To Paskal

CRISPR-BASED SYNTHETIC TRANSLATIONAL REGULATION USING NON-CONVENTIONAL YEAST

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

MATERIALS SCIENCE AND NANOTECHNOLOGY

By

DAMLA ALBAYRAK

JULY 2025

CRISPR-BASED SYNTHETIC TRANSLATIONAL REGULATION USING NON-CONVENTIONAL YEAST

By Damla Albayrak

July 2025

We certify that we have read this dissertation and that in our opinion it is fully adequate,in scope and in quality, as a thesis for the degree of Master of Science.

_____
Urartu Özgür Şafak Şeker (Advisor)

_____
Sreeparna Banerjee

_____
Fulya Türker

Approved for the Graduate School of Engineering and Science:

_____

Orhan Arıkan
Director of the Graduate School

ABSTRACT

CRISPR-BASED SYNTHETIC TRANSLATIONAL REGULATION USING NON-CONVENTIONAL YEAST

Damla Albayrak

M.Sc in Materials Science and Nanotechnology

Advisor: Urartu Özgür Şafak Şeker

July, 2025

Efficient and programmable gene expression systems are essential for improving recombinant protein production in non-conventional yeast hosts such as *Pichia pastoris*. In this thesis, a synthetic gene regulation platform was established in *P. pastoris* by integrating rationally engineered GAP promoters with CRISPR/dCas9-based transcriptional activation modules. The aim was to convert the native constitutive GAP promoter into a tunable element capable of both activation and repression through gRNA-guided recruitment of effector domains, thus paving the way for orthogonal and context-specific control of gene expression. Two synthetic promoter variants (version 1 and version 2) were designed by introducing targeted mutations to create novel gRNA binding sites without disrupting core promoter function. These promoters were cloned upstream of an eGFP reporter and integrated into the genome of *P. pastoris*. Colony screening under various carbon sources (glucose, glycerol, ethanol, and methanol)

revealed that most mutant promoters retained expression levels comparable to the wild-type PGAP, while certain clones displayed elevated eGFP production due to multiple gene integrations. Quantitative PCR analysis was employed to identify single-copy integrants for further use. Subsequently, a CRISPRa system comprising dCas9, MS2-binding scaffold RNAs, and the VP64 activation domain was introduced into selected single-copy clones. Ten custom-designed gRNAs (five for each promoter version) were tested under four carbon conditions to assess their activation potential. Notably, version 1 demonstrated robust transcriptional activation with specific gRNAs, especially v1-g2-c1, which significantly enhanced eGFP expression across all tested conditions. In contrast, version 2 failed to elicit notable activation, possibly due to unfavorable gRNA positioning or inhibitory mutations within the promoter sequence.

This work introduces a modular and orthogonal transcriptional regulation system in *P. pastoris*, offering dynamic control over synthetic promoters using CRISPRa components. The approach establishes a foundation for future metabolic engineering strategies and recombinant protein expression systems that are independent of traditional inducible promoters and adaptable to various industrial contexts.

Keywords: Synthetic promoter engineering, CRISPRa / CRISPRi, *Pichia pastoris*, Gene expression regulation, Recombinant protein production

ÖZET

CRISPR Tabanlı Sentetik Translasyonel Regülasyonun Konvansiyonel Olmayan Maya Kullanılarak Gerçekleştirilmesi

Damla Albayrak

Malzeme Bilimi ve Nanoteknoloji, Yüksek Lisans

Tez Danışmanı: Urartu Özgür Şafak Şeker

Temmuz, 2025

Verimli ve programlanabilir gen ekspresyon sistemleri, *Pichia pastoris* gibi model olmayan maya konakçılarında rekombinant protein üretiminin iyileştirilmesi için kritik öneme sahiptir. Bu tez kapsamında, *P. pastoris*'te rasyonel olarak tasarlanmış GAP promotorları ile CRISPR/dCas9 tabanlı transkripsiyonel aktivasyon modüllerinin entegrasyonu yoluyla sentetik bir gen düzenleme platformu kurulmuştur. Bu çalışmanın temel amacı, doğal olarak konstitutif ifade gösteren GAP promotorunu, gRNA aracılığıyla efektör bölgelerin yönlendirilmesiyle hem aktivasyon hem de represyon yapılabilecek şekilde ayarlanabilir bir yapıya dönüştürmektir. Böylece, ortogonal ve bağlama-özgü kontrol sağlanması hedeflenmiştir. Bu amaçla, GAP promotorunun çekirdek işlevini bozmadan yeni gRNA bağlanma bölgeleri oluşturmak üzere iki farklı sentetik versiyon (versiyon 1 ve versiyon 2) rasyonel mutasyonlarla tasarlanmıştır. Bu promotorlar, eGFP raporlayıcı geninin önüne klonlanarak *P. pastoris* genomuna

entegre edilmiştir. Glukoz, gliserol, etanol ve metanol gibi çeşitli karbon kaynakları altında yapılan koloni taraması sonucunda, çoğu mutant promotorun doğal tip PGAP promotoru ile karşılaştırılabilir düzeyde ekspresyon sağladığı gözlenmiştir. Bazı kolonilerde ise yüksek eGFP üretimi gözlenmiş ve bu durum, qPCR analizi ile doğrulanan çoklu gen kopya sayısına bağlanmıştır. Bu nedenle, deneylerin sağlıklı yorumlanabilmesi için tek kopya taşıyan koloniler seçilerek sonraki aşamalarda kullanılmıştır. Ardından, dCas9, MS2-bağlayıcı scaffold RNA'ları ve VP64 aktivasyon domenini içeren CRISPRa sistemi, seçilmiş tek kopya taşıyan kolonilere aktarılmıştır. Her promotor versiyonuna yönelik beşer adet olmak üzere toplam on özel tasarlanmış gRNA, dört farklı karbon kaynağı altında test edilmiştir. Özellikle versiyon 1, belirli gRNA'larla, en dikkat çekici olarak da v1-g2-c1 ile, tüm koşullarda anlamlı düzeyde eGFP aktivasyonu sağlamıştır. Buna karşın, versiyon 2 ile anlamlı bir aktivasyon elde edilememiş; bu durum, muhtemelen bağlanma bölgesinin konumu ya da promotor dizisindeki inhibe edici mutasyonlara bağlanmıştır.

Bu çalışma, *P. pastoris*'te sentetik promotorlar üzerinde CRISPRa bileşenleri kullanılarak dinamik kontrol sağlayan modüler ve ortogonal bir transkripsiyonel düzenleme sistemini tanıtmaktadır. Sunulan yaklaşım, geleneksel indüklenebilir promotorlara alternatif olarak, gelecekteki metabolik mühendislik stratejileri ve endüstriyel rekombinant protein üretim uygulamaları için özelleştirilebilir ve kontrol edilebilir gen ekspresyon sistemlerinin geliştirilmesine zemin hazırlamaktadır.

Anahtar Kelimeler: Sentetik promotor tasarımı, CRISPRa / CRISPRi sistemleri, *Pichia pastoris,* Gen ekspresyonu düzenlemesi, Rekombinant protein üretimi

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr. Urartu Özgür Şafak Şeker for accepting me into his research group and supporting me throughout my thesis work. After transitioning from another research group, I was welcomed into the Synthetic Biology Laboratory with openness and encouragement. This opportunity provided me with a dynamic environment where I could grow both scientifically and personally. I appreciate his guidance during this process and the opportunities he provided.

I would also like to thank my valuable jury members, Prof. Sreeparna Banerjee and Asst. Prof. Fulya Türker, for spending time on my study to review and improve.

I would also like to thank the members of the Synthetic Biology Laboratory for their help and companionship throughout this process. Whether it was discussing experiments, solving small problems together, or just sharing the everyday lab routine, their presence made things easier. I'm glad I had the chance to work with them.

I also thank Dr. Burcu Gündüz Ergün for identifying the mutations and gRNA binding sites used in the promoter design process, and Abdurahman Atilla for his assistance with some of the repressor construct clonings. Their support at key steps of the experimental workflow is sincerely appreciated.

I would like to express my heartfelt thanks to Senem Şen. She has been the most precious part of this journey. Her love, care, friendship, and unwavering support—both emotionally and academically—have meant the world to me. I truly don’t think I could

have made it through this process without her by my side. She is one of the greatest gifts life has given me. Thank you for everything, Senem. I'm so lucky to have you.

I would also like to thank Şimal Şenocak and Damla Karadavut, who have been by my side since my undergraduate years. I'm truly grateful for their constant support and for reminding me that we can get through any challenge together. I'm so glad to have you both in my life.

Finally, I would like to express my deepest appreciation to my family—especially to my sister Beyzanur Demirören, to my soon-to-be-born nibling whose arrival I eagerly await, to my brother Kağan Albayrak, and to my parents Münüre and Semih Albayrak. Their constant love, trust, and encouragement have been the foundation of my strength throughout this journey. I am truly grateful to have such a supportive and caring family behind me.

This study was supported by TUBITAK, project number 222Z229.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

1.1 Non-Conventional Yeast- *Pichia pastoris*

Highly efficient and controllable gene expression systems are essential for the biotechnological industry. Past several decades yeast biotechnology has been used for manufacturing industrial biocatalysts, biopharmaceuticals, food and beverage products, and bioethanol. Yeast is the most preferred eukaryotic host organism for recombinant protein production due to its ability to multiply rapidly, be easily genetically manipulated, and perform post-translational modifications. Thus far, strains of the traditional baker's yeast, *Saccharomyces cerevisiae*, have been the primary cell factories used in the bioconversion processes. It is mainly because of is classified as Generally Regarded as Safe (GRAS) organism [1]. Forestry, agricultural, industrial, and other low-cost, renewable carbon sources are transformed into a variety of bioproducts, such as biofuels, bulk and fine biochemicals, and biomaterials, in next-generation manufacturing processes [2]. However, limited production capacity and complex transcriptional regulation mechanisms limit the industrial applications of yeast. *S. cerevisiae* lacks many of the characteristics necessary for the effective conversion of these alternative substrates into products of interest, and such processes necessitate a variety of distinct cell factories designed for the particular goals. For instance, strains of *S. cerevisiae* of the wild type do not easily exploit numerous carbon sources at once and instead prefer hexose sugars like glucose. This reduces the yeast's ability to ferment mixed-sugar substrates like lignocellulosic hydrolysates, which have

large concentrations of pentoses in addition to hexoses. Although *S. cerevisiae* produces good amounts of alcohols, esters, and organic acids, its uses as a cell factory are limited because it does not naturally collect large amounts of intracellular lipids or use other residual carbon sources such short chain fatty acids [3]. Despite *S. cerevisiae* is the most commonly used yeast model organism, many different organisms are used in microbial chemical production such as *Escherichia coli* or *Aspergillus genus*. *E. coli* is widely known for producing vitamins and fine compounds, while *Aspergillus* is well known for producing enzymes and citric acid [4]. As in *cerevisiae*, there are difficulties in the production processes in these host organisms. For example, glycosylated biopharmaceutical products, complexly assembled proteins, and proteins with a large number of disulfide links cannot be produced by *E. coli* [5]. In a similar way the filamentous growth of *Aspergillus* in industrial tanks or bioreactors is linked to numerous issues, making the operation more laborious and costlier. The list of issues that need to be resolved can also include costly feedstocks, excessive energy and water use, productivity losses brought on by contamination, and downstream separation expenses[4]. Due to these problems, there has recently been an increasing interest in non-traditional microorganisms for the production of therapeutic proteins due to their specific properties such as their unique metabolic chemistry or high-capacity biochemical pathway engineering. Non-traditional or more commonly known as non-conventional organisms essentially means that it denotes any microorganism distinct from the traditional industrial workhorses previously stated. Moreover, non-conventional yeasts may offer numerous benefits over *S. cerevisiae* regarding metabolic pathway needs, product profiles, and growth physiology [6]. The application of non-conventional yeasts could effectively address several of the previously identified difficulties. The physiology, metabolic pathways, and regulatory mechanisms of these

non-conventional yeast strains significantly surpass those *of S. cerevisiae*, hence enhancing their applicability in diverse processes such as heterologous protein production [6]. A wide range of carbon sources found in renewable feedstock, including monosaccharides (xylose, galactose, fructose, arabinose), disaccharides (lactose, cellobiose, xylobiose), and even polysaccharides (xylan and cellulose), can be used by many non-conventional yeasts to grow and ferment under a variety of process conditions [3]. While some non-conventional yeasts exhibit tolerance to inhibitors generated during various bioprocesses, others may tolerate harsh environmental conditions such high temperatures and osmotic pressure [4]. Therefore, non-conventional yeasts are commonly employed as cell factories, although only a small number of them have been characterized in comparatively considerable detail such as *Pichia pastoris* (also known as *Komagataella phaffii*), *Yarrowia lipolytica* and *Kluyveromyces marxianus* [3]. Alexandre Guilliermond initially isolated the methylotrophic yeast *Pichia pastoris* from a chestnut tree in France and named it *Zygosaccharomyces pastori* [7]. Herman Phaff dubbed the species *Pichia pastoris* in the 1950s after isolating further strains from black oak trees in California, USA [8]. *P. pastoris* can thrive on methanol as its sole form of carbon and energy source because it is a methylotrophic yeast. An effective protein manufacturing system was created using the methanol utilization-related genes' incredibly high and methanol-inducible expression level, specifically the alcohol oxidase gene *AOX1* [9]. Today, *Pichia pastoris* is the most widely used yeast cell in academia and industry for recombinant protein production, and more than 5000 protein molecules have been produced by the methylotrophic yeast *P. pastoris* to date [10]. High yields of recombinant proteins with glycosylation that closely resembles that of mammalian cells can be produced by *P. pastoris* [11]. In contrast to mammalian cells, *P. pastoris* have a eukaryotic protein

synthesis pathway, are very straightforward to alter genetically, and do not require a sophisticated growth medium or culture conditions. Moreover, for several reasons, such as ease of genetic manipulation (e.g., gene targeting), high-frequency DNA transformation, cloning by functional complementation, high levels of intracellular or extracellular protein expression, and the capacity to carry out higher eukaryotic protein modifications, this methylotrophic yeast is especially well-suited to the expression of foreign proteins [12]. *P. pastoris*, in contrast to *E. coli*, uses vesicular trafficking and eukaryotic folding chaperones to process complex proteins (Figure 1), although these pathways also need to be optimized for the expression of particular proteins [10]. Furthermore, *P. pastoris* culture conditions that are best suited for heterologous protein expression in this yeast system include pH 5–6, temperatures 25–30°C, the presence of antifoam agents, aeration of 0.1–1.0 v.v.m., and agitation speeds of 500 and 1500 r.p.m. for bioreactors (250–300 r.p.m. for shake flask) [10].

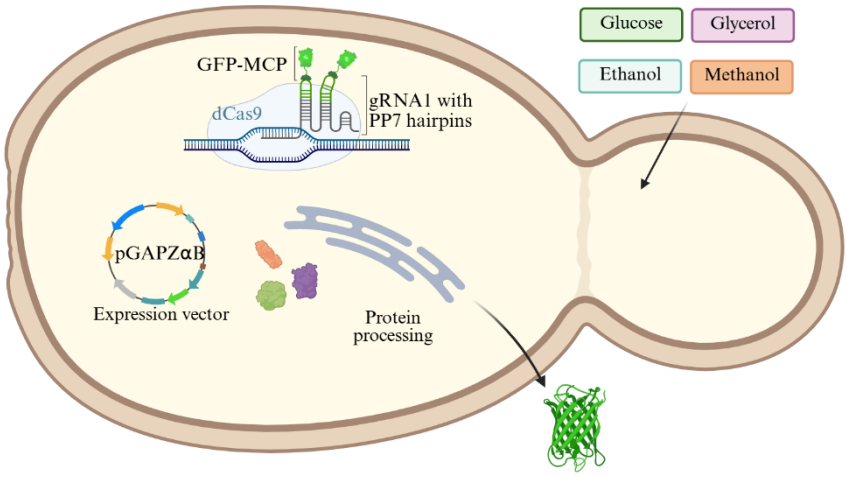

Figure 1: Schematic overview of CRISPR-based gene expression and protein production in *Pichia pastoris*. The engineered yeast strain carries a pGAPZαB expression vector for recombinant protein production. Additionally, a dCas9-based synthetic transcriptional regulation system is employed, where a gRNA containing PP7

hairpins recruits the GFP-tagged MCP protein to enhance gene expression. Different carbon sources (glucose, glycerol, ethanol, and methanol) are used to modulate expression levels. The synthesized protein undergoes processing in the endoplasmic reticulum and is secreted out of the cell. Illustration drawn using BioRender.com.

1.2 CRISPR-dCAS9 System for Genome Editing in *Pichia pastoris*

The restricted genetic tools currently available restrict the use of *P. pastoris* in metabolic engineering and synthetic biology, despite the growing interest in this yeast for the manufacture of bioproducts. The metabolic pathways for cell proliferation and the synthesis of the intended product frequently compete in strains that produce biotechnological goods. To optimize new flux patterns and guarantee the synthesis of the intended product in these situations, it is crucial to have easy control over the metabolic enzyme sets, particularly bidirectional (on/off) control [13]. Eukaryotic cells' complex and sophisticated phenotypes are an example of regulatory networks made up of several layers and distinct expression programs that involve the transcription of numerous genes. By binding to transcription factors binding sites, where regulatory information is encoded in promoter genes through metabolic signals, and by interacting either antagonistically or synergistically with one another, eukaryotic transcription factors (TFs), the most crucial parameters in this network, regulate patterns of gene expression [14]. Using synthetic or natural TFs to reengineer the endogenous transcription network is a potent method for controlling cell phenotypic and studying cell activities [15]. In earlier research, effector molecules (activators and repressors) were directed to regulatory areas of eukaryotic genes to regulate their transcription by the use of native DNA binding domains (DBD), such as TetR, LacI, and LexA, which

are primarily obtained from bacterial sources. Nevertheless, since DBD-specific operator sections must be added to the cis-regulatory regions of promoter genes, this strategy is costly and time-consuming, particularly when addressing the regulation of several genes. Additionally, orthogonal transcription factors are necessary for the construction and regulation of complex transcriptional networks, which can be extended to other systems. Only a few number of orthogonal natural DBD variations have been thoroughly described, and it has been demonstrated that changing their specificity is quite challenging [16]. This limits the use of natural DBD-based TFs for the reorganization of synthetic gene circuits and intricate transcription networks. Synthetic TFs based on zinc fingers (ZFs) and transcriptional activator-like effectors (TALEs) have been created to get around these restrictions [17]. ZFs and TALEs are desirable DBDs for creating orthogonal TF libraries because their specificity can be programmed to target any gene. Synthetic TFs based on ZF and TALE have been demonstrated to work in a variety of eukaryotic cells, such as mammalian, plant, and yeast cells [18]. But getting a specific TF necessitates laborious selection procedures and a multi-step DNA assembly approach [19]. Furthermore, the amount of TFs encoded at the same time and the metabolic load produced in the cell may limit the scale of control that these TFs can accomplish [20]. The field of synthetic biology has revolutionized the reprogramming of cells by constructing genetic parts or modules into more complex genetic circuits. This approach provides tools applicable to many organisms to control cell response in new and predictable ways. To achieve this, genetic parts are often taken from various organisms and used to create orthogonal, independent functions within the cell. Because they are not found in the host organism naturally, orthogonal control components of genetic circuits can operate independently of or in response to host signaling pathways.

One element of the bacterial and archaeal immune systems is CRISPR (clustered regularly interspaced short palindromic repeats)–Cas (CRISPR associated protein). Many bacteria use CRISPR-based immune systems to degrade the genetic material of invading phages [21]. The endonuclease protein (Cas9) is guided to the invasive genetic material in these systems by the expression of short RNAs from CRISPR loci. Here, it is possible to produce double- (or single-) strand breaks in DNA by targeting the Cas9 endonuclease to certain loci in living cells using specially made short guide RNAs (gRNA). Error-prone or template-based repair mechanisms are triggered when DNA breaks develop, producing variations of the original target loci. The type II CRISPR/Cas system, which needs the production of only one Cas protein—typically *Streptococcus pyohegenes* Cas9 (SpCas9)—is the most commonly utilized system in biotechnology, despite the fact that six distinct CRISPR-Cas systems, classified into two classes, have been identified to date. D10A and H841A mutations in the RuvC and HNH nuclease domains produced dCas9 (deficient Cas9), poor in nuclease action which can be modified to create novel transcription factors [22]. In contrast to zinc finger proteins and TAL effectors, dCas9-based transcription factors may generate new DNA binding sites without requiring the assembly of several lengthy DNA sequences. Rather, the 20-bp nucleotide sequences that come after the PAM sequence are what sgRNAs are made to match. Furthermore, dCas9 or the sgRNA itself can be directly coupled with activation or repression domains (Figure 2). Both natural and manufactured synthetic promoters can be targeted and regulated with the help of these transcription factors.

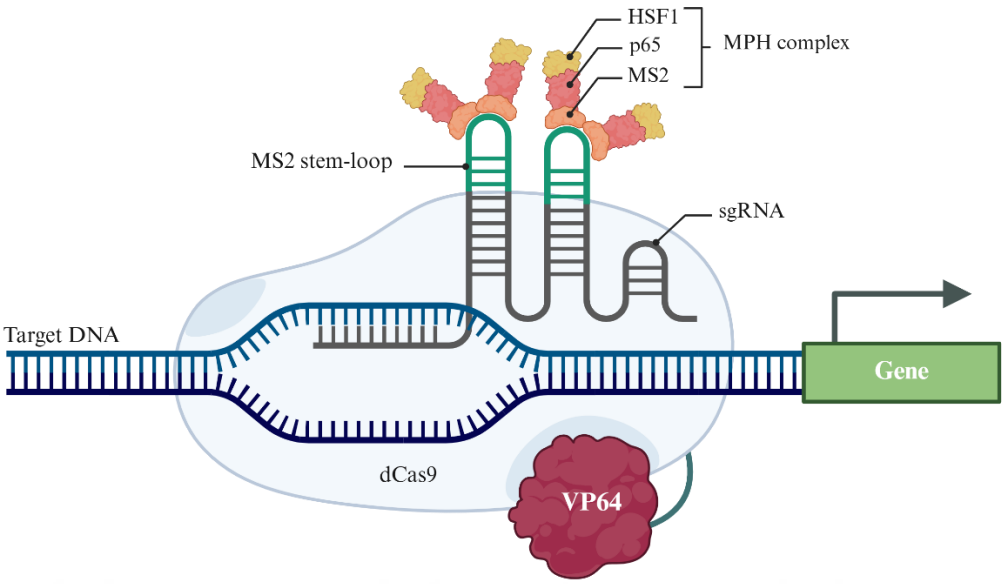

Figure 2: Schematic representation of the CRISPR-dCAS9 transcriptional activation system. A nuclease-dead Cas9 (dCas9) is guided to a specific genomic locus by a single guide RNA (sgRNA). The sgRNA is engineered to contain MS2 stem-loop structures, which are recognized by the MS2 coat protein fused to p65 and HSF1 (MPH complex). In parallel, the dCas9 is fused to the VP64 activation domain. Together, these components synergistically recruit the transcriptional machinery to the promoter region, enhancing the expression of the target gene. Illustration drawn using BioRender.com.

Based on the bacterial CRISPR/Cas system, methods for modifying synthetic and natural promoters with synthetic transcription factors have been created to get over the difficulties in creating novel transcriptional gene circuits. A programmable "CRISPRi" tool for gene silencing in *Escherichia coli* bacteria has been developed using dSpCas9 [22]. dCas9 can suppress the expression of a target gene by blocking the progression of RNA polymerase when it is directed to a certain promoter or *ORF* (open reading frame) gene. Fusing the omega subunit of RNA polymerase to dCas9 has been demonstrated

to produce transcriptional activation in *E. coli* in addition to transcriptional repression [23]. Codon-optimized *S. pyogenes* dCas9, SV40 nuclear localization signal (NLS), and four copies of the tandem Herpes Simplex Viral Protein 16 (VP64, commonly used eukaryotic activator domain) genes were combined to develop dSpCas9-VP64 in order to modify the CRISPR-TF system for the yeast *S. cerevisiae* [24]. GFP expression that was cloned using the minimal CYC1 promoter was used to test the activity of dSpCas9-VP64. Interestingly, dSpCas9-VP64 directed close to the TATA-box and transcription start site inhibited GFP production, whereas dSpCas9-VP64 directed to the upstream sequence of the TATA box provided activation [24]. Preventing the development of the transcription initiation complex is most likely how dSpCas9-VP64 inhibits. Expression rose dramatically as the number of synthetic target sites (up to 12) added to the promoter upstream of the TATA box to which dSpCas9-VP64 would bind increased. However, there was no statistically significant difference observed between promoters with three to twelve synthetic target binding positions [24]. In a related study, binding sites with a range of 2 to 16 were randomly inserted upstream of the minimal CYC1 promoter induced by dSpCas9-VP64. The promoter with 16 binding sites exhibited very low activity, and there was no correlation between the number of binding sites and promoter activity [25]. The reduced promoter activity in this study could be due to the close proximity of the dCas9-VP64 binding sites (6 nucleotides) to one another. Gilbert et al. (2013) used dCas9 in conjunction with activator and repressor domains to accomplish CRISPR-mediated gene activation and repression in yeast and mammalian cells. In this work, dCas9 was coupled with the yeast histone deacetylase Sin3 homolog, which interacts with the mammalian transcription repressor Mxi1 [26]. The endogenous *TEF1* gene in *S. cerevisiae* was shown to be suppressed 18 times when dCas9 was directed to it alone, and 53 times when dCas9-Mxi1 was addressed [27]. The findings imply that

dCas9 inhibits other transcription factors sterically in order to suppress transcription. dCas9 combined with the chromatin remodeling protein Mxi1 created orthogonal and digital responsive complex NOR circuits in *S. cerevisiae* yeast [28]. In this work, seven gRNAs were used to build logic circuits with seven-layer repression cascades. Only one-way regulation is possible in the system created when the Cas9 protein fuses with activator or repressor proteins. The design of regulation programs that concurrently target and induce distinct loci is limited by this circumstance. In order to get over this restriction, Zalatan et al. (2015) expanded the CRISPR single-guide RNA sequence with modular RNA sequences that will produce RNA-binding proteins, creating the scaffold RNA (scRNA) structure. By encoding both a gRNA that targets a particular locus and an aptamer region like MS2, PP7, and com to add the regulatory function, modular scRNA physically binds DNA binding and effector delivering functions. Both the target locus and the information indicating which regulatory function will be carried out at that locus are encoded by a single scRNA molecule. With this method, many target genes can be simultaneously activated and repressed inside a single regulation program, enabling multi-way control. Three orthogonal RNA-protein binding modules (MS2, PP7, and com) in *S. cerevisiae* were used to test scRNA designs [29]. When the effector binding sequence was appended to the 5' end of the sgRNA in scRNA designs, no activity was seen. Furthermore, a diminished regulating impact was noted when the linker length between the sgRNA and protein binding hairpin domains exceeded two bases (up to twenty bases). Moreover, compared to the dCas9-VP64 fusion, the CRISPR-TF method produced noticeably higher gene activation when the VP64 activator was added via scRNA. Using chromatin modifiers and the VP64 activator and KRAB transcription repressor domains, scRNA was used to regulate the expression of five distinct genes in the violasein pathway in *S. cerevisiae* yeast. Eight distinct on/off

combinations were used to produce distinct product profiles [29]. Nevertheless, it was shown that the fusion proteins MCP-KRAB and PCP-KRAB were ineffective at inhibiting transcription. This might be because the dimer structure of PCP and MCP inhibits the potential function of KRAB. Using the CRISPRa system based on scRNA, VP64 activator domain, and the synthetic GAL7 promoter in *S. boulardii* cells—a probiotic yeast—a two-step orthogonality transactivation method was recently created [30].

In the yeast *Pichia pastoris*, the CRISPR/Cas9 system has recently been demonstrated to be a programmable tool for genome editing [31]. In order to create a CRISPR/Cas9 system for *P. pastoris*, Weninger et al. (2016) methodically examined a number of modules, including the Cas9 coding sequence, gRNA sequences, and gRNA constructions (like ribozyme sequences). Of the 95 constructs assessed for genome editing, only six demonstrated functionality [32]. The efficiency of deleting the ADE2 locus has been demonstrated to increase from 10% to 80% when panARS is substituted for the PARS1 replication origin sequence in episomal CRISPR/Cas9 plasmids [33]. In order to generate several sgRNAs from a single cassette (multiplex), Dalvie et al. (2020) investigated *P. pastoris* RNA PolIII promoters (tRNA, snoRNA, and srpRNA) and tRNA sequences. Up to 40% simultaneous genome editing efficiency was attained in *P. pastoris* for three distinct targets using the multiplex gRNA technique, which targets three distinct heterologous gene integrations for the humanized glycosylation process [34]. The primary method for fixing double-stranded DNA breaks in the yeast *Pichia pastoris* is non-homologous end-joining (NHEJ). The integration efficiency at two and three loci was reported to be between 58 and 70 percent and 13 and 32 percent, respectively, in the *P. pastoris* strain with non-homologous end-joining mechanism removed (Δku70) for multiplex gene integration with CRISPR/Cas9 [35]. Another

study sought to improve the efficiency of CRISPR-based targeted gene editing by deleting some genes involved in the *P. pastoris* gene recombination mechanism and overexpressing others, particularly *RAD52*. The efficiency of homologous recombination and the simultaneous integration of numerous fragments were enhanced by lowering the minimal homologous gene arms for donor DNA to 50 bases [36]. Few investigations have been conducted on synthetic gene regulation in *P. pastoris*, despite the fact that CRISPR-based gene regulation concepts have been developed for CHO and *S. cerevisiae*, which are often used hosts for the generation of recombinant antibodies. Baumschabl et al. (2020) used the scaffold RNA system and VP64 activator module to activate the suppressed THI11 promoter in the presence of thiamine. Additionally, they used the CRISPR-scRNA technique to increase *P. pastoris RIB1* expression, which in turn improved riboflavin production [37]. By combining the dCas9 protein with activator (VPR) or repressor (Mxi1) modules, Liao et al. (2021) created a system that uses a multiplex gRNA system to simultaneously activate, repress, and regulate genes in *P. pastoris*. However, attempting to use the same dCas9-VPR fusion protein for all three of these purposes leads to issues with the system's ability to function simultaneously. This work shows that dCas9-VPR improves the level of eGFP expression when it is directed to the GAP promoter and decreases mCherry expression when it is directed to the CYC1 promoter. The use of the Cas9 system, which contains the same effector molecule as both an activator and a repressor, is illogical, and this conclusion runs counter to the literature on VPR activators. The absence of control data pertaining to the targeting of Cas9 in the absence of an effector molecule is another flaw in the results presentation. A true CRISPR-based simultaneous activation and repression system has not yet been created, as the result shown as repression data is most likely the consequence of dCas9-VPR reducing transcription through steric

hindrance of RNA Polymerase or other transcription initiator complex modules [38]. Recently, a synthetic regulation mechanism called the SynPic-X platform was created using CRISPR-based activator and repressor modules to create a system for producing recombinant proteins in *P. pastoris* without the use of methanol [39]. The core module of the *P. pastoris* AOX1 promoter and the *E. coli* LacO operator were combined to create a hybrid promoter. The activation domain of the Mit1 TF was fused to the LacI protein, and the hybrid transcription activator was used to induce strong gene expression activation. The created hybrid promoter and hybrid activator system not only creates a potent expression system but also results in extremely leaky expression. Repressor and activator methods based on CRISPR have been developed to regulate expression [40].

1.3 The Aim of Study

The use of antigen-specific agents together with other drugs improves the prognosis of many diseases from cancer to infectious diseases. In diseases that cannot be treated (such as hereditary diseases), it provides improvement of symptoms. However, antibody production with today's recombinant technologies is quite expensive and is performed with sensitive mammalian cell lines. Purification of the antibodies obtained by the virus transfection method used in the production process by cleaning them from viral agents increases the costs per antibody dose. Making the production costs of antibodies more accessible by reducing them is essential for widespread use of treatments. *Pichia pastoris* is a candidate that can be an alternative to mammalian cell lines for antibody production due to its higher secretion capacity and humanized glycosylation pathway compared to other yeasts. The fact that the medium is cheap,

DNA transfection and genome modification tools have been developed, and the ability to produce more extracellular protein per liter are some of the features that make *P. pastoris* yeast advantageous in the production of recombinant antibodies compared to mammalian cell lines. Despite the advantages it provides, it has been less studied than the evolutionarily similar conventional yeast *S. cerevisiae* and therefore the few characterized genetic regulation systems, the limited number of characterized transcription factor/promoter pairs and the fact that their use is condition-dependent create different limitations in the inducible protein production system of *P. pastoris*. Environmental conditions often trigger a change in metabolic state by repressing some enzyme sets while activating previously repressed enzyme sets, leading to new metabolic fluxes. These complex multi-locus, multi-directional expression programs are encoded by transcription activators, transcription repressors, or other regulatory patterns that associate with specific regions in the genome. Creating a different cell type or state requires reprogramming the cell by precisely targeting the expression of many sets of genes. Transcriptional regulation is central to the complex behavior of biological systems and synthetic gene circuits. Engineering transcriptional control is a prerequisite for cells to perform new functions and precisely modulate existing properties. Platforms for scalable, tunable, and simple modulation of transcription enable new methods of investigating natural systems and the translation of artificial capabilities into living organisms.

The aim of the thesis presented is to create a new and orthogonal inducible protein production system in *P. pastoris* with a synthetic regulation system. Synthetic transcription factors are DNA-binding proteins that are not naturally transcription factors but have been made capable of controlling transcription through different engineering methods. The aim of this thesis is to develop and characterize control

systems based on transcription repressor and initiator protein modules using the non-catalytic Cas9 protein (dCas9) and to develop an efficient antibody production system using these. In line with these objectives, our thesis objectives are as follow: Firstly, design a synthetic GAP promoter (Psyn-GAP) based on promoter engineering principles to achieve double-layer orthogonality, and to integrate it into the *P. pastoris* genome at the AOX1tt locus together with the eGFP sequence. The expression level of eGFP will then be compared to that driven by the wild-type GAP promoter in order to assess the performance of the synthetic construct. Secondly, developing a multipurpose, programmable, and versatile synthetic CRISPR-based transcription factor (CRISPR-TF) library targeting both natural and synthetic promoters in *P. pastoris* gene regulation by functionalizing dCas9 with scRNA containing effector domains. Thirdly, the VP64, p65AD, and VPR eukaryotic activator modules will be fused with the MS2 coat protein sequence and integrated into the *P. pastoris* genome from the RGI2 locus as a single gene fragment together with scRNA expression cassettes designed by fusion of sgRNA and MS2 aptamer targeting different positions of dCas9 and Psyn-GAP. The effect of the CRISPRa library on the regulation of Psyn-GAP will be determined by measuring the amount of eGFP. Fourthly, the fusion of Mxi1, KRAB, and TPL repressor domains with Com protein will be performed, and the scRNA expression cassettes designed by fusion of dCas9 and sgRNA targeting different positions of Psyn-GAP and Com aptamer will be integrated into the *P. pastoris* genome from the RGI2 locus as a single gene fragment. The effect of the CRISPRi library on the regulation of Psyn-GAP will be determined by measuring the amount of eGFP (Figure 3). In summary, the overarching aim of this thesis is to address the current limitations in gene expression control tools in *P. pastoris*, which has recently emerged as a prominent host for the production of therapeutic and prophylactic proteins. Despite its increasing importance,

the available genetic tools remain limited, and modifications to existing systems have not fully resolved the associated challenges. In light of this critical gap, the thesis proposes not to optimize current elements, but rather to develop a bidirectional gene expression control system in *P. pastoris* by employing an innovative CRISPR-based transcription factor library in combination with synthetic GAP promoter designs, thereby achieving double-layer orthogonality.

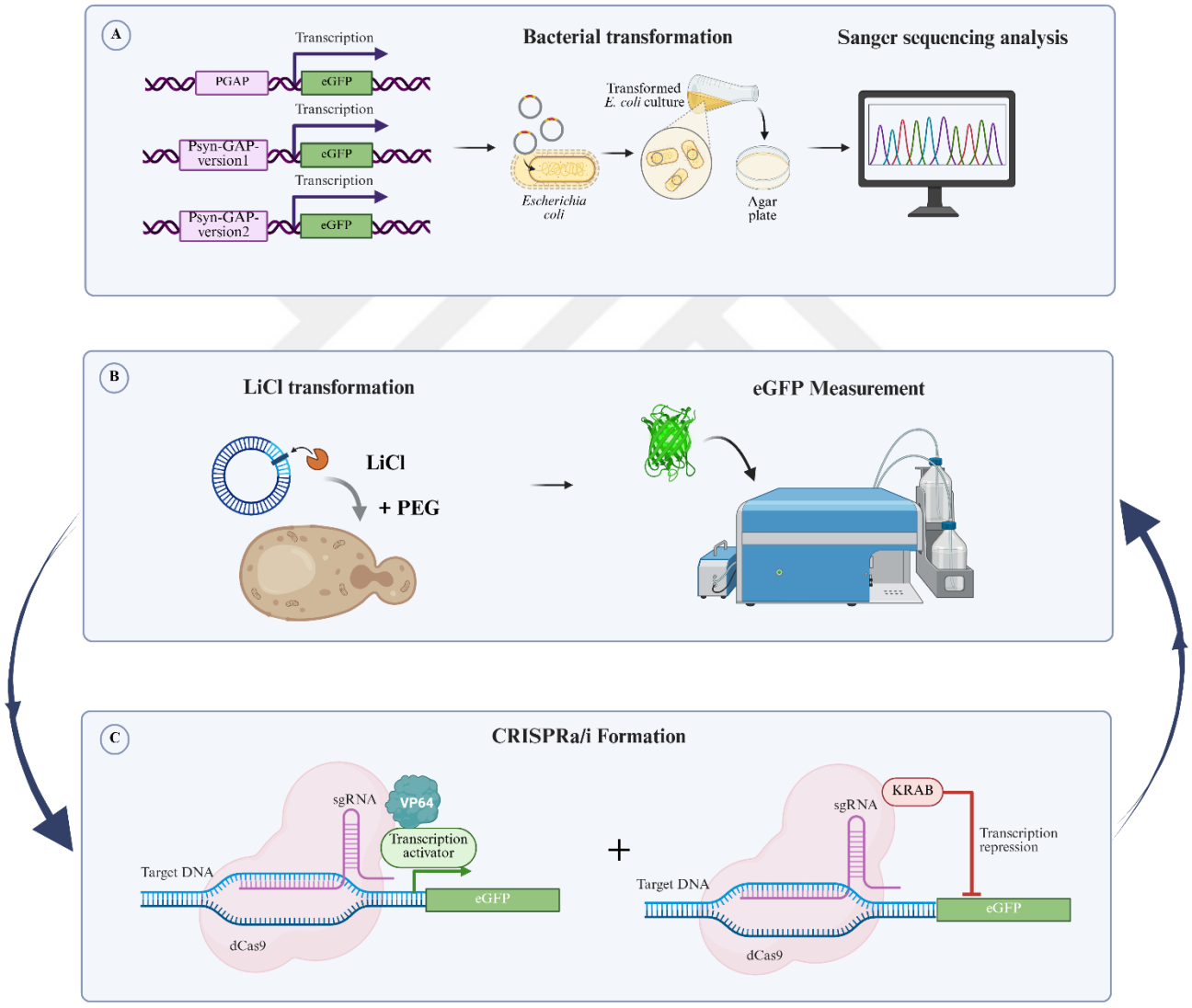

Figure 3: Schematic representation of the aim of the study. General workflow of this thesis. Illustration drawn using BioRender.com.

CHAPTER II: MATERIALS AND METHODS

2.1 Design of the Synthetic GAP Promoter and Construction of the *E. coli* Strain Containing the Psyn-GAP-eGFP Expression Cassettes

In this thesis, the *P. pastoris* GAP promoter region [41] was identified as a 477-nucleotide portion commonly used in commercial plasmids (pGAPZα-B) (This plasmid was kindly requested from BOKU, Vienna.). Potential transcription factor binding sites (TFBS) on *P. pastoris* PGAP were identified using the MatInspector tool in the analysis. A database of *Pichia pastoris* transcription factor binding sites has not yet been developed. The only database study on this subject is done with Yeastract Plus. When the Yeastract+ TF binding site consensus list is examined, it is seen that there is information for only 1 transcription factor for *P. pastoris*. The most in-depth and comprehensive studies on transcriptional regulation in yeast have been performed with *Saccharomyces cerevisiae*, and for this reason, while determining potential cis-acting elements on PGAP, the *Saccharomyces cerevisiae* TFBS database was used and regions with a matrix similarity score of at least 0.75 were accepted as potential TFBS regions. Analysis of PGAP with MatInspector revealed 42 different potential TFBS regions with a score above 0.75. The 42 different TFBSs found were grouped into 4 subgroups as carbon source consumption (glucose consumption and glucose repression), stress response elements, amino acid metabolism, and general transcription factor elements.

While designing the synthetic GAP promoter, the determined TFBSs were examined in terms of promoter engineering and mutations in regions that could potentially have positive and negative effects were avoided [42]. New PAM sequences were added by selecting regions that did not have an effect on regulation. Within the scope of the thesis, 10 different promoter loci were targeted with the CRISPRa/i system and 10 different gRNA binding sites were obtained. In previous studies, GAP promoter libraries were designed to identify and enhance the strength of cis-acting elements of the *P. pastoris* GAP promoter by random mutation [43] and deletion or insertion of potential TFBSs found by in silico tools [42]. In addition to the MatInspector results, among the mutations made in previous PGAP library designs, nucleotides that do not affect the regulation and potency of PGAP were identified and by changing these nucleotide sequences, a synthetic PGAP promoter (Psyn-GAP) library was designed for orthogonal control by CRISPR-based transcription factors. Potential guide RNA recognition sites on Psyn-GAP were determined using the Chopchop program, and gRNA binding sites were determined among the binding sites that do not recognize natural PGAP, and single guide RNA (sgRNA) designs were made based on these sites. Enhanced green fluorescent protein (eGFP) was used as a reporter protein to determine the transcription control activity of CRISPRa/i modules in *P. pastoris*. For this purpose, the *eGFP* gene sequence was obtained by PCR with Q5 polymerase from a pCat8-L2-eGFP plasmid (plasmid available in our laboratory) using DA-21 and DA-22 primers. In this way, a strain containing a wild type GAP promoter (pGAPZα-B) and *eGFP* sequence was created (wtPGAP-eGFP) (Figure 4).

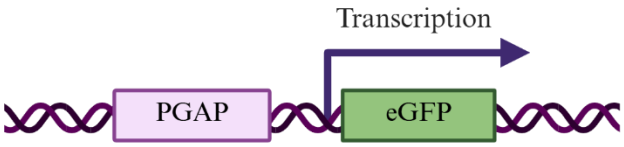

Figure 4: Schematic representation of the wtPGAP-eGFP construct. The constitutive PGAP promoter drives the transcription of the *enhanced green fluorescent protein (eGFP)* gene. The arrow indicates the direction of transcription. This construct was designed for expression analysis in the host organism. Illustration drawn using BioRender.com.

The mutation sites to be made on the GAP promoter to create the designed synthetic PGAP (Psyn-GAP-eGFP) variant were determined and divided into two (version 1 and version 2). Using the wtPGAP-eGFP plasmid, 5 new gRNA binding sites were created thanks to the 12 mutation sites specified on the Version 1 GAP promoter (the name of this new plasmid is Psyn-GAP-eGFP-version1), while 5 new gRNA binding sites were created with the 8 mutation sites specified on the Version 2 synthetic GAP promoter (the name of this new plasmid is Psyn-GAP- eGFP-version2) (Figure 5). The mutations made are stated in the Appendix A, Table A1.

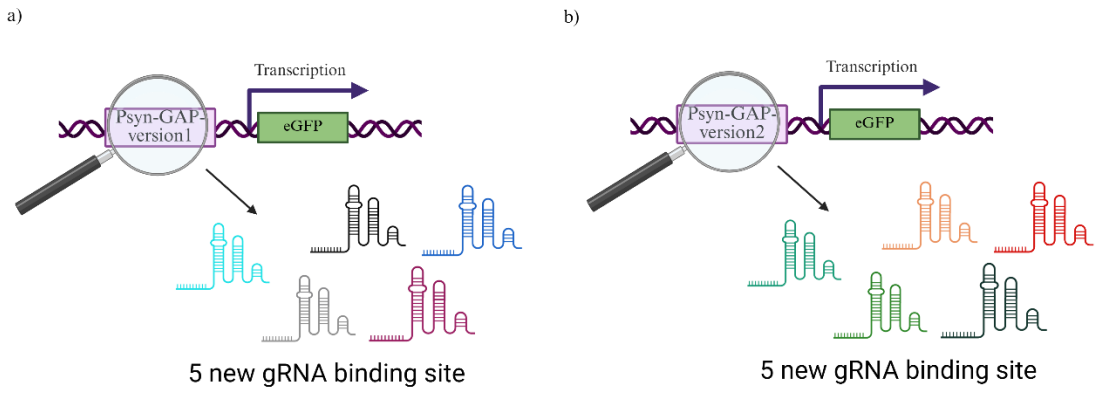

Figure 5: Synthetic promoter constructs driving eGFP expression. Two distinct synthetic promoter variants (Psyn-GAP version1 and version2) were designed to drive the expression of the *enhanced green fluorescent protein* (*eGFP*) gene. The constructs are shown with a schematic representation of transcriptional direction (purple arrow). The fluorescence signal profiles below each construct illustrate the expression outputs resulting from each promoter variant. Illustration drawn using BioRender.com.

In order to create mutations on the Psyn-GAP-eGFP-version1 plasmid, primers named DA-26, DA-27, DA-28 and DA-29 were used by PCR with Q5 polymerase. In order to create mutations on the Psyn-GAP- eGFP-version2 plasmid, primers named DA-30, DA-31, DA-41 and DA-42 were used by PCR with Q5 polymerase. SYBR™ Safe DNA gel dye was added to a 1% agarose gel containing PCR results, and the gel was electrophoresed for 45 minutes at 120V. As directed by the manufacturer, the Nucleospin Gel and PCR Clean-up kit (Macherey-Nagel) was used to cut off the anticipated bands on the agarose gel for DNA extraction. Then, concentration was measured with Thermo Scientific NanoDrop™ 2000 Spectrophotometer. The obtained linear DNA fragments were assembled using the in-house Gibson Assembly (reaction mix protocol is provided in Appendix E, Table E8) to assemble from overhangs at a concentration of 100 ng. The Gibson assembly reaction was performed at 50°C for 1

hour. Following the Gibson assembly reaction, the mixture was transformed into 100 µl chemically competent *Escherichia coli* DH5α cells. (Cells were injected from stocks stored at -80°C to develop overnight in LB medium supplemented with the proper antibiotics in order to produce chemically competent *E. coli* DH5α cells. The cells were then diluted 1:100 and cultivated in LB media with antibiotics until their OD600 was between 0.4 and 0.6. After that, they incubated ten minutes on ice. The cells and supernatant were separated after 10 minutes at 4°C and 3000 x g centrifugation. The leftover cell pellet was reconstituted in Transformation and Storage (TSS) Buffer at a dilution of 1:10. For future transformations, 100 µL aliquots of chemically competent cell solution were stored at -80°C.) The mixture added to competent cells was kept on ice for 30 minutes. Then, incubated at 42 degrees for 45 seconds to apply heat shock and then kept on ice for 2 minutes. 250 µL LB medium was added to the cells waiting on ice. Then, incubated at 37 degrees for 30 minutes for recovery time. Incubated cells were plated on selective LB agar medium containing 100 µg/ml Zeocin™ and incubated overnight at 37 degrees. Cells that have grown under appropriate incubation conditions on selective LB agar medium containing 100 µg/ml Zeocin™ will be passaged to fresh selective medium, and strains containing *eGFP* genes will be screened using colony polymerase chain reaction with TAQ polymerase using DA-21 and DA-49 primers. 3 colonies will be selected for each strain from potential positive colonies and plasmids will be purified with purification kit (Thermo Scientific GeneJet Plasmid Miniprep Kit) according to the manufacturer's recommendations. Then, plasmids were directed to Sanger sequence verification (BM laboratuvar sistemleri, Ankara) of the expression cassette. Cell stocks of *E. coli* strains containing the confirmed recombinant plasmids were prepared and stored at −80 °C. Therefore, after this step; *P. pastoris* + wtPGAP-

eGFP, *P.pastoris* + Psyn-GAP-eGFP-version1, and *P.pastoris* + Psyn-GAP-eGFP-version2 strains were obtained.

2.2 Transformations of Constructed *E. coli* Plasmids to *P. pastoris*

In this thesis, *Pichia pastoris* CBS7435 strain (This strain was kindly requested from BOKU, Vienna), which has high-quality whole genome sequence information and more detailed annotation of genes, was used. During the transfection stage in *P. pastoris*, the random integration of the designed expression cassette into the cell genome may cause damage to the essential genes of the yeast, disrupting the vital activities of the cell or negatively affecting the production of recombinant proteins [45]. To ensure stable expression of *eGFP* genes in *P. pastoris* and to prevent changes in physiological properties among colonies due to random integration process, the designed recombinant plasmids were integrated into the AOX1 transcription terminator (AOX1tt) region of the *P. pastoris* genome by homologous recombination process. The linearization process of the recombinant plasmids before transfection was carried out by cutting the AOX1tt region on the plasmid with the BsiWI restriction enzyme from the middle (reaction mix protocol is provided in Appendix E, Table E6), so that gene fragments homologous to the AOX1tt region were present at the 5' and 3' ends of the linear expression cassette. Furthermore, it was noted that during the typical transformation procedure, endogenous gene segments from *E. coli* were incorporated into the genome of *P. pastoris* [45]. In order to clean the gene fragments that were cut incorrectly or that could be transmitted from *E. coli* during plasmid purification, the linearized plasmids were run on agarose gel (for 45 minutes at 120V), then the relevant band was cut from

the gel and purified with a gel extraction kit (the Nucleospin Gel and PCR Clean-up kit (Macherey-Nagel) according to the manufacturer's recommendations, and prepared for transfection. Gene transfer to *P. pastoris* was performed by chemical transformation using lithium chloride [46] (Figure 6); using YPD media and shaking at 200 rpm, a 50 ml culture of the *Pichia pastoris* CBS7435 strain was grown to an OD600 of 0.8 to 1.0 (around 108 cells/ml) at 30°C. After that, cells were harvested (centrifuged at 4500 × g for 5 min at room temperature), washed with 25 ml sterile water and centrifuged at 1500 × g for 10 min at room temperature. The water was decanted and the cells were resuspended in 1 ml of 100 mM LiCl. The cell suspension was transferred to a 1.5 ml microcentrifuge tube. Cells were centrifuged for 30 s at maximum speed and supernatant (LiCl) was removed with a pipette. Cells were resuspended in 400 µl of 100 mM LiCl. For each transformation, 50 µl of the cell suspension was decanted into a 1.5 ml microcentrifuge tube and used immediately. Since 3 transformations (1- wtPGAP-eGFP, 2- Psyn-GAP-eGFP-version1, 3- Psyn-GAP-eGFP-version2) will be made at this stage, it was divided into 3 tubes. While performing these operations, 10 mg/ml single-stranded DNA (Salmon Sperm DNA) was boiled at 99 degrees for 5 minutes. microcentrifuge tubes containing 50 µl of cell suspension were centrifuged at maximum speed for 30 seconds. For each transformation example, the following reagents were added to the cells in the order given. The importance of adding them in this order is that the PEG protects the cells from the damaging effects of high LiCl concentrations. 1-240 µl 50% PEG, 2- 36 µl 1 M LiCl, 3- 5 µl 10 mg/ml single-stranded DNA (boiled), 4- Plasmid DNA (1- wtPGAP-eGFP, 2- Psyn-GAP-eGFP-version1, 3- Psyn-GAP-eGFP-version2) (5-10 µg) in 70 µl sterile water. Each tube was vigorously vortexed until the cell pellet was completely mixed (~1 min). The tubes were incubated at 30°C for 30 minutes without shaking. After that, heat shock was applied for 20-25

minutes at 42°C. The tubes were centrifuged at 8000 rpm and the transformation solution was taken with a pipette. Pellets were resuspended in 1 ml YPD medium and at this stage, cells were transferred to 15 ml falcon tubes, then incubated at 30°C with shaking (200 rpm). After 4 hours, 50, 100, 200, 300 µl of cells were seeded onto YPD agar plates containing 100 µg/ml Zeocin™ and the plates were incubated at 30°C for 3-4 days. In order to select single colonies, the growing colonies were re-streaked into selective YPD agar plates containing 100 µg/ml Zeocin™ 3 times. Colony PCR was performed with TAQ polymerase enzyme to check whether the colonies contained plasmid DNA. For this, DA-21 and DA-49 primers that recognize the eGFP region were used. PCR products were run on agarose gel and colonies that gave positive results were grown in selective medium (YPD medium containing 100 µg/ml Zeocin™) and -80 stocks were taken.

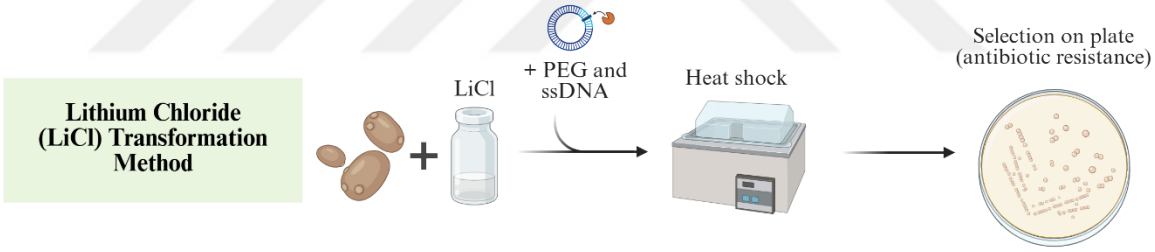

Figure 6: Overview of the Lithium Chloride (LiCl) yeast transformation method. The schematic illustrates the steps involved in the LiCl-based transformation of yeast cells. Yeast is first incubated with lithium chloride (LiCl), followed by the addition of plasmid DNA, polyethylene glycol (PEG), and single-stranded carrier DNA (ssDNA). The cells are then subjected to heat shock to facilitate DNA uptake. Finally, transformed cells are selected on antibiotic-containing plates based on resistance markers. Illustration drawn using BioRender.com.

2.3 Growth Media, Screening Conditions and Determination of Gene Copy Number

To determine the strength and regulation of the engineered Psyn-GAP promoter, different carbon sources (excess glucose, excess glycerol, ethanol and methanol) were used to compare how much eGFP produced compared to the wild type GAP promoter. To perform intracellular eGFP screening (Figure 7), yeast colonies were grown in 5 ml YP medium (containing μg/ml Zeocin™) for 24 hours at 30 degrees and 220 rpm. Afterwards, cells were cultured in ASMv6 medium with a final OD600 of 0.1 (For glucose and glycerol OD600 value should be 0.1, for ethanol and methanol it should be 1.0.). The recipe created for each carbon source is given in the Appendix E, Table E14, Table E15, Table E16, and Table E17. The cultured cells were inoculated at 30 degrees and 220 rpm for 24 hours. After 24 hours, cells were diluted in 1M PBS to a final OD600 of 0.4, then a flow cytometer was used for eGFP measurement. While making this comparison, *P. pastoris* CBS7435 strain was used as negative control, while *P. pastoris* + wtPGAP-eGFP was used as positive control.

Some of the yeast colonies containing version 1, version 2 and wild type GAP promoter were selected. gDNA isolate of these yeasts was made according to the kit (Macherey-Nagel). The *eGFP* gene copy number contained in the selected recombinant *P. pastoris* strains was determined by real-time quantitative PCR (qPCR) method. qPCR was performed with the isolated genomic DNAs using Luna® Universal qPCR Master Mix (NEB). While determining the *eGFP* gene copy number, *Arginine* gene (*ARG1*) was used as housekeeping gene and the *eGFP* gene copy number was calculated relative to

the *Arginine* gene copy number. *P. pastoris* strains carrying different *eGFP* gene copies were stored in glycerol stock at -80 °C.

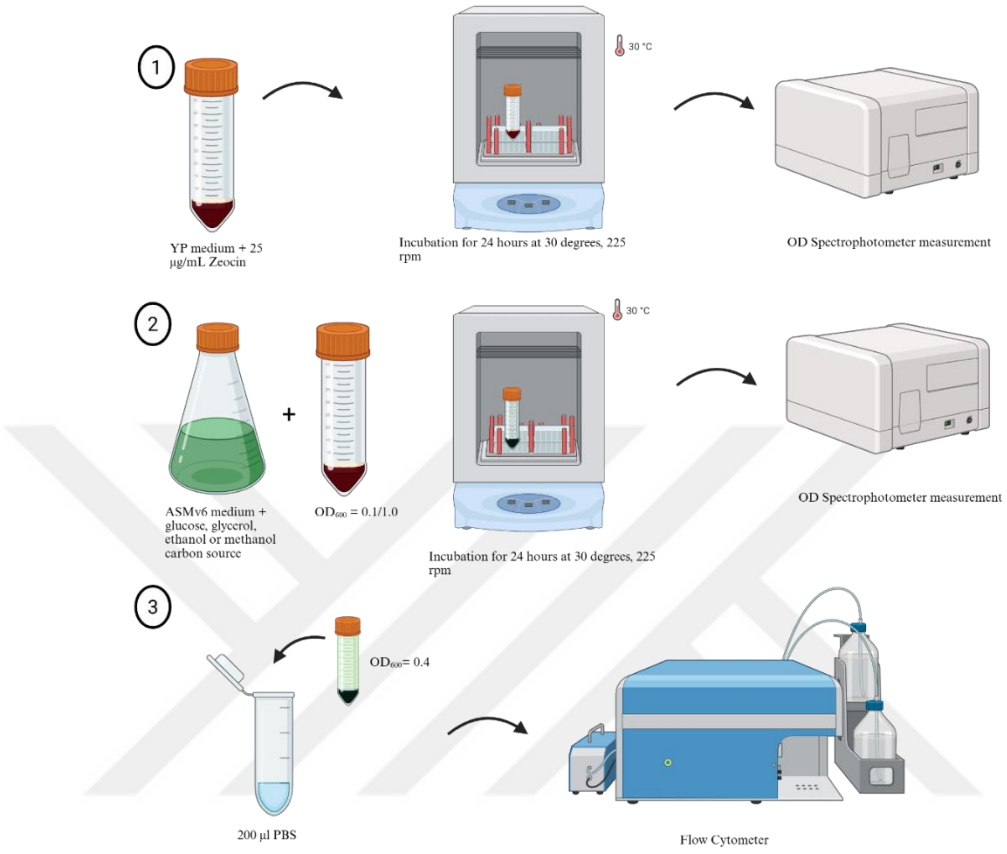

Figure 7: Schematic representation of eGFP screening and flow cytometry analysis performed in *Pichia pastoris*. Illustration drawn using BioRender.com.

2.4 Design and Cloning of CRISPR-Based Transcription Factors (CRISPR-TF: CRISPRa/i) into *E. coli*

A variety of modules were used to create the minimum sgRNAs utilized in prior studies. A 20-nucleotide variable DNA-targeting sequence and two structural RNA modules—the trans-activating tracrRNA domain and the dCas9 binding domain—are necessary

for proper structure and Cas9 binding [44]. Within the scope of the thesis, the CRISPR sgRNA sequence was modified by combining it with modules that will allow the binding of RNA-binding proteins. Thus, the sgRNA was converted into scaffold RNA (scRNA) that physically binds to DNA and associates the activities of incorporating proteins into the mechanism. It was observed that when VP64 was incorporated into the system via scRNA, it provided significantly more gene activation than when VP64 was directly combined with the dCas9 protein (dCas9-VP64) [29]. In CRISPRa/i designs, scRNAs modified by adding MS2 and Com aptamers were used to create combinations of different activator and repressor modules. MS2 and Com aptamers were combined with the corresponding MS2 coat protein (MCP) to create different activator modules (Figure 8) and Com protein to create different repressor modules (Figure 9).

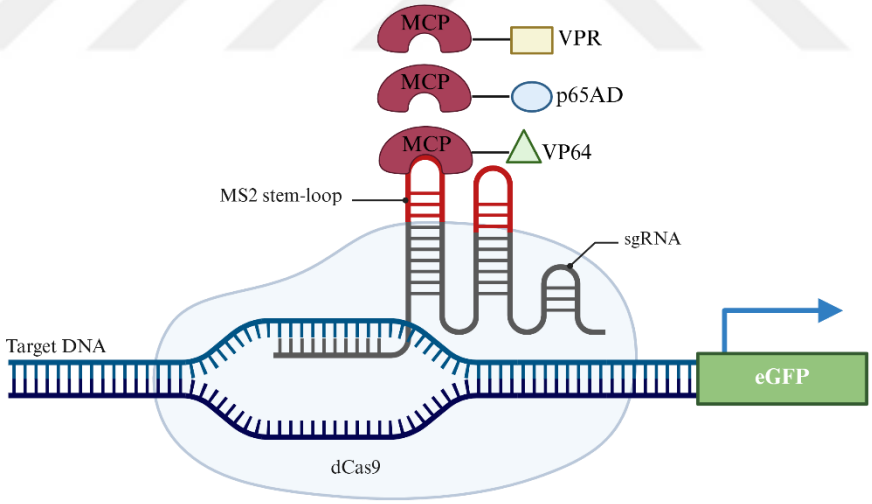

Figure 8: Schematic representation of a dCas9-MCP-based gene activation system.A catalytically inactive Cas9 (dCas9) is guided to the target DNA sequence by a single guide RNA (sgRNA). The sgRNA is engineered to include MS2 stem-loop structures that recruit MS2 coat proteins (MCP). MCP is fused to various transcriptional activators

(represented by colored shapes) to promote transcription of the downstream eGFP reporter gene. This system enables programmable and modular transcriptional activation at a specific genomic locus. Illustration drawn using BioRender.com.

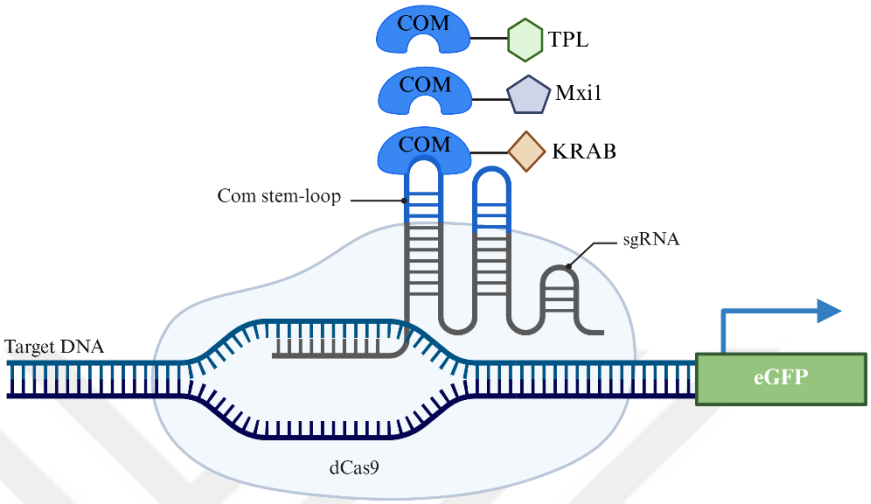

Figure 9: Schematic representation of a COM-based synthetic transcriptional activation system. A catalytically inactive Cas9 (dCas9) is guided to a specific genomic locus by a single guide RNA (sgRNA). The sgRNA is engineered to include multiple COM stem-loop motifs, which are recognized and bound by COM RNA-binding proteins. These COM proteins are fused to different transcriptional activators (represented by various colored shapes), allowing recruitment of multiple effectors to the target site. The assembly of this complex promotes transcriptional activation of a downstream eGFP reporter gene. Illustration drawn using BioRender.com.

When designing the scRNA structure, it was observed that the linker part between the sgRNA and recruitment hairpin domains being longer than 2 nucleotides weakened the effect of CRISPRa on gene expression [29]. In scRNA design, viral MS2 and Com RNA sequences were added to the 3’ end of the sgRNA with a 2-base linker in a hairpin

structure for binding of MS2 coat protein and Com protein. This approach allowed for multi-directional regulation (simultaneous activation and repression of different target genes within the same regulatory program). In the protein design, SV40 nuclear localization signal peptide (NLS) was added to the N'-termini to locate the activator and repressor domains in the nucleus. In the development of transcription activating CRISPRa systems, a nuclear localization signal (NLS) was added to the N-terminus of the MS2 coat protein, and transcription activation domains, namely VP64 [24], p65AD [45], and VP64-p65-Rta (VPR) [45], which were formed by the sequential addition of three activator domains, were added to the C-terminus. In the development of transcriptional repressor CRISPRi systems, NLS was added to the N-terminus of the Com protein, and the repressor protein domains Mxi1 [46], KRAB [46], and TPL [47] were added to the C-terminus, respectively. To investigate the effect of CRISPRa/i modules on *P. pastoris* transcription regulation, scRNA and dCas9 complexes were bound to different points on the promoter with sgRNA designs containing different PAM sequences. The regions targeted for binding of CRISPRa/i modules were determined by examining the promoter architecture. Binding sites of synthetic transcription factors were determined at 3 different positions on the core promoter containing the transcription start site and 3 different positions in the upstream activating sequence region located between -200 and -477 bp on the promoter, considering the promoter architecture, and sgRNA sequences were designed. Since transcription control performance increases when more than one sgRNA targets the same gene [48], it was decided to test the performance of sgRNA CRISPRa/i modules, which alone create significant changes, on transcription regulation with binary sgRNA vectors encoding two different protospacers.

In *P. pastoris* cells, the dCas9 was designed to be integrated into the *P. pastoris* genome from the RGI2 locus as a single expression cassette containing the PTEF1 promoter control, the activator or repressor modules under PGAP control, and the *Nurseotrichin* (*clonNat*) resistance gene under scRNA PTEF control. To construct the activator and repressor modules, the following plasmids were obtained from the non-profit organization Addgene: THI-VP64-pGAP-T1 (#161488), pAG414GPD-dCAS9-VPR (#63801), pJZC74 (#62342), pTDH3-dCAS9-Mxi (#46921), and pSumo-His6-Sumo-AtTPL(1-209) (#177856). Sequence information of the activator and repressor proteins used in the thesis is given in the Appendix A, Table A1.

First, the gRNA region in the THI-VP64-pGAP-T1 plasmid was replaced with a scrambled gRNA. To do this, DA-78 and DA-79 primers were used and the Gibson Assembly method was used as described in 2.1, followed by transformation into chemically competent *E. coli* DH5α cells. First of all, the reason for creating scrambled gRNA is to obtain negative control. A guide RNA sequence that has been purposefully changed to avoid targeting any particular location in the genome is known as a scrambled gRNA. In CRISPR research, it is frequently employed as a negative control to evaluate background effects and confirm that reported phenotypes are caused by particular gRNA activity [49]. Five colonies were randomly selected from the colonies formed and sent for Sanger sequencing (BM laboratuvar sistemleri, Ankara) to verify the gRNA sequence. -80 stocks of colonies that gave positive results were taken. This construct was named THI-VP64-scrgRNA. Later, in order to replace VP64 in this plasmid with other activator protein sequences, the VP64 sequence was removed using PCR with Q5 polymerase using DA-72 and DA-88 primers. The resulting PCR product was named THI-scrgRNA. The p65AD protein sequence was obtained by PCR with Q5 polymerase using primers DA-73 and DA-74 from the pAG414GPD-dCas9-VPR

plasmid, which could be added to the THI-scrgRNA PCR product by the Gibson Assembly method. Similarly, DA-76 and DA-77 primers were used to obtain the VPR protein sequence. The obtained PCR products (for p65AD and VPR) were inserted into plasmids by the Gibson Assembly method, with the THI-scrgRNA PCR product as the backbone. Then, the plasmids were transformed into chemically competent *E. coli* DH5α cells as described in 2.1. Five colonies were randomly selected from the colonies formed and sent for NGS sequencing (Intergen, Ankara) to verify the gRNA sequence. -80 stocks of colonies that gave positive results were taken. As a result, THI-p65AD-scrgRNA and THI-VPR-scrgRNA plasmids were created.

The THI-VP64-scrgRNA plasmid was used to construct the above-mentioned version 1 and version 2 gRNAs. Primer numbers used to generate these gRNAs are as given in the Table 1.

Table 1: Primer pairs for different gRNAs Primer pairs for different gRNAs.

	gRNAs	Primer pairs
Version 1	gRNA-1	DA-45 & DA-46
	gRNA-2	DA-50 & DA-51
	gRNA-3	DA-52 & DA-53
	gRNA-4	DA-54 & DA-55
	gRNA-5	DA-56 & DA-57
Version 2	gRNA-1	DA-58 & DA-59
	gRNA-2	DA-60 & DA-61
	gRNA-3	DA-62 & DA-63
	gRNA-4	DA-65 & DA-71
	gRNA-5	DA-66 & DA-67

While the THI-VP64-scrgRNA plasmid was used as a template, PCR with Q5 polymerase was performed using the primers in Table 1 to generate each gRNA. Then, Gibson Assembly method was used as described in 2.1, followed by transformation into chemically competent *E. coli* DH5α cells. Five colonies were randomly selected from the colonies formed and sent for Sanger sequencing (BM laboratuvar sistemleri, Ankara) to verify the gRNA sequence. -80 stocks of colonies that gave positive results were taken.

To construct repressor cassettes, MS2 and VP64 sequences from the THI-VP64-scrgRNA plasmid were excised by PCR with Q5 polymerase using primers DA-80 and DA-88. In order to add the COM-KRAB sequence to the resulting PCR product (THI-w/out MS2+VP64-scrgRNA), the COM-KRAB sequence was first obtained from the pJZC74 plasmid using primers DA-81 and DA-82, and then these two products were combined with Gibson assembly method was used as described in 2.1, followed by transformation into chemically competent *E. coli* DH5α cells. Five colonies were randomly selected from the colonies formed and sent for NGS sequencing (Intergen, Ankara) to verify the sequence. -80 stocks of colonies that gave positive results were taken. Later, to replace the MS2 stem loop sequence in the THI-COM-KRAB-scrgRNA plasmid with the COM stem loop sequence, firstly the MS2 stem loop was removed using PCR with Q5 polymerase using DA-89 and DA-90 primers, then the COM step loop sequence was added using PCR with Q5 polymerase using this PCR product and DA-91 and DA-92 primers. Then, Gibson Assembly method was used as described in 2.1, followed by transformation into chemically competent *E. coli* DH5α cells. Five colonies were randomly selected from the colonies formed and sent for NGS

sequencing (Intergen, Ankara) to verify the sequence. -80 stocks of colonies that gave positive results were taken. As a result, the THI-COM-KRAB-COM stem loop-scrgRNA plasmid was obtained. In order to be able to add other receptor modules, the KRAB sequence from this plasmid was extracted using PCR with Q5 using primers DA-83 and DA-88. To obtain the Mxi1 protein sequence, the pTDH3-dCas9-Mxi plasmid was used and the DA-84 and DA-85 primers were used to obtain the Mxi1 sequence using PCR with Q5 polymerase. Then, Gibson Assembly method was used as described in 2.1, followed by transformation into chemically competent *E. coli* DH5α cells. Five colonies were randomly selected from the colonies formed and sent for NGS sequencing (Intergen, Ankara) to verify the gRNA sequence. Moreover, to obtain the TPL protein sequence, the pSumo-His6-Sumo-AtTPL(1-209) plasmid was used and the DA-86 and DA-87 primers were used to obtain the TPL sequence using PCR with Q5 polymerase. Then, Gibson Assembly method was used as described in 2.1, followed by transformation into chemically competent *E. coli* DH5α cells. Five colonies were randomly selected from the colonies formed and sent for NGS sequencing (Intergen, Ankara) to verify the gRNA sequence. After that, in order to obtain gRNAs for KRAB repressor cassettes, THI-COM-KRAB-COM stem loop-scrgRNA plasmid was used as a template, PCR with Q5 polymerase was performed using the primers in Table 1 to generate each gRNA. Then, Gibson Assembly method was used as described in 2.1, followed by transformation into chemically competent *E. coli* DH5α cells. Five colonies were randomly selected from the colonies formed and sent for Sanger sequencing (BM laboratuvar sistemleri, Ankara) to verify the gRNA sequence. -80 stocks of colonies that gave positive results were taken.

2.5 Transformations of Constructed CRISPR-TF- *E. coli* Plasmids to *P. pastoris*

Vectors carrying CRISPRa/i modules were integrated into the genome of *P. pastoris* cells (cells with a known gene copy number of 1) carrying the Psyn-GAP-version1, Psyn-GAP-version2 cassettes from the RGI2 locus. scrgRNA was integrated into both versions, while version 1 gRNAs were integrated into *Pichia pastoris* carrying the version 1 plasmid, and version 2 gRNAs were integrated into *Pichia pastoris* carrying the version 2 plasmid. In order to integrate from the RGI2 locus, these vectors were linearized using the AscI restriction enzyme. In order to clean the gene fragments that were cut incorrectly or that could be transmitted from *E. coli* during plasmid purification, the linearized plasmids were run on agarose gel (for 45 minutes at 120V), then the relevant band was cut from the gel and purified with a gel extraction kit (the Nucleospin Gel and PCR Clean-up kit (Macherey-Nagel) according to the manufacturer's recommendations, and prepared for transfection. The linearized vectors were transferred using the LiCl method as described in 2.2. Since the new strain to be formed contained both *Zeocin* and *Nurseotrichin* resistance genes, cells were plated on YPD Agar plates containing both (100 µg/ml Zeocin™ and 25 µg/ml Nurseotrichin) and 3 rounds of re-streak were performed in order to select single colonies. Cells with both resistance genes were grown and minus 80 stocks were taken from these cells.

2.6 Growth Media, Screening Conditions and Determination of Gene Copy Number for Constructed CRISPR-TF- *P. pastoris* plasmids

To determine how much the designed gRNAs altered the strength and regulation of the Psyn-GAP promoter with activator and repressor modules, we compared how much eGFP was produced using different carbon sources (excess glucose, excess glycerol, ethanol and methanol) compared to scrgRNA containing modules. Growth medium, screening conditions and determination of gene copy number were performed as described in 2.3. The gRNAs that gave the best results were determined as a result of the change in the amount of eGFP produced according to scrgRNA.

For each gRNA, plasmids containing VP64 in the activator module were screened. The gRNAs that gave the best results as a result of this screening were created with the appropriate primers mentioned before instead of scrgRNA to *E. coli* plasmids containing the other activator protein (p65AD).

CHAPTER III: RESULTS AND DISCUSSION

3.1 Cloning of Psyn-GAP-eGFP

To create the synthetic GAP promoter, mutation sites were designed as described in 2.1. The mutations made are as shown in Figure 10 and 11.

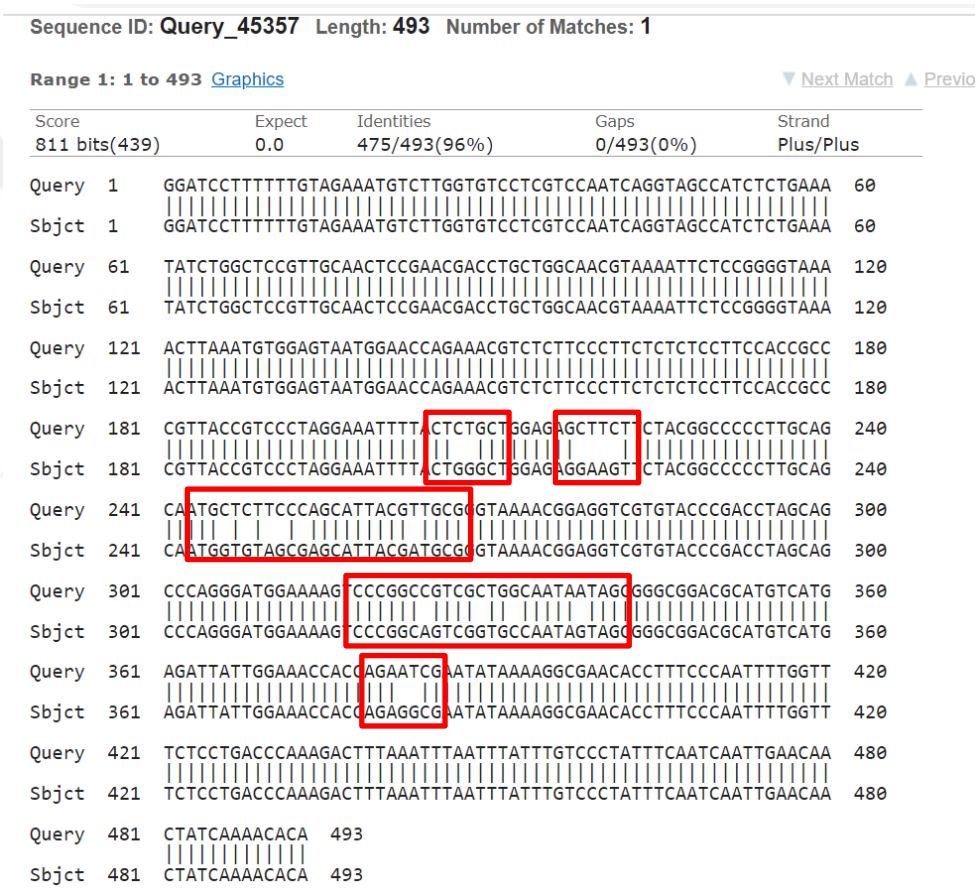

Figure 10: Sequence alignment of the synthetic GAP promoter version 1 (Query) with the wild-type GAP promoter (wild-type sequence) using BLAST. The synthetic Psyn-GAP version 1 promoter sequence (Query) was aligned with the wild-type *Pichia pastoris* GAP promoter (wild-type sequence) using NCBI BLAST. The alignment showed 96% identity over 493 base pairs, with no insertions or deletions. The observed

mismatches correspond to rationally introduced mutations in the promoter sequence, which were designed to create new CRISPRa target sites without significantly altering the core promoter function. These modifications served as the foundation for testing transcriptional modulation via synthetic gRNAs.

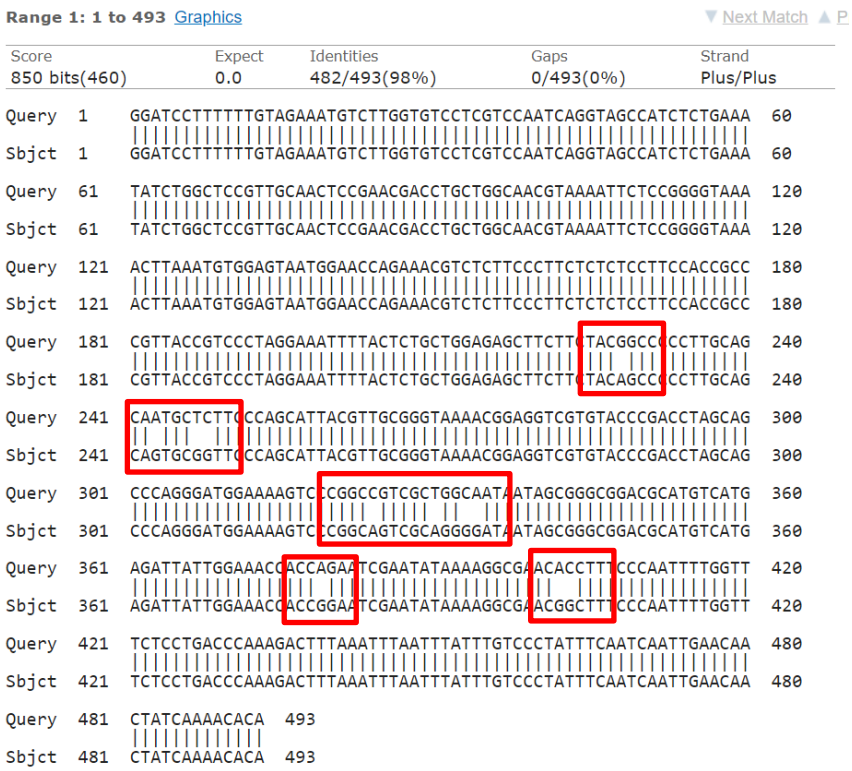

Figure 11: Sequence alignment of the synthetic GAP promoter version 2 (Query) with the wild-type GAP promoter (wild-type sequence) using BLAST. The synthetic Psyn-GAP version 2 promoter sequence (Query) was aligned against the native *Pichia pastoris* GAP promoter (wild-type sequence) using NCBI BLAST. The alignment revealed 98% identity over a 493 bp region with no insertions or deletions, indicating high sequence conservation. The observed base mismatches represent intentional modifications introduced during the design of version 2 to create novel gRNA binding

sites for CRISPRa targeting. These alterations aim to modulate promoter activity without compromising its core functionality.

Then, pGAPZα-B plasmid cut with EcoRI and XbaI to remove MCS. Resulting in an expected band of 2819 bases, as shown in Figure 12.

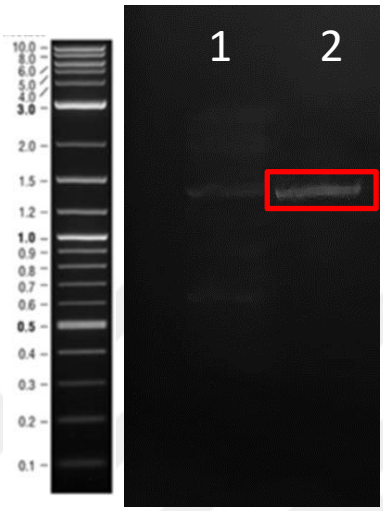

Figure 12: Double digestion of pGAPZα-B. Agarose gel electrophoresis result of double digestion of pGAPZα-B plasmid (line 2). Lane 1 shows the DNA marker (NEB).

The eGFP region from pCat8-L2-eGFP plasmid was amplified by PCR, resulting in an expected band of 790 bases, as shown in Figure 13.

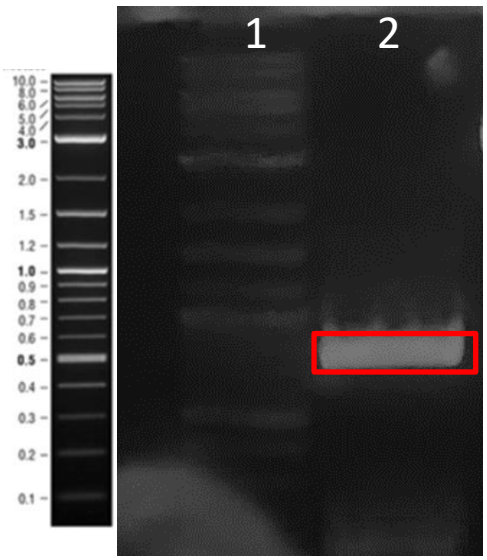

Figure 13: eGFP amplification. Agarose gel electrophoresis result of PCR-amplified 790 bp (lane 2) from the pCat8-L2-eGFP plasmid. Lane 1 shows the DNA marker (NEB). The eGFP gene was amplified to have overhang regions suitable for further Gibson assembly reaction.

A Gibson assembly reaction was performed with the digested backbone and the PCR-amplified eGFP insert at a 1:3 molar ratio. The plasmid map of the resulting construct is provided in Appendix C, Figure C2. And colony PCR result of this plasmid shown in Figure 14.

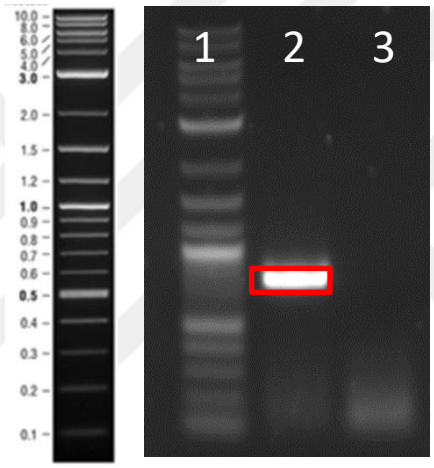

Figure 14: Colony PCR result of wtGAP-eGFP plasmid. Line 2 shows positive colony result with the expected 790 bp while line 3 shows negative colony result. Lane 1 shows the DNA marker (NEB).

Next, PCR was performed to generate mutation sites in the Psyn-GAP-eGFP-version1 and Psyn-GAP-eGFP-version2 plasmids, resulting in an expected band of 137 and 170 bases, respectively, as shown in Figure 15.

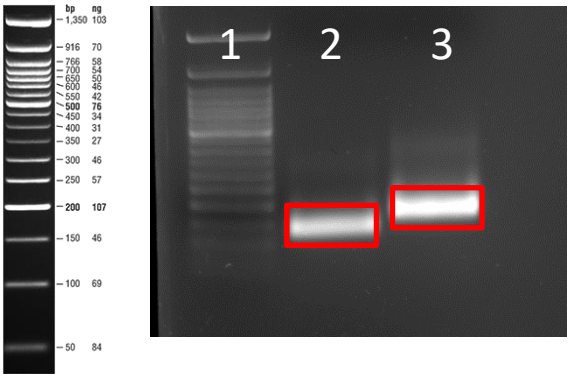

Figure 15: PCR for synthetic GAP promoters. Agarose gel electrophoresis result of 137 bp (lane 2) and 170 bp (lane 3). Lane 1 shows the DNA marker (NEB).

After that, a Gibson assembly reaction was performed. The plasmid map of the resulting constructs is provided in Appendix C, Figure C3 and C4. And colony PCR results of Psyn-GAP-eGFP-version1 and Psyn-GAP-eGFP-version2 plasmids shown in Figure 16.

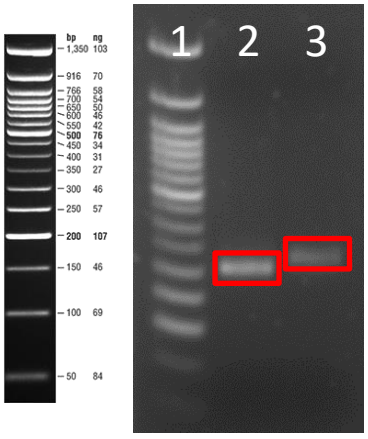

Figure 16: Colony PCR result of Psyn-GAP-eGFP-version1 and Psyn-GAP-eGFP-version2 plasmids. Line 2 shows positive colony result for version1 with the expected 311 bp while line 3 shows positive colony result for version 2 with the expected 328 bp. Lane 1 shows the DNA marker (NEB).

For linearization, wtGAP-eGFP, Psyn-GAP-eGFP-Version1 and Psyn-GAP-eGFP-Version2 plasmids were digested with BsiWI, and the expected band of 3545 bp, corresponding to the digested plasmids, is shown in Figure 17.

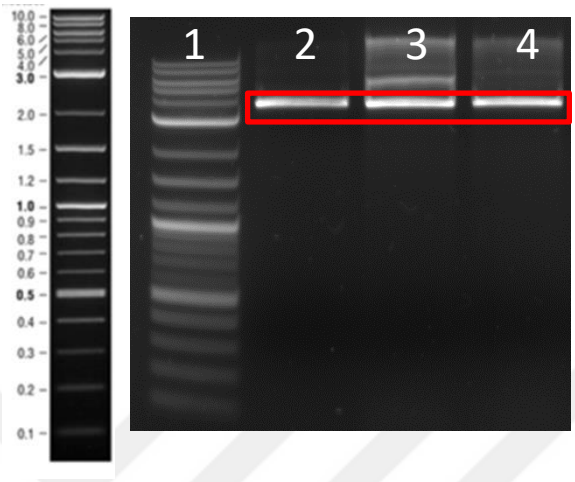

Figure 17: Single digestion of wtGAP-eGFP, Psyn-GAP-eGFP-Version1 and Psyn-GAP-eGFP-Version2 plasmids. Agarose gel electrophoresis result showing the wtGAP-eGFP, Psyn-GAP-eGFP-Version1 and Psyn-GAP-eGFP-Version2 plasmids digested with BsiWI (lane 2,3,4 respectively), resulting in a linear fragment of 3545 bases. Lane 1 shows the DNA marker (NEB).

3.2 Transformation of *Pichia pastoris*

The linearized plasmids were transformed into *Pichia pastoris* CBS7435 strain as described in 2.2. And colony PCR results of wtGAP-eGFP shown in Figure 18, Psyn-GAP-eGFP-Version1 shown in Figure 19, and Psyn-GAP-eGFP-Version2 shown in Figure 20.

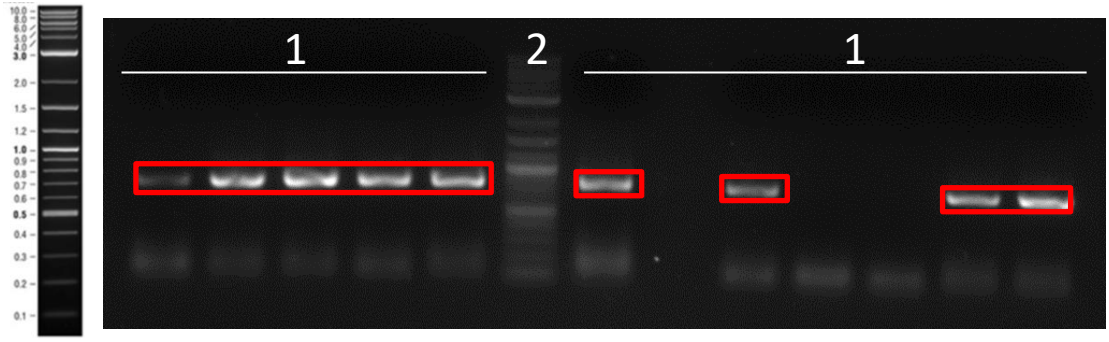

Figure 18: *Pichia pastoris* + wtGAP-eGFP plasmid. Number 1 shows positive colonies with the expected 790 bp. Number 2 shows the DNA marker (NEB).

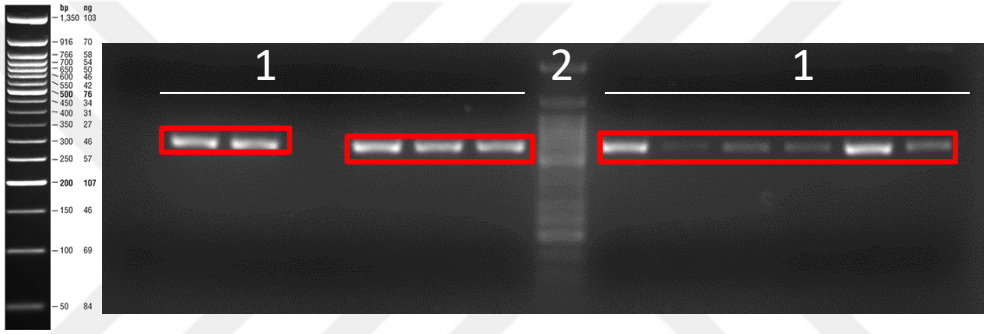

Figure 19: *Pichia pastoris* + Psyn-GAP-eGFP-Version1 plasmid. Number 1 shows positive colonies with the expected 563 bp. Number 2 shows the DNA marker (NEB).

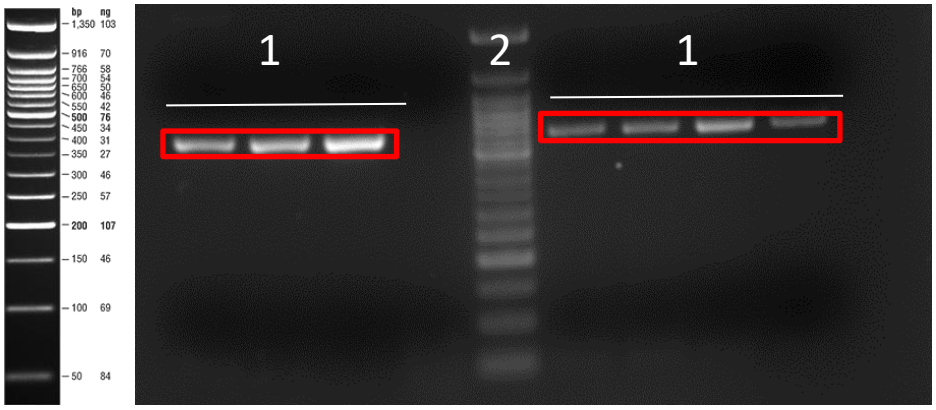

Figure 20: *Pichia pastoris* + Psyn-GAP-eGFP-Version2 plasmid. Number 1 shows positive colonies with the expected 563 bp. Number 2 shows the DNA marker (NEB).

3.3 Evaluation of eGFP Expression by Flow Cytometry

To evaluate the strength and regulatory properties of the modified Psyn-GAP promoter, different carbon sources—including excess glucose, excess glycerol, ethanol, and methanol—were used to compare eGFP expression levels relative to the wild-type GAP promoter. A colony screening experiment was conducted as described in Section 2.3. The results, presented in Figure 21, illustrate the percentage of eGFP production by various colonies under each carbon condition.

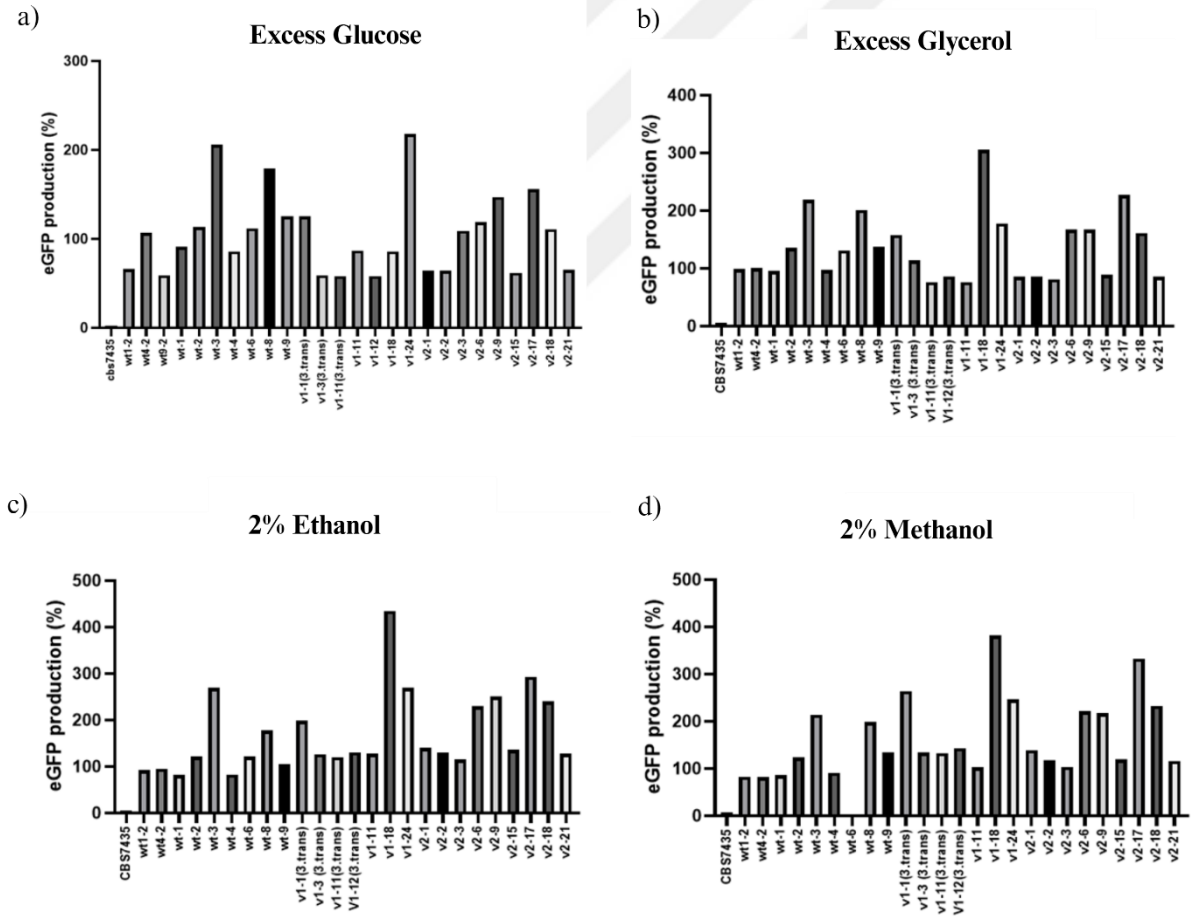

Figure 21: Screening of eGFP production under various carbon sources in *P. pastoris* transformants harboring wild-type or synthetic GAP promoters. Bar graphs represent the percentage of eGFP production in individual colonies cultivated under (a) excess

glucose, (b) excess glycerol, (c) 2% ethanol, and (d) 2% methanol conditions. Each bar corresponds to a different colony containing either the wild-type GAP promoter (wt), synthetic GAP version 1 or version 2 constructs (v1, v2), or re-transformed clones. CBS7435 was used as the negative control. The data reflect relative eGFP fluorescence normalized to the colony with the highest expression in each condition.

During data analysis, singlet yeast populations were identified from the flow cytometry data by selecting cells exhibiting a linear relationship between FSC-H and FSC-A within the population gated on the FSC-A vs. SSC-A plot. This ensured accurate discrimination of single cells for further quantification. The amount of eGFP produced per unit cell volume was calculated following the method described by Hohenblum, Borth, and Mattanovich (2003) [50]. Specifically, the geometric means of eGFP fluorescence and FSC-H values from the gated population were used to estimate the specific eGFP synthesis levels. For each measurement, fluorescence signals from 100,000 individual cells were analyzed.

As shown in Figure 21, most colonies exhibited eGFP production levels in the range of approximately 60–80%. However, a subset of colonies displayed markedly higher expression levels, such as v1-1 (3.trans) with 144% and v1-18 with 298% relative production. These unusually high expression levels were hypothesized to result from multiple gene integration events during transformation. To validate this, quantitative PCR (qPCR) was performed on selected colonies to determine gene copy numbers. Quantitative real-time PCR experiment was performed with selected positive samples to determine the *P. pastoris* strain that was integrated into the genome in the correct

position and contained a single copy Psyn-GAP-eGFP-version1 and Psyn-GAP-eGFP-version2 cassettes in the genome. qPCR results were as shown in Figure 22.

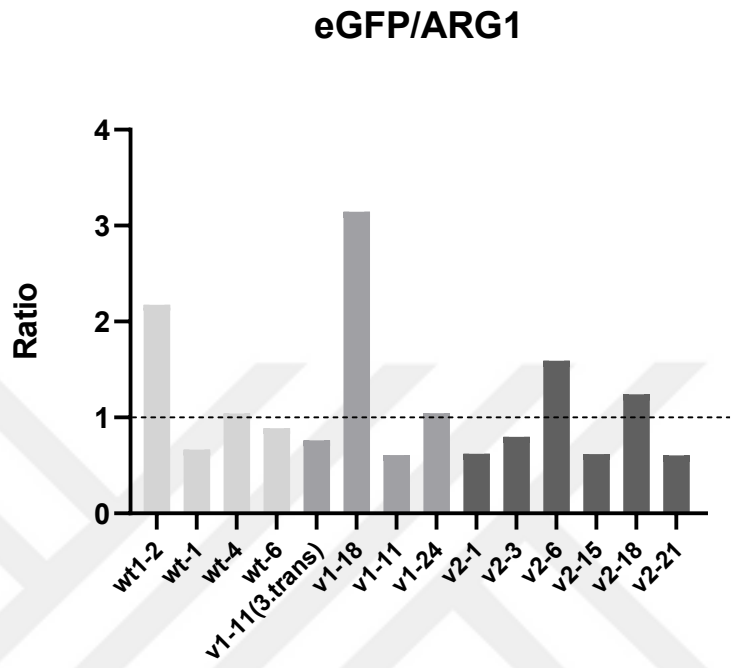

Figure 22: Determination of gene copy number in selected *P. pastoris* transformants using quantitative real-time PCR. The eGFP/ARG1 ratio was calculated for colonies harboring either wild-type PGAP-eGFP or synthetic Psyn-GAP-eGFP (version 1 or version 2) constructs. *ARG1* was used as a single-copy reference gene. Colonies with an approximate ratio of 1 were considered to possess a single integrated copy of the eGFP construct. Higher ratios (e.g., v1-18) indicated multiple gene integrations. This analysis was performed to identify suitable single-copy colonies for subsequent CRISPRa experiments.

As shown in Figure 22, the eGFP/ARG1 ratio varied among the analyzed colonies, indicating differences in gene copy number. Colonies wt-1 and wt-6, which harbor the *P. pastoris* + wtPGAP-eGFP construct, exhibited a ratio close to 1, suggesting single-

copy integration. Similarly, colonies v1-11 (3.trans), v1-11, v2-1, v2-3, v2-15, and v2-21, all carrying Psyn-GAP-eGFP version 1 or version 2 constructs, were also determined to contain a single gene copy. In contrast, several other colonies demonstrated elevated eGFP/ARG1 ratios, implying multiple gene integration events. Notably, v1-18 displayed a ratio around 3, which correlates with its exceptionally high eGFP expression observed in the screening experiment. These findings confirm that gene copy number significantly influences expression levels and must be considered when evaluating promoter activity. Importantly, the results also showed that the introduced mutations in the wild-type GAP promoter did not markedly weaken its transcriptional strength. Colonies v1-11 (version 1) and v2-3 (version 2), which were confirmed to harbor single-copy integrations, were selected for downstream experiments to assess the effect of synthetic gRNA-mediated regulation.

3.4 Cloning of CRISPR-TF: CRISPRa/i

Our aim in the thesis is to turn off the transcription of the promoter with the CRISPRi system and to increase its transcription with the CRIPRa system. It is known that the wild type GAP promoter generally exhibits a constitutive expression pattern, with some variation with different carbon sources. However, within the scope of the thesis, it is aimed to design a regulated transcription system by developing a system that produces more expression with CRISPRa and whose expression is stopped with CRISPRi. The reason for choosing constitutive PGAP as the target is that the effectiveness of the CRISPRi system can be investigated under different conditions (not only in the activation state). Synthetic GAP promoters were created as mentioned above so that the 10 different gRNAs could bind to the GAP promoter and change its strength. The reason

why only the activator protein is not used to increase the promoter strength here is that when VP64 is added to the system via scRNA, significantly more gene activation is observed compared to the direct combination of VP64 and dCas9 protein (dCas9-VP64) [29]. For this reason, activator and repressor modules were created on *E. coli* as stated in 2.4. First of all, PCR was performed to generate THI-VP64-scrgRNA plasmid, resulting in an expected band of 10847 bp as shown in Figure 23.

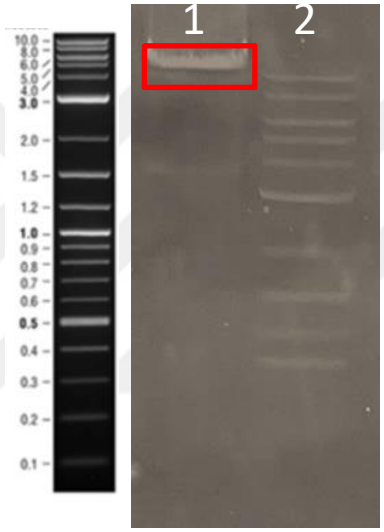

Figure 23: THI-VP64-scrgRNA plasmid. Agarose gel electrophoresis result showing the THI-VP64-scrgRNA plasmid with the expected 10847 bp (line 1). Line 2 shows the DNA marker (NEB).

Next, PCR was performed for other gRNAs for VP64 module (Figure 24, 25 and 26).

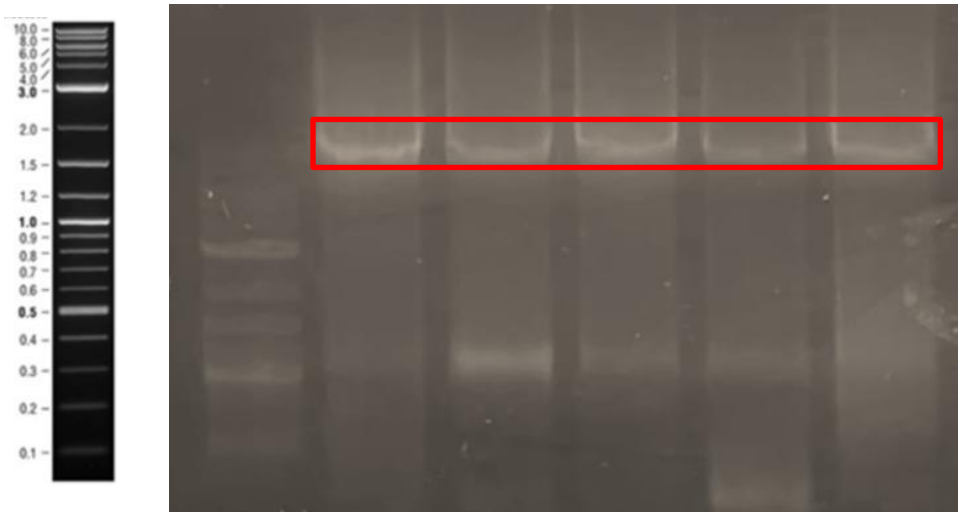

Figure 24: THI-VP64- ver1- gRNA1, 2, 3, 4 and 5. Agarose gel electrophoresis result showing the THI-VP64-ver1-gRNA1 plasmid (line 2), THI-VP64-ver1-gRNA2 plasmid (line 3), THI-VP64-ver1-gRNA3 plasmid (line 4), THI-VP64-ver1-gRNA4 plasmid (line 5), THI-VP64-ver1-gRNA5 plasmid (line 6), with the expected 10850 bp. Line 2 shows the DNA marker (NEB).

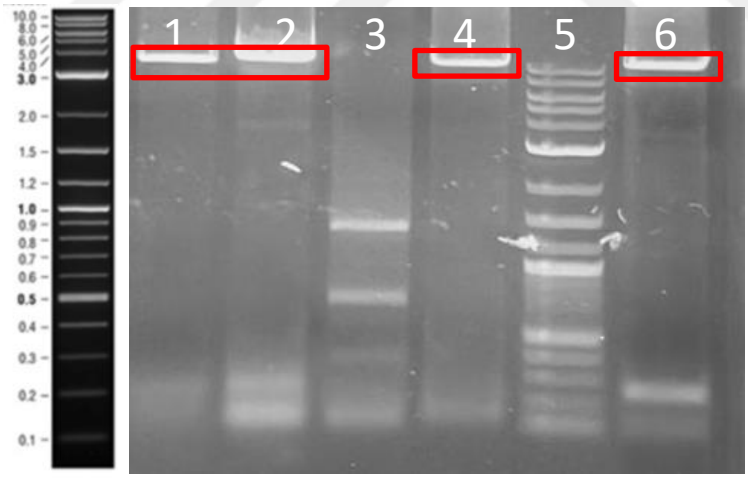

Figure 25: THI-VP64- ver2- gRNA1, 2, 4 and 5. Agarose gel electrophoresis result showing the THI-VP64-ver2-gRNA1 plasmid (line1), THI-VP64-ver2-gRNA2 plasmid (line 2), THI-VP64-ver2-gRNA4 plasmid (line 4), THI-VP64-ver2-gRNA5 plasmid (line 6), with the expected 10850 bp. Line 5 shows the DNA marker (NEB).

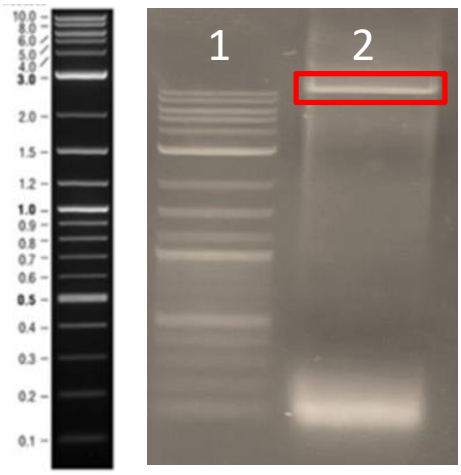

Figure 26: THI-VP64- ver2- gRNA3. Agarose gel electrophoresis result showing the THI-VP64-ver2-gRNA3 plasmid (line2), with the expected 10850 bp. Line 1 shows the DNA marker (NEB).

After that, PCR was performed to generate THI-VPR-scrgRNA plasmid and THI-p65AD-scrgRNA plasmid, resulting in an expected band of 12260 bp and 11039 bp, respectively as shown in Figure 27.

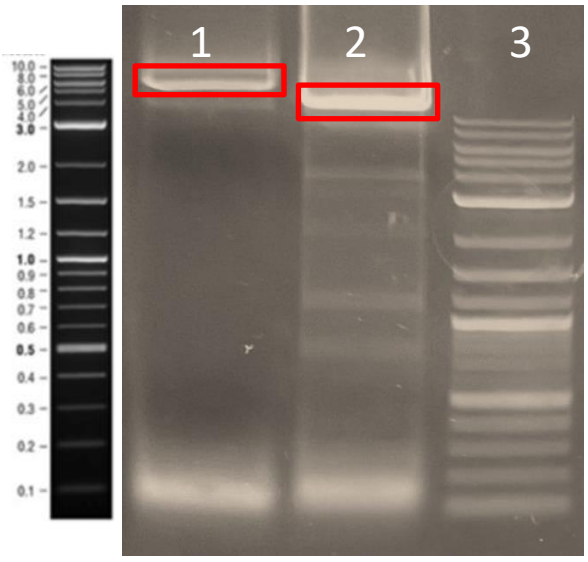

Figure 27: THI-VPR-scrgRNA and THI-p65AD-scrgRNA plasmid.Agarose gel electrophoresis result showing the THI-VPR-scrgRNA and THI-p65AD-scrgRNA

plasmids with the expected 12260 bp (line 1) and 11039 bp (line 2). Line 3 shows the DNA marker (NEB).

Next, a Gibson assembly reaction was performed. The plasmid map of the resulting constructs is provided in Appendix C, Figure C5, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19, C7 and C6. And colony PCR results of THI-VP64-scrgRNA, THI-VP64- ver1- gRNA1, 2, 3, 4 and 5, and THI-VP64- ver2- gRNA1, 2, 3, 4 and 5, and THI-VPR-scrgRNA and THI-p65AD-scrgRNA plasmids shown in Figure 28, 29 and 30.

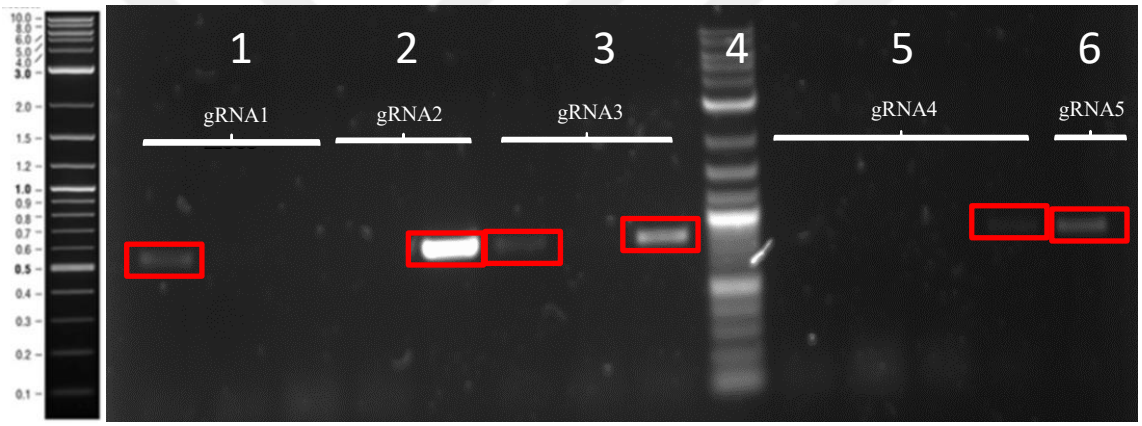

Figure 28: Colony PCR result of THI-VP64- ver1- gRNA1, 2, 3, 4 and 5. Line 1 shows positive colony result for gRNA1, line 2 shows positive colony result for gRNA2, line 3 shows positive colony result for gRNA3, line 5 shows positive colony result for gRNA4, line 6 shows positive colony result for gRNA5, with the expected 846 bp while Lane 4 shows the DNA marker (NEB).

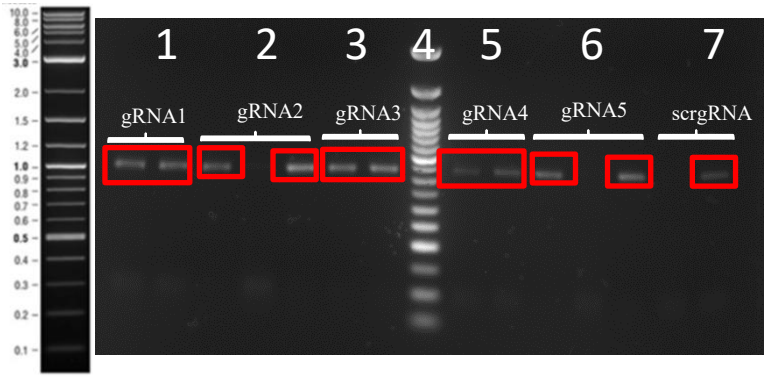

Figure 29: Colony PCR result of THI-VP64- ver2- gRNA1, 2, 3, 4, 5, and THI-VP64-scrgRNA. Line 1 shows positive colony result for gRNA1, line 2 shows positive colony result for gRNA2, line 3 shows positive colony result for gRNA3, line 5 shows positive colony result for gRNA4, line 6 shows positive colony result for gRNA5, line 7 positive colony result for scrgRNA, with the expected 846 bp while Lane 4 shows the DNA marker (NEB).

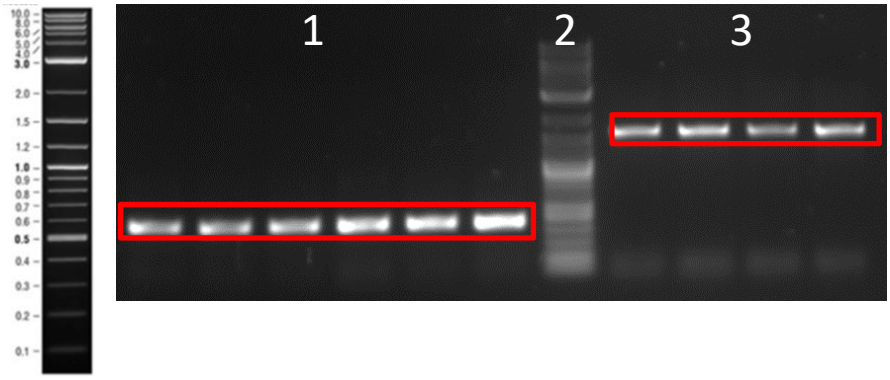

Figure 30: Colony PCR result of THI-p65AD-scrgRNA and THI-VPR-scrgRNA. Line 1 shows positive colony result for p65AD, with the expected 418 bp, line 3 shows positive colony result for VPR, with the expected 1693 bp. Lane 2 shows the DNA marker (NEB).

Next, PCR was performed to generate THI-w/out MS2 + VP64-scrgRNA plasmid, resulting in an expected band of 10252 bp as shown in Figure 31.

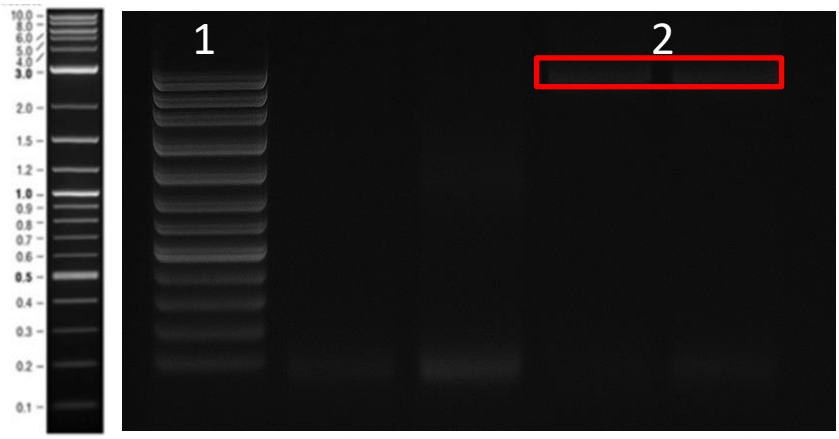

Figure 31: PCR result of THI-w/out MS2 + VP64-scrgRNA plasmid. Agarose gel electrophoresis result showing the THI-w/out MS2 + VP64-scrgRNA plasmid with the expected 10252 bp (line 2). Lane 1 shows the DNA marker (NEB).

Then, using this plasmid, THI-COM-KRAB-Com stem loop-scrgRNA plasmid was created, resulting in an expected band of 10729 bp as shown in Figure 32.

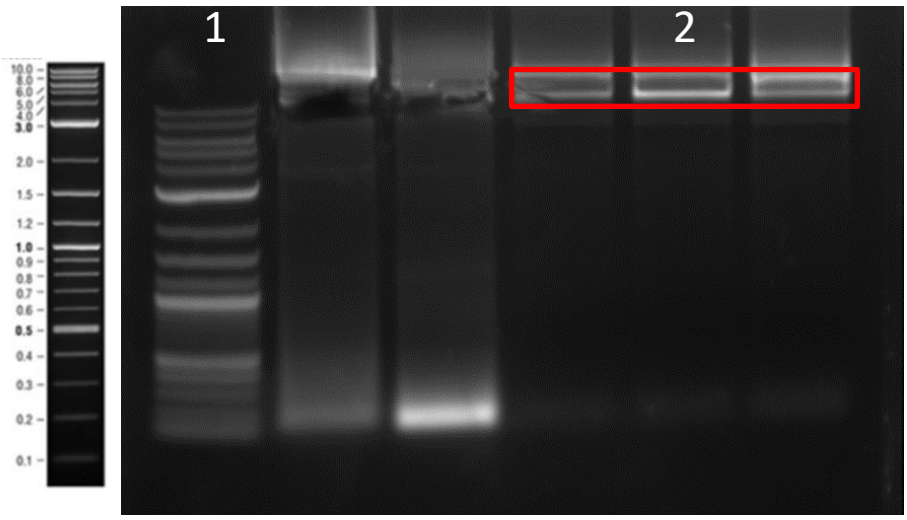

Figure 32: PCR result of THI-COM-KRAB-Com stem loop-scrgRNA plasmid.Agarose gel electrophoresis result showing the THI-COM-KRAB-Com stem loop-scrgRNA plasmid with the expected 10729 bp (line 2). Lane 1 shows the DNA marker (NEB).

Next, PCR was performed for other gRNAs for KRAB module (Figure 33).

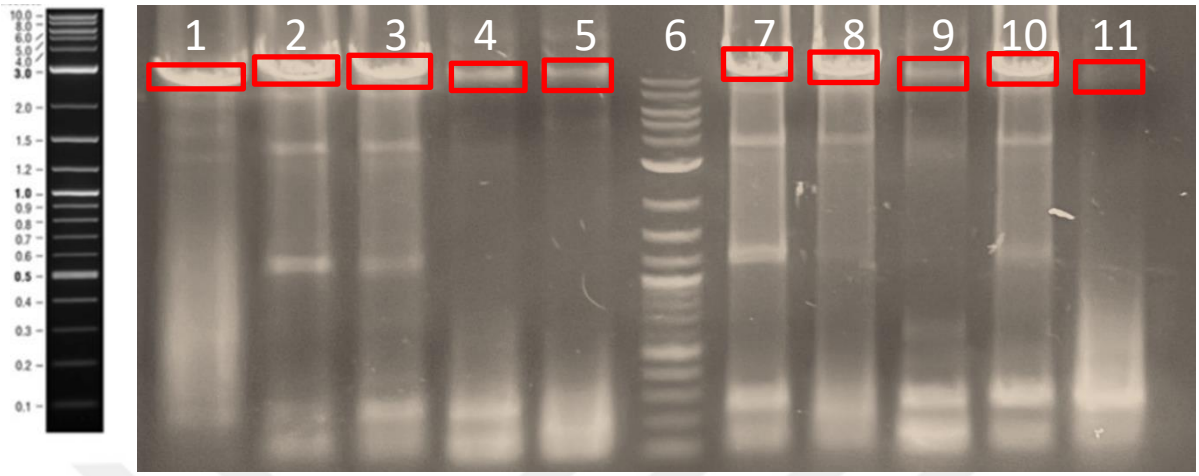

Figure 33: THI-COM-KRAB-VER1-gRNA1(2,3,4,5) + Com stem loop and THI-COM-KRAB-VER2-gRNA1(2,3,4,5) + Com stem loop. Agarose gel electrophoresis result showing the THI-COM-KRAB-VER1-gRNA1+Com stem loop (line1), THI-COM-KRAB-VER1-gRNA2+Com stem loop (line2), THI-COM-KRAB-VER1-gRNA3+Com stem loop (line3), THI-COM-KRAB-VER1-gRNA4+Com stem loop (line4), THI-COM-KRAB-VER1-gRNA5+Com stem loop (line5), THI-COM-KRAB-VER2-gRNA1+Com stem loop (line7), THI-COM-KRAB-VER2-gRNA2+Com stem loop (line8), THI-COM-KRAB-VER2-gRNA3+Com stem loop (line9), THI-COM-KRAB-VER2-gRNA4+Com stem loop (line10), THI-COM-KRAB-VER2-gRNA5+Com stem loop (line11), with the expected 10732 bp. Line 6 shows the DNA marker (NEB).

Next, a Gibson assembly reaction was performed. The plasmid map of the resulting constructs is provided in Appendix C, Figure C9. And colony PCR results of THI-COM-KRAB-Com stem loop-scrgRNA plasmid shown in Figure 34.

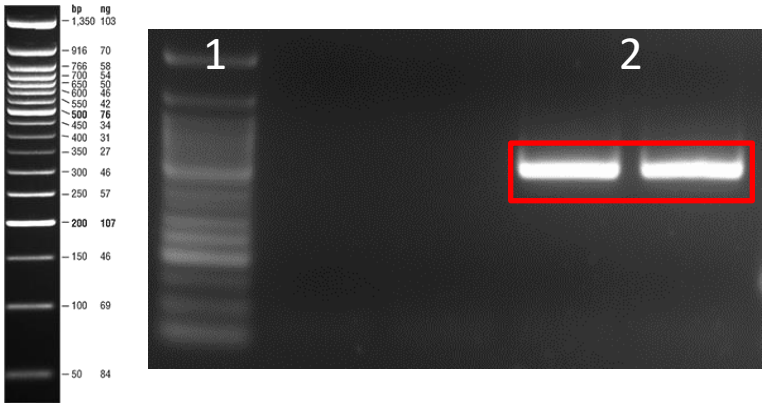

Figure 34: Colony PCR result of THI-COM-KRAB-Com stem loop-scrgRNA plasmid. Line 2 shows positive colony result for THI-COM-KRAB-Com stem loop-scrgRNA, with the expected 515 bp. Lane 1 shows the DNA marker (NEB).

The accuracy of gRNAs in the COM-KRAB module was verified by Sanger sequencing. The results are given in the Appendix D, Figure D13, D14, D15, D16, D17, and D18.

For linearization, THI-VP64-scrgRNA, THI-VP64-Ver1-gRNA1,2,3,4,5 and THI-VP64-Ver2-gRNA1,2,3,4,5 plasmids were digested with AscI restriction enzyme. PCR clean-up was performed after single digestion to get rid of PCR reagents and not to lose DNA.

3.5 Transformation of CRISPR-TF: CRISPRa/i into *Pichia pastoris*

The linearized plasmids (THI-VP64-scrgRNA and THI-VP64-ver1-gRNA1,2,3,4,5) were transformed into *Pichia pastoris* + Psyn-GAP-eGFP-version 1 colony V1-11, and (THI-VP64-scrgRNA and THI-VP64-ver1-gRNA1,2,3,4,5) were transformed into

Pichia pastoris + Psyn-GAP-eGFP-version 2 colony V2-3 as described in 2.2. As a result of this transformation, only those with two different resistance genes could form colonies, so colony PCR was not performed separately. In addition, since the colonies contained the *eGFP* gene, -80 stocks were taken only from colonies that gave green color.

3.6 Evaluation of eGFP Expression under Different Carbon Sources and CRISPR-TF Using Flow Cytometry

In order to understand how the VP64 activator module changes eGFP production with the designed gRNAs, a screening experiment was performed as described in 2.3. In this thesis, the transcriptional activation potential of two synthetically engineered GAP promoter variants in *Pichia pastoris* was evaluated using a CRISPR activation (CRISPRa) system composed of dCas9-VP64 and MS2-based RNA scaffolds. The promoter variants (version 1 and version 2) were designed with specific mutations that enabled the integration of five novel gRNA target sites per version. These synthetic promoters were introduced into yeast strains carrying a single-copy plasmid harboring *eGFP* as a reporter gene. To assess activation efficiency, each engineered strain was transformed with plasmids encoding the appropriate gRNA, dCas9-VP64, and MS2 scaffold modules. Scrambled gRNA (scrgRNA), which does not bind to any promoter region, was used as a negative control. Expression levels were quantified via flow cytometry and normalized to the scrambled control to calculate fold changes. Gene copy number was confirmed by qPCR using the *ARG1* housekeeping gene to validate that the observed differences in gene expression originated from transcriptional regulation rather than differences in plasmid dosage.

3.6.1 Promoter Variant-Specific Activation Patterns

Among the two promoter versions, version 1 demonstrated significant transcriptional activation in response to several gRNAs. In particular, gRNA v1-g2-c1 resulted in substantial upregulation of eGFP expression under glucose ($p < 0.001$), glycerol ($p < 0.0001$), methanol ($p < 0.0001$), and ethanol conditions. Additional but lower levels of activation were also observed for gRNA v1-g2-c2 and gRNA v1-g3-c5**,** particularly under glucose and glycerol. These results are shown in Figure 35, which presents fold change data for version 1 under different carbon sources.

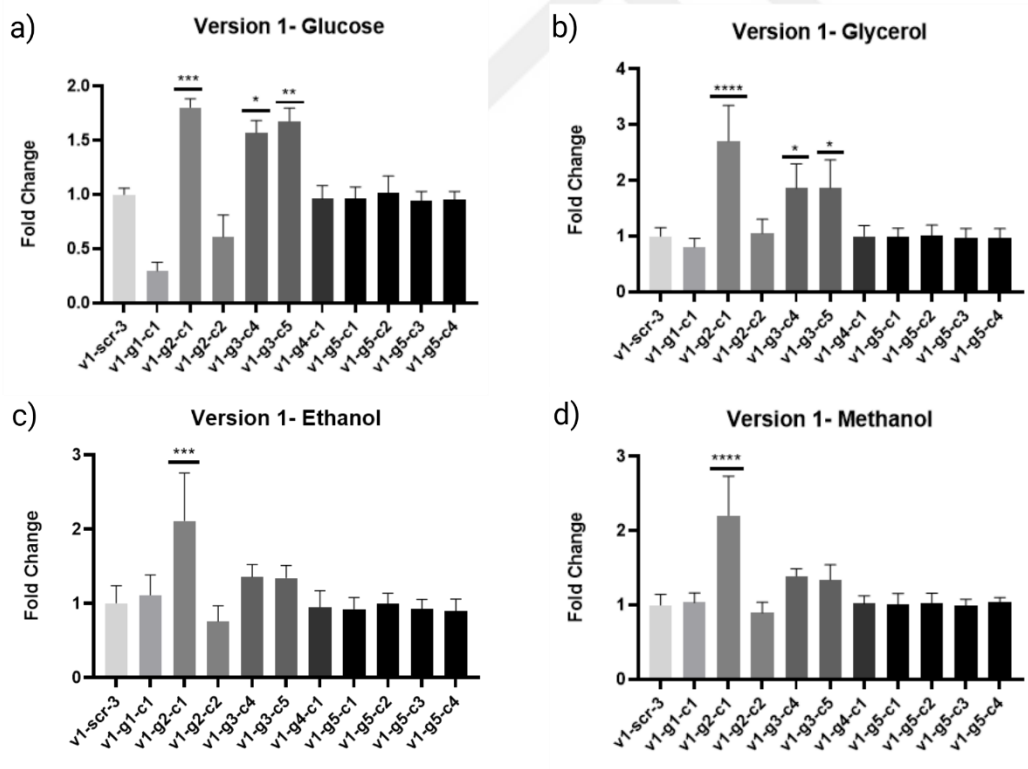

Figure 35: Fold change in eGFP expression for P. pastoris + THI-VP64-ver1-eGFP under different carbon sources. Data were normalized to scrambled gRNA control. Significant gRNA effects are indicated.

In contrast, version 2 did not exhibit significant activation under any tested condition. Fold change values remained close to baseline for all gRNAs and carbon sources, indicating limited responsiveness of this promoter variant to CRISPRa modulation. The data for version 2 is presented in Figure 36. This difference may be explained by the disruption of essential cis-regulatory elements or suboptimal positioning of the gRNA target sites relative to the transcription start site (TSS).

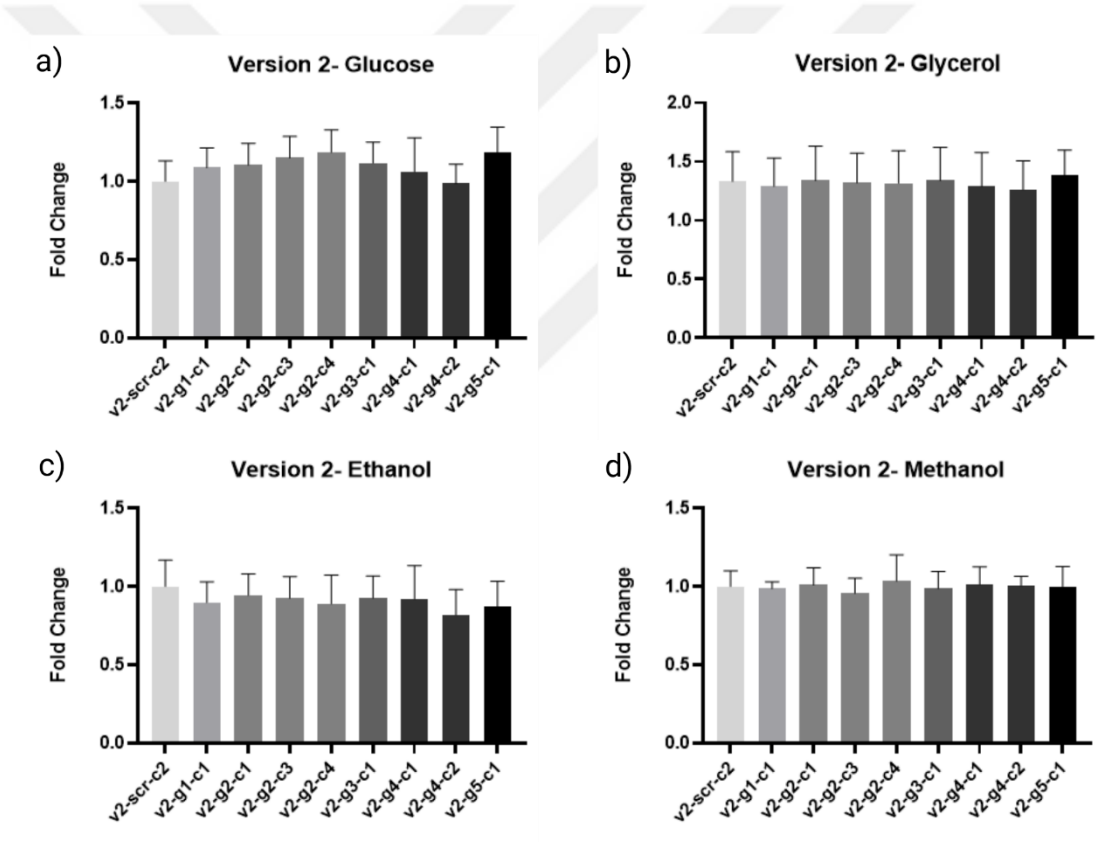

Figure 36: Fold change in eGFP expression for P. pastoris + THI-VP64-ver2-eGFP under different carbon sources. Data were normalized to scrambled gRNA control. No significant activation was observed.

These findings are consistent with the work of Baumschabl et al. (2020) [37], which emphasized the importance of gRNA positioning in relation to the TSS for effective transcriptional activation. It was reported that gRNA sites located approximately −70 to −100 bp upstream of the TSS result in the highest activation, while those closer to the TSS may block RNA polymerase binding and reduce expression. Based on the strong fold changes observed with v1-g2-c1, it is likely that this gRNA is positioned within this optimal activation window.

Further support is provided by Omachi & Miner (2022) [51], who demonstrated that MS2-scaffold-based recruitment of VP64, particularly when combined with multiple gRNAs, enhances activation efficiency significantly compared to single VP64-fusion systems. Although only a single gRNA was used per construct in this study, the presence of MS2-based recruitment may have amplified local activator concentration and contributed to the strong transcriptional output observed for version 1.

3.6.2 Effect of Carbon Sources on CRISPRa Performance

The carbon source–specific performance of gRNAs was also evaluated. The most responsive gRNA, v1-g2-c1, consistently induced high expression under glucose, glycerol, methanol, and ethanol. This broad activity profile indicates that the synthetic GAP version 1 retained functional accessibility under both fermentative and respiratory metabolic conditions. These observations support the findings of Baumschabl et al. (2020) [37], which showed that CRISPRa performance may vary depending on the carbon source, potentially due to changes in cellular transcriptional capacity or differences in promoter strength.

The observation that CRISPRa-mediated activation from a GAP-based promoter was sustained under methanol and ethanol, carbon sources typically associated with AOX1-based inducible expression systems, is particularly noteworthy. These results suggest that synthetic versions of constitutive promoters can be engineered to enable broad-range expression control, increasing the flexibility of bioprocess design in *P. pastoris*.

3.6.3 Gene Copy Number Validation

To confirm that differences in eGFP expression were not due to unequal plasmid integration, qPCR-based gene copy number analysis was performed using *ARG1* as a single-copy reference gene. The results indicated that all transformants used in the evaluation harbored a single copy of the expression cassette. Therefore, the observed differences in fluorescence intensity could be attributed to transcriptional regulation rather than variation in gene dosage. These findings are illustrated in Figure 37.

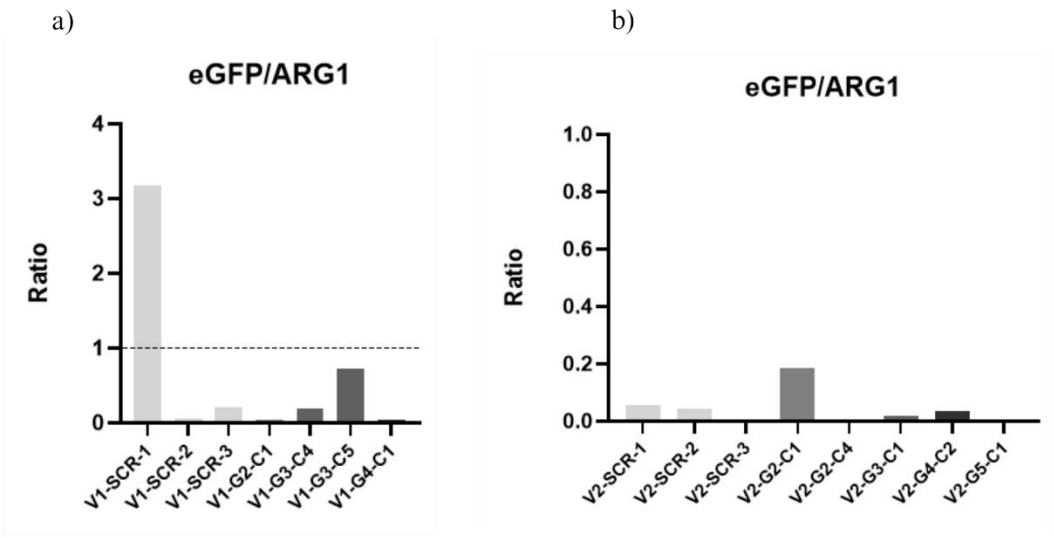

Figure 37: Relative *eGFP* gene copy number (normalized to *ARG1*) determined by qPCR for selected transformants from version 1 and version 2 constructs.

CHAPTER IV: CONCLUSION AND FUTURE PERSPECTIVES

4.1 Conclusion

In this study, a modular and tunable transcriptional regulation platform was established in Pichia pastoris by integrating CRISPRa components with synthetically engineered GAP promoters. The general objective was to generate a controllable gene expression system based on a traditionally constitutive promoter (PGAP), capable of both transcriptional repression (CRISPRi) and activation (CRISPRa), depending on the regulatory modules applied.

Initially, two different versions of the GAP promoter (version 1 and version 2) were rationally designed by introducing sequence-specific mutations to create novel CRISPR gRNA binding sites. These promoters were cloned upstream of an eGFP reporter and transformed into *P. pastoris*. A colony screening experiment under four different carbon sources (glucose, glycerol, ethanol, and methanol) revealed that the majority of mutant promoters retained comparable strength to the wild-type GAP promoter. Some colonies demonstrated markedly higher expression levels, which were subsequently linked to increased gene copy number, as confirmed by qPCR analysis. To ensure accurate evaluation of transcriptional activity, colonies harboring single-copy integration events were selected for downstream experiments.

In the second phase of the study, CRISPRa modules comprising dCas9-VP64, MS2 scaffold, and custom gRNAs were introduced into selected colonies. A systematic screening of ten gRNAs—five targeting version 1 and five targeting version 2

promoters—was conducted across various carbon conditions. Notably, version 1 exhibited significant transcriptional activation with several gRNAs, particularly v1-g2-c1, which consistently led to increased eGFP production under glucose, glycerol, methanol, and ethanol. These results highlighted the context-dependency of gRNA-mediated activation and emphasized the importance of binding site location relative to core promoter elements. In contrast, version 2 failed to exhibit notable activation under any tested condition, suggesting the presence of inhibitory mutations or suboptimal gRNA positioning.

Collectively, the findings of this thesis provide strong evidence that transcriptional activation of a constitutive promoter can be precisely modulated in *P. pastoris* by combining rational promoter engineering with CRISPRa systems. The use of MS2-based scaffold recruitment was shown to significantly enhance transcriptional output, in agreement with previous literature. Furthermore, the modular promoter design enabled carbon-source-independent regulation, offering a versatile platform for synthetic biology and metabolic engineering applications in yeast. The strategy introduced here paves the way for future developments in tunable expression systems that move beyond inducible promoters and toward customizable regulatory architectures tailored to specific metabolic and industrial needs.

4.2 Future Perspectives

While promising results were obtained using the VP64-based CRISPRa system, future studies may explore the efficiency of alternative transcriptional activators such as p65AD and VPR, which have already been cloned into constructs containing scrambled

gRNAs. The most effective gRNA identified in this study, v1-g2-c1, can be further tested with these stronger activators in *P. pastoris* strains harboring single-copy version 1 and version 2 promoter variants. Subsequent transformation, screening, and flow cytometry analysis would enable a direct comparison of transcriptional activation efficiencies among different activator domains.

In addition to transcriptional activation, the groundwork was also laid for a CRISPRi system to enable promoter repression. A repressor module was constructed by replacing the MS2 stem loop with a COM stem loop, the MS2 coat protein with a COM-binding protein, and the VP64 activation domain with the KRAB repression domain. Scrambled and target-specific gRNA versions of these repressor constructs were successfully assembled but not yet transformed into *P. pastoris*. Therefore, the efficiency of CRISPRi-mediated repression remains to be experimentally validated. Future experiments could involve transformation of these repressor modules into version 1 and version 2 single-copy strains followed by phenotypic and fluorescence-based expression analysis.

Overall, this study introduces a synthetic gene regulation toolkit for *P. pastoris* that can be expanded to include stronger transcriptional activators, repressor systems, and potentially multiplexed gRNA strategies. These advancements will contribute to the development of more sophisticated, programmable, and tightly controlled gene expression systems for diverse applications in biotechnology and synthetic biology.

BIBLIORGRAPHY

1. Sewalt V, Shanahan D, Gregg L, La Marta J, Carrillo R. The Generally Recognized as Safe (GRAS) Process for Industrial Microbial Enzymes. Industrial Biotechnology. 2016;12(5):295-302.

2. Kavšček M, Stražar M, Curk T, Natter K, Petrovič U. Yeast as a cell factory: current state and perspectives. Microbial Cell Factories. 2015;14(1):94.

3. Geijer C, Ledesma-Amaro R, Tomás-Pejó E. Unraveling the potential of non-conventional yeasts in biotechnology. FEMS Yeast Research. 2022;22(1).

4. Navarrete C, Martínez JL. Non-conventional yeasts as superior production platforms for sustainable fermentation based bio-manufacturing processes. AIMS Bioengineering. 2020;7(4):289-305.

5. Pontrelli S, Chiu T-Y, Lan EI, Chen FYH, Chang P, Liao JC. Escherichia coli as a host for metabolic engineering. Metabolic Engineering. 2018;50:16-46.

6. Rebello S, Abraham A, Madhavan A, Sindhu R, Binod P, Karthika Bahuleyan A, et al. Non-conventional yeast cell factories for sustainable bioprocesses. FEMS Microbiology Letters. 2018;365(21).

7. Guilliermond A. Zygosaccharomyces Pastori, nouvelle espèce de levures à copulation hétérogamique: Société mycologique de France; 1920.

8. Csoma H, Ács-Szabó L, Papp LA, Sipiczki M. Application of different markers and data-analysis tools to the examination of biodiversity can lead to different results: a case study with Starmerella bacillaris (synonym Candida zemplinina) strains. FEMS Yeast Res. 2018;18(5).

9. Zahrl RJ, Peña DA, Mattanovich D, Gasser B. Systems biotechnology for protein production in Pichia pastoris. FEMS Yeast Research. 2017;17(7).

10. Wang N, Xiao Y, Liu X, Li Y, Yu D, Guo J, et al. Structural Homology Fails to Predict Secretion Efficiency in Pichia pastoris: Divergent Responses of Architecturally Similar scFvs to Multi-Parametric Genetic Engineering. International Journal of Molecular Sciences. 2025;26(10):4922.

11. Karbalaei M, Rezaee SA, Farsiani H. Pichia pastoris: A highly successful expression system for optimal synthesis of heterologous proteins. Journal of Cellular Physiology. 2020;235(9):5867-81.

12. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. Heterologous protein production using the Pichia pastoris expression system. Yeast. 2005;22(4):249-70.

13. Fidan O, Zhan J. Recent advances in engineering yeast for pharmaceutical protein production. RSC Advances. 2015;5(105):86665-74.

14. Hahn S, Young ET. Transcriptional Regulation in Saccharomyces cerevisiae: Transcription Factor Regulation and Function, Mechanisms of Initiation, and Roles of Activators and Coactivators. Genetics. 2011;189(3):705-36.

15. Dossani ZY, Reider Apel A, Szmidt-Middleton H, Hillson NJ, Deutsch S, Keasling JD, et al. A combinatorial approach to synthetic transcription factor-promoter combinations for yeast strain engineering. Yeast. 2018;35(3):273-80.

16. Urlinger S, Baron U, Thellmann M, Hasan MT, Bujard H, Hillen W. Exploring the sequence space for tetracycline-dependent transcriptional activators: Novel mutations yield expanded range and sensitivity. Proceedings of the National Academy of Sciences. 2000;97(14):7963-8.

17. Khalil Ahmad S, Lu Timothy K, Bashor Caleb J, Ramirez Cherie L, Pyenson Nora C, Joung JK, et al. A Synthetic Biology Framework for Programming Eukaryotic Transcription Functions. Cell. 2012;150(3):647-58.

18. Rahman SH, Maeder ML, Joung JK, Cathomen T. Zinc-Finger Nucleases for Somatic Gene Therapy: The Next Frontier. Human Gene Therapy. 2011;22(8):925-33.

19. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. Nature Biotechnology. 2012;30(5):460-5.

20. Glick BR. Metabolic load and heterologous gene expression. Biotechnology Advances. 1995;13(2):247-61.

21. Horvath P, Barrangou R. CRISPR/Cas, the Immune System of Bacteria and Archaea. Science. 2010;327(5962):167-70.

22. Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nature Protocols. 2013;8(11):2180-96.

23. Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Research. 2013;41(15):7429-37.

24. Farzadfard F, Perli SD, Lu TK. Tunable and Multifunctional Eukaryotic Transcription Factors Based on CRISPR/Cas. ACS Synthetic Biology. 2013;2(10):604-13.

25. Machens F, Balazadeh S, Mueller-Roeber B, Messerschmidt K. Synthetic Promoters and Transcription Factors for Heterologous Protein Expression in Saccharomyces cerevisiae. Frontiers in Bioengineering and Biotechnology. 2017;Volume 5 - 2017.

26. Chen J, Willingham T, Margraf LR, Schreiber-agus N, Depinho RA, Nisen PD. Effects of the MYC oncogene antagonist, MAD, on proliferation, cell cycling and the

malignant phenotype of human brain tumour cells. Nature Medicine. 1995;1(7):638-43.

27. Qi Lei S, Larson Matthew H, Gilbert Luke A, Doudna Jennifer A, Weissman Jonathan S, Arkin Adam P, et al. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell. 2013;152(5):1173-83.

28. Gander MW, Vrana JD, Voje WE, Carothers JM, Klavins E. Digital logic circuits in yeast with CRISPR-dCas9 NOR gates. Nature Communications. 2017;8(1):15459.

29. Zalatan Jesse G, Lee Michael E, Almeida R, Gilbert Luke A, Whitehead Evan H, La Russa M, et al. Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. Cell. 2015;160(1):339-50.

30. Jeong Y-J, Jeongin C, Jina K, Hoon SY, and Kang B-C. Immortalization of primary marmoset skin fibroblasts by CRISPR-Cas9-mediated gene targeting. Animal Cells and Systems. 2022;26(6):266-74.

31. Brady JR, Whittaker CA, Tan MC, Kristensen II DL, Ma D, Dalvie NC, et al. Comparative genome-scale analysis of Pichia pastoris variants informs selection of an optimal base strain. Biotechnology and Bioengineering. 2020;117(2):543-55.

32. Weninger A, Hatzl A-M, Schmid C, Vogl T, Glieder A. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast Pichia pastoris. Journal of Biotechnology. 2016;235:139-49.

33. Gu Y, Gao J, Cao M, Dong C, Lian J, Huang L, et al. Construction of a series of episomal plasmids and their application in the development of an efficient CRISPR/Cas9 system in Pichia pastoris. World Journal of Microbiology and Biotechnology. 2019;35(6):79.

34. Dalvie NC, Lorgeree T, Biedermann AM, Love KR, Love JC. Simplified Gene Knockout by CRISPR-Cas9-Induced Homologous Recombination. ACS Synthetic Biology. 2022;11(1):497-501.

35. Liu Q, Shi X, Song L, Liu H, Zhou X, Wang Q, et al. CRISPR–Cas9-mediated genomic multiloci integration in Pichia pastoris. Microbial Cell Factories. 2019;18(1):144.

36. Cai P, Duan X, Wu X, Gao L, Ye M, Zhou Yongjin J. Recombination machinery engineering facilitates metabolic engineering of the industrial yeast Pichia pastoris. Nucleic Acids Research. 2021;49(13):7791-805.

37. Baumschabl M, Prielhofer R, Mattanovich D, Steiger MG. Fine-Tuning of Transcription in Pichia pastoris Using dCas9 and RNA Scaffolds. ACS Synthetic Biology. 2020;9(12):3202-9.

38. Liao X, Li L, Jameel A, Xing X-H, Zhang C. A versatile toolbox for CRISPR-based genome engineering in Pichia pastoris. Applied Microbiology and Biotechnology. 2021;105(24):9211-8.

39. Liu Q, Song L, Peng Q, Zhu Q, Shi X, Xu M, et al. A programmable high-expression yeast platform responsive to user-defined signals. Science Advances. 2022;8(6):eabl5166.

40. Li T, Ma J, Xu Z, Wang S, Wang N, Shao S, et al. Transcriptomic Analysis of the Influence of Methanol Assimilation on the Gene Expression in the Recombinant Pichia pastoris Producing Hirudin Variant 3. Genes. 2019;10(8):606.

41. Waterham HR, Digan ME, Koutz PJ, Lair SV, Cregg JM. Isolation of the Pichia pastoris glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. Gene. 1997;186(1):37-44.

42. Ata Ö, Prielhofer R, Gasser B, Mattanovich D, Çalık P. Transcriptional engineering of the glyceraldehyde-3-phosphate dehydrogenase promoter for improved heterologous protein production in Pichia pastoris. Biotechnology and Bioengineering. 2017;114(10):2319-27.

43. Qin X, Qian J, Yao G, Zhuang Y, Zhang S, Chu J. GAP Promoter Library for Fine-Tuning of Gene Expression in Pichia pastoris. Applied and Environmental Microbiology. 2011;77(11):3600-8.

44. Anders C, Niewoehner O, Duerst A, Jinek M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature. 2014;513(7519):569-73.

45. Chen M, Qi LS. Repurposing CRISPR System for Transcriptional Activation. In: Li L-C, editor. RNA Activation. Singapore: Springer Singapore; 2017. p. 147-57.

46. Gilbert Luke A, Larson Matthew H, Morsut L, Liu Z, Brar Gloria A, Torres Sandra E, et al. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. Cell. 2013;154(2):442-51.

47. Ke J, Ma H, Gu X, Thelen A, Brunzelle JS, Li J, et al. Structural basis for recognition of diverse transcriptional repressors by the TOPLESS family of corepressors. Sci Adv. 2015;1(6):e1500107.

48. Replogle JM, Norman TM, Xu A, Hussmann JA, Chen J, Cogan JZ, et al. Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. Nature Biotechnology. 2020;38(8):954-61.

49. Zhang J-P, Li X-L, Neises A, Chen W, Hu L-P, Ji G-Z, et al. Different Effects of sgRNA Length on CRISPR-mediated Gene Knockout Efficiency. Scientific Reports. 2016;6(1):28566.

50. Hohenblum H, Borth N, Mattanovich D. Assessing viability and cell-associated product of recombinant protein producing Pichia pastoris with flow cytometry. Journal of Biotechnology. 2003;102(3):281-90.

51. Omachi K, Miner JH. Comparative analysis of dCas9-VP64 variants and multiplexed guide RNAs mediating CRISPR activation. PLOS ONE. 2022;17(6):e0270008.

APPENDIX A

DNA SEQUENCES OF CONSTRUCTS

Table A.1: Gene sequences and regulatory elements used in this study

Gene Name	Sequence
eGFP	ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGC CCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAA GTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTAC GGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCT GCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACG GCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCA GCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCC AGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAA GACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTG AACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACG GCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAG CCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGC ATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACG GCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCC CATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACC TGAGCACCCAGTCCAAGCTGAGCAAAGACCCCAACGAGAA GCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCG GGATCACTCTCGGCATGGACGAGCTGTACAAGTAA
Wild type GAP promoter	TTTTTGTAGAAATGTCTTGGTGTCCTCGTCCAATCAGGTAG CCATCTCTGAAATATCTGGCTCCGTTGCAACTCCGAACGAC CTGCTGGCAACGTAAAATTCTCCGGGGTAAAACTTAAATGT GGAGTAATGGAACCAGAAACGTCTCTTCCCTTCTCTCTCCT TCCACCGCCCGTTACCGTCCCTAGGAAATTTTACTCTGCTG GAGAGCTTCTTCTACGGCCCCCTTGCAGCAATGCTCTTCCC AGCATTACGTTGCGGGTAAAACGGAGGTCGTGTACCCGAC CTAGCAGCCCAGGGATGGAAAAGTCCCGGCCGTCGCTGGC AATAATAGCGGGCGGACGCATGTCATGAGATTATTGGAAA CCACCAGAATCGAATATAAAAGGCGAACACCTTTCCCAATT TTGGTTTCTCCTGACCCAAAGACTTTAAATTTAATTTATTTG TCCCTATTTCAATCAATTGAACAACTAT
PsynGAP-version 1 promoter	TTTTTGTAGAAATGTCTTGGTGTCCTCGTCCAATCAGGTAG CCATCTCTGAAATATCTGGCTCCGTTGCAACTCCGAACGAC CTGCTGGCAACGTAAAATTCTCCGGGGTAAAACTTAAATGT GGAGTAATGGAACCAGAAACGTCTCTTCCCTTCTCTCTCCT TCCACCGCCCGTTACCGTCCCTAGGAAATTTTACTGGGCTG GAGAGGAAGTTCTACGGCCCCCTTGCAGCAATGGTGTAGC GAGCATTACGATGCGGGTAAAACGGAGGTCGTGTACCCGA CCTAGCAGCCCAGGGATGGAAAAGTCCCGGCAGTCGGTGC

	CAATAGTAGCGGGCGGACGCATGTCATGAGATTATTGGAA ACCACCAGAGGCGAATATAAAAGGCGAACACCTTTCCCAA TTTTGGTTTCTCCTGACCCAAAGACTTTAAATTTAATTTATT TGTCCCTATTTCAATCAATTGAACAACTAT
PsynGAP-version 2 promoter	TTTTTGTAGAAATGTCTTGGTGTCCTCGTCCAATCAGGTAG CCATCTCTGAAATATCTGGCTCCGTTGCAACTCCGAACGAC CTGCTGGCAACGTAAAATTCTCCGGGGTAAAACTTAAATGT GGAGTAATGGAACCAGAAACGTCTCTTCCCTTCTCTCTCCT TCCACCGCCCGTTACCGTCCCTAGGAAATTTTACTCTGCTG GAGAGCTTCTTCTACAGCCCCCTTGCAGCAGTGCGGTTCCC AGCATTACGTTGCGGGTAAAACGGAGGTCGTGTACCCGAC CTAGCAGCCCAGGGATGGAAAAGTCCCGGCAGTCGCAGGG GATAATAGCGGGCGGACGCATGTCATGAGATTATTGGAAA CCACCGGAATCGAATATAAAAGGCGAACGGCTTTCCCAAT TTTGGTTTCTCCTGACCCAAAGACTTTAAATTTAATTTATTT GTCCCTATTTCAATCAATTGAACAACTAT
scrgRNA	TCCAAAGACTCCCTGTTTTC
Version1-gRNA1	AGATTATTGGAAACCACCAGAGG
Version1-gRNA2	GCAGTCGGTGCCAATAGTAGCGG
Version1-gRNA3	CCCCCTTGCAGCAATGGTGTAGC
Version1-gRNA4	GTGTAGCGAGCATTACGATGCGG
Version1-gRNA5	GGGCTGGAGAGGAAGTTCTACGG
Version2-gRNA1	AATCGAATATAAAAGGCGAACGG
Version2-gRNA2	ATGAGATTATTGGAAACCACCGG
Version2-gRNA3	GCAGTCGCAGGGGATAATAGCGG
Version2-gRNA4	AAAAGTCCCGGCAGTCGCAGGGG
Version2-gRNA5	ACAGCCCCCTTGCAGCAGTGCGG
MS2	GCGTCTAACTTCACCCAGTTCGTTCTGGTTGACAACGGTGG TACCGGTGACGTTACCGTTGCGCCGTCTAACTTCGCGAACG GTGTTGCGGAATGGATCTCTTCTAACTCTCGTTCTCAGGCG TACAAAGTTACCTGCTCTGTTCGTCAATCCTCCGCGCAGAA CCGTAAATACACCATCAAAGTTGAAGTTCCGAAAGTTGCG ACCCAGACCGTTGGTGGTGTTGAACTGCCGGTTGCGGCGTG GCGTTCTTACCTGAACATGGAACTGACCATCCCGATCTTCG CGACCAACTCTGACTGCGAACTGATCGTTAAAGCGATGCA GGGTCTGCTGAAAGACGGTAACCCGATCCCGTCTGCGATCG CGGCGAACTCTGGTATCTAC
MS2 Stem Loop	ACATGAGGATCACCCATGT

VP64	GATGCCCTTGACGATTTTGACTTGGACATGTTGGGTTCTGA CGCTTTGGATGACTTTGATCTTGATATGCTTGGTTCCGACGC TCTAGATGATTTCGACTTGGATATGCTGGGATCCGATGCCT TGGACGATTTCGACTTGGATATGTTG
p65AD	CCTACACAGGCCGGCGAGGGCACACTGTCTGAAGCTCTGCT GCAGCTGCAGTTCGACGACGAGGATCTGGGAGCCCTGCTG GGAAACAGCACCGATCCTGCCGTGTTCACCGACCTGGCCA GCGTGGACAACAGCGAGTTCCAGCAGCTGCTGAACCAGGG CATCCCTGTGGCCCCTCACACCACCGAGCCCATGCTGATGG AATACCCCGAGGCCATCACCCGGCTCGTGACAGGCGCTCA GAGGCCTCCTGATCCAGCTCCTGCCCCTCTGGGAGCACCAG GCCTGCCTAATGGACTGCTGTCTGGCGACGAGGACTTCAGC TCTATCGCCGACATGGACTTCTCCGCACTGCTG
VPR	GGACGGGCTGACGCATTGGACGATTTTGATCTGGATATGCT GGGAAGTGACGCCCTCGATGATTTTGACCTTGACATGCTTG GTTCGGATGCCCTTGATGACTTTGACCTCGACATGCTCGGC AGTGACGCCCTTGATGATTTCGACCTGGACATGCTGATTAA CTCTAGAAGTTCCGGATCTCCGAAAAAGAAACGCAAAGTT GGTAGCCAGTACCTGCCCGACACCGACGACCGGCACCGGA TCGAGGAAAAGCGGAAGCGGACCTACGAGACATTCAAGAG CATCATGAAGAAGTCCCCCTTCAGCGGCCCCACCGACCCTA GACCTCCACCTAGAAGAATCGCCGTGCCCAGCAGATCCAG CGCCAGCGTGCCAAAACCTGCCCCCCAGCCTTACCCCTTCA CCAGCAGCCTGAGCACCATCAACTACGACGAGTTCCCTACC ATGGTGTTCCCCAGCGGCCAGATCTCTCAGGCCTCTGCTCT GGCTCCAGCCCCTCCTCAGGTGCTGCCTCAGGCTCCTGCTC CTGCACCAGCTCCAGCCATGGTGTCTGCACTGGCTCAGGCA CCAGCACCCGTGCCTGTGCTGGCTCCTGGACCTCCACAGGC TGTGGCTCCACCAGCCCCTAAACCTACACAGGCCGGCGAG GGCACACTGTCTGAAGCTCTGCTGCAGCTGCAGTTCGACGA CGAGGATCTGGGAGCCCTGCTGGGAAACAGCACCGATCCT GCCGTGTTCACCGACCTGGCCAGCGTGGACAACAGCGAGT TCCAGCAGCTGCTGAACCAGGGCATCCCTGTGGCCCCTCAC ACCACCGAGCCCATGCTGATGGAATACCCCGAGGCCATCA CCCGGCTCGTGACAGGCGCTCAGAGGCCTCCTGATCCAGCT CCTGCCCCTCTGGGAGCACCAGGCCTGCCTAATGGACTGCT GTCTGGCGACGAGGACTTCAGCTCTATCGCCGACATGGACT TCTCCGCACTGCTGGGTAGCGGATCGGGATCTCGGGATTCC AGGGAAGGGATGTTTTTGCCGAAGCCTGAGGCCGGCTCCG CTATTAGTGACGTGTTTGAGGGCCGCGAGGTGTGCCAGCCA AAACGAATCCGGCCATTTCATCCTCCAGGAAGTCCATGGGC CAACCGCCCACTCCCCGCCAGCCTCGCACCAACACCAACCG GTCCAGTACATGAGCCAGTCGGGTCACTGACCCCGGCACC AGTCCCTCAGCCACTGGATCCAGCGCCCGCAGTGACTCCCG AGGCCAGTCACCTGTTGGAGGATCCCGATGAAGAGACGAG CCAGGCTGTCAAAGCCCTTCGGGAGATGGCCGATACTGTG ATTCCCCAGAAGGAAGAGGCTGCAATCTGTGGCCAAATGG ACCTTTCCCATCCGCCCCCAAGGGGCCATCTGGATGAGCTG ACAACCACACTTGAGTCCATGACCGAGGATCTGAACCTGG ACTCACCCCTGACCCCGGAATTGAACGAGATTCTGGATACC

	TTCCTGAACGACGAGTGCCTCTTGCATGCCATGCATATCAG CACAGGACTGTCCATCTTCGACACATCTCTGTTT
COM	GGTAGTATGAAATCAATTCGCTGTAAAAACTGCAACAAAC TGTTATTTAAGGCGGATAGTTTTGATCACATTGAAATCAGG TGTCCGCGTTGCAAACGTCACATCATAATGCTGAATGCCTG CGAGCATCCCACGGAGAAACATTGTGGGAAAAGAGAAAAA ATCACGCATTCTGACGAAACCGTGCGTTATGGCTCCACTAG TGGGCACAAGCTTAATGGCGGTGGCGGAGGGATGGATGCT AAGTCACTAACTGCCTGGTCC
COM Stem Loop	CTGAATGCCTGCGAGCATC
KRAB	CGGACACTGGTGACCTTCAAGGATGTATTTGTGGACTTCAC CAGGGAGGAGTGGAAGCTGCTGGACACTGCTCAGCAGATC GTGTACAGAAATGTGATGCTGGAGAACTATAAGAACCTGG TTTCCTTGGGTTATCAGCTTACTAAGCCAGATGTGATCCTCC GGTTGGAGAAGGGAGAAGAGCCC
Mxi1	ATGGAACGTGTGAGAATGATTAATGTGCAAAGGCTGTTAG AAGCCGCAGAGTTTTTAGAAAGAAGAGAAAGAGAATGCGA ACACGGGTATGCCAGTTCTTTCCCTAGCATGCCCTCTCCCA GA
TPL	ATGAGCAGCCTGAGCCGCGAACTGGTGTTCCTGATCTTACA ATTCCTGGACGAGGAGAAGTTCAAGGAGACCGTGCACAAG CTGGAGCAGGAGAGTGGTTTCTTCTTTAACATGAAGTACTT CGAAGACGAAGTGCACAACGGCAACTGGGACGAGGTGGA GAAATACCTGAGCGGCTTCACCAAGGTTGACGACAACCGC TATAGCATGAAGATTTTCTTCGAGATCCGCAAGCAGAAGTA CCTGGAGGCCCTGGACAAGCACGATCGCCCGAAGGCCGTG GACATCCTGGTGAAGGACCTGAAGGTGTTCAGCACCTTCAA CGAAGAGCTGTTCAAGGAGATCACCCAACTGCTGACCCTG GAGAACTTCCGCGAGAATGAGCAGCTGAGCAAGTACGGCG ACACAAAGAGCGCCCGTGCCATCATGCTGGTGGAGCTGAA GAAACTGATCGAGGCAAACCCGCTGTTCCGCGACAAGTTA CAATTCCCGACACTGCGCAACAGCCGCCTGCGTACCCTGAT CAACCAGAGCCTGAACTGGCAGCACCAGCTGTGCAAGAAC CCGCGCCCGAACCCGGATATCAAGACCCTGTTCGTGGACCA CAGTTGCGGCCCTCCGAACGGT
AOX1 Terminator	TCAAGAGGATGTCAGAATGCCATTTGCCTGAGAGATGCAG GCTTCATTTTTGATACTTTTTTATTTGTAACCTATATAGTAT AGGATTTTTTTTGTCATTTTGTTTCTTCTCGTACGAGCTTGC TCCTGATCAGCCTATCTCGCAGCTGATGAATATCTTGTGGT AGGGGTTTGGGAAAATCATTCGAGTTTGATGTTTTTCTTGG TATTTCCCACTCCTCTTCAGAGTACAGAAGATTAAGTGAGA
RGI2	ATACCCAGCATTGACCTTTGGTATGAGCATCTGAAAAACAA CCAGGTGTTGCAAAGTTAAACATCCTTCTTTGTTCATATAG AACCCACTATTCATGGTACTCCCCAATCGAATTTCACATTC TGGTTTTGAAATTACACACCACGTTAGCTTATAAGATTTCA TATAACTTATTGATATACGGTTTCCATTGTTCGAATAGTTGA GGTTGTATGTAATTCGATTGAAGGGGCCATTTTTGTTTCCTA CTTTTCCTGGGAGCTTATCCGATGCGCTTCAAAGCTGGAAT TGTAAATATAGAGAAAAAGAAGGATGTTGTTTTATTCTTGA

	AAGAGTATAATTTTACTTCTAGCAACTCTCCCACTTCGCTTG ACTTCATTTATTTCTTGGGCACATAGGCGTAGTAATCTAGA CCAACAGATAATTTGCCGGAATGATATAGCGATTGGAAAA TGAACTGAAATTTTTTGCTGTCTTTCAATTTGACGGGCAGTT CATCAGGCGCGCCGTGACCGACCATATAAATACGTTGAGA ATGTTATTCTTCCTCGTAGTTGAAGTGGCTTCATAATTTCAG AACTCAATAGATAAACTAGGATGTTTTAAAGCAATTAATGC TCACAAGTAAGGAGCGACTCTCTTGCTTTTCGAATACTAAA AGTATCGTCCCAACCCAGAAAAAAAGACCTCTTAACTGCA AAATAAACTCTATATATTTCTTCTAAAACAGTTTCAGGTTG GATAGTATCGCATTCTCATCACTTCTAACTAGTAGGCCATG AGATATATTAACGTTTACTTGAGTTCTAAGTTCTCCGAATT AGATGCACAGCACAAACAAGATTAGGTTTCACTTGGTACA AAATACGAACAGAGTTTAAGGTCGTAATTTCATTTCGTTAT TGATCCCCACAATCTATTCTTATCACAGTCATCAGATAGTC GCGAAAAAGCATGCAGAAAAGGGGGTCGTCCCTATCTAAG TTGTAGCATTACAACAAATATGAC

APPENDIX B

LIST OF PRIMERS

Table B.1: Primer sequences used for cloning of constructs and qPCR experiments

Primer Name	**Sequence (5' -- 3')**
DA-21	CAATTGAACAACTATTTCGAAACGAGGAATTCATGGTGA GCAAGGGCGAGGAG
DA-22	TTCAGATCCTCTTCTGAGATGAGTTTTTGTGTTCTAGATT ACTTGTACAGCTCGTCCATGCCG
DA-26	GGGCTGGAGAGGAAGTTCTACGGCCCCCTTGCAGCAATG GTGTAGCGAGCATTACGATGCGGGTAAAACGG
DA-27	AGTCGGTGCCAATAGTAGCGGGCGGACGCATGTCATGA GATTATTGGAAACCACCAGAGGCGAATATAAAAG
DA-28	AGCCCCCTTGCAGCAGTGCGGTTCCCAGCATTACG
DA-29	AGTCGCAGGGGATAATAGCGGG
DA-30	GGAATCGAATATAAAAGGCGAACGGCTTTCCCAATTTTG
DA-31	TCGTAATGCTCGCTACACCATTGCTGCAAGGGGGCCGTA GAACTTCCTCTCCAGCCCAGTAAAATTTC
DA-41	GAAATTTTACTCTGCTGGAGAGCTTCTTCTACAGCCCCCT TGCAGCAGTGCGGTTCCCAGCATTACGTTGCGG
DA-42	GTAGAAGAAGCTCTCCAGCAGAGTAAAATTTCCTAGGGA CGGTAACGGG
DA-45	GACGAGCTTACTCGTTTCGTCCTCACGGACTCATCAGTCT AATCATGGTGTTTTGATAGTTG

DA-46	GTCCGTGAGGACGAAACGAGTAAGCTCGTCAGATTATTG GAAACCACCAGAGGGTTTTAGAGCTAGAAATAGC
DA-49	TTACTTGTACAGCTCGTCCATGC
DA-50	GTCCGTGAGGACGAAACGAGTAAGCTCGTCGCAGTCGGT GCCAATAGTAGCGGGTTTTAGAGCTAGAAATAGC
DA-51	GACGAGCTTACTCGTTTCGTCCTCACGGACTCATCAGCG TCAGCATGGTGTTTTGATAGTTG
DA-52	GTCCGTGAGGACGAAACGAGTAAGCTCGTCCCCCCTTGC AGCAATGGTGTAGCGTTTTAGAGCTAGAAATAGC
DA-53	GACGAGCTTACTCGTTTCGTCCTCACGGACTCATCAGGG GGGACATGGTGTTTTGATAGTTG
DA-54	GTCCGTGAGGACGAAACGAGTAAGCTCGTCGTGTAGCG AGCATTACGATGCGGGTTTTAGAGCTAGAAATAGC
DA-55	GACGAGCTTACTCGTTTCGTCCTCACGGACTCATCAGCA CATCCATGGTGTTTTGATAGTTG
DA-56	GTCCGTGAGGACGAAACGAGTAAGCTCGTCGGGCTGGA GAGGAAGTTCTACGGGTTTTAGAGCTAGAAATAGC
DA-57	GACGAGCTTACTCGTTTCGTCCTCACGGACTCATCAGCC CGACCATGGTGTTTTGATAGTTG
DA-58	GTCCGTGAGGACGAAACGAGTAAGCTCGTCAATCGAAT ATAAAAGGCGAACGGGTTTTAGAGCTAGAAATAGC
DA-59	GACGAGCTTACTCGTTTCGTCCTCACGGACTCATCAGTTA GCTCATGGTGTTTTGATAGTTG
DA-60	GTCCGTGAGGACGAAACGAGTAAGCTCGTCATGAGATTA TTGGAAACCACCGGGTTTTAGAGCTAGAAATAGC

DA-61	GACGAGCTTACTCGTTTCGTCCTCACGGACTCATCAGTACTCTCATGGTGTTTTGATAGTTG
DA-62	GTCCGTGAGGACGAAACGAGTAAGCTCGTCGCAGTCGCAGGGGATAATAGCGGGTTTTAGAGCTAGAAATAGC
DA-63	GACGAGCTTACTCGTTTCGTCCTCACGGACTCATCAGCGTCAGCATGGTGTTTTGATAGTTG
DA-65	GACGAGCTTACTCGTTTCGTCCTCACGGACTCATCAGTTTTCACATGGTGTTTTGATAGTTG
DA-66	GTCCGTGAGGACGAAACGAGTAAGCTCGTCACAGCCCCCTTGCAGCAGTGCGGGTTTTAGAGCTAGAAATAGC
DA-67	GACGAGCTTACTCGTTTCGTCCTCACGGACTCATCAGTGTCGGCATGGTGTTTTGATAGTTG
DA-71	GAAACGAGTAAGCTCGTCAAAAGTCCCGGCAGTCGCAGGGGGTTTTAGAGCTAGAAATAGC
DA-72	GGAGCCGTAGATACCAGAGTTC
DA-73	GATCGCGGCGAACTCTGGTATCTACGGCTCCCCTACACAGGCCGGCGAG
DA-74	GTTTATCCTACAGAATTAAACAGAAGCTCACAGCAGTGCGGAGAAGTC
DA-76	GTTTATCCTACAGAATTAAACAGAAGCTCAAAACAGAGATGTGTCGAAGATG
DA-77	GATCGCGGCGAACTCTGGTATCTACGGCTCCGGACGGGCTGACGCATTG
DA-81	CATGCCCAAAAAGAAAAGAAAAGTGGGTAGTGGTAGTATGAAATCAATTCGCTG

DA-82	GTTTATCCTACAGAATTAAACAGAAGCTCATTTGGGCTC TTCTCCCTTC
DA-83	GGACCAGGCAGTTAGTGACTTAG
DA-84	GATGGATGCTAAGTCACTAACTGCCTGGTCCATGGAACG TGTGAGAATG
DA-85	GTTTATCCTACAGAATTAAACAGAAGCTCATCTGGGAGA GGGCATGC
DA-86	TGGATGCTAAGTCACTAACTGCCTGGTCCATGAGCAGCC TGAGCCGC
DA-87	GTTTATCCTACAGAATTAAACAGAAGCTCAACCGTTCGG AGGGCCG
DA-88	TGAGCTTCTGTTTAATTCTGTAGGATAAAC
DA-89	GCACCGACTCGGTGCCAC
DA-90	TTTTGGCCGGCATGGTCCC
DA-91	GATGCTCGCAGGCATTCAGGCACCGACTCGGTGCCACTT TTTCAAGTTGA
DA-92	AATGCCTGCGAGCATCTTTTGGCCGGCATGGTCCCAGCC TCCTCGCTGGC
eGFP For	GGACGACGGCAACTACAAGA
eGFP Rev	CCTTGATGCCGTTCTTCTGC
ARG For	GGTGAGTTGATTGGTCGTGG
ARG Rev	CCGGGCATCAAGACGTCTAT

APPENDIX C

PLASMID MAPS OF CONSTRUCTS

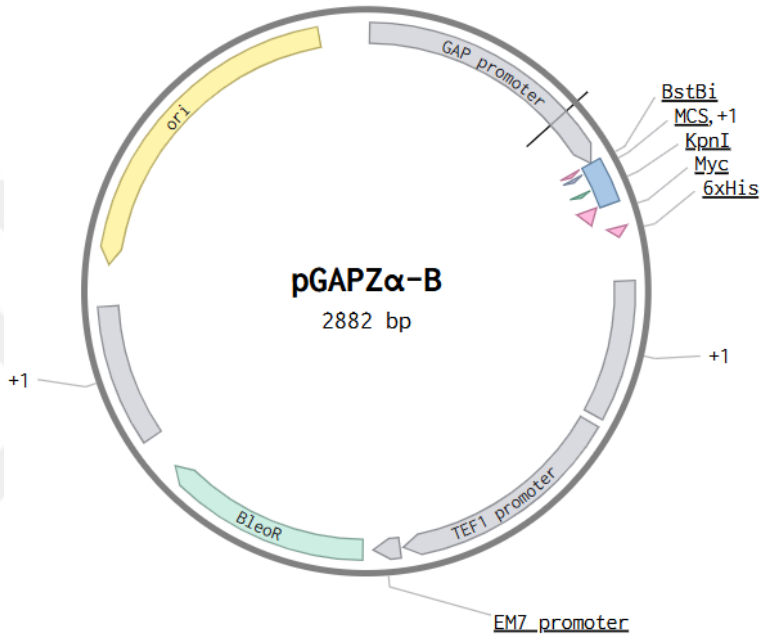

Figure C. 2: The schematic illustration of pGAPZα-B plasmid

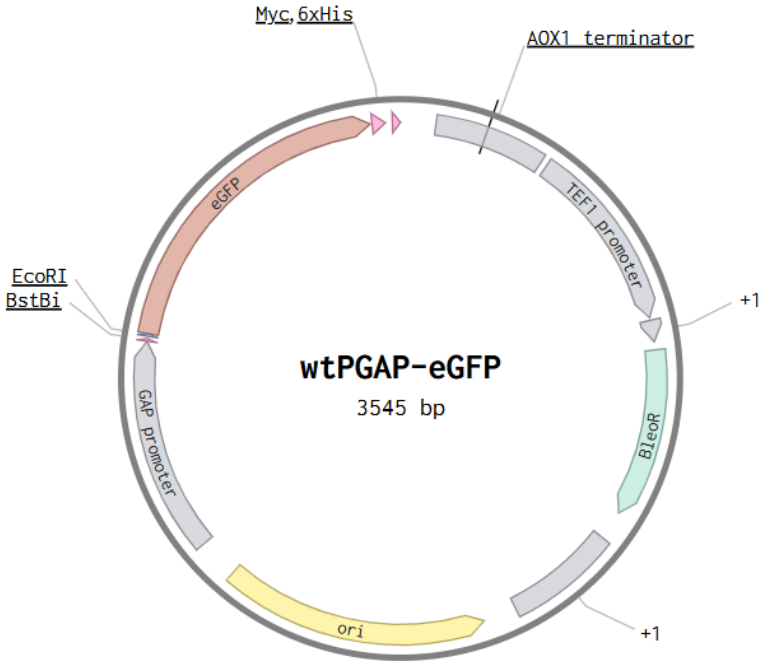

Figure C. 2: The schematic illustration of wtPGAP-eGFP plasmid

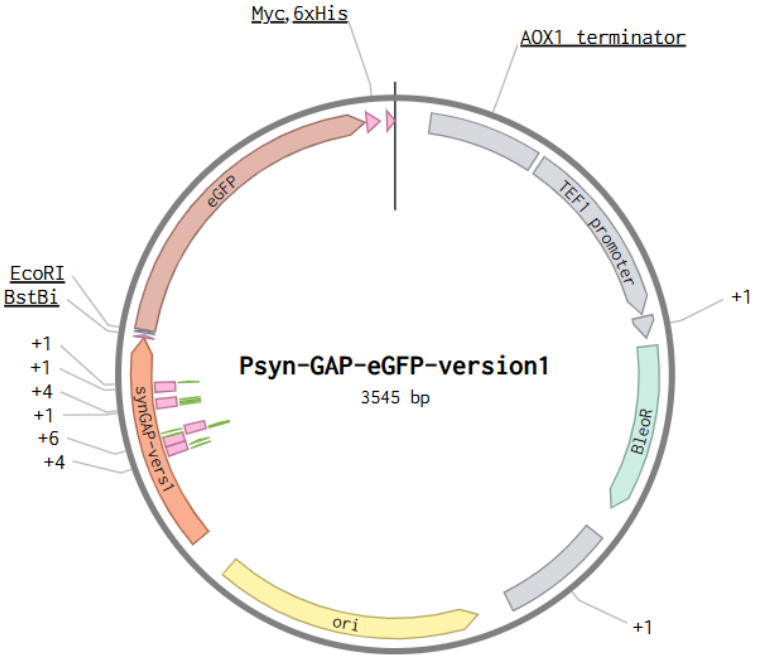

Figure C. 3: The schematic illustration of Psyn-GAP-eGFP-version1 plasmid

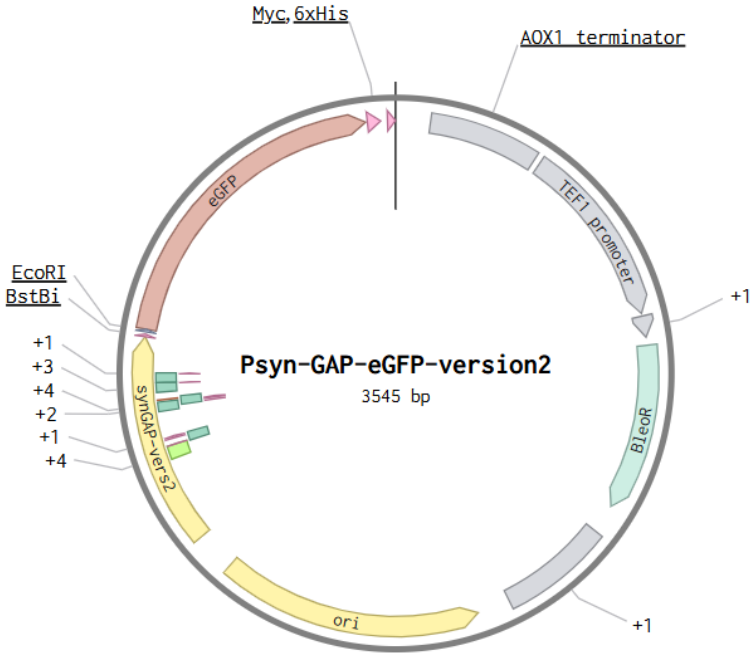

Figure C. 4: The schematic illustration of Psyn-GAP-eGFP-version2 plasmid

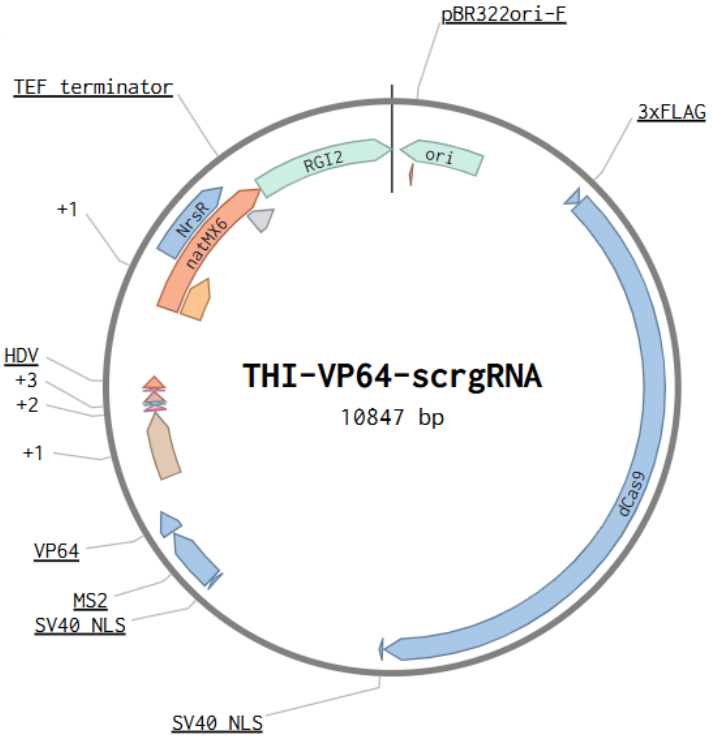

Figure C. 5: The schematic illustration of THI-VP64-scrgRNA plasmid

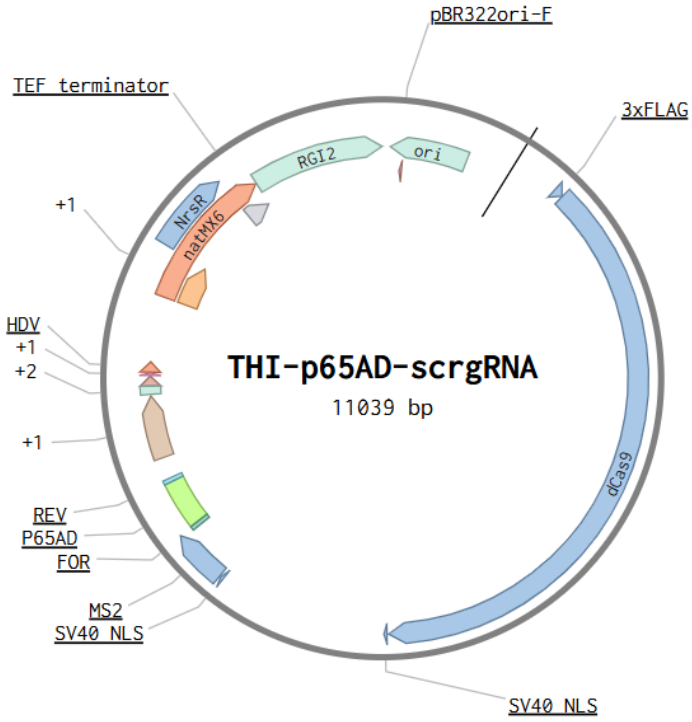

Figure C. 6: The schematic illustration of THI-p65AD-scrgRNA plasmid

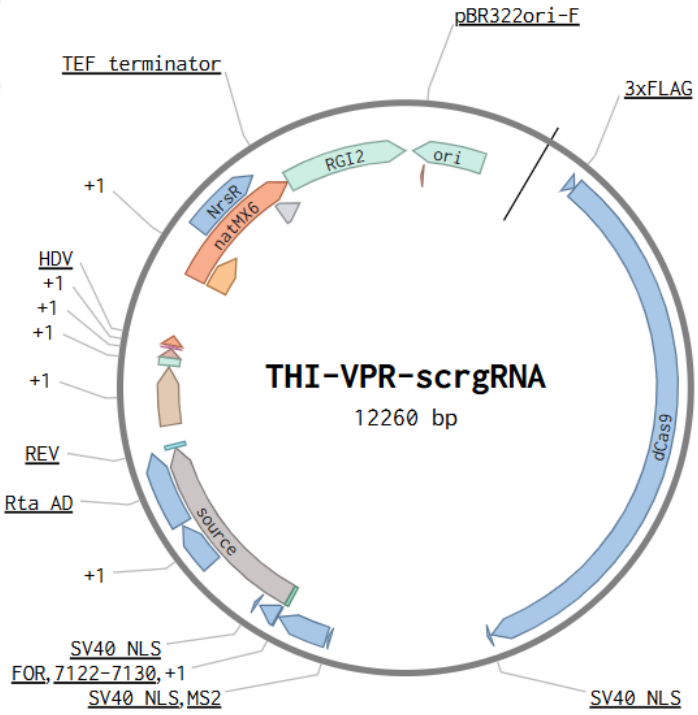

Figure C. 7: The schematic illustration of THI-VPR-scrgRNA plasmid

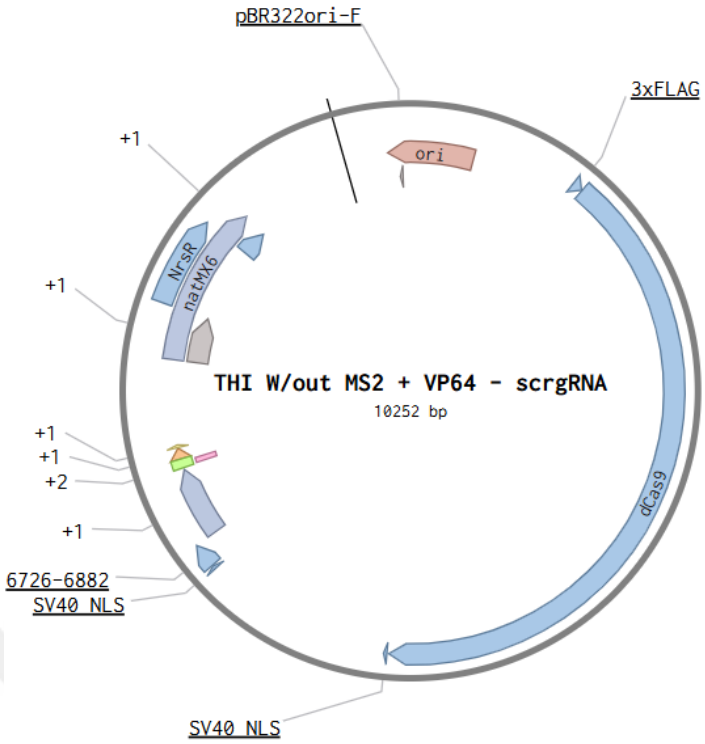

Figure C. 8: The schematic illustration of THI W/out MS2 + VP64 – scrgRNA plasmid

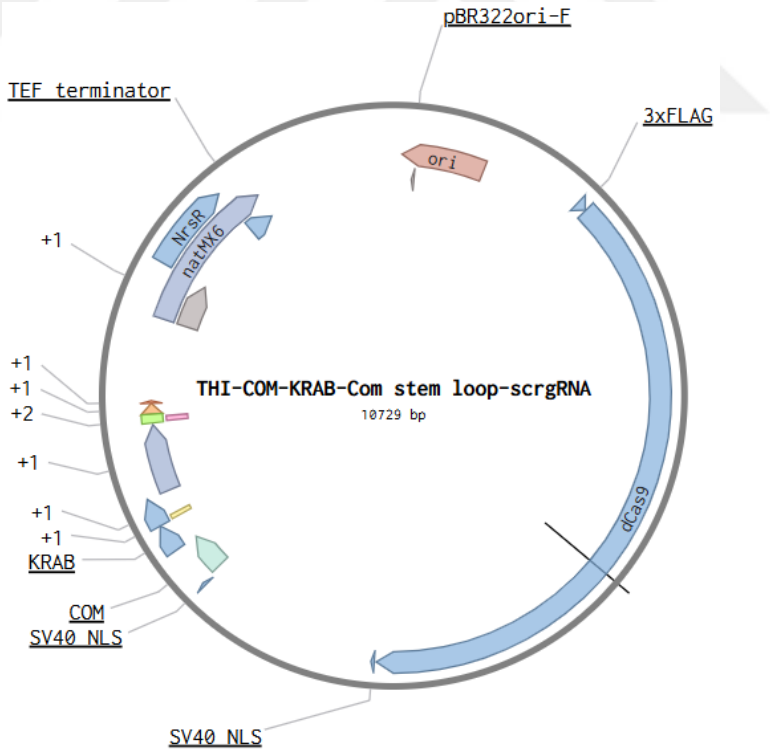

Figure C. 9: The schematic illustration of THI-COM-KRAB-Com stem loop-scrgRNA plasmid

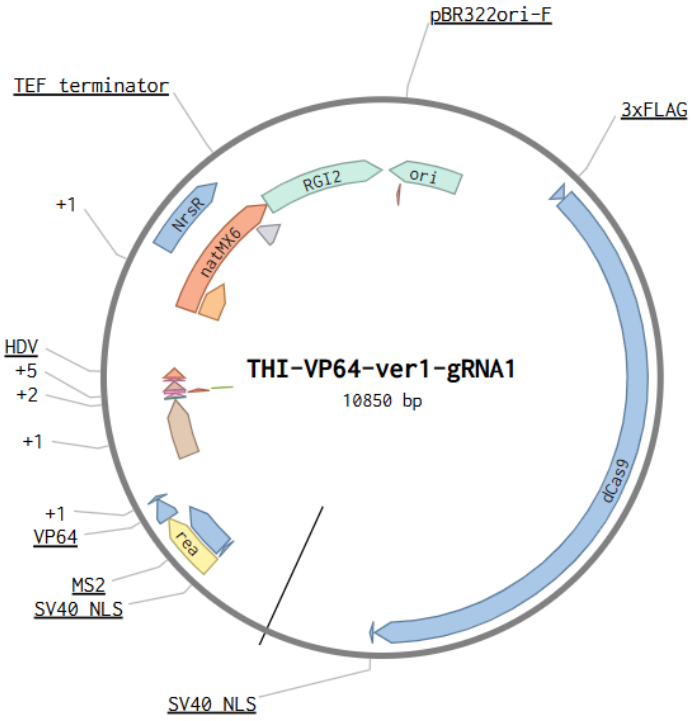

Figure C. 10: The schematic illustration of THI-VP64-ver1-gRNA1 plasmid

Since the backbone plasmid for other gRNAs is the same, only the gRNA regions are shown schematically.

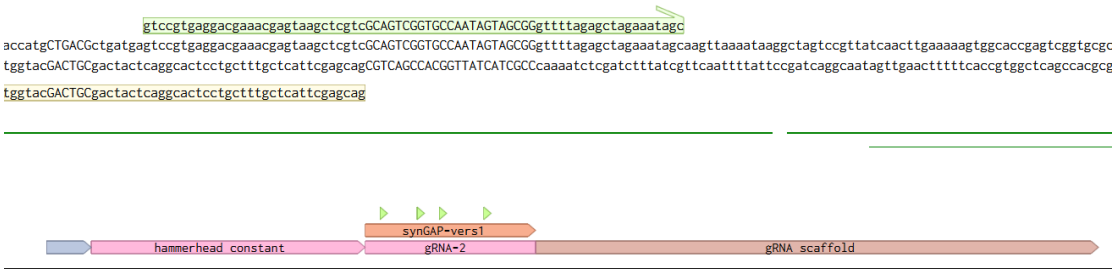

Figure C. 11: The schematic illustration of THI-VP64-ver1-gRNA2

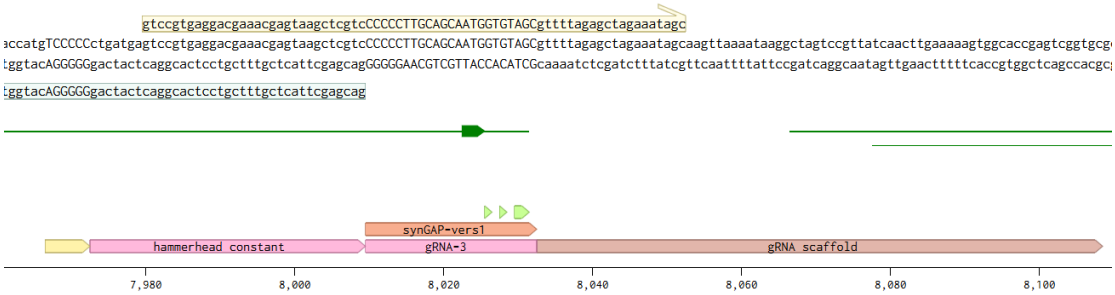

Figure C. 12: The schematic illustration of THI-VP64-ver1-gRNA3

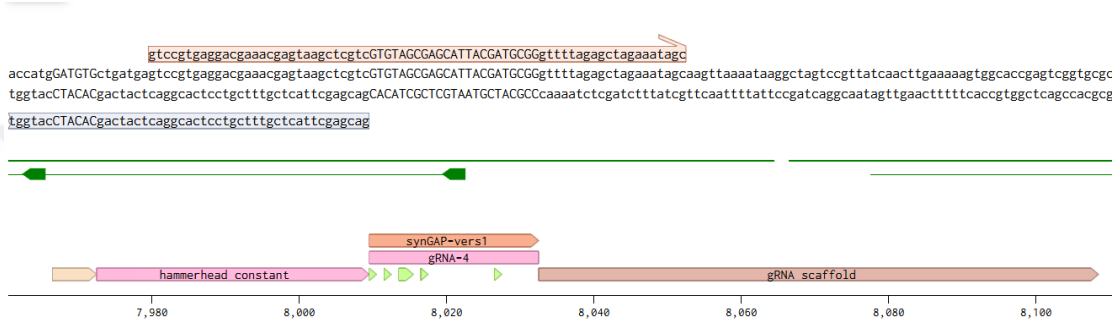

Figure C. 13: The schematic illustration of THI-VP64-ver1-gRNA4

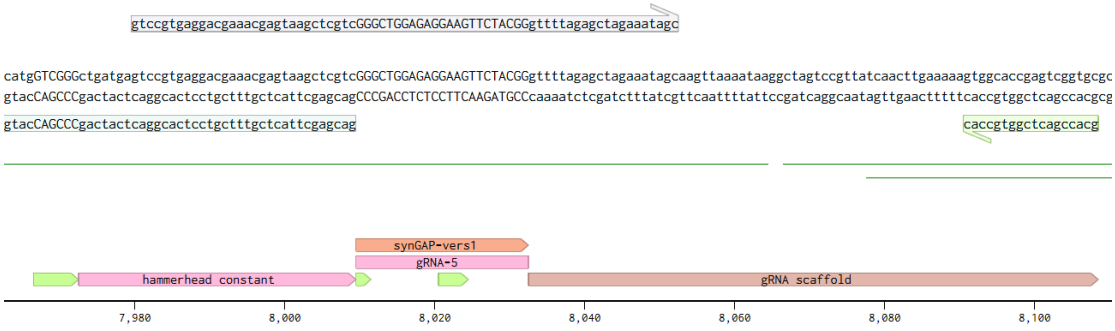

Figure C. 14: The schematic illustration of THI-VP64-ver1-gRNA5

Figure C. 15: The schematic illustration of THI-VP64-ver2-gRNA1

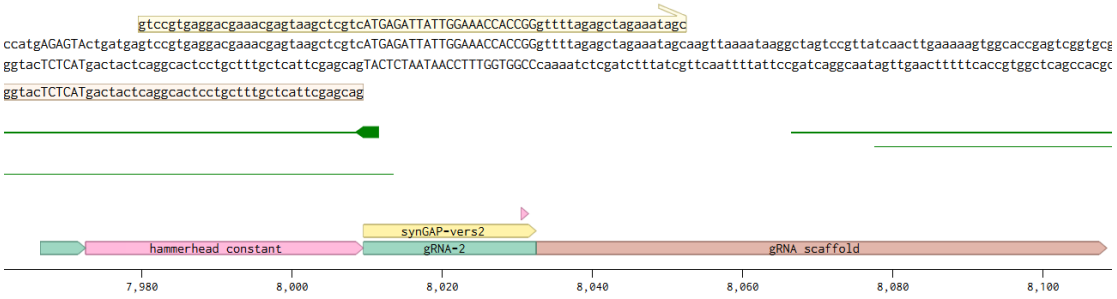

Figure C. 16: The schematic illustration of THI-VP64-ver2-gRNA2

Figure C. 17: The schematic illustration of THI-VP64-ver2-gRNA3

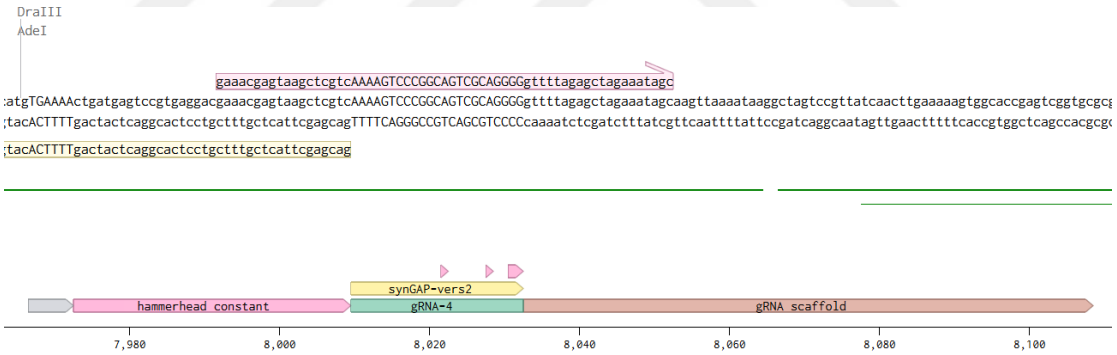

Figure C. 18: The schematic illustration of THI-VP64-ver2-gRNA4

Figure C. 19: The schematic illustration of THI-VP64-ver2-gRNA5

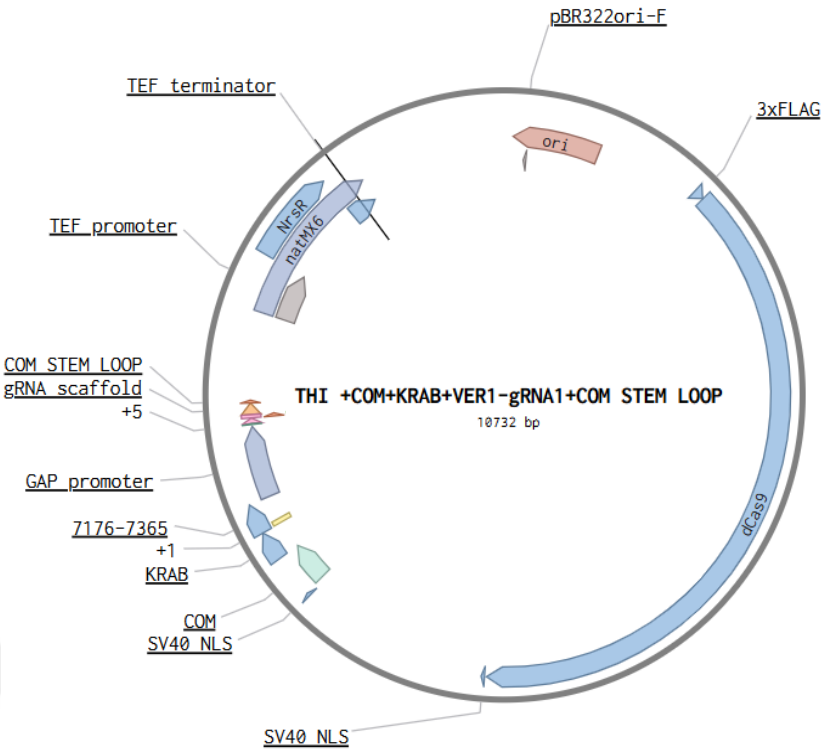

Figure C. 20: The schematic illustration of THI+COM+KRAB+VER1-gRNA1+Com Stem loop plasmid

The sequences containing other gRNAs for the COM-KRAB module are the same as those in the VP64 module.

APPENDIX D

SEQUENCING RESULTS

For sequence analysis, Sanger sequencing was used. The consensus sequence obtained from Sanger results was uploaded to Benchling in order to compare the plasmid sequence's matching status and make it easier to identify mismatched bases.

Figure D. 1: Sequencing result of Psyn-GAP-eGFP-Version1 plasmid

Figure D. 2: Sequencing result of Psyn-GAP-eGFP-Version2 plasmid

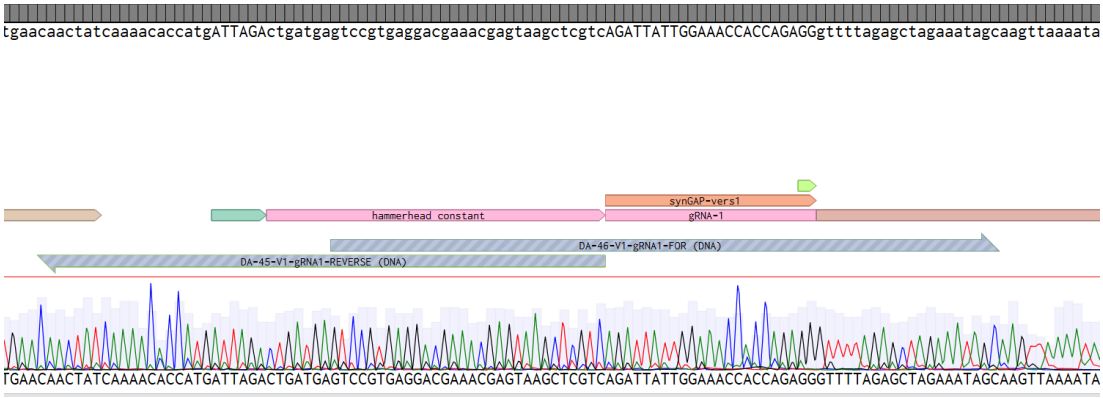

Figure D. 3: Sequencing result of THI-VP64-Ver1-gRNA1 plasmid

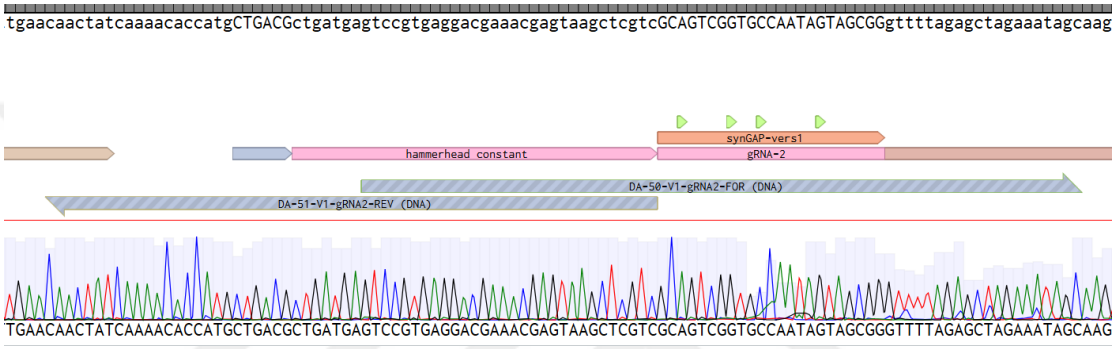

Figure D. 4: Sequencing result of THI-VP64-Ver1-gRNA2 plasmid

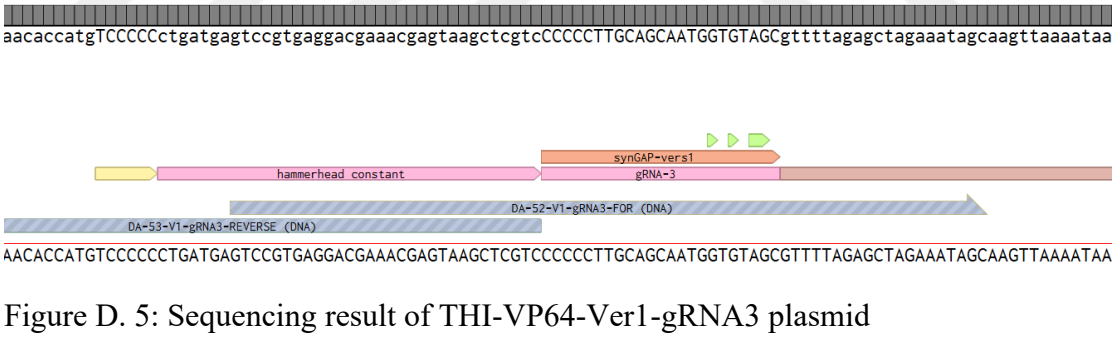

Figure D. 5: Sequencing result of THI-VP64-Ver1-gRNA3 plasmid

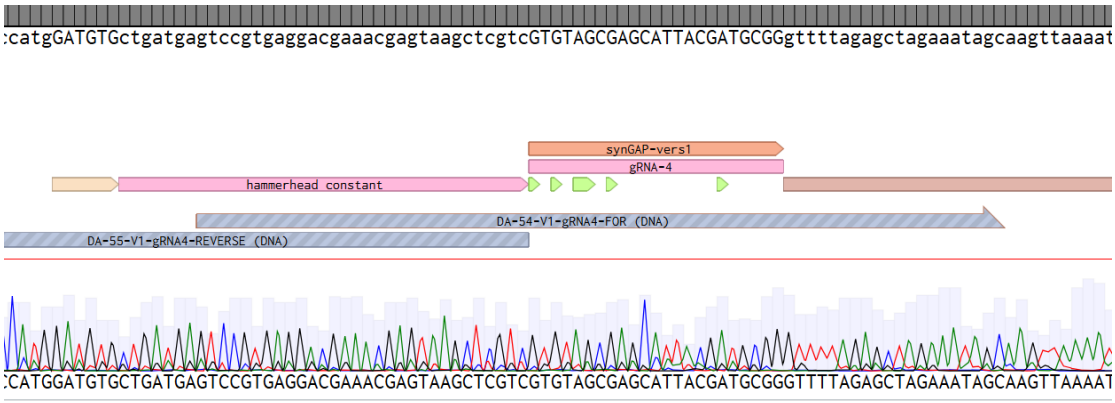

Figure D. 6: Sequencing result of THI-VP64-Ver1-gRNA4 plasmid

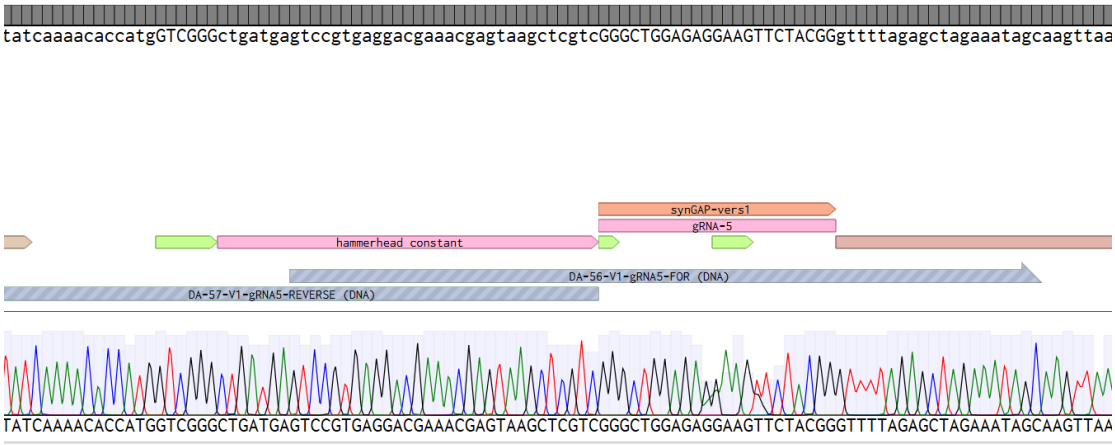

Figure D. 7: Sequencing result of THI-VP64-Ver1-gRNA5 plasmid

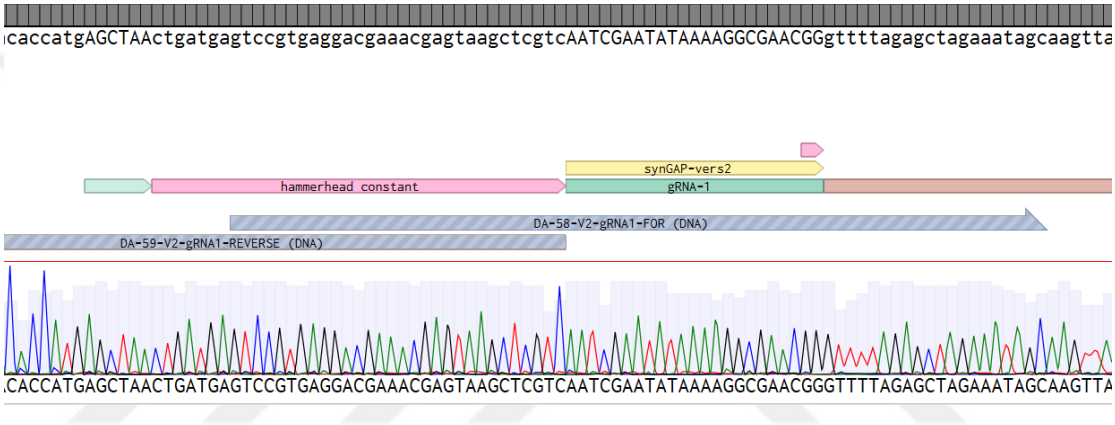

Figure D. 8: Sequencing result of THI-VP64-Ver2-gRNA1 plasmid

Figure D. 9: Sequencing result of THI-VP64-Ver2-gRNA2 plasmid

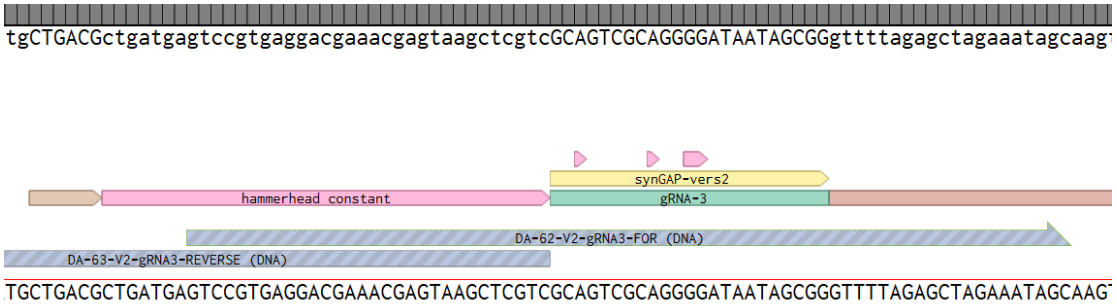

Figure D. 10: Sequencing result of THI-VP64-Ver2-gRNA3 plasmid

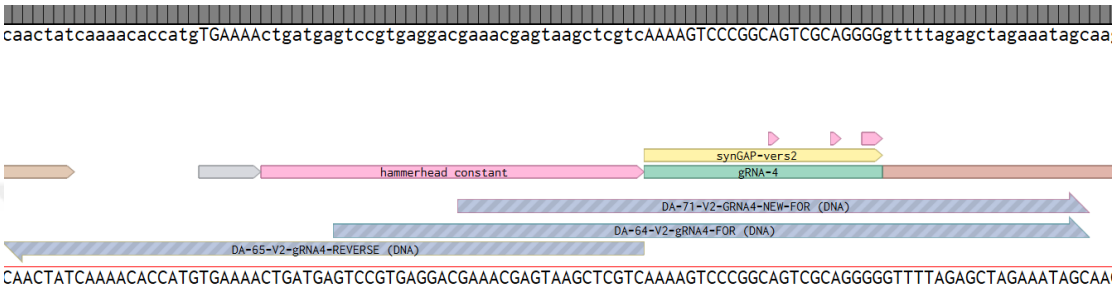

Figure D. 11: Sequencing result of THI-VP64-Ver2-gRNA4 plasmid

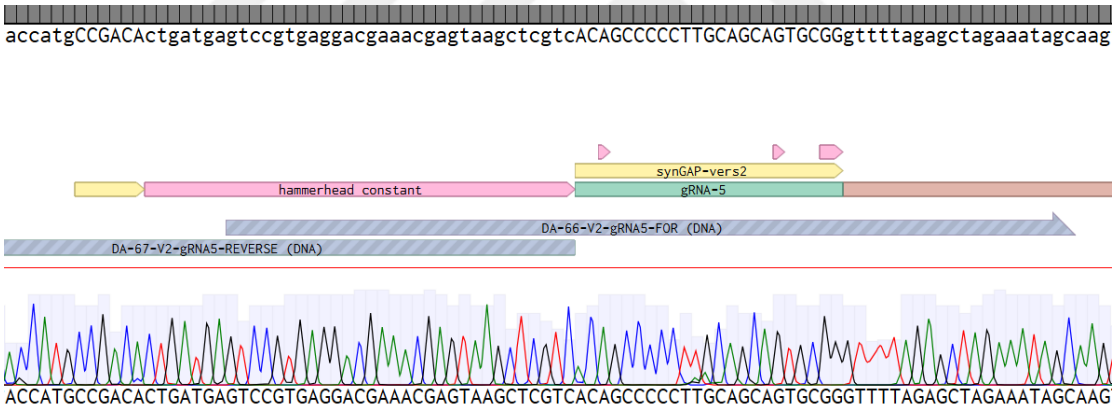

Figure D. 12: Sequencing result of THI-VP64-Ver2-gRNA5 plasmid

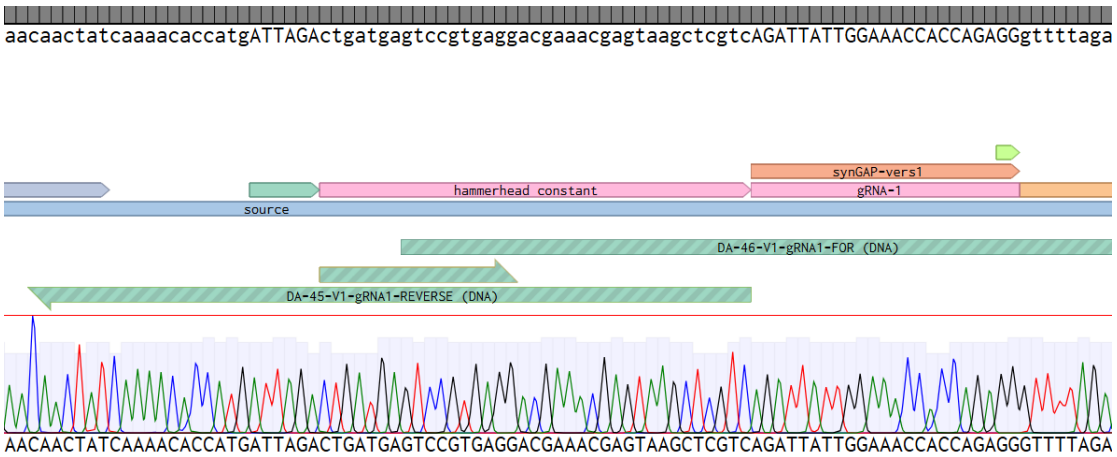

Figure D. 13: Sequencing result of THI-COM-KRAB-Ver1-gRNA1 plasmid

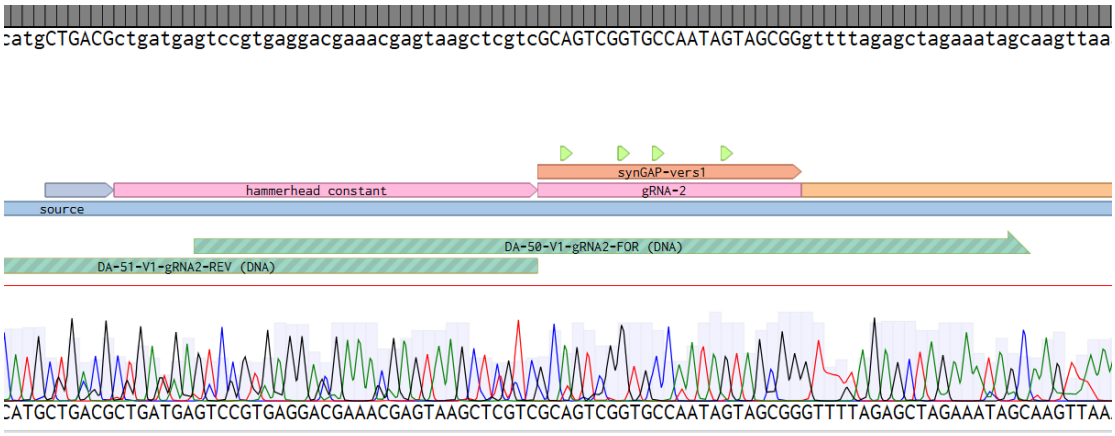

Figure D. 14: Sequencing result of THI-COM-KRAB-Ver1-gRNA2 plasmid

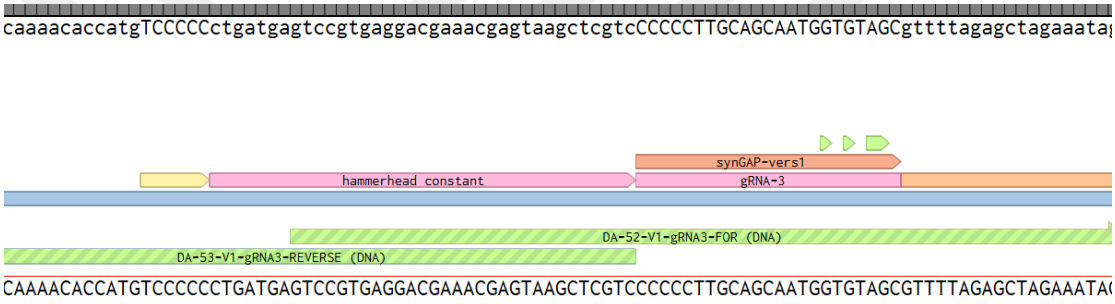

Figure D. 15: Sequencing result of THI-COM-KRAB-Ver1-gRNA3 plasmid

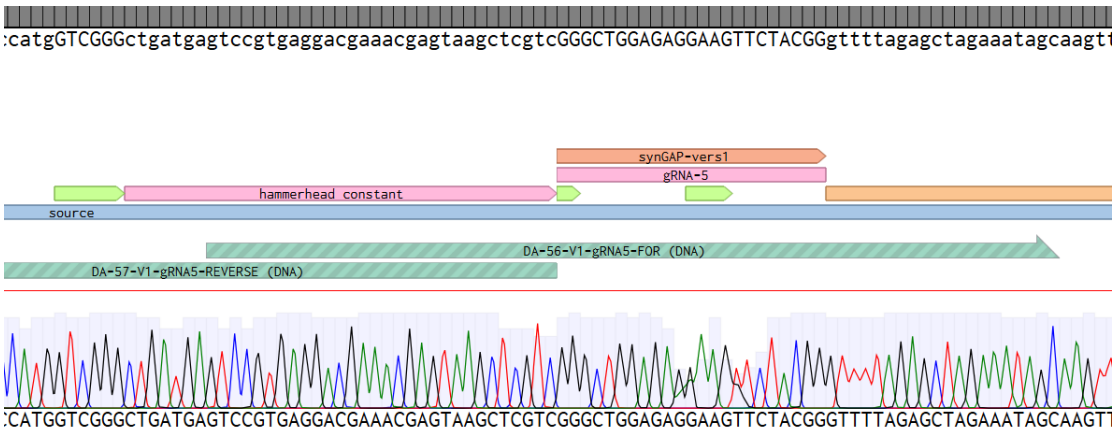

Figure D. 16: Sequencing result of THI-COM-KRAB-Ver1-gRNA5 plasmid

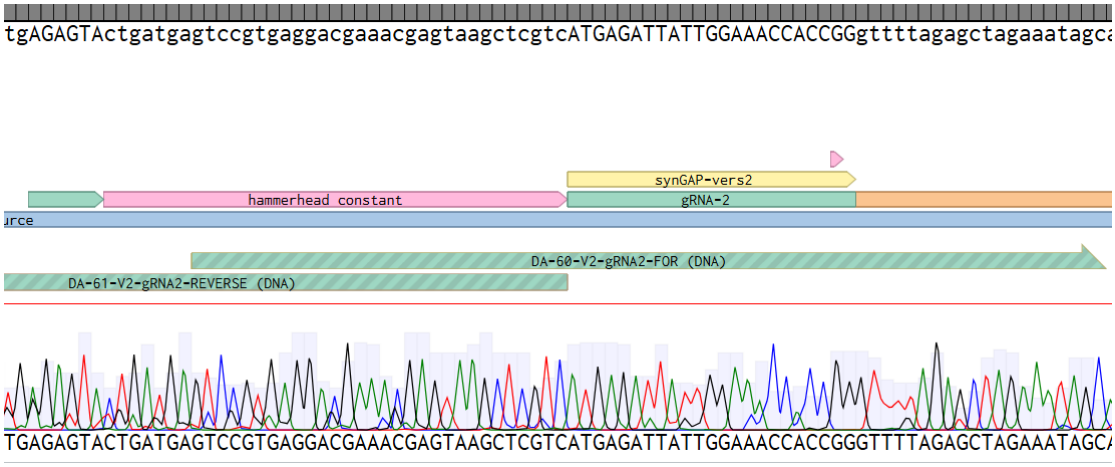

Figure D. 17: Sequencing result of THI-COM-KRAB-Ver2-gRNA2 plasmid

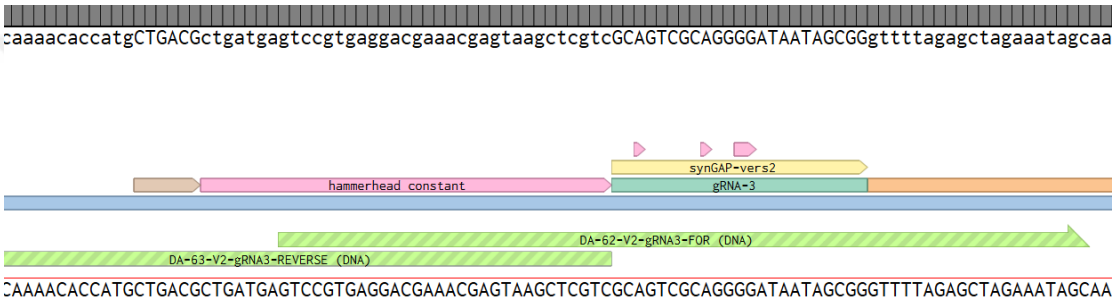

Figure D. 18: Sequencing result of THI-COM-KRAB-Ver2-gRNA3 plasmid

Sequence analysis was done using next-generation sequencing (NGS). In order to compare the matching status of the plasmid sequence, the consensus sequence obtained from the NGS data was aligned using the Geneious Application and then uploaded to Benchling.

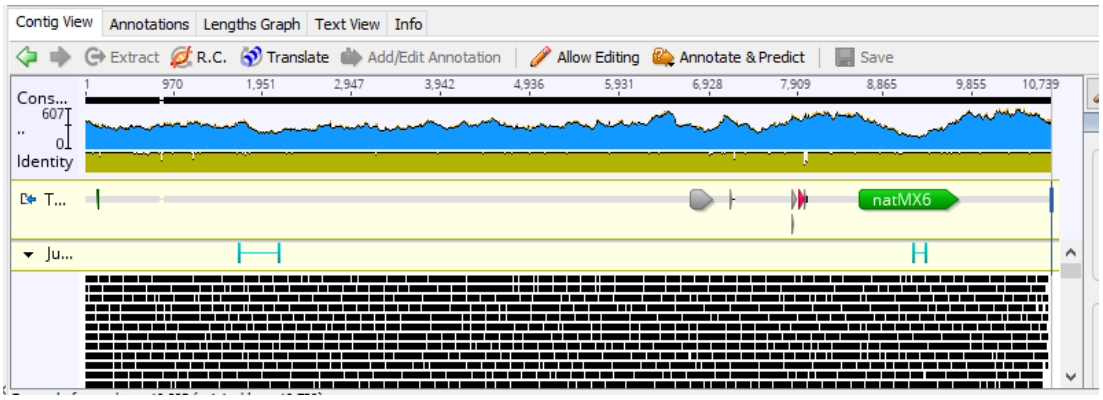

Figure D. 19: Sequencing result of THI-VP64-scrgRNA plasmid

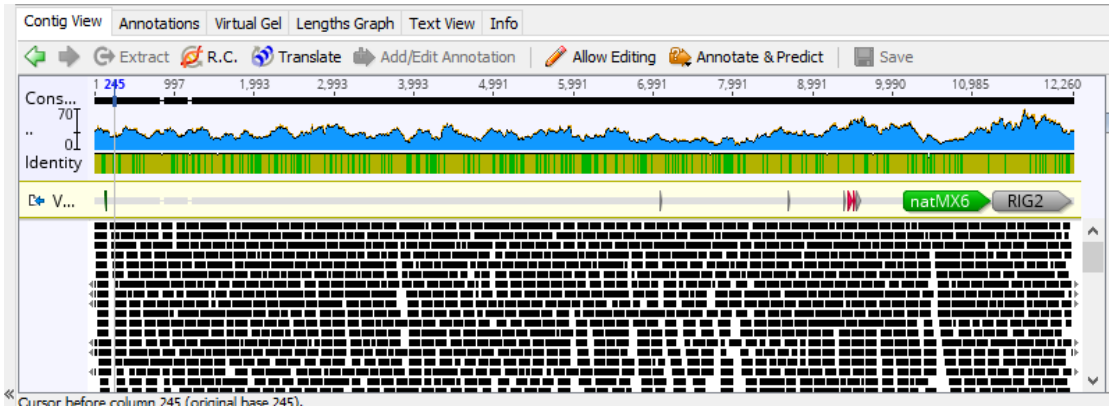

Figure D. 20: Sequencing result of THI-p65AD-scrgRNA plasmid

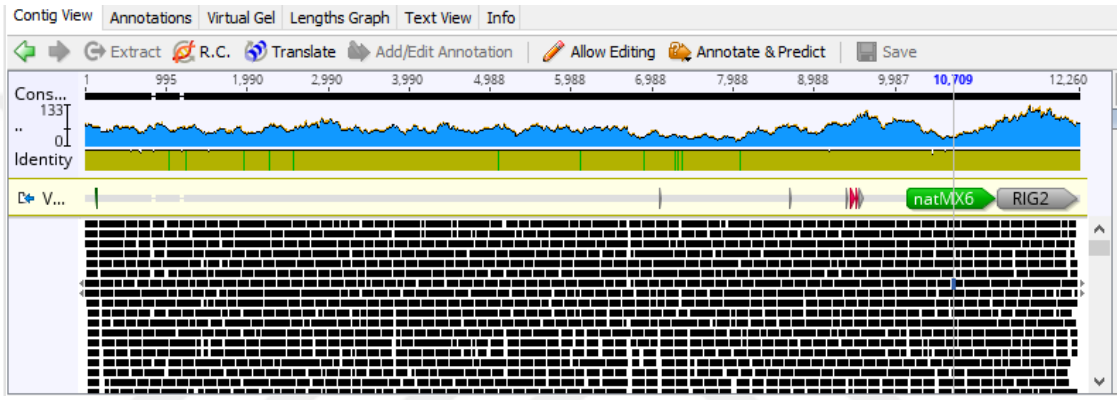

Figure D. 21: Sequencing result of THI-VPR-scrgRNA plasmid

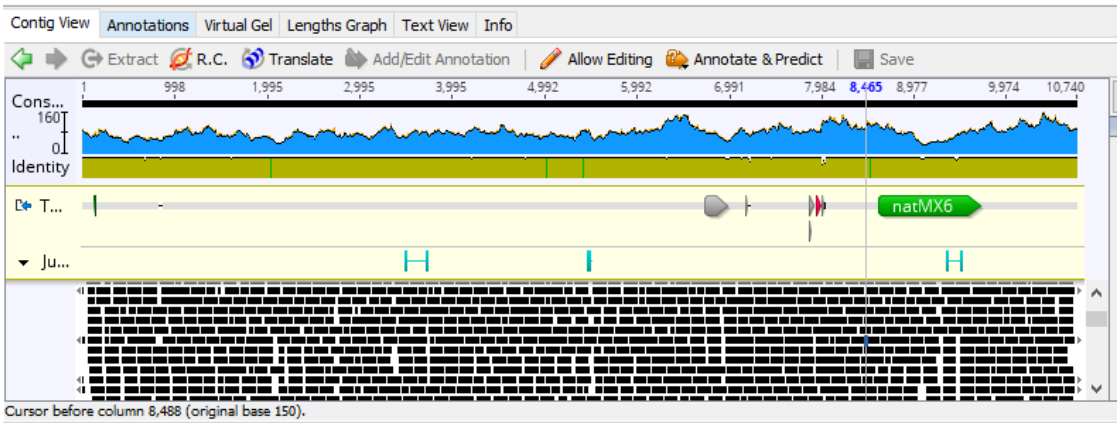

Figure D. 22: Sequencing result of THI-COM-KRAB-Com Stem Loop-scrgRNA plasmid

APPENDIX E

ADDITIONAL REACTION RECIPES

Table E. 1 Lysogeny Broth (LB) Growth Medium and LB Agar Recipe

Component	**Amount**
Yeast Extract	5 g
Sodium Chloride (NaCl)	10 g
Tryptone	10 g
Agar (for LB agar)	15 g
Water	1000 ml

Table E. 2 Yeast Extract Peptone Dextrose (YPD) Growth Medium and YPD Agar Recipe

Component	**Amount**
Yeast Extract	10 g
D-Glucose	20 g
Tryptone	20 g
Agar (for YPD agar)	15 g
Water	1000 ml

Table E.3 TSS Buffer Recipe

Component	**Amount**
PEG 8000	5 g
1M $MgCl_2$	1.5 ml
2.5 ml DMSO	2.5 ml
LB Medium	Up to 50 ml

Table E. 4 Glycerol stock recipe for bacteria sample for -80 degrees

Component	**Amount**
50% Glycerol stock	500 µl
Bacterial sample stock	500 µl

Table E. 5 Glycerol stock recipe for yeast sample for -80 degrees

Component	**Amount**
20% Glycerol + YPD stock	1000 µl
Yeast colony	A single colony is taken and dissolved in solution.

Table E. 6 Single Digestion Reaction

Component	**Amount (for 50 µl)**
10x Cut Smart Buffer	5 µl
DNA	2.5 µg
Restriction Enzyme 1	1 µl
Nuclease Free Water	Up to 50 µl

Table E. 7 Double Digestion Reaction

Component	**Amount (for 50 µl)**
10x Cut Smart Buffer	5 µl
DNA	2.5 µg
Restriction Enzyme 1	1 µl
Restriction Enzyme 2	1 µl
Nuclease Free Water	Up to 50 µl

Table E. 8 Gibson Assembly Mixture (1.33X)

Component	**Amount**
Taq Ligase (40 U/µl)	50 µl
5X Isothermal Buffer	100 µl
T5 Exonuclease (1 U/µl)	2 µl
Phusion Polymerase (2 U/µl)	6.25 µl
Nuclease Free Water	216.75 µl

Table E. 9 PCR reaction mix recipe with Q5 DNA Polymerase

Component	**Amount (for 25 µl)**
5X Q5 Reaction Buffer	5 µl
10 mM dNTP	0.5 µl
10 µM Forward Primer	1.25 µl
10 µM Reverse Primer	1.25 µl
Template DNA	0.5 ng – 0.5 µg

Q5 DNA Polymerase	0.5 µl
Nuclease Free Water	Up to 25 µl

Table E. 10 PCR Reaction Conditions of Q5 DNA Polymerase

Step	**Temperature**	**Duration**
Initial Denaturation	98 °C	30 seconds
Denaturation	98 °C	5 - 10 seconds
Annealing	50 - 72 °C	10 - 30 seconds
Extension	72 °C	30 seconds / kb
Final Extension	72 °C	2 minutes
Hold	4 °C	∞

Table E. 11 PCR reaction mix recipe with TAQ DNA Polymerase

Component	**Amount (for 25 µl)**
10X TAQ Reaction Buffer	2,5 µl
10 mM dNTP	0.5 µl
10 µM Forward Primer	0.5 µl
10 µM Reverse Primer	0.5 µl
Template DNA	0.5 ng – 0.5 µg
Q5 DNA Polymerase	0.3 µl
Nuclease Free Water	Up to 25 µl

Table E. 12 PCR Reaction Conditions of TAQ DNA Polymerase

Step	**Temperature**	**Duration**
Initial Denaturation	95 °C	3 minutes
Denaturation	95 °C	30 seconds
Annealing	45 - 68 °C	10 - 30 seconds
Extension	72 °C	1 minute / kb
Final Extension	72 °C	15 minutes
Hold	4 °C	∞

Table E. 23 2X ASMv6 Screening Medium Recipe

Component	**Amount**
Diammonium-H-Phosphate	6.3 g
Ammonium Sulphate	0.8 g
Magnesium Sulphate- Heptahydrate	0.49 g
Potassium Chloride	2.64 g
Calcium Chloride- Dihydrate	0.0535 g
Citric Acid Monohydrate	22 g

25% NH_4OH	20 mL
Water	Up to 500 mL

Table E. 34 Excess Glucose Growing Medium Recipe for Screening Experiment

Component	**Amount (1 mL)**
2X ASMv6	0.5 mL
10X Glucose	0.1 mL
PTM1	0.002 mL
Water	0.398 mL

Table E. 45 Excess Glycerol Growing Medium Recipe for Screening Experiment

Component	**Amount (1 mL)**
2X ASMv6	0.5 mL
10X Glycerol	0.2 mL
PTM1	0.002 mL
Water	0.298 mL

Table E. 56 2% Ethanol Growing Medium Recipe for Screening Experiment

Component	**Amount (1 mL)**
2X ASMv6	0.5 mL
Ethanol	0.02 mL
PTM1	0.002 mL
Water	0.478 mL

Table E. 67 2% Methanol Growing Medium Recipe for Screening Experiment

Component	**Amount (1 mL)**
2X ASMv6	0.5 mL
Methanol	0.02 mL
PTM1	0.002 mL
Water	0.478 mL

Table E. 78 PEG MW 3350 (50% w/v) Recipe

Component	**Amount**
PEG 3350	12.5 g
Water	Up to 25 mL

Table E. 89 qPCR Mix Recipe

Component	**Amount**
Luna Universal qPCR Master Mix	10 µl
10 µM Forward Primer	0.5 µl
10 µM Reverse Primer	0.5 µl
Template DNA	< 100 ng
Nuclease-free Water	Up to 20 µl

Table E. 20 qPCR Reaction Conditions

Cycle Step	**Temperature**	**Time**	**Cycles**
Initial Denaturation	95 °C	60 seconds	1
Denaturation	95 °C	15 seconds	40-45
Extension	60 °C	30 seconds (+plate read)	
Melt Curve	60-95 °C	various	1