



# THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Towards combinatorial transcriptional engineering

**Citation for published version:**

Mehrotra, R, Renganaath, K, Kanodia, H, Loake, G & Mehrotra, S 2017, 'Towards combinatorial transcriptional engineering' *Biotechnology Advances*, vol 35, no. 3, pp. 390-405. DOI: 10.1016/j.biotechadv.2017.03.006

**Digital Object Identifier (DOI):**

[10.1016/j.biotechadv.2017.03.006](https://doi.org/10.1016/j.biotechadv.2017.03.006)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

*Biotechnology Advances*

**Publisher Rights Statement:**

© 2017 Elsevier Inc. All rights reserved

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Accepted Manuscript

Towards combinatorial transcriptional engineering

Rajesh Mehrotra, Kaushik Renganaath, Harsh Kanodia, Gary J Loake, Sandhya Mehrotra



PII: S0734-9750(17)30024-1  
DOI: doi: [10.1016/j.biotechadv.2017.03.006](https://doi.org/10.1016/j.biotechadv.2017.03.006)  
Reference: JBA 7110  
To appear in: *Biotechnology Advances*  
Received date: 2 November 2016  
Revised date: 22 February 2017  
Accepted date: 9 March 2017

Please cite this article as: Rajesh Mehrotra, Kaushik Renganaath, Harsh Kanodia, Gary J Loake, Sandhya Mehrotra , Towards combinatorial transcriptional engineering. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Jba(2017), doi: [10.1016/j.biotechadv.2017.03.006](https://doi.org/10.1016/j.biotechadv.2017.03.006)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Towards Combinatorial Transcriptional Engineering

Rajesh Mehrotra <sup>a¶\*</sup>, Kaushik Renganaath <sup>b¶§</sup>, Harsh Kanodia <sup>a c ¶ §</sup>, Gary J Loake <sup>d</sup>,  
Sandhya Mehrotra <sup>a</sup>

<sup>a</sup> Department of Biological Sciences, Birla Institute of Technology and Science (BITS) Pilani, Pilani - 333031, Rajasthan (India)

<sup>b</sup> Email: kaushik.renganaath@gmail.com

<sup>c</sup> Email: kanodia.harsh.hk@gmail.com

<sup>d</sup> Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, The King's Buildings, Max Born Crescent, Edinburgh EH9 3BF.

\*Corresponding author: **Tel: +91-9983734134, Email: rmehrotra@pilani.bits-pilani.ac.in, rajmeh25@hotmail.com**

¶ - Co-First Author

§ - Equal Contribution

.....  
.....

**Abstract:** The modular nature of the transcriptional unit makes it possible to design robust modules with predictable input-output characteristics using a ‘parts- off a shelf’ approach. Customized regulatory circuits composed of multiple such transcriptional units have immense scope for application in diverse fields of basic and applied research. Synthetic transcriptional engineering seeks to construct such genetic cascades. Here, we discuss the three principle strands of transcriptional engineering: promoter and transcriptional factor engineering, and programming inducibility into synthetic modules. In this context, we review the scope and limitations of some recent technologies that seek to achieve these ends. Our discussion emphasizes a requirement for rational combinatorial engineering principles and the promise this approach holds for the future development of this field.

**Key Words:** Promoters, Transcription Factors, Gene Expression, Synthetic Biology, Transcriptional engineering, TALE, CRISPR.

.....  
.....

## Abbreviations

1. ARE - androgen responsive element; 2. aTF- Allosteric Transcription Factor; 3. BSR – base specifying residue; 4. CMV- Cytomegalovirus; 5. CR - chromatin regulator; 6. DBD - DNA binding domain; 7. dCas9 - endonuclease deficient Cas9; 8. G6PDH - Glucose-6 phosphate dehydrogenase; 9. gRNA - guide RNA; 10. HBGS - homology based gene silencing; 11. HPL - hybrid promoter libraries; 12. IPTG - Isopropyl  $\beta$ -D-1-thiogalactopyranoside; 13. KRAB- Krüppel-associated box; 14. OPEN - oligomerized pool engineering; 15. PAM – protospacer adjacent motif; 16. RE - response elements; 17. RNAP – RNA Polymerase; 18. RVD - repeat variable di-residues; 19. SPL - synthetic promoter library; 20. sTF - synthetic TF; 21. TALE - Transcription activator-like effector; 22. TALOR - TALE orthogonal repressor; 23. TEF - translation elongation factor; 24. TF – Transcription Factor; 25. TFBS - TF binding site; 26. TSS - transcription start site; 27. UAS - upstream activating sequences; 28. ZNF – Zinc Fingers.

.....  
 .....

## 1. Introduction

In recent years, transcription in eukaryotes has been examined in great detail as a component of protein synthesis. This has uncovered a highly complex machinery that regulates transcription of individual genes. *Cis*-acting promoter elements including suppressors, silencers and enhancers provide sites for the binding of *trans*-acting activators and repressors (Bhattacharjee et al., 2013). These elements form discrete transcriptional modules, which may be combined to form a complete transcriptional regulatory unit. Synthetic biologists now seek to exploit the modular feature of transcriptional units to design customized regulatory circuits.

Eukaryotic transcriptional programs are highly integrated networks that wire together multiple promoter elements to specific cellular pathways. These transcription modules are implicated in the tuning of molecular noise, recruitment of transcription factor (TF) complexes and in controlling nucleosomal remodeling, among others (Hahn and Young, 2011; Pedraza and van Oudenaarden, 2005; Rosenfeld et al., 2005). The details of the regulatory pathways are yet to be understood clearly, however, the design of synthetic circuits using a ‘parts-off a shelf’ approach will expedite our understanding of these basic frameworks *in-vivo* (Guido et al., 2006; Vilar, 2006). Additionally, custom transcriptional networks may be employed to advance metabolic engineering and optimization, which are useful in industry, therapeutics and crop improvement (Alper and Fischer, 2005; Jensen and Hammer, 1998a; Le Bec and Douar, 2006; Nandagopal and Elowitz, 2011; Smolke and Silver, 2011).

The future of transcriptional network engineering demands building a repository of *cis*-regulatory modules that can be knit into synthetic promoters. Although a large library of native promoters is available, this resource does not encompass a wide range of promoter strengths over a continuous range (Mehrotra et al., 2011). Sometimes, this basic library may itself be limited as in new or obscure model organisms (Blazek et al., 2012; Siegl et al., 2013). The modular

promoter elements required for construction of novel expression cassettes are sufficiently large in *Escherichia coli* (Andrianantoandro and Basu, 2006; Nandagopal and Elowitz, 2011). These standard biological parts have also been opened up for public access (Bio FAB Group et al., 2006). However, promoter engineering efforts have only been partially successful in eukaryotes, despite the availability of strong native promoters in organisms like *Saccharomyces cerevisiae* (Da Silva and Srikrishnan, 2012).

In some eukaryotic systems like the budding yeast (*S. cerevisiae*), the repeated use of the same promoter modules is problematic due to homologous recombination events that decrease the stability of the artificially introduced expression cassettes (Gibson et al., 2008). In plant systems, the introduction of multiple promoters or transgenes bearing homology to host *cis*-regulatory modules leads to homology based gene silencing (HBGS) at the transcriptional or post-transcriptional level (Meyer and Saedler, 1996). Native promoter modules also use the natural TF population in the cell, which increases the probability of inducing off-target genes, whose promoters contain the binding sites for the same TF (Siegl et al., 2013). Such transcriptional noise overburdens the cell and adversely affects its survival. Therefore, this necessitates the development of modules with novel TF binding sites (TFBS) and of synthetic TFs (sTF) with low off-target binding. In synthetic biological terms, this concept is called orthogonality (Rao, 2012).

In complex multicellular systems like humans or plants, systemic effects are undesirable, particularly in the fields of therapeutics and transgenic technology. Thus, synthetic transcriptional networks with intended application in these areas will need to have high spatiotemporal resolution for targeted action in specific tissues. Prolonged expression of target genes from potent activator systems with supra-optimal regulatory activity may burden the cell and reduce its viability. In other cases, it may be desirable to express a particular gene for a limited duration only. Therefore, there is need for high spatial and temporal control of expression from synthetic transcriptional modules. This can be achieved by placing synthetic modules under control of external inducers whose concentration in the extracellular environment may be tightly regulated by the user.

Efforts towards developing new transcription modules have focused on addressing these requirements by one or more of the following methods: (a) creating promoter libraries with reduced homology to native *cis*-regulatory elements; (b) design of novel expression cassettes either stronger (for promoting high output) or weaker than wild-type promoters (for reducing the expression of gene products that could be toxic); (c) achieving orthogonality by designing custom-made TFs that are specific to the target promoters; and, (d) making synthetic promoters or TFs responsive to physical and chemical inducers.

Ultimately, the aim is to design complete transcriptional networks that are easily tunable, dynamic, robust, orthogonal and simple to handle with predictable input-output characteristics. This requires rational engineering that combines designer promoters, transcription factors and the ability to regulate them externally. In this review, we address these three strands of synthetic transcriptional engineering. We initially present some recent tools for creating synthetic promoter libraries followed by a discussion of three major technologies that have revolutionized

TF engineering: Zinc Fingers (ZNFs), Transcription Activator-like Effectors (TALEs) and CRISPR/Cas9. In each of these sections we separately discuss the property of inducibility to underscore the differences in the required principles. In our discussion, we make the case for the utility of rational combinatorial engineering. The latter is important for overcoming the limitations of individual technologies and for designing transcriptional networks with improved future performance.

## 2. Promoter Engineering

The basic eukaryotic promoter is the seat for assembly of TFs. It consists of two regions – the core promoter and the upstream promoter elements. The core promoter element lies about 40 bp upstream of the transcription start site (TSS) and contains -10 TATA box (Molina et al., 2005), which is bound by the basal TFs that recruit RNA Polymerase (RNAP) II (Lee and Young, 2000). Promoters lacking a TATA-box are also known; these promoters use upstream or downstream activating sequences to assemble TFs a few base pairs (bps) upstream of the TSS. The basal transcription rate from the core promoter is minimal and varies depending upon the nature of core promoter motifs (Blazek et al., 2013). However, transcription rate may be enhanced or suppressed through the presence of additional regulatory elements, like enhancers or repressors. The core TATA box motif itself promotes gene transcription independent of the upstream regulator regions (Mogno et al., 2010), and without altering the gene specificity conferred by them. An experimenter may thus combine these motifs with other activator or repressor modules to produce a chimeric promoter that caters to his/her given requirements.

Proximal or distal *cis*-regulatory elements may be used to develop either constitutive or inducible promoters. Constitutive promoters do not require additional factors such as inducers, enhancers, repressors etc. Thus, they can offer relatively stable transcriptional outputs in diverse conditions. Among eukaryotes, many constitutive promoters have been characterized in yeast; most of them are associated with glycolytic pathway genes. Examples include the promoters of phosphoglycerate kinase 1  $P_{PGK1}$  (Holland and Holland, 1978; Ogden et al., 1986), pyruvate decarboxylase  $PPD_{C1}$  (Kellermann et al., 1986), triose phosphate isomerase  $PTPI1$  (Alber and Kawasaki, 1982), alcohol dehydrogenase I  $P_{ADH1}$  (Denis et al., 1983; Hitzeman et al., 1981), glyceraldehyde-3-phosphate dehydrogenase  $P_{TDH3(GAP491)}$  or  $P_{GPD}$  (Bitter and Egan, 1984; McAlister and Holland, 1985) and pyruvate kinase  $P_{PYK1}$  (Nishizawa et al., 1989). The promoter encoding translation elongation factor 1 (TEF1) is also commonly used for the purpose of promoter engineering and is well characterized (Gatignol et al., 1990). The minimal Cauliflower Mosaic Virus (CMV) 35S core promoter is frequently used for promoter engineering efforts in plants. It provides strong transcription initiation while displaying minimum basal transcription in monocots and dicots (Liu and Stewart, 2016). The 126 bp ZmUbi1 promoter provides low basal transcription and has been recently characterized for specific use in monocots (Kumar et al., 2015).

Expression from inducible promoters is driven by specific environmental or developmental signals. For example, the two yeast promoters for *GAL1*, *GAL7* and *GAL10* genes (Bassel and Mortimer, 1971; DOUGLAS and HAWTHORNE, 1964) are selectively induced by addition of galactose in the growth medium and are repressed in the presence of glucose. Inducible

promoters may also be used to program tissue or cell-type specific expression, which has wide implications in therapeutics. For example, the ApoB and ornithine transcarbamoylase gene promoter elements have been engineered to express only in liver cells (Lemken et al., 2005). Cell-type specific expression in myeloid cells has also been achieved by randomly inserting myeloid-specific elements for PU.1, C/EBPalpha, AML-1 and by inserting myeloid-associated elements for Sp1 and AP-1, upstream of the p47-phox minimal promoter (He et al., 2006).

Characterised native promoters to date do not cover a wide range of promoter strengths in a continuously graded fashion. They are thus unsuitable for applications in metabolic optimization and engineering. The latter requires fine-tuning of gene expression. Constitutive promoters are usually preferred for promoter engineering since they are easy to manipulate. Inducible promoters may be useful for continuous control of gene expression at the macroscopic level. But practically, they are limited by factors like prohibitive inducer costs, hypersensitivity to inducer concentrations and transcriptional heterogeneity at the single cell level (Alper and Fischer, 2005; Mnaimneh et al., 2004; Siegele and Hu, 1997). More recently, hybrid promoters have been designed to overcome this limitation. Upstream activating sequences (UAS) have also been used to modulate inducibility (Blazek et al., 2012).

Some novel promoter engineering strategies have recently been developed, these have been broadly grouped into three classes: (a) random mutagenesis; (b) generation of hybrid promoters; (c) de-novo designing of promoters. Computational modeling and massive high-throughput DNA sequencing/synthesis have simplified the design, optimization, construction and screening of novel and robust synthetic promoters. An interdisciplinary approach has made *cis*-engineering more efficient in the present and it promises to advance the frontiers of this field in the future (Mehrotra et al., 2011).

### 2.1. Random Mutagenesis

It has been shown that the intervening spacer between the -10 and -35 consensus sequences in *E. coli* contributes to the strength of bacterial promoters (Jensen and Hammer, 1998b). Keeping this in view, the intervening spacer sequences were randomly altered creating synthetic promoter libraries (SPLs) containing promoters of variable strength. The random alteration of spacers was achieved mainly by one or both of the following two methods: (a) use of randomized oligonucleotide primers in a single PCR reaction, preceding a region of homology to target gene (Figure 1a) (Solem and Jensen, 2002); or, (b) mutagenic PCR of the promoter region of bacteriophage  $\lambda$  (Figure 1b) (Alper and Fischer, 2005). While the former method made it simpler at the stage of screening the mutant colonies, the screening of the promoter libraries created by error prone PCR was more laborious. For instance, in one case, the mutagenic PCR strategy generated 30,000 transformed *E. coli* colonies, out of which 27 (possessing 22 distinct promoter sequences) gave uniform altered, promoter activities at single cell level (Alper and Fischer, 2005). The strength of these promoters could be as high as 196-fold relative to the native promoter, as observed using single cell GFP fluorescence measurements, RT-PCR as well as the measurement of regulation of a second *cat* reporter gene. The promoter variants showed consistent results in a wide range of conditions and in a variety of host strains, underscoring the constitutive and universal nature of the SPL. In contrast, the earlier methods employing the use

of degenerate primers gave a variable promoter strength in 50-90% of the cases, thereby decreasing the load during screening of transformants (Hammer et al., 2006; Solem and Jensen, 2002).

Techniques similar to the above have been employed successfully in eukaryotes including budding yeast (Alper and Fischer, 2005). In this study, a TEF1 promoter template from yeast was used to generate a similar SPL with a wide range of gene regulatory capacities. In another study, 11 yeast clones (range 8%-120% of native promoter strength) that were generated from the SPL were validated through GFP fluorescence and RT-PCR measurements (Alper and Fischer, 2005). Random mutagenesis of the oxygen responsive DAN1 gene promoter in *S. cerevisiae* generated two mutants with decreased anaerobic inducibility. These promoters could be used to regulate gene expression during cell growth and proliferation as transgenes could be expressed upon growth induced oxygen depletion (Nevoigt et al., 2007).

Degenerate primer based random mutagenesis was also used to successfully generate similar SPLs in other studies. For instance, in budding yeast, Jeppsson et al., 2003, used this technique to achieve variable Glucose-6 phosphate dehydrogenase (G6PDH) activities, which ranged from 0% to 179% of the wild type. The degenerate primer method for generating SPLs has also been employed in relatively less common prokaryotic model systems like the actinomycetes (Siegl et al., 2013). The strength of the activity of SPL generated promoters ranged from 2% and 319%, relative to native promoter.

Recent high throughput synthesis approaches have also been used to create more exhaustive promoter libraries. For instance, in one recent study, mutant promoters were synthesized in the form of DNA oligonucleotides on a programmable microarray (Patwardhan et al., 2009). The library consisting of these mutants consisted of promoters involving variants at each of the 35 nucleotide positions (23 bp upstream and 12 bp downstream of the TSS). This library was then used for *in-vitro* transcription to evaluate the promoter strengths, achieved by DNA sequencing the tagged downstream region containing multiple 20 nucleotide barcoded sequences. The study was successful in assessment of mutations at each position for each of the 3 bacteriophages and 3 mammalian promoters used. The use of similar advanced technologies will assist future experiments that aim to use longer template regions for mutagenesis.

## 2.2. Designing Hybrid Promoters

The SPLs generated using random mutagenesis approaches have two major limitations: (i) The mutant promoters range in strength on either side of that of wild-type modules, and are more often weaker rather than stronger relative to the native promoter; however, a multitude of applications require only up- regulation of expression of specific gene products. For instance, the SPLs generated in budding yeast are weaker than the constitutive wild-type  $P_{GPD}$  and the inducible  $P_{GAL}$ , both being the strongest known natural promoters in the organism (Blazek et al., 2012; Da Silva and Srikrishnan, 2012). (ii) SPLs generally do not carry bi-functional modules, which may either activate or repress transcription output depending on the nature of inducible signals.



To overcome the above limitations, hybrid promoter libraries have been synthesized by stitching together multiple modular elements from different sources (Figure 2). The choice of modular elements also permits the designer to introduce a degree of orthogonality, which is necessary in transcription network design. The strategy had been used earlier in budding yeast (de Boer et al., 1983; Guarente et al., 1984; Guarente and Hoar, 1984; Guarente and Ptashne, 1981; Johnston et al., 1994; Rosenberg and Tekamp-olson, 1992; Sengstag and Hinnen, 1988). The recent advances in the field have expanded these HPLs and shed light into newer principles of natural and synthetic promoter architecture.

### 2.2.1. Use of UAS in hybrid promoters

A series of studies were recently conducted to increase the activities of the strongest native core promoters in *S. cerevisiae* and *Yarrowia lipolytica* (Blazec et al., 2013, 2012, 2011). These studies involved fusing single or multiple copies of UAS modules. UAS modules regulate TF localization around the core promoter and influence transcription initiation from them. These promoter chimeras can be tuned by either varying the strengths of the core promoters or the number of UAS repeats. Similar studies were previously conducted in *Y. lipolytica*, where 1 to 4 UAS elements, dubbed UAS1B, were fused to the minimal LEU2 promoter. This led to an almost linear up-regulation in the transcriptional output as the number of UAS1B repeats increased (Madzak et al., 2000). Following the same logic, Blazec et al., 2011 constructed promoter hybrids using LEUM core promoter and 1 to 32 UAS1B sequences. The highest recorded strength was 8-fold that of the native promoters. When the LEUM promoter was swapped with the stronger TEF1 core module, the strength of the chimeras was higher than that in the former. Similar trends were obtained when budding yeast was used as the host (Blazec et al., 2012). In this study involving budding yeast, an increase in UAS units resulted in a 2.5-fold increase in transcriptional output of  $P_{GPD}$ . This study also validated the design of hybrid UAS modules, which involved combining of different UAS motifs. The hybrid UAS element  $UAS_{CLB}-UAS_{CIT}-UAS_{TEF}$  yielded maximum  $\beta$ -galactosidase activity at 1,386 Miller Units, when compared with basal  $P_{GPD}$  promoter or chimeric promoters involving  $P_{GPD}$  along with 1-3 copies of  $UAS_{CLB}$  elements.

The full length TEF1 core promoter, which naturally consists of UAS1B repeats, also showed an increase in reporter gene activity, when more UAS1B elements were added upstream (Blazec et al., 2011). Additionally, these chimeras showed extraordinary integrity and stability even after several generations. The use of chimeric UAS elements along with the core promoter, precludes the possibility of homologous recombination and consequent loss of stability of synthetic a promoter chimera (Blazec et al., 2012). Collectively, these strategies show promise in designing robust hybrid synthetic promoters that are stronger than the strongest native promoters available today.

Traditional hybrid promoter construction depends on native modular elements. Their scope is thus limited in relatively obscure organisms like *Y. lipolytica*, where only a few modular elements are well characterized. Therefore, one needs to discover many similar activating elements. Recently, Blazec et al pioneered such a study for the *Y. lipolytica* system (Blazec et al., 2013). Similar studies are necessary to expedite progress in this area.

Specific and non-specific upstream activating sequence (UAS) elements have also been used to regulate synthetic promoter inducibility in *S. cerevisiae* (Blazeck et al., 2012). The use of constitutive UAS modules like UAS<sub>CLB</sub> and UAS<sub>CIT</sub> along with the inducible GAL1 promoter led to a decrease in glucose-mediated repressive effects; this decrease was proportional to the number of UAS repeats. On the other hand, the use of one or more inducible UAS<sub>GAL</sub> elements along with the GAL1 promoter led to a 7% dip in transcription output when compared to native levels under glucose repression. Also, this led to a 15% increase in gene expression in the presence of galactose. An important finding of this study was that the chimeric promoters created using different constitutive core elements and inducible UAS elements, showed variable results. When the core element P<sub>LEUM</sub> was used, the behavior of the promoter was similar to the GAL1 promoter in inducible conditions. Other constitutive promoters like the P<sub>GPD</sub> and P<sub>TEF</sub> showed different results in that they showed significant transcription activity in glucose rich media, which increased in galactose inducible conditions. This result showed that different levels of inducibility and repressibility may be achieved by using different core promoter and UAS modules for promoter engineering.

### 2.2.2. Use of operators for developing hybrid promoters

HPLs have also been generated by fusing tetO and Olac operators along with the core GAL1 promoter, resulting in predictable patterns of expression, when tested *in-vivo* (Ellis et al., 2009). One of the major problems of inducible systems is transcriptional noise. In the uninduced state, some amount of leaky transcription is evident in various systems. One study tested the leakiness of the core GAL1 promoter in relation to the number and positioning of tetO<sub>2</sub> operators with respect to the TATA box (Murphy et al., 2007). The repression ability of promoters increases, when operators are positioned closer to the TATA box, so that the basal transcriptional output decreases. However, the transcription noise was also found to increase, as opposed to popular belief. Transcription noise can be programmed and fine-tuned through creation of HPLs, thus throwing light on the importance of this phenomena in cellular homeostasis and function; this will also be critical to the success of synthetic network design by ensuring accurate temporal regulation of the various elements of the network.

### 2.3. Regulating promoter inducibility

It is possible to program and regulate the inducibility of promoters using promoter hybridization strategies. Inducible HPLs are generated by either fusing the core promoters with specific enhancers or repressors upstream, or operators downstream. These modules serve as TFBSs in an inducer dependent manner, although the core promoters may or may not be themselves inducer regulated.

In *S. cerevisiae*, it was observed that the fusion of galactose responsive UAS to the core promoter led to an increase in galactose induced transcription activation and glucose induced repression. The core promoter may itself be inducible or constitutive, thereby augmenting or counteracting the inducibility of the hybrid promoter construct (Blazeck et al., 2012). In another study, it was possible to design galactose inducible promoters with outputs comparable to the native yeast GAL1 promoter by fusing the core promoter with a condensed galactose responsive

UAS (Redden et al., 2015). Rajkumar et al., 2016 engineered synthetic promoters responsive to low pH by fusing the core promoter with multiple TFBSs that are active in acidic environments. Using this strategy, they could program low-pH inducibility into the promoter of the CCW14 gene that was originally unresponsive to low pH.

Various native inducible promoters have been isolated in plants, and have been shown to respond to different environmental cues (i.e. light, cold and heat stress), biotic and abiotic stress (pathogens, wounding, insects, drought and salinity), hormones (i.e. ethylene, auxin, abscisic and salicylic acid) and chemicals (i.e. tetracycline, copper, estradiol and dexamethasone) (Venter, 2007). These promoters are of great interest for programming inducibility in synthetic promoter engineering (for a review, see Mehrotra et al., 2011). Liu et al., 2011 constructed a phytosensitive promoter composed of *cis*-regulatory elements that are responsive to salicylic acid, jasmonic acid and ethylene, along with multiple phytopathogenic bacteria. They used this promoter fused to a fluorescent reporter to create a transient phytosensitive system responsive to infection by various phytopathogenic bacteria. In a subsequent study, this promoter system was applied to create stable phytosensitive transgenic plants, which may be used to create reliable plant phytosensors that allow us to closely monitor host-pathogen interactions (Liu et al., 2013).

The same individual synthetic promoter may be programmed with inducible activation and repression in response to different environmental cues. Such bi-functional promoters will possess both the features discussed above – inducible response elements (REs) upstream and operators downstream of the core promoter element. The REs and operators respond to different environmental cues and execute a corresponding change in function. (Mazumder and McMillen, 2014) created a hybrid promoter containing the following elements: (i) five human androgen responsive elements (ARE), (ii) the TATA box region of the minimal cytochrome C promoter and (iii) the lac operator. The lac operator is the binding site for the lacI repressor in the absence of lac inducer. In the presence of the lac inducer or its chemical analogue isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG), the repression is lifted. The transcription output is further enhanced in the presence of the hormone testosterone, which leads to recruitment of transcriptional activators to the AREs. The response curves varied with different concentrations of the cognate inducers, either IPTG or testosterone. Further research on this module will effect fine-tuning of expression from this hybrid promoter and increase its utility in synthetic biological applications. In a more recent study, promoters that were responsive to tetracycline and light were synthesized by hybridizing the TET operator and UASs that bind the blue- light responsive trans-activator GAVPO (Chen et al., 2016). These promoters were used to create synergistic and antagonistic dual-input circuits in mice. Under synergy, the system behaved as an OR logic gate, and it functioned as an AND gate in the antagonistic design. Further, the variation of relative numbers of TET operators and the light-REs affected the intensity of regulation by either environmental cue. This shows that it is possible to vary the quantity of output in multi-input circuits without changing the concentration of the inducer. Inducer costs and inducer-mediated toxicity are key challenges to the application of inducer circuits. Hybrid promoter engineering makes it easier to regulate the output of synthetic circuits without having to alter the concentration of the applied inducer.

The leakiness of independent modules in complex engineered networks may effect undesired cross-talk between them. This decreases the orthogonality and consequently the efficiency of the system. As shown in the study by Chen et al., 2016, multi-input circuits display greater stringency and reduced leakiness in regulation. Similar circuits may be useful to program greater orthogonality and spatiotemporal resolution into our designs of synthetic transcriptional networks.

The power endowed using artificial promoters that can switch functions in the presence of different signals is enormous with respect to metabolic engineering and systems' design. Multi-input transcriptional circuits are extremely useful in environments characterized by frequent environmental perturbations as indicated by previous theoretical studies (Ang et al., 2010; Ang and McMillen, 2013). When combined with the design of allosteric synthetic transcriptional regulators, the power of such promoter systems is immense. We will underscore this principle again in a later section of this review by revisiting this property of inducibility in the light of synthetic transcription factor design.

## **2.4. De-Novo Promoter Design Technologies**

Most of the current promoter design methodologies rely on native promoter modules, which are then utilized to develop synthetic promoters by random mutagenesis or hybrid promoter approaches. Sometimes this becomes a limitation. In extensively researched systems, the use of synthetic promoters bearing homology to the native modules brings in additional complications such as HBGS in plants and homologous recombination in yeast systems.

More recently, advances in computational biology and the presence of high-throughput data associated with promoter architecture have facilitated *de novo* development of promoters. This area of research holds great promise in future. Here, we illustrate this by taking the examples of two recently developed models for promoter engineering *de-novo*: (i) design of synthetic promoters using a nucleosome occupancy model and (ii) design of artificial promoters based on RNAP Binding Affinity model. Both these principles use a primary computational backbone to construct novel *cis*-regulatory elements *ab-initio*.

### **2.4.1. Nucleosome Occupancy Models for Promoter Engineering**

The close relationship between chromatin architecture and promoter strength, as well as the links between nucleosome occupancy and transcription output, have been validated by past studies (Lam et al., 2008; Sharon et al., 2012). Curran et al., (2014) sought to test the relationship between promoter strength and nucleosome occupancy in *S. cerevisiae*. They used a previously developed hidden Markov model (Xi et al., 2010) to predict the nucleosome affinity of arbitrary DNA sequences *de-novo*. After validating the conjecture that promoter strength varied inversely with nucleosome affinity, the scientists sought to create strong promoters by indirectly designing sequences with low nucleosome binding affinity. After successive rounds of minimization of the cumulative affinity score for nucleosome binding, the scientists could direct specific mutations in 4 native yeast promoters; this experiment increased the strengths of these wild type promoters beyond native levels in single and alternative genetic contexts. To take this approach a step further, the group used this method for the generation of *de-novo* synthetic promoters from base

scaffolds containing common yeast glycolytic TFBS embedded in random spacer sequences. This generated a library of 6 artificial promoters that had a similar or higher reporter gene expression profile *in-vivo*, when compared to the native CYC1 promoter in yeast. Additionally, the method ensured minimum sequence homology between these promoters and the native *cis*-regulatory elements, making them extremely stable systems *in-vivo*. Similar models will help in advancing synthetic engineering technologies in eukaryotic systems, which are complicated by the presence of nucleosomes in and around regulatory regions.

#### **2.4.2. RNA-Polymerase Binding Affinity Models in Promoter Design**

This approach was used in prokaryotic systems, but we will briefly present it here to highlight its potential scope in synthetic biology. Multiple parameters including the copy numbers of TFs acting on the target gene, the strength of ribosome binding sites, RNAP binding affinity, and decay rates of the gene product, control gene expression output. Two of these parameters were used by Brewster et al., (2012) for promoter engineering: RNAP binding affinity and TF copy number (in a repressive context). Here the native *lac* and *lacUV5* promoters from *E. coli* and introduced directed mutations to generate promoters with widely varying RNAP binding affinities calculated per a computationally derived equation. A wide variety of promoters were generated with expression outputs ranging from 50-fold lower to 10-fold higher than native promoter levels. Subsequently, the promoters were designed per the desired expression level based on a thermodynamic model which evaluated the RNAP binding probability to a given sequence. Previous high throughput mutagenesis experiments generated RNAP binding affinity matrices (Kinney et al., 2010), which were used in this thermodynamic model. When the promoter designs were tested *in-vivo*, the transcription outputs correlated well with the theoretically predicted values. Such quantitative approaches to synthetic promoter design may be used to substitute random mutagenesis based approaches to obtain the desired gene output. The *lacI* based repressive system was modeled in a similar fashion, but faced poor correlation with experimental data due to the absence of refined theoretical models that accurately represent such systems. This posits a need for progress in theoretical and computational biology, to drive further advancement of similar methods for synthetic promoter design in more complex systems.

### **3. Transcription Factor Engineering**

TFs play an important role in inducing transcription and in determining the specificity of transcription. Thus, TFs can be suitably engineered to serve any specific purpose including the need to prevent non-specific transcription. The latter is referred to as orthogonality (Rao, 2012). One of the major challenges posed by the current methods of promoter engineering is the utilization of cells native TF machinery by the the given synthetic promoter. This increases the propensity for firing secondary targets, which is undesirable due to multiple reasons. Activation of off-target modules increases the metabolic burden for cells; this hampers their survivability and growth. It is desirable to minimize transcriptional noise produced by extraneous cross-talk between independent modules of engineered transcriptional networks as these affect their robustness and predictability. Products of off-target gene expression may interfere with transcription from the artificial networks thereby reducing their efficiency. Therefore, it is imperative to engineer orthogonality in the artificial systems for increasing their on-field success

and utility. In the previous section, we discussed approaches to engineer orthogonality in artificial networks with respect to *cis*-engineering. In this section, we will briefly revisit this subject in the light of modern TF engineering.

One of the traditional ways to engineer orthogonality in eukaryotic systems is the use of bacterial TF – promoter pairs (Lu et al., 2009; Weber and Fussenegger, 2009) as the prokaryotic and the eukaryotic transcriptional machineries are orthogonal to each other. However, these approaches are limited by the low flexibility of bacterial TFs and their propensity to oligomerize cooperatively when bound. Both these properties are closely interrelated thereby making it laborious to fine-tune the activities of prokaryotic TFs; it requires repeated rounds of re-engineering to create a wide repertoire of TFs required for synthetic biology (Khalil et al., 2012). Recently, three more methodologies have evolved for creating orthogonal sTFs – (a) Zinc finger TF (ZnF) engineering, (b) use of TALE TFs for customized synthetic engineering; (c) CRISPR/Cas9 based methods for TF engineering. These technologies have significantly expanded the application of synthetic promoters. Therefore, we shall briefly discuss them here.

### 3.1. Artificial ZNFs

sTFs have been engineered in recent times using the characteristic Cys<sub>2</sub>-His<sub>2</sub> ZNF domains. These modular DNA binding domains (DBD) are capable of binding to specific sequences in engineered promoters (Figure 3). When fused with specific activation domains for recruiting transcription initiation machinery along with protein-protein interaction domains for achievement of cooperativity, these sTFs can be used for fine-tuning gene expression from specific *cis*-regulatory elements. ZNFs are naturally used in native systems to solve the problems of combinatorial binding; they bind to promoters in tandem arrays (Pabo et al., 2001). In the wake of recent methods such as oligomerized pool engineering (OPEN) (Maeder et al., 2008) and other contextual techniques (Sander et al., 2011), it has been possible to engineer a large repertoire of ZNFs with defined specificities towards signature *cis*-regulatory motifs.

A recent study by Khalil et al., (2012) illustrated the power of ZnF engineering in programming orthogonality into regulatory contexts. Using the OPEN framework, the study could engineer ZNF-DNA interaction specificity and consequently create a large library of interacting partners. Using the same approach, 19, 3-finger arrays, which possessed mutually orthogonal binding specificities were identified. These sTFs were shown to produce a transcription output 1.3-6.6-fold native levels, when bound to their respective cognate promoters; cross reactivity was absent. It was also found that this output could be increased by up-regulating the number of operator sequence repeats, a result that was observed in many of the previously discussed promoter engineering methodologies (Blazek et al., 2013, 2012, 2011). Specific properties to sTFs were also engineered including cooperativity and alternate functional states. The latter hinted at the diversity of functions (activating, repressive, or neutral) that different orthogonal sTF combinations could provide in two-input systems. Such systems are useful tools in systems' engineering, where the same TFs or *cis*-regulatory elements may be used to achieve alternate functionalities. This reduces the number of modules that need to be introduced and consequently eases the metabolic burden of the cells. Another novelty of this study was the use of three finger arrays instead of the previously used two-finger arrays (Wang et al., 2001; Wolfe et al., 2000).

This permits the recognition of 18 bp long DNA elements. This is long enough to be potentially unique in eukaryotic and more specifically in mammalian genomes. Engineering specific phosphate backbone mutations in the ZNF DBDs conferred greater specificity on the sTFs; this was apparent in lesser off-target induction by mutant sTFs.

McIsaac et al., (2013) employed three or four finger ZNF arrays for the design of sTFs. They were able to elicit fast, graded responses in *S. cerevisiae* without significant off-target responses unlike previous Gal4p based systems (McIsaac et al., 2011). The four finger arrays used in this study are an example of rational TF engineering; they were designed to bind to a cognate promoter sequence that is not present in the yeast genome. The study also showed how similar engineered sTFs could be used to decipher transcriptional networks. This is proof again of the importance of synthetic biology in basic research.

Activation domains may be used for multiple purposes. In both the above examples, they were used to confer inducible properties in the sTFs. In another study, the basic ZNF DBDs were fused to chromatin regulator (CR) domains. These TF chimeras were then utilized to study the roles of over 223 known CR classes in transcriptional regulation. Similar ZNF based targeting strategies could help in designing unique regulatory modules that focus on epigenomic rather than genomic targeting (Keung et al., 2014). In future, efficient strategies for protein-promoter interaction engineering like the ones followed by Khalil et al. (2012) could allow us to silence or induce specific *cis*-regulatory modules through a purely epigenome centered approach.

### 3.2. TALEs

Another class of proteins that have recently gained importance in synthetic biology are the TALEs. Originally identified in *Xanthomonas sp.*, these proteins contain 33-35 amino acid repeat domains that possess a single nucleotide specificity by amino acid identities at two positions dubbed as repeat variable di-residues (RVDs). The TALE repeats for each of the four nucleotides is known. With this knowledge, one can custom design TALE based sTFs for almost any DNA sequence based on a simple motif-to-base recognition principle. This increases the programmability of sTFs for extended stretches of DNA (Boch, 2011; Bogdanove and Voytas, 2011; Reyon et al., 2012). In sTFs the TALE repeat arrays are used to create the DNA binding domain, which may be coupled to a trans-activator or repressor domain (Figure 4). Previously, Garg et al., (2012) used TALE DBDs fused with a KRAB repressor module to achieve 36 to 97-fold repression of a fluorescent reporter. Further, these sTFs could achieve near complete silencing of target gene expression when co-expressed with shRNAs. Combining traditional sTF technologies with post-transcriptional regulation may enhance the efficiency of transcriptional control *in-vivo*, as we shall see in our subsequent discussion. Alternately, TALE based sTFs without trans-repressor fusions can antagonize transcription by sterically hindering the assembly of the basal transcriptional machinery (Li et al., 2015).

When combined with approaches to generate better and more efficient SPLs, TALE engineering presents enormous promise in programming orthogonality and fine tuning of gene expression in various systems. As a proof of concept, a study was performed in *S. cerevisiae* by Blount et al., (2012); synthetic promoters were designed from a PFY1 core by a two-step approach- (a)

random mutagenesis of a 48 bp fragment of the PFYp core promoter; and, (b) fusion of TetR operator regions to program TetR mediated repression (iPFYp). To make the design more robust, TALE orthogonal repressors (TALORs) were designed for both, PFYp and iPFYp. These TALORs showed low off-target induction. It was noted that the TALORs designed for one kind of promoter couldn't regulate the other despite their target recognition sequences sharing a 9 bp homology. Such a degree of specificity is critical for designing synthetic networks, where orthogonality between various component modules is indispensable. This study also illustrates the scope of combinatorial engineering of TALEs in synthetic biology. To quote the authors, "... with the combination of the synthetic promoter libraries and TALORs, any limits on the complexity of synthetic networks due to the lack of orthogonal regulated promoters would be effectively removed" (Blount et al., 2012).

However, the specificity of TALE based sTFs has come under scrutiny as some computational and experimental studies have reported off-target activity for TALE based nucleases in more complex genomes including the human genome (Fine et al., 2014; Osborn et al., 2013). Although the data from these studies concerns TALE nucleases, the authors emphasized the need for a detailed analysis of the targeting specificity of TALE based sTFs. Also, the engineered TALEs are sensitive to 5-methylcytosine (5mC) and the RVD corresponding to cytosine does not bind to the 5mC (Bultmann et al., 2012). This is significant in complex genomes like the human genome because they are abundant with CpG islands containing 70-80% methylated cytosines. Valton et al., (2012) and Deng et al., (2012) reported that the thymidine cognate RVD, NG was able to bind to the methylated cytosine. Here, targeting of methylated cytosines comes at the cost of specificity. Thus, there is a need for better characterization of TALEs to improve the predictability of TALE based systems. This is essential for their application in complex organisms.

TALE-like proteins have been discovered in different species (de Lange et al., 2014; De Lange et al., 2015; Stella et al., 2014). Although these proteins are conserved in sequences proximal to the base specifying residue (BSR), they show polymorphism in the remaining spaces of the repeat domains. Such polymorphism is absent in TALEs derived from *Xanthomonas* sp. The TALE-like proteins isolated from marine bacteria, dubbed MO<sub>r</sub>TL1 and MO<sub>r</sub>TL2, showed significant variation around the BSR also. These polymorphisms show interactions of varying strengths with the DNA (De Lange et al., 2015). Such natural TALE like proteins will help in engineering TALE based sTFs with an enhanced ability to fine tune their function. De Lange et al., (2015) also showed that chimeras of MO<sub>r</sub>TL1 repeats and traditional TALE or Bat repeat arrays (sourced from *Burkholderia rhizoxinca*) show high orthogonality. Use of such chimeras allows control of stringency as different repeats make a distinct contribution in the preferential selection of one binding site over another. In future, such chimeric protein libraries with variable binding degeneracy may be used to regulate a family of genes sharing highly similar TFBSs like, for example, the Hox gene family, which have high utility in developmental biology research.

Despite these opportunities offered by TALE based sTFs, some key issues remain to be addressed. One significant problem associated with the application of TALEs is their delivery into target cells. TALEs are neutral proteins unlike ZNFs (Gaj et al., 2012), precluding their



transport across the plasma membrane. However, TALEs conjugate them with positively charged peptide sequences to circumvent this issue (Liu et al., 2014). Secondly, the repetitive nature of TALEs decreases the genomic stability of their corresponding DNA sequences due to their high recombination propensity (Holkers et al., 2013). Therefore, the most preferred technique for introducing TALEs *in-vivo* is the transient delivery of their corresponding mRNA or DNA. The extensive structural and functional characterization of ZNFs has expanded their application in rational engineering. Future research on TALEs in this direction will expand their use in synthetic biology.

### 3.3. CRISPR/Cas9 based sTFs

Although extremely useful in the design of orthogonal transcription networks, both the ZNF and TALE centered approaches involve tedious selection processes or complex DNA hybrid assembly procedures (Reyon et al., 2012; Sanjana et al., 2012). The success of these protocols is also limited by the number of different TFs that can be simultaneously synthesized by the cells without imposing a significant burden. As an alternative to these two methods, the RNA-guided bacterial CRISPR/Cas9 system has been used extensively to engineer highly specific genome regulation.

Using small customizable guide RNAs (gRNAs), the Cas9 endonuclease may be targeted to different regions in the genome, where it introduces double stranded breaks upstream of the protospacer adjacent motif (PAM). Cong et al., (2013) used this mechanism to introduce specific insertion/deletion (indel) mutations in the genome. Such a mechanism may have high utility for advanced promoter engineering approaches, whereby insertions and deletions may be programmed into *cis*-regulatory regions. Similar genome editing principles have been used for *cis*-engineering in plant systems (Upadhyay et al., 2013).

RNA guided sTFs can be created by fusing endonuclease deficient Cas9 (dCas9) (Qi et al., 2013) with different activator or repressor domains. These chimeric TFs are then targeted to the target *cis*-regulatory motifs to facilitate gene regulation (Figure 5) (Gilbert et al., 2013). Due to its large size, dCas9 gives a further advantage in simple organisms like *E. coli* and *S. cerevisiae*. It can significantly repress transcription without needing repressor fusions by sterically obstructing the assembly of transcription factors at the proximal promoter or the movement of the RNAP II during elongation (Kabadi and Gersbach, 2014). However, significant repression requires heterochromatin recruiting repressor domains like KRAB in mammalian cells. As far as activation is concerned, dCas9 fused to VP64 transactivation domain leads to approximately three-fold upregulation of target gene expression in *S. cerevisiae* (Farzadfard et al., 2013). But the activation efficiencies of the simple dCas9 fusion is lower than that of ZNFs and TALEs in mammalian systems (Jusiak et al., 2016). There is a need to simultaneously recruit multiple trans-activators for significant activation.

This can be achieved by tiling multiple gRNAs around the TSS or proximal promoter of the target gene, which increases the density of dCas9 fusion proteins upstream of the gene provides a strategy to recruit multiple trans-activators. In an earlier study, it was observed that an increase in number of gRNA target operator sequences upstream to the target gene led to an increase in transcription output in yeast and mammalian cells. In HEK293T cells, a 56 -fold increase in output was observed corresponding to a 3-fold increase in gRNA repeat number (Farzadfard et al., 2013). Although this approach is straight-forward, it limits the ability of the system to

simultaneously regulate multiple gene targets – a requirement as far as complex synthetic gene networks are concerned.

Recent studies have increased CRISPR based activation by creating activator systems through engineering multimodal TF interactions (Maeder et al., 2013; Perez-Pinera et al., 2013). The strength of the CRISPR activator (gRNA/dCas9-VP64) could be increased by recruiting additional transactivation domains to MS2 binding RNA aptamers that were engineered into the gRNA (Konermann et al., 2015). Tanenbaum et al., (2014) increased the number of VP64 activators recruited to the dCas9 by fusing it with peptide repeats dubbed SunTag. Zalatan et al., (2015) increased the recruitment of the VP64 fusion protein to the target locus using an RNA scaffold that simultaneously recruits dCas9 and RNA-aptamer binding proteins like MS2 or PP7 fused to transactivator domains, while targeting the protospacer through its 3'-end. This study also showed the power of CRISPR/Cas activators and repressors in regulating the outputs from metabolic gene expression cascades.

CRISPR/Cas based sTFs show variable regulation of target loci depending on where they are targeted relative to the TSS. dCas9-VP64 TFs could be used to program activation and repression of target promoters in *S. cerevisiae* and HEK293T cells (Farzadfard et al., 2013). The study revealed that the TF chimera repressed transcription from the core promoter when they were targeted to downstream sequences and showed activation when gRNAs were specific to sequences upstream of the core motif.

Thus, CRISPR/Cas based TFs make for good bidirectional regulators that may be used to switch on or switch off a gene depending on where they are targeted (Jusiak et al., 2016). Farzadfard et al., (2013) noted that the gRNA mediated targeting of a hybrid TF was extremely sensitive to the target sequence. Even a single nucleotide mutation in the gRNA target motif led to complete loss of regulation by a dCas9-VP64 hybrid. The specificity of CRISPR based sTFs is greater than that of ZNFs, TALEs or even the Cas9 nucleases, as indicated by RNA-Seq studies in mammalian systems (Gilbert et al., 2013; Konermann et al., 2015). Jusiak et al., (2016) suggested that this may be due to the high position-specific activity of CRISPR based TFs. The high sensitivity and modularity of the CRISPR/Cas9 based sTFs can be exploited in regulating complex transcriptional networks, where each component of the circuit is controlled by a unique gRNA/sTF combination. It is possible therefore to orthogonally control the various genes of such a network by targeting a unique gRNA to each locus involved. In this regulatory design, the dCas9 is common to all modules being regulated. Zalatan et al., (2015) could control the final output of a multi-branched bacterial biosynthetic pathway in yeast by independently targeting three of the five genes involved. In another study, the simultaneous activation of up to 10 genes mediated by co-transfecting ten gRNAs into cell culture has been reported (Konermann et al., 2015)

It is also possible to selectively program different parts of a synthetic circuit using Cas9 orthologs with altered PAM specificities (Esvelt et al., 2013). The orthogonal Cas9 could be used to independently target different genes in the same cell by expressing them from separate inducible promoters. Such a system facilitates a seamlessly switch between different cellular states that are each dependent on the expression of a specific Cas9 variant. The combinatorial engineering of promoters and transcriptional factors will assist in programming greater dynamism and specificity into synthetic transcriptional networks. Further, it is significantly

easier to design dynamic circuits with CRISPR/Cas9 based sTFs relative to those based on ZNFs or TALEs, as the turnover rate of cellular RNAs is faster than the rate of protein degradation (Kabadi and Gersbach, 2014).

Further, the expression and design of gRNAs is faster and more convenient when compared to the design and assembly of complete proteins like ZNF or TALEs. But the true power of synthetic circuits is unlocked when they can be switched on or off by specific physical or chemical triggers. The common strategy for engineering induction is expression from inducible RNAP II promoters. In the context of CRISPR/Cas9 based systems, this is a major challenge as the gRNAs expressed from inducible RNAP II promoter systems are not retained within the nucleus – a critical requirement for their activity in eukaryotes. The nuclear retention of gRNAs, which is currently achieved through their expression from constitutive RNAP III promoters comes at the cost of a low induction potential of the CRISPR/Cas9 systems (Jusiak et al., 2016). This major drawback has been addressed by either regulating the expression of the dCas9 based TF or through innovative gRNA expression systems, which are discussed in the next section.

### 3.4. Programming induction of gene expression

Programming induction potential into sTFs increase their utility in synthetic biology by enhancing their properties. sTFs may be designed for promoters with less sequence homology to native systems. They can variably regulate expression levels in response to varying inducer concentrations and signal strengths. This makes a given circuit element more dynamic. Also, they display limited cross-reactivity by selectively responding to specific inducers. sTFs can be made responsive to physical stimuli including light (Nihongaki et al., 2015; Polstein and Gersbach, 2015) or chemical ligands (Rogers et al., 2015). Such inducible sTFs are also termed allosteric transcription factors (aTFs) in keeping with the nomenclature of the native inducible proteins. Herein, we shall use these terms interchangeably.

Allosteric transcription factors (aTF) are very important regulatory proteins present in the bacterial system. Recent studies have shown that naturally occurring aTFs can be engineered to bind and thus respond to effectors beyond the natural inducer-aTF pairs (Taylor et al., 2016). Taylor et al., (2016) engineered aTFs responsive to new ligands, fucose, lactic acid and sucralose, by using computational protein design, single residue saturation mutagenesis or random mutagenesis coupled with multiplex assembly. The experimental results showed that the engineered variants had induction potential and with a specificity comparable to the wild type LacI. Similar libraries of variants created through combinatorial design can be engineered to have a broad range of specificities and inducer responsiveness. This expands their utility in multiplexed transcriptional circuits.

Protein domains responsive to specific inducer(s) can be fused to a sTF to make them inducible. Mercer et al., (2014) fused a TALE based activator to ligand binding domains from steroidal receptors. The activity of these sTFs could be tuned by varying the concentration of the respective inducer molecule. Polstein and Gersbach, (2012) designed a light inducible sTF that had a ZNF DBD. Using photo inducible proteins from *Arabidopsis thaliana*, they synthesized bipartite sTFs whose individual subunits combined to form a fully functional activator upon light

exposure. The transcriptional output of this systems could be tuned by varying the intensity of input signal.

CRISPR/Cas9 based sTFs may be made inducible by making either the Cas9 or gRNA inducer responsive. Modular approaches like the ones mentioned above are being explored for Cas9 also. Oakes et al., (2016) identified regions in Cas9 that are capable of tolerating insertions without affecting its function significantly. With this knowledge, they designed an allosterically regulated dCas9, which had an estrogen receptor ligand binding domain. Nihongaki et al., (2015) synthesized a light inducible synthetic activator by fusing dCas9 with photo-sensitive cryptochrome 2.

Inducible expression of gRNA synthesis is challenging. gRNA that is transcribed using RNAP II cannot bind to the target DNA sequence as RNAP II transcripts are exported out of the nucleus. This prompts the use of constitutive RNAP III promoters for expressing gRNA. Some novel techniques have been developed for inducible CRISPR/Cas systems. One of these is the incorporation of the gRNA into the introns of a protein coding gene that is inducible (Nissim et al., 2014). Herein, the gRNA sequence is flanked by splicing signals and is subsequently spliced out and would become available for targeting DNA. Alternately, the gRNA could be flanked by self-cleaving ribozymes (Kiani et al., 2014) or Csy4 endonuclease sites which would result in cleavage post transcription, with their subsequent availability for targeting. The Csy4 endonuclease is not found endogenously in the eukaryotic system. Thus, a CRISPR based system may be induced by inducible expression of Csy4, providing an additional level of control (Nissim et al., 2014).

Collectively, these data imply that induction potential is a very important feature in transcriptional engineering, as spatial and temporal regulation of expression is a key requirement to make the fullest application of synthetic biological systems. To achieve complexity that parallels natural systems it becomes essential to control multiple nodes of a branched network. In depth characterization is necessary to ensure orthogonality and limited cross talk between multiple elements of a circuit and also with the host genome. Designing sTFs responsive to molecules other than native inducers can ensure specificity of the sTFs. Further, the addition of similar inducible domains to different sTFs that target different genes of a complex circuit, could lead to simultaneous regulation of multiple nodes in response to a single inducer. Also, sTFs that need multiple inducers to activate have caught the eye of synthetic biologists due to the complexity that can be engineered into circuits. Similar designs for programming induction potential advance the abstraction of components of synthetic circuits. This would lead to a drastic increase in the complexity of the resulting circuits as designers would now have a wide range of elements, inducible by a plethora of potential inputs.

However, we note that such systems require rigorous protein engineering in most cases. This is challenging due to the time and labor involved. Extensive characterization is required to test the orthogonality, specificity and induction potential of these designed sTFs. Also, the input-output characteristic curves for each expression system and their corresponding inducer(s) requirements to be evaluated. Such characterization may be difficult to generalize for an inducible element as in the case of the modified LacI based aTFs responsive to other inducers. Every inducer –

inducible element pair will be expected to show a different response characteristic and each one needs to be studied carefully. However, the promise of such systems have prompted research in the field and a number of these modules have already been characterized for input – output properties (Rogers et al., 2015). With a growing numbers of associated libraries available, opportunities are increasing rapidly in this area. Table 1 briefly summarizes the scope and limitations of the transcriptional engineering techniques discussed above.

Technique	Scope	References	Limitations	References
<b>Promoter Engineering</b>				
<b>Synthetic Promoter libraries</b>	Universal across strains and conditions	(Alper and Fischer, 2005)	Involves laborious screening of libraries	(Alper and Fischer, 2005; Hammer et al., 2006; Solem and Jensen, 2002)
	Wide range of expression levels	(Alper and Fischer, 2005; Siegl et al., 2013)	Lower expression levels than native promoter usually. Bifunctional modules absent.	(Blazeck et al., 2012; Da Silva and Srikrishnan, 2012)
	Several SPLS are available with advent of high throughput techniques.	(Patwardhan et al., 2009)		
<b>Hybrid Promoter Libraries</b>	Wide range of expression levels, with high expression modules available too.	(Blazeck et al., 2013, 2012, 2011; Madzak et al., 2000)	Limited application in less characterised organisms	(Blazeck et al., 2013)
	Orthogonality and Bi-functionality	(Blazeck et al., 2012; Chen et al., 2016; Mazumder and McMillen, 2014)		
	Increased stability and lowered recombination propensity	(Blazeck et al., 2012)		
	Fine tuning of transcriptional noise	(Chen et al., 2016; Murphy et al., 2007)		
<b>De novo promoter design</b>	Lowered homologous recombination propensity	(Curran et al., 2014)	Low correlation with experimental data in some cases requiring better models and computational tools.	(Brewster et al., 2012)
	Wide range of expression activity obtained	(Brewster et al., 2012)		
<b>sTF engineering</b>				
<b>ZNF based sTFs</b>	No significant off-target activity.	(Khalil et al., 2012)	Require protein engineering and laborious selection and DNA hybrid assembly	(Khalil et al., 2012; Reyon et al., 2012; Sanjana et al., 2012)
	Recognition of long sequences confers orthogonality.	(Wang et al., 2001; Wolfe et al., 2000)		
<b>TALE based sTFs</b>	Simple motif-to-base recognition principle	(Boch, 2011; Bogdanove and Voytas, 2011; Reyon et al.,	Require protein engineering and laborious selection and	(Reyon et al., 2012; Sanjana et al., 2012)

		2012)	DNA hybrid assembly	
	Orthogonal and highly specific	(Blount et al., 2012)	Off-target activity of TALENs in complex genomes	(Bultmann et al., 2012; Deng et al., 2012; Fine et al., 2014; Osborn et al., 2013; Valton et al., 2012)
	Can antagonize transcription via steric hindrance without trans-repressor fusions	(Li et al., 2015)		
	TALE-like proteins allow enhanced tunability and control of stringency for recognition	(de Lange et al., 2015, 2014; Stella et al., 2014)	Challenging to transport protein across cell membrane	(Gaj et al., 2012; Liu et al., 2014)
			High recombination propensity of the respective DNA sequence.	(Holkers et al., 2013)
<b>CRISPR/Cas9 based sTFs</b>	Target selection requires design of short nucleotide sequences.	(Qi et al., 2013)	Activation efficiency of dCas9 fusions is reportedly lower than that of ZNFs and TALEs in mammalian systems.	(Farzadfard et al., 2013; Jusiak et al., 2016; Konermann et al., 2015; Maeder et al., 2013; Perez-Pinera et al., 2013; Tanenbaum et al., 2014; Zalatan et al., 2015)
	Can cause repression by steric hindrance, without the need for repressor fusions	(Kabadi and Gersbach, 2014)		
	Specificity of CRISPR/dCas9 based systems is higher than ZNFs, TALEs, Cas9 nucleases.	(Gilbert et al., 2013; Konermann et al., 2015)		
	Orthogonal control of various genes is possible	(Konermann et al., 2015; Yang et al., 2015; Zalatan et al., 2015)	Conferring inducibility to CRISPR/ Cas9 based systems is challenging	(Jusiak et al., 2016; Kiani et al., 2014; Nihongaki et al., 2015; Nissim et al., 2014; Oakes et al., 2016; Shechner et al., 2015)
	Cas9 orthologs allow orthogonal control in synthetic circuits	(Esvelt et al., 2013)		
	Dynamic in nature due to more rapid RNA turnover rate	(Kabadi and Gersbach, 2014)		

Table 1: Scope and limitations of various promoter engineering and transcription factor engineering techniques

## 4. Discussion

Every synthetic transcriptional circuit is composed of six core features that influence its utility: (a) Predictability and tunability of transcriptional outputs, (b) Amenability or the dynamic nature of the modules, (c) Orthogonality, (d) Specificity, (e) Scalability, and; (f) Ease of Handling. In this review, we have discussed some modern methodologies for engineering synthetic transcriptional modules in the light of these six features.

There is a need to shift towards more rational ways of engineering biological modules while carefully considering these six parameters. A combinatorial approach seeking to integrate promoter and transcription factor engineering will promote this cause. A proof of this concept is

the study by Blount et al., (2012), which combined random mutagenesis based approaches, hybrid promoter construction and TALE factor design principles to construct robust transcription modules in *S. cerevisiae*. In another study, a library of synthetic promoters was constructed with a common 18 bp motif that binds a designer TALE. The various promoters in the library could be used as orthogonal modules in plant systems yielding a broad range of expression strengths in comparison to the CMV promoter (Brückner et al., 2015). It is possible to rationally progress towards stronger expression systems by iteratively engineering and optimizing the design of promoters and their corresponding TFs (Leavitt et al., 2016).

Synthetic promoters and TFs contain one or more invariant sequences whose biological functions have been well characterized. Presently, most design algorithms stitch together these sequence spaces with randomized spacers. The spacers are subsequently varied to generate a library of synthetic promoters and TFs. The libraries are screened either computationally or experimentally to deduce the most optimal modules that fit the experimenter's requirements. The efficiency of the design is affected by the length of the variant sequence spaces. By minimizing these lengths, it is possible to construct minimal expression elements with predictable input-output characteristics. Vogl et al., (2014) used this principle to construct short and purely synthetic core promoters in *Pichia pastoris*. However, the construction of similar modules requires an extensive library of well-characterized motifs - TFBS for promoters and protein structural domains for TFs. Therefore, greater knowledge of promoter and TF architecture will promote the development of more rational design algorithms in future. Many groups have focused on research in this direction. Mehrotra et al., (2013) noted that plant promoters containing ACGT repeat architecture are implicated in stress responses. The regulation by these promoters is sensitive to the length of the spacer between repeats. Efforts towards designing synthetic stress responsive promoters in plants may incorporate this conserved architectural feature to achieve better performance. De Lange et al., (2015) have characterized TALE-like proteins and showed that their repeat architecture may be used to enhance orthogonality and specificity of synthetic TALEs. In future, it may be possible to use the structural and functional knowledge of TALE-like proteins to engineer specific features into TALE based TFs.

Promoter and transcription factor engineering efforts are still largely based on hybridizing independent modules whose functions are well characterized. Whether it is the hybridization of promoters with different activator and operator sequences, the design of chimeric transcription factors composed of specific DNA binding, regulator and inducer responsive domains, or the incorporation of protein binding RNA aptamers into gRNAs – the principle is the same. Therefore, their success depends on the availability of well characterized libraries of modular elements that may be used interchangeably in engineering synthetic modules. However, this 'parts-off a shelf' approach is challenging in relatively new or obscure systems where such libraries are limited. Novel module characterization protocols like that adopted by Blazeck et al. (2013) in *Y. lipolytica* need to be advanced. Further, de-novo design principles may assist in circumventing similar problems by reducing the need for pre-existing modular elements. In plants, the availability of well characterized core promoter modules is limited (Liu and Stewart, 2016). Using rational engineering approaches, it may be possible to create purely synthetic cores by hybridizing various core promoter elements. Such approaches have already been

demonstrated in higher eukaryotes (Juven-Gershon et al., 2006) and fungi (Vogl et al., 2014). Synthetic promoters and TFs designed de-novo are orthologous to the host genomes. Consequently, such systems are more robust than those assembled from native modules as they are resistant to mechanisms like homologous recombination in yeast and HBGS in plants.

Additionally, engineering synthetic promoters and transcriptional factors through such hybridization-based strategies is constrained by challenges in assembling the modules into a complete synthetic unit. Recently developed methods like Golden Gate assembly have made it easier to assemble complete transcriptional units into the cell (Agmon et al., 2015). Mitchell et al., (2015) extended the Golden Gate strategy to assemble four to six gene pathways in yeast using a method called the Versatile Genetic Assembly System, VEGAS. Such powerful tools have been well complemented with computational toolkits (Yang et al., 2016) that help the experimenter to easily design and assemble complete transcriptional units *in-vivo*. It is also possible to carry out this assembly *in-vitro* without repeated rounds of DNA isolation, digestion, ligation and cloning into cells. Lundqvist et al., (2015) could assemble over 150 two to four-part DNA systems using a solid phase cloning strategy. Such systems may be automated to efficiently and easily assemble arrays of synthetic transcriptional units, each of which code for a gene belonging to a complex genetic cascade.

While engineering transcriptional pathways and networks, one needs to ensure that the individual modules show minimal cross-reactivity with each other and the host. sTFs are key to addressing this requirement. The emergence of the ZNF, TALE and CRISPR/Cas9 technologies has simplified the synthesis of artificial TFs. In this review, we briefly discussed these technologies, their specific advantages, applications and current limitations with respect to synthetic engineering. Recent developments to these technologies have assisted in achieving greater specificity and orthogonality in synthetic circuits. It is possible to fine-tune the specificity of ZNFs, TALEs and CRISPR/Cas9 based TFs through rational engineering principles that are devised based on our knowledge of similar proteins. For example, a recent study had identified polymorphisms in TALE-like proteins, which influence their specificity by modifying their interactions with DNA (De Lange et al., 2015). The knowledge of such polymorphisms may be used to rationally design tunable effectors with desired target specificities. The engineering of three and four part ZNFs that show limited off-target activity in the host is another illustration of the power and need for rational engineering approaches (Khalil et al., 2012; McIsaac et al., 2013). Computational tools may aid in increasing the specificity of sTF technologies, as recently illustrated by Didovyk et al., (2016).

The design of sTF proteins is constrained by several limitations in structure and practical synthesis. The parts of sTF must be combined together in specific orders, orientations and spatial positions to perform the intended functions properly and efficiently (Purcell et al., 2014). An important part of rational engineering is establishing principles or basic blue prints for efficient assembly of a chimeric transcriptional factor. Recently, Purcell et al. (2014) developed a grammar for designing sTFs. This grammar library was constructed by taking into consideration a set of basic constraints on the positioning and functioning of the various domains. Based on this initial library, eleven permissible designs were identified for eukaryotic systems. Such



grammar can be continuously updated and improved as more modular components are identified and characterized. Consequently, it may be possible for experimenters to have a standard dictionary of sTF design principles, which can be used as a reference for assembling chimeric TFs that meet their specific requirements. These grammars can be implemented in computer based design aids to enable users to design sTF and organize experimental libraries. To date, only experts can take advantage of the rapid progress in the field of sTF engineering. This is because only a few researchers are familiar with the constraints in the design and assembly of sTFs from the modular components of nuclear localization signal, DBD, effector domain, linker domain, cleavage domain, reporter and protein interaction domain. The availability of such grammar in open source will expedite future work in this field.

Any synthetic designer technology will ideally enable control of multiple targets simultaneously as this is critical for assembling complex transcriptional circuits. This property is called scalability. The synthesis of multiple ZNFs and TALEs is challenging due to the large amounts of time and labor required for traditional protein engineering. Combinatorial technologies have simplified this process by promoting cell-free assembly and screening of proteins. Recently, Blackburn et al., (2016) could synthesize and characterize 400 ZNFs in a short duration by integrating gene synthesis and microfluidic analysis platforms. When augmented by efficient delivery systems like the nanoparticle based technology devised by Patel et al., (2014), similar technologies will expand the application of ZNFs and TALEs in synthetic biology. An easier alternative for controlling multiple nodes of a complex genetic circuit is the CRISPR/Cas9 system due to the relative simplicity in designing and targeting gRNAs to different genomic targets.

In the context on clinical and industrial application, synthetic circuits need to function as inducible switches exhibiting high spatiotemporal resolution. By programming responsiveness to multiple inducers, it is possible to design multi-functional modules that can switch between alternate functional states in response to different inducers. Multimodal regulatory models are necessary (Chen et al., 2016; Farzadfard et al., 2013; Keung et al., 2014; Mazumder and McMillen, 2014) as an increase in the number of synthetic modules engineered into cells leads to a decrease in cell survival due to metabolic overload. Multifunctional synthetic modules will also enable easy switching in regulatory systems that are sensitive to frequent environmental perturbations. Inducible TFs are easily synthesized in the case of ZNFs and TALEs by incorporating an inducer responsive domain. However, there are major challenges to programming expression induction into CRISPR/Cas9 based systems as the gRNA expressed from inducible RNAP II promoters is exported outside the nucleus (Jusiak et al., 2016). In future, innovative gRNA expression and delivery systems will need to be developed to expand the utility of the CRISPR/Cas9 system in synthetic biology (Kiani et al., 2014; Nissim et al., 2014; Shechner et al., 2015). Alternately, we could fuse the dCas9 based transcriptional regulator with an inducible domain. But this would undo the specific advantage of the CRISPR/Cas9 system over its protein-centric counterparts. It may however be possible to make the CRISPR/Cas9 system inducible while retaining its RNA based targeting advantage by post-transcriptionally regulating the activity of the gRNA; this could be done by fusing the gRNA with RNA aptamers, which would function as riboswitches that respond to specific inducers.

The efficiency of a synthetic module is affected by parameters across different levels of genomic organization. De-novo designer technologies that factor in contributions from one or more of these parameters will produce more robust modules. In this review, we have discussed at least two examples that highlight this point (Brewster et al., 2012; Curran et al., 2014). The next generation synthetic networks must leverage the regulatory mechanisms present across different organizational levels to improve the design of regulatory switches. The objective is to design switches that can target the epigenome, genome and the transcriptome, and consequently synthesize a more robust cellular cascade. Recent studies have sought to separately engineer epigenomic (Keung et al., 2014; Ma et al., 2014) and post-transcriptional regulation (Green et al., 2014; Kiani et al., 2014; Nissim et al., 2014), showing initial promise towards scaling this long-term objective. However, the challenge is to integrate these individual regulatory mechanisms at each level to produce a combinatorial regulatory model.

Addressing this challenge would require an interdisciplinary effort to circumvent the current limitations and expand the application of designer promoter and TF technologies. We note that the future of rational and *de-novo* engineering of promoters and TFs depends on the progress made in disciplines like computational biology and bioinformatics, theoretical and systems biology. In this review, we presented some recent studies (Brewster et al., 2012; Curran et al., 2014) that support this fact. De-novo designer technologies are currently handicapped by the vast gaps in knowledge sourced from these fields. For example, the computational models of repressive systems used by Brewster et al. (2012) were insufficient to explain the experimental observations and showed poor correlation. Further, there is a need to expand the repertoire of well-characterized parts that may be used for combinatorial engineering. The use of functional genomics and similar high throughput technologies will assist in characterizing specific modular components in living systems. Studies exploiting similar strategies have identified myriad putative promoter elements in *E. coli* and *S. cerevisiae* (Mendoza-Vargas et al., 2009). If characterized, these may become available for synthetic manipulation. The use of high throughput technologies for improving rational engineering principles has been seen in recent studies by Curran et al., (2014) and Blount et al., (2012). Schlabach et al., (2010) used DNA based microarrays for efficient generation of extensive enhancer libraries that showed cell-line specific activity. Such libraries may be useful in tissue specific transgene expression, making them useful tools in gene therapy. Microarray and ChIP technologies may be used along with probabilistic motif detection algorithms like Gibbs Sampling and MEME to aid in detection of overrepresented *cis*-regulatory motifs (Venter and Warnich, 2009); these will point to specific motifs implicated in disease and may be used in therapeutics. To take it a step further, similar approaches may be used to extend the application of synthetic circuits into personalized therapeutics.

Advances in promoter and sTF engineering will further our current understanding of basic biological architectures. Synthetic ZNF TFs enabled the detection of multiple genes that were co-expressed or repressed upon induction of the native GCN4 networks in *S. cerevisiae* (McIsaac et al., 2013). The study revealed 327 genes that were up-regulated by >2 fold, and 255 genes that were repressed >2 fold upon GCN4 expression. Of these only 116 and 5 genes in the former and latter classes were respectively detected as direct GCN4p targets using ChIP based assays in the

past. In a similar way, synthetic engineering may assist in dissecting complex functional networks *in-vivo*. In another study, a synthetic biology approach to study 223 CRs, helped dissect their roles in transcriptional regulation and the principles of combinatorial control by multiple regulator proteins (Keung et al., 2014). The different roles played by these CRs, when bound upstream and downstream of specific gene modules, was unknown before the study. Hybrid promoter engineering strategies used by Blazeck et al. (2011, 2012, and 2013) indicated that even the strongest promoters in native systems are enhancer-limited. Synthetic engineering approaches have also shown that an increase in *cis*-regulatory motif copy number corresponds to greater transcriptional outputs. The over representation of specific *cis*-regulatory motifs such as the ACGT (Mehrotra et al., 2013) in certain plant promoter architectures may now be explained based on similar principles. The experiments by Blazeck et al. (2011, 2012, and 2013) explain the increased activity from artificial promoter chimeras based on increased TF co-localization around the promoter by the *cis*-regulatory motifs. This points to one of the possible mechanisms that may explain how a TF can easily find its target promoter out of countless possible binding sites dispersed across the genome.

## Conclusions

Synthetic genetic modules are important toolkits with vast scope for application in industrial, biological clinical and diagnostic research. Ultimately, the aim of this approach is to construct transcriptional logic circuits with predictable input-output characteristics enabling efficient applications. Through the course of this review, we have addressed the emerging technologies that have contributed significantly towards this goal. There have been many previous reviews that discuss the scope, application and limitations of one or more of these technologies (Blazeck and Alper, 2013; Choo and Isalan, 2000; Jusiak et al., 2016; Kabadi and Gersbach, 2014; Moore et al., 2014). However, to the extent of our knowledge, there is no comprehensive presentation of all the three strands of combinatorial design: transcription factors, designer promoters and external inducers of gene expression. In this review, we have sought to address this deficiency, promoting both the requirement for and scope of combinatorial design. In aggregate, we have shown that these key technologies are individually handicapped by limitations that may be addressed via combinatorial principles. In conclusion, the six salient features of a highly functional synthetic transcriptional network: tunability, amenability, specificity, orthogonality, scalability and handling simplicity can be individually and collectively enhanced by rational combinatorial engineering. Moving forward, we envisage this approach will facilitate significant advances in basic research, human therapeutics and agriculture.

## 5. Acknowledgements

We would like to thank Prof. PK Gupta (CCS University, Meerut, India) for proof reading the manuscript and providing valuable suggestions. Research in the Loake lab is supported by the BBSRC. RM is thankful to Indian National Science academy and Royal society of Edinburgh for support as bilateral exchange fellow to visit Loake lab. RM and SM are thankful to Department of Biological Sciences, BITS, Pilani for infrastructure support.

## 6. References

- Agmon, N., Mitchell, L.A., Cai, Y., Ikushima, S., Chuang, J., Zheng, A., Choi, W.-J., Martin, J.A., Caravelli, K., Stracquadanio, G., Boeke, J.D., 2015. Yeast Golden Gate (yGG) for the Efficient Assembly of *S. cerevisiae* Transcription Units. *ACS Synth. Biol.* 4, 853–859. doi:10.1021/sb500372z
- Alber, T., Kawasaki, G., 1982. Nucleotide sequence of the triose phosphate isomerase gene of *Saccharomyces cerevisiae*. *J. Mol. Appl. Genet.* 1, 419–34.
- Alper, H., Fischer, C., 2005. Tuning genetic control through promoter engineering. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12678–12683. doi:10.1073/pnas.0504604102
- Andrianantoandro E.a Basu, S.. K.D.K.. W.R.. b c, 2006. Synthetic biology: New engineering rules for an emerging discipline. *Mol. Syst. Biol.* 2, 2006.0028. doi:10.1038/msb4100073
- Ang, J., Bagh, S., Ingalls, B.P., McMillen, D.R., 2010. Considerations for using integral feedback control to construct a perfectly adapting synthetic gene network. *J. Theor. Biol.* 266, 723–738. doi:10.1016/j.jtbi.2010.07.034
- Ang, J., McMillen, D.R., 2013. Physical constraints on biological integral control design for homeostasis and sensory adaptation. *Biophys. J.* 104, 505–15. doi:10.1016/j.bpj.2012.12.015
- Bassel, J., Mortimer, R., 1971. Genetic order of the galactose structural genes in *Saccharomyces cerevisiae*. *J. Bacteriol.* 108, 179–183.
- Bhattacharjee, S., Renganaath, K., Mehrotra, R., Mehrotra, S., 2013. Combinatorial control of gene expression. *Biomed Res. Int.* doi:10.1155/2013/407263
- Bio FAB Group, Baker, D., Church, G., Collins, J., Endy, D., Jacobson, J., Keasling, J., Modrich, P., Smolke, C., Weiss, R., 2006. Engineering life: building a fab for biology. *Sci. Am.* 294, 44–51. doi:10.1038/scientificamerican0606-44
- Bitter, G. a, Egan, K.M., 1984. Expression of heterologous genes in *Saccharomyces*

- cerevisiae* from vectors utilizing the glyceraldehyde-3-phosphate dehydrogenase gene promoter. *Gene* 32, 263–274.
- Blackburn, M.C., Petrova, E., Correia, B.E., Maerkl, S.J., 2016. Integrating gene synthesis and microfluidic protein analysis for rapid protein engineering. *Nucleic Acids Res.* 44, e68. doi:10.1093/nar/gkv1497
- Blazeck, J., Alper, H.S., 2013. Promoter engineering: recent advances in controlling transcription at the most fundamental level. *Biotechnol. J.* 8, 46–58. doi:10.1002/biot.201200120
- Blazeck, J., Garg, R., Reed, B., Alper, H.S., 2012. Controlling promoter strength and regulation in *Saccharomyces cerevisiae* using synthetic hybrid promoters. *Biotechnol. Bioeng.* 109, 2884–2895. doi:10.1002/bit.24552
- Blazeck, J., Liu, L., Redden, H., Alper, H., 2011. Tuning gene expression in *Yarrowia lipolytica* by a hybrid promoter approach. *Appl. Environ. Microbiol.* 77, 7905–7914. doi:10.1128/AEM.05763-11
- Blazeck, J., Reed, B., Garg, R., Gerstner, R., Pan, A., Agarwala, V., Alper, H.S., 2013. Generalizing a hybrid synthetic promoter approach in *Yarrowia lipolytica*. *Appl. Microbiol. Biotechnol.* 97, 3037–3052. doi:10.1007/s00253-012-4421-5
- Blount, B.A., Weenink, T., Vasylechko, S., Ellis, T., 2012. Rational diversification of a promoter providing fine-tuned expression and orthogonal regulation for synthetic biology. *PLoS One* 7. doi:10.1371/journal.pone.0033279
- Boch, J., 2011. TALEs of genome targeting. *Nat. Biotechnol.* 29, 135–6. doi:10.1038/nbt.1767
- Bogdanove, a. J., Voytas, D.F., 2011. TAL Effectors: Customizable Proteins for DNA Targeting. *Science* (80-. ). 333, 1843–1846. doi:10.1126/science.1204094
- Brewster, R.C., Jones, D.L., Phillips, R., 2012. Tuning Promoter Strength through RNA Polymerase Binding Site Design in *Escherichia coli*. *PLoS Comput. Biol.* 8.

doi:10.1371/journal.pcbi.1002811

Brückner, K., Schäfer, P., Weber, E., Grützner, R., Marillonnet, S., Tissier, A., 2015. A library of synthetic transcription activator-like effector-activated promoters for coordinated orthogonal gene expression in plants. *Plant J.* 82, 707–16.

doi:10.1111/tpj.12843

Bultmann, S., Morbitzer, R., Schmidt, C.S., Thanisch, K., Spada, F., Elsaesser, J., Lahaye, T., Leonhardt, H., 2012. Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Res.* 40, 5368–77. doi:10.1093/nar/gks199

Chen, X., Li, T., Wang, X., Du, Z., Liu, R., Yang, Y., 2016. Synthetic dual-input mammalian genetic circuits enable tunable and stringent transcription control by chemical and light. *Nucleic Acids Res.* 44, 2677–90. doi:10.1093/nar/gkv1343

Choo, Y., Isalan, M., 2000. Advances in zinc finger engineering. *Curr. Opin. Struct. Biol.* 10, 411–6.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., Zhang, F., 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–23. doi:10.1126/science.1231143

Curran, K. a, Crook, N.C., Karim, A.S., Gupta, A., Wagman, A.M., Alper, H.S., 2014. Design of synthetic yeast promoters via tuning of nucleosome architecture. *Nat. Commun.* 5, 4002. doi:10.1038/ncomms5002

Da Silva, N.A., Srikrishnan, S., 2012. Introduction and expression of genes for metabolic engineering applications in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 12, 197–214. doi:10.1111/j.1567-1364.2011.00769.x

de Boer, H.A., Comstock, L.J., Vasser, M., 1983. The tac promoter: a functional hybrid derived from the trp and lac promoters. *Proc. Natl. Acad. Sci. U. S. A.* 80, 21–5.

de Lange, O., Wolf, C., Dietze, J., Elsaesser, J., Morbitzer, R., Lahaye, T., 2014.

- Programmable DNA-binding proteins from Burkholderia provide a fresh perspective on the TALE-like repeat domain. *Nucleic Acids Res.* 42, 7436–49.  
doi:10.1093/nar/gku329
- De Lange, O., Wolf, C., Thiel, P., Krüger, J., Kleusch, C., Kohlbacher, O., Lahaye, T., 2015. DNA-binding proteins from marine bacteria expand the known sequence diversity of TALE-like repeats. *Nucleic Acids Res.* 43, 10065–10080.  
doi:10.1093/nar/gkv1053
- Deng, D., Yin, P., Yan, C., Pan, X., Gong, X., Qi, S., Xie, T., Mahfouz, M., Zhu, J.-K., Yan, N., Shi, Y., 2012. Recognition of methylated DNA by TAL effectors. *Cell Res.* 22, 1502–4. doi:10.1038/cr.2012.127
- Denis, C.L., Ferguson, J., Young, E.T., 1983. mRNA levels for the fermentative alcohol dehydrogenase of *Saccharomyces cerevisiae* decrease upon growth on a nonfermentable carbon source. *J. Biol. Chem.* 258, 1165–71.
- Didovyk, A., Borek, B., Hasty, J., Tsimring, L., 2016. Orthogonal Modular Gene Repression in *Escherichia coli* Using Engineered CRISPR/Cas9. *ACS Synth. Biol.* 5, 81–88. doi:10.1021/acssynbio.5b00147
- DOUGLAS, H.C., HAWTHORNE, D.C., 1964. ENZYMATIC EXPRESSION AND GENETIC LINKAGE OF GENES CONTROLLING GALACTOSE UTILIZATION IN *SACCHAROMYCES*. *Genetics* 49, 837–44.
- Ellis, T., Wang, X., Collins, J.J., 2009. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat. Biotechnol.* 27, 465–71.  
doi:10.1038/nbt.1536
- Esvelt, K.M., Mali, P., Braff, J.L., Moosburner, M., Yaung, S.J., Church, G.M., 2013. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods* 10, 1116–21. doi:10.1038/nmeth.2681
- Farzadfard, F., Perli, S.D., Lu, T.K., 2013. Tunable and Multifunctional Eukaryotic

- Transcription Factors Based on CRISPR/Cas. *ACS Synth. Biol.* 2, 604–613.  
doi:10.1021/sb400081r
- Fine, E.J., Cradick, T.J., Zhao, C.L., Lin, Y., Bao, G., 2014. An online bioinformatics tool predicts zinc finger and TALE nuclease off-target cleavage. *Nucleic Acids Res.* 42, e42. doi:10.1093/nar/gkt1326
- Gaj, T., Guo, J., Kato, Y., Sirk, S.J., Barbas, C.F., 2012. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. *Nat. Methods* 9, 805–7.  
doi:10.1038/nmeth.2030
- Garg, A., Lohmueller, J.J., Silver, P.A., Armel, T.Z., 2012. Engineering synthetic TAL effectors with orthogonal target sites. *Nucleic Acids Res.* 40, 7584–7595.  
doi:10.1093/nar/gks404
- Gatignol, A., Dassain, M., Tiraby, G., 1990. Cloning of *Saccharomyces cerevisiae* promoters using a probe vector based on phleomycin resistance. *Gene* 91, 35–41.
- Gibson, D.G., Benders, G.A., Axelrod, K.C., Zaveri, J., Algire, M.A., Moodie, M., Montague, M.G., Venter, J.C., Smith, H.O., Hutchison, C.A., 2008. One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proc. Natl. Acad. Sci. U. S. A.* 105, 20404–9.  
doi:10.1073/pnas.0811011106
- Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., Qi, L.S., Kampmann, M., Weissman, J.S., 2014. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* 159, 647–61. doi:10.1016/j.cell.2014.09.029
- Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., Lim, W.A., Weissman, J.S., Qi, L.S., 2013. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442–51. doi:10.1016/j.cell.2013.06.044



- Green, A.A., Silver, P.A., Collins, J.J., Yin, P., 2014. Toehold switches: De-novo-designed regulators of gene expression. *Cell* 159, 925–939.  
doi:10.1016/j.cell.2014.10.002
- Guarente, L., Hoar, E., 1984. Upstream activation sites of the CYC1 gene of *Saccharomyces cerevisiae* are active when inverted but not when placed downstream of the &quot;TATA box&quot;. *Proc. Natl. Acad. Sci. U. S. A.* 81, 7860–4.
- Guarente, L., Lalonde, B., Gifford, P., Alani, E., 1984. Distinctly regulated tandem upstream activation sites mediate catabolite repression of the CYC1 gene of *S. cerevisiae*. *Cell* 36, 503–511. doi:10.1016/0092-8674(84)90243-5
- Guarente, L., Ptashne, M., 1981. Fusion of *Escherichia coli* lacZ to the cytochrome c gene of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 78, 2199–203.
- Guido, N.J., Wang, X., Adalsteinsson, D., McMillen, D., Hasty, J., Cantor, C.R., Elston, T.C., Collins, J.J., 2006. A bottom-up approach to gene regulation. *Nature* 439, 856–860. doi:10.1038/nature04473
- Hahn, S., Young, E.T., 2011. Transcriptional regulation in *Saccharomyces cerevisiae*: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. *Genetics* 189, 705–36. doi:10.1534/genetics.111.127019
- Hammer, K., Mijakovic, I., Jensen, P.R., 2006. Synthetic promoter libraries--tuning of gene expression. *Trends Biotechnol.* 24, 53–5. doi:10.1016/j.tibtech.2005.12.003
- He, W., Qiang, M., Ma, W., Valente, A.J., Quinones, M.P., Wang, W., Reddick, R.L., Xiao, Q., Ahuja, S.S., Clark, R.A., Freeman, G.L., Li, S., 2006. Development of a synthetic promoter for macrophage gene therapy. *Hum. Gene Ther.* 17, 949–59.  
doi:10.1089/hum.2006.17.949
- Hitzeman, R.A., Hagie, F.E., Levine, H.L., Goeddel, D. V., Ammerer, G., Hall, B.D., 1981. Expression of a human gene for interferon in yeast. *Nature* 293, 717–722.  
doi:10.1038/293717a0

- Holkers, M., Maggio, I., Liu, J., Janssen, J.M., Miselli, F., Mussolino, C., Recchia, A., Cathomen, T., Gonçalves, M.A.F. V, 2013. Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Res.* 41, e63. doi:10.1093/nar/gks1446
- Holland, M.J., Holland, J.P., 1978. Isolation and identification of yeast messenger ribonucleic acids coding for enolase, glyceraldehyde-3-phosphate dehydrogenase, and phosphoglycerate kinase. *Biochemistry* 17, 4900–7.
- Jensen, P.R., Hammer, K., 1998a. Artificial promoters for metabolic optimization. *Biotechnol. Bioeng.* 58, 191–5.
- Jensen, P.R., Hammer, K., 1998b. The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Appl. Environ. Microbiol.* 64, 82–7.
- Jeppsson, M., Johansson, B., Jensen, P.R., Hahn-Hägerdal, B., Gorwa-Grauslund, M.F., 2003. The level of glucose-6-phosphate dehydrogenase activity strongly influences xylose fermentation and inhibitor sensitivity in recombinant *Saccharomyces cerevisiae* strains. *Yeast* 20, 1263–72. doi:10.1002/yea.1043
- Johnston, M., Flick, J.S., Pexton, T., 1994. Multiple mechanisms provide rapid and stringent glucose repression of GAL gene expression in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 14, 3834–41.
- Jusiak, B., Cleto, S., Perez-Piñera, P., Lu, T.K., 2016. Engineering Synthetic Gene Circuits in Living Cells with CRISPR Technology. *Trends Biotechnol.* 34, 535–47. doi:10.1016/j.tibtech.2015.12.014
- Juven-Gershon, T., Cheng, S., Kadonaga, J.T., 2006. Rational design of a super core promoter that enhances gene expression. *Nat. Methods* 3, 917–922. doi:10.1038/nmeth937
- Kabadi, A.M., Gersbach, C.A., 2014. Engineering synthetic TALE and CRISPR/Cas9

transcription factors for regulating gene expression. *Methods* 69, 188–97.

doi:10.1016/j.ymeth.2014.06.014

Kellermann, E., Seeboth, P.G., Hollenberg, C.P., 1986. Analysis of the primary structure and promoter function of a pyruvate decarboxylase gene (PDC1) from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 14, 8963–77.

Keung, A.J., Bashor, C.J., Kiriakov, S., Collins, J.J., Khalil, A.S., 2014. Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. *Cell* 158, 110–20. doi:10.1016/j.cell.2014.04.047

Khalil, A.S., Lu, T.K., Bashor, C.J., Ramirez, C.L., Pyenson, N.C., Joung, J.K., Collins, J.J., Andrianantoandro, E., Basu, S., Karig, D.K., Weiss, R., Bashor, C.J., Horwitz, A.A., Peisajovich, S.G., Lim, W.A., Basu, S., Gerchman, Y., Collins, C.H., Arnold, F.H., Weiss, R., Becskei, A., Serrano, L., Beerli, R.R., Barbas, C.F., Blancafort, P., Chen, E.I., Gonzalez, B., Bergquist, S., Zijlstra, A., Guthy, D., Brachat, A., Brakenhoff, R.H., Quigley, J.P., Erdmann, D., Barbas, C.F., Bogdanove, A.J., Voytas, D.F., Bryant, G.O., Ptashne, M., Cormack, B.P., Bertram, G., Egerton, M., Gow, N.A., Falkow, S., Brown, A.J., Craven, S.E., Brecht, D.S., Ellis, T., Wang, X., Collins, J.J., Elowitz, M., Lim, W.A., Elowitz, M.B., Leibler, S., Elrod-Erickson, M., Rould, M.A., Nekludova, L., Pabo, C.O., Elrod-Erickson, M., Benson, T.E., Pabo, C.O., Fischer, J.A., Giniger, E., Maniatis, T., Ptashne, M., Foley, J.E., Yeh, J.R., Maeder, M.L., Reyon, D., Sander, J.D., Peterson, R.T., Joung, J.K., Friedland, A.E., Lu, T.K., Wang, X., Shi, D., Church, G.M., Collins, J.J., Gardner, T.S., Cantor, C.R., Collins, J.J., Guet, C.C., Elowitz, M.B., Hsing, W., Leibler, S., Hahn, S., Young, E.T., Harris, B.Z., Lim, W.A., Harris, B.Z., Hillier, B.J., Lim, W.A., Joung, J.K., Le, L.U., Hochschild, A., Joung, J.K., Koepp, D.M., Hochschild, A., Khalil, A.S., Collins, J.J., Kramer, B.P., Fischer, C., Fussenegger, M., Lu, T.K., Khalil, A.S., Collins, J.J., Ma, J., Ptashne, M., Ma, J., Przibilla, E., Hu, J., Bogorad, L., Ptashne, M., Maeder, M.L., Thibodeau-Beganny, S., Osiak, A., Wright, D.A., Anthony, R.M., Eichinger, M., Jiang, T., Foley, J.E., Winfrey, R.J., Townsend, J.A., al., et, Maeder,

M.L., Thibodeau-Beganny, S., Sander, J.D., Voytas, D.F., Joung, J.K., Maxon, M.E., Herskowitz, I., Mukherji, S., Oudenaarden, A. van, Murphy, K.F., Balázsi, G., Collins, J.J., Nandagopal, N., Elowitz, M.B., Nissim, L., Bar-Ziv, R.H., Pabo, C.O., Peisach, E., Grant, R.A., Pavletich, N.P., Pabo, C.O., Pedraza, J.M., Oudenaarden, A. van, Pomerantz, J.L., Wolfe, S.A., Pabo, C.O., Ptashne, M., Ptashne, M., Ptashne, M., Gann, A., Ptashne, M., Gann, A., Reyon, D., Tsai, S.Q., Khayter, C., Foden, J.A., Sander, J.D., Joung, J.K., Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S., Elowitz, M.B., Rubin-Bejerano, I., Mandel, S., Robzyk, K., Kassir, Y., Sander, J.D., Dahlborg, E.J., Goodwin, M.J., Cade, L., Zhang, F., Cifuentes, D., Curtin, S.J., Blackburn, J.S., Thibodeau-Beganny, S., Qi, Y., al., et, Sebastiano, V., Maeder, M.L., Angstman, J.F., Haddad, B., Khayter, C., Yeo, D.T., Goodwin, M.J., Hawkins, J.S., Ramirez, C.L., Batista, L.F., al., et, Smolke, C.D., Silver, P.A., Townsend, J.A., Wright, D.A., Winfrey, R.J., Fu, F., Maeder, M.L., Joung, J.K., Voytas, D.F., Wang, B.S., Pabo, C.O., Wang, B.S., Grant, R.A., Pabo, C.O., Weber, W., Fussenegger, M., Webster, N., Jin, J.R., Green, S., Hollis, M., Chambon, P., Wolfe, S.A., Ramm, E.I., Pabo, C.O., Wright, D.A., Thibodeau-Beganny, S., Sander, J.D., Winfrey, R.J., Hirsh, A.S., Eichinger, M., Fu, F., Porteus, M.H., Dobbs, D., Voytas, D.F., Joung, J.K., Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R., Benenson, Y., Zou, J., Maeder, M.L., Mali, P., Pruetz-Miller, S.M., Thibodeau-Beganny, S., Chou, B.K., Chen, G., Ye, Z., Park, I.H., Daley, G.Q., al., et, 2012. A synthetic biology framework for programming eukaryotic transcription functions. *Cell* 150, 647–58. doi:10.1016/j.cell.2012.05.045

Kiani, S., Beal, J., Ebrahimkhani, M.R., Huh, J., Hall, R.N., Xie, Z., Li, Y., Weiss, R., 2014. CRISPR transcriptional repression devices and layered circuits in mammalian cells. *Nat. Methods* 11, 723–6. doi:10.1038/nmeth.2969

Kinney, J.B., Murugan, A., Callan, C.G., Cox, E.C., 2010. Using deep sequencing to characterize the biophysical mechanism of a transcriptional regulatory sequence. *Proc. Natl. Acad. Sci. U. S. A.* 107, 9158–63. doi:10.1073/pnas.1004290107

- Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., Nureki, O., Zhang, F., 2015. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 517, 583–8. doi:10.1038/nature14136
- Kumar, S., AlAbed, D., Whitteck, J.T., Chen, W., Bennett, S., Asberry, A., Wang, X., DeSloover, D., Rangasamy, M., Wright, T.R., Gupta, M., 2015. A combinatorial bidirectional and bicistronic approach for coordinated multi-gene expression in corn. *Plant Mol. Biol.* 87, 341–53. doi:10.1007/s11103-015-0281-6
- Lam, F.H., Steger, D.J., O’Shea, E.K., 2008. Chromatin decouples promoter threshold from dynamic range. *Nature* 453, 246–50. doi:10.1038/nature06867
- Le Bec, C., Douar, A.M., 2006. Gene therapy progress and prospects--vectorology: design and production of expression cassettes in AAV vectors. *Gene Ther.* 13, 805–13. doi:10.1038/sj.gt.3302724
- Leavitt, J.M., Tong, A., Tong, J., Pattie, J., Alper, H.S., 2016. Coordinated transcription factor and promoter engineering to establish strong expression elements in *Saccharomyces cerevisiae*. *Biotechnol. J.* 11, 866–76. doi:10.1002/biot.201600029
- Lee, T.I., Young, R.A., 2000. Transcription of eukaryotic protein-coding genes. *Annu. Rev. Genet.* 34, 77–137. doi:10.1146/annurev.genet.34.1.77
- Lemken, M.-L., Wybranietz, W.-A., Schmidt, U., Graepler, F., Armeanu, S., Bitzer, M., Lauer, U.-M., 2005. Expression liver-directed genes by employing synthetic transcriptional control units. *World J. Gastroenterol.* 11, 5295–302. doi:10.3748/wjg.v11.i34.5295
- Li, Y., Jiang, Y., Chen, H., Liao, W., Li, Z., Weiss, R., Xie, Z., 2015. Modular construction of mammalian gene circuits using TALE transcriptional repressors. *Nat. Chem. Biol.* 11, 207–213. doi:10.1038/nchembio.1736
- Liu, J., Gaj, T., Patterson, J.T., Sirk, S.J., Barbas III, C.F., Gaj, T., Gersbach, C., 3rd,

C.B., Segal, D., Meckler, J., Kanaar, R., Hoeijmakers, J., Gent, D. van, Perez, E., Wang, J., Miller, J., Jouvenot, Y., Kim, K., Provasi, E., Genovese, P., Lombardo, A., Magnani, Z., Liu, P., Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., Hockemeyer, D., Wang, H., Kiani, S., Lai, C., Gao, Q., Mali, P., Yang, L., Esvelt, K., Aach, J., Guell, M., Holt, N., Wang, J., Kim, K., Friedman, G., Wang, X., Li, L., Krymskaya, L., Wang, J., Henley, J., Rao, A., Beumer, K., Bhattacharyya, G., Bibikova, M., Trautman, J., Carroll, D., Wood, A., Lo, T., Zeitler, B., Pickle, C., Ralston, E., Lo, T., Pickle, C., Lin, S., Ralston, E., Gurling, M., Tesson, L., Usal, C., Menoret, S., Leung, E., Niles, B., Geurts, A., Cost, G., Freyvert, Y., Zeitler, B., Miller, J., Wang, H., Yang, H., Shivalila, C., Dawlaty, M., Cheng, A., Doyon, Y., McCammon, J., Miller, J., Faraji, F., Ngo, C., Bedell, V., Wang, Y., Campbell, J., Poshusta, T., Starker, C., Hwang, W., Fu, Y., Reyon, D., Maeder, M., Tsai, S., Meng, X., Noyes, M., Zhu, L., Lawson, N., Wolfe, S., Tan, W., Carlson, D., Lancto, C., Garbe, J., Webster, D., Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Moscou, M., Bogdanove, A., Cermak, T., Doyle, E., Christian, M., Wang, L., Zhang, Y., Holkers, M., Maggio, I., Liu, J., Janssen, J., Miselli, F., Wu, Z., Asokan, A., Samulski, R., Asokan, A., Schaffer, D., Samulski, R., Tendeloo, V. Van, Willems, R., Ponsaerts, P., Lenjou, M., Nijs, G., Gaj, T., Guo, J., Kato, Y., Sirk, S., 3rd, C.B., Chen, Z., Jaafar, L., Agyekum, D., Xiao, H., Wade, M., Ru, R., Yao, Y., Yu, S., Yin, B., Xu, W., Mino, T., Mori, T., Aoyama, Y., Sera, T., Miller, J., Tan, S., Qiao, G., Barlow, K., Wang, J., Reyon, D., Tsai, S., Khayter, C., Foden, J., Sander, J., Guo, J., Gaj, T., 3rd, C.B., Guschin, D., Waite, A., Katibah, G., Miller, J., Holmes, M., Vives, E., Brodin, P., Lebleu, B., Derossi, D., Joliot, A., Chassaing, G., Prochiantz, A., Choi, J., Ahn, M., Chae, W., Jung, Y., Park, J., Pooga, M., Hallbrink, M., Zorko, M., Langel, U., Mak, A., Bradley, P., Cernadas, R., Bogdanove, A., Stoddard, B., Deng, D., Yan, C., Pan, X., Mahfouz, M., Wang, J., Stella, S., Molina, R., Yefimenko, I., Prieto, J., Silva, G., Kumar, P., Ban, H., Kim, S., Wu, H., Pearson, T., Kim, S., Subramanya, S., Peer, D., Shimaoka, M., Shankar, P., Jin, L., Bahn, J., Eum, W., Kwon, H., Jang, S., Mitsui, H., Inozume, T., Kitamura, R., Shibagaki, N.,

- Shimada, S., Cong, L., Ran, F., Cox, D., Lin, S., Barretto, R., 2014. Cell-Penetrating Peptide-Mediated Delivery of TALEN Proteins via Bioconjugation for Genome Engineering. *PLoS One* 9, e85755. doi:10.1371/journal.pone.0085755
- Liu, W., Mazarei, M., Rudis, M.R., Fethe, M.H., Peng, Y., Millwood, R.J., Schoene, G., Burris, J.N., Stewart, C.N., 2013. Bacterial pathogen phytosensing in transgenic tobacco and Arabidopsis plants. *Plant Biotechnol. J.* 11, 43–52. doi:10.1111/pbi.12005
- Liu, W., Mazarei, M., Rudis, M.R., Fethe, M.H., Stewart, C.N., Stewart, C., Stewart, C., Touraev, A., Citovsky, V., Tzfira, T., Rushton, P., Reinstedler, A., Lipka, V., Lippok, B., Somssich, I., Gurr, S., Rushton, P., Venter, M., Botha, F., Mazarei, M., Teplova, I., Hajimorad, M., Stewart, C., Alieva, N., Konzen, A., Field, S., Meleshkevitch, E., Hunt, M., BeltranRamirez, V., Miller, D., Wiedenmann, J., Salih, A., Matz, M., Bhaskar, P., Venkateshwaran, M., Wu, L., Ane, J., Jiang, J., Figueiredo, J., Romer, P., Lahaye, T., Graham, J., White, F., Jones, J., McIntosh, K., Hulm, J., Young, L., Bonham-Smith, P., Santos-Rosa, M., Poutaraud, A., Merdinoglu, D., Mestre, P., Sheludko, Y., Sindarovska, Y., Gerasymenko, I., Bennikova, M., Kuchuk, N., Sparkes, I., Runions, J., Kearns, A., Hawes, C., Hoorn, R. Van der, Laurent, F., Roth, R., Wit, P. De, VanderGheynest, J., Guo, H., Simmons, C., Wroblewski, T., Tomczak, A., Michelmores, R., Yang, Y., Li, R., Qi, M., Yasmin, A., Debener, T., Zottini, M., Barizza, E., Costa, A., Formentin, E., Ruberti, C., Carimi, F., Schiavo, F., Cazzonelli, C., Velten, J., Zahur, M., Maqbool, A., Irfan, M., Barozai, M., Qaiser, U., Rashid, B., Husnain, T., Riazuddin, S., Liu, S., Bao, Y., Raventos, D., Jensen, A., Rask, M., Casacuberta, J., Mundy, J., Segundo, B.S., Collmer, A., Badel, J., Charkowski, A., Deng, W., Fouts, D., Ramos, A., Rehm, A., Anderson, D., Schneewind, O., Dijk, K. van, Alfano, J., Hermes, S., Seehaus, K., Koehle, H., Conrath, U., Huang, W., Huang, L., Preston, G., Naylor, M., Carr, J., Li, Y., Singer, A., Whiteley, A., Wang, H., Shan, L., Thara, V., Martin, G., Zhou, J., Tang, X., Durbin, R., Jung, H., Lim, C., Hwang, B., Palanichelvam, K.,

- Cole, A., Shababi, M., Schoelz, J., Janssen, B., Gardner, R., Leckie, B., Stewart, C., Merle, C., Perret, S., Lacour, T., Jonval, V., Hudaverdian, S., Garrone, R., Ruggiero, F., Theisen, M., Voinnett, O., Rivas, S., Mestre, P., Baulcombe, D., Hajdukiewicz, P., Svab, Z., Maliga, P., Millwood, J.R., Halfhill, D.M., Harkins, D., Rusotti, R., Stewart, C., 2011. Rapid in vivo analysis of synthetic promoters for plant pathogen phytosensing. *BMC Biotechnol.* 11, 108. doi:10.1186/1472-6750-11-108
- Liu, W., Stewart, C.N., 2016. Plant synthetic promoters and transcription factors. *Curr. Opin. Biotechnol.* 37, 36–44. doi:10.1016/j.copbio.2015.10.001
- Lu, T.K., Khalil, A.S., Collins, J.J., 2009. Next-generation synthetic gene networks. *Nat. Biotechnol.* 27, 1139–50. doi:10.1038/nbt.1591
- Lundqvist, M., Edfors, F., Sivertsson, Å., Hallström, B.M., Hudson, E.P., Tegel, H., Holmberg, A., Uhlén, M., Rockberg, J., 2015. Solid-phase cloning for high-throughput assembly of single and multiple DNA parts. *Nucleic Acids Res.* 43, e49. doi:10.1093/nar/gkv036
- Ma, A.-N., Wang, H., Guo, R., Wang, Y.-X., Li, W., Cui, J., Wang, G., Hoffman, A.R., Hu, J.-F., 2014. Targeted gene suppression by inducing de novo DNA methylation in the gene promoter. *Epigenetics Chromatin* 7, 20. doi:10.1186/1756-8935-7-20
- Madzak, C., Tréton, B., Blanchin-Roland, S., 2000. Strong hybrid promoters and integrative expression/secretion vectors for quasi-constitutive expression of heterologous proteins in the yeast *Yarrowia lipolytica*. *J. Mol. Microbiol. Biotechnol.* 2, 207–16.
- Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., Joung, J.K., 2013. CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods* 10, 977–9. doi:10.1038/nmeth.2598
- Maeder, M.L., Thibodeau-Beganny, S., Osiak, A., Wright, D.A., Anthony, R.M., Eichinger, M., Jiang, T., Foley, J.E., Winfrey, R.J., Townsend, J.A., Unger-Wallace,



- E., Sander, J.D., Müller-Lerch, F., Fu, F., Pearlberg, J., Göbel, C., Dassie, J.P., Pruetz-Miller, S.M., Porteus, M.H., Sgroi, D.C., Iafrate, A.J., Dobbs, D., McCray, P.B., Cathomen, T., Voytas, D.F., Joung, J.K., 2008. Rapid &quot;open-source&quot; engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol. Cell* 31, 294–301. doi:10.1016/j.molcel.2008.06.016
- Mazumder, M., McMillen, D.R., 2014. Design and characterization of a dual-mode promoter with activation and repression capability for tuning gene expression in yeast. *Nucleic Acids Res.* 42, 9514–9522. doi:10.1093/nar/gku651
- McAlister, L., Holland, M.J., 1985. Differential expression of the three yeast glyceraldehyde-3-phosphate dehydrogenase genes. *J. Biol. Chem.* 260, 15019–27.
- McIsaac, R.S., Oakes, B.L., Wang, X., Dummit, K.A., Botstein, D., Noyes, M.B., 2013. Synthetic gene expression perturbation systems with rapid, tunable, single-gene specificity in yeast. *Nucleic Acids Res.* 41, e57. doi:10.1093/nar/gks1313
- McIsaac, R.S., Silverman, S.J., McClean, M.N., Gibney, P.A., Macinskas, J., Hickman, M.J., Petti, A.A., Botstein, D., 2011. Fast-acting and nearly gratuitous induction of gene expression and protein depletion in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 22, 4447–59. doi:10.1091/mbc.E11-05-0466
- Mehrotra, R., Gupta, G., Sethi, R., Bhalothia, P., Kumar, N., Mehrotra, S., 2011. Designer promoter: an artwork of cis engineering. *Plant Mol. Biol.* 75, 527–536. doi:10.1007/s11103-011-9755-3
- Mehrotra, R., Sethi, S., Zutshi, I., Bhalothia, P., Mehrotra, S., 2013. Patterns and evolution of ACGT repeat cis-element landscape across four plant genomes. *BMC Genomics* 14, 203. doi:10.1186/1471-2164-14-203
- Mendoza-Vargas, A., Olvera, L., Olvera, M., Grande, R., Vega-Alvarado, L., Taboada, B., Jimenez-Jacinto, V., Salgado, H., Juárez, K., Contreras-Moreira, B., Huerta, A.M., Collado-Vides, J., Morett, E., 2009. Genome-wide identification of

- transcription start sites, promoters and transcription factor binding sites in *E. coli*. PLoS One 4, e7526. doi:10.1371/journal.pone.0007526
- Mercer, A.C., Gaj, T., Sirk, S.J., Lamb, B.M., Barbas, C.F., 2014. Regulation of endogenous human gene expression by ligand-inducible TALE transcription factors. ACS Synth. Biol. 3, 723–730. doi:10.1021/sb400114p
- Meyer, P., Saedler, H., 1996. HOMOLOGY-DEPENDENT GENE SILENCING IN PLANTS. Annu. Rev. Plant Physiol. Plant Mol. Biol. 47, 23–48. doi:10.1146/annurev.arplant.47.1.23
- Mitchell, L.A., Chuang, J., Agmon, N., Khunsriraksakul, C., Phillips, N.A., Cai, Y., Truong, D.M., Veerakumar, A., Wang, Y., Mayorga, M., Blomquist, P., Sadda, P., Trueheart, J., Boeke, J.D., 2015. Versatile genetic assembly system (VEGAS) to assemble pathways for expression in *S. cerevisiae*. Nucleic Acids Res. 43, 6620–30. doi:10.1093/nar/gkv466
- Mnaimneh, S., Davierwala, A.P., Haynes, J., Moffat, J., Peng, W.-T., Zhang, W., Yang, X., Pootoolal, J., Chua, G., Lopez, A., Trochesset, M., Morse, D., Krogan, N.J., Hiley, S.L., Li, Z., Morris, Q., Grigull, J., Mitsakakis, N., Roberts, C.J., Greenblatt, J.F., Boone, C., Kaiser, C.A., Andrews, B.J., Hughes, T.R., 2004. Exploration of essential gene functions via titratable promoter alleles. Cell 118, 31–44. doi:10.1016/j.cell.2004.06.013
- Mogno, I., Vallania, F., Mitra, R.D., Cohen, B.A., 2010. TATA is a modular component of synthetic promoters. Genome Res. 20, 1391–7. doi:10.1101/gr.106732.110
- Molina, C., Grotewold, E., Smale, S., Smale, S., Kadonaga, J., Burley, S., Roeder, R., Lagrange, T., Kapanidis, A., Tang, H., Reinberg, D., Ebright, R., Tsai, F., Sigler, P., Davuluri, R., Zhang, M., Shahmuradov, I., Gammerman, A., Hancock, J., Bramley, P., Solovyev, V., Seki, M., Narusaka, M., Yamaguchi-Shinozaki, K., Carninci, P., Kawai, J., Hayashizaki, Y., Shinozaki, K., Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T., Nakajima, M., Enju, A., Akiyama, K., Oono, Y.,

- Haas, B., Delcher, A., Mount, S., Wortman, J., Smith, R., Hannick, L., Maiti, R., Ronning, C., Rusch, D., Town, C., Rapp, P., Carlson, J., Michael, T., McClung, C., Gross, R., Janaki, C., Joshi, R., Sinden, R., Potaman, V., Oussatcheva, E., Pearson, C., Lyubchenko, Y., Shlyakhtenko, L., Kozak, M., Fujimori, S., Washio, T., Higo, K., Ohtomo, Y., Murakami, K., Matsubara, K., Kawai, J., Carninci, P., Hayashizaki, Y., Kikuchi, S., Richards, E., Ausubel, F., Bilaud, T., Koering, C., Binet-Brasselet, null, Ancelin, K., Pollice, A., Gasser, S., Gilson, E., Manevski, A., Bertoni, G., Bardet, C., Tremousaygue, D., Lescure, B., Tremousaygue, D., Manevski, A., Bardet, C., Lescure, N., Lescure, B., Tremousaygue, D., Garnier, L., Bardet, C., Dabos, P., Herve, C., Lescure, B., Ohler, U., Liao, G., Niemann, H., Rubin, G., Burke, T., Kadonaga, J., Burke, T., Willy, P., Kutach, A., Butler, J., Kadonaga, J., Mantovani, R., Nakamura, M., Tsunoda, T., Obokata, J., Achard, P., Lagrange, T., El-Zanaty, A.-F., Mache, R., Butler, J., Kadonaga, J., Ohler, U., Niemann, H., Bailey, T., Elkan, C., Roth, F., Hughes, J., Estep, P., Church, G., Crooks, G., Hon, G., Chandonia, J., Brenner, S., Thijs, G., Lescot, M., Marchal, K., Rombauts, S., Moor, B. De, Rouze, P., Moreau, Y., 2005. Genome wide analysis of Arabidopsis core promoters. *BMC Genomics* 6, 25. doi:10.1186/1471-2164-6-25
- Moore, R., Chandrabhas, A., Bleris, L., 2014. Transcription Activator-like Effectors: A Toolkit for Synthetic Biology. *ACS Synth. Biol.* 3, 708–716. doi:10.1021/sb400137b
- Murphy, K.F., Balázs, G., Collins, J.J., 2007. Combinatorial promoter design for engineering noisy gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 104, 12726–31. doi:10.1073/pnas.0608451104
- Nandagopal, N., Elowitz, M.B., 2011. Synthetic biology: integrated gene circuits. *Science* 333, 1244–8. doi:10.1126/science.1207084
- Nevoigt, E., Fischer, C., Mucha, O., Matthäus, F., Stahl, U., Stephanopoulos, G., 2007. Engineering promoter regulation. *Biotechnol. Bioeng.* 96, 550–8. doi:10.1002/bit.21129

- Nihongaki, Y., Kawano, F., Nakajima, T., Sato, M., 2015. Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nat. Biotechnol.* 33, 755–760.  
doi:10.1038/nbt.3245
- Nishizawa, M., Araki, R., Teranishi, Y., 1989. Identification of an upstream activating sequence and an upstream repressible sequence of the pyruvate kinase gene of the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 9, 442–51.
- Nissim, L., Perli, S.D., Fridkin, A., Perez-Pinera, P., Lu, T.K., 2014. Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells. *Mol. Cell* 54, 698–710.  
doi:10.1016/j.molcel.2014.04.022
- Oakes, B.L., Nadler, D.C., Flamholz, A., Fellmann, C., Staahl, B.T., Doudna, J.A., Savage, D.F., 2016. Profiling of engineering hotspots identifies an allosteric CRISPR-Cas9 switch. *Nat. Biotechnol.* 34, 646–51. doi:10.1038/nbt.3528
- Ogden, J.E., Stanway, C., Kim, S., Mellor, J., Kingsman, A.J., Kingsman, S.M., 1986. Efficient expression of the *Saccharomyces cerevisiae* PGK gene depends on an upstream activation sequence but does not require TATA sequences. *Mol. Cell. Biol.* 6, 4335–43. doi:10.1128/MCB.6.12.4335
- Osborn, M.J., Starker, C.G., McElroy, A.N., Webber, B.R., Riddle, M.J., Xia, L., DeFeo, A.P., Gabriel, R., Schmidt, M., von Kalle, C., Carlson, D.F., Maeder, M.L., Joung, J.K., Wagner, J.E., Voytas, D.F., Blazar, B.R., Tolar, J., 2013. TALEN-based gene correction for epidermolysis bullosa. *Mol. Ther.* 21, 1151–9.  
doi:10.1038/mt.2013.56
- Pabo, C.O., Peisach, E., Grant, R.A., 2001. Design and selection of novel Cys2His2 zinc finger proteins. *Annu. Rev. Biochem.* 70, 313–40.  
doi:10.1146/annurev.biochem.70.1.313
- Patel, S., Jung, D., Yin, P.T., Carlton, P., Yamamoto, M., Bando, T., 2014. NanoScript :

- A Nanoparticle-Based Artificial Transcription Factor for. *ACS Nano* 8, 8959–8967. doi:10.1021/nm501589f
- Patwardhan, R.P., Lee, C., Litvin, O., Young, D.L., Pe'er, D., Shendure, J., 2009. High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. *Nat. Biotechnol.* 27, 1173–5. doi:10.1038/nbt.1589
- Pedraza, J.M., van Oudenaarden, A., 2005. Noise propagation in gene networks. *Science* 307, 1965–9. doi:10.1126/science.1109090
- Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., Guilak, F., Crawford, G.E., Reddy, T.E., Gersbach, C.A., 2013. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* 10, 973–6. doi:10.1038/nmeth.2600
- Polstein, L.R., Gersbach, C.A., 2015. A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.* 11, 198–200. doi:10.1038/nchembio.1753
- Polstein, L.R., Gersbach, C.A., 2012. Light-Inducible Spatiotemporal Control of Gene Activation by Customizable Zinc Finger Transcription Factors. *J. Am. Chem. Soc.* 134, 16480–16483. doi:10.1021/ja3065667
- Purcell, O., Peccoud, J., Lu, T.K., 2014. Rule-based design of synthetic transcription factors in eukaryotes. *ACS Synth. Biol.* 3, 737–744. doi:10.1021/sb400134k
- Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., Lim, W.A., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173–83. doi:10.1016/j.cell.2013.02.022
- Rajkumar, A.S., Liu, G., Bergenholm, D., Arsovska, D., Kristensen, M., Nielsen, J., Jensen, M.K., Keasling, J.D., 2016. Engineering of synthetic, stress-responsive yeast promoters. *Nucleic Acids Res.* 44, e136. doi:10.1093/nar/gkw553
- Rao, C. V., 2012. Expanding the synthetic biology toolbox: engineering orthogonal

regulators of gene expression. *Curr. Opin. Biotechnol.* 23, 689–94.

doi:10.1016/j.copbio.2011.12.015

Redden, H., Alper, H.S., Keaveney, M., Struhl, K., Tirosh, I., Barkai, N., Verstrepen, K., Teo, W.S., Chang, M.W., Liang, J., Ning, J.C., Zhao, H., Afonso, B., Silver, P.A., Ajo-Franklin, C.M., Lubliner, S., Keren, L., Segal, E., Smanski, M.J., Singh, V., Redden, H., Morse, N., Alper, H.S., Zhang, G., Darst, S.A., Cramer, P., Bushnell, D.A., Kornberg, R.D., Alper, H., Fischer, C., Nevoigt, E., Stephanopoulos, G., Du, J., Yuan, Y., Si, T., Lian, J., Zhao, H., Nevoigt, E., Blount, B.A., Weenink, T., Vasylechko, S., Ellis, T., Jeppsson, M., Johansson, B., Jensen, P.R., Hahn-Hägerdal, B., Gorwa-Grauslund, M.F., Ligr, M., Siddharthan, R., Cross, F.R., Siggia, E.D., Iyer, V., Struhl, K., Raveh-Sadka, T., Curran, K.A., Sharon, E., Blazeck, J., Garg, R., Reed, B., Alper, H.S., Khalil, A.S., Basehoar, A.D., Zanton, S.J., Pugh, B.F., Zhang, Z., Dietrich, F.S., Struhl, K., Hahn, S., Young, E.T., Hahn, S., Leuther, K.K., Bushnell, D.A., Kornberg, R.D., Carninci, P., Rosenkrantz, M., Kell, C.S., Pennell, E.A., Webster, M., Devenish, L.J., Curran, K.A., Karim, A.S., Gupta, A., Alper, H.S., Slyke, C. Van, Grayhack, E.J., Teixeira, M.C., Stewart, A.J., Hannehalli, S., Plotkin, J.B., Harbison, C.T., Erb, I., Nimwegen, E. van, Mohibullah, N., Hahn, S., Bhaumik, S.R., Green, M.R., Dudley, A.M., Rougeulle, C., Winston, F., Xi, L., Lee, W., Lee, W., Ganapathi, M., Badis, G., Leavitt, J., Alper, H., Curran, K., Gietz, R.D., Schiestl, R.H., Lee, S.M., Jellison, T., Alper, H.S., 2015. The development and characterization of synthetic minimal yeast promoters. *Nat. Commun.* 6, 7810. doi:10.1038/ncomms8810

Reyon, D., Tsai, S.Q., Khayter, C., Foden, J.A., Sander, J.D., Joung, J.K., 2012. FLASH assembly of TALENs for high-throughput genome editing. *Nat. Biotechnol.* 30, 460–5. doi:10.1038/nbt.2170

Rogers, J.K., Guzman, C.D., Taylor, N.D., Raman, S., Anderson, K., Church, G.M., 2015. Synthetic biosensors for precise gene control and real-time monitoring of metabolites. *Nucleic Acids Res.* 43, 7648–7660. doi:10.1093/nar/gkv616

- Rosenberg, S., Tekamp-olson, P., 1992. Enhanced yeast transcription employing hybrid GAPDH promoter region constructs.
- Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S., Elowitz, M.B., 2005. Gene regulation at the single-cell level. *Science* 307, 1962–5. doi:10.1126/science.1106914
- Sander, J.D., Dahlborg, E.J., Goodwin, M.J., Cade, L., Zhang, F., Cifuentes, D., Curtin, S.J., Blackburn, J.S., Thibodeau-Beganny, S., Qi, Y., Pierick, C.J., Hoffman, E., Maeder, M.L., Khayter, C., Reyon, D., Dobbs, D., Langenau, D.M., Stupar, R.M., Giraldez, A.J., Voytas, D.F., Peterson, R.T., Yeh, J.-R.J., Joung, J.K., 2011. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat. Methods* 8, 67–9. doi:10.1038/nmeth.1542
- Sanjana, N.E., Cong, L., Zhou, Y., Cunniff, M.M., Feng, G., Zhang, F., 2012. A transcription activator-like effector toolbox for genome engineering. *Nat. Protoc.* 7, 171–92. doi:10.1038/nprot.2011.431
- Schlabach, M.R., Hu, J.K., Li, M., Elledge, S.J., 2010. Synthetic design of strong promoters. *Proc. Natl. Acad. Sci. U. S. A.* 107, 2538–43. doi:10.1073/pnas.0914803107
- Sengstag, C., Hinnen, A., 1988. A 28-bp segment of the *Saccharomyces cerevisiae* PHO5 upstream activator sequence confers phosphate control to the CYC1-lacZ gene fusion. *Gene* 67, 223–8.
- Sharon, E., Kalma, Y., Sharp, A., Raveh-Sadka, T., Levo, M., Zeevi, D., Keren, L., Yakhini, Z., Weinberger, A., Segal, E., 2012. Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat. Biotechnol.* 30, 521–30. doi:10.1038/nbt.2205
- Shechner, D.M., Hacısuleyman, E., Younger, S.T., Rinn, J.L., 2015. Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. *Nat. Methods* 12, 664–670. doi:10.1038/nmeth.3433

- Siegele, D.A., Hu, J.C., 1997. Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. *Proc. Natl. Acad. Sci. U. S. A.* 94, 8168–72.
- Siegl, T., Tokovenko, B., Myronovskiy, M., Luzhetskyy, A., 2013. Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes. *Metab. Eng.* 19, 98–106. doi:10.1016/j.ymben.2013.07.006
- Smolke, C.D., Silver, P.A., 2011. Informing biological design by integration of systems and synthetic biology. *Cell* 144, 855–9. doi:10.1016/j.cell.2011.02.020
- Solem, C., Jensen, P.R., 2002. Modulation of gene expression made easy. *Appl. Environ. Microbiol.* 68, 2397–403. doi:10.1128/AEM.68.5.2397-2403.2002
- Stella, S., Molina, R., Bertonatti, C., Juillerrat, A., Montoya, G., 2014. Expression, purification, crystallization and preliminary X-ray diffraction analysis of the novel modular DNA-binding protein BurrH in its apo form and in complex with its target DNA. *Acta Crystallogr. Sect. F, Struct. Biol. Commun.* 70, 87–91. doi:10.1107/S2053230X13033037
- Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., Vale, R.D., 2014. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* 159, 635–46. doi:10.1016/j.cell.2014.09.039
- Taylor, N.D., Garruss, A.S., Moretti, R., Chan, S., Arbing, M., Cascio, D., Rogers, J.K., Isaacs, F.J., Kosuri, S., Baker, D., Fields, S., Church, G.M., Raman, S., 2016. Engineering an allosteric transcription factor to respond to new ligands. *Nat. Methods* In press, 1–11. doi:10.1038/nmeth.3696
- Upadhyay, S.K., Kumar, J., Alok, A., Tuli, R., 2013. RNA-guided genome editing for target gene mutations in wheat. *G3 (Bethesda)*. 3, 2233–8. doi:10.1534/g3.113.008847
- Valton, J., Dupuy, A., Daboussi, F., Thomas, S., Maréchal, A., Macmaster, R., Melliand,



- K., Juillerat, A., Duchateau, P., 2012. Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. *J. Biol. Chem.* 287, 38427–32. doi:10.1074/jbc.C112.408864
- Venter, M., 2007. Synthetic promoters: genetic control through cis engineering. *Trends Plant Sci.* 12, 118–24. doi:10.1016/j.tplants.2007.01.002
- Venter, M., Warnich, L., 2009. In silico promoters: modelling of cis-regulatory context facilitates target prediction. *J. Cell. Mol. Med.* 13, 270–8. doi:10.1111/j.1582-4934.2008.00371.x
- Vilar, J.M.G., 2006. Modularizing gene regulation. *Mol. Syst. Biol.* 2, 2006.0016. doi:10.1038/msb4100058
- Vogl, T., Ruth, C., Pitzer, J., Kickenweiz, T., Glieder, A., 2014. Synthetic Core Promoters for *Pichia pastoris*. *ACS Synth. Biol.* 3, 188–191. doi:10.1021/sb400091p
- Wang, B.S., Grant, R.A., Pabo, C.O., 2001. Selected peptide extension contacts hydrophobic patch on neighboring zinc finger and mediates dimerization on DNA. *Nat. Struct. Biol.* 8, 589–93. doi:10.1038/89617
- Weber, W., Fussenegger, M., 2009. Engineering of synthetic mammalian gene networks. *Chem. Biol.* 16, 287–97. doi:10.1016/j.chembiol.2009.02.005
- Wolfe, S.A., Ramm, E.I., Pabo, C.O., 2000. Combining structure-based design with phage display to create new Cys(2)His(2) zinc finger dimers. *Structure* 8, 739–50.
- Xi, L., Fondufe-Mittendorf, Y., Xia, L., Flatow, J., Widom, J., Wang, J.-P., 2010. Predicting nucleosome positioning using a duration Hidden Markov Model. *BMC Bioinformatics* 11, 346. doi:10.1186/1471-2105-11-346
- Yang, K., Stracquadanio, G., Luo, J., Boeke, J.D., Bader, J.S., 2016. BioPartsBuilder: a synthetic biology tool for combinatorial assembly of biological parts. *Bioinformatics* 32, 937–9. doi:10.1093/bioinformatics/btv664

Zalatan, J.G., Lee, M.E., Almeida, R., Gilbert, L.A., Whitehead, E.H., La Russa, M., Tsai, J.C., Weissman, J.S., Dueber, J.E., Qi, L.S., Lim, W.A., 2015. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* 160, 339–50. doi:10.1016/j.cell.2014.11.052

ACCEPTED MANUSCRIPT

Sl. No.	Figure Caption	Figure Location
1	<b>Graphical Abstract</b>	<b>After Abstract</b>
2	<p><b>Figure 1. Random Mutagenesis Based Strategies</b> – (a) Library generation using randomized oligonucleotide primers. The randomized sequences are indicated as ‘NNN’, with each member of the oligonucleotide primer identified by a unique color. Two sets of primers are used amplify the sequence upstream and downstream of the target sequence (black and red respectively). The two fragments are then put together and amplified in a single PCR reaction to generate the full length promoter mutated at the location of choice. (b) Library Generation using Mutagenic PCR – The template is amplified by Error prone PCR to generate random mutants modified at various positions within the sequence. The number of mutations may be controlled by altering the PCR reaction composition and conditions. Additionally, the mutations may also be selected for specific nucleotide replacements by altering the ratio of dNTPs in the initial reaction mix. The primers for amplification of template are indicated in red and each mutation is labelled as ‘N’ and identified by a unique color.</p>	<b>Page 4</b>
3	<p><b>Figure 2. Promoter engineering via Hybrid Promoter Strategies</b> – The promoter hybrid is constructed by a parts-off a shelf approach. There are 3 main components that may be exploited for the same – Upstream Regulatory elements (URE), Core promoter, and the Operator (Op) sequence. Their relative positions with respect to the protein coding sequence is indicated in the top panel. These may be derived from different sources according to the requirements of the experiment (Indicated as Source A-C). The promoter hybrids always contain the core promoter and vary in the presence of the other two elements. Mono-hybrids (A+B, and A+C) are constructed by fusing the core promoter with either of the URE or Op sequences. They are single function modules and may act to either increase or decrease core promoter activity when compared to wild type levels. Additionally, bi-functional hybrids may be constructed by combining all 3 components (A+B+C). In these modules the regulatory properties of the URE and Op sequences are opposite.</p>	<b>Page 5</b>
4	<p><b>Figure 3. Transcriptional regulation by artificial ZNF TFs</b> – (A) Transcriptional activation using artificial TFs composed of a ZNF DBD and a trans-activator domain. The trans-activator domain (shown in yellow) promotes gene transcription by stabilizing the binding of RNAP II basal transcriptional machinery (shown in blue) to the core promoter (shown in orange). (B) Transcriptional repression using artificial TFs containing a ZNF DBD and a repressor domain. The repressor domain (shown in brown) destabilizes the binding of RNAP II basal transcription machinery to the core promoter thereby antagonizing gene transcription.</p> <p>Here, the ZNF DBD is a three finger array (indicated in the top panel) with each finger (shown as rectangular blocks) binding to a trinucleotide sequence (indicated in capital letters). The trinucleotide sequences are specific to the fingers used in the construction of the DBD. The spacer sequence between the cis-regulatory motif bound by the ZNF DBD and the core promoter is variable (indicated by ‘n’).</p>	<b>Page 10</b>
5	<b>Figure 4. Transcriptional regulation by TALE based TFs</b> – (A) Transcriptional activation using TFs composed of a TALE DBD	<b>Page 11</b>

	<p>(shown in green) and a trans-activator domain. The trans-activator domain (shown in yellow) promotes gene transcription by stabilizing the binding of RNAP II basal transcriptional machinery (shown in blue) to the core promoter (shown in orange). (B) Transcriptional repression using TFs containing a TALE DBD and a repressor domain. The repressor domain (shown in brown) destabilizes the binding of RNAP II basal transcription machinery to the core promoter thereby antagonizing gene transcription.</p> <p>The TALE DBD is composed of repeat domains that vary at positions called repeat variable di-residues (RVD). Each nucleotide is bound by a unique RVD repeat (shown in top panel). Here, each RVD repeat is indicated by a different shade of green. The base specifying residues of the TALE protein vary depending on the source of the protein. The residues indicated here are characteristic of TALEs isolated from <i>Xanthomonas sp.</i> The spacer sequence between the cis-regulatory motif bound by the DBD and the core promoter is variable (indicated by 'n').</p>	
6	<p><b>Figure 5. Transcriptional regulation by CRISPR/Cas9 based TFs in Eukaryotes</b> - (A) Transcriptional repression using CRISPR/Cas9 system. The sgRNA targets the dCas9/repressor fusion protein upstream to the transcriptional start site. Repressor domains like KRAB (shown in yellow patterned with brown stripes) recruit repressive chromatin modifiers that silence the promoter and suppress transcription of the target gene. (B) Transcriptional activation using CRISPR/Cas9 system. The dCas9 protein fused to a transactivator (shown in yellow) is guided upstream of the target gene by the sgRNA. The transactivators like VP64 upregulate gene expression by stabilizing the basal transcriptional machinery. All the components of the CRISPR/Cas9 system are shown in the top panel.</p>	Page 12

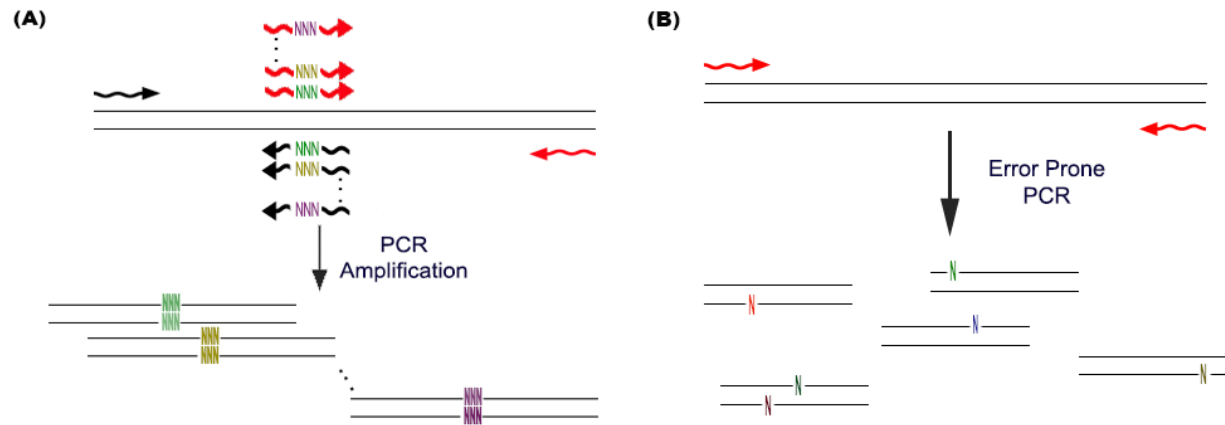


Fig. 1

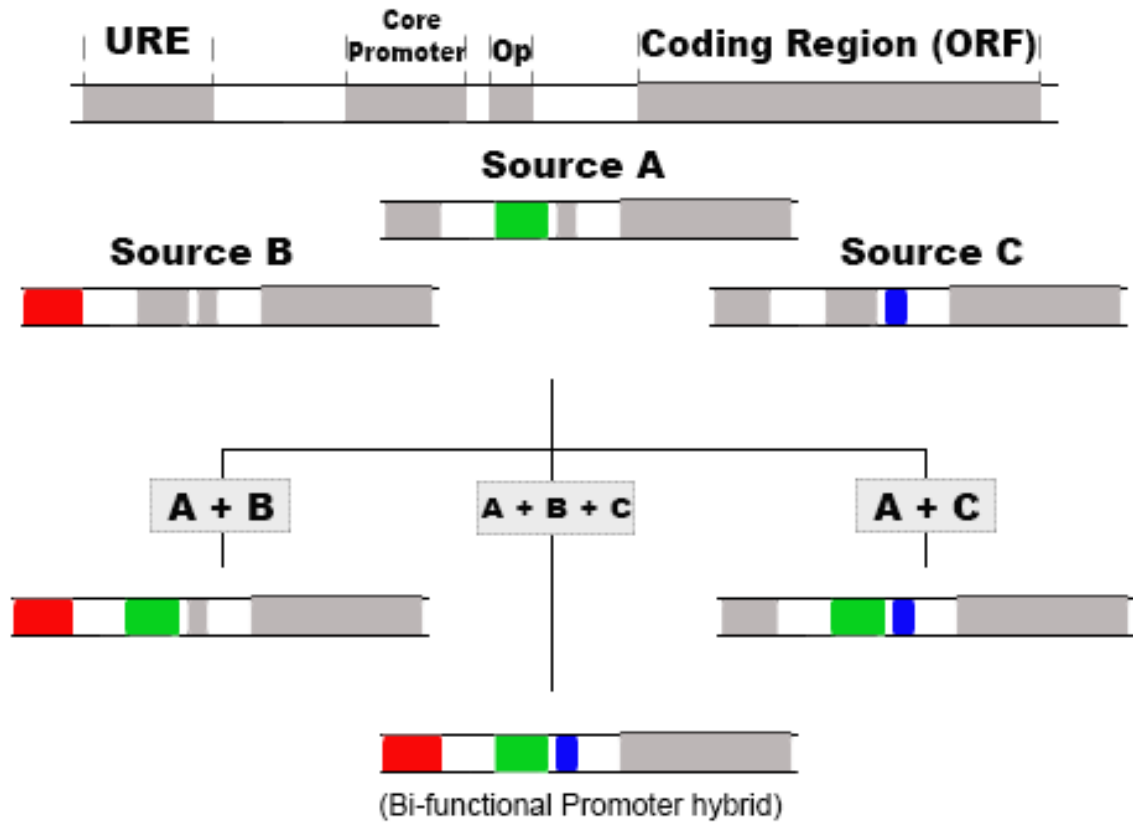


Fig. 2

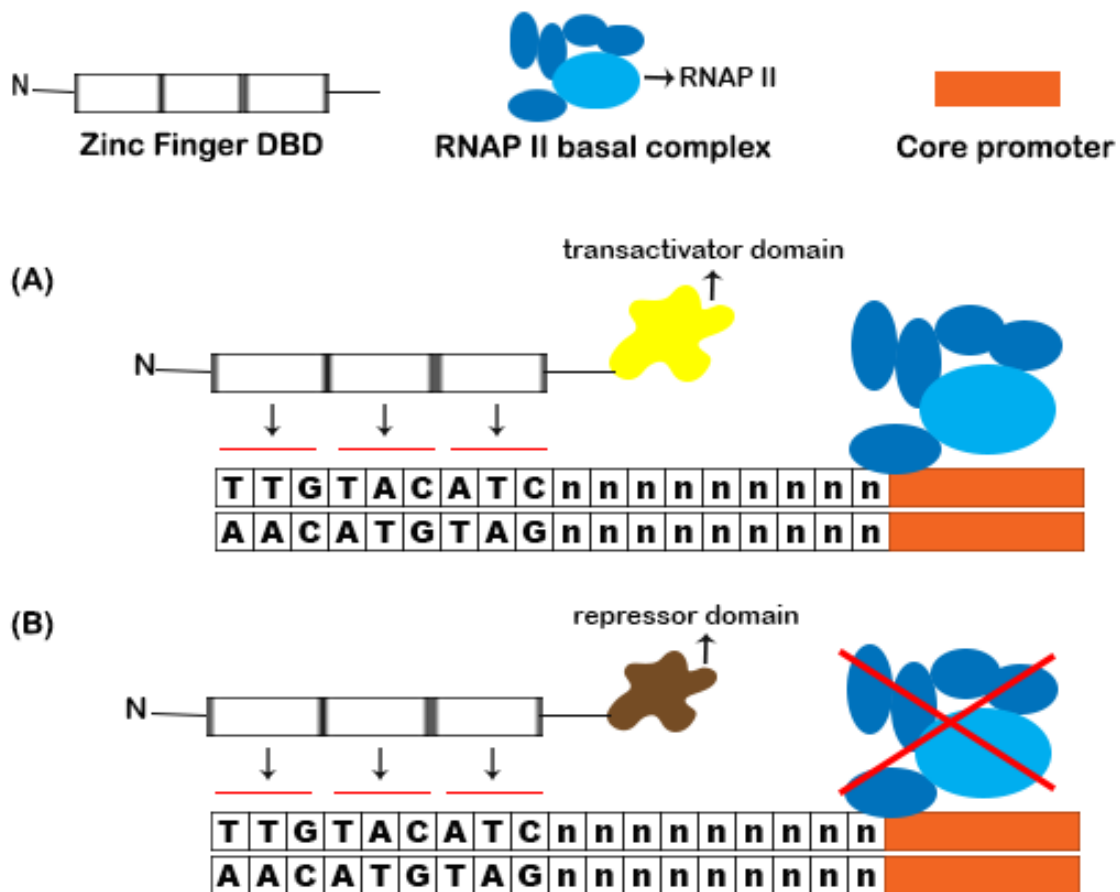


Fig. 3

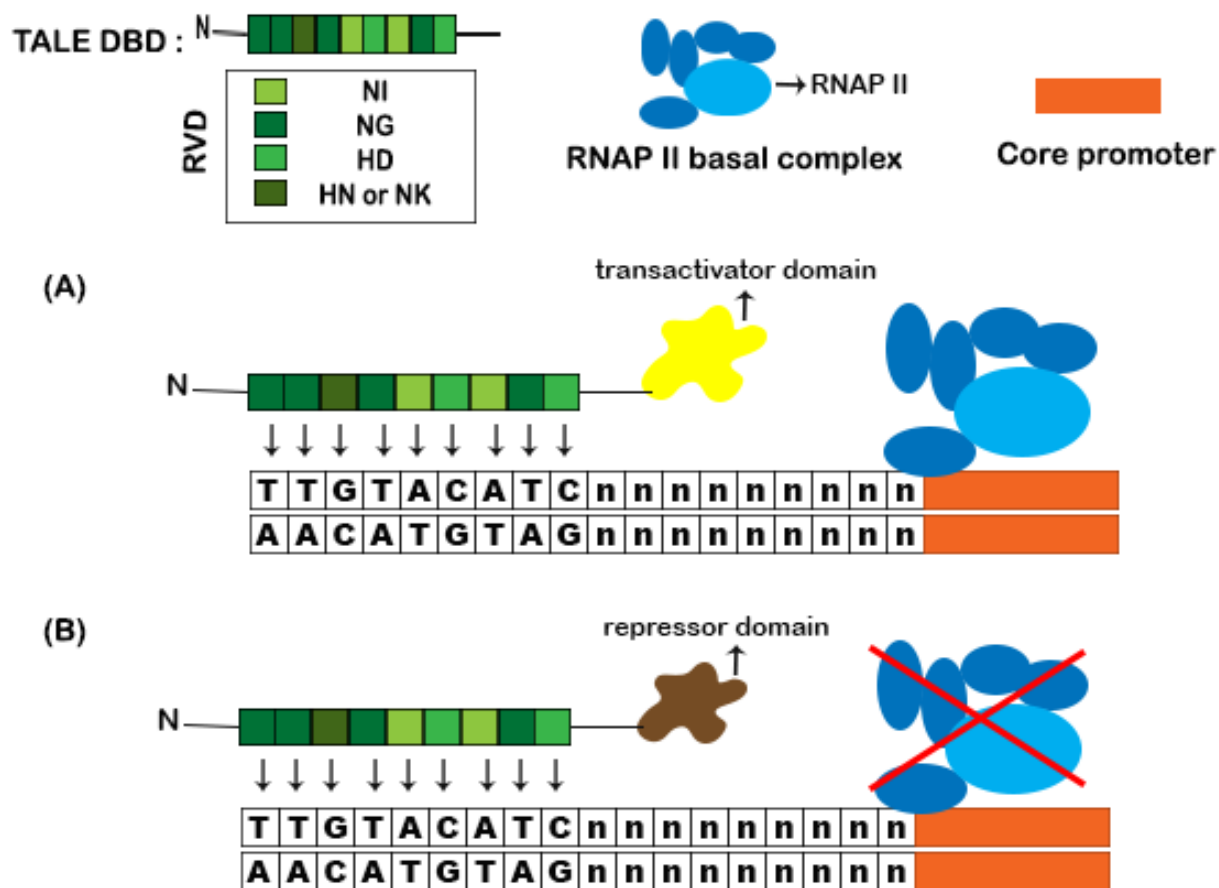


Fig. 4



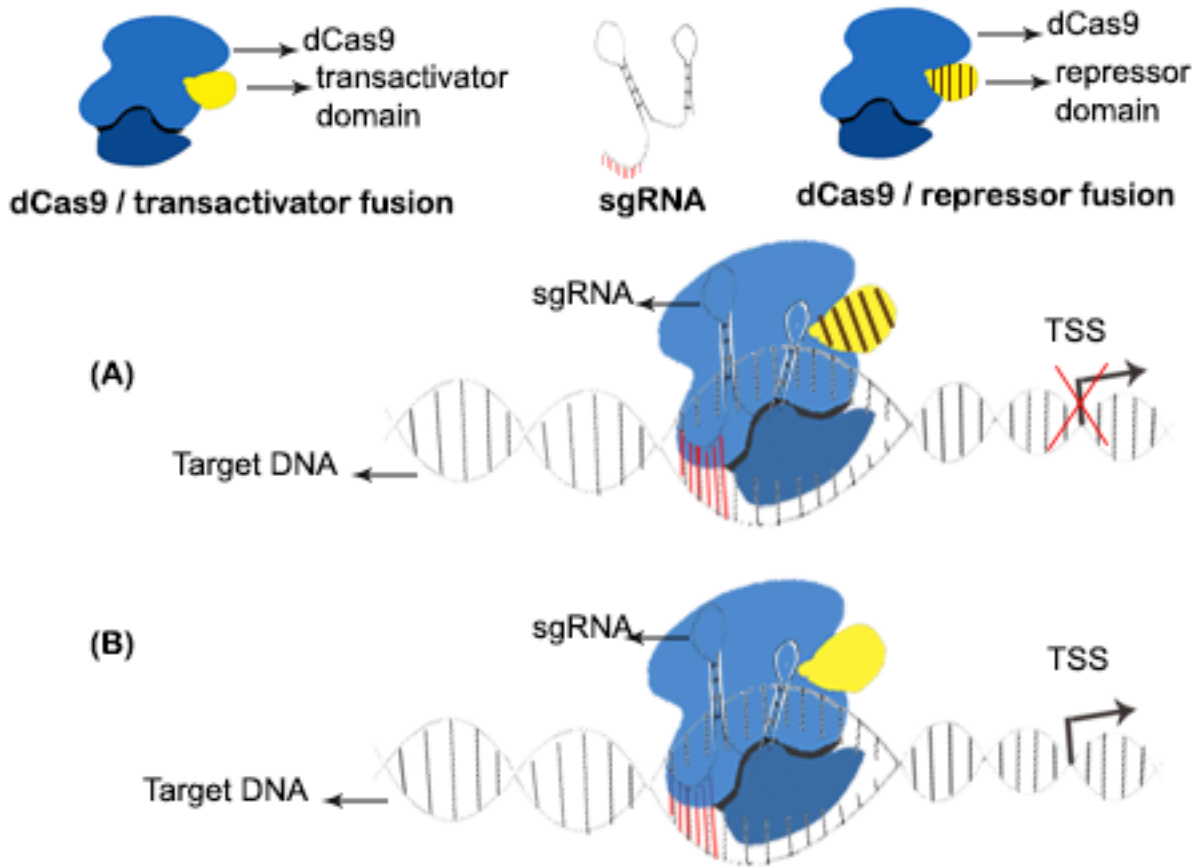
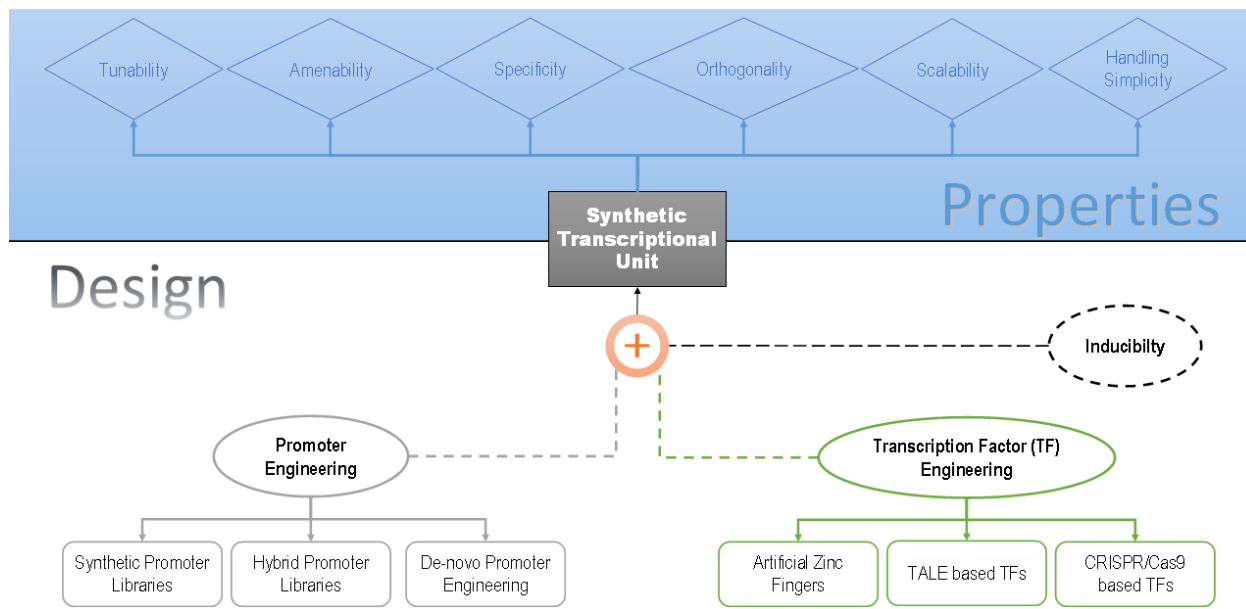


Fig. 5



Graphical abstract