TATA boxes. Interestingly, previous studies demonstrated that while TATA-less promoters also require TBPs for PIC assembly [23], TATA-containing promoters are highly dependent on TBP-targeted mechanisms which closely correlate to stress responses and these promoters often elicit a higher transcriptional activity [22,24,25]. TATA boxes with different sequence information affect promoter activities [26,27]. For example, Mogno et al. [24] found that the activity of a promoter containing the strong TATA box (TATATAAA) was 2.56-fold higher than the weaker one (CATTTAAA), or 4.9-fold higher than the activity of promoters without any TATA box. It was reported that the TATA box sequence TATAAA was necessary for P_{HIS3} activity and almost all single-base mutations were deleterious to its activity [28]. Besides nucleotide sequences, the location of TATA boxes is also a key determinant of its corresponding promoter activity. The activity of synthetic P_{PDC1} was higher when the TATA box docked between 88 and 66 bp upstream of the TSS, compared to between 65 and 39 bp upstream of the TSS. The TATA box was not functional when it was located 29, 19 or 9 bp upstream of the TSS, or at 19 or 9 bp downstream of the TSS [29]. Lubliner et al. [29] also found that the addition of some random flanking bases around the TATA box affected P_{PDC1} promoter activity, ranging from 24% to 132%. These results demonstrated the impacts of TATA sequences, their location and flanking bases, on the regulation of gene expression, indicating that TATA boxes can be an effective element for fine-tuning promoter activities.

The TSS region is the consensus sequence $A(A_{rich})_5NPy\underline{A}(A/T)NN(A_{rich})_6$ presenting in both TATA-containing and TATA-less promoters, and the transcription is often initiated at the underlined adenosine site [30]. In *S. cerevisiae*, the location of the TSS varies from 40 bp to 120 bp downstream of the TATA box [31,32]. Between the TATA box and TSS region, a PIC region for PIC localization and a scanning region for TSS scanning by pol-II were also discovered in *S. cerevisiae* [29]. The distance variation between the TATA box and TSS mainly affects the length of the scanning region, and the PIC region is relatively constant. For example, P_{GAL1} and P_{GAL10} have different lengths between the TATA box and TSS, which are 84 bp and 114 bp, respectively, while both of their PIC regions span about 20 bp downstream of the TATA box [33]. Even though the length of the PIC

region has little variation, its nucleotide sequence has a strong influence on the promoter activities. Compared to G/C-rich sequences, promoters containing A/T- or T/C-rich PIC regions possess higher activities [29]. A previous study showed that overly long scanning regions have negative effects on the promoter activity, because pol-II may need more time to search the TSS and may fall off anytime during the scanning, and so will cease the transcription initiation. The study also demonstrated that a low T content of the scanning region was negatively correlated with the promoter strength, whereas A-rich sequences overlapping with or slightly downstream of the TSS had positive effects [34].

The core promoter is one of the major determinants for the regulation of gene transcriptional levels in *S. cerevisiae*, and each of the abovementioned regions can affect the strength of its corresponding core promoter.

2.2. UAS and URS

The upstream activating sequence (UAS) is located upstream of the core promoter and serves as a binding site for specific transcription activators. The UAS is a crucial region of promoters which enhances gene expression. While the core promoter is responsible for PIC recruitment and assembly, the UAS provides additional stability and regulation of PIC formation [35]. All promoters recognized by pol-II may require one or more UASs for regulated gene expression [36,37]. For example, galactose-inducible promoters have various conserved UAS_{GAL} sites, a 17-bp consensus sequence 5′-CCGNNNNNNNNNNNNCGG-3′ that is recognized by the transcription activator Gal4p to control their expression activities [38]. The promoters P_{GAL1}, P_{GAL2} and P_{GAL10} contain four UAS_{GAL} sites, while P_{GAL7} only has one. During the induction of galactose, the production of Gal4p can improve the P_{GAL1} activity 1000-fold [37]. The upstream repressing sequence (URS) is a binding site of transcription repressors which inhibits the promoter activities. For instance, the consensus sequence 5′-SYGGGG-3′ is the recognition site of the transcription repressor Mig1p which is involved in glucose repression [39]. The activities of 5′-SYGGGG-3′ containing promoters, such as P_{GAL1}, P_{HXT2}, P_{SUC2}, P_{JEN1}, are inhibited in the presence of glucose [40–43]. Common UASs and URSs in *S. cerevisiae* are summarized in Table 1. The length of the UAS and URS vary from 5 to 30 bp and are typically 10 bp in both prokaryotic and eukaryotic organisms, including *S. cerevisiae*, because overly long sequences may have more mutational risks for their inactivation while too-short sequences may result in nonspecific genomic binding [44].

Table 1. Common UASs and URSs in *S. cerevisiae*.

UAS Sequence (5′-3′)	Transcription Factor	Promoters	Inducer	Function	Reference
CGGRNNRCYNYNCNCCG	Gal4p	*GAL1/2/7/10, MEL1*	Galactose	Regulation of galactose metabolism	[38]
ATGACTCTT	Gcn4p	*ARG1, ARG4, HIS4, CPA2*	Amino acid starvation	Regulation of amino acid biosynthetic genes	[45]
TTACTAA	Yap1p/2p	*GSH1, TRX2, YCF1, GLR1*	Oxidative stress such as H_2O_2	Regulation of genes expressed in response to environmental changes	[46]
TG(A/C)GCCNC	Crz1p	*PMC1, PMR1, FKS2*	Calcium	Calcineurin-responsive transcription factor	[47,48]
CGGNBNVMHGGA	Cat8p	*FBP1, PCK1, ACR1, IDP2, JEN1*	Non-fermentative growth conditions	Derepression of gene expression under non-fermentative growth conditions	[43]
PyPuCACCCPu	Aft1p	*FRE1, FTR1, FTH1*	Iron deprivation	Iron utilization and homeostasis	[49]
TGAAAC	Ste12p	*TEC1, FLO11*	Pheromone	Involved in mating and invasive growth	[50]
GAATGT	Tec1p	*TEC1, FLO11*	*n*/A	Ste12p cofactor	[50,51]
CAGCGTG	Hac1p	*KAR2, PDI1, EUG1, FKB2*	Unfolded/misfolded proteins	Regulates the unfolded protein response	[52]
NGAAN	Hsf1p	*HSP82, HSP26, HSP104, HSP26,*	Heat shock	Trimeric heat shock transcription factor	[53]
SYGGGG	Mig1p	*GAL1, HXT2, SUC2, JEN1*	Glucose	Involved in glucose repression	[54]
TGACGTCA	Aca1p	*GRE2, COS8*	*n*/A	Basic leucine zipper (bZIP) transcription factor involved in carbon source utilization	[55]

Table 1. *Cont.*

UAS Sequence (5'-3')	Transcription Factor	Promoters	Inducer	Function	Reference
CGGN₃TNAN₉₋₁₂CCG	Oaf1p	*CTA1, FOX1/2/3, FAA2, PAS8, PAS10*	Oleate	Involved in fatty acid beta-oxidation	[56]
TCCGCGGA	Pdr1/3p	*SNQ2, PDR5*	Pleiotropic drug	Pleiotropic drug response	[57–59]
GGTGGCAAA	Rpn4p	*RPT2/3/6*	Patulin	Regulation of proteasome genes	[60,61]
DNCRCAAAW	Ndt80p	*SMK1, SPR3*	Sporulation	Required for full meiotic recombination and middle sporulation	[62]
CCAAT	Hap4p	*CYC1*	Heme	Global regulator of respiratory gene expression	[63]
TGACGTCA	Sko1p	*SUC2, MSN2, ROX1, PTP3*	Osmotic stress	Involved in osmotic and oxidative stress responses	[40,64]
GcCTCGA(G/A)G(C/A)g(a/g)	Xbp1p	*CLN1, CYS3, SMF2*	Stress or starvation	Transcriptional repressor	[65]
CAC(A/G)T(T/G)	Pho4p	*HIS4, PHO5*	Phosphate limitation	Regulation of the purine and histidine biosynthesis pathways	[66]
ACCYYNAAGGT	Zap1p	*ZRT1, ZRT2*	Zinc	Zinc-regulated transcription factor	[67]
ACTACTA(T/A)₄TAG	Smp1p	*STL1, CWP1*	Osmotic stress	Osmotic stress response	[68]
CTA(T/A)₄TAG	Rlm1p	*HKR1, KTR2, HSP150, FLO1*	n/A	Maintenance of cell integrity	[69]
TTGGRG	Adr1p	*ADH2, ALD4, ALD6, POX1*	n/A	Carbon source responsive transcription factor	[70]
AATCA-N₈-TGAYT	Vhr1p	*VHT1, BIO5*	Biotin	Response to low biotin concentrations	[71]
AAACTGTGG	Met31p	*MET25, MET14, MET3*	n/A	Sulfur amino acid metabolism	[72]
CCCCT	Msn2/4p	*CTT1, DDR2, HSP12*	Various stress	Response to multiple stress conditions	[73]
CCRTYCRTCCG	Sip4p	*FBP1, PKC1, ICL1*	n/A	Positive regulation of gluconeogenesis	[74]
CGGANNA	Rgt1p	*HXT2, HXT4*	Glucose	Glucose-responsive transcription factor	[75,76]
CTTCC	Gcr1p	*ENO1, TPI1, TDH3*	n/A	Transcriptional activator involved in the regulation of glycolysis	[77]
RRRTAACAAGAG	Rox1p	*HEM13, COX5B, ANB1, CYC7*	Heme	Heme-dependent repressor of hypoxic genes	[78]

Note: *n*/A, not available.

The binding affinity, quantity, and location of the UAS and URS affects promoter activities. The sequence variations of UAS and URS result in different binding affinities with their corresponding transcription factors and this has been studied in great detail using site-directed mutagenesis [79–81]. UAS$_{GAL}$ is characterized by the presence of the CGG triplets at both ends, which are separated by 11 bp nucleotides. This allows for the existence of a total of 4^{11} theoretical UAS$_{GAL}$ sites with different binding affinities. It was reported that the four UAS$_{GAL}$ sites of *S. cerevisiae* P$_{GAL1}$ demonstrated different activities: UAS$_{GAL}$4 had the lowest activity, while UAS$_{GAL}$2 and UAS$_{GAL}$3 showed the highest activities [38]. It was reported that the affinities of UAS$_{GAL}$ measured in an in vitro assay were inconsistent with the in vivo data [80]. Thus, the UAS$_{GAL}$ mutant library must be characterized in vivo to understand the properties of the UAS$_{GAL}$ sites. The saturation mutagenesis of the transcription factor Gcn4p's binding site (5'-ATGACTCTT-3') within the *HIS3* promoter found that almost all mismatch mutants reduced the P$_{HIS3}$ activity significantly and only one mutant with the sequence 5'-ATGACTCAT-3' increased the binding affinity of Gcn4p and improved the P$_{HIS3}$ activity [82]. It has been shown that regulatory regions containing multiple UAS or URS sites for binding the same transcription factor could enhance their activation or repression of transcription. In a test of 15 transcription factors, such as Gal4p, Gcn4p, Bas1p, increasing the number of their UAS sites improved promoter activities; similarly, promoters with multiple URS sites showed a stronger repression, such as Matα2p-Mcm1p. It has also been shown that this accumulation effect will saturate in the presence of a certain number of UASs or URSs. [83,84]. It is known that UAS sites are often docked several hundred base pairs upstream of the core promoter in *S. cerevisiae* [85]. Previous studies showed that promoter activity decreased with the increasing distance of the UAS site from the core promoter in

S. cerevisiae [83,86]. Thus, there are clear relationships between the binding affinity, quantity, location of these two regulatory sequences and the transcription level of their corresponding promoters.

The coexistence of various UASs or URSs in the same promoter could bring about the combinatorial and dynamic regulation of its transcription. Promoters of many genes related to carbon source metabolism have both URS sites, which are commonly suppressed by glucose, and UAS sites induced by other carbon sources. For example, the *GAL1* promoter contains a URS_{MIG} and four UAS_{GAL}. URS_{MIG} mainly represses the activity of *GAL1* promoters under glucose growth conditions, while UAS_{GAL} induces the activity of *GAL1* promoters under galactose growth conditions. Under the condition of glucose and galactose fermentation, the regulation of *GAL1* promoters correlated to the ratio of glucose and galactose [87]. In another case, the promoter of Ime1p, a transcription factor that participates in meiosis, has a more complex regulation mechanism which is regulated by at least six URSs and four UASs for binding multiple transcription factors such as Msn1p, Msn2p, Rem1p, Sok2p, Yhp1p, and Sum1p [86]. Even though many methods have been developed for analyzing the function and interaction of UASs and URSs, the understanding of the synergistic regulation mechanisms among these different types of UAS and URS sites remains incomplete, which leads to difficulties in predicting the exact synthetic promoter phenotypes [88–92].

2.3. Nucleosomes Disfavoring Sequences at Gene Promoters

Promoters' chromatin structure plays an essential role in transcription regulation. A low nucleosome occupancy facilitates transcription and improves mRNA abundance, while increasing the nucleosome occupancy of promoters tends to decrease their mRNA levels [93,94]. Previous studies showed that a high nucleosome occupancy can be found in numerous open reading frames and the promoters of some genes whose expression was repressed, while a low nucleosome occupancy was observed in the promoters of genes with a high expression [95–97]. Many studies revealed that a low nucleosome occupancy facilitates the binding of the transcription factor to the regulatory DNA sequences for the regulation of promoter activity [98,99]. For example, Gal4p interacted with its naked DNA binding sites at picomolar concentrations, while at least 100-fold more Gal4p was required to bind to the nucleosomal DNA [100,101].

Nature has evolved two main strategies to decrease nucleosome occupancy in order to regulate gene expression. One strategy involves the assistance of transcription factors, such as Rap1p, Reb1p, Abf1p and the SWI/SNF complex, which can release nucleosomes from DNA [83,101,102]. It was reported that the transcription factor Rap1p not only depleted the nucleosome from its own binding site of the *HIS4* promoter, but also reduced a nearby nucleosome to increase the accessibility of other transcription factors, including Gcn4p, Bas1p, Bas2p [102]. Another strategy is the distribution of the nucleosome-disfavoring sequences, which results in an efficient accessibility for the transcription factors [103,104]. For instance, the *PHO5* promoter has two binding sites of transcription factor Pho4p. One is a low-affinity site located in a nucleosome-free region and another is a high-affinity site occupied by a nucleosome. Under induction conditions, Pho4p interacts with the low-affinity binding site first rather than the high-affinity binding site [105,106].

The poly(dA:dT) tract, a homopolymeric stretch of deoxyadenosine nucleotides, is a well-known nucleosome-disfavoring sequence in eukaryotic organisms. Its length ranges from 10 to 20 bp, or is even greater in some cases [107]. The poly(dA:dT) tract has a low-affinity for nucleosome formation which results in nucleosome-free regions or a low nucleosome occupancy to stimulate transcription. In general, the poly(dA:dT) tract is considered an upstream activating element, not through its interaction with transcription factors, but by depleting nucleosomes [104]. Many native promoters in *S. cerevisiae* are controlled by poly(dA:dT) tracts, such as P_{HIS3}, P_{URA3}, P_{ADH2}, P_{RPS28A} [108–111]. Modification of the poly(dA:dT) tract by changing its length, sequence information and location next to transcription factor binding sites will affect gene expression [83,112,113]. For example, P_{RPS28A} contains a poly(dA:dT) tract located 7 bp downstream of the transcription factor Abf1p binding site. Mutants of this poly(dA:dT) tract had a lower transcription of the *RPS28A* gene and the nucleosome moved closer toward the

Abf1p binding site [110]. Raveh-Sadka et al. [113] has systematically analyzed the function of the poly(dA:dT) tract on promoter activities and the results showed that perfect poly(dA:dT) tracts increased transcription compared to those with two mismatches, a longer poly(dA:dT) tract with 22 bp worked better than short tracts, and an increase of the distance between the poly(dA:dT) tract and transcription factor binding site decreased the promoter activities. Thus, engineering nucleosome-disfavoring sequences, such as poly(dA:dT) tracts, may be an effective strategy for the construction of synthetic promoters with high activities to fine-tune gene expression.

3. Promoter Engineering Approaches

Several different methods, such as a random mutagenesis by error-prone PCR, saturated mutagenesis, hybrid-promoter engineering, have been used for promoter engineering and have been reviewed previously [114–116]. More recently, with the introduction of machine learning into synthetic biology, great progress has been made in predicting the structure, function and interactions of biological macromolecules such as nucleic acids and proteins, and the procedure is shown in Figure 2 [117–120]. Based on big data of promoter–protein interactions, machine learning provides a new strategy for rational design and increases the predictability of promoter engineering.
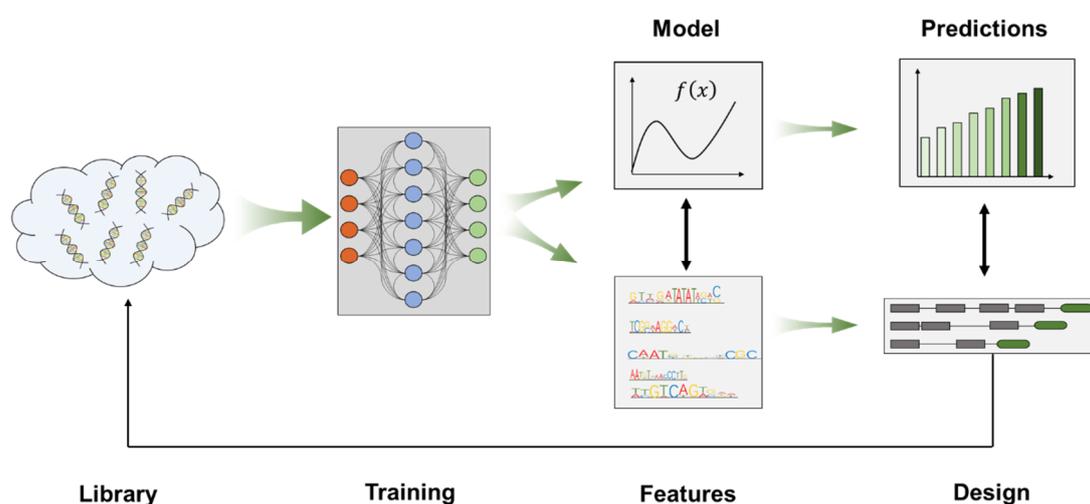

Figure 2. Machine learning procedures including database building, training, feature extraction (modeling), design (prediction) and verification.

Currently, machine learning is mainly used for promoter engineering in *Escherichia coli*. De Mey et al. [121] applied a partial least squares (PLS) regression method to analyze the relationship between promoter sequences and strength in *E. coli*, demonstrating the prospects of predictive and rational promoter designs. However, the accuracy for the prediction still needs improvement. Artificial neural network (ANN) models can better represent the complex and nonlinear interactions within promoter sequences and have been successfully applied in the rational design of promoters. A series of promoters with different strengths was obtained by randomly mutating the Trc promoter and its ribosome binding site in *E. coli*, which were then used for training and testing using the ANN mathematical model. Sixteen novel artificial elements were generated in silico and their predicted expression levels showed a good correlation with the experimental results, indicating that the model could be used for synthetic promoter designs with specific properties [122]. The generative adversarial network (GAN) and convolutional neural network (CNN) were also used to generate artificial promoters *de novo* and predict their expression levels, and three highly active synthetic promoters were identified in *E. coli*, among the predicted strong promoters, by experimental validation [123]. In *S. cerevisiae*, promoter libraries were created based on native promoters, including the constitutive P_{TDH3} and the inducible P_{ZEV}, and a reliable prediction model was trained with the CNN based on the promoter sequence-activity data

collected from these libraries, and the activity of a synthetic promoter (predicted from the *TDH3* promoter) increased by 37% and the activity of a *ZEV* promoter mutant also increased by *β*-estradiol induction, and its basal expression was reduced [124]. However, at present, few studies have applied machine learning in promoter engineering in *S. cerevisiae*.

Libraries created by random or saturated mutagenesis are dependent on the transformation efficiency of the strain. The maximal transformation efficiency of *S. cerevisiae* is about 10^8, therefore the library capacity will not exceed the transformation efficiency. Hybrid-promoter engineering requires a lot of testing, and it is difficult to achieve a high-throughput selection due to the heavy construction workload. Machine learning could construct quantitative models based on a limited database to analyze the data distribution characteristics of the designed promoter library, helping us better understand the underlying interaction principle. The result of machine learning is to predict a serial of promoters that function as expected. This predictability can reduce the workload and increase the accuracy of testing. Therefore, it is foreseeable that machine learning methods will gain momentum in the near future and fuel the development of accurate and customer-tailored engineered promoters.

4. Promoter Engineering for Diverse Synthetic Promoters and Their Applications

4.1. Synthetic Promoters for Expanding Dynamic Ranges

Multiple enzymes are often introduced and overexpressed in metabolic engineering to boost the yield of the desired product. However, an unbalanced expression of these enzymes would accumulate intermediates and result in unnecessary metabolic burdens or toxicities [125]. Therefore, it is necessary to precisely control enzymatic activities with the help of promoters with a wide dynamic range to ensure a balanced flux for pathway optimization. [126]. However, this field is often limited in techniques to construct a promoter set with a wide dynamic range. Currently, many studies have been focused on endogenous promoter engineering to expand the yeast promoter library to overcome these limitations. A random mutagenesis library of existing promoters has proven to be an effective method for the construction of synthetic promoters (Figure 3a). Alper et al. [15] created a library based on the *TEF1* promoter and obtained a series of synthetic promoters with a wide range of activities; the best candidate showed a two-fold higher activity than the native P_{TEF1}. These promoters were used to regulate efficient glycerol production by driving the rate-limiting enzyme expression in *S. cerevisiae* [127]. Other native promoters such as P_{ENO2} and P_{PDC2} were also engineered by random mutagenesis. Synthetic P_{ENO2} and synthetic P_{PDC2} obtained from their corresponding mutagenesis libraries improved the expression of recombinant proteins cellobiose transporter and β-glucosidase for cellobiose degradation by 24.4-fold and 3.0-fold, respectively. The recombinant strain had a higher cellobiose consumption and ethanol production than its parent strain, reported as 6.41-fold and 6.36-fold more, respectively [128].
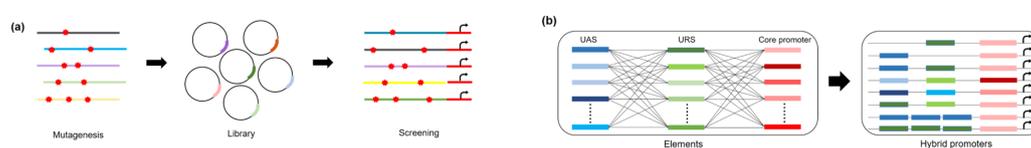

Figure 3. Promoter engineering for the development of synthetic promoters with dynamic ranges in *S. cerevisiae*. (**a**) Engineering promoter by random mutagenesis; (**b**) Combination of each element for hybrid promoters.

The rational combination of different promoter elements is another efficient method to improve the dynamic range of promoters (Figure 3b). Blazeck et al. [16] created a synthetic hybrid promoter by combining the strong *TDH3* promoter with three UAS_{CLB}s, the 240-bp UAS sequence of the mitotic cyclin (CLB2), and this hybrid promoter enhanced the transcription level 2.5-fold compared to *TDH3* promoter. In addition, they built a series of heterozygous galactose-inducible promoters by fusing UAS_{GAL} with

different core promoters and the resulting synthetic promoters had a continuous strength gradient which could achieve a fine-tuned gene expression. A series of post-diauxic phase inducible synthetic promoters were also constructed by combining core promoters and UASs, and their activities were further improved by the optimization of the UAS number [129]. More recently, engineering the nucleosome architecture of promoters has received increasing attention. Synthetic promoters with different strengths were produced by altering the properties of the poly(dA:dT) tract, including making changes to its length, composition and distance from the UAS [113]. The introduction of nucleosome-disfavoring sequences into promoters, such as P_{CYC1}, P_{HIS5}, P_{HXT7}, and P_{TEF1}, increased the strength of the resulting synthetic promoters. However, this method did not work on strong promoters such as P_{TDH3} and P_{GAL1}, which may have been evolutionarily optimized for their nucleosome architecture in nature [17]. Based on such designs, synthetic promoters could acquire some expected properties such as a high activity and inducibility.

4.2. Synthetic Promoters for Reducing Homologous Recombination

It is well-known that *S. cerevisiae* has a strong homologous DNA recombinant capacity, which has been widely utilized in genome insertion, deletion and replacement. Multiple uses of the same promoters or their elements are prone to generating homologous recombinations and lead to the instability of synthetic pathways in *S. cerevisiae*. Thus, the development of synthetic promoters with sequence orthogonality to avoid homologous recombination is an attractive research field. Exploiting the naturally evolved diversity of heterologous promoters in *S. cerevisiae* is an effective strategy. Peng et al. [130] compared 11 galactose-inducible promoters from *Saccharomyces* species and found that most of them are stronger than ScP$_{GAL1}$, especially SeP$_{GAL2}$ and SkP$_{GAL2}$ from *Saccharomyces eubayanus* and *Saccharomyces kudriavzevii*, respectively. Recently, the design of artificial promoters without the utilization of a native promoter as a scaffold has been exploited. Based on a computational approach, Curran et al. [17] predicted active promoter sequences and created six artificial promoters with a 20-fold dynamic range of transcription. The activities of these synthetic promoters were comparable to native promoters P_{CYC1} and P_{HXT7}. These promoters, although still small in number, will be useful for synthetic biology applications, especially for industrial applications which are more dependent on stable gene expression.

4.3. Synthetic Promoters with Minimal Size

In bacteria, such as *E. coli*, the lengths of promoters are typically less than 100 bp, whereas native yeast promoters usually span hundreds of nucleotides. The long nucleotide sequences not only decrease the efficiency of biosynthetic pathway construction, but also complicate the regulation of these pathways. The construction of minimal promoters could overcome these limitations. The truncation of endogenous promoters to remove non-essential bases is one strategy for minimal promoter construction (Figure 4a) [131,132]. *S. cerevisiae* P_{TEF1}, a strong promoter commonly used in recombinant expression, was used as a model for the study of minimal promoter constructions. The results indicated that a 69 bp essential sequence can sustain detectable transcriptional activities. A series of short synthetic promoters were developed through a combination of a UAS and this short essential sequence, which achieved an 80% activity of P_{TEF1} [132]. However, the minimal promoters produced by this method also contain elements from endogenous promoters and suffer from the risk of homologous recombination. Thus, saturation mutagenesis may be a preferable method for creating minimal promoters (Figure 4b) [133,134]. In *S. cerevisiae*, to achieve minimal core promoters, sequences with different sizes (20 bp, 25 bp and 30 bp) between the TATA box and TSS were selected for saturation mutagenesis, and it was found that only the N30 library generated partially active promoters. Thirteen (13) of these functional core promoters were isolated. Furthermore, minimal constitutive UASs of 10 bp were also selected by a similar process. By combining the minimal core sequences and minimal constitutive UASs, minimal promoters were created and the activity of one of them reached 70% of the strong P_{TDH3}, with approximately 20% of its original length. In addition, minimal galactose-inducible

promoters were also developed by using these minimal core promoters and their activities were comparable to wild type P_{GAL1} [135]. Synthetic promoters with minimal sizes work well in *S. cerevisiae* and may have strong potential in large-scale synthetic biology applications.
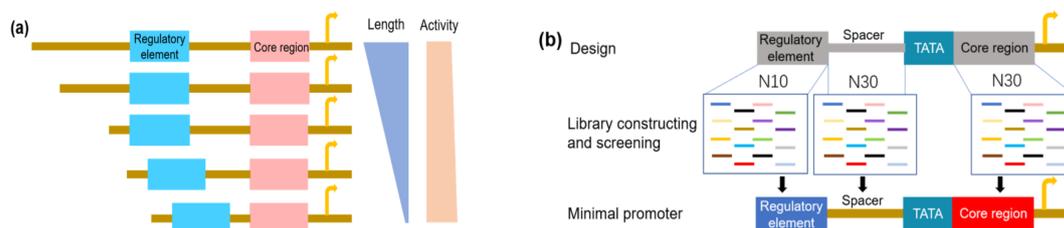

Figure 4. Methods for minimal promoter construction. (**a**) Rational truncation of endogenous promoters; (**b**) Artificial construction *de novo* by constructing and screening the promoter library. N30 and N10 mean the random mutation libraries of 30 and 10 consensus nucleotides, respectively.

4.4. Synthetic Promoters for Multi-Host Application

Commonly used microbial hosts in synthetic biology, such as prokaryotic *E. coli* and *Bacillus subtilis* and eukaryotic *S. cerevisiae* and *Pichia pastoris*, often have certain genetic features which make them suitable hosts for specific genetic circuits and biosynthetic pathways. However, the characterization of these circuits and pathways in different hosts always needs promoter substitution. Thus, the creation of broad-spectrum synthetic promoters for applications in diverse hosts could expand the synthetic biology toolbox to avoid promoter redesign and facilitate host selection [136]. The construction of synthetic, broad-spectrum promoters requires inter-species combinations of all basic and efficient promoter elements to enable the transcription of downstream genes in different hosts. For example, the strong synthetic minimal promoter of *S. cerevisiae* and the conserved −35 and −10 boxes from *E. coli* and *B. subtilis* were rationally combined to develop P_{bs}, which could be used to drive gene expression in all three hosts. The activity of P_{bs} was much stronger than the *E. coli* strong promoter P_{J23119}, comparable to the *S. cerevisiae* minimal promoter P_{min}, and approximately 75% of the *B. subtilis* strong promoter P_{cdd}. UAS mutations renders this promoter with an activity gradient, which is useful in biosynthetic pathway optimization [137]. Based on the information of the nucleotide distribution of *S. cerevisiae* core promoters, Portela et al. [138] successfully constructed a series of universal core promoters utilizing computational designs and library selection, and these core promoters could be used in different yeast species, including *S. cerevisiae* and *P. pastoris*. The reported broad-spectrum promoters have been used in several hosts, and more promoters of this type with a dynamic activity range are required to simultaneously drive gene expression in more diverse hosts, which would facilitate the examination of synthetic pathways in different hosts.

4.5. Synthetic Promoters for Constructing Biosensors

To improve the robustness of biological pathways, it is essential to design a feedback control network for gene expression regulated by metabolic intermediates or other stress factors. Gene circuits with feedback regulation usually have one or more biosensors which can respond to physical or chemical signals to realize the real-time dynamic transcription regulation, and thus automatically adjust the state of gene expression. Two key components of biosensors are transcription factors, which can detect either intracellular or environmental signals, and promoters, which can receive signals from effectors and generate an output. Thus, promoter engineering is a common method in biosensor development and optimization.

Most native promoters have some defects in terms of being a part of an excellent biosensor, therefore, engineering native promoters could increase the sensitivity and the overall performance of biosensors. Cytosolic NADPH/NADP⁺ ratios are important to maintain redox homeostasis and cell fitness. The activity of the native *TRX2* promoter, which is regulated by the transcription factor Yap1p,

can be altered by sensing NADPH/NADP$^+$ ratios, but its sensitivity is low. Increasing the number of the UAS-containing Yap1p binding sites greatly enhanced the cascade response effect, and this novel biosensor was useful at selecting cell populations with higher NADPH/NADP$^+$ ratios [139]. Similarly, engineering the *YGP1* and *CCW14* promoters by optimizing the core promoter and the number and type of UASs successfully created a set of strong synthetic promoters for sensing low pH; these low-pH sensing promoters improved the production of lactic acid by 10-fold under low-pH fermentation compared to the native *TEF1* promoter [140]. In addition, engineering transcription factors could also change the promoter performance. In the galactose regulatory network, Gal3p responds to galactose and activates the transcription activator Gal4p to induce the transcription of galactose-inducible promoters. Gopinarayanan et al. [141] found a Gal3p mutant that could sense xylose and then regulate all galactose-inducible promoters under xylose growth conditions, allowing a better utilization and growth when using xylose as the sole carbon source.

The utilization of heterologous regulatory elements can increase the number and types of biosensors which do not interfere with the native cellular regulation of *S. cerevisiae*. The bacterial tetracycline operator (*tetO*), a DNA binding sequence of tetracycline-inducible repressors (*tetR*), has been used in numerous biosensor applications. Hybrid promoters created by the fusing of one or more *tetO* to a yeast native promoter, such as a *CYC1* promoter, were controllable under different tetracycline concentrations, and the overexpression levels of β-galactosidase comparable to P$_{GAL1}$ are reached [142]. The number and location of the heterologous regulatory DNA sequence affect the induction ratio and the dose-response curve of biosensors, and thus their engineering is key for sensitive and robust biosensor construction. An analysis of the *tetO* locations between the TATA box and TSS found that the closer the *tetO* was located to the TATA box, the stronger the transcriptional repression; the location and number of the *tetO* together determined the dose-response curve [143]. Camphor is an inexpensive small molecule that binds to the repressor CamR, a distant homolog of the Tet repressor, to inhibit gene expression. An URS site for binding CamR was embedded within the core promoter of P$_{CYC1}$, and a camphor-off switch was successfully constructed, which activated gene expression without camphor and repressed gene expression through the addition of micromolar concentrations of camphor, and this sensor was applied to complement the adenine-auxotrophy in a camphor-dependent manner [144]. The bacterial FapR transcriptional repressors and their cognate *fapO* DNA binding sites have also been widely used to develop various biosensors for the detection of many metabolites in *S. cerevisiae*, such as fatty acid intermediates, such as malonyl-CoA [145,146]. By combining malonyl-CoA sensors with a genome-wide overexpression library, the titer of 3-hydroxypropionic acid produced from malonyl-CoA was enhanced by 120% [146]. The successful application of heterologous elements into promoters endowed the biosensor with new properties to respond to more metabolites or chemicals, which makes it able to be used for the accurate regulation of gene expression in biosynthetic pathways.

The rational combination of different regulatory elements enables signal integration to perform the combinatorial effects of biosensors. For instance, five binding sites of the androgen receptor, which respond to steroid hormones, were placed upstream of the TATA box of the *CYC1* promoter. A lactose inhibitor LacI binding site was positioned downstream of the TATA box, and the hybrid promoter led to a wide range of dual-mode promoter outputs under the regulation of testosterone and isopropyl β-ᴅ-1-thiogalactopyranoside IPTG [147]. In a similar vein, a combination of multiple promoters with different functions can form a sophisticated biosensor system. Quorum sensing was tuned by the pheromone-responsive *FUS1* promoter to sense α-pheromone levels. The *ARO9* promoter was responsive to aromatic amino acids in the cultivation medium and was used to drive α-pheromone expression, and thus the pheromone quorum sensing could be fine-tuned by the aromatic amino acid concentration [148]. The synergistic effect of multiple regulatory elements or promoters contributed to the multilayer regulation of biosensors.

In conclusion, numerous synthetic promoters have been engineered and their characteristics are shown in Table 2.

Table 2. The characteristics of engineered promoters.

Application	Note (Elements or Parts)	Approach	Expression Range (fold)	Product/Inducers	Reference
Expanding dynamic ranges	P_{TEF1}	Random mutation	0–2.0	*n*/A	[15]
	P_{TEF1}	Random mutation	0.08–1.2	Increase glycerol 3-phosphate dehydrogenase activity	[127]
	P_{ENO2}; P_{PDC1}	Random mutation	24.4; 3.0	Obtain a higher cellobiose consumption rate (6.41-fold) and ethanol productivity (6.36-fold)	[128]
	$UAS_{CLB(3X)}$-P_{TDH3}; UAS_{GAL1}-P_{LEUM}/P_{CYC}/P_{GAL1}	Hybrid	2.5; 50-fold dynamic range	*n*/A	[16]
	$UAS_{ENO2(3X)}$-P_{TEF1}; UAS_{HXK2}-P_{TEF1}/UAS_{HSP30}-P_{TEF1}	Hybrid	2.0; 8-fold induction range	UAS_{HXK2}-P_{TEF1} and UAS_{HSP30}-P_{TEF1} are post-diauxic phase-induced promoters	[129]
	P_{HIS3}	Manipulating poly(dA:dT) tracts	3-fold dynamic range	*n*/A	[113]
	P_{CYC1}	Tuning of nucleosome architecture	6.0	*n*/A	[17]
	PTDH3	Machine learning	1.37	*n*/A	[124]
Reducing homologous recombination	Galactose-inducible promoters	Heterologous expression	2.5-fold to 99-fold induction ratio	Producing 11.5 mg/L lycopene	[130]
	Psynth promoters	*De novo*	20-fold dynamic range	*n*/A	[17]
Minimal promoters	UAS_A/UAS_C/UAS_{FEC}, P_{TEF1}	Truncation and hybrid	20-fold dynamic range	*n*/A	[132]
	UAS_{EXP1}/UAS_{GPD}, P_{N30}	Saturation mutagenesis and hybrid	*n*/A	5.5-fold enhancement of lycopene–carotene transformation; producing *β*-carotene 7.4 mg/g DCW	[133]
	UAS_{N10}; P_{N30}	*De novo* by saturation mutagenesis	0.7	achieve 70% of the strength of the strongest *TDH3* promoter	[135]
Multi-host suitable	P_{min}	Random mutation and hybrid	*n*/A	P_{bs} was much stronger than E. coli P_{J23119}; 75% of that of P_{cdd} in B. subtilis; lower than that of the strong promoter P_{TDH3}	[137]
	CRM; P_{AOX1}	Computational design and hybrid	200-fold dynamic range	0.3% to 70.6% of the wild type P_{AOX1} level	[138]
Biosensor	P_{TRX2}	Hybrid	100-fold dynamic range	NADPH/NADP+ ratio	[139]
	P_{YGP1}; P_{CCW14}	Hybrid	6.0; 16.0	Enabling a 10-fold increased production of lactic acid; low pH	[140]
	P_{CYC1}	Hybrid	1000-fold induction ratio	Tetracycline	[142]
	P_{CYC1}	Hybrid	*n*/A	Camphor	[141]
	P_{GAL1}	Hybrid	*n*/A	Fatty acid/fatty acyl-CoA	[145]
	P_{GPM1}	Hybrid	1-fold to 4.17-fold induction ratio	Enhancing 3-hydroxypropionic acid titer by 120%; Malonyl-CoA biosensor	[146]
	P_{CYC1}	Hybrid	8-fold induction ratio	IPTG and testosterone dual induction	[147]

Note: *n*/A, not available.

5. Perspective

Although numerous synthetic promoters have been created to fine-tune gene circuits and metabolic pathways for a greater compatibility and production improvements, the complexity of biological systems still require more sophisticated and elaborate artificial promoters. Although studies have focused on understanding promoter architecture, there remains much to be learned about the interactions of multiple promoter elements to make promoter engineering easier. Promoter designs of the past were unpredictable, requiring laborious screening and testing, which is expensive and time-consuming. In addition, although multiple synthetic promoters with stronger activities were created, they did not exceed the strength of P_{GAL1}. Machine learning provides a new design method for synthetic promoters and helps to further understand the mechanism of genome-wide gene expression regulation. Machine learning models may accurately predict synthetic promoter activities and quickly lead to target characteristics. As such, machine learning will likely become a powerful tool for promoter engineering and synthetic biology as a whole.

Author Contributions: Writing, H.T., Y.W. (Yanling Wu), J.D. and N.C.; proof reading, H.T., Z.Z., Y.W. (Yongjun Wei), X.L. and J.D.K. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: X.L. has a financial interest in Demetrix. J.D.K. has a financial interest in Amyris, Lygos, Demetrix, Constructive Biology, Napigen, and Maple Bio.

References

1. Nevoigt, E. Progress in Metabolic Engineering of *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 379–412. [CrossRef]

2. Scalcinati, G.; Knuf, C.; Partow, S.; Chen, Y.; Maury, J.; Schalk, M.; Daviet, L.; Nielsen, J.; Siewers, V. Dynamic control of gene expression in *Saccharomyces cerevisiae* engineered for the production of plant sesquitepene α-santalene in a fed-batch mode. *Metab. Eng.* **2012**, *14*, 91–103. [CrossRef] [PubMed]

3. Latimer, L.N.; Lee, M.E.; Medina-Cleghorn, D.; Kohnz, R.A.; Nomura, D.K.; Dueber, J.E. Employing a combinatorial expression approach to characterize xylose utilization in *Saccharomyces cerevisiae*. *Metab. Eng.* **2014**, *25*, 20–29. [CrossRef] [PubMed]

4. Keren, L.; Zackay, O.; Lotan-Pompan, M.; Barenholz, U.; Dekel, E.; Sasson, V.; Aidelberg, G.; Bren, A.; Zeevi, D.; Weinberger, A.; et al. Promoters maintain their relative activity levels under different growth conditions. *Mol. Syst. Biol.* **2013**, *9*, 701. [CrossRef] [PubMed]

5. Lipson, D.; Raz, T.; Kieu, A.; Jones, D.R.; Giladi, E.; Thayer, E.; Thompson, J.F.; Letovsky, S.; Milos, P.; Causey, M. Quantification of the yeast transcriptome by single-molecule sequencing. *Nat. Biotechnol.* **2009**, *27*, 652–658. [CrossRef] [PubMed]

6. Partow, S.; Siewers, V.; Bjørn, S.; Nielsen, J.; Maury, J. Characterization of different promoters for designing a new expression vector in *Saccharomyces cerevisiae*. *Yeast* **2010**, *27*, 955–964. [CrossRef] [PubMed]

7. Sun, J.; Shao, Z.; Zhao, H.; Nair, N.; Wen, F.; Xu, J.-H.; Zhao, H. Cloning and characterization of a panel of constitutive promoters for applications in pathway engineering in *Saccharomyces cerevisiae*. *Bitechnol. Bioeng.* **2012**, *109*, 2082–2092. [CrossRef]

8. Reider Apel, A.; D'Espaux, L.; Wehrs, M.; Sachs, D.; Li, R.A.; Tong, G.J.; Garber, M.; Nnadi, O.; Zhuang, W.; Hillson, N.J.; et al. A Cas9-based toolkit to program gene expression in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **2016**, *45*, 496–508. [CrossRef]

9. Kim, S.; Lee, K.; Bae, S.-J.; Hahn, J.-S. Promoters inducible by aromatic amino acids and γ-aminobutyrate (GABA) for metabolic engineering applications in *Saccharomyces cerevisiae*. *Appl. Mircrobiol. Bitechnol.* **2015**, *99*, 2705–2714. [CrossRef]

10. Weinhandl, K.; Winkler, M.; Glieder, A.; Camattari, A. Carbon source dependent promoters in yeasts. *Microb. Cell Factories* **2014**, *13*, 5. [CrossRef]

11. Ottoz, D.S.; Rudolf, F. Constitutive and Regulated Promoters in Yeast: How to Design and Make Use of Promoters in *S. cerevisiae*. *Synth. Biol. Parts Devices Appl.* **2018**, *10*, 14.

12. Ro, D.-K.; Paradise, E.M.; Ouellet, M.; Fisher, K.J.; Newman, K.L.; Ndungu, J.M.; Ho, K.A.; Eachus, R.A.; Ham, T.S.; Kirby, J.; et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **2006**, *440*, 940–943. [CrossRef]

13. Luo, X.; Reiter, M.A.; d'Espaux, L.; Wong, J.; Denby, C.M.; Lechner, A.; Zhang, Y.; Grzybowski, A.T.; Harth, S.; Lin, W.; et al. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* **2019**, *567*, 123–126. [CrossRef] [PubMed]

14. Wang, F.; Lv, X.; Xie, W.; Zhou, P.; Zhu, Y.; Yao, Z.; Yang, C.; Yang, X.; Ye, L.; Yu, H. Combining Gal4p-mediated expression enhancement and directed evolution of isoprene synthase to improve isoprene production in *Saccharomyces cerevisiae*. *Metab. Eng.* **2017**, *39*, 257–266. [CrossRef] [PubMed]

15. Alper, H.; Fischer, C.; Nevoigt, E.; Stephanopoulos, G. Tuning genetic control through promoter engineering. *Proc. Natl. Acad.Sci. USA* **2005**, *102*, 12678–12683. [CrossRef] [PubMed]

16. Blazeck, J.; Garg, R.; Reed, B.; Alper, H.S. Controlling promoter strength and regulation in *Saccharomyces cerevisiae* using synthetic hybrid promoters. *Bitechnol. Bioeng.* **2012**, *109*, 2884–2895. [CrossRef]

17. Curran, K.A.; Crook, N.C.; Karim, A.S.; Gupta, A.; Wagman, A.M.; Alper, H.S. Design of synthetic yeast promoters via tuning of nucleosome architecture. *Nat. Commun.* **2014**, *5*, 4002. [CrossRef]

18. Zhu, J.; Zhang, M.Q. SCPD: A promoter database of the yeast *Saccharomyces cerevisiae*. *Bioinformatics* **1999**, *15*, 607–611. [CrossRef]

19. Hertzberg, L.; Zuk, O.; Getz, G.; Domany, E. Finding Motifs in Promoter Regions. *J. Comput. Biol.* **2005**, *12*, 314–330. [CrossRef]

20. Teixeira, M.C.; Monteiro, P.; Jain, P.; Tenreiro, S.; Fernandes, A.R.; Mira, N.P.; Alenquer, M.; Freitas, A.T.; Oliveira, A.L.; Sá-Correia, I. The YEASTRACT database: A tool for the analysis of transcription regulatory associations in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **2006**, *34*, D446–D451. [CrossRef]

21. Teixeira, M.C.; Monteiro, P.T.; Palma, M.; Costa, C.; Godinho, C.P.; Pais, P.; Cavalheiro, M.; Antunes, M.; Lemos, A.; Pedreira, T.; et al. YEASTRACT: An upgraded database for the analysis of transcription regulatory networks in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **2017**, *46*, D348–D353. [CrossRef] [PubMed]

22. Basehoar, A.D.; Zanton, S.J.; Pugh, B.F. Identification and distinct regulation of yeast TATA box-containing genes. *Cell* **2004**, *116*, 699–709. [CrossRef]

23. Pugh, B.F.; Tjian, R. Transcription from a TATA-less promoter requires a multisubunit TFIID complex. *Genes Dev.* **1991**, *5*, 1935–1945. [CrossRef] [PubMed]

24. Mogno, I.; Vallania, F.; Mitra, R.D.; Cohen, B.A. TATA is a modular component of synthetic promoters. *Genome Res.* **2010**, *20*, 1391–1397. [CrossRef]

25. Bae, S.-H.; Han, H.W.; Moon, J. Functional Analysis of the Molecular Interactions of TATA Box-Containing Genes and Essential Genes. *PLoS ONE* **2015**, *10*, e0120848. [CrossRef]

26. Wobbe, C.R.; Struhl, K. Yeast and human TATA-binding proteins have nearly identical DNA sequence requirements for transcription in vitro. *Mol. Cell. Biol.* **1990**, *10*, 3859–3867. [CrossRef]

27. Stewart, J.J.; Stargell, L.A. The stability of the TFIIA-TBP-DNA complex is dependent on the sequence of the TATAAA element. *J. Biol. Chem.* **2001**, *276*, 30078–30084. [CrossRef]

28. Chen, W.; Struhl, K. Saturation mutagenesis of a yeast his3″ TATA element″: genetic evidence for a specific TATA-binding protein. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 2691–2695. [CrossRef]

29. Lubliner, S.; Regev, I.; Lotan-Pompan, M.; Edelheit, S.; Weinberger, A.; Segal, E. Core promoter sequence in yeast is a major determinant of expression level. *Genome Res.* **2015**, *25*, 1008–1017. [CrossRef]

30. Zhang, Z.; Dietrich, F.S. Mapping of transcription start sites in *Saccharomyces cerevisiae* using 5′ SAGE. *Nucleic Acids Res.* **2005**, *33*, 2838–2851. [CrossRef]

31. Buratowski, S.; Hahn, S.; Guarente, L.; Sharp, P.A. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **1989**, *56*, 549–561. [CrossRef]

32. Struhl, K. Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. *Cell* **1987**, *49*, 295–297. [CrossRef]

33. Giardina, C.; Lis, J. DNA melting on yeast RNA polymerase II promoters. *Science* **1993**, *261*, 759–762. [CrossRef]

34. Lubliner, S.; Keren, L.; Segal, E. Sequence features of yeast and human core promoters that are predictive of maximal promoter activity. *Nucleic Acids Res.* **2013**, *11*, 5569–5581. [CrossRef] [PubMed]

35. Redden, H.R. *Developing Synthetic, Minimal Promoters in Saccharomyces Cerevisiae*; The University of Texas at Austin: Austin, TX, USA, 2017.

36. Bitter, G.A.; Chang, K.K.; Egan, K.M. A multi-component upstream activation sequence of the *Saccharomyces cerevisiae* glyceraldehyde-3-phosphate dehydrogenase gene promoter. *Mol. Gen. Genet. MGG* **1991**, *231*, 22–32. [CrossRef]

37. West, R.; Yocum, R.R.; Ptashne, M. *Saccharomyces cerevisiae* GAL1-GAL10 divergent promoter region: Location and function of the upstream activating sequence UASG. *Mol. Cell. Biol.* **1984**, *4*, 2467–2478. [CrossRef]

38. Giniger, E.; Varnum, S.M.; Ptashne, M. Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell* **1985**, *40*, 767–774. [CrossRef]

39. Gancedo, J.M. Yeast carbon catabolite repression. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 334–361. [CrossRef]

40. Bu, Y.; Schmidt, M.C. Identification of cis-acting elements in the SUC2 promoter of *Saccharomyces cerevisiae* required for activation of transcription. *Nucleic Acids Res.* **1998**, *26*, 1002–1009. [CrossRef]

41. Ruiz, A.; Serrano, R.; Ariño, J. Direct regulation of genes involved in glucose utilization by the calcium/calcineurin pathway. *J. Biol. Chem.* **2008**, *283*, 13923–13933. [CrossRef]

42. Nehlin, J.O.; Carlberg, M.; Ronne, H. Control of yeast GAL genes by MIG1 repressor: A transcriptional cascade in the glucose response. *EMBO J.* **1991**, *10*, 3373–3377. [CrossRef] [PubMed]

43. Bojunga, N.; Entian, K.-D. Cat8p, the activator of gluconeogenic genes in *Saccharomyces cerevisiae*, regulates carbon source-dependent expression of NADP-dependent cytosolic isocitrate dehydrogenase (Idp2p) and lactate permease (Jen1p). *Mol. Gen. Gent. MGG* **1999**, *262*, 869–875. [CrossRef] [PubMed]

44. Stewart, A.J.; Hannenhalli, S.; Plotkin, J.B. Why transcription factor binding sites are ten nucleotides long. *Genetics* **2012**, *192*, 973–985. [CrossRef] [PubMed]

45. Rawal, Y.; Chereji, R.V.; Valabhoju, V.; Qiu, H.; Ocampo, J.; Clark, D.J.; Hinnebusch, A.G. Gcn4 Binding in Coding Regions Can Activate Internal and Canonical 5′ Promoters in Yeast. *Mol. Cell* **2018**, *70*, 297–311.e294. [CrossRef]

46. Fernandes, L.; Rodrigues-Pousada, C.; Struhl, K. Yap, a novel family of eight bZIP proteins in *Saccharomyces cerevisiae* with distinct biological functions. *Mol. Cell. Biol.* **1997**, *17*, 6982–6993. [CrossRef]

47. Stathopoulos, A.M.; Cyert, M.S. Calcineurin acts through the CRZ1/TCN1-encoded transcription factor to regulate gene expression in yeast. *Genes Dev.* **1997**, *11*, 3432–3444. [CrossRef]

48. Roque, A.; Petrezsélyová, S.; Serra-Cardona, A.; Ariño, J. Genome-wide recruitment profiling of transcription factor Crz1 in response to high pH stress. *BMC Genom.* **2016**, *17*, 662. [CrossRef]

49. Yamaguchi-Iwai, Y.; Stearman, R.; Dancis, A.; Klausner, R.D. Iron-regulated DNA binding by the AFT1 protein controls the iron regulon in yeast. *EMBO J.* **1996**, *15*, 3377–3384. [CrossRef]

50. Oehlen, L.; Cross, F. The mating factor response pathway regulates transcription of TEC1, a gene involved in pseudohyphal differentiation of *Saccharomyces cerevisiae*. *FEBS Lett.* **1998**, *429*, 83–88. [CrossRef]

51. Dorrity, M.W.; Cuperus, J.T.; Carlisle, J.A.; Fields, S.; Queitsch, C. Preferences in a trait decision determined by transcription factor variants. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E7997–E8006. [CrossRef]

52. Mori, K.; Kawahara, T.; Yoshida, H.; Yanagi, H.; Yura, T. Signalling from endoplasmic reticulum to nucleus: Transcription factor with a basic-leucine zipper motif is required for the unfolded protein-response pathway. *Genes Cells* **1996**, *1*, 803–817. [CrossRef] [PubMed]

53. Eastmond, D.L.; Nelson, H.C.M. Genome-wide analysis reveals new roles for the activation domains of the *Saccharomyces cerevisiae* heat shock transcription factor (Hsf1) during the transient heat shock response. *J. Biol. Chem.* **2006**, *281*, 32909–32921. [CrossRef]

54. Nehlin, J.O.; Ronne, H. Yeast MIG1 repressor is related to the mammalian early growth response and Wilms' tumour finger proteins. *EMBO J.* **1990**, *9*, 2891–2898. [CrossRef] [PubMed]

55. Garcia-Gimeno, M.A.; Struhl, K. Aca1 and Aca2, ATF/CREB activators in *Saccharomyces cerevisiae*, are important for carbon source utilization but not the response to stress. *Mol. Cell. Biol.* **2000**, *20*, 4340–4349. [CrossRef] [PubMed]

56. Luo, Y.; Karpichev, I.V.; Kohanski, R.A.; Small, G.M. Purification, identification, and properties of a *Saccharomyces cerevisiae* oleate-activated upstream activating sequence-binding protein that is involved in the activation of POX1. *J. Biol. Chem.* **1996**, *271*, 12068–12075. [CrossRef]

57. Balzi, E.; Chen, W.; Ulaszewski, S.; Capieaux, E.; Goffeau, A. The multidrug resistance gene PDR1 from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **1987**, *262*, 16871–16879.

58. Subik, J.; Ulaszewski, S.; Goffeau, A. Genetic mapping of nuclear mucidin resistance mutations in *Saccharomyces cerevisiae*. A new pdr locus on chromosome II. *Curr. Genet.* **1986**, *10*, 665–670. [CrossRef]

59. Coorey, N.V.C.; Matthews, J.H.; Bellows, D.S.; Atkinson, P.H. Pleiotropic drug-resistance attenuated genomic library improves elucidation of drug mechanisms. *Mol. Biosyst.* **2015**, *11*, 3129–3136. [CrossRef]

60. Shirozu, R.; Yashiroda, H.; Murata, S. Identification of minimum Rpn4-responsive elements in genes related to proteasome functions. *FEBS Lett.* **2015**, *589*, 933–940. [CrossRef]

61. Guerra-Moreno, A.; Hanna, J. Induction of proteotoxic stress by the mycotoxin patulin. *Toxicol. Lett.* **2017**, *276*, 85–91. [CrossRef]

62. Fingerman, I.M.; Sutphen, K.; Montano, S.P.; Georgiadis, M.M.; Vershon, A.K. Characterization of critical interactions between Ndt80 and MSE DNA defining a novel family of Ig-fold transcription factors. *Nucleic Acids Res.* **2004**, *32*, 2947–2956. [CrossRef] [PubMed]

63. Stebbins, J.L.; Triezenberg, S.J. Identification, mutational analysis, and coactivator requirements of two distinct transcriptional activation domains of the *Saccharomyces cerevisiae* Hap4 protein. *Eukaryot. Cell* **2004**, *3*, 339–347. [CrossRef] [PubMed]

64. Proft, M.; Gibbons, F.D.; Copeland, M.; Roth, F.P.; Struhl, K. Genomewide identification of Sko1 target promoters reveals a regulatory network that operates in response to osmotic stress in *Saccharomyces cerevisiae*. *Eukaryotic. Cell* **2005**, *4*, 1343–1352. [CrossRef] [PubMed]

65. Mai, B.; Breeden, L. Xbp1, a stress-induced transcriptional repressor of the *Saccharomyces cerevisiae* Swi4/Mbp1 family. *Mol. Cell. Biol.* **1997**, *17*, 6491–6501. [CrossRef]

66. Barbaric, S.; Münsterkötter, M.; Goding, C.; Hörz, W. Cooperative Pho2-Pho4 Interactions at thePHO5 Promoter Are Critical for Binding of Pho4 to UASp1 and for Efficient Transactivation by Pho4 at UASp2. *Mol. Cell. Biol.* **1998**, *18*, 2629–2639. [CrossRef]

67. Zhao, H.; Butler, E.; Rodgers, J.; Spizzo, T.; Duesterhoeft, S.; Eide, D. Regulation of zinc homeostasis in yeast by binding of the ZAP1 transcriptional activator to zinc-responsive promoter elements. *J. Biol. Chem.* **1998**, *273*, 28713–28720. [CrossRef] [PubMed]

68. De Nadal, E.; Casadome, L.; Posas, F. Targeting the MEF2-like transcription factor Smp1 by the stress-activated Hog1 mitogen-activated protein kinase. *Mol. Cell. Biol.* **2003**, *23*, 229–237. [CrossRef]

69. Dodou, E.; Treisman, R. The *Saccharomyces cerevisiae* MADS-box transcription factor Rlm1 is a target for the Mpk1 mitogen-activated protein kinase pathway. *Mol. Cell. Biol.* **1997**, *17*, 1848–1859. [CrossRef]

70. Tachibana, C.; Yoo, J.Y.; Tagne, J.-B.; Kacherovsky, N.; Lee, T.I.; Young, E.T. Combined Global Localization Analysis and Transcriptome Data Identify Genes That Are Directly Coregulated by Adr1 and Cat8. *Mol. Cell. Biol.* **2005**, *25*, 2138–2146. [CrossRef] [PubMed]

71. Weider, M.; Machnik, A.; Klebl, F.; Sauer, N. Vhr1p, a new transcription factor from budding yeast, regulates biotin-dependent expression of VHT1 and BIO5. *J. Biol. Chem.* **2006**, *281*, 13513–13524. [CrossRef]

72. Blaiseau, P.L.; Thomas, D. Multiple transcriptional activation complexes tether the yeast activator Met4 to DNA. *EMBO J.* **1998**, *17*, 6327–6336. [CrossRef]

73. Estruch, F.; Carlson, M. Two homologous zinc finger genes identified by multicopy suppression in a SNF1 protein kinase mutant of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **1993**, *13*, 3872–3881. [CrossRef]

74. Vincent, O.; Carlson, M. Sip4, a Snf1 kinase-dependent transcriptional activator, binds to the carbon source-responsive element of gluconeogenic genes. *EMBO J.* **1998**, *17*, 7002–7008. [CrossRef] [PubMed]

75. Ozcan, S.; Leong, T.; Johnston, M. Rgt1p of *Saccharomyces cerevisiae*, a key regulator of glucose-induced genes, is both an activator and a repressor of transcription. *Mol. Cell. Biol.* **1996**, *16*, 6419–6426. [CrossRef] [PubMed]

76. Kim, J.-H.; Polish, J.; Johnston, M. Specificity and regulation of DNA binding by the yeast glucose transporter gene repressor Rgt1. *Mol. Cell. Biol.* **2003**, *23*, 5208–5216. [CrossRef] [PubMed]

77. Tornow, J.; Zeng, X.; Gao, W.; Santangelo, G.M. GCR1, a transcriptional activator in *Saccharomyces cerevisiae*, complexes with RAP1 and can function without its DNA binding domain. *EMBO J.* **1993**, *12*, 2431–2437. [CrossRef] [PubMed]

78. Lowry, C.V.; Cerdán, M.E.; Zitomer, R.S. A hypoxic consensus operator and a constitutive activation region regulate the ANB1 gene of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **1990**, *10*, 5921–5926. [CrossRef]

79. Fields, D.S.; He, Y.Y.; Al-Uzri, A.Y.; Stormo, G.D. Quantitative specificity of the Mnt repressor. *J. Mol. Biol.* **1997**, *271*, 178–194. [CrossRef]

80. Vashee, S.; Xu, H.; Johnston, S.; Kodadek, T. How do "Zn2 cys6" proteins distinguish between similar upstream activation sites? Comparison of the DNA-binding specificity of the GAL4 protein in vitro and in vivo. *J. Biol. Chem.* **1993**, *268*, 24699–24706.

81. Skowron, P.M.; Harasimowicz, R.; Rutkowska, S.M. GCN4 eukaryotic transcription factor/FokI endonuclease-mediated 'Achilles' heel cleavage': Quantitative study of protein-DNA interaction. *Gene* **1996**, *170*, 1–8. [CrossRef]

82. Hill, D.E.; Hope, I.A.; Macke, J.P.; Struhl, K. Saturation mutagenesis of the yeast his3 regulatory site: Requirements for transcriptional induction and for binding by GCN4 activator protein. *Science* **1986**, *234*, 451–457. [CrossRef] [PubMed]

83. Sharon, E.; Kalma, Y.; Sharp, A.; Raveh-Sadka, T.; Levo, M.; Zeevi, D.; Keren, L.; Yakhini, Z.; Weinberger, A.; Segal, E. Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat. Bitechnol.* **2012**, *30*, 521. [CrossRef] [PubMed]

84. Giniger, E.; Ptashne, M. Cooperative DNA binding of the yeast transcriptional activator GAL4. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 382–386. [CrossRef] [PubMed]

85. Hahn, S.; Young, E.T. Transcriptional regulation in *Saccharomyces cerevisiae*: Transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. *Genetics* **2011**, *189*, 705–736. [CrossRef] [PubMed]

86. Dobi, K.C.; Winston, F. Analysis of transcriptional activation at a distance in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **2007**, *27*, 5575–5586. [CrossRef] [PubMed]

87. Escalante-Chong, R.; Savir, Y.; Carroll, S.M.; Ingraham, J.B.; Wang, J.; Marx, C.J.; Springer, M. Galactose metabolic genes in yeast respond to a ratio of galactose and glucose. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 1636–1641. [CrossRef]

88. Mustonen, V.; Kinney, J.; Callan, C.G.; Lässig, M. Energy-dependent fitness: A quantitative model for the evolution of yeast transcription factor binding sites. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 12376–12381. [CrossRef]

89. Moses, A.M.; Chiang, D.Y.; Kellis, M.; Lander, E.S.; Eisen, M.B. Position specific variation in the rate of evolution in transcription factor binding sites. *BMC Evol. Biol.* **2003**, *3*, 19. [CrossRef]

90. Nagalakshmi, U.; Wang, Z.; Waern, K.; Shou, C.; Raha, D.; Gerstein, M.; Snyder, M. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **2008**, *320*, 1344–1349. [CrossRef]

91. Hu, M.; Yu, J.; Taylor, J.M.; Chinnaiyan, A.M.; Qin, Z.S. On the detection and refinement of transcription factor binding sites using ChIP-Seq data. *Nucleic Acids Res.* **2010**, *38*, 2154–2167. [CrossRef]

92. Shultzaberger, R.K.; Maerkl, S.J.; Kirsch, J.F.; Eisen, M.B. Probing the Informational and regulatory plasticity of a transcription factor DNA–binding domain. *PLoS Gent.* **2012**, *8*, e1002614. [CrossRef]

93. Lee, W.; Tillo, D.; Bray, N.; Morse, R.H.; Davis, R.W.; Hughes, T.R.; Nislow, C. A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Gent.* **2007**, *39*, 1235–1244. [CrossRef]

94. Juan, L.-J.; Walter, P.; Taylor, I.; Kingston, R.; Workman, J. Nucleosome cores and histone H1 in the binding of GAL4 derivatives and the reactivation of transcription from nucleosome templates in vitro. In *Cold Spring Harbor Symposia on Quantitative Biology*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 1993; Volume 58, pp. 213–223.

95. Ercan, S.; Simpson, R.T. Global chromatin structure of 45,000 base pairs of chromosome III in a-and α-cell yeast and during mating-type switching. *Mol. Cell. Biol.* **2004**, *24*, 10026–10035. [CrossRef]

96. Ercan, S.; Carrozza, M.J.; Workman, J.L. Global nucleosome distribution and the regulation of transcription in yeast. *Genome Biol.* **2004**, *5*, 243. [CrossRef] [PubMed]

97. Bernstein, B.E.; Liu, C.L.; Humphrey, E.L.; Perlstein, E.O.; Schreiber, S.L. Global nucleosome occupancy in yeast. *Genome Biol.* **2004**, *5*, R62. [CrossRef] [PubMed]

98. Segal, E.; Fondufe-Mittendorf, Y.; Chen, L.; Thåström, A.; Field, Y.; Moore, I.K.; Wang, J.-P.Z.; Widom, J. A genomic code for nucleosome positioning. *Nature* **2006**, *442*, 772–778. [CrossRef] [PubMed]

99. Li, B.; Carey, M.; Workman, J.L. The role of chromatin during transcription. *Cell* **2007**, *128*, 707–719. [CrossRef]

100. Liang, S.D.; Marmorstein, R.; Harrison, S.C.; Ptashne, M. DNA sequence preferences of GAL4 and PPR1: How a subset of Zn2 Cys6 binuclear cluster proteins recognizes DNA. *Mol. Cell. Biol.* **1996**, *16*, 3773–3780. [CrossRef]

101. Cote, J.; Quinn, J.; Workman, J.L.; Peterson, C.L. Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science* **1994**, *265*, 53–60. [CrossRef]

102. Yu, L.; Morse, R.H. Chromatin Opening and Transactivator Potentiation by RAP1 in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **1999**, *19*, 5279–5288. [CrossRef]

103. Wang, X.; Bryant, G.O.; Floer, M.; Spagna, D.; Ptashne, M. An effect of DNA sequence on nucleosome occupancy and removal. *Nat. Struct. Mol. Biol.* **2011**, *18*, 507. [CrossRef] [PubMed]

104. Workman, J.L. Nucleosome displacement in transcription. *Genes Dev.* **2006**, *20*, 2009–2017. [CrossRef] [PubMed]

105. Svaren, J.; Schmitz, J.; Hörz, W. The transactivation domain of Pho4 is required for nucleosome disruption at the PHO5 promoter. *EMBO J.* **1994**, *13*, 4856–4862. [CrossRef] [PubMed]

106. Venter, U.; Svaren, J.; Schmitz, J.; Schmid, A.; Hörz, W. A nucleosome precludes binding of the transcription factor Pho4 in vivo to a critical target site in the PHO5 promoter. *EMBO J.* **1994**, *13*, 4848–4855. [CrossRef] [PubMed]

107. Segal, E.; Widom, J. Poly (dA: dT) tracts: Major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.* **2009**, *19*, 65–71. [CrossRef]

108. Verdone, L.; Camilloni, G.; Di Mauro, E.; Caserta, M. Chromatin remodeling during *Saccharomyces cerevisiae* ADH2 gene activation. *Mol. Cell. Biol.* **1996**, *16*, 1978–1988. [CrossRef]

109. Filetici, P.; Aranda, C.; Gonzalez, A.; Ballario, P. GCN5, a Yeast Transcriptional Coactivator, Induces Chromatin Reconfiguration ofHIS3Promoter in Vivo. *Biochem. Biophys. Res. Commun.* **1998**, *242*, 84–87. [CrossRef]

110. Lascaris, R.F.; Groot, E.d.; Hoen, P.-B.t.; Mager, W.H.; Planta, R.J. Different roles for abf1p and a T-rich promoter element in nucleosome organization of the yeast RPS28A gene. *Nucleic Acids Res.* **2000**, *28*, 1390–1396. [CrossRef]

111. Tanaka, S.; Livingstone-Zatchej, M.; Thoma, F. Chromatin structure of the yeast URA3 gene at high resolution provides insight into structure and positioning of nucleosomes in the chromosomal context. *J. Mol. Biol.* **1996**, *257*, 919–934. [CrossRef]

112. Iyer, V.; Struhl, K. Poly (dA: dT), a ubiquitous promoter element that stimulates transcription via its intrinsic DNA structure. *EMBO J.* **1995**, *14*, 2570–2579. [CrossRef]

113. Raveh-Sadka, T.; Levo, M.; Shabi, U.; Shany, B.; Keren, L.; Lotan-Pompan, M.; Zeevi, D.; Sharon, E.; Weinberger, A.; Segal, E. Manipulating nucleosome disfavoring sequences allows fine-tune regulation of gene expression in yeast. *Nat. Gent.* **2012**, *44*, 743. [CrossRef] [PubMed]

114. Jin, L.-Q.; Jin, W.-R.; Ma, Z.-C.; Shen, Q.; Cai, X.; Liu, Z.-Q.; Zheng, Y.-G. Promoter engineering strategies for the overproduction of valuable metabolites in microbes. *Appl. Mircrobiol. Bitechnol.* **2019**, *103*, 8725–8736. [CrossRef]

115. Xu, N.; Wei, L.; Liu, J. Recent advances in the applications of promoter engineering for the optimization of metabolite biosynthesis. *World J. Mircrobiol. Bitechnol.* **2019**, *35*, 33. [CrossRef]

116. Blazeck, J.; Alper, H.S. Promoter engineering: Recent advances in controlling transcription at the most fundamental level. *Bitechnol. J.* **2013**, *8*, 46–58. [CrossRef] [PubMed]

117. Bhardwaj, N.; Langlois, R.E.; Zhao, G.; Lu, H. Kernel-based machine learning protocol for predicting DNA-binding proteins. *Nucleic Acids Res.* **2005**, *33*, 6486–6493. [CrossRef] [PubMed]

118. Abeel, T.; Saeys, Y.; Bonnet, E.; Rouzé, P.; Van de Peer, Y. Generic eukaryotic core promoter prediction using structural features of DNA. *Genome Res.* **2008**, *18*, 310–323. [CrossRef]

119. Zakov, S.; Goldberg, Y.; Elhadad, M.; Ziv-Ukelson, M. Rich parameterization improves RNA structure prediction. *J. Comput. Biol.* **2011**, *18*, 1525–1542. [CrossRef]

120. Cuperus, J.T.; Groves, B.; Kuchina, A.; Rosenberg, A.B.; Jojic, N.; Fields, S.; Seelig, G. Deep learning of the regulatory grammar of yeast 5′ untranslated regions from 500,000 random sequences. *Genome Res.* **2017**, *27*, 2015–2024. [CrossRef]

121. De Mey, M.; Maertens, J.; Lequeux, G.J.; Soetaert, W.K.; Vandamme, E.J. Construction and model-based analysis of a promoter library for *E. coli*: An indispensable tool for metabolic engineering. *BMC Bitechnol.* **2007**, *7*, 34. [CrossRef]

122. Meng, H.; Wang, J.; Xiong, Z.; Xu, F.; Zhao, G.; Wang, Y. Quantitative design of regulatory elements based on high-precision strength prediction using artificial neural network. *PLoS ONE* **2013**, *8*, e60288. [CrossRef]

123. Wang, Y.; Wang, H.; Wei, L.; Li, S.; Liu, L.; Wang, X. Synthetic promoter design in Escherichia coli based on a deep generative network. *Nucleic Acids Res* **2020**, *48*, 6403–6412. [CrossRef] [PubMed]

124. Kotopka, B.J.; Smolke, C.D. Model-driven generation of artificial yeast promoters. *Nat. Commun.* **2020**, *11*, 2113. [CrossRef]

125. Chen, X.; Gao, C.; Guo, L.; Hu, G.; Luo, Q.; Liu, J.; Nielsen, J.; Chen, J.; Liu, L. DCEO Biotechnology: Tools To Design, Construct, Evaluate, and Optimize the Metabolic Pathway for Biosynthesis of Chemicals. *Chem. Rev.* **2018**, *118*, 4–72. [CrossRef] [PubMed]

126. Zhou, Y.; Li, G.; Dong, J.; Xing, X.-h.; Dai, J.; Zhang, C. MiYA, an efficient machine-learning workflow in conjunction with the YeastFab assembly strategy for combinatorial optimization of heterologous metabolic pathways in *Saccharomyces cerevisiae*. *Metab. Eng.* **2018**, *47*, 294–302. [CrossRef] [PubMed]

127. Nevoigt, E.; Kohnke, J.; Fischer, C.R.; Alper, H.; Stahl, U.; Stephanopoulos, G. Engineering of promoter replacement cassettes for fine-tuning of gene expression in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **2006**, *72*, 5266–5273. [CrossRef] [PubMed]

128. Yuan, Y.; Zhao, H. Directed evolution of a highly efficient cellobiose utilizing pathway in an industrial *Saccharomyces cerevisiae* strain. *Bitechnol. Bioeng.* **2013**, *110*, 2874–2881. [CrossRef] [PubMed]

129. Wang, J.; Zhai, H.; Rexida, R.; Shen, Y.; Hou, J.; Bao, X. Developing synthetic hybrid promoters to increase constitutive or diauxic shift-induced expression in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **2018**, *18*, foy098. [CrossRef]

130. Peng, B.; Wood, R.J.; Nielsen, L.K.; Vickers, C.E. An Expanded Heterologous GAL Promoter Collection for Diauxie-Inducible Expression in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **2018**, *7*, 748–751. [CrossRef]

131. Shabbir Hussain, M.; Gambill, L.; Smith, S.; Blenner, M.A. Engineering Promoter Architecture in Oleaginous Yeast Yarrowia lipolytica. *ACS Synth. Biol.* **2016**, *5*, 213–223. [CrossRef]

132. Decoene, T.; De Maeseneire, S.L.; De Mey, M. Modulating transcription through development of semi-synthetic yeast core promoters. *PLoS ONE* **2019**, *14*, e0224476. [CrossRef]

133. Liu, R.; Liu, L.; Li, X.; Liu, D.; Yuan, Y. Engineering yeast artificial core promoter with designated base motifs. *Microb. Cell Factories* **2020**, *19*, 38. [CrossRef] [PubMed]

134. Jin, L.; Nawab, S.; Xia, M.; Ma, X.; Huo, Y.-X. Context-dependency of synthetic minimal promoters in driving gene expression: A case study. *Microb. Bitechnol.* **2019**, *12*, 1476–1486. [CrossRef]

135. Redden, H.; Alper, H.S. The development and characterization of synthetic minimal yeast promoters. *Nat. Commun.* **2015**, *6*, 7810. [CrossRef] [PubMed]

136. Katzke, N.; Knapp, A.; Loeschcke, A.; Drepper, T.; Jaeger, K.-E. Novel Tools for the Functional Expression of Metagenomic DNA. In *Metagenomics: Methods and Protocols, Streit, W.R., Daniel, R., Eds.*; Springer New York: Midtown Manhattan, NY, USA, 2017; pp. 159–196. [CrossRef]

137. Yang, S.; Liu, Q.; Zhang, Y.; Du, G.; Chen, J.; Kang, Z. Construction and Characterization of Broad-Spectrum Promoters for Synthetic Biology. *ACS Synth. Biol.* **2018**, *7*, 287–291. [CrossRef] [PubMed]

138. Portela, R.M.C.; Vogl, T.; Kniely, C.; Fischer, J.E.; Oliveira, R.; Glieder, A. Synthetic Core Promoters as Universal Parts for Fine-Tuning Expression in Different Yeast Species. *ACS Synth. Biol.* **2017**, *6*, 471–484. [CrossRef] [PubMed]

139. Zhang, J.; Sonnenschein, N.; Pihl, T.P.B.; Pedersen, K.R.; Jensen, M.K.; Keasling, J.D. Engineering an NADPH/NADP+ Redox Biosensor in Yeast. *ACS Synth. Biol.* **2016**, *5*, 1546–1556. [CrossRef] [PubMed]

140. Rajkumar, A.S.; Liu, G.; Bergenholm, D.; Arsovska, D.; Kristensen, M.; Nielsen, J.; Jensen, M.K.; Keasling, J.D. Engineering of synthetic, stress-responsive yeast promoters. *Nucleic Acids Res.* **2016**, *44*, e136. [CrossRef]

141. Gopinarayanan, V.E.; Nair, N.U. A semi-synthetic regulon enables rapid growth of yeast on xylose. *Nat. Commun.* **2018**, *9*, 1233. [CrossRef]

142. Garí, E.; Piedrafita, L.; Aldea, M.; Herrero, E. A Set of Vectors with a Tetracycline-Regulatable Promoter System for Modulated Gene Expression in *Saccharomyces cerevisiae*. *Yeast* **1997**, *13*, 837–848. [CrossRef]

143. Murphy, K.F.; Balázsi, G.; Collins, J.J. Combinatorial promoter design for engineering noisy gene expression. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 12726–12731. [CrossRef]

144. Ikushima, S.; Zhao, Y.; Boeke, J.D. Development of a tightly controlled off switch for *Saccharomyces cerevisiae* regulated by camphor, a low-cost natural product. *G3 Genes Genomes Gent.* **2015**, *5*, 1983–1990. [CrossRef] [PubMed]

145. Teo, W.S.; Hee, K.S.; Chang, M.W. Bacterial F ad R and synthetic promoters function as modular fatty acid sensor-regulators in S accharomyces cerevisiae. *Eng. Life Sci.* **2013**, *13*, 456–463. [CrossRef]

146. Li, S.; Si, T.; Wang, M.; Zhao, H. Development of a synthetic malonyl-CoA sensor in *Saccharomyces cerevisiae* for intracellular metabolite monitoring and genetic screening. *ACS Synth. Biol.* **2015**, *4*, 1308–1315. [CrossRef] [PubMed]

147. Mazumder, M.; McMillen, D.R. Design and characterization of a dual-mode promoter with activation and repression capability for tuning gene expression in yeast. *Nucleic Acids Res.* **2014**, *42*, 9514–9522. [CrossRef] [PubMed]

148. Williams, T.C.; Nielsen, L.K.; Vickers, C.E. Engineered Quorum Sensing Using Pheromone-Mediated Cell-to-Cell Communication in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **2013**, *2*, 136–149. [CrossRef] [PubMed]