Expanding CRISPR-based toolbox for enhanced production of β-carotene in *S. cerevisiae*

Dissertation zur Erlangung des akademischen Grades
Doctor rerum naturalium (Dr. rer. nat.)

eingereicht am Fachbereich Chemie der Fakultät für Mathematik, Informatik und Naturwissenschaften der Universität Hamburg

vorgelegt von

Klaudia Ciurkot, M. Sc.

geboren in Tychy (Polen)

Disputation hold on 25th of March 2022

Hamburg

2021

Preface

Die in dieser Dissertation präsentierte Arbeit wurde im Zeitraum von Februar 2018 bis Juni 2021 innerhalb der DSM Food Specialties am DSM Biotechnology Center in Delft, Niederlande, und an der University of Bristol, Großbritannien, durchgeführt. Dieses Projekt wurde von Dr. René Verwaal, Dr. Johannes A. Roubos, Dr. Thomas E. Gorochowski und Prof. Dr. Zoya Ignatova betreut. Dieses Projekt wurde vom Forschungs- und Innovationsprogramm Horizont 2020 der Europäischen Union im Rahmen der Marie-Skłodowska-Curie-Finanzhilfevereinbarung Nr. [764591] gefördert.

The work presented in this dissertation was conducted from February 2018 until June 2021 at the DSM Food Specialties at the DSM Biotechnology Center in Delft, the Netherlands and the University of Bristol, the United Kingdom. The supervision of this project was provided by Dr. René Verwaal, Dr. Johannes A. Roubos, Dr. Thomas E. Gorochowski and Prof. Dr. Zoya Ignatova. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No [764591].

Dissertationsgutachter / dissertation reviewers:

1. Gutachterin / first reviewer:	Prof. Dr. Zoya Ignatova
	Institute of Biochemistry and Molecular Biology, Universität Hamburg
2. Gutachter / second reviewer:	Dr. Thomas E. Gorochowski
	School of Biological Sciences, University of Bristol

Declaration

Hiermit erkläre ich, dass es sich bei dieser Dissertation um meine eigene Arbeit handelt und ich keine anderen als die aufgeführten Quellen verwendet habe.

I hereby declare that this doctoral dissertation is my own work and that I have not used any sources other than those listed.

London, 20 December 2021 Klaudia Ciurkot

Teile dieser Arbeit wurden in folgenden Publikationen veröffentlicht:

Parts of this work were published in the following scientific articles:

Ciurkot, K., Gorochowski, T. E., Roubos, J. A. & Verwaal, R. Efficient multiplexed gene regulation in *Saccharomyces cerevisiae* using dCas12a. *Nucleic Acids Res.* **49**, 7775–7790 (2021) doi:10.1093/nar/gkab529.

Ciurkot, K., Vonk, B., Gorochowski, T. E., Roubos, J. A. & Verwaal, R. CRISPR/Cas12a Multiplex Genome Editing of *Saccharomyces cerevisiae* and the Creation of Yeast Pixel Art. *J. Vis. Exp.* **147**, e59350 (2019) doi:10.3791/59350.

Content

List of Tables ..... **iv**
List of Figures ..... **iii**
Abbreviations ..... **vi**
Zusammenfassung ..... **ix**
Abstract ..... **x**

1 Introduction **1**
1.1 Microbial cell factories 1
Saccharomyces cerevisiae as a host organism 2
Production burden 3
Molecular tools to mitigate production burden 4
1.2 CRISPR in nature and molecular biology 5
Class 2 type V – Cas12a (Cpf1) 8
Class 2 type VI – Cas13 9
CRISPR regulation 10
1.3 β-carotene importance and production 12
1.4 Thesis aim 15

2 Evaluation of growth and carotenoids production by engineered *S. cerevisiae* strains **17**
2.1 Introduction 17
2.2 Materials and methods 18
2.3 Results 20
Cellular growth and metabolism 20
Carotenoids production 22
Colour trait 24
Genetic instability 24
2.4 Discussion 26

3 CRISPR-Cas12a Multiplex Genome Editing of *Saccharomyces cerevisiae* and the Creation of Yeast Pixel Art **28**
3.1 Introduction 28
3.2 Protocol 31
3.3 Results 38
3.4 Discussion 41

4 Construction of carotenogenic strains using Cas12a as a foundation for CRISPRi **43**
4.1 Introduction 43
4.2 Materials and methods 43
4.3 Results 44
Elucidating control node in β-carotene biosynthesis 44
Balancing expression of *crt* genes 46
4.4 Discussion 47

5 Efficient multiplexed gene regulation in *Saccharomyces cerevisiae* using dCas12a **49**
5.1 Introduction 49
5.2 Material and methods 50
5.3 Results 55
Influence of NLS position on dCas12a-based CRISPRi efficiency 55
Evaluation of repression domains 57

eGFP downregulation targeting additional promoters 59
Controllable expression of gRNA from RNAP II promoter 61
Multiplexed downregulation 62
Stringent regulation of heterologous β-carotene production 64
5.4 Discussion 67

6 Exploration of Cas13 ribonucleases for a transcript knockdown 69

6.1 Introduction 69
6.2 Materials and methods 70
6.3 Results 73
Expression of Cas13 proteins in *S. cerevisiae* 73
Collateral effect in *S. cerevisiae* 75
6.4 Discussion 78

7 Biosensing stress induced by carotenoids production 80

7.1 Introduction 80
7.2 Materials and methods 81
7.3 Results 83
Cell response to the production of β-carotene 83
Selection of promoters for biosensing production stress 85
gRNA expression from promoters acting as a biosensor 88
7.4 Discussion 90

8 Feedback control system 92

8.1 Introduction 92
8.2 Materials and methods 93
8.3 Results 95
Construction of the feedback control system 95
Influence of the feedback control system on performance of carotenogenic strains 97
8.4 Discussion 101

9 General discussion 103

9.1 Feedback control system to mitigate burden in β-carotene production 103
Multiplex gene downregulation mediated by dCas12a-CRISPRi 103
Monitoring burden related to the production of β-carotene 104
Burden-driven control of β-carotene production 104
Post-transcriptional pathway control by Cas13-CRISPR 105
Designing the regulatory circuit 105
9.2 Expanding the CRISPR-based toolbox to optimize *S. cerevisiae* as a production chassis 106

10 References 107
11 List of the hazardous substances used by GHS 11-1
12 Acknowledgments 12-2
13 Supplementary Information 13-4

List of Tables

Table 1-1. Main features of selected class 2 CRISPR-Cas systems 8
Table 2-1 Strains characterized in Chapter 2 19
Table 3-1. Editing efficiency of the multiplex genome editing approach 39
Table 4-1. Promoters used for expression of *crt* genes 44
Table 4-2. Promoters' strengths selected for expression of *crt* genes 45
Table 5-1. Comparison of dCas12a efficiency when expressed from plasmid or integrated into the genome 58
Table 6-1. Cas13 expression cassettes in strains used in Chapter 6 71
Table 6-2. Abundance of ribosomal RNA (18S and 25S) under inducible and repressive conditions for LwaCas13a NLS expression 76
Table 7-1. Genotype of engineered strains used for testing biosensing promoters 82
Table 7-2. Normalized number of transcripts measured in carotenoids producing strain CAR-034 and wild type strain 86
Table 8-1. Genotype of strains tested with the feedback control system 93
Table 8-2. Plasmids used in Chapter 8 95
Table 8-3. Carotenoids production by strains CAR-002 and CAR-034 after 48 h of fermentation in rich and mineral medium 100

List of Figures

Figure 1-1. Adaptive immunity of the CRISPR-Cas system 6
Figure 1-2. gRNA architecture of CRISPR-Cas9, CRISPR-Cas12a and CRISPR-Cas13 systems 7
Figure 1-3. Carotenoids biosynthesis in *X. dendrorhous* 13
Figure 1-4. Concept of the feedback control system for dynamic regulation of β-carotene production. 16
Figure 2-1. β-carotene synthesis in *S. cerevisiae* by overexpression of *crtE*, *crtYB* and *crtI* from *X. dendrorhous* 17
Figure 2-2. Growth of carotenogenic strains during the batch fermentation 21
Figure 2-3. Ethanol and glucose throughout fermentation in mineral medium for the wild type strain and CAR-002, CAR-020 and CAR-021 22
Figure 2-4. Total carotenoids production by *S. cerevisiae* strains in mineral medium and rich medium 23
Figure 2-5. Phytoene, lycopene and β-carotene production profiles by strains CAR-002 and CAR-020 23
Figure 2-6. Genetic instability of carotenogenic strains 25
Figure 2-7. Mutations in *crt* expression cassettes of strains CAR-002 and CAR-020 26
Figure 3-1. Workflow of the protocol for CRISPR-Cas12a multiplex genome editing in *S. cerevisiae* 29
Figure 3-2. Scheme of CRISPR-Cas12a multiplex genome editing using a single crRNA array 30
Figure 3-3. PCR verifying the genome editing experiments 39
Figure 3-4. Plates of *S. cerevisiae* transformations using the multiplex genome editing approach. 39
Figure 3-5. PCR verifying integration of the donor DNA expression cassettes at the intended loci within the genomic DNA 40
Figure 3-6. Yeast pixel art of Rosalind Franklin 40
Figure 4-1. Carotenogenic *S. cerevisiae* strains constructed with CRISPR-Cas12a 45
Figure 4-2. Production of total carotenoids by constructed strains 45
Figure 4-3. Carotenoids production profile of constructed strains 46
Figure 5-1. Effect of NLS number and position in dCas12a E925A Mxi1 fusions 56
Figure 5-2. Comparison of repression domains and mutations defining nuclease-deficiency in dCas12a 58
Figure 5-3. Downregulation of *eGFP* expression by targeting dCas12a E925A Mxi1 to promoters controlling *eGFP* or *eGFP* ORF 60
Figure 5-4. gRNA expression from RNAP III promoter and RNAP II promoter 62
Figure 5-5. Simultaneous downregulation of fluorescent proteins with dCas12a E925A NLS Mxi1 and a single crRNA array 63
Figure 5-6. Simultaneous downregulation of β-carotene production with dCas12a E925A NLS Mxi1 and single crRNA array 66
Figure 6-1. Implementation of CRISPR-Cas13 systems in *S. cerevisiae* 74
Figure 6-2. Collateral effect of CRISPR-LwaCas13a NLS in *S. cerevisiae* 75
Figure 6-3. Lethality of CRISPR-LwaCas13a system in *S. cerevisiae*. 77
Figure 6-4. Conceptualization of CRISPR-Cas13a based growth control device 79
Figure 7-1. Gene upregulation in strain producing β-carotene in comparison with the wild type *S. cerevisiae* strain 83

Figure 7-2. Impact of β-carotene production on ergosterol biosynthesis superpathway in *S. cerevisiae*.....84
Figure 7-3. Volcano plots illustrating gene upregulation induced by production of carotenoids after 6 h and 24 h of cultivation.....86
Figure 7-4. Activity test for promoters with biosensing capacity.....87
Figure 7-5. CRISPR interference of *RFP* with gRNAs driven by *ERG3* and *PDR5* promoters.....88
Figure 7-6. Tunability of biosensing promoters evaluated by CRISPRi mediated downregulation of *RFP*...89
Figure 8-1. CRISPRi-based feedback control system for carotenoids downregulation in *S. cerevisiae*.....96
Figure 8-2. Growth experiment evaluating feedback control system.....97
Figure 8-3. Production of carotenoids in production strain CAR-034 and strain CAR-042 with the feedback control system.....98
Figure 8-4. Stability of β-carotene biosynthesis pathway in strain with (CAR-042) and without the PDR5p feedback control system (CAR-034) after cultivation in mineral medium for 73 h. 100

Abbreviations

ABC	ATP-binding cassette (domain)
BFP	blue fluorescent protein
bp	base pair
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRISPRa	CRISPR activation
CRISPRi	CRISPR interference
crRNA	CRISPR RNA
DMAP	dimethylallyl diphosphate
dsDNA	double stranded DNA
eGFP	enhanced green fluorescent protein
FACS	fluorescence-activated cell sorting
FCS	feedback control system
FPP	farnesyl pyrophosphate
G418	geneticin
GGPP	geranylgeranyl pyrophosphate
HEPN	Higher Eukaryotes and Prokaryotes Nucleotide-binding (domain)
IPP	iso-pentenyl diphosphate
gRNA	guide RNA
KRAB	Krüppel associated box (domain)
mRNA	messenger RNA
NLS	nuclear localization signal
nt	nucleotide
NTC	nourseothricin
OD	optical density
ORF	open reading frame
PAM	protospacer adjacent motif
PFS	protospacer flanking site
qPCR	quantitative polymerase chain reaction
RFP	red fluorescent protein
RNAP	RNA polymerase
RNA-seq	RNA sequencing

rRNA	ribosomal RNA
sgRNA	single guide RNA
snoRNA	small nucleolar RNA
ssDNA	single stranded DNA
ssRNA	single stranded RNA
TPM	transcripts per million
tracrRNA	trans-activating CRISPR RNA
tRNA	transfer RNA
UPR	unfolded protein response
WT	wild type

Zusammenfassung

β-Carotin ist ein natürliches Pigment mit antioxidativen Eigenschaften, das in der Lebens- und Futtermittelproduktion, der Pharma- und Schönheitsindustrie breite Verwendung findet. Die Marktnachfrage nach β-Carotin steigt stetig, was die Entwicklung neuer Produktionsverfahren wie der Biosynthese in mikrobiellen Zellfabriken motiviert, die eine nachhaltige Alternative zur chemischen Synthese darstellt. Allerdings führt die β-Carotin-Biosynthese im Modellwirtsorganismus, *Saccharomyces cerevisiae*, zu einem Ungleichgewicht in Stoffwechselwegen (z.B. Ergosterol-Biosynthese), veränderter Morphologie (z.B. Zusammensetzung der Zellmembran) und einer Aktivierung der Stressreaktion (z.B. pleiotrope Wirkstoffresistenztransporter), was eine Hemmung des Zellwachstums und eine Instabilität der Gene des Produktionswegs zur Folge hat.

Diese Dissertation untersucht die CRISPR-Systeme als neuartige molekulare Werkzeuge zur Optimierung der β-Carotin-Biosynthese in *S. cerevisiae* und zur Minderung der damit verbundenen Produktionsbelastung. Hierbei wird das Protokoll für die Multiplex-Manipulation des Hefegenoms - unter Verwendung des CRISPR-Cas12a-Systems - in Kombination mit crRNA-Arrays eingesetzt, um eine schnelle Konstruktion von Carotinoiden produzierenden Stämmen zu ermöglichen. Der Nuklease-defiziente CRISPR-dCas12a-Effektor wurde dabei zur gezielten Herunterregulierung der Genexpression, der sogenannten CRISPR-Interferenz, eingesetzt. Dies wurde durch eine systematische Untersuchung der wichtigsten Designmerkmale für eine effiziente dCas12a-vermittelte Gen-Herunterregulierung, wie Anzahl und Position des nuklearen Lokalisierungssignals, Fusion an eine zusätzliche Repressionsdomäne, Position des Targets und Architektur der gRNA, unterstützt. Das optimierte dCas12a-System wurde erfolgreich für die Multiplex-Herunterregulierung der β-Carotin-Produktion eingesetzt.

Um die Belastung durch die Produktion von β-Carotin in *S. cerevisiae* zu mindern, wurde das CRISPRi-dCas12a-System verwendet, um ein Feedback-Kontrollsystem (FCS) zu entwickeln. Das FCS regelt bei auftretender Produktionsbelastung die β-Carotin-Produktion temporär herunter, damit die Zelle ihre Fitness wiederherstellen kann. Um eine dynamische Regulation zu erreichen, wurde gRNA von *ERG3*- oder *PDR5*-Promotoren exprimiert. Diese Promotoren fördern die Expression von Genen, die bei der Carotinogenese hochreguliert sind, wie mittels RNA-Sequenzierung gezeigt werde konnte. Die Carotinoide produzierenden Stämme - mit und ohne Feedback-Kontrollsystem - wurden dann hinsichtlich Wachstum, Carotinoid-Produktion und genetischer Stabilität der Expressionskassetten, die Gene kodieren, die für die β-Carotin-Biosynthese verantwortlich sind, verglichen. Obwohl gezeigt wurde, dass FCS die β-Carotin-Produktion effektiv unterdrückt, ist eine weitere Optimierung des Systems erforderlich, um dem Wirtsorganismus ein synthetisches Regulationssystem bereitzustellen, das verbessertes Wachstum, genetische Stabilität und folglich eine höhere Produktausbeute zeigen würde.

Schließlich wurden die CRISPR-Cas13a- und -Cas13b-Systeme auf spezifische Transkript-Knockdowns in *S. cerevisiae* untersucht. Die Aktivierung von Cas13 hatte jedoch einen Kollateraleffekt (unspezifische Spaltung der in der Zelle vorhandenen RNA) zur Folge, der das Zellwachstum reversibel blockiert, allerdings nicht zum Zelltod führt. Basierend auf diesen Erkenntnissen wurde vorgeschlagen, das CRISPR-Cas13a-System als molekulares Werkzeug zur Kontrolle des Wachstums von *S. cerevisiae* zu verwenden.

Abstract

β-carotene is a natural pigment with antioxidant properties, widely used in food and animal feed production, pharmaceutical and beauty industries. The market demand for β-carotene is constantly rising which motivates development of new production processes such as biosynthesis in microbial cell factories, a sustainable alternative for the chemical synthesis. However, β-carotene biosynthesis in the model host organism – *Saccharomyces cerevisiae* leads to an imbalance in metabolic pathways (*e.g*. ergosterol biosynthesis), altered morphology (*e.g.* composition of the cellular membrane), activation of stress response (*i.e.* pleiotropic drug resistance transporters) and consequently hampered cell growth and instability of the production pathway genes.

This thesis explores the CRISPR systems as novel molecular tools for the optimization of β-carotene biosynthesis in *S. cerevisiae* and mitigation of the related production burden. The protocol for multiplex manipulation of the yeast genome using the CRISPR-Cas12a system in combination with crRNA arrays was proposed to enable fast construction of carotenoids-producing strains. The nuclease deficient CRISPR-dCas12a effector was then implemented for specific downregulation of gene expression, so-called CRISPR interference. Development of CRISPRi system was supported by a systematic study of the key design features for efficient dCas12a mediated gene downregulation, such as the amount and position of the nuclear location signal, fusion to an additional repression domain, position of the target and architecture of the gRNA. The optimized dCas12a system was successfully applied for multiplex downregulation of β-carotene production.

To mitigate the burden caused by the production of β-carotene in *S. cerevisiae*, the CRISPRi-dCas12a system was used to develop a feedback control system (FCS). In principle, the FCS temporarily downregulates β-carotene production when the production burden occurs so that the cell can restore its fitness. To achieve dynamic regulation, gRNA was expressed from *ERG3* or *PDR5* promoters. According to the conducted RNA sequencing, these promoters drive expression of genes which are highly upregulated upon carotenogenesis. The carotenoids producing strains with and without the feedback control system were then compared in terms of growth, carotenoids production and genetic stability of the expression cassettes encoding genes responsible for β-carotene biosynthesis. Although the FCS was shown to effectively repress β-carotene production, further optimization of the system is required to provide the host organism with the synthetic regulatory system which would exhibit improved growth, genetic stability and consequently higher product yield.

Finally, CRISPR-Cas13a and -Cas13b systems were explored for specific mRNA transcript knockdown in *S. cerevisiae*. However, activation of Cas13 led to the collateral effect (unspecific cleavage of RNA present in the cell) which was shown to reversibly block cellular growth but did not lead to the cells death. Based on these learnings, CRISPR-Cas13a system was proposed to be used as a molecular tool to control growth of *S. cerevisiae*.

1 Introduction

1.1 Microbial cell factories

Environmental concerns, depletion of petroleum resources and the increasing world population accelerate development of new methods for the production of valuable compounds, including solutions offered by biotechnology. Microorganisms have been used for the production of fermented foods and beverages reaches over 8000 years. In the 20th century the repertoire of the products generated in fermentation of microorganisms was expanded to chemicals (*e.g.* acetone–butanol–ethanol fermentation), food ingredients (*e.g.* citric acid) and pharmaceuticals (*e.g.* penicillin)[1]. However, only with the advances in recombinant DNA technology in the 80's, researchers learnt to harness microorganisms to produce a compound of choice by expressing non-native genes in easily cultivable organisms, so-called microbial cell factories[2,3]. The rapid pace of biotechnology development in the past decades led to the emerge of new fields such as genetic and metabolic engineering, synthetic and systems biology which underpin cell factory advances. For example, using genome editing a heterologous gene can be integrated into the genome of the host organism to ensure stable expression, whereas optimization of a native production pathway can be mediated by selective gene up- or downregulation and gene knock-out based on a metabolic model.

Construction of a cell factory relies on the design-build-test-learn cycle. Firstly, elements of a production pathway (*i.e.* genes and genetic parts required for gene expression) are determined and designed. Next, the required homologous and/or heterologous genes are introduced to a host organism by either integration to the genome or provided as a vector. At this phase a native pathway is optimized by elucidation of targets in a metabolic pathway for modulation of expression levels, a gene deletion or integration of additional copy of a native or heterologous gene to boost availability of precursors and cofactors. The design is next characterized *in vivo* by fermentation of the constructed strain in terms of product yield, formation of intermediates and the impact of the production on the host organism. These learnings are applied in the next cycle to further improve the cell factory. Analogously, the cell can be optimized for utilization of a cheap, renewable feedstock to ensure that the production process is sustainable and cost-effective[4–6].

A cell factory can be constructed using molecular tools for genome manipulation which allow to genomically integrate a desired gene (*e.g.* a heterologous pathway gene or an additional copy of a native gene) and provide stable expression ideally in a marker-free manner. A huge advance in the genome editing field was achieved by harnessing Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) for programmable manipulation of the genome[7]. Furthermore, a crucial role in establishing a cell factory is played by tools enabling control of the production and metabolic pathways[1,6]. Unbalanced expression of heterologous enzymes catalysing a multi-step reaction might cause accumulation of toxic intermediates and excessive transcriptional and translational load on the cell. Modulation of native metabolism allows to accommodate the requirement for precursors and cofactors to synthesize the heterologous product and to redirect the flux of competing pathways or to disable undesired conversion of the final product[8]. The current stage of a molecular toolbox offers a wide range of elements for a gene expression, *e.g.* inducible and tunable

promoters. In addition, a flexible control over gene expression can be obtained by implementation of CRISPR interference and activation systems[9,10]. Due to the major focus of this thesis on CRISPR-based tools, this technology will be discussed in details in the next section.

Finally, functional microbial production relies on the selection of the host organism. The choice of the host organism is dependent on the properties and complexity of the product of interest but generally microorganisms with a well-established molecular tools and extensive experience are preferred. Furthermore, a favourable host organism exhibits a fast reproduction cycle and low cultivation requirements. To ensure the safety of the compounds obtained in a microbial fermentation, the host organisms are assessed by agencies such as the American Food and Drug Administration (FDA) or the European Medicines Agency (EMEA).

Saccharomyces cerevisiae as a host organism

The abovementioned requirements for a cell factory are met by *Saccharomyces cerevisiae*, also known as baker's yeast due to its use as a leavening agent in bakery products. The vast application of *S. cerevisiae* in food industry can be further exemplified by the production of fermented beverages such as beer and wine due to the anaerobic metabolism of *S. cerevisiae* which generates ethanol. Conversion of the glycolysis product – pyruvate, to ethanol by yeast cells allows to regenerate NAD^+ required for glycolysis and production of ATP[11]. Under oxidative conditions pyruvate is converted to acetyl-CoA, which subsequently enters Krebs cycle and is further oxidized to carbon dioxide while reducing equivalents (NADH and $FADH_2$) are regenerated. Overall, respiration yields more energy in the form of ATP than fermentation, however when glucose source is depleted but ethanol generated during the respiration is available, yeast cells can adapt metabolism to fermentative, a phenomenon known as diauxie.

Limited growth requirements make *S. cerevisiae* a commonly selected host organism for a microbial production. In contrast to bacterial cell factories, *S. cerevisiae* has ability to grow in acidic pH (reduces the risk of bacterial contamination), grow at a lower temperature (*e.g.* compared to *Escherichia coli*) and has the ability to reach high culture density[4,12–14]. In addition, yeast cells exhibit higher tolerance to by-products and acids and are sustainable to the phage infection[4,12,13]. The larger size of yeast cells compared to bacteria allows to easily separate cells from the production broth during the downstream processing[4,13]. Unlike prokaryotes, *S. cerevisiae* can mediate post-translational modifications such as phosphorylation and glycosylation which motivates selection of this organism when the production fails in the favourable bacterial system *E. coli*[13]. The research by Prinz *et al.* showed that 70% of recombinant proteins which failed during production in a prokaryotic host were successfully achieved in baker's yeast[15]. Although this limitation is not valid for higher eukaryote such as mammalian cells, their growth requirements are more sophisticated than for *S. cerevisiae*. The growth medium for yeast cells is relatively simple in composition and low in cost[6]. Furthermore, the doubling time of yeast cells oscillates around 1.5 h which is rapid in comparison with other eukaryotic host organisms (*e.g.* Chinese hamster ovary cells). Finally, *S. cerevisiae* is a unicellular eukaryote, which serves as a simple model system to study human biology and diseases, thus a wide range of molecular tools is offered[5]. These characteristics motivate the status of *S. cerevisiae* as a workhorse of biotechnology and the vast application of *S. cerevisiae* as a cell

factory for the production of chemicals (glycerol[16], isoprene[17]), pharmaceuticals (insulin[18], artemisinin precursor[19]), food ingredients (vanillin[20], β-carotene[21]) and biofuels[22].

Production burden

Microorganism have limited capability to produce high quantities of a single compound or protein. This is due to restricted pool of resources mediating transcription (*i.e.* RNA polymerase, nucleotides, transcriptional activators) and translation (*i.e.* ribosomes, tRNAs) that results in competition between the production and native pathways[23–27]. Estimated energetic demand for translation accounts for 50% in bacteria and 30% in mammalian cells[28]. A lower contribution of energy to protein synthesis in mammalian cells can be explained by higher competition for transcriptional than translational resources[29], as also shown for yeast[30]. Nevertheless, excessive protein production can cause aggregation and accumulation of unfolded or misfolded proteins in endoplasmic reticulum, triggering unfolded protein response (UPR). By activating UPR cell can adjust protein folding capacity and if this requirement cannot be fulfilled - programmed cell death is induced. Furthermore, high concentration of an enzyme in the cell contributes to the metabolic imbalance and impairs regulation of cellular processes[31]. To better understand protein burden in *S. cerevisiae*, Eguchi *et al.* analysed impact of overexpressing 29 glycolytic proteins and a heterologous protein (green fluorescent protein - GFP)[31,32]. The majority of analysed proteins (19 out of 29 and GFP) induced cellular toxicity when overexpressed to 15% of total proteins content, regardless of protein function, while expression of another 7 proteins triggered stress at lower level.

Apart from the restriction in availability of the gene expression machinery, competition between pathways can also include metabolic precursors required for building biomass (*e.g.* limited pool of farnesyl diphosphate for ergosterol biosynthesis due to the competing carotenoids production pathway). In addition, some compounds require a multi-step conversion catalysed by different enzymes. Designing the right expression system for each enzyme in a multi-step pathway is a challenging task due to a potential overload, accumulation of toxic intermediates and bottleneck in the pathway if enzymes expression is insufficient. Furthermore, the cellular burden can be escalated by the toxicity of the intermediates and product of the desired biosynthesis (*e.g.* lycopene during production of carotenoids[33]).

Eventually, cellular overload defined as burden, hampers growth and impairs fitness. As a consequence, the productivity of the cell factory is decreased and predictability of the production circuit is reduced. For example, the genetic stability of the introduced genes is diminished and the loss-of-function mutants can be promoted as the production ceases and cellular growth is restored[23,25,33,34]. Problems with predictability of the genetic circuit can also arise from coupling an orthogonal production pathway with native metabolic pathways by creating the competition for resources unforeseen during the design phase[27]. From the perspective of large scale production processes, burden leads to the variability in the cell population in terms of optimal growth conditions and productivity which causes issues in fermentation scale-up[35].

Molecular tools to mitigate production burden

In the recent years, more attention is given to the mitigation of the production burden and development of host aware circuits to reflect cell's production capability[23,25]. For this purpose, a production pathway is controlled either in a static manner *i.e.* by careful selection of elements driving gene expression such as promoter and ribosome binding site (RBS) or a dynamic manner, which includes self-regulating circuits. Selection of a weaker promoter or RBS reduces gene expression and the energetic requirements and resources required for the biosynthesis, however, such a system does not provide to the cell ability to adapt to a sudden perturbation. Instead, dynamic regulators are developed to *sense* environment or cellular fitness and act when a certain stimulus occurs *e.g.* nutrient concentration, cell density or pH. The principle of these molecular devices often mimics feedback repression in nature which helps cells to effectively allocate resources and prevent excessive biosynthesis. The natural feedback regulation is most commonly achieved by the control of enzyme's activity (*i.e.* the excessive product – ligand, binds the enzyme at its allosteric site and induces conformational change which inactivates the enzyme) or repression of gene transcription. The latter process utilizes transcriptional repressors which in free form bind to the specific DNA site and blocks gene transcription and lose the repressing capability when bound to the ligand.

Inspired by this process, Zhang *et al.* developed the first synthetic dynamic sensor-regulator system (DSRS) to control production of biodiesel (in the form of fatty acid ethyl ester) in *E. coli*[36]. The burden sensor was based on the fatty-acid sensing protein FadR and the synthetic promoter which blocked production of fatty acids upon their accumulation in the cell and activated production of acyl-coA and ethanol. The DSRS increased biodiesel yield by 3-fold compared to the production strains without the dynamic regulator and improved strain stability. Another approach which prevents cellular overload is separation of the growth phase from the production phase which could otherwise have a negative impact on biomass formation. In *S. cerevisiae* this can be achieved by selection of promoters induced or repressed by sugar source (*e.g. HXT1*, *HXT7*, *GAL1*, *GAL10* promoters) for the expression of production pathway genes[37]. This approach was applied in *S. cerevisiae* for the production of fatty alcohols[38], α-santalene[39] and carotenoids[40] which was activated when glucose was depleted and sugar source was switched.

An even more sophisticated regulatory system developed by Moser *et al*. allowed for metabolic control in regards to the consumption of sugar source (glucose), external environment (dissolved oxygen) and accumulation of by-product (acetate)[41]. These sensors were constructed using natural *E. coli* transcription factors and synthetic promoters generated from the constitutive core element and a range of operators in different positions flanked by random sequences. To expand the sensor dynamic range this library was screened using expression of a fluorescent protein reported in presence or absence of the inducer and fluorescence-activated cell sorting (FACS). Subsequently, biosensors were used to construct logic gates AND (gives output only when in presence of glucose and acetate) and NAND (activated by availability of glucose and low oxygen). As an output, these gates activated transcriptional repression based on dCas9-CRISPR interference (principles are discussed in the next section) and protein abundance by activation of protases. This system allowed to regulate acetate production which is catalysed by enzymes encoded by endogenous genes *pta* (in exponential phase) and *poxB* (during stationary phase). At the initial growth phase NAND gate represses *pta* to counteract natural upregulation of this gene under low oxygen,

while upon entry to the stationary phase gate AND blocks *poxB* to prevent accumulation of toxic levels of acetate, overall resulting in the temporarily and balanced production.

Apart from dynamic systems based on transcriptional repression, CRISPRi and proteases, new regulatory approaches were established in the recent years, including riboswitches[42] and ribosome actuator[43]. Still, sensing burden maintains a major challenge. Ceroni *et al*. proposed a burden monitor comprised of a fluorescent reporter gene (*GFP*) integrated into a defined genomic locus, which expression indicates cell's production capacity. This fluorescence-based monitor provides an easy way to assess burden and explore production capabilities of the host organism, however it cannot be coupled to a regulatory unit to create a dynamic control system. Therefore, the monitoring unit in the dynamic regulatory systems is often based on sensing capability of transcription factors which act on promoters, as presented in the abovementioned examples. Screening for the promoter that is regulated by transcription factors which sense a desired state such as burden can be supported by differential expression analysis of mRNA, as shown for the general burden sensor developed for *E. coli*[24]. In a different approach, the *sensing* promoter is determined from a library of native[44] or engineered promoters[41] which are exposed to the desired stimulus and screened for expression strength and tightness. Finally, to sense a concentration of a non-native compound artificial transcription factors and promoter pairs can be created. This can be exemplified by VanR biosensor from *Caulobacter crescentus* which was used to detect production of vanillin in *S. cerevisiae*. Such artificial transcription factor/promoter pair can be employed in the dynamic regulatory system to prevent accumulation of toxic product.

1.2 CRISPR in nature and molecular biology

Prokaryotes and archaeons are constantly challenged by the risk of a virial infection which imposes evolutionary pressure on defence mechanisms. To counteract foreign mobile genetic elements, microorganisms developed innate immunity strategies exemplified by abortive infection and restriction modification systems[45]. However, high mutation and recombination rate of viruses helps to escape bacterial defence, thus the host immunity is supplemented by the adaptive defence system. Clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated proteins (Cas) create the CRISPR-Cas system which allows the host to acquire immunity against a new invader. The protection is provided in the three steps process: acquisition, expression and interference. During the acquisition, the foreign nucleic acid is recognized and incorporated into the CRISPR array as a spacer that in most of the CRISPR-Cas systems is mediated by Cas1 and Cas2 (Figure 1-1i). Subsequent expression results in a pre-CRISPR RNA (pre-crRNA) array that encodes multiple spacers separated by direct repeats unique for each CRISPR-Cas system. The pre-crRNA array is processed to individual crRNAs encoding a single spacer and a direct repeat by either Cas protein with inherent RNase activity (*e.g.* Cas6, Cas12a) or if the systems lacks such activity (*e.g.* Cas9) – a non-Cas ribonuclease (*e.g.* RNase III)(Figure 1-1ii). In the interference phase Cas protein (or proteins, depending on the class of the system) forms a complex with crRNA and upon recognition of the nucleic acid of the invading species (protospacer) is bound to the complementary region of the crRNA and degraded by Cas nuclease (Figure 1-1iii). To distinguish the native sequence from the virial and prevent lethal cleavage of the host genetic information, Cas endonuclease exhibits specificity to the nucleotides

proximal to the protospacer, so-called Protospacer Adjacent Motif (PAM). Furthermore, identification of the DNA target and cleavage requires complementarity between the seed region (for Cas9 10–12 nt at the 3' end) in the gRNA and the target[46]. Any mutation in the seed region can reduce or even abolish effectiveness of the genetic manipulation, especially when the mismatch occurs in the proximity to the PAM[46].

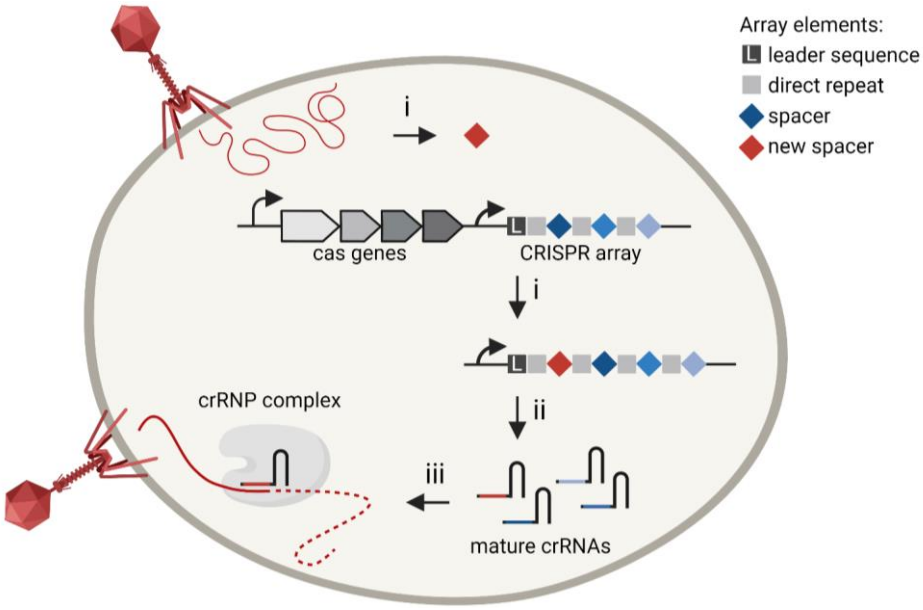


Figure 1-1. Adaptive immunity of the CRISPR-Cas system. Action of the CRISPR-Cas system is comprised of three phases: **i.** *acquisition* – upon recognition of virial nucleic acid its fragment is acquired into the CRISPR array. The invading DNA (protospacer) is incorporated as a spacer at the proximal to the leader sequence in the CRISPR array. **ii.** *expression* – CRISPR array is transcribed to the pre-crRNA which is subsequently processed into individual crRNAs encoding a single target and a direct repeat. **iii.** *interference* – mature crRNA forms a ribonucleoprotein (crRNP) complex with the effector Cas protein. When the virus re-attacks the cell, the foreign nucleic acid is recognized by the complementary binding of the protospacer and spacer in the crRNP and it is cleaved. The cluster of Cas genes schematically represents genes encoding Cas proteins involved in different phases of the immunity. Figure adapter from [45,47].

The exact mechanism of the CRISPR-Cas immunity is dependent on the class and type of the CRISPR-Cas system. The two main classes are distinguish based on the Cas proteins involved in the interference phase – in class 1 this step is mediated by a multi-subunit crRNA-effector complex, whereas in class 2 by a single-effector nuclease. The ability to specifically cleave the nucleic acid with only a single Cas nuclease is attractive when the CRISPR-Cas system is applied in a non-native host, which justifies selection of CRISPR-Cas9 system belonging to the class 2 in the first genome editing reports[48,49]. Cas9 acts as a endonuclease due to the possession of two nuclease domains (HNH and RuvC)[7]. Each of these domains cleaves one of the strands in the double stranded DNA resulting in the blunt product (Figure 1-2A). In type II system the pre-crRNA maturation requires hybridization of crRNA and additional element – trans-activating crRNA (tracrRNA), in the process mediated by a non-CRISPR associated RNase III. In addition, tracrRNA facilitates binding between crRNA and Cas9. To simplify the design of CRISPR-Cas9 experiments, Jinek *et al.* proposed a single guide crRNA (sgRNA) comprised of crRNA, tracrRNA and a synthetic tetraloop which ensures correct secondary structure of the chimeric RNA[7]. Cas9 lacks ribonuclease activity to process

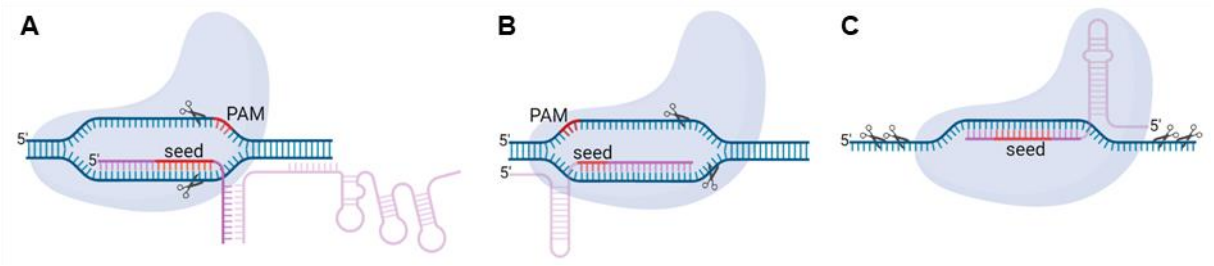


Figure 1-2. gRNA architecture of A. CRISPR-Cas9. Cas9 forms a complex with an sgRNA comprised of a spacer (20 nt, at the 5' end, dark purple) and a tracrRNA (79 nt, at the 3' end, light purple). The complex binds target DNA (blue) at the distal to the PAM (5'-NGG-3', red) and introduces cleavage at the 5' target strand (mediated by RuvC domain) and 3' target strand (mediated by HNH domain), resulting in the blunt ends. The spacer contains a seed region (10-12 nt at the 3' end, red) which is sensitive to mismatches. **B. CRISPR-Cas12a**. gRNA of Cas12a system contains a direct repeat (20-21 nt, light purple) and spacer (20-23 nt, dark purple), in the reversed orientation compared to the sgRNA of Cas9. Cas12a recognizes 5'-TTN-3' PAM and it is sensitive to seed region at the 5' end of the spacer (5 nt, red). Cas12a-gRNA complex cleaves both DNA strands with RuvC domain, resulting in sticky ends. **C. CRISPR-Cas13 system**. Cas13 targets RNA and apart from the target cleavage, it can exhibit non-specific collateral effect. gRNA architecture is orthologue dependent (in terms of elements length and orientation) but it always comprised of a spacer (dark purple) and direct repeat (light purple). The cleavage can be PFS dependent, then PFS is located on the either or both site(s) of the target. Orthologues with no PFS preference are known. Cas13 gRNA has a seed region in the central region of the spacer (11-18 nt from 5' end of the spacer[50], red). Blue sequence – DNA or RNA target; red sequence – position of PAM and seed; dark purple sequence – spacer; light purple sequence – tracrRNA (Cas9) or direct repeat (Cas12a and Cas13). Scissors indicate cleavage site. Elements length is given according to the design used in this thesis. Figure adapted from [51–53].

CRISPR arrays, therefore, a simultaneous manipulation of multiple loci using CRISPR-Cas9 requires one of the following approaches for expression of a gRNA:

i) encoding gRNAs as individual expression cassettes controlled by different promoters[54,55],
ii) co-expressing additional endoribonuclease to process an array encoding multiple gRNAs (*e.g.* Csy4[56]),
iii) equipping a gRNA array with additional sites ensuring proper excision (*e.g.* self-cleaving ribozyme[57] or pre-tRNA recognized and cleaved by endogenous RNase P/RNase Z[58]).

Harnessing the CRISPR-Cas9 system for genome editing revolutionized molecular biology by serving as an easy to design and yet precise tool suitable for a vast range of organisms[48,49]. The flexibility of this system relies on the straightforward adaptation of the genomic target by a simple modification of a spacer in a gRNA , as long as the 3' end of the target is flanked by the G-rich PAM sequence (5'-NGG-3'). Furthermore, manipulation of the genome of the organism of interest using CRISPR-Cas9 required expression of only two elements: Cas9 gene and gRNA or alternatively, supply of the cells with ribonucleoprotein Cas9-gRNA complex[59]. CRISPR-Cas9 systems cleaves the genome at the desired position which is programmed in the gRNA, and the cleaved genomic DNA must subsequently be repaired by one of the cell's rescue mechanisms:

- Non-homologous end-joining (NHEJ);
- Microhomology-mediated end joining (MMEJ);
- Homologous-Directed Repair (HDR)[60].

NHEJ relies on rejoining blunts ends in the DNA region with the double stranded break which may lead to the introduction of 1-10 nt indels in a form of insertion or a deletion. Occurrence of an indel in the open-reading-frame can cause a frameshift and expression of a dysfunctional protein by for example introduction of a premature stop codon. Therefore, CRISPR-mediated cleavage followed by NHEJ repair is useful to study

gene knock-outs. In contrast, DNA repair via HDR requires a repair template with regions homologous to the cleaved DNA region. HDR provides accurate reconstruction of the damaged DNA region as indels are not incorporated and it is therefore advantageous for knock-in of heterologous DNA. Although NHEJ is generally the predominant DNA repair mechanism in eukaryotes, in *S. cerevisiae* genome integrity is mainly ensured via HDR. To promote HDR in organisms which use NHEJ as the main DNA repair mechanism, genes involved in NHEJ can be manipulated to reduce efficiency of this process (*e.g.* in *Yarrowia lipolytica* – knockout of ku70[61]). Collectively, CRISPR-Cas cleavage of double stranded DNA (dsDNA) in the genome and the repair mechanism allow to manipulate the genome of the organism of interest by precisely inserting or deleting a sequence to/in the desired locus.

CRISPR technology spans systems from both classes and various types (distinguished based on the signature protein) which offer different characteristics. This diversity of CRISPR-Cas systems can be exemplified by three Cas proteins representing class 2, mainly: Cas9, Cas12a and Cas13 (Table 1-1). While both Cas9 and Cas12a cleave dsDNA, these two proteins have a different architecture of the nuclease domain. In addition, Cas12a exhibits ribonuclease activity which allows to process a crRNA array to mature crRNAs without any additional elements[62]. Therefore, Cas12a is commonly used for multiplex manipulation of genomic regions[62–67]. In contrast to CRISPR-Cas9 and Cas12a, a class 2 type VI system represented by Cas13 effector protein acts on RNA rather than DNA and can be applied as a specific transcript knockdown tool[68–70]. CRISPR-Cas12a and Cas13 systems were explored in the presented work as a new molecular tools for *S. cerevisiae*, therefore both systems are discussed in the following part of this thesis.

Table 1-1. Main features of selected class 2 CRISPR-Cas systems

Effector protein:	**Cas9**	**Cas12a (Cpf1)**	**Cas13**
Type:	II	V	VI
Acts on:	dsDNA	dsDNA, ssDNA	ssRNA
Nuclease domain:	RuvC and HNH	RuvC	HEPN
Processing gRNA array:	Mediated by RNase III	Inherent RNase activity	Inherent RNase activity
tracrRNA:	Required	-	-
PAM or PFS:	5'-NGG-3'	5'-TTN-3' or 5'-TTTV-3'	Varies between orthologues

Class 2 type V – Cas12a (Cpf1)

CRISPR systems belonging to class 2 type V are characterized by the signature protein Cas12 containing a single RuvC nuclease domain which allows to cleave dsDNA in a PAM-dependent manner and to unspecifically degrade single-stranded DNA (ssDNA)[71,72]. Among CRISPR-Cas12 systems, the first and most extensively characterized subsystem is Cas12a, initially named as Cpf1[73]. Due to its unique features, CRISPR-Cas12a has extended the CRISPR toolbox for the genome manipulation. Firstly, the PAM sequence of Cas12a is T-rich and depends on the orthologue, for instance Cas12a from *Francisella novicida* U112 (FnCas12a) favours 5'-TTN-3', whereas orthologues derived from *Acidaminococcus* sp. BV3L6 (AsCas12a) and

Lachnospiraceae bacterium ND 2006 (LbCas12a) recognize a more specific PAM sequence 5'-TTTV-3'. The preference for T-rich PAMs of Cas12a compensates the limitation of G-rich PAM of Cas9 and expands the position of a genomic target and allows for the manipulation of organisms with high AT content in the genome (*e.g. Plasmodium falcipracum*)[74]. Secondly, the inherent RNase activity of this effector allows for processing pre-crRNA into individual crRNAs[74,75]. In contrast to Cas9, multiplex manipulation using Cas12a is achieved with an array comprised solely of direct repeats and spacers (without additional sequences indicating pre-crRNA processing sites) or exogenous nucleases. The record number of crRNAs processed from a single array by Cas12a is currently set at an impressive number of 25 targets[76]. Thirdly, the Cas12a gRNA is comprised of 42 to 44 nt and lacks tracrRNA (Figure 1-2B), resulting in a shorter DR-spacer region than for Cas9 (~ 100 nt). In the type V system maturation of the pre-crRNA is facilitated by Cas12a which recognises the pseudoknot formed by the repeat fragment of the pre-crRNA and acts as a ribonuclease, producing mature crRNA. Empirical experiments investigating length of gRNA elements revealed functionality of a spacer and a direct repeat truncated to 18 nt and 17 nt, respectively[62,63,74]. As a consequence of shortening the spacer below 20 nt, the position of the cleavage is shifted downstream to 14 nt distal to PAM[77]. In contrast, selection of the full length direct repeat of pre-crRNA (36 nt) decreased efficiency of Cas12a interference in comparison with the truncated form (19 nt)[62,63]. Based on these studies, Cas12a gRNAs and crRNA arrays are typically designed to contain 20-23 nt spacer and 20-21 nt DR. Alike Cas9, Cas12a is sensitive to mismatches between the target sequence and the seed region of crRNA, which spans the first 5 nt at the 5' end of the spacer[74]. Finally, the double strand break introduced by Cas12a is distal from PAM (18-23 nt downstream) resulting in sticky ends which may enhance integration of a donor DNA in host organisms with low HDR frequency[74,78]. According to the hypothesis behind this statement, the blunt end generated by Cas9 action can be repaired by NHEJ and cause formation of an indel in the target region which disables the match between spacer and target and prevents from re-cutting the genome by Cas9. Instead, double-stranded break distant from the PAM introduced by Cas12a remains the target sequence and PAM which allows for the recurring cleavage by Cas12a and increased the chance for the desired HDR repair.

The CRISPR-Cas12a system is functional in non-native prokaryotic organisms as demonstrated by plasmid clearance in *E. coli*[74] and Cas12a mediated DNA recombination in various bacterial species[79]. While expression of CRISPR-Cas9 system in bacteria negatively affects growth, CRISPR-Cas12a exhibits mild toxicity and can be applied as an alternative genome editing tool in bacteria[79]. Finally, Cas12a-mediated genome editing was shown to be functional in numerous eukaryotic organisms, including yeast species[65,80,81], mammalian cells[62,76], plants[82] and higher eukaryotes[83,84].

Class 2 type VI – Cas13

The family of Cas13 proteins is distinct based on the presence of two HEPN (higher eukaryotes and prokaryotes nucleotide-binding) domains which mediate single stranded RNA (ssRNA) cleavage. This feature is the only similarity between effector proteins belonging to different subtypes of type VI. The firstly discovered type VI effector Cas13a from *Leptotrichia shahii* (LshCas13a) was explored as a gRNA programmable tool for RNA manipulation in prokaryotes and eukaryotes[68,70]. While LshCas13a provided

selective knockdown of the desired transcript in mammalian cells[70], this system caused a growth defect of *E. coli*[68]. To understand the source of the toxicity in *E. coli*, Abudayyeh *et al.* conducted programmable *in vitro* degradation of RNA using LshCas13a[68]. In addition to the expected RNA product, random degradation of non-complementary RNA was observed, caused by collateral activity of LshCas13a. Collateral effect is defined as a trans activity of Cas effector induced upon conformational change in the protein structure as a consequence of complex formation between crRNA and Cas protein and introduction of the first cleavage in the desired target (cis activity)[85]. In nature, virial infection triggers activation of CRISPR-Cas13 and cellular dormancy which provides herd immunity to the unaffected bacterial population and prevents the virus from gaining evolutionary advantage[86]. Although other Cas13 orthologues and type VI were shown to exhibit the collateral effect in some hosts[68,87,88], CRISPR-Cas13 serves as an effective yet specific platform for transcript knockdown in higher eukaryotic cells[70,89–91], which are not susceptible to the collateral activity.

CRISPR-Cas13b belongs to another subtype of type V systems but similarly to Cas13a, cleaves ssRNA using a mechanism mediated by HEPN domains. RNA knockdown in mammalian cells was found to be the most efficient and specific using Cas13b from *Prevotella* sp. P5–125 (PspCas13b) among 43 different orthologues of Cas13a, Cas13b and Cas13c systems[89], including previously top-ranked LshCas13a and LwaCas13a[68,70]. Analysis of the CRISPR gene cluster of Cas13b systems revealed presence of accessory proteins Csx27 and Csx28 that control activity of Cas13b (repress and enhance, respectively). This feature can be attractive for construction of tightly controlled tool for regulation of the transcript abundance.

The architecture of gRNA is dependent on the subtype; naturally Cas13a utilizes a direct repeat of 28 nt to 36 nt (orthologue specific) and a 28 nt spacer, which in case of LwaCas13a is also functional in truncated form (20 nt)(Figure 1-2C). Elements of Cas13b gRNA are generally longer – the spacer is comprised of 30 nt while the direct repeat of 36 nt. Interestingly, the orientation of the direct repeat in a gRNAs of Cas13a and Cas13b systems in regards to the spacer is opposite. Alike type V system, Cas13 effectors possess pre-crRNA processing activity which allows for multiplex targeting of RNA in knockdown experiments[69,92]. While Cas systems acting on dsDNA discriminate an invading sequence based on the presence of PAM, in Cas13 systems the protospacer can be bordered by a few nucleotides on the 5' and/or 3' end defined as protospacer flanking site (PFS)[68]. In contrast to PAM, the function of PFS is not to prevent degradation of native RNA as the CRISPR-Cas13 system exhibits the collateral activity. PFS architecture is diverse and orthologue-dependent; for instance, LshCas13a and BzoCas13b (derived from *Bergeyella zoohelcum*) discriminate a single nucleotide at the 3' and 5' site of the spacer, respectively[68,89,92]. In both systems PFS is rather flexible as the effector protein accepts any nucleotide other than guanine for LshCas13a and cytosine in case of BzoCas13b. PspCas13b requires a double-sites PFS, at the 5' DD (not cytosine) and at the 3' NAN or NNA[89]. Finally, Cas13a from *Leptotrichia wadei* (LwaCas13a) lacks PFS requirements[70,89].

CRISPR regulation

The repertoire of the CRISPR-based tools was recently expanded by repurposing the Cas protein for regulation of the gene expression[9,10]. Such system requires inactivation of the nuclease domain in the Cas protein (so-called *dead* Cas or dCas) which disables introduction of the double-stranded break in the target

DNA whereas the capability of the dCas protein to bind the desired locus remains. The inactivation of Cas9 is achieved by a mutation of the RuvC and HNH (histidine-asparagine-histidine) domains, whereas a single point mutation is effective in disabling the RuvC domain of Cas12a. Binding of the dCas protein to the promoter region or the open-reading-frame (ORF) of a gene-of-interest sterically blocks the access to RNA polymerase (RNAP) and prevents gene transcription. This mechanism defined as CRISPR interference (CRISPRi) can be further enhanced by a recruitment of an effector domain from a transcription factor[93]. For gene downregulation commonly used repression domains are the chromatin modifying protein KRAB (Krüppel-associated box) domain of Kox1 and the transcriptional repressor Mxi1[10,94–96]. While in prokaryotes nuclease dead Cas9 can provide up to 1000-fold repression[97], in eukaryote the efficiency of dCas9 effector without an additional repression domain provided only 18-fold which was improved upon the fusion to KRAB (53-fold repression). Furthermore, the repression range of the CRISPRi system can be controlled by adaptation of the spacer position, mismatch or selection of a non-conventional PAM[24,98].

Analogously, a dCas effector can be coupled with a transcriptional activator and targeted to the promoter region to enhance the transcription of desired gene, so-called CRISPR activation (CRISPRa). Gene upregulation can be provided by a transcriptional activator VP64 which is a tandem of four VP16 domains (Viral Protein 16 from *Herpes simplex*) or p65AD which is a subunit of the nuclear factor-κB (NF-κB). The CRISPRa/i system can be provided either as a fusion of the dCas protein and the regulatory domain[9,93] or a recruitment of the regulatory protein to the dCas9 or an RNA scaffold built in the gRNA[99]. The latter approach benefits from the flexibility of the direction in regulating gene expression (*i.e.* both activation and inhibition is possible by a simple change of the gRNA-scaffold construct). Yet another regulatory CRISPR system was reported by Kiani *et al.* based on the observation that Cas9 nuclease activity is disabled when the spacer region in the gRNA is truncated to ≤ 16 nt without affecting the capability to recognize and bind the target[100]. The potential of this method lies in the ability to leverage Cas9 with functional nuclease domains for both engineering and regulation (activation and interference) solely by a control of gRNA length. A similar system was established for CRISPR-Cas12a in *Y. lipolytica* utilizing a 16 nt spacers for repression of gene expression and a 23-25 nt spacer for gene disruption[80] and for mice (spacer comprised of 15 nt and 20 nt, respectively)[101].

Regulation of the gene expression by CRISPRa/i is a reversible process and the modification in the transcription efficiency is not inherited[9]. Specificity, reversibility and simplicity of the experimental design make CRISPR-based regulatory systems particularly attractive for optimization of the microbial production. For instance, Lian *et al*. established a tri-functional system based on CRISPR-based knockout, up- and downregulation (CRISPR-AID) to achieve higher production of β-carotene in *S. cerevisiae*[94]. Each of the modulation units utilized a different Cas orthologue (and SaCas9 for deletion, dLbCpf1-VP for CRISPRa, dSpCas9-RD1152 for CRISPRi) to ensure orthogonality of the system. As described in the previous section, pathway modulation can act dynamically to reflect cellular fitness. To this day, several feedback control systems based on CRISPRi-dCas9 were reported, including a proof-of-concept study mitigating burden imposed by expression of a synthetic circuit in *E. coli*[24]. In this paper, the transcriptional response to a synthetic construct was evaluated via RNA-seq for three different cases (expression of lacZ reporter, expression of a heterologous protein and expression of an operon responsible for a metabolic pathway),

revealing upregulation of *htpG* gene (encodes chaperone protein) as a common response to the expression of burdensome circuits. The promoter of *htpG* was used for an expression of gRNA mediating downregulation of the production pathway by CRISR-dCas9 system. Implementation of this burden-driven regulator for the production of VioB–mCherry improved total product, presumably due to increased production capacity and growth rate. Similarly, Moser *et al.* utilized CRISPRi-dCas9 to create an AND and NAND gate which linked the regulatory system with sensors for the external environment (O_2), availability of the feedstock (glucose) and the production of the final product, acetate, cytotoxic at high concentration. In addition to dCas9, the regulatory unit contained proteases which act on *poxB* and *pta* genes catalyzing conversion of pyruvate to acetate[41]. Selective control of the protein abundance is attractive due to the rapid response, however, *Mycoplasma florum* Lon protease (mf-LON) applied by Moser *et al*. was shown to be less effective than CRISPRi[41]. In another approach applied in *Bacillus subtilis*, Wu *et al*. combined a CRISPRi based NOT gate and CRISPRa for the regulation of N-acetylglucosamine biosynthesis, a nutraceutical for arthritic treatment[102]. The production of N-acetylglucosamine was boosted by redirectioning flux from competing pathways with the downregulating unit and enhancing expression of genes responsible for the product formation. The common feature of the abovementioned feedback control systems is connection of the sensing unit (which reflects the well-being of the host) with the regulatory unit (acting on the burdensome pathway).

1.3 β-carotene importance and production

Carotenoids represent the most abundant group of pigments in the nature which exceeds 600 compounds[103]. Carotenoids, yellow to red coloured compounds, are comprised of isoprene units reaching 30-40 carbon atoms and resulting in conjugated double bounds which imparts the property to absorb light at 400-500 nm[104]. Depending on the presence of oxygen atom in the structure, carotenoids are classified as carotenes (hydrocarbons, *e.g.* β-carotene, lycopene) and xanthophylls (derivatives with oxygen, *e.g.* lutein, astaxanthin).

Although carotenoids can be found across all the kingdoms their physiological role varies among species[105]; in photosynthetic organisms (*i.e.* plants, algae, cynanobacteria) carotenoids play two crucial roles - light-harvest and the transfer of the light energy to chlorophyll and serve as photoprotective agents; non-photosynthetic yeast and fungi utilize carotenoids as antioxidant to prevent oxidative stress caused by reactive oxygen species[106]. Although animals cannot synthesize carotenoids *de novo*, they require carotenoids as a precursor of vitamin A which is important for cell differentiation and development, immune response and proper vision[103]. Different carotenoids can serve as vitamin A precursors, however the main compound in this process is β-carotene which is cleaved to two retinal molecules (one of the forms of vitamin A) during digestion. The importance of β-carotene in this process was discovered by Paul Karrer who for this work was awarded the Nobel Prize in Chemistry in 1937. Furthermore, carotenoids are important for human health due to its antioxidant property and could help in prevention of cancer, cardiovascular diseases and delaying aging[107]. Due to the yellow to red colour, carotenoids are also used commercially as food colourants, including orange β-carotene. Cumulatively, the demand for β-carotene in

pharmaceutical, food and animal feed industries contributes to the global market exceeding $526 million in 2020 and expected increase to $780 million by 2027[108].

The main approach for the production of β-carotene includes extraction from the natural sources (plants and algae), however it provides low yields and is season-dependent. To provide a robust source of β-carotene, chemical synthesis was developed. Alternative production approach employs microbial cell factories and offers a sustainable source of this industrially relevant compound. This approach can be employed using either microorganisms that naturally produce β-carotene or introduce the β-carotene biosynthesis pathway in a platform microorganism such as *E. coli* and *S. cerevisiae*. For example, the red yeast *Xanthophyllomyces dendrorhous* contains a carotenoids biosynthesis pathway, resulting in β-carotene and astaxanthin (Figure 1-3). Efforts in improving the production of astaxanthin in *X. dendrorhous* were mostly focused on adaptation of the growth environment while the number of papers reporting molecular modifications is minimal due to limited knowledge about regulation of carotenogenesis in this species[109], premature state of molecular tools (*e.g.* CRISPR-mediated genome editing was reported only in 2021[110])

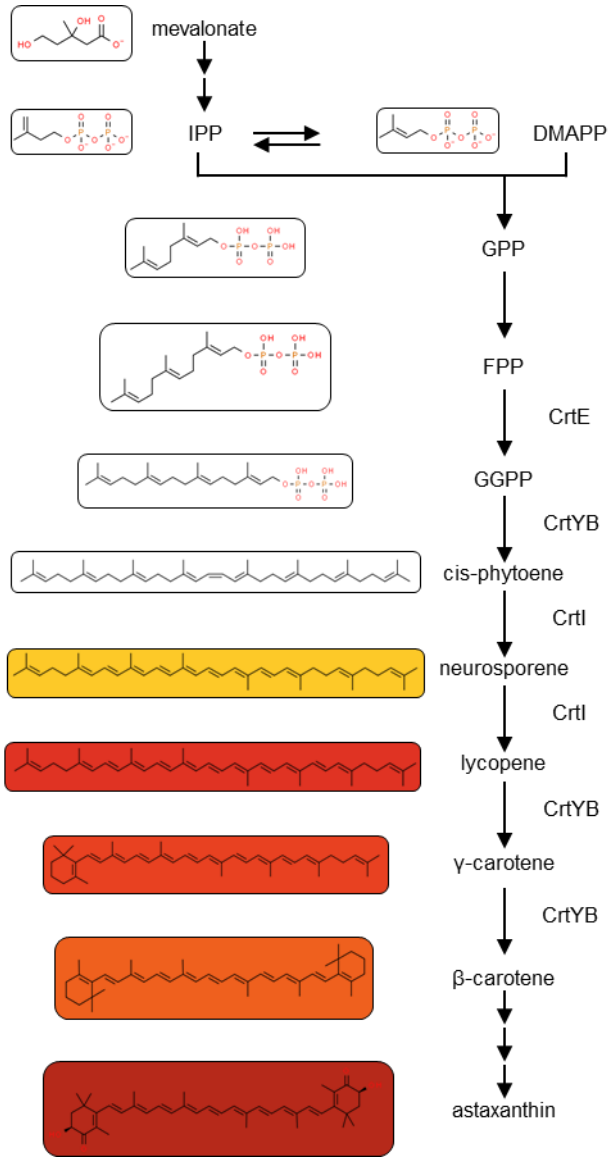


Figure 1-3. Carotenoids biosynthesis in *X. dendrorhous* (example of production in the natural host). Biosynthesis starts with the mevalonate pathway which converts three acetyl-coA molecules to mevalonate (not dipicted). Subsequently, mevalonate is transformed to isopentenyl diphosphate (IPP) in a multi-step reaction. IPP is isomerized to dimethylallyl diphosphate (DMAP) and these two molecules are condensed to geranyl diphosphate (GPP). GPP is further converted with IPP molecule to farnesyl pyrophosphate (FPP). The following step is catalysed by the first of the enzymes required for carotenogenesis – GGPP synthase (CrtE) which generates geranylgeranyl pyrophosphate (GGPP) from FPP and IPP. Two molecules of GGPP are then condensed to the first carotenoid product – cis-phytoene in the reaction driven by the bi-functional enzyme phytoene synthase and lycopene cyclase (CrtYB). The next conversions are catalysed by phytoene desaturase (CrtI) resulting in red lycopene. Lycopene is further converted to γ-carotene and then β-carotene by the action of CrtYB. Carotenoids production in *X. dendrorhous* exceeds biosynthesis of β-carotene and can further generate other carotenoids *e.g.* astaxanthin which represents xanthophylls. Colour of the compound is indicated by the background. This figure was modified from [109,252]. Chemical structures were retrieved from ChemSpider.

and expression systems[111]. The fungus *Blakeslea trispora* and microalgae *Dunaliella salina* are used for commercial production of β-carotene. The production using *B. trispora* is a two phases process which initially requires separate cultivation of a (+) and (-) mating types followed by the joint fermentation of both types in specific medium[112]. Similarly, β-carotene production in *D. salina* is separated to two steps – firstly fermentation conditions are optimal for the growth and then switched to enhance carotenogenesis[113].

Yield of carotenoids obtained in biosynthesis using natural producers cannot yet compete with chemical synthesis while there is increasing demand for carotenoids[114]. This motivates the increasing interest in establishing β-carotene production in model microorganism such as *E. coli*, *S. cerevisiae* and *Y. lipolytica* which are characterized by simple growth requirements, availability of molecular toolbox and vast experience to further improve the product yield[21,115,116]. However, achieving high productivity of β-carotene is challenged by the toxicity of pathway compounds to the host organism. This is due to the accumulation of carotenoids in cell membranes which leads to improper composition of the membrane *i.e.* reduced content of unsaturated fatty acids and consequently decreased membrane fluidity[33,117–119]. The imbalance in membrane composition is further escalated by the competition for farnesyl diphosphate (FPP) which serves as a precursor in β-carotene and ergosterol biosynthesis[119]. Ergosterol is an essential component of the cell membrane in fungi and its shortage has a severe impact on the cellular processes which require cell membrane such as endocytosis, cell fusion and polarization and construction of cell wall[120]. The toxicity of carotenoids was also demonstrated by the analysis of transcriptome which revealed activation of the pleiotropic drug resistance transporters as the cell's attempt to secrete carotenoids[21]. Consequently, alteration in cellular metabolism and physiology due to the production of carotenoids hampers cellular growth and compromises productivity[33,121].

To mitigate the burden triggered by the biosynthesis of carotenoids and maximize productivity researchers optimized expression of genes involved in the production pathway and affected host's metabolic pathways by adjustment of the promoter strength[122]. Alternatively, production of carotenoids was decoupled from the growth phase by the inducible expression of *crt* genes[40,123]; in this system presence of glucose in the medium triggered expression of ERG9 and enabled biosynthesis of ergosterol while blocking carotenogenesis. Next, depletion of glucose and switch to consumption of galactose activated expression of *crt* genes but blocked ERG9. However, the first of the described approaches is static while the second works in the ON/OFF mode, therefore none of the systems can act when perturbation occurs (*i.e.* excessive accumulation of a toxic compounds). The dynamic controller for carotenoids production in *E. coli* was developed by Farmer and Liao[124] and was based on an artificial regulon which sensed levels of acetyl phosphate reflecting the metabolic state (*i.e.* availability of glucose) and acted on *idi* and *psp* expression – genes encoding enzymes involved in generating precursors for lycopene biosynthesis. Despite of the new developments in molecular biology, there were no reports of dynamic circuits for burden-driven regulation of carotenogenesis since this publication. Lian *et al.* implemented a combinatorial system for simultaneous up-/down-regulation and knockdown to boost production of carotenoids in *S. cerevisiae* using CRISPR technology, however this system was also static[94].

1.4 Thesis aim

The work presented in this thesis attempted to expand the molecular toolbox based on CRISPR technology which can be used to improve production of β-carotene in *S. cerevisiae*. A variety of elements was used to develop a dynamic feedback control system to mitigate burden caused by production of carotenoids. It is known that microbial production of this compound negatively affects the host organism which is exemplified by hampered growth and unbalanced metabolism caused by redirection of cellular resources towards the production pathway[33,34,117,118,121,125]. To understand the toxic effect of β-carotene production in *S. cerevisiae*, we first evaluated a set of strains with varying productivity in terms of growth, β-carotene production and genetic stability (**Chapter 2**). Based on these learnings, we were able to select parameters characterizing strain performance and production which could be improved with a feedback control system (*i.e.* microbial growth and genetic stability of the production pathway).

Next, we developed a detailed protocol describing a genome editing method based on CRISPR-Cas12a (**Chapter 3**). The inherent RNase activity of Cas12a confers the capability to process gRNA arrays encoding multiple targets, which allows to simultaneously cleave multiple genomic targets and to perform multiple genomic modifications in parallel. This method was used to create β-carotene producing strains with the expression cassettes integrated into different genomic loci. The set of carotenogenic strains was further extended to create strains with differentially expressed *crt* genes, responsible for β-carotene biosynthesis, to elucidate the bottleneck in the pathway (**Chapter 4**). The limiting step in the pathway can be next used to downregulate β-carotene production by targeting a single *crt* gene with dCas12a, which simplified the design in the first CRISPRi experiments. Furthermore, the set of carotenogenic strains created with CRISPR-Cas12a contained heterologous promoters in contrast to strains characterized in Chapter 2 containing native *S. cerevisiae* promoters in the expression cassettes of the production pathway genes. Control of *crt* genes from heterologous promoters allowed to use these strains in CRISPRi experiments without affecting expression of native genes. This was a crucial consideration in CRISPRi experiments as targeting a gene (ORF) sequence represses expression less effectively than locating dCas12a to the promoter region. Furthermore, generation of strains with non-native promoters in the expression cassettes served as a foundation for implementation of the CRISPRi-dCas12a system to downregulate carotenoids production, evaluate promoters for burden-sensing capabilities and to eventually establish the feedback control system.

A minimal system for feedback regulation must be comprised of two element; mainly a monitor for the cell well-being which will activate a regulatory system when the production leads to elevated stress, and the regulatory system itself which will act on the production pathway to mitigate the burden (Figure 1-4). Starting with the regulatory system, we established CRISPR interference for dCas12a and elucidated the key design parameters to achieve efficient downregulation (**Chapter 5**). Furthermore, CRISPRi-dCas12a was used to control transcription of genes involved in β-carotene biosynthesis. In **Chapter 6** CRISPR-Cas13 was assessed for transcript knockdown in *S. cerevisiae* as a new regulatory tool which could be used in the feedback control system.

To determine whether any native *S. cerevisiae* promoter can *sense* a burden related to the formation of β-carotene by enhanced transcription, RNA sequencing was conducted (**Chapter 7**). Promoters of genes strongly upregulated under production of carotenoids were selected and further assessed for biosensing properties and ability to express a functional gRNA. Finally, the feedback control system was created by combining CRISPRi-dCas12a and a gRNA expressed from a *biosensing* promoter (**Chapter 8**). The created device was then evaluated for dynamic optimization of β-carotene biosynthesis.

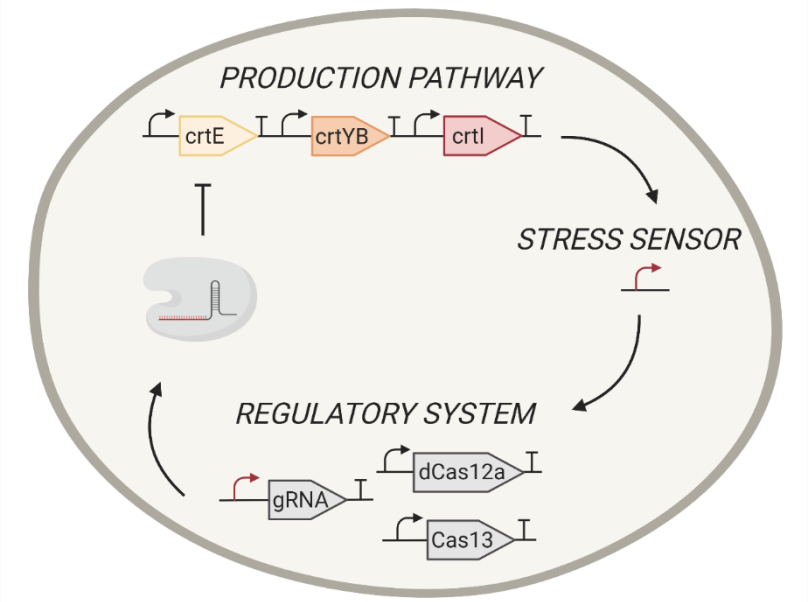


Figure 1-4. Concept of the feedback control system for dynamic regulation of β-carotene production. Construction of the feedback control system comprises of three main steps: 1. Creating a production strain with heterologous promoters controlling *crt* genes to enable CRISPRi without affecting expression of native genes (**Chapter 3, 4**); 2. Determination of native promoters sensitive to the production of β-carotene that can be used to *sense* production stress (**Chapter 7**); 3. Establishing tools to regulate gene expression based on CRISPR, specifically CRISPRi-dCas12a repressing gene transcription (**Chapter 5**) and Cas13 acting on post-transcriptional level (**Chapter 6**). Ultimately, the biosensing promoter was used to express a gRNA which with either dCas12a or Cas13 mediates downregulation of the production pathway in the carotenoids producing strain (**Chapter 8**).

The CRISPR-based tools developed in this thesis to establish a dynamic regulatory system (*i.e.* multiplex genome editing with CRISPR-Cas12a, CRISPRi-dCas12a and CRISPR-Cas13) are versatile can be applied individually to construct or optimize a pathway other than β-carotene biosynthesis, thereby expanding the molecular toolbox for engineering *S. cerevisiae*.

2 Evaluation of growth and carotenoids production by engineered *S. cerevisiae* strains

This Chapter describes characterization of carotenoids producing *S. cerevisiae* strains constructed by Nathalie Buiting-Wiessenhaan (DSM) as a part of European Union's Horizon 2020 ITN DD-DeCaF grant [686070]. All experiments included here were conducted by me except for the operation of HPLC: carotenoids quantification was supported by Reza Maleki Seifar (DSM) and sugar content was performed by Miriam van Steendijk Daele (DSM). Nanopore sequencing was conducted by me under supervision of Thomas E. Gorochowski (University of Bristol) and the data processing was done by Irsan Kooi (DSM). The work was supervised by René Verwaal (DSM), Thomas E. Gorochowski and Johannes A. Roubos (DSM).

2.1 Introduction

Carotenoids are naturally produced by a vast variety of organisms, such as bacteria (*Paracoccus sp.*), algae (*Haematococcus pluvialis*), fungi (*Phycomyces blakesleeanus*), and yeast (*Xanthophyllomyces dendrorhous*) to confer protection against reactive oxygen species[34]. Limited knowledge, lack of molecular tools, demanding growth conditions, and concerns over the safety of non-conventional organisms promotes the selection of organisms that are well-characterized but naturally unable to synthesize carotenoids as a production chassis. Production of carotenoids has been successfully implemented in model organisms: *Escherichia coli* and *Saccharomyces cerevisiae* by overexpressing genes derived from natural producers (Figure 2-1)[21,115]. *S. cerevisiae* is particularly suitable for the production of carotenoids due to efficient isoprenoid metabolism, reducing the limitation in precursors availability[117]. The drawback of heterologous production in a non-native host organism is the potential burden related to the reallocation of cellular resources and competition with native processes[26]. As a consequence, cell growth is often hampered and productivity decreased which can further promote the formation of escape mutants (cells with a loss-of-function mutations or the full loss of genes from the production pathway) and reduced ability to predict the behaviour of synthetic circuits[23]. In addition, the production of certain non-native compounds or biosynthesis of excessive amounts of these can lead to an accumulation of toxic metabolites in the cell. An example of such a pathway in *S. cerevisiae* is carotenogenesis (biosynthesis of carotenoids), which has been reported to impair growth[33,121], affect the stability of the produc-

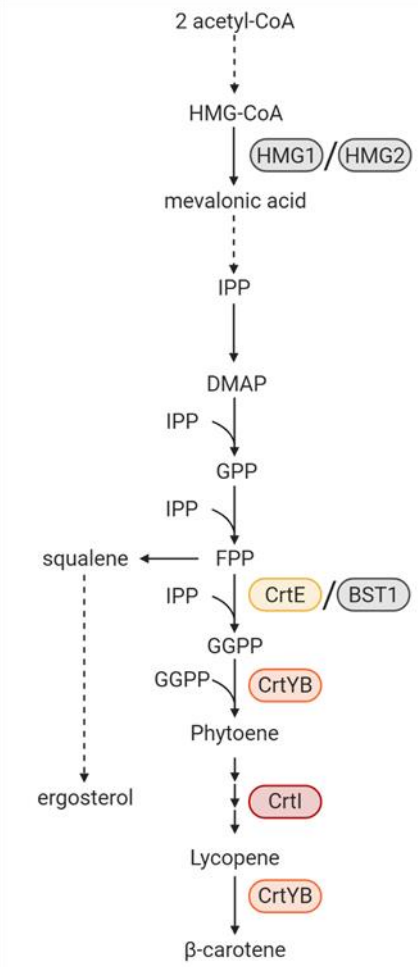


Figure 2-1. β-carotene synthesis in *S. cerevisiae* by overexpression of *crtE*, *crtYB* and *crtI* from *X. dendrorhous*. IPP, iso-pentenyl diphosphate; DMAP, dimethylallyl diphosphate; GPP, geranyl diphosphate; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate. Figure adapted from [21].

tion genes[33,34] and cellular membrane[33,117,118] and activate pleiotropic drug resistance transporters[125]. Here, we characterized a set of six *S. cerevisiae* strains producing different levels of carotenoids in a batch fermentation and elucidated the stress response related to the formation of carotenoids. For this purpose we analysed cellular growth, metabolites profile, carotenoids production and tested stability of the β-carotene biosynthesis pathway to determine the optimal production level which can be obtained without compromising cellular growth and fitness.

2.2 Materials and methods

Cultivation conditions

Strains were characterized in a batch fermentation in mineral and rich media monitored for 4 days. Rich medium provides nutrients which can be directly incorporated into an anabolic pathway, whereas growth on mineral medium requires synthesis of metabolic precursors further increasing energetic requirements[126]. The rich medium was composed of yeast extract (1%, Bacto™ Becton–Dickinson, Franklin Lakes, NJ, USA), phytone peptone (2% Difco™ BD), and glucose (4%, Sigma Aldrich, St Louis, MO, USA). The mineral medium contained potassium phosphate monobasic (3%, Sigma Aldrich), magnesium sulfate heptahydrate (0.5%, Merck, Darmstadt, Germany), urea (2.3%, Merck), glucose (4%, Sigma Aldrich), vitamins, and trace elements[127]. Cultures were set up using 0.5 L shake flasks filled with 0.1 L medium and inoculated with a starting OD_{600} of 0.1 using pre-cultures grown overnight in the same type of medium. Growth was evaluated by optical density and dry weight. To measure the dry weight, ca. 7.5 g of the culture was collected in a pre-dried and weighted polypropylene tube (VWR, Radnor, PA, USA). The cell pellet was separated from the spent medium by spinning down (5300 rpm, 10 min) and dried at 104 °C for 24 h. The experiment was conducted in biological duplicates.

Strains

S. cerevisiae strains were obtained from the DSM Strain Conservation Unit. Carotenoids production was achieved by genomic integration of *crt* genes encoding geranylgeranyl diphosphate synthase (*crtE*), bifunctional lycopene cyclase/phytoene synthase (*crtYB*), and phytoene desaturase (*crtI*) from *Xanthophyllomyces dendrorhous*[21]. Except for strain CAR-001, *crt* genes were codon-optimized for expression in *S. cerevisiae*. Strains CAR-002, CAR-005, and CAR-009 varied in the strength of promoters used to express *crt* genes, resulting in different levels of carotenoids production (Table 2-1). Strains CAR-020 and CAR-021 contained an N-terminal truncated gene of hydroxymethylglutaryl-CoA reductase 1 (*tHMG1*) in addition to the *crt* expression cassettes with high strength constitutive promoters, to boost the production of mevalonate which is the precursor for carotenogenesis (Figure 2-1). The productivity of strains CAR-020 and CAR-021 was further improved by the integration of additional copy of *crtI* in both strains and *crtYB* in CAR-021. The genotype of analysed strains can be found in SI Table 2-1.

Table 2-1 Strains characterized in Chapter 2. For further details about expression cassettes see SI Table 2-1.

Strain name	Description	Expected carotenoids production	Origin
CEN.PK113–7D	Wild type	No	[128]
CAR-001	Non-codon optimized *crtE, crtYB, crtI* expressed from high strength constitutive promoters	Low	[21]
CAR-002	Codon optimized *crtE, crtYB, crtI* expressed from high strength constitutive promoters	Medium	Internal DSM
CAR-005	Codon optimized *crtE, crtYB, crtI* expressed from medium strength constitutive promoters	Low	Internal DSM
CAR-009	Codon optimized *crtE, crtYB, crtI* expressed from low strength constitutive promoters	Very low	Internal DSM
CAR-020	Codon optimized *crtE, crtYB, crtI* expressed from high strength constitutive promoters; additional copy of *tHMG1* and *crtI*	High	Internal DSM
CAR-021	Codon optimized *crtE, crtYB, crtI* expressed from high strength constitutive promoters; additional copy of *tHMG1, crtI* and *crtYB*	Very high	Internal DSM

Quantification of metabolites and carotenoids

The concentration of glucose and ethanol in cultures was measured with a Dionex Ultimate 3000 HPLC system (Thermo Fisher Scientific) and Aminex HPX-87H column (300 x 7.8 mm, Bio-Rad, Hercules, CA, USA) according to the manufacturer's protocol.

For quantification of carotenoids (*i.e.* phytoene, lycopene, and β-carotene), 1 mL of cell culture corresponding to an OD_{600}=20 was collected and separated from the spent medium by centrifugation (14,000 rpm, 8 min). The cell pellet was resuspended in tetrahydrofuran (THF, Merck) and homogenized using glass beads and Precellys homogenizer (Bertin, Montigny-le-Bretonneux, France) in two cycles (6,800 rpm, 20 sec) with a 30 sec break. The resulting carotenogenic fraction was then separated from debris by centrifugation (14,000 rpm, 8 min, 4 °C) and analysed with UHPLC Ultimate 3000 system (Thermo Fisher Scientific). The UHPLC system was equipped with Waters XBridge C18 column (3.5 µm, 2.1mm x 50 mm) and an Ultimate 3000 photodiode array detector (Thermo Fisher Scientific). Phytoene was detected at a wavelength of 286 nm, lycopene at 475 nm, and β-carotene at 450 nm. Data was collected and processed with Chromeleon software. All solvents used contained 100 mg/L butylated hydroxytoluene (BHT). The mobile phases and the gradient used for the separation of carotenoids in UHPLC are in SI Table 2-2.

Analysis of genetic instability

To evaluate stability of the carotenoid pathway, carotenogenic strains were grown for 72 h in either rich or mineral media until stationary phase. Cultures were then diluted in the same type of media to an OD_{600}=0.1 and cultured for a further 72 h. After incubation, cultures were diluted to ca. $3 \cdot 10^4$ cells/mL and plated on rich or mineral agar, such that the type of media was kept the same or switched (*i.e.* from mineral to rich and *vice versa*). Plates were incubated at 30 °C for two days and visually inspected for colonies that

suspectedly lost the potency to produce carotenoids. The experiment was performed twice on independent days.

Nanopore DNA sequencing was conducted to evaluate if the inability to produce carotenoids in colonies that lost colour was caused by mutations in cassettes encoding *crt* genes. For this purpose, genomic DNA was isolated from selected clones of strains CAR-002 and CAR-0020 using Zymolyase (Zymo Research, Irvine, CA, USA), as described before[66]. Subsequently, genomic DNA was used as a template to amplify the expression cassettes integrated into locus INT1 for sequencing. Amplification of DNA was conducted with Phusion polymerase (New England Biolabs, Ipswich, MA, USA) and primers encoding unique barcodes (SI Table 2-3) to enable multiplexing of samples. The resulting library was prepped using the Ligation Sequencing Kit (SQK LSK 108, Oxford Nanopore Technologies, Oxford, UK) and sequenced on a FLO-MIN106 (9.4 SpotON) flow cell and the MinION MIN-101B device. Data was acquired with MinKnow2.0 software and the collected reads were processed with Porechop and compared with the template sequences using Minimap2 and Freebayes variant calling to identify point mutations.

2.3 Results

Cellular growth and metabolism

Cellular growth was monitored by the optical density and the dry weight throughout the fermentation (Figure 2-2). After 24 h of cultivation in a mineral medium, strains with moderate and high production (CAR-002, CAR-020, CAR-021) reached significantly lower optical density compared to strains producing modest amounts of carotenoids (CAR-001, CAR-005, CAR-009) and the wild type cells (Figure 2-2A). This deviation was also observed after 37 h for strains CAR-020 and CAR-021. However, after 48 and 72 h, the growth of these strains was comparable to the strains with lower production and the wild type. Restrained growth of medium/high producing strains (CAR-002, CAR-020) in the initial phase of fermentation was not observed during cultivation in a rich medium. However, some hindrance was noted for strain CAR-021 throughout the experiment (Figure 2-2B). The difference in growth between low and medium/high producing strains was also confirmed by the cell dry weight analysis. In mineral medium, strains CAR-002, CAR-020, and CAR-021 displayed lower dry weight for all time points, whereas in the rich medium the difference between strains with low versus high production was less striking; reduced dry weight was observed only for CAR-002 and CAR-021 in comparison to the wild type strain.

The analysis of metabolites in the broth revealed the presence of glucose after 24 h of fermentation in the mineral medium for strains CAR-002, CAR-020, and CAR-021 (Figure 2-3). Contrary, after 24 h all glucose was consumed by the wild type strain, which shifted to growth on ethanol. Diauxic growth of strains CAR-020 and CAR-021 was delayed until after 37 h of fermentation.

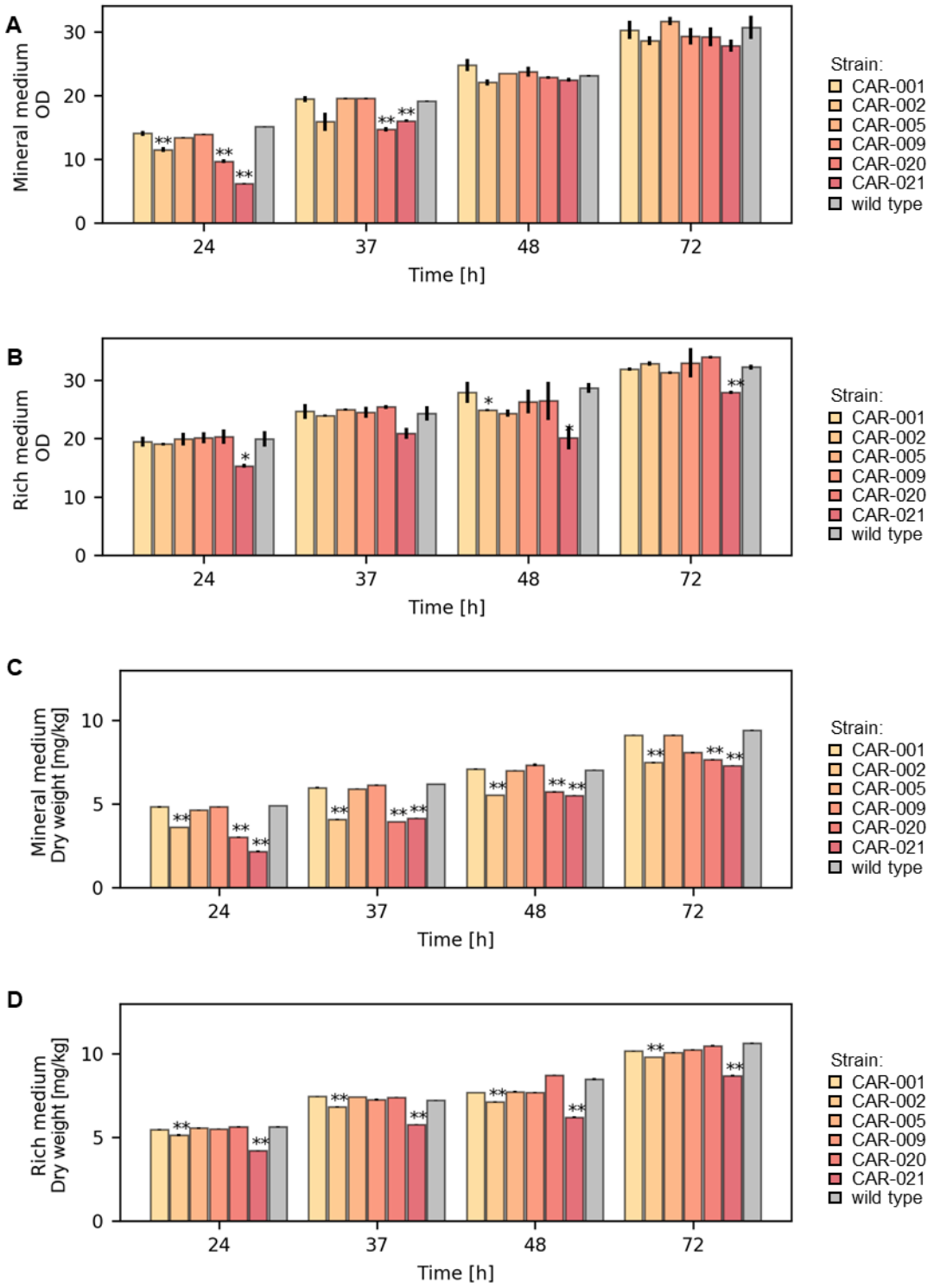


Figure 2-2. Growth of carotenogenic strains during the batch fermentation. A. Optical density (OD) during growth on the mineral medium and **B.** rich medium. **C.** Cellular dry weight (DW) during growth on the mineral medium and **D.** rich medium. Data represents mean ± standard deviation (n=2). * p<0.05, ** p<0.01, two-tailed t-test for carotenogenic and wild type strain.

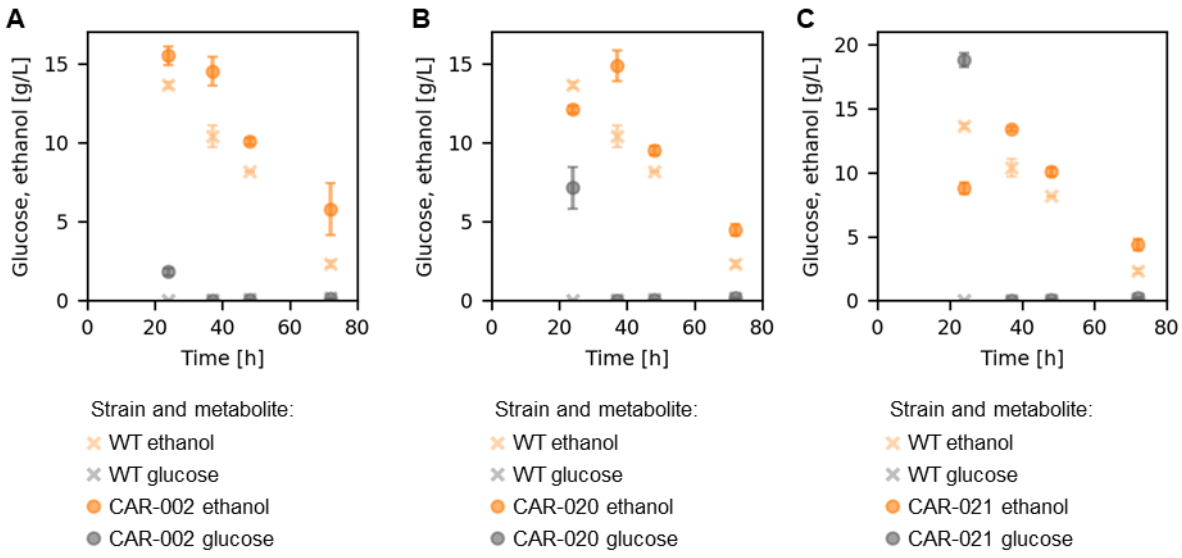


Figure 2-3. Ethanol and glucose throughout fermentation in mineral medium for the wild type strain and A. CAR-002, **B**. CAR-020 and **C**. CAR-021. Data represents mean ± standard deviation (n=2).

Carotenoids production

Quantification of phytoene, lycopene, and β-carotene confirmed the expected productivity of strains CAR-001 – CAR-020, represented as total carotenoids production (*i.e.* the sum of phytoene, lycopene and β-carotene, Figure 2-4). The lowest total production was formed for strain CAR-001 with native (non-codon optimized) *crt* genes from *X. dendrorhous*. Gene expression from promoters classified as low, medium, and high strength resulted in gradual improvement in productivity (strains CAR-009, CAR-005, and CAR-002, respectively). Integration of *tHMG1* and an additional copy of *crtI* further improved the productivity of strains CAR-020 and CAR-021 in comparison to CAR-002. Although strain CAR-021 was expected to achieve the highest carotenoids production due to integration of two copies of *crtYB* expression cassette, the level of total carotenoids produced by strain CAR-021 was comparable to CAR-020 with a single *crtYB* expression cassette. In comparison to strain CAR-020, higher production of total carotenoids by strain CAR-021 was observed only after 72 h of incubation during the growth in a mineral medium (1.6 times higher), while in a rich medium the amount of total carotenoids was significantly lower in the initial phase of fermentation (24 h and 37 h, 0.9 times lower).

Notably, the production profile varied between the mineral and rich medium. We compared phytoene, lycopene and β-carotene profiles of two selected strains: CAR-002 and CAR-020, which displayed slower growth but different production level of total carotenoids (Figure 2-5, profiles of the remaining strains are enclosed in SI Figure 2-1). Fermentation on the mineral medium resulted in higher amounts of β-carotene and improved conversion of phytoene to the subsequent intermediates in the first days of fermentation. β-carotene production in strains CAR-002 and CAR-020 was at least 1.5- and 2-fold higher throughout the fermentation in the mineral medium in comparison with the rich medium. The highest change in β-carotene production by strain CAR-002 was observed after 37 h when 19.9±4.9 μg/OD was produced in mineral medium and 9.3±0.1 μg/OD in rich medium, whereas in strain CAR-020, β-carotene after 37 h reached 28.6±0.7 μg/OD and 6.3±0.1 μg/OD in mineral and rich medium, respectively.

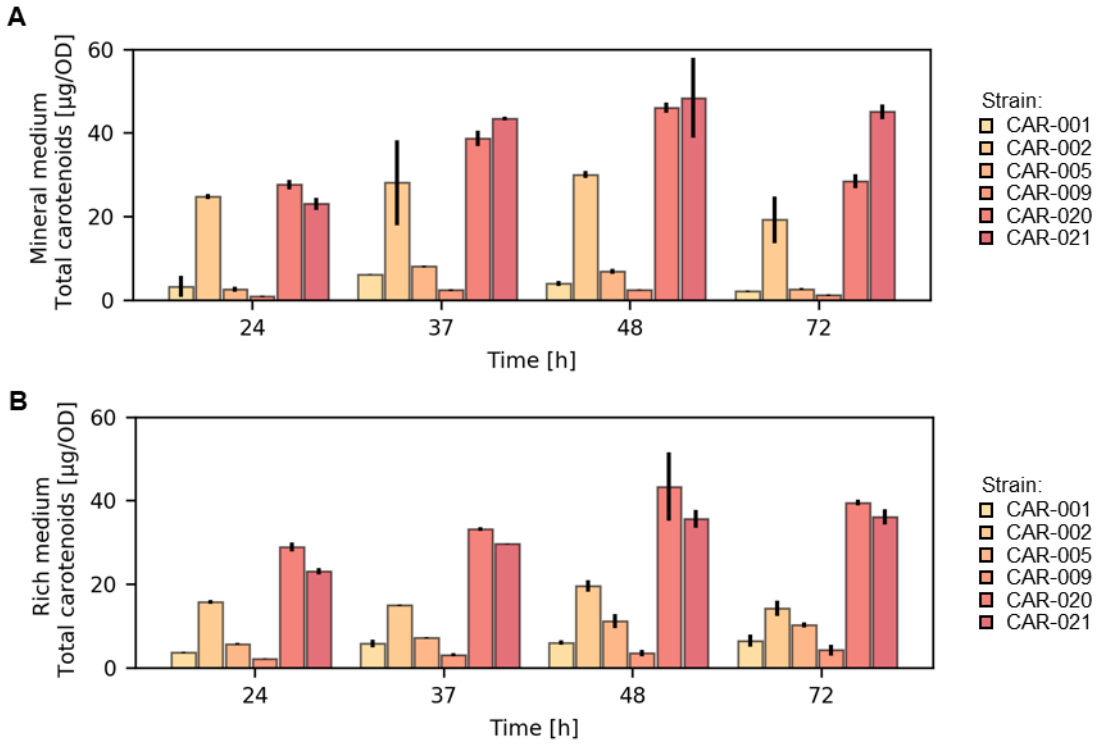


Figure 2-4. Total carotenoids production by *S. cerevisiae* strains in A. mineral medium and **B**. rich medium. Total carotenoids production is the sum of phytoene, lycopene and β-carotene. Data represents mean ± standard deviation (n=2).

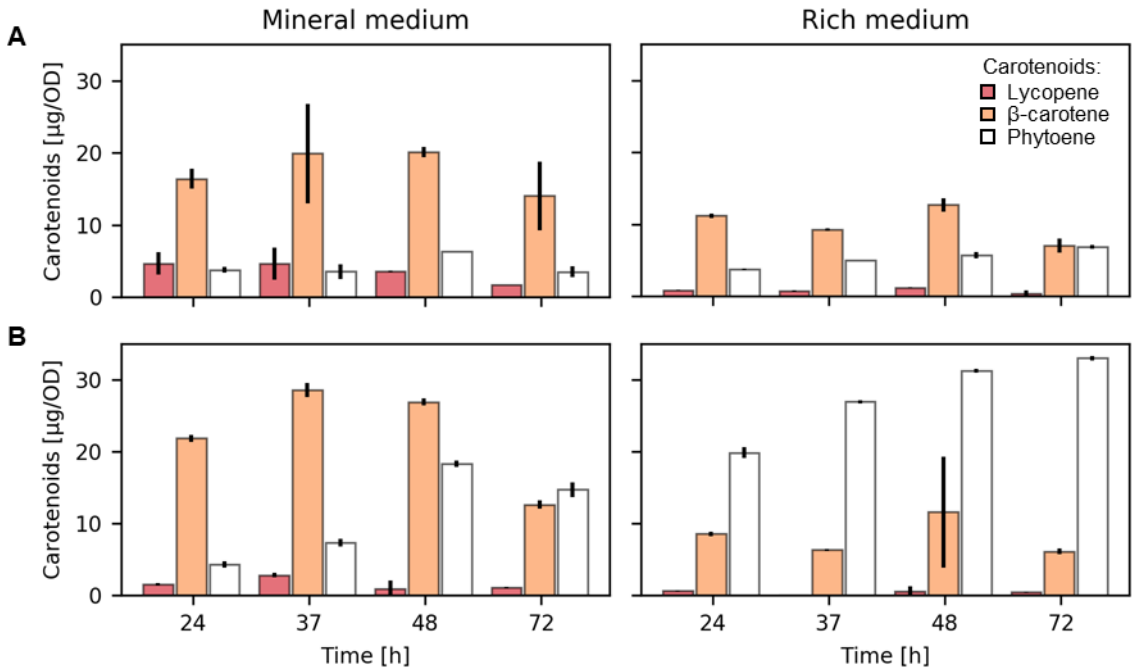


Figure 2-5. Phytoene, lycopene and β-carotene production profiles by strains A. CAR-002 and **B**. CAR-020 in mineral (left) and rich medium (right). Data represents mean ± standard deviation (n=2).

Strain CAR-002 produced 3.2-5.6 times more β-carotene than phytoene in the mineral medium whereas in the rich medium this ratio dropped to 2.4 for 24–48 h and equal amounts of β-carotene and phytoene after 72 h. Cultivation of CAR-002 in a mineral medium led to accumulation of lycopene which accounted for 9-19% of total carotenoids produced, compared to 2-6% during fermentation in a rich medium. In the first day of fermentation in the mineral medium, the amount of β-carotene produced by strain CAR-020 was 5-times higher than phytoene (21.8±0.3 and 4.3±0.3 μg/OD, respectively). However, the ratio between β-carotene and phytoene decreased as the fermentation proceeded to similar levels after 72 h (12.6±0.4 μg/OD β-carotene and 14.7±0.7 μg/OD phytoene). In contrast, strain CAR-020 cultivated in the rich medium produced 2-times more phytoene than β-carotene. A slight lycopene accumulation was also observed for CAR-020 (2-7% of total carotenoids in mineral and <2% in rich medium).

Colour trait

As natural pigments, carotenoids exhibit different colours ranging from yellow, orange to red. These colours are a consequence of conjugated double bonds in the molecule[105]. Microbial production of β-carotene in *S. cerevisiae* results in a variety of coloured colonies ranging from yellow to orange, with the intensity of the colour linked to the production level. Strain CAR-009 which produced low levels of carotenoids displayed a pale yellow colour after culturing and plating on rich agar (Figure 2-6, first column). Equipping strain CAR-005 with medium strength promoters to control *crt* genes resulted in higher production of carotenoids and dark yellow coloured colonies. Colonies formed by medium and high producing strains (CAR-002 and CAR-020, respectively) were orange in colour. A similar colour of these strains and lack of increased colour intensity of strain CAR-020 in comparison with CAR-002 can be explained by the nature of the produced carotenoids. β-carotene has an orange colour whereas the intermediate of its biosynthesis pathway, lycopene is red[105]. Therefore, red colonies would indicate accumulation of lycopene and insufficient conversion to β-carotene[118]. Phytoene is an intermediate in β-carotene biosynthesis pathway, which is colourless and therefore it has no impact on colonies colour. Finally, strain CAR-021 producing high amounts of carotenoids formed smaller colonies than any other strain. Carotenoids production by strains CAR-020 and CAR-021 was similar, therefore difficulty in colony formation might be due to protein expression rather than carotenoids cytotoxicity. Overall, two additional days of incubation were required for strains plated on mineral agar to obtain the same colony sizes as for the rich medium.

Genetic instability

A production of a heterologous compound can have a negative impact on the growth of a host organism through imbalance in cellular processes and it can impose additional load if toxic compounds accumulate. To overcome production burden, host organisms can evolve loss-of-function mutations or a complete loss of the genes in the production pathway, a phenomenon referred as genetic instability[23,25,126,129,130]. With the colour trait of carotenoids producing cells, indication of a genetic instability can be visually assessed. Cultivation in a mineral medium prior to plating on agar caused the appearance of white colonies among the orange colonies of medium and high carotenoids producing strains CAR-002 and CAR-020 (Figure 2-6). White colonies formed on rich agar after cultivation in mineral medium accounted for ~75% of total colonies for strain CAR-002 and <20% for CAR-020. In addition, formation of white colonies of strain CAR-

002 was also observed when rich medium and rich agar was used (3%). The higher degree of instability of strain CAR-002 as compared to CAR-020 may be caused by the difference in genotype, as the number of *crtI* cassettes was increased from one to two in strain CAR-020. Cultivation of strain CAR-002 in a mineral medium followed by plating on the same type of agar resulted in a single orange colony among pale-pink colonies, which may indicate imbalance between different carotenoids, such as lycopene. Change of media from rich to mineral challenged adaptation of strain CAR-002 as only a single colony was obtained. In contrast, strain CAR-020 produced numerous orange colonies moving from rich to mineral media. The instability of the carotenogenic pathway was not observed for strains producing low amounts of carotenoids (CAR-005 and CAR-009). However, strain CAR-001 with *crt* genes expressed from high strength promoters but non-codon optimized for expression in *S. cerevisiae* formed some white colonies (1.3-4.8%) under all tested conditions. In summary, prolonged cultivation (72 h) in the mineral liquid medium of strains producing moderate (CAR-002) and high (CAR-020) amounts of carotenoids caused instability in the production displayed by a change in the colonies' colour from orange to almost white.

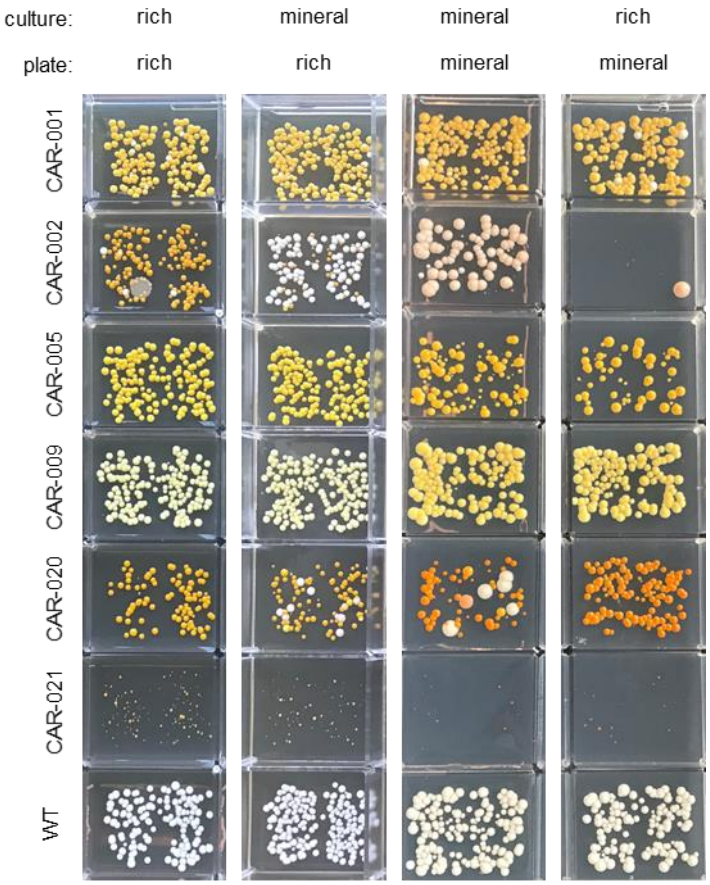


Figure 2-6. Genetic instability of carotenogenic strains. Strains were cultured for 72 h in either mineral or rich medium and plated on mineral or rich agar in equal dilutions. Experiment performed twice on independent days.

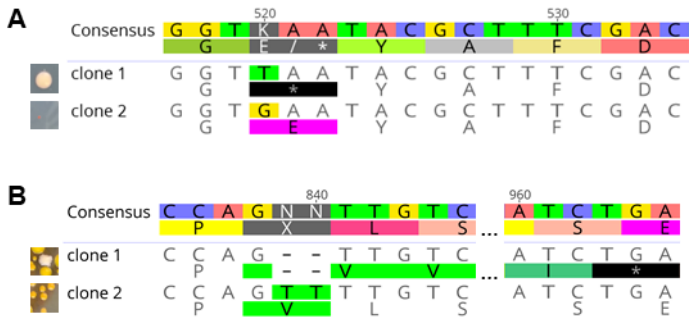


Figure 2-7. Mutations in *crt* expression cassettes of strains A. CAR-002 and **B**. CAR-020. On the left: photography of colonies subjected to sequence analysis of expression cassettes encoding *crt* genes. For each strain one pale yellow and one orange colony was analysed. On the right: sequence of *crtYB* gene analysed by nanopore sequencing. Upper sequence: mutated sequence. Lower sequence: native sequence.

To investigate whether the growth conditions that resulted in a loss of colour of the carotenoid producing colonies led to mutations in genes involved in carotenoids biosynthesis, we sequenced the genomic regions in which the *crt* expression cassettes were integrated. Four independent colonies displaying different colours (orange and pale-yellow) of strains CAR-002 and CAR-020 were subjected to sequencing (SI Figure 2-2). Analysis of pale-coloured clones revealed mutations within the *crtYB* gene resulting in a premature stop codon in both strains (Figure 2-7). The clone of CAR-002 exhibited a substitution 520G>T, whereas the clone of CAR-020 contained a double deletion 839_840delTT causing a frameshift and the consequent occurrence of 29 premature stop codons (SI Table 2-4). No additional mutations were found in *crtE* nor *crtI* for clones with defective *crtYB*, although mutations in *crtE* and *crtI* could be found if more colonies were analysed.

2.4 Discussion

Fermentation in a mineral medium led to a worse growth of strains with medium (CAR-002) and high (CAR-020, CAR-021) production of carotenoids which was not observed for the strains producing low amounts of carotenoids (CAR-001, CAR-005, and CAR-009), reaching similar optical density and dry weight as the wild type strain. When a rich medium was used, growth defects were only noted for strain CAR-021. Improved performance of strains CAR-002 and CAR-020 in a rich medium and shorter duration of exponential phase when compared to a mineral medium can be explained by the availability of biosynthetic precursors which can be directly metabolized but must be synthesized during cultivation in mineral medium[126]. These results suggest that the production of high amounts of carotenoids in *S. cerevisiae* compromises microbial growth. A growth defect has also been observed after attempting to improve lycopene and astaxanthin production in *S. cerevisiae*[33,121]. Furthermore, glucose consumption of strains with moderate and high production decreased in a mineral medium, indicating slower metabolism compared to the unmodified wild type strain.

Production of carotenoids can be regulated by the strength of the promoters expressing the *crt* genes and the copy number of these genes. Substitution of weak promoters controlling *crtE, crtYB,* and *crtI* in strain CAR-009 with medium strength in CAR-005 increased total carotenoids by 64% and replacement

with strong promoters in CAR-002 improved the production by 93% (based on growth in mineral media for all time points). Carotenoids belong to isoprenoids which are synthesised from precursors produced by the mevalonate pathway. Thus, the yield of microbial production can be improved by optimizing the mevalonate pathway and increasing the pool of IPP and DMAP (Figure 2-1). This approach was used here to construct strains CAR-020 and CAR-021 which increased carotenoids by 27% and 36%, respectively when compared to strains not overexpressing *tHMG1* or only including a single copy of carotenogenic genes (CAR-002, mineral medium, all time points). Although the use of a mineral medium impacted growth and metabolism of strains with medium/high production, this growth condition led to higher levels of β-carotene and reduced accumulation of phytoene. Cultivation of CAR-002 in a mineral medium caused accumulation of lycopene which was reduced when two copies of *crtYB* were integrated into the genome in strain CAR-020, indicating insufficient conversion rate of lycopene into β-carotene in CAR-002. Balancing metabolic pathways to reduce the accumulation of intermediates enhances the yield of the final product and prevents the potential toxicity of these compounds. Among the intermediates in β-carotene biosynthesis, previous reports indicated lycopene as particularly toxic to yeast cells[33,118]. Lycopene overproduced in *S. cerevisiae* is accumulated in cell membrane resulting in membrane stress[33,118]. Evaluation of the cell membrane of β-carotene producing *S. cerevisiae* showed decreased cell membrane fluidity caused by a reduction in unsaturated fatty acid and ergosterol content[117]. The toxicity of carotenoids was also demonstrated by activation of transporters belonging to the pleiotropic drug resistance family which facilitate the export of harmful compounds from the cell[125]. Furthermore, membrane physiology may be imbalanced as a consequence of competition for FPP between biosynthesis of β-carotene and ergosterol, an essential compound of the yeast cell membrane.

Production burden and/or toxicity can also be illustrated by the stability of a heterologous production pathway genes. Carotenoids production in *S. cerevisiae* results in orange coloured cells, therefore a visual inspection of colour loss can be the first indication of loss or mutation in the pathway genes. Cultivation in a mineral medium led to the instability in the production of carotenoids observed for strains producing medium and high levels of carotenoids. Co-expression of the *tHMG1* gene and *crtI* from two gene copies (CAR-020) mitigated the instability. However, formation of smaller colonies was observed when an additional copy of *crtYB* was included (CAR-021). Alteration of colony colour to white has been previously reported for *S. cerevisiae* strains overproducing lycopene and astaxanthin[33,34], however, the source of defective production was not elucidated in these studies. We analysed the DNA sequence of *crt* genes in two pale-yellow clones which revealed premature stop codons in *crtYB* blocking biosynthesis of phytoene. Mutations in other genes that would disable the production pathway are possible and could be found if a higher number of colonies was analysed. Mutants that lose the potency for production may restore fitness and these mutants may overgrow a cellular population with unfavourable fitness but good productivity, a situation that poses a threat in a large scale fermentation.

In this chapter, a set of *S. cerevisiae* strains with different levels of β-carotene production were characterized in terms of growth, metabolism, and physiology. A growth defect and instability of one of the introduced pathway genes was observed for strains with at least moderate production. Our findings provide an insight into the challenges of the microbial production of carotenoids which will be tackled with a feedback control system described in the next chapters.

3 CRISPR-Cas12a Multiplex Genome Editing of *Saccharomyces cerevisiae* and the Creation of Yeast Pixel Art

This Chapter was published at the Journal of Visualized Experiments[66] and it is accompanied by a video available at https://www.jove.com/embed/player?id=59350&t=1&s=1&fpv=1 . The experimental work was conducted by me and Brenda Vonk (DSM). Operation of the acoustic liquid handler was supported by Jeffrey van Wijk (DSM). The script for transforming a photography in a full colour scale to a 4-colours picture was written by Zi Di (DSM). The work was supervised by René Verwaal (DSM), Thomas E. Gorochowski (University of Bristol) and Johannes A. Roubos (DSM). This work describes a detailed protocol for CRISPR-Cas12a (Cpf1) mediated genome editing of *S. cerevisiae*, previously published by René Verwaal, Johannes A. Roubos *et al.*[65].

3.1 Introduction

CRISPR-Cas enzymes have unquestionably revolutionized molecular biology and been widely adopted as tools for engineering genomes at a speed that was previously unfeasible[131]. The first modification of a *Saccharomyces cerevisiae* genome by the CRISPR-Cas9 genome editing system was reported by DiCarlo *et al.*[49], demonstrating successful gene knock-out and making point mutations using externally introduced oligonucleotides. Further yeast CRISPR toolbox developments included: transcriptional regulation by fusion of catalytically inactive dead Cas9 (dCas9) with transcriptional effector domains to enable activation and silencing of transcription[93], application for both genome editing and regulatory functions for metabolic pathway engineering by simultaneous activation, repression and deletion[94], deletion of large fragments from the *S. cerevisiae* genome[132], and multiple-chromosome fusions[133]. CRISPR-Cas genome editing systems find their origin in adaptive immune systems of bacteria and archaea and these systems have been adapted by molecular biologists for genome editing. Their functionality is based on the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) DNA regions encoding RNA responsible for the recognition of the foreign DNA or RNA and the CRISPR associated genes (Cas) which encodes RNA-guided endonucleases[7,68,131,134]. Based on the recent genome analysis of CRISPR-Cas systems it was proposed to divide the CRISPR-Cas systems into two classes, five types and 16 subtypes[73]. The two classes are distinguished based on the organization of effector complexes involved in target cleavage. Typically, CRISPR-Cas systems with a multi-subunit organization are categorized as class 1, whereas single subunit effector complexes belong to class 2[47,73]. In this paper, we explore the class 2 type V Cas12a, formerly called Cpf1[73,74], which is an alternative to the class 2 type II Cas9. Although Cas9 is well-characterized and widely used in research, Cas12a offers additional features[74]. Firstly, Cas12a forms a complex with CRISPR RNA (crRNA) of 42 to 44 nucleotides without requiring an additional trans-activating CRISPR RNA (tracrRNA). Therefore, a shorter guide RNA can be utilized in genome editing with CRISPR-Cas12a systems compared to CRISPR-Cas9. Secondly, the unique endonuclease and endoribonuclease activity of Cas12a enables maturation of its pre-crRNA[75]. This RNase activity allows for the encoding of multiple crRNAs on a single CRISPR crRNA array, whereas Cas9 requires the separate expression of each so-called single guide RNA (sgRNA) or alternatively for example expression of an additional endonuclease (*e.g.* Csy4) in combination with recognition motifs for Csy4 surrounding each sgRNA[135,136]. Thirdly, Cas12a target site recognition requires a protospacer adjacent motif (PAM) at the 5' end from the target and cleaves after the +18/+23

position from its PAM resulting in cleaved DNA with sticky ends, whereas Cas9 requires a PAM located on the 3' end from the target and cleaves after the -3 position creating blunt end cuts in the DNA[74]. Fourthly, the consensus nucleotide sequence of the PAM differs between Cas12a ((T)TTV) and Cas9 (NGG), which makes Cas12a a promising candidate for targeting T-rich promoter and terminator sequences[137]. Finally, a recent study reported greater target specificity for Cas12a than for the native Cas9[138]. We present a protocol for using the CRISPR-Cas12a system for genome editing of *S. cerevisiae* with a particular focus on the introduction of multiple DNA expression cassettes into independent genomic loci simultaneously (multiplex genome editing) using a single crRNA array. The key steps of the protocol are depicted in Figure 3-1.

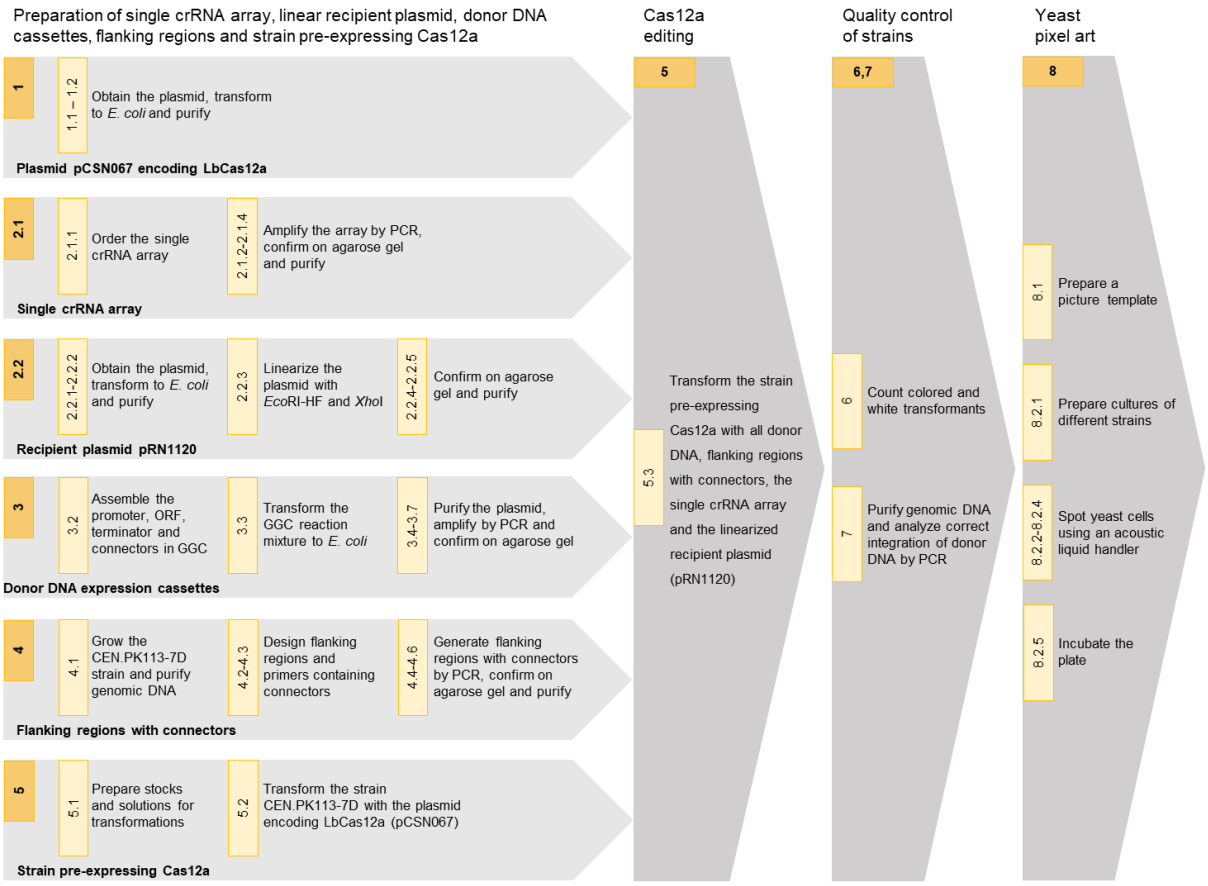


Figure 3-1. Workflow of the protocol for CRISPR-Cas12a multiplex genome editing in *S. cerevisiae*. The workflow includes crucial steps of the presented method. For details see the Protocol.

As a proof of concept, the CRISPR-Cas12a system was applied for introduction of three expression cassettes into the genome of *S. cerevisiae* which enable the production of β-carotene[21] as schematically shown in Figure 3-2. Production of β-carotene affects the phenotype of *S. cerevisiae*: *i.e.* upon successful introduction of all three heterologous genes required for carotenoids biosynthesis, the white *S. cerevisiae* cells turn yellow or orange, depending on the expression strength of each gene's promoter. Due to the simple visual read-out of this pathway, it has been introduced to develop advanced CRISPR-based systems and methods for genome editing[65,139]. In this work, expression cassettes encoding the carotenoid genes *crtE*, *crtYB* and

crtI have been constructed using a Golden Gate cloning (GGC) approach[140] with heterologous promoters and homologous terminators used to drive expression of the genes. The expression cassettes are surrounded by unique 50 base pairs (bp) sequences, called connectors, that allow for *in vivo* assembly with integration flank DNA sequences (flanking regions) with the same 50 bp sequences, and subsequent integration into the genomic DNA of yeast at the position determined by the flanking regions. By using different promoter strengths, strains with different levels of carotenoids production were obtained resulting in variation in colour of the cells. These strains - inspired by the "Yeast Art Project"[141] - were used in a spotting setup with an acoustic liquid handler to create a 4-colour high-resolution "yeast photograph" of Rosalind Franklin. Franklin (1920-1958) was an English chemist and X-ray crystallographer well known for her contribution to the discovery of the DNA structure by Photo 51[142–144].

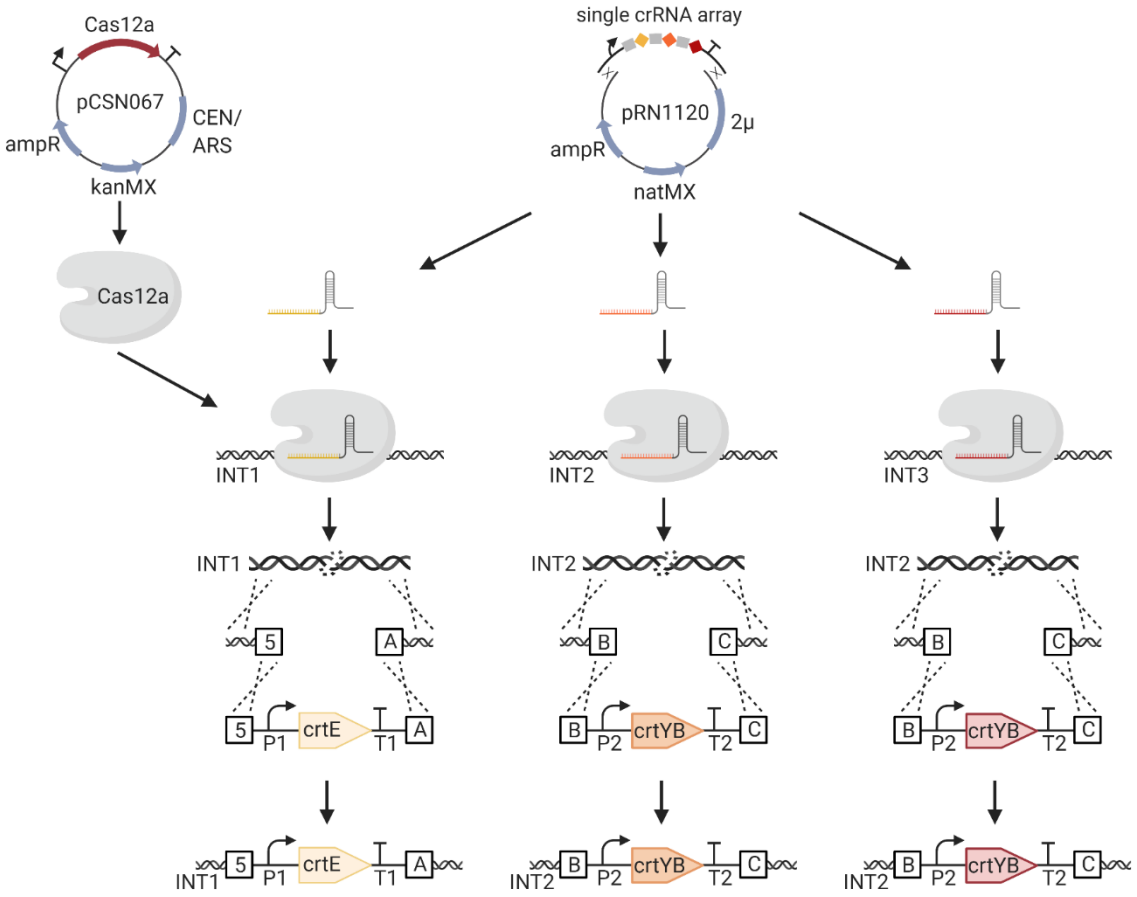


Figure 3-2. Scheme of CRISPR-Cas12a multiplex genome editing using a single crRNA array. The single crRNA array is composed of three crRNAs units in their mature form, a 20 nt direct repeat specific for LbCas12a (grey squares) with a 23 nt guide sequence (coloured diamonds). Expression of the crRNA array is enabled by the *SNR52* promoter and *SUP4* terminator. Transformation of *S. cerevisiae* with a linearized pRN1120 and the single crRNA array expression cassette containing homology with pRN1120 (diagonal stripes) allows for *in vivo* recombination into a circular plasmid in cells pre-expressing LbCas12a. The single crRNA array is subsequently processed by Cas12a. Cas12a is directed to the intended INT1, INT2 and INT3 genomic target sites and creates double stranded breaks. In the transformation mixture, donor DNA consisting of flanking regions and the carotenoid gene expression cassette were included. Donor DNA assemblies were targeted to one stretch of DNA in genomic DNA around the INT1 (*crtE*), INT2 (*crtYB*) and INT3 (*crtI*) loci by *in vivo* recombination due to the presence of 50 bp homologous connectors sequences, indicated as 5, A, B, C, D or E. P1–P3, different promoters; T1–T3, different terminators. This figure has been modified from [65].

3.2 Protocol

1. Preparation of the *Cas12a* plasmids

Note: The plasmid containing the *Lachnospiraceae bacterium* ND2006 *Cas12a* (*LbCpf1*, pCSN067) codon optimized for expression in *S. cerevisiae*, was previously constructed[65], deposited at a plasmid repository (see SI Table 3-1). This is a single-copy episomal *S. cerevisiae*/*E. coli* shuttle plasmid containing a KanMX resistance marker gene to allow for selection of *S. cerevisiae* transformants on geneticin (G418).

1.1. Amplify pCSN067 plasmid to obtain a high amount.

1.1.1. Transform 25 µL of purchased chemically competent *E. coli* cells with plasmid pCSN067 according to manufacturer's protocol. Dilute the transformation mix 10 and 50 times in 2x peptone-yeast (PY). Plate out 10 and 50 times dilutions on 2x PY agar plates containing ampicillin (0.1 g/L) and incubate overnight at 37 °C.

1.1.2. Pick 2 to 3 colonies and inoculate each colony in 3 mL of 2x PY and grow overnight at 37 °C in a shaking incubator at 180 rpm.

1.1.3. Purify the plasmid using a plasmid purification kit according to manufacturer's instructions.

2. Preparation of the single crRNA array expression cassette

2.1. Prepare the single crRNA array

Note: The single crRNA array comprises an *SNR52* RNA polymerase III promoter from *S. cerevisiae*[49], a direct repeat specific for LbCas12a and a spacer (genomic target sequence), together repeated for each target[65] and ends with a *SUP4* terminator from *S. cerevisiae*[49]. The single crRNA array is assembled by *in vivo* recombination into the linearized plasmid pRN1120 to generate a circular plasmid, thus regions homologous to plasmid pRN1120 must be present at the start and end of the single crRNA array (see Figure 3-2). It is recommended to in advance evaluate the functionality of a number of designed crRNAs separately[65]. This information is subsequently used to select most functional crRNAs to combine these into the direct repeat and spacer sequences to create a single crRNA array for the multiplexing purpose.

2.1.1. Order the single crRNA array for multiplex genome editing experiments as synthetic DNA (see the DNA sequence of the single crRNA array in SI Table 3-2).

2.1.2. Amplify the ordered single crRNA array, *e.g.* using primers KC-101 and KC-102 (SI Table 3-3). Prepare the PCR amplification mix containing: 0.5 µL of DNA polymerase, 10 µL of 5x buffer required for the DNA polymerase, 1 µL of 10 mM dNTPs, 2.5 µL of 10 µM forward primer, 2.5 µL of 10 µM reverse primer, 2 µL of DNA template at a concentration of 5 ng/µL and ultrapure H_2O up to a total volume of 50 µL.

Perform the reaction in a thermocycler using the following program: (i) 98 °C for 3 min, (ii) 98 °C for 10 sec, (iii) 60 °C for 20 sec, (iv) 72 °C for 15 sec – repeat steps (ii) to (iv) 30 times, (v) 72 °C for 5 min (vi) hold at 12 °C until further analysis.

2.1.3. Analyse the PCR products by electrophoresis by running the samples on a 0.8% agarose gel at 5 V/cm for 40 min using a DNA loading dye and DNA ladder with DNA fragments in a range of 100 to 10,000 bp.

2.1.4. Purify the PCR products using a PCR purification kit according to the instructions of the manufacturer.

2.2. Prepare the single crRNA array recipient plasmid

Note: The single crRNA array is expressed from the *S. cerevisiae*/*E. coli* shuttle plasmid pRN1120[65] (see SI Table 3-1). This multi-copy plasmid contains a NatMX resistance marker gene to allow selection of *S. cerevisiae* transformants on nourseothricin (NTC).

2.2.1. Obtain pRN1120 plasmid.

2.2.2. Amplify pRN1120 plasmid to obtain a high amount.

2.2.2.1. Transform 25 µL of purchased chemically competent *E. coli* cells with plasmid pRN1120 according to manufacturer's protocol. Dilute the transformation mix 10 and 50 times in 2x PY. Plate out 10 and 50 times dilutions on 2x PY agar plates containing ampicillin (0.1 g/L) and incubate overnight at 37 °C.

2.2.2.2. Pick 2 to 3 colonies and inoculate each colony in 3 mL of 2 x PY and grow overnight at 37 °C in a shaking incubator at 180 rpm.

2.2.2.3. Purify the plasmid using a plasmid purification kit according to manufacturer's instructions.

2.2.3. Linearize plasmid pRN1120 with EcoRI-HF and XhoI. For this, prepare a digestion mix composed of 1 µg of pRN1120, 5 µL of 10x buffer (1x buffer contains 50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 100 µg/ml bovine serum albumin [BSA]; pH 7.9), 1 µL of EcoRI-HF (20 U), 1 µL XhoI (20 U) and ultrapure H_2O up to a total volume of 50 µL. Incubate the digestion mix at 37 °C for 2 hours and inactivate at 65 °C for 20 min.

2.2.4. Analyse the linearized plasmid by electrophoresis on an agarose gel (0.8%, 40 min, 5 V/cm) using a DNA loading dye and DNA ladder with DNA fragments in a range of 100 to 10,000 bp. As a control include circular plasmid in the analysis.

2.2.5. Purify the linearized plasmid using a PCR purification kit according to the instructions of the manufacturer.

3. Preparation of Promoter-ORF-Terminator (POT) donor DNA constructs

3.1. Order a set of promoter (P) of different strength, open reading frame (O) and terminator (T) sequences as synthetic DNA such that each element contains standardized 4 bp recognition sequences that are flanked by BsaI sites to enable Golden Gate Cloning (GGC) assembly[145] (see the detailed designs in SI Table 3-4 and sequences in SI Table 3-5).

3.2. Assemble POT expression cassettes composed of a promoter, open reading frame, terminator and connectors sequences via a 4-part assembly using a GGC reaction[140], into a destination vector that already contains pre-specified 50 bp connectors sequences (see SI Table 3-5 and references [145,146]).

3.2.1. Measure the concentration of DNA parts using a spectrophotometer. Dilute each DNA part in ultrapure H_2O to a final concentration of 15 fmol/µL.

3.2.2. Prepare a reaction mix composed of DNA fragments: 2 µL of promoter, 2 µL of open reading frame, 2 µL of terminator and 2 µL backbone (Level 1 destination vectors as described in [145]), 4 µL of 5x T4 DNA ligase buffer, 2.5 µL of 1U/µL T4 DNA Ligase, 1.5 µL of 20 U/µL BsaI-HF and ultrapure H_2O up to a total volume of 20 µL.

3.2.3. Perform the GGC reaction in a thermocycler using the following program: (i) 37 °C for 2 min, (ii) 16 °C for 5 min – repeat steps (i) and (ii) 50 times, (iii) 50 °C for 60 min, (iv) 80 °C for 45 min, (v) hold at 12 °C until further analysis.

3.3. Transform 25 µL of purchased chemically competent *E. coli* [147] cells with 3 µL of the GGC reaction mix according to manufacturer's protocol. Dilute the transformation mix 10 and 50 times in 2x PY. Plate out 10 and 50 times dilutions on 2 x PY agar plates containing ampicillin (0.1 g/L) and incubate overnight at 37 °C.

3.4. Pick 2 to 3 colonies and inoculate each colony in 3 mL of 2 x PY and grow overnight at 37 °C in a shaking incubator at 180 rpm.

3.5. Purify the plasmids using a plasmid purification kit according to manufacturer's instructions.

3.5.1. Check if POT expression cassettes were assembled correctly in the GGC reaction by PCR. Design primers complementary to the connector sequence present at the start and the end of each expression cassette (see Figure 3-2). For connectors chosen in this protocol use primers KC-103 to KC-108 (see SI Table 3-3).

3.5.2. Prepare PCR amplification mixes for each plasmid containing: 0.5 µL of proofreading DNA polymerase, 10 µL of 5x buffer required for the DNA polymerase, 1 µL of 10 mM dNTPs, 2.5 µL of 10 µM forward primer, 2.5 µL of 10 µM reverse primer, 2 µL of DNA template with a concentration of 5 ng/µL, and ultrapure H_2O up to a total volume of 50 µL.

3.5.3. Perform the PCR reaction in a thermocycler using the following program: (i) 98 °C 3 min, (ii) 98 °C 10 sec, (iii) 60 °C 20 sec, (iv) 72 °C 2 min 30 sec – repeat steps (ii) to (iv) 30 times, (v) 72 °C 5 min, (vi) hold at 12 °C until further analysis.

Note: Resulting PCR products consist of 50 bp of the 5' connector, promoter, open reading frame, terminator and 50 bp of the 3' connector.

3.6. Analyse the PCR products by electrophoresis by running samples on a 0.8% agarose gel at 5 V/cm for 40 min using a DNA loading dye and DNA ladder with DNA fragments in a range of 100 to 10,000 bp.

4. Preparation of integration flank DNA sequences containing connectors sequences

4.1. Purify genomic DNA from wild type *S. cerevisiae* CEN.PK113-7D[128].

4.1.1. Grow the strain in a 500 mL shake flask filled with 100 mL yeast extract peptone dextrose (YEPD, 2% glucose) medium at 30 °C and shaking at 250 rpm for 48 hours.

4.1.2. Harvest the cells by centrifugation of 2 mL broth at 16,000 g for 1 min and discard the supernatant. Resuspend the cells in physiological salt (200 µL; 0.85% NaCl solution) with RNase (10 µL, 10 mg/mL) and yeast lytic enzyme (4 µL). Incubate the cell suspension at 37 °C for 15 min.

4.1.3. Add 300 µL of cell lysis solution (see SI Table 3-1) and vortex shortly.

4.1.4. Add 168 µL of protein precipitation solution (see SI Table 3-1) and vortex vigorously for 20 s.

4.1.5. Separate the protein fraction by centrifugation at 16,000 g and 4 °C for 10 min. Collect 600 µL of supernatant in a new tube and mix with 600 µL of isopropanol and vortex shortly.

4.1.6. Recover DNA by spinning down at 16,000 g at room temperature for 10 min. Discard the supernatant and keep the pellet.

4.1.7. Wash the pellet with 200 µL of ethanol (70%). Centrifuge at 16,000 g at room temperature for 10 min and remove the supernatant. Evaporate the ethanol by incubating the tube at room temperature for 10 min with the lid opened.

Note: If liquid in the tube is still visible, repeat the step 4.1.8. Do not dry the pellet for longer than 10 min to prevent decreased solubility of the DNA.

4.1.8. Dissolve DNA in 50 µL of TE buffer. Store purified DNA at –4 °C.

4.2. For each integration site, design integration flank DNA sequences (approx. 500 bp) such that approximately 1000 bp of genomic DNA will be removed upon introduction of donor DNA (see the schematic design in Figure 3-2 and sequences in SI Table 3-5).

4.3. Design primers to generate the flanking regions by PCR.

4.3.1. For the **left flanking region**, design forward and reverse primers to amplify approximately 500 bp of the genomic DNA region positioned 5' (left) of the integration site of interest.

Note: The forward primer includes 20 nt of homology with the intended flanking region. The reverse primer includes 20 nt with homology with the intended flanking region and contains the desired 50 nt connector sequence to enable *in vivo* assembly in the Cas12a editing on the genome later on.

4.3.2. For the **right flanking region**, design forward and reverse primers to amplify approximately 500 bp of the genomic DNA region positioned 3' (right) of the integration site of interest.

Note: The forward primer includes 20 nt with homology with the intended flanking region and contains the desired 50 nt connector sequence to enable *in vivo* assembly in the Cas12a editing on the genome later on. The reverse primer includes 20 nt of homology with the intended flanking region.

4.4. Amplify the flanking regions with the designed primers (*e.g.* primers KC-109 to KC-120 enclosed in SI Table 3-3).

4.4.1. Measure the concentration of purified genomic DNA that will serve as template in the PCR. Adjust the DNA concentration to 50 ng/µL.

4.4.2. Prepare PCR amplification mixes composed of genomic DNA (1 – 4 µL of 50 ng/µL genomic DNA dilution) purified in step 4.1, forward and reverse primer (10 µM each), 1 µL of 10 mM dNTPs, 10 µL of 5 x buffer required for the DNA polymerase, 0.5 µL of DNA polymerase (1.0 U), and ultrapure H_2O up to total volume of 50 µL.

4.4.3. Perform PCRs in a thermocycler using following program: (i) 98 °C 3 min, (ii) 98 °C 20 sec, (iii) 60 °C 20 sec, (iv) 72 °C 15 sec, repeat steps (ii) to (iv) 30 times, (v) 72 °C 5 min, (vi) hold at 12 °C until further analysis.

4.5. Analyse the PCR products by electrophoresis on a 0.8% agarose gel at 5 V/cm for 40 min using a DNA loading dye and DNA ladder with DNA fragments in a range of 100 to 10,000 bp.

4.6. Purify the correct PCR products using a PCR purification kit according the instructions of the manufacturer.

5. Transformation to *S. cerevisiae*

Note: Perform transformation using a protocol based on the methods developed by Gietz *et al*. (1995)[148] and Hill *et al.* [149] which can be used for various strains of *S. cerevisiae*. The protocol described below is sufficient for 1 transformation.

5.1. Prepare solutions required for transformation.

5.1.1. Prepare the following stock solutions and filter-sterilize.

10x TE buffer containing 100mM Tris-HCl (pH 7.5), 10 mM EDTA, total volume of 50 mL.

1 M LiAc at pH 7.5, total volume of 50 mL.

50% PEG 4000, total volume of 100 mL.

Note: Always check that PEG 4000 stock is at pH 5. This stock should not be stored longer than one month.

5.1.2. Prepare the following solutions using stocks:

Prepare LiAc-TE solution containing 0.1 M LiAc, 10 mM Tris-HCl, 1 mM EDTA, total volume of 0.5 mL.

Prepare PEG-LiAc-TE solution containing 40% PEG 4000, 0.1 M LiAc, 10 mM Tris-HCl, 1 mM EDTA, total volume of 1 mL.

Note: It is crucial for successful transformation that PEG-LiAc-TE and LiAc-TE solutions are freshly prepared.

5.2. First transformation round (prepare the strain pre-expressing Cas12a).

Note: In all the transformation steps, use water with a pH higher than 5. It is recommended to use demineralized water in all the steps of the transformation.

5.2.1. Prepare a pre-culture by growing strain CEN.PK113-7D in a 100 mL shake flask containing 20 mL of YEPD (2% glucose) medium and incubate overnight at 30 °C with shaking at 250 rpm.

5.2.2. Measure the OD_{600} of the pre-culture (OD_{pc}). Calculate the dilution factor (*df*) between the volume of pre-culture and the volume of fresh medium required for preparation of the cells pre-expressing Cas12a to be used in the transformation (transformation culture). In the calculations assume the optical density of the transformation culture (OD_{tc}) to be 1.0 after the incubation step described in 5.2.3 (*ti*).

$$df = \frac{\mathrm{OD_{pc}}}{\mathrm{OD_{tc}}} \cdot 2^{\frac{ti}{\tau}},$$

where *ti* and τ are the incubation time and doubling time, respectively.

5.2.2.1. Calculate the volume of the pre-culture (V_i) required for inoculation of the transformation culture (V_{tc}) based on the dilution factor.

$$V_i = \frac{V_{tc}}{df}$$

5.2.3. Prepare the transformation culture by inoculation of 20 mL of YEPD (2% glucose) (V_{tc}) with the volume of pre-culture determined in the previous step (V_i). Incubate at 30 °C with shaking at 250 rpm.

5.2.4. Measure the OD_{600} of the transformation culture until an OD_{600} of 1.0 is reached.

5.2.5. Harvest the cells by centrifugation of the 20 mL broth at 2,500 g for 5 min. Discard the supernatant and wash the cells in 20 mL room temperature demineralized water. Repeat the centrifugation step and keep the cell pellet.

5.2.6. Resuspend the cells in 100 µL LiAc-TE solution and transfer to a microcentrifuge tube.

5.2.7. Add 5 µL of single-stranded carrier DNA (10 mg/mL salmon sperm DNA) and mix by pipetting.

5.2.8. Pipette 1 µg of plasmid pCSN067 to the microcentrifuge tube.

Note: The total volume of the DNA mixture should not exceed 100 µL to prevent a lower transformation efficiency.

5.2.9. Add 600 µL of PEG-LiAc-TE solution and mix by pipetting. Incubate for 30 min at 30 °C while shaking at 450 rpm in a table top heat block.

5.2.10. Add 70 µL of DMSO (100%) to the transformation mixture and mix by pipetting. Perform heat-shock by incubating the transformation mixture at 42 °C for 15 minutes in a water bath.

5.2.11. Recover the cells by transferring the mixture to a 15 mL round bottom tube and add 10 mL YEPD (2% glucose) to the tube. Incubate overnight at 30 °C with shaking at 250 rpm.

5.2.12. Centrifuge the transformation mix at 2,500 g for 5 min. Discard the supernatant and resuspend the cell pellet in approximately 200 µL of the remaining solution.

5.2.13. Plate out 150 µL of the transformation mix and a 20 x dilution in YEPD (2% glucose) of transformation mix on YEPD (2% glucose) agar plates supplemented with 0.2 g/L G418. Incubate the plates at 30 °C for 48 – 72 hours.

5.2.14. Pick a single transformant and re-streak on a YEPD (2% glucose) agar plate supplemented with 0.2 g/L G418 to obtain single colonies.

5.3. Second transformation round (perform multiplex genome editing with CRISPR-Cas12a).

5.3.1. Prepare a pre-culture by growing the strain pre-expressing Cas12a, created in the first transformation round (step 5.2), in a 100 mL shake flask containing 20 mL YEPD (2% glucose) medium. Incubate overnight at 30 °C with shaking 250 rpm.

Note: For multiple transformations, adapt the volume of the pre-culture.

5.3.2. Follow the steps described in Sections 5.2.2 to 5.2.7 for the first transformation round.

Note: For multiple transformations, adapt the volumes of required solutions and culture of the strain pre-expressing Cas12a.

5.3.3. Pipette 1 µg of the single crRNA array, 1 µg of the linearized recipient plasmid for the crRNA array, 1 µg of donor DNA and 1 µg of each flanking region (step 4.3) in a microcentrifuge tube.

Note: The total volume of the DNA mixture should not exceed 100 µL to prevent a lower transformation efficiency.

5.3.4. Prepare the following controls for the transformation: negative control (ultrapure H_2O); positive control for determination of the transformation efficiency (1 µg circular pRN1120); a control verifying if introduction of donor DNA is conducted via CRISPR editing (1 µg circular pRN1120, 1 µg of all donor DNA expression cassettes and 1 µg flanking regions but no single crRNA array); control verifying if donor DNA can be integrated outside of target (1 µg linearized pRN1120, 1 µg of donor DNA expression

cassettes and 1 µg of the single crRNA array but no flanking regions); a control verifying full linearization of pRN1120 (1 µg linearized pRN1120).

5.3.5. Follow the steps described in Sections 5.2.9 to 5.2.12 for the first transformation round.

5.3.6. Plate out 150 µL of the transformation mix and 20 x dilution in YEPD (2% glucose) of transformation mix on YEPD (2% glucose) agar supplemented with 0.2 g/L G418 and 0.2 g/L NTC. Plate out controls on YEPD (2% glucose) agar supplemented with the appropriate selection (G418 and/or NTC or no selection). Incubate the plates at 30 °C for 48 – 72 hours.

5.3.7. Pick a single coloured transformant and re-streak on a YEPD (2% glucose) agar plate to obtain single coloured colonies.

6. Evaluation of the genome editing efficiency

6.1. Count the number of coloured colonies and white colonies on the transformation plates.

6.2. Calculate genome editing efficiency by dividing the number of coloured colonies by the total number of colonies (both white and coloured), as shown in Table 3-1.

7. Confirmation of integration of donor DNA at the intended loci

7.1. Re-streak a coloured single colony from a transformation plate on a YEPD (2% glucose) agar plate without G418 and NTC selection and incubate for 48 hours at 30 °C.

7.2. Pick a single colony and inoculate a 500 mL shake flask filled with 100 mL of YEPD (2% glucose) medium. Incubate for 48 hours at 30 °C and shaking at 250 rpm.

7.3. Isolate the genomic DNA as described in Section 4.1.

Note: Alternatively, use a protocol for preparation of yeast for colony PCR previously proposed by Looke *et al.* [150]. In this case, growth in liquid medium (Section 7.2) can be skipped.

7.4. Verify correct integration by amplification of two fragments per integrated expression cassette.

7.4.1. Design primers which anneal to genomic DNA outside of the transformed flanking regions and the gene of interest (see examples in SI Table 3-3, KC-121 to KC-132). When using primers KC-121 to KC-132, set the annealing temperature in the PCR program to 62 °C.

7.4.2. Amplify region of interest as described in Section 4.4.2. Adapt the PCR program, specifically adjust the time of the extension step in PCR according to the length of the template and manufacturer's recommendations for the DNA polymerase.

7.5. Check the size of the PCR products by electrophoresis on an agarose gel (0.8%, 40 min, 5 V/cm) using a DNA loading dye and DNA ladder with DNA fragments in a range of 100 to 10,000 bp.

8. Creation of yeast pixel art using an acoustic liquid handler

8.1. Prepare a picture template for the yeast pixel art.

8.1.1. Resize the original RGB picture (220 × 280 pixels, see the representative results), *e.g.* using ImageJ to create a final 64 × 96 pixels (width × height) grey-scale image visualized in intended colours (Results).

8.1.2. Convert the RGB picture into grey-scale using this formula:

$$I_{\mathrm{gr}} = \frac{(I_{\mathrm{r}}+I_{\mathrm{g}}+I_{\mathrm{b}})}{3},$$

where I_{gr}, I_r, I_g, I_b are the grey, red, green and blue intensities, respectively.

8.1.3. In order to categorize the pixels, develop an ImageJ plugin applying the following rules: (a) If I_{gr} is ≤ 64, use the dark orange yeast (strain 1, SI Table 3-4) for this pixel. (b) If 64 < I_{gr} ≤ 128, use the orange yeast (strain 2, SI Table 3-4) for this pixel. (c) If 128 < I_{gr} ≤ 192, use the yellow yeast (strain 3, SI Table 3-4) for this pixel. (d) If I_{gr} > 192, use the white yeast (CEN.PK113-7D) for this pixel.

8.2. Spot yeast cells to create the yeast pixel art.

8.2.1. Inoculate 500 mL shake flasks containing 100 mL of YEPD (2% glucose) medium with three differently coloured carotenoid producing *S. cerevisiae* strains and wild type CEN.PK113-7D. Incubate cultures overnight at 30 °C with shaking at 250 rpm.

8.2.2. Transfer 0.5 mL of the overnight culture to a tube filled with 0.5 mL of sterile non-ionic density gradient medium (see SI Table 3-1). Mix by vortexing briefly.

8.2.3. Transfer the cell suspension to a qualified reservoir, 2 × 3 well. Perform spotting using an acoustic liquid handler instrument from a qualified reservoir source plate to a microplate (see SI Table 3-1) containing 50 mL YEPD (2% glucose) agar. To simplify plating, define wells on plate, *e.g.* use a microplate as a 6144 well plate (64 × 96).

8.2.4. Spot 25 nL of each *S. cerevisiae* strain from the 2 x 3 well reservoir source plate using a .csv file with the fluid calibration setting 6RES_AQ_GPSA2 onto the destination microplate. Define each of these 25 nL droplets as a pixel in the 64 x 96 grid which is translated to the well positions (A01, B01, C01 etc.).

8.2.5. Incubate the microplate at 30 °C for 48 hours. To intensify the colours of the strains store the agar plate at 4 °C for at least 72 hours.

3.3 Results

The protocol for multiplex genome editing using CRISPR-Cas12a was demonstrated by constructing three carotenoid producing *S. cerevisiae* strains expressing the *crtE*, *crtYB* and *crtI* genes using heterologous promoters of high, medium and low strength: strain 1, 2 and, 3 respectively (SI Table 3-4). Construction of these strains required generation of three donor DNA expression cassettes and six flanking regions per strain for targeting to three different loci in genomic DNA (shown in Figure 3-2). As described herein, promoter, open reading frame, terminator and two contiguous 50 bp connectors sequences were assembled into an expression cassette via a Golden Gate Cloning reaction and the assembly was verified by PCR (Figure 3-3A). The single crRNA array was ordered as a synthetic DNA fragment and was amplified by PCR (Figure 3-3B). The recipient plasmid for the single crRNA array (plasmid pRN1120) was linearized with EcoRI-HF and XhoI and linearization was confirmed by electrophoresis (Figure 3-3C). The design and nucleotide sequences of the introduced donor DNA expression cassettes and flanking regions are shown in SI Table 3-5. The sequence of single crRNA array expression cassettes is provided in SI Table 3-2. Functionality of the spacers included in the single crRNA array was tested beforehand by singleplex genome editing with individual crRNAs[65].

The efficiency of genome editing using Cas12a was firstly evaluated based on the number of coloured colonies obtained after transformation (Table 3-1, Figure 3-4). The editing efficiency of the three

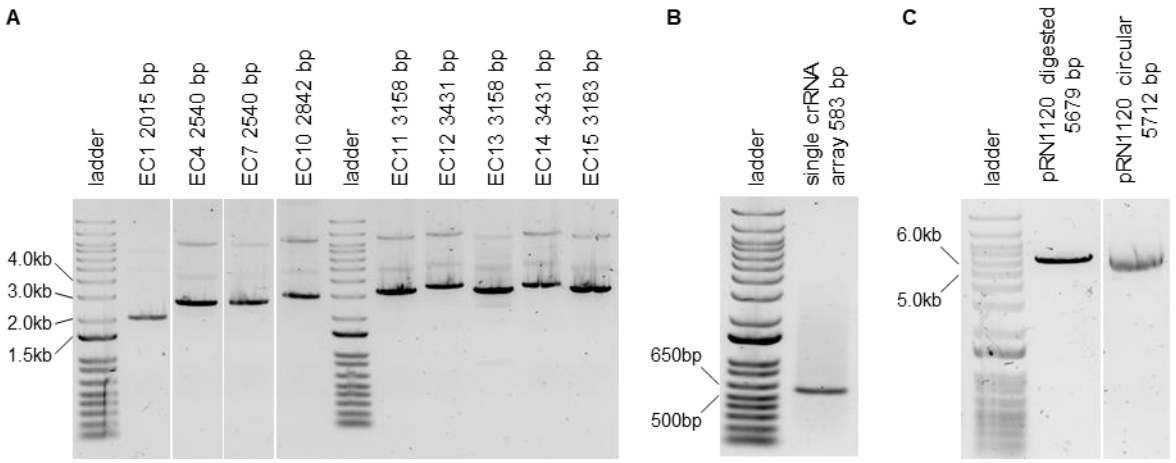


Figure 3-3. PCR verifying the genome editing experiments. **A**. Verification of Golden Gate Cloning reactions of assembled donor DNA cassettes. Obtained results are in agreement with expected lengths. **B**. PCR of the single crRNA array. **C**. Linearization of plasmid pRN1120.

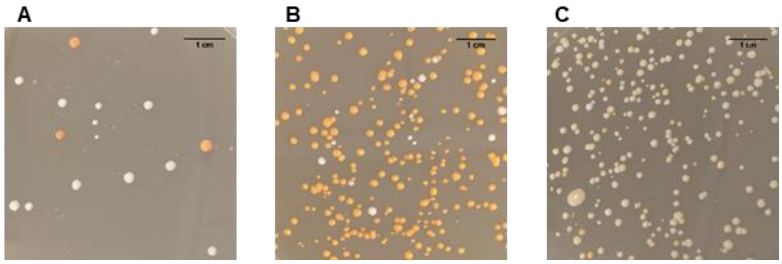


Figure 3-4. Plates of *S. cerevisiae* transformations using the multiplex genome editing approach. A. Strain 1 expressing *crtE*, *crtYB* and *crtI* from three strong promoters (dark orange colonies). **B**. Strain 2 expressing *crtE*, *crtYB* and *crtI* from three medium strength promoters (orange colonies). **C**. Strain 3 expressing *crtE*, *crtYB* and *crtI* from three low strength promoters (yellow colonies).

Table 3-1. Editing efficiency of the multiplex genome editing approach.

	Strain 1	Strain 2	Strain 3
Coloured colonies	16	279	220
White colonies	16	18	18
Total colonies	32	297	238
Efficiency	50%	94%	92%

constructed strains varied from 50% to 94%. Notably, introduction of expression cassettes used to generate strain 1 displayed the lowest editing efficiency, possibly caused by the nature of the donor DNA, *i.e.* these expression cassettes encode *crtE*, *crtYB* and *crtI* from three high strength promoters.

Secondly, correct integration of the three donor DNA expression cassettes at the intended loci on the genomic DNA was confirmed by PCR (Figure 3-5). Primers were designed in such a way that PCR products were obtained when correct integration of donor DNA at the intended locus occurred. For each transformation experiment, eight colonies were picked from the transformation plate and tested (note that only three are presented in Figure 3-5). In general, out of 8 colonies tested per donor DNA, correct

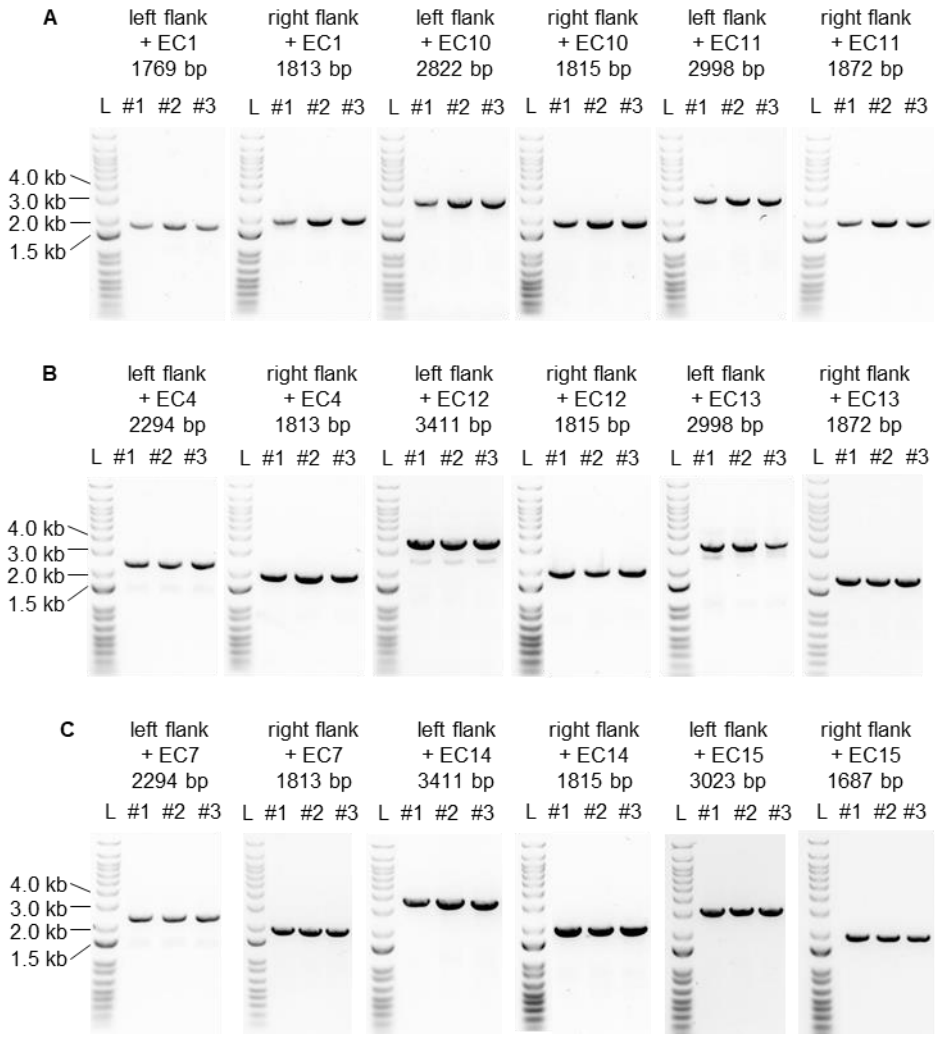


Figure 3-5. PCR verifying integration of the donor DNA expression cassettes at the intended loci within the genomic DNA. **A**. Verification of three colonies of the strain 1. **B**. Verification of three colonies of the strain 2. **C**. Verification of three colonies of the strain 3.

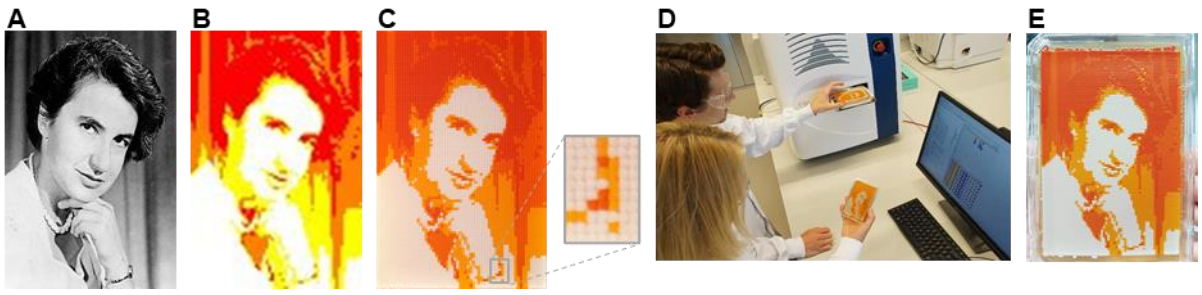


Figure 3-6. Yeast pixel art of Rosalind Franklin. **A**. Black and white RGB photo of 220 × 280 pixels of Rosalind Franklin that was used as a template. **B**. Computer conversion of the black and white photo of Rosalind Franklin into a 4-colour 64 × 96 pixel list. **C**. Photo of yeast pixel art with 64 × 96 yeast colonies with a zoomed-in section. **D.** Photo of an acoustic liquid handler with two full grown plates. **E.** Photo of a full grown microplate with 64 × 96 yeast colonies.

integration of the *crtE* donor DNA at the INT1 locus, *crtYB* at the INT2 locus and *crtI* at the INT3 locus was confirmed in >90% of the transformants. These results demonstrate the CRISPR-Cas12a system in combination with a single crRNA array enables efficient multiplex editing of the *S. cerevisiae* genome at multiple loci simultaneously. Additionally, we demonstrate the creation of "yeast pixel art" using the three carotenoid producing strains that were constructed together with a non-coloured wild-type strain. Starting from a black and white picture of Rosalind Franklin (Figure 3-6A), a 4-colour picture (Figure 3-6B) and spotting list was created which was then used to spot the four different yeast strains on an agar microplate using an acoustic liquid handler, resulting in a high-resolution "yeast painting" of Rosalind Franklin (Figure 3-6C-E).

3.4 Discussion

The provided protocol describes multiplex genome editing of *S. cerevisiae* using Cas12a from *Lachnospiraceae bacterium* ND2006 in combination with a single crRNA array and donor DNA. Design of the single crRNA array and donor DNA is explained in detail. In contrast to the well-established CRISPR-Cas9 system, the CRISPR-Cas12a has the unique additional ability of processing multiple crRNAs expressed from a single crRNA array[75,151]. Due to this feature, simultaneous editing of multiple targets is easier to set up and can be achieved in a single transformation. This single crRNA array approach was demonstrated before by Zetsche *et al.*[62] who simultaneously edited up to four genes in mammalian cells using AsCas12a, and by Świat *et al.*[63] for the introduction of four DNA fragments into a yeast genome using FnCas12a. To our knowledge, a higher number of simultaneous genomic modifications using a Cas12a system has not been reported and the maximal limit of targets per single array for Cas12a is yet to be determined. Further research utilizing single crRNA arrays in combination with Cas12a includes multiplex transcriptional regulation in a wide range of organisms[151–153].

Critical steps in the presented protocol:

- Carefully design all DNA sequences that are involved in the Cas12a genome editing experiment, especially in case when novel DNA sequences are introduced.
- Determine the functionality of new spacer sequences part of a crRNA, for example by a singleplex genome editing experiment as described by Verwaal *et al.*[65] before combining them into a single crRNA array.
- Follow the recommendations for the preparation of transformation buffer solutions used in the Cas12a editing experiment to achieve a good transformation efficiency of yeast.

Optional modifications of the technique:

- It is recommended to use 1 µg of each donor DNA, linearized pRN1120 or single crRNA array expression cassette in the transformation, although the use of a lower DNA amount is also expected to result in a

satisfactory transformation efficiency. Perform a test transformation to determine whether lower DNA amounts can be used.

- The transformation of *S. cerevisiae* might be performed using a different method than the one described in this protocol, for example the protocol described by Gietz *et al.* (2007) [154].
- The guide RNA recipient plasmid pRN1120 is suitable for the expression of a single crRNA and single crRNA array of different Cas12a variants (*e.g.* from *Acidaminococcus spp.* BV3L6 or *Francisella novicida* U112) as well as for expression of sgRNA in combination with Cas9[65].
- Donor DNA does not need to be limited to carotenoid gene expression cassettes and flanking regions that target donor DNA to the described INT1, INT2 and INT3 sites in genomic DNA. Any DNA of interest can be introduced, in a multiplex manner, into genomic DNA of the host by the design principles described in this protocol, or alternatively donor DNA can be used to delete DNA from a host genome.
- The modular structure of single crRNA array facilitates easy adjustment of spacer and direct repeat sequences. Modification of spacer sequences allows for a change of the intended integration locus which can be designed by one of the tools for identification of a genomic target site, *e.g.* GuideScan software 1.0[155].
- Instead of using large flanking sequences that contain connectors sequences, 50 bp of the flanking region can be included in the donor DNA sequences by incorporating these 50 bp flanking region sequences in the primers used in the PCR. In this case, in total just three instead of nine donor DNA fragments are required for a successful multiplex genome editing experiment.

In summary, this protocol provides step-by-step directions to perform multiplex genome editing in *S. cerevisiae* using Cas12a in combination with a single crRNA array approach. The protocol was demonstrated by multiplex genome editing using 9 donor DNA fragments and single crRNA array coding for three gRNAs. We show high overall editing frequencies between 50 and 94% for the three strain designs reported here. Concluding, the unique feature of Cas12a is the ability to process a single crRNA array into individual crRNAs in a cell, which makes Cas12a an excellent tool to enable multiplex genome editing and develop transcriptional regulation modules targeting multiple expression cassettes in one go. In the end, three strains were obtained producing carotenoids at a different level and colours in shades between yellow and orange. With those strains and a wild-type strain, we showed how an acoustic liquid handler can be used straightforwardly to make yeast pixel art – this in honour of Rosalind Franklin who contributed to the discovery of the DNA structure 65 years ago by her famous photo 51[142].

4 Construction of carotenogenic strains using Cas12a as a foundation for CRISPRi

This Chapter described characterization of carotenoids producing *S. cerevisiae* strains constructed by me based on the method enclosed in Chapter 3. Carotenoids quantification using UHPLC system was supported by Sandra Pous Torres (DSM), whereas operation of the acoustic liquid hander was supervised by Jeffrey van Wijk. All remaining experiments were conducted by me under supervised by René Verwaal (DSM), Thomas E. Gorochowski (University of Bristol) and Johannes A. Roubos (DSM).

4.1 Introduction

Production of β-carotene is a multistep conversion catalysed by three enzymes: geranylgeranyl diphosphate synthase (CrtE), bifunctional lycopene cyclase/phytoene synthase (CrtYB) and phytoene desaturase (CrtI). Proper balancing of the multi-enzyme pathway is crucial for providing sufficient turnover of intermediates, high productivity and preventing the accumulation of potentially toxic compounds in the cell. We sought to learn about expression levels of *crt* genes required for optimal β-carotene production and to elucidate a possible bottleneck in the pathway. With this knowledge, the feedback control system could be developed to dynamically balance expression of *crt* genes to maximize product formation while minimizing accumulation of toxic intermediated. Furthermore, determination of a single step which can completely block formation of β-carotene would simplify pathway control with CRISPR interference, allowing for a use of a single gRNA instead of an crRNA array in the initial tests of the system, simplifying experimental design. For this purpose, we created a library of carotenogenic *S. cerevisiae* strains with differentially expressed *crtE, crtYB* and *crtI* genes using promoters of various strengths. All promoters selected for expression of *crt* genes were heterologous so any of the promotes can be targeted with CRISPRi without disturbing cellular processes. The set of strains constructed in this Chapter allowed to test CRISPRi-dCas12a mediated downregulation of β-carotene biosynthesis (Chapter 5) and to evaluate native *S. cerevisiae* promoters as a burden monitor (Chapter 7), fundamental steps in the construction of the feedback control system.

4.2 Materials and methods

Cultivation conditions

During the construction of carotenogenic strains, rich YEPD (yeast extract peptone dextrose) medium with glucose (2% w/v, Chapter 2) was used, supplemented with nourseothricin or geneticin, when selection of *Cas12a* and gRNA plasmids (pCSN067 and pRN1120, respectively) was required. To characterize the production of carotenoids, constructed strains were grown in the mineral medium containing glucose (4% w/v, Chapter 2) in two steps. Pre-cultures were prepared in 20 mL of mineral medium and after 24 h diluted in 100 mL of fresh mineral medium to the OD_{600} of 0.1 and grown for additional 48 h at 30 °C and 250 rpm.

Strain construction

Carotenogenic strains (SI Table 4-1) were constructed by integrating expression cassettes encoding *crtE*, *crtYB* and *crtI* intro three independent loci using CRISPR-Cas12a. The detailed protocol can be found in Chapter 3. To control expression level of *crt* genes we selected nine heterologous promoters were classified as strong, medium and weak according to previous data (unpublished) (Table 4-1).

Table 4-1. Promoters used for expression of *crt* genes

Promoter strength	Promoter controlling *crtE*	Promoter controlling *crtYB*	Promoter controlling *crtI*
Strong (S)	*S. bayanus TDH3*	*S. mikatae TEF1*	*K. lactis ENO1*
Medium (M)	*K. lactis PGK1*	*K. lactis TEF2*	*K. lactis OLE1*
Weak (W)	*K. lactis TDH2*	*K. lactis YDR1*	*K. lactis LEU2*

Carotenoids quantification

Yeast culture corresponding to 1 mL of OD=20 was harvested after 24 and 48 h of fermentation and pellet was formed by centrifugation (14,000 rpm, 8 min). Cell pellets were resuspended in tetrahydrofuran (THF, Merck) and subjected to homogenization with glass beads and Precellys homogenizer (Bertin, Montigny-le-Bretonneux, France). Homogenization was performed in two cycles of 20 sec at 6,800 rpm with a 30 sec break. Next, the carotenoids fraction was collected (14,000 rpm, 8 min, 4 °C) and quantified with UHPLC Ultimate 3000 system (Thermo Fisher Scientific). Carotenoids were separated on Waters XBridge C18 column (3.5 µm, 2.1mm x 50 mm) with a gradient of mobile phases (SI Table 2-2) and detected using Ultimate 3000 photodiode array (Thermo Fisher Scientific). β-carotene, lycopene and phytoene were detected at 450 nm, 475 nm and 286 nm, respectively. Data collection and analysis was conducted with Chromeleon software.

Creation of yeast pixel art

The constructed strains were plated on rich agar using liquid acoustic handler according to Chapter 3.

4.3 Results

Elucidating control node in β-carotene biosynthesis

In this chapter we constructed a set of carotenoids producing strains with *crtE*, *crtYB* and *crtI* controlled by heterologous promoters of different strength (*i.e.* strong, medium or weak, Table 4-2) to determine the bottleneck in the production pathway. Expression cassettes encoding the production pathway genes were assembled in Golden Gate reaction and integrated into the genomic loci using CRISPR-Cas12a, as described in Chapter 3.

Carotenoids formation confers to the cells a colour trait which is directly associated with the production. To visualize the colour trait of constructed strains, we used an acoustic liquid handler robot (Figure 4-1), according to the protocol enclosed in Chapter 3. Expression of *crt* genes from a strong, moderate or weak promoter affected the colour of formed colonies, conferring pale yellow to dark orange shade. As expected, use of exclusively strong promoters conferred orange colonies (SSS), slightly brighter for medium strength promoters (MMM) and bright yellow for three weak promoters (WWW). Among the strains with one gene expressed from either a weak or medium promoter and the remaining two *crt* genes expressed from a strong promoters (xSS, SxS or SSx, where x stands for a promoter of any strength), the most striking change in colour was observed for strains SSW and SSM. Expression of *crtE* and *crtYB* from a weak promoter (WSS, SWS) resulted in intensively orange colonies. This observation suggests that in the tested settings, phytoene desaturation catalysed by CrtI is the limiting step in the carotenogenesis.

Figure 4-1. Carotenogenic *S. cerevisiae* strains constructed with CRISPR-Cas12a.

Table 4-2. Promoters' strengths selected for expression of *crt* genes.

Strain		Promoter strength of *crtE*	Promoter strength of *crtYB*	Promoter strength of *crtI*
CAR-022	SSM	Strong	Strong	Medium
CAR-023	SSW	Strong	Strong	Weak
CAR-024	SMS	Strong	Medium	Strong
CAR-025	SWS	Strong	Weak	Strong
CAR-026	SMM	Strong	Medium	Medium
CAR-027	SWW	Strong	Weak	Weak
CAR-028	MSS	Medium	Strong	Strong
CAR-029	WMM	Weak	Medium	Medium
CAR-030	WSS	Weak	Strong	Strong
CAR-031	MWW	Medium	Weak	Weak
CAR-033	-SS	no *crtE*	Strong	Strong
CAR-034	SSS	Strong	Strong	Strong
CAR-035	MMM	Medium	Medium	Medium
CAR-036	WWW	Weak	Weak	Weak

Next, we sought to confirm the visual trait with carotenoids quantification. The strains were cultivated in a mineral medium which results in better production of β-carotene and improved conversion of phytoene in comparison with a rich medium (Chapter 2). To prevent instability of the pathway during the fermentation, cultivation conditions were adjusted (SI Figure 4-1). Carotenoids quantification confirmed gradual decrease in the production of total carotenoids (*i.e.* the sum of phytoene, lycopene and β-carotene) when three strong promoters were substituted with medium strength and weak promoters (Figure 4-2). Among the strains with one of the *crt* genes expressed from either weak or medium strength promoter and the remaining two from strong promoters (SSx, SxS or SSx), the lowest level of carotenoids was detected for SSW and SSM. These results prove that the formation of carotenoids is affected by insufficient expression of *crtI*. With the reduced expression of *crtE*, strains MSS and WSS produced a similar levels of carotenoids to the high-producing strains SSS. Substitution of a strong promoter to control *crtYB* with medium or weak led to an increase in total carotenoids due to higher formation of β-carotene but also accumulation of lycopene and phytoene (Figure 4-3).

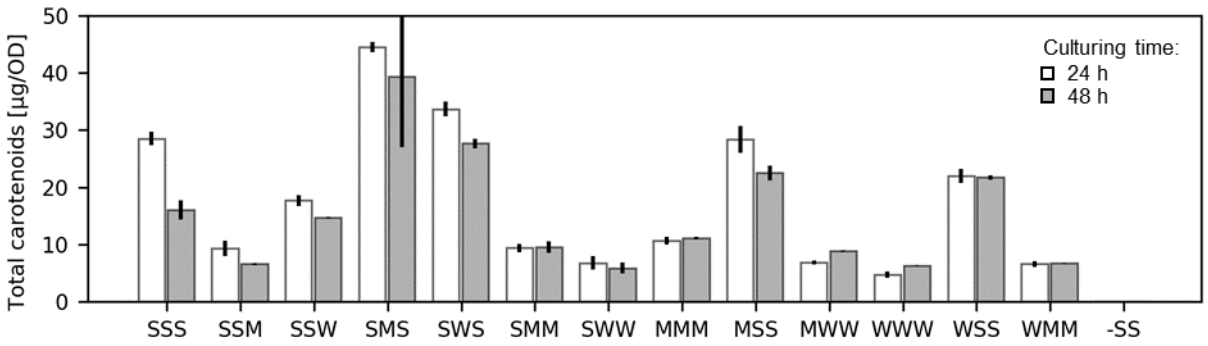


Figure 4-2. Production of total carotenoids by constructed strains. Total carotenoids production is the sum of phytoene, lycopene and β-carotene. Data represents mean ± standard deviation (n=2).

Balancing expression of *crt* genes

Selection of strong promoters for the expression of *crt* genes (strain SSS) resulted in insufficient conversion of phytoene and lycopene to β-carotene after 24 h of fermentation, which may indicate insufficient conversion by CrtI (Figure 4-3). This observation is in line with the previous reports in the literature[156,157] which indicated that production of high yields of β-carotene requires strong expression of *crtI* but weaker expression of *crtE* and *crtYB*. Strains MMM and WWW produced 6 and 16 times less of β-carotene comparing to strain SSS, respectively. Although, no lycopene was detected for these strains, phytoene levels exceeded β-carotene by 3.5 times. Phytoene is converted to lycopene in the step catalysed by CrtI, thus increasing pool of this enzyme, *e.g.* by additional gene copy might have a beneficial impact on the production. Very high accumulation of phytoene was observed when *crtI* was expressed from a weak promoter (strain SSW), posing 88% and 75% of total carotenoids produced after 24 h and 48 h, respectively. Although blocking an influx of the first intermediate in the pathway could be expected to affect the formation of the final product, we observed the contrary for strains WSS and MSS with a moderate expression of *crtE*. CrtE catalyses formation of geranylgeranyl diphosphate (GGPP) which is the first intermediate in the carotenoids synthesis. The level of β-carotene produced by strains WSS and MSS

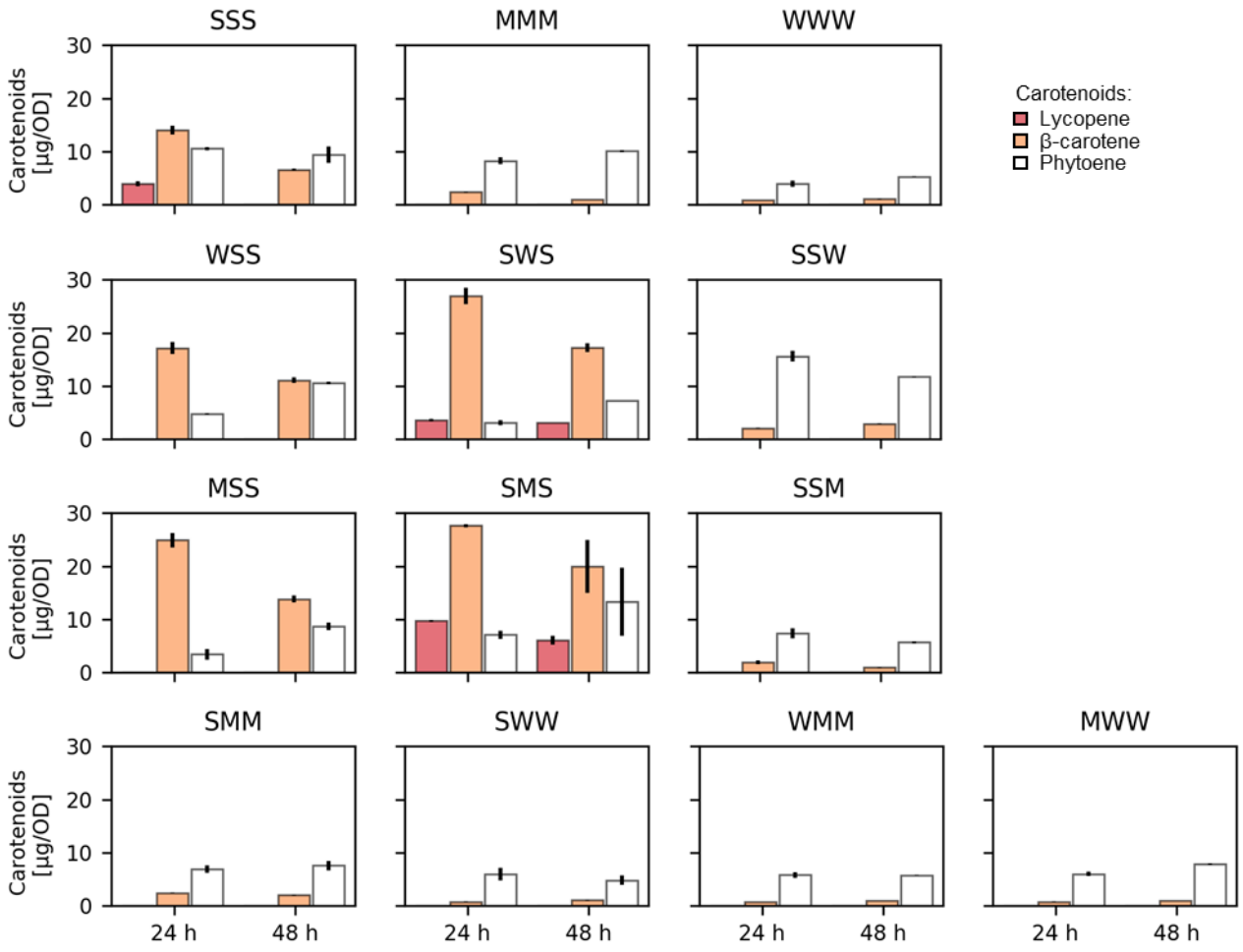


Figure 4-3. Carotenoids production profile of constructed strains. Strain names indicate promoter strengths used for expression of *crtE*, *crtYB* and *crtI* genes (S – strong, M – moderate and W – weak promoter). Data represents mean ± standard deviation (n=2).

exceeded the production by strain SSS and reduced accumulation of phytoene. Although *S. cerevisiae* contains a native geranylgeranyl diphosphate synthase encoded by BTS1, which potentially could balance low expression of *crtE*, we found that strain missing *crtE* gene (strain -SS) was unable to produce carotenoids. Selection of a weaker expression system for *crtYB* while maintaining high levels of *crtE* and *crtI* resulted in the highest production of β-carotene among tested strains. However, strains SMS and SWS exhibited high accumulation of lycopene, up to 22% of total carotenoids (SMS, 24 h).

Finally, we tested four strains (SMM, SWW, WMM, MWW) with *crtYB* and *crtI* expressed from either weak or moderate promoters and differentially expressed *crtE*. These strains produced from 0.8 to 2.5 μg of β-carotene which indicating that high productivity cannot be achieved with low expression of *crtYB* and *crtI*. A similar level of phytoene was noted for strains SMM, SWW, WMM, MWW, presumably due to insufficient conversion to lycopene. Substitution of a weak promoter with a strong one in front of *crtE* had a minimal impact on β-carotene yield when *crtYB* and *crtI* expression was insufficient.

4.4 Discussion

Orchestrating the correct gene expression in a multi-step pathway is crucial for maintaining optimal flux of intermediates and to prevent perturbation of native metabolism[158]. Reconstruction of a metabolic pathway using a promoter library is a common approach for a fine-tuning of the production[145,159,160]. Here we created a set of carotenogenic strains with differentially expressed *crt* genes responsible for β-carotene production. Instead of selecting the same strong, moderate or weak promoter to control all three *crt* genes, we selected three different promoters for each of the genes. Although this approach may lead to unequal expression of *crt* genes when different promoters of similar strength are selected, it allows us to reuse constructed strains in CRISPRi experiments and to target individually promoters of *crt* genes.

We determined that the expression level of *crtI* has a severe effect on the production pathway, consistently with the previous results[21,118,160]. To overcome the pathway limitation by *crtI* caused by insufficient expression or enzyme turn-over, production could be improved by selection of strong expression system, additional gene copy or selection of phytoene desaturase from a different organism[118]. Fine-tuning of the mevalonate pathway and ergosterol biosynthesis can further improve productivity by boosting precursor availability[40,161], however modification of the native *S. cerevisiae* metabolism lays out of the scope of this study.

We noted the highest β-carotene yields for strains with a high expression of *crtI* and lycopene accumulation in strains with strong promoters of both *crtE* and *crtI*. Lycopene is toxic to the yeast cells and its accumulation is known to affect cells fitness[33,118]. Therefore, an optimal design should provide high yields of β-carotene and efficient conversion of intermediates to prevent cellular accumulation of phytoene and lycopene. From the tested designs, the combination of a strongly expressed *crtYB* and *crtI* with moderate expression of *crtE* (MSS) provided the desired features.

In this chapter we elucidated the control node for β-carotene production which can be targeted with CRISPRi to effectively silence the pathway. The created strains can be used as a platform for testing

CRISPRi downregulation of carotenogenesis by targeting heterologous promoters and to assess sensitivity of native *S. cerevisiae* promoters to carotenoids production to create a burden biosensor.

5 Efficient multiplexed gene regulation in *Saccharomyces cerevisiae* using dCas12a

This Chapter was published at the Nucleic Acids Research[67] and it is presented here in unchanged form. Several people were involved in the process of retrieving the data. Nanopore sequencing was performed by me during a stay at the University of Bristol under supervision of Thomas E. Gorochowski (University of Bristol). Generated data was processed by Thomas E. Gorochowski. The tool used for verification of off-targets for gRNAs was developed by Jurgen Nijkamp (DSM). Experiments using flow cytometry were assisted by Bianca Gielesen (DSM) and script for recalculating fluorescence units for data obtained from the flow cytometry was prepared by Thomas E. Gorochowski. UHPLC system for carotenoids quantification was operated by Reza Maleki Seifar (DSM). Measurements using the fluorescence microscopy were supported by Sam Ruinard (DSM). Statistical analysis was guided by Abel Folch-Fortuny (DSM). Liesbeth Veenhoff provided valuable input on dCas12a localization. The remaining experiments were performed by me under the supervision of René Verwaal (DSM), Thomas E. Gorochowski and Johannes A. Roubos (DSM)

5.1 Introduction

Regulation of gene expression underpins numerous cellular processes, from the control of cell cycle progression to the dynamic adaptation of the proteome in response to environmental challenges[162]. The ability to precisely and dynamically manipulate gene expression is also a crucial part of metabolic engineering because the proper control of a cell factory's function requires the regulation of selected genes, often in a timely manner, and the balancing of other native metabolic pathways to accommodate heterologous production of a compound of interest[1,135]. This demand has resulted in the development of many molecular parts able to regulate gene expression[163–165].

Recently, the bacterial and archaeal adaptive immune system CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) has been adapted for activation (CRISPRa) and interference (CRISPRi) of gene expression[9,93,166]. The easily programmable nature of the CRISPRa/i system, where a gene is targeted through the simple modification of a guide RNA (gRNA) to include a short ~20 nucleotide (nt) complementary sequence, has resulted in CRISPRa/i becoming the predominant method for genome-wide studies of gene regulation[10,96,167]. To date, the majority of CRISPRa/i systems have made use of dCas9[9,10,93,96,167,168], which is sufficient to modulate transcription in prokaryotic cells without further modification. However, in eukaryotic cells performance can be enhanced by fusing the effector protein with a repression domain such as the Krüppel associated box (KRAB) domain of Kox1 involved in the formation of heterochromatin or the mammalian Mxi1 domain which interacts with the yeast histone deacetylase Sin3 homolog[93,169,170]. It has been shown that fusing dCas9 with Mxi1 can greatly enhance repression, enabling a 53-fold decrease in constitutive GFP expression compared to just an 18-fold decrease when dCas9 alone was used to target the same fluorescent reporter[93]. Analogously, CRISPR activation enhances gene expression through the fusion of transcriptional activator such as VPR or VP64 and the targeting of regions in close proximity to the promoter of a gene[168].

Beyond Cas9, CRISPRa/i systems have also been developed using class 2 type V Cas12a nucleases, formerly also known as Cpf1[93,94,98,151,168,171]. Cas12a nucleases are characterized by a different preference for the protospacer adjacent motif (PAM) sequence. Specifically, while the widely used SpCas9 recognizes NGG PAM sequences at the 3' end of a genomic target, LbCas12a favours TTTV PAM sequences at the 5' end

of a genomic target (where V is either A, C or G). Furthermore, unlike Cas9, Cas12a does not require trans-activating CRISPR RNA (tracRNA) and DNA targeting can be achieved with CRISPR RNA (crRNA) as short as 42 nt[74]. Ribonuclease activity of Cas12a also enables processing of multiple gRNAs encoded within a single crRNA array, enabling simultaneous modification of multiple genetic loci[62]. Recognition of T-rich PAMs by Cas12a is especially useful for targeting promoter regions, which are generally AT-rich[137]. A previous evaluation of dCas12a mediated downregulation of genes in *E. coli* showed no strand bias within the promoter region, thus PAM sequences on either DNA strand can be used, further increasing the number of potential targets[98]. In native yeast promoters, an enrichment of Ts within the core region has also been linked to the frequency of translation initiation events[172].

Saccharomyces cerevisiae is one of the most versatile and extensively used microorganism in industry and central to the production of pharmaceuticals, enzymes and food additives[14]. Although CRISPR-Cas12a has been applied for genome editing of *S. cerevisiae*, its application for gene regulation has been limited to a selection of organisms, excluding baker's yeast[63,65,81,83,95,98,173]. Furthermore, there has been a lack of systematic studies that characterise the many design choices when implementing dCas12a based CRISPRi systems in this host.

Here, we have addressed this issue by providing the first comprehensive assessment of dCas12a-based CRISPRi systems in *S. cerevisiae*. We vary and assess all key aspects of this system, including the position of nuclear localisation signals (NLS) fused to the dCas12a protein, the potential for enhanced repression through fusion of repression domains, the effectiveness of targeting different positions in promoters and ORFs, development of controllable gRNA expression regulation through the use of an RNA polymerase II (RNAP II) promoter and implementation of a dCas12a system for simultaneous downregulation of multiple genes. Finally, we demonstrate the potential application of this refined system for the stringent control of a heterologous β-carotene production pathway. This work provides valuable insight into the design constraints for effective CRISPRi when using dCas12a in *S. cerevisiae* and opens new avenues for stringent multiplexed control of gene expression and metabolic processes in this industrially important host.

5.2 Material and methods

Strains and cultivation conditions

Strains constructed in this study were generated using *S. cerevisiae* CEN.PK113-7D as a parent strain, except for strains FR013 and FR014 which were constructed in the CEN.PK113-9D background giving a possibility for using auxotrophic markers[128]. Cultures were grown in complex medium (YEPD) comprised of 2% Difco™ phytone peptone (Becton–Dickinson, Franklin Lakes, NJ, USA), 1% Bacto™ yeast extract (BD) and either D-glucose (2%, Sigma Aldrich, St Louis, MO, USA) or galactose (2%, Sigma Aldrich) when induction of the *GAL10* promoter was required. Selection was achieved using nourseothricin (NTC, 200 µg/mL, Jena Bioscience, Germany) or geneticin (G418, 200 µg/mL, Sigma Aldrich), when appropriate. Solid medium was prepared by addition of Difco™ granulated agar (BD) to the medium to a final concentration of 2% (w/v). Propagation of plasmids was performed using *Escherichia coli* NEB 10-beta cells (New England BioLabs, Ipswich, MA, USA). Bacterial cultures were prepared in 2*PY medium comprised of tryptone peptone (1.6%,

BD), Bacto™ yeast extract (1.0%, BD) and NaCl (0.5%, Sigma Aldrich) and containing ampicillin (100 µg/mL, Sigma Aldrich) or neomycin (50 µg/mL, Sigma Aldrich).

Strain construction

Yeast strains were constructed by genome editing of a single locus (INT4, SI Table 5-1) using CRISPR-Cas9 or multiple loci (INT1,2,3, SI Table 5-1) using CRISPR-LbCas12a in combination with a single crRNA array encoding three targets, as described previously[65,66]. Briefly, strain CSN001 pre-expressing Cas9 from plasmid pCSN061 or strain CSN004 pre-expressing LbCas12a from plasmid pCSN067 (SI Tables 5-2 and 5-3) were transformed with 1000 ng of the gRNA expression cassette and 100 ng of each DNA part and 100 ng of the gRNA recipient plasmid pRN1120 using the LiAc/ssDNA/PEG method[148]. Reagents required for yeast transformation were obtained from Sigma Aldrich (lithium acetate dihydrate (LiAc) and deoxyribonucleic acid sodium salt from salmon testes (ssDNA)) and Merck (polyethylene glycol 4000 (PEG), Darmstadt, Germany). The recipient plasmid used in combination with the sgRNA of Cas9 was amplified from pRN1120 using primers pKC003-004 (SI Table 5-4) whereas to generate the backbone for the Cas12a single crRNA array primers pKC007–008 were used. sgRNA for Cas9, crRNA and a single crRNA array for Cas12a (SI Tables 5-5 – 5-7) were obtained as synthetic DNA and amplified with different sets of primers (pKC001–002 and pKC005–006 for Cas9 and LbCas12a, respectively). Prior to the studies of gene expression regulation, plasmids used for the genome editing (pCSN060, pCSN067 and pRN1120) were removed by a sequential re-streaking and culturing of a strain and verification of the absence of resistance to markers used for selection of these plasmids.

A set of reporter strains with fluorescent protein genes integrated into genomic DNA (FR003, FR007-009, FR013; SI Table 5-2) was created to study transcriptional silencing with plasmid borne *dCas12a*. Fluorescent protein genes were expressed from heterologous promoters to enable targeting of dCas12a without affecting the expression of any native gene. To create these reporter strains, expression cassettes for *BFP*, *eGFP* and *mCherry* were assembled via Golden Gate cloning. DNA fragments coding for the genes of interest (open-reading frames/ORFs) of the fluorescent protein genes were obtained as synthetic DNA (BaseClear, Leiden, the Netherlands) and BsaI recognition sites were added by a PCR reaction using primers pKC009–014. Each expression cassette was comprised of a heterologous promoter, an ORF and a terminator, flanked by 50 bp connector sequence on the 5' and 3' end to facilitate assembly via *in vivo* recombination in *S. cerevisiae*. Additionally, two flanking regions of ~500 bp encoding the left and right region of an integration site were amplified from the genomic DNA of the host strain (primers pKC015–029,037). The same connector sequences as found in the expression cassettes were added to these flanking regions. Upon cleavage of double stranded DNA by a Cas protein and subsequent integration of the expression cassette, a sequence of ~1000 bp between the flanking regions was removed. Strain FR003 contained expression cassettes of *mCherry* and *eGFP* integrated into the INT4 locus. In strains FR007–009 *mCherry* controlled by the *FBA1* promoter from *S. cerevisiae* was integrated into INT3 whereas for *eGFP* three expression cassettes with different promoters were constructed and introduced into INT2. To prevent a lethal double strand break at INT1 when the single crRNA array was used in combination with Cas12a, a donor DNA encoding for a connector sequence and flanks of INT1 was supplied as the used array encodes three targets (INT1, INT2 and INT3). Strain FR013 was constructed to study multiplex silencing by

integrating *BFP*, *mCherry* and *eGFP* controlled by three heterologous promoters into three independent loci (INT1,2,3). This strain was further engineered using Cas9 to integrate the dCas12a E925A NLS Mxi1 expression cassette in INT4 resulting in strain FR014. Genomic DNA was isolated from constructed strains using Zymolyase (Zymo Research, Irvine, CA, USA)[66] and the integrated construct was amplified to generate a sequencing template (primers pKC030-037). Integrated constructs were verified by Sanger sequencing using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (ThermoFisher Scientific, USA) and NucleoSEQ columns for dye terminator removal (Macherey-Nagel, Düren, Germany).

Strain DC001 expressing functional Cas12a and the set of strains DC002–013 (SI Table 5-2) were created to compare the mutations D823A and E925A which confer nuclease deficiency to Cas12a and the functionality of five fused repression domains: Mxi1, KRAB, MIG1, TUP1 and UME6. To enhance stable expression, *dCas12a* was integrated into the genome at a single locus (INT4) together with *mCherry* and *eGFP*[66]. Constructs encoding *dCas12a* with a C-terminal NLS sequence and repression domains were obtained as synthetic DNA (Invitrogen, Carlsbad, CA, USA). A cloning free strategy was used to generate expression cassettes of integrated genes[174]. Promoters and terminators to be assembled in each transcriptional unit were first amplified to attach a 50 bp homology region to an integration site, connector, or a gene (primers pKC038–058). These homology regions facilitated *in vivo* recombination of nine parts to form expression cassettes of *dCas12a*, *mCherry* and *eGFP* separated by a 50 bp connector sequence. Constructs integrated in the edited locus of the created strains were sequence-verified using nanopore sequencing on a MinION device (Oxford Nanopore Technologies, UK). Genomic DNA was isolated as stated before and the sequencing template was amplified using a set of primers with unique 25 bp barcodes (pKC059–095). A combination of 5 reverse primers and 32 forward primers with unique barcodes allowed for 160 templates to be tagged which were subsequently pooled in equivalent amounts. Samples were prepped with a ligation kit (SQK-LSK 109, Oxford Nanopore Technologies, UK) and sequenced in a single run using a FLOW-MIN106 (9.4 SpotON) flow cell and the MinION MIN-101B sequencer.

The β-carotene producing strain CAR-034 was constructed by integrating the yeast codon optimized genes *crtE, crtYB* and *crtI* from *Xanthophyllomyces dendrorhous*[21] into three independent genomic loci using CRISPR-Cas12a, as reported in [66]. To test downregulation of β-carotene biosynthesis we integrated dCas12a E925A NLS Mxi1 into the INT4 locus of strain CAR-034, resulting in strains CAR-041 (*dCas12a* expressed from the *TEF1* promoter) and CAR-042 (*dCas12a* expressed from the *PGI1* promoter). Genome editing was conducted with CRISPR-Cas9, as described in the previous section. Strain CAR-041 was used in single- and multiplex CRISPRi experiments with gRNA targeting either each of the *crt* ORFs or a single crRNA array encoding three spacers targeting heterologous promoters in front of all the *crt* genes. Expression of the *crt* genes was controlled by Sbay_TDH3 p, Smik_TEF1p and Kl_ENO1p promoters (as used in strain FR013 to control expression of fluorescent protein genes) (SI Table 5-2). Strain CAR-042 was used in a spotting assay to assess the toxicity of dCas12a. Cultures grown overnight were diluted to OD_{600} of 1 and subsequently plated on YEPD agar plates in a series of 10-fold dilutions (*i.e.*, 10^{0}, 10^{-1}, 10^{-2}, 10^{-3} fold dilutions).

Plasmid assembly

Plasmids pC-NLS and pC-NLS-Mxi1 (Figure 5-1A, SI Table 5-3) used to express *dCas12a* were assembled by *in vivo* recombination in *S. cerevisiae* [66]. The backbone was amplified from plasmid pCSN061 with primers pKC096 and pKC097 such that the *Cas9* expression cassette was removed. Constructs of *dCas12a* E925A NLS without or with a Mxi1 domain (pC-NLS and pC-NLS-Mxi1, respectively) were obtained by PCR using genomic DNA from strains DC003 and DC010 as template with primers containing homology to the pCSN backbone (pKC098,099). To construct plasmids pC-Mxi1, pC-Mxi1-NLS, pC-NLS-Mxi1-NLS (Figure 5-1A, SI Table 5-3) synthetic DNA encoding NLS and the repression domain Mxi1 was ordered from IDT (Leuven, Belgium) and assembled in the pCSN backbone with *dCas12a* E925A via Gibson assembly [175]. The pCSN backbone with *dCas12a* E925A was amplified from plasmid pC-NLS with primers pKC100 and pKC101. To construct a functional dCas12a-eGFP fusion protein, a 60 bp linker was used[176]. *dCas12a-eGFP* was assembled in a Gibson reaction using a PCR fragment encoding *dCas12a* in a pCSN backbone amplified from plasmid pC-NLS with primers KC043,101 and fragment encoding *eGFP* amplified with primers pKC103-104.

Molecular biology techniques

DNA parts subjected to cloning and sequencing were amplified using Q5 polymerase (NEB) and primers were obtained from IDT (Leuven, Belgium). PCR products were purified with the Wizard SV gel and PCR clean up kit (Promega, Madison, WI, USA). Golden Gate cloning was conducted using BsaI-HF v2 (NEB) and T4 DNA Ligase (Invitrogen). Plasmid isolation from bacterial cultures was performed using QIAprep Miniprep (QIAGEN, Venlo, the Netherlands) whereas for yeast cultures Zymoprep Yeast Plasmid Miniprep II (Zymo Research) was used.

gRNA design

Genomic targets were designed using the Benchling website. Specificity of the designed crRNAs was tested against the CEN.PK113-7D genome using Burrows–Wheeler Aligner (BWA) and visualized with SAMtools [177,178]. Specific features of gRNA such as the length of the spacer, direct repeat and PAM sequence varied between nucleases and purpose (SI Table 5-5). crRNAs for gene downregulation with dCas12a were designed to target either the template or non-template DNA strand for the promoter region, but only target the template strand of a gene sequence (SI Table 5-6). The distance of a spacer from the transcription start site (TSS) was calculated according to Smith *et al.*[96] and using the YeasTSS database to elucidate the TSS of heterologous promoters used in this study[179]. Spacers selected for single crRNA arrays are listed in SI Tables 5-7, 5-8. gRNAs were expressed using the *SNR52* promoter (SNR52p) and *SUP4* terminator (SUP4t) from a multicopy plasmid pRN1120 (NatMX marker)[49,65] with the exception of the experiment where gRNA expression from a promoter processed by RNAP II was tested (SI Table 5-9). In this case, gRNA was expressed using the galactose-inducible *GAL10* promoter and the *GND2* terminator. A linear recipient plasmid fragment was obtained in a PCR reaction using primers pKC003,004 for Cas9 sgRNA and primers pKC007,008 for (d)Cas12a crRNA or single crRNA array. gRNAs were obtained from Twist Bioscience or IDT as expression cassettes flanked with 50 bp homology regions to the recipient plasmid pRN1120.

Micro-fermenter settings

Growth experiments were conducted in a BioLector (m2p-labs, Baesweiler, Germany). Cultures were prepared in YEPD medium with the appropriate antibiotic selection from individual colonies and grown overnight at 30°C, 250 rpm. The following day cultures were diluted in the medium to an OD_{600} of 0.05 and 1 mL was transferred to MTP-R48-B FlowerPlate (m2p-labs). For each strain three biological replicates were assayed. Cultures were grown in the BioLector under controlled conditions (30°C, 800 rpm, 80% rh, O_2 20.95%) with biomass (em. 620 nm/ex. 620 nm and gain 20) and fluorescence measured at 15 min intervals for a total duration of 72 h. eGFP was detected with a filter gain of 100 and wavelengths em. 488 nm / ex. 520 nm while RFP at a gain of 50 and wavelengths em. 589 nm / ex. 610 nm.

Flow cytometry

For analysis by flow cytometry, cultures were prepared from four colonies picked from transformation plates and inoculated in YEPD media supplemented with NTC (and G418 for dCas12a expressed from a plasmid) followed by incubation at 30°C, 550 rpm and 80% rh for two days to reach full saturation. Subsequently, cultures were diluted 20x in physiological salt and analysed with a BD FACSAria Fusion (Becton–Dickinson). Detection of events was set such that 20,000 events were measured for single cells and double cells were excluded from the analysis. The signal of fluorescent proteins was detected with a bandpass filter set at 530/30 nm for eGFP, 450/50 nm for BFP and 610/20 nm for mCherry. The data was recorded using BD FACSDiva 8.0.2 software to retrieve the geometric mean of the fluorescence distribution which was averaged for quadruplicates. Fluorescence obtained for eGFP, mCherry and BFP in arbitrary units was converted to molecules of equivalent fluorophores using Rainbow calibration beads with 8-peaks (BioLegend, London, UK) and the FlowCal Python package[180]. Specifically, the fluorescence of eGFP was expressed in Molecules of Equivalent FLuorescein (MEFL), mCherry in Molecules of Equivalent Phycoerythrin-TR (MEPTR) and BFP in Molecules of Equivalent BFP (MEBFP) using values of calibration beads detected with channels ECD, FITC and BFP for mCherry, eGFP and BFP, respectively.

Fluorescence microscopy

Nuclear DNA was stained with Vybrant® DyeCycle™ Violet (Invitrogen) and yeast cells were imagined with an Olympus BX53 microscope using CellSens software (Tokyo, Japan).

Carotenoid quantification

Isolation of carotenoids from cell pellets and subsequent quantification of phytoene, lycopene and β-carotene was conducted as described before [181]. Briefly, cells in 0.15–0.45 mL of carotenogenic culture were pelletized and resuspended in 1 mL of tetrahydrofuran (Merck). Carotenoids were extracted using Precellys® homogeniser (Bertin, Montigny-le-Bretonneux, France) in two 20 second cycles at 6,800 rpm with a 30 second pause in between. The resulting homogenised mixture was centrifuged at a speed of 13,000 rcf for 8 minutes at 4°C and the supernatant was collected and analysed using a UHPLC Ultimate 3000 (Thermo Fisher Scientific, Waltham, MA, USA), using a Waters XBridge® C18 column (3.5 µm, 2.1mm x 50 mm, Milford, MA, USA) for the separation of products. Three mobile phases of ethyl acetate, water and acetonitrile were used for separation of carotenoids (SI Table 2-2). Samples were stored at 12°C upon

injection into the system (10 µL). Carotenoids were detected with Ultimate 3000 photodiode array detector (Thermo Fisher Scientific) at wavelengths 286 nm, 475 nm and 450 nm to measure phytoene, lycopene and β-carotene, respectively. Data was recorded and analysed with Chromeleon software. Amounts of carotenoids were calculated based on standards: E/Z-Phytoene (Sigma Aldrich, #78903), lycopene (Sigma Aldrich, #75051) and β-carotene (U.S. Pharmacopeial, North Bethesda, MD, USA, #1065480).

Data analysis

Gene downregulation was expressed as fold repression by calculating the ratio between the fluorescence of a reporter gene upon expression of targeting gRNA (or a single crRNA array) and a non-targeting control gRNA (or single crRNA array). The standard error (SE) for these ratios was calculated as

$$SE = \frac{\sqrt{var\left(\frac{x}{y}\right)}}{\sqrt{n}},$$

where *x* indicates fluorescence of a reporter gene when targeted with gRNA, *y* corresponds to fluorescence of a reporter upon treatment with non-targeting gRNA and *n* is the number of tested samples. To check the statistical validity of a difference between obtained results, an unpaired t-test was used. Plots were generated using custom Python scripts and the Matplotlib package.

5.3 Results

Influence of NLS position on dCas12a-based CRISPRi efficiency

Fusion of a nuclear localization signal (NLS) to a Cas protein provides instructions for nuclear import and ultimately ensures the Cas protein is located within the same compartment as its target genomic DNA. In reported experiments using eukaryotes, Cas9 and Cas12a are most commonly fused to a single NLS at the C-terminus, which is often followed by a repression domain such as Mxi1 or KRAB to improve performance [48,49,93,94]. The NLS composition (*i.e.* monopartite vs bipartite), their number, and position at either N- or C-terminus was reported to affect the genome editing efficiency of AsCas12a in mammalian cells as well as the linker used to fuse Cas9 to an NLS[83,182].

For our dCas12a system we selected the SV40 NLS and began by assessing the effect of its position on the performance of CRISPRi. We constructed five plasmids (pC-Mxi1, pC-Mxi1-NLS, pC-NLS-Mxi1-NLS, pC-NLS, pC-NLS-Mxi1) encoding *dCas12a* E925A and a Mxi1 repression domain with a single NLS or multiple NLSs at various positions (Figure 5-1A). These plasmids were evaluated using two strains (FR003, FR009) which contain an *eGFP* expression cassette integrated at a different locus in the genome (INT4 and INT2, respectively for FR003 and FR009). Two gRNAs (gENO2 and gENO6) targeting two different positions of the promoter in front of *eGFP* were selected to mediate downregulation. Fluorescence of the eGFP reporter protein was measured by flow cytometry and fluorescence was converted to calibrated units through the use of an external standard (Material and Methods). Calibration of the fluorescence measurements improved the robustness of our results (ensuring technical errors were avoided) and has been demonstrated to enable better data reproducibility[183].

From these experiments (Figures 5-1B,C, SI Figure 5-4), the strongest repression of *eGFP* integrated at locus INT4 was seen when dCas12a was fused to a single C-terminal NLS and Mxi1 repression domain

(pC-NLS-Mxi1) and dCas12a-Mxi1 fusion missing an NLS (pC-Mxi1) (4.5 ± 0.3 and 4.1 ± 0.1 fold repression, respectively). In addition to constructs pC-NLS-Mxi1 and pC-Mxi1, which strongly downregulated *eGFP* at locus INT2 (7.1 ± 0.6 and 7.3 ± 1.6 fold repression, respectively), we observed an improved performance of pC-Mxi1-NLS (gENO2) harbouring a single NLS at the 3' end of the Mxi1 domain (7.8 ± 0.5 fold repression). In contrast, a reduction in repression was seen for designs containing two C-terminal NLSs (pC-NLS-Mxi1-NLS), regardless of the target locus. Addition of the Mxi1 repression domain to dCas12a with a C-terminal NLS (pC-NLS, pC-NLS-Mxi1) in combination with gENO2 enhanced *eGFP* downregulation in both strains, matching findings in other organisms[93,94]. Overall, fold repression of *eGFP* at locus INT2 was higher than for INT4, which is likely caused by the accessibility of the target locus. Based on these findings, we proceeded to use dCas12a-NLS-Mxi1 fusion for further experiments.

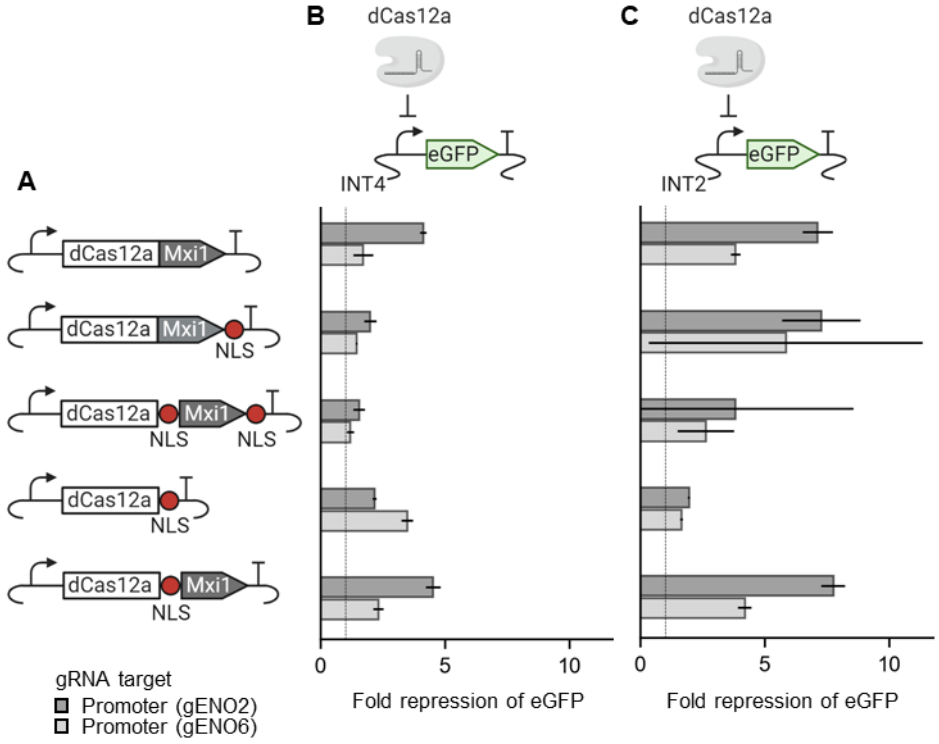


Figure 5-1. Effect of NLS number and position in dCas12a E925A Mxi1 fusions. **A**. Genetic diagram of plasmids used for expression of dCas12a E925A Mxi1 with NLS fusions (pC-Mxi1, pC-Mxi1-NLS, pC-NLS-Mxi1-NLS, pC-NLS, pC-NLS-Mxi1). **B**. Repression pattern in strain FR003 with *eGFP* integrated into INT4 and **C**. strain FR009 with *eGFP* integrated into INT2. CRISPRi was assessed using two gRNAs targeting Kl_ENO1 promoter controlling *eGFP* expression. Bars represent fold repression between targeting and non-targeting gRNA ± 1 standard error (n = 4). Dashed line indicates 1-fold change (*i.e.* no repression).

The ability for Cas9 lacking an NLS to edit yeast cells was reported previously[184], although editing efficiency was lower than when Cas9 was equipped with an NLS. To better understand our observation of gene downregulation using dCas12a without an NLS (pC-Mxi1), we constructed a fusion of dCas12a and eGFP to elucidate its cellular localization (SI Figure 5-1). Fluorescence microscopy revealed that the dCas12a-eGFP fusion construct was located in both the cytoplasm and nucleus explaining the downregulation in gene expression observed. Functionality of the dCas12a-eGFP fusion construct was

demonstrated in combination with the use of an effective gRNA to downregulate carotenoid production (SI Figure 5-1C).

Evaluation of repression domains

CRISPR Type V effector proteins such as Cas12a contain a single nuclease RuvC domain to sequentially cleave both strands of targeted DNA[137], thus endonuclease deficiency can be conferred by a single mutation. Evaluation of the dCas12a CRISPRi system in *E. coli* revealed that a single mutation in the nuclease domain of Cas12a, either D832A or E925A, outperformed the combination of these mutations[98]. Thus, we sought to compare dCas12a harbouring these single mutations by creating dCas12a mutants D832A and E925A. Gene downregulation can be further enhanced by equipping dCas12a with a repression domain, however, it has been shown that certain repressors can also abolish the activity of a dCas protein[95]. To our knowledge, in previous studies repression domains used in combination with dCas12a were limited to KRAB and Mxi1[95,135,185]. We therefore extended this set by including the additional well-characterized MIG1, TUP1 and UME6 domains that have been shown to be effective in combination with dCas9[94].

While *dCas12a* gene expression from a plasmid may provide variation in terms of copy number, integration of a single copy of the expression construct into the genome ensures stable expression during cell propagation. Expression of dCas12a solely or dCas12a-Mxi1 fusion from the same high strength promoter integrated into the genome resulted in at least three-fold repression of *eGFP* in comparison with plasmid borne expression of the same construct, which is possibly a consequence of a selection pressure to maintain only the gRNA plasmid (Table 5-1). For this reason, we constructed 12 strains with genome-integrated *dCas12*, with and without fusion to the five different repression domains (strains DC002-013). This set of strains was constructed using CRISPR-Cas9 mediated genome edition in combination with assembly of the P-O-T cassettes of *dCas12a*, *mCherry* and *eGFP* via *in vivo* recombination guided by 50 bp homology flanks annealed to the individual parts. Subsequently, integration of *dCas12a* and the fluorescent reporter was verified by nanopore sequencing (Material and Methods). The combination of these two approaches for strain construction reduced the total construction time by omitting cloning steps and sequence verification of individual samples. A set of unique barcodes was used to generate a sequencing template allowing the resulting PCR amplicons to be pooled and analysed in a single sequencing run (similar to the approach in [186]).

dCas12a fusions with repression domains displayed varied efficiency in downregulating *eGFP* expression (Figure 5-2, SI Figure 5-5). As observed for the previously tested plasmid pC_NLS, a reduction in eGFP fluorescence was noted for dCas12 without an additional repression domain. A similar repression level was observed for dCas12a D832A and dCas12a E925A targeting the promoter controlling *eGFP* expression (gENO2 and gENO4). The repression strength of dCas12a was further improved by a fusion to the MIG1 domain leading to almost a full blockage of *eGFP* repression (for dCas12a D832A 98% ± 4% and for dCas12a E925A 96% ± 1% decrease in eGFP fluorescence), in combination with the most efficient gRNA (gENO4). Fusion of the Mxi1 domain had a beneficial impact on the performance of dCas12a, particularly for gRNAs targeting *eGFP* ORF, which were ineffective for dCas12a alone. The strongest gRNA gENO4 and dCas12a D832A Mxi1 repressed eGFP fluorescence by 93% ± 27%, whereas for dCas12a D832A Mxi1 by 97% ± 5%.

Table 5-1. Comparison of dCas12a efficiency when expressed from plasmid (pC-NLS, pC-NLS-Mxi1) or integrated into the genome

	Fold repression		eGFP fluorescence (10^3 MEFL)		
	gENO2	gENO6	gENO2	gENO6	gNone
Integrated *dCas12a*-NLS	8 ± 0.7	14.9 ± 1.1	21.6 ± 3.1	11.5 ± 1.8	171.7 ± 17.3
Plasmid-based *dCas12a*-NLS	2.2 ± 0.1	3.5 ± 0.2	37.7 ± 2.4	23.5 ± 3.3	81.8 ± 2.6
Integrated *dCas12a*-NLS-Mxi1	12.6 ± 0.7	8.2 ± 0.5	13.2 ± 0.8	20.3 ± 1.3	166.5 ± 18.7
Plasmid-based *dCas12a*-NLS-Mxi1	4.5 ± 0.3	2.3 ± 0.2	14.1 ± 1.1	27.3 ± 4.5	63.6 ± 8.1

Fold repression of *eGFP* expressed as an average of values normalized by a non-targeting gRNA (gNone) for four biological replicates ± 1 standard error.

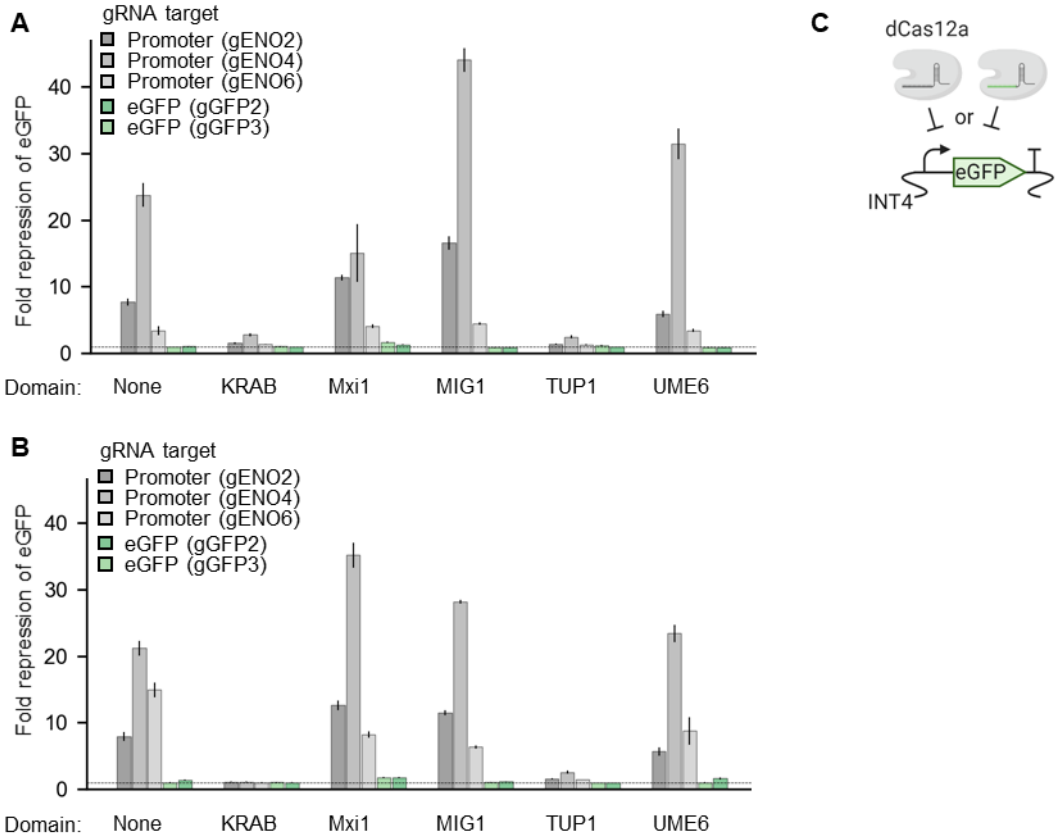


Figure 5-2. Comparison of repression domains and mutations defining nuclease-deficiency in dCas12a. **A**. Downregulation efficiency of dCas12a D832A and repression domains genome integrated in strains DC002, DC004-008. **B**. Downregulation efficiency of dCas12a E925A and repression domains genome integrated in strains DC003, DC009-013. Bars represent fold repression between targeting and non-targeting gRNA ± 1 standard error (n = 4). Dashed line indicates 1-fold change (*i.e.* no repression). **C**. dCas12a variants were assessed in downregulation of genome integrated *eGFP* (locus INT4) in strains DC002-013 using gRNAs targeting the Kl_ENO1 promoter or *eGFP* ORF.

The combination of KRAB and Cas12a E925a abolished any repression and significantly reduced the efficiency of Cas12a D832A, which is in line with previous findings[95]. Similarly, a general transcriptional repressor TUP1 diminished dCas12a mediated downregulation, although not to the extent observed for

Cas12a E925A KRAB (2.6- and 1.1-fold repression for TUP1 and KRAB in combination with gRNA gENO4, respectively). No clear benefit was noted for fusion of the UME6 domain as an improved repression level was observed only for one gRNA (gENO4) and for the remaining gRNAs resembled repression levels of dCas12a lacking a repression domain. The presented results show boundaries of transferability of the design rules for *S. cerevisiae* elucidated for dCas9 to dCas12a. Although fusion of a KRAB domain to dCas9 improved downregulation, such effect was not observed for dCas12a used in *Yarrowia lipolytica*[95]. In the study conducted by Lian *et al*.[94] the Mxi1 domain was outperformed by TUP1 in a fusion with dCas9, however, we observe an opposite behaviour for dCas12a. Our evaluation of repression domains using five gRNAs provides evidence for the advantage of using Mxi1 and MIG1 for dCas12a mediated CRISPRi.

eGFP downregulation targeting additional promoters

Understanding the requirements for effective gRNA design is complicated by a large number of different factors thought to affect efficiency. These include chromatin accessibility, target occupancy by the nucleosome and transcriptional factors, the target distance relative to the TSS and the DNA strand encoded by the spacer[95,96,98,187]. We sought to test three different heterologous promoter targets integrated into the same locus to elucidate common features of functional gRNAs. For this purpose, we designed constructs where *eGFP* expression was controlled by either the *TDH3* promoter from *S. bayanus*, the *TEF1* promoter from *S. mikatae* or the *ENO1* promoter from *K. lactis,* all integrated into the genome at the INT2 locus. A library of gRNAs was then designed for each strain with 4 to 6 gRNAs targeting the different heterologous promoters and 4 gRNAs targeting the *eGFP* ORF (Figure 5-3, SI Table 5-6, SI Figures 5-2, 5-6). The efficiency of these gRNAs for gene repression was then tested with a plasmid carrying dCas12a-NLS-Mxi1 (pC-NLS-Mxi1), which we previously found achieved considerable repression when evaluating the effect of NLS configuration.

We found that targeting the promoter driving expression of *eGFP* rather than the ORF resulted in higher repression for functional gRNAs in all of the strains. The strongest repression was seen for gRNAs targeting promoters and resulted in 4.2-fold repression for Sbay_TDH3p (gTDH1), 4.0-fold for Smik_TEF1p (gTEF4) and 20-fold for Kl_ENO1p (gENO3). However, while all gRNAs were designed in the same way (Materials and Methods), when targeting the promoter or ORF at least one non-functioning gRNA was found per target (Figure 5-3). Notably, no transformants were obtained for gRNA gTEF2 targeting the Smik_TEF1 promoter. Due to the similarity between the *TEF1* promoter from *S. mikatae* and *S. cerevisiae*, gTEF2 contained 18 complementary nucleotides to the promoters of both origins. The remaining two nucleotides mismatched the *S. cerevisiae TEF1* promoter within the seed region and were therefore not expected to exhibit lethal activity towards the native promoter[188,189]. The lack of transformants might be caused by a potential silencing effect of the native *TEF1* gene. The target position of gRNAs exhibiting the strongest repression varied between promoters. The most efficient gRNAs targeting the Kl_ENO1 promoter were located 106-319 nt upstream the TSS, whereas for Sbay_TDH3p and Smik_TEF1p optimal targets were positioned 297 nt and 31 nt upstream TSS, respectively. Our observation leads to the conclusion that targets located at certain promoter regions are more effective, however, this region varies between promoters. Finally, no strand bias was found for gRNAs targeting Kl_ENO1 promoter nor other features such as PAM preference.

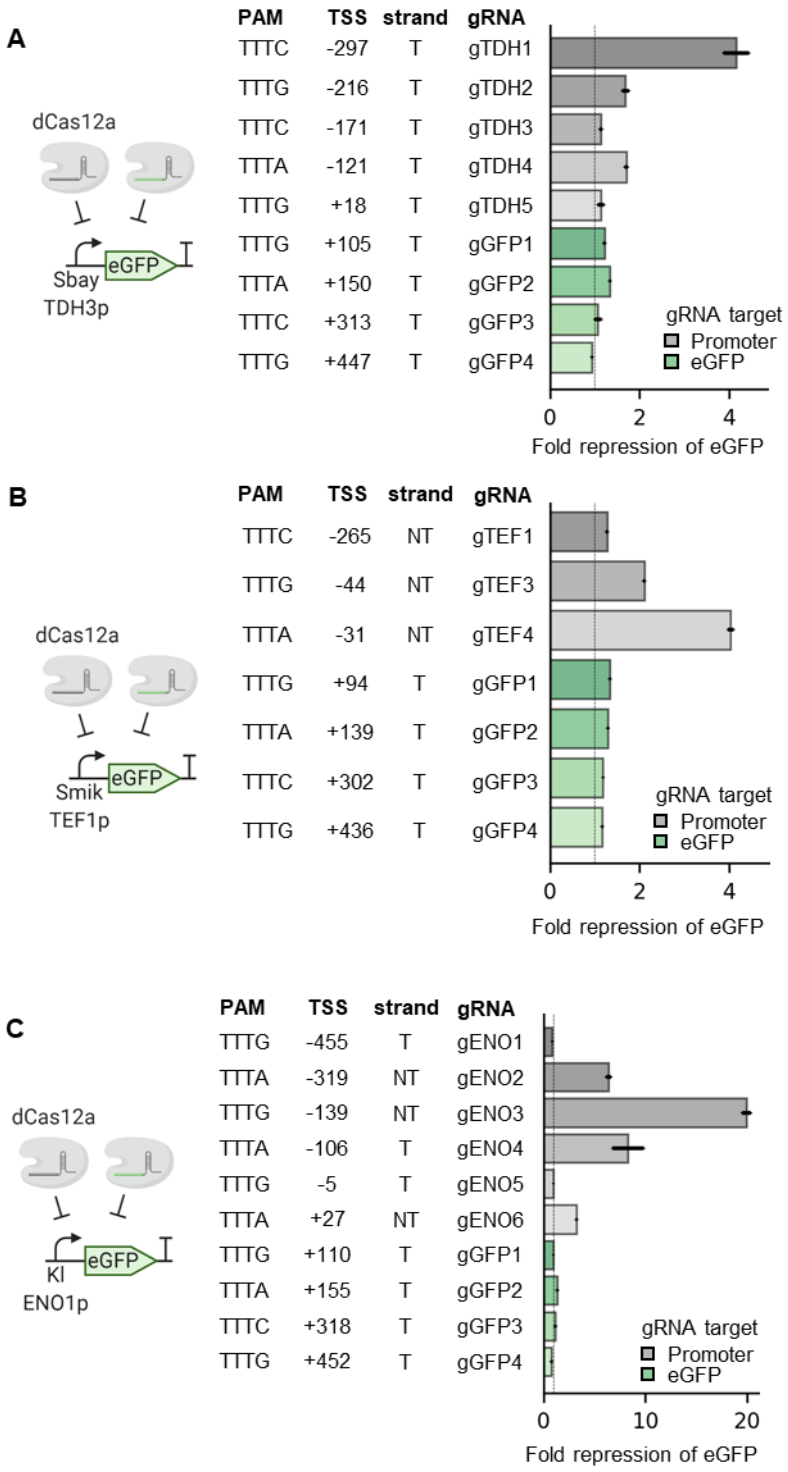


Figure 5-3. Downregulation of *eGFP* expression by targeting dCas12a E925A Mxi1 to promoters controlling *eGFP* or *eGFP* ORF. **A**. Repression pattern in strain FR007 with *eGFP* expressed from Sbay_TDH3p, **B**. strain FR0008 with *eGFP* expressed from Smik_TEF1p, **C**. strain FR009 with *eGFP* expressed from Kl_ENO1p. gRNAs targeting promoter in front of *eGFP* are depicted in grey and gRNAs targeting *eGFP* ORF in green. NT indicates non-template strand and T – template strand. Bars represent fold repression between targeting and non-targeting gRNA ± 1 standard error (n = 4). Dashed line indicates 1-fold change (*i.e.* no repression).

Controllable expression of gRNA from RNAP II promoter

In order to produce a functional gRNA the transcript must be precisely excised at specific sites, lack post-transcriptional modification of the 5' and 3' ends and should not be exported from the nucleus. These requirements are generally met by RNAP III promoters, exemplified by the snoRNA *SNR52* promoter[49]. An alternative approach is to combine the use of an RNAP II promoter with a gRNA flanked by self-cleaving ribozymes to cleave any signals for post-transcriptional modification or nuclear export[57]. The intrinsic RNase activity of Cas12a also allows for an additional design of RNAP II expressed gRNAs, whereby a spacer is flanked by two direct repeats. In this case, the two direct repeats are recognized by Cas12a and the pre-crRNA is processed into its mature form of a spacer and a single direct repeat[81]. Promoters recognized by RNAP II are generally well-characterized, exhibit a wide range of strengths and many are able to dynamically respond to changes in environmental conditions or the presence of an inducer molecule allowing for gRNA expression to be dynamically controlled[190]. A recent study on Cas12a-mediated genome editing showed that driving gRNA expression from a RNAP II promoter increased gRNA availability and improved editing efficiency[81].

As inducible gRNA expression for dCas12a in *S. cerevisiae* has not been demonstrated before, we selected a crRNA encoding a gRNA that targeted the *ENO1* promoter from *K. lactis* and compared repression when expressed from RNAP II and III promoters. For the RNAP II construct, the crRNA was expressed from an inducible *GAL10* promoter which is activated by the presence of galactose. We tested three designs of galactose inducible gRNAs containing one direct repeat of Cas12a (gGAL1), a spacer flanked by two direct repeats (gGAL2) or self-cleaving ribozymes with a single direct repeat (gGAL3) (Figure 5-4A, SI Figure 5-7). The RNAP III construct was expressed from the *SNR52* promoter and contained a single direct repeat either with or without ribozymes (gRNA gSNR1 and gSNR2, respectively). Strong downregulation of *eGFP* was achieved with GAL10p expressed gRNAs with either two direct repeats (gGAL2) or ribozymes (gGAL3) which under inducing conditions displayed 35.7 ± 3.7 and 41.2 ± 3 fold repression, respectively. gGAL1 lead to much lower repression (1.8 ± 0.2) in comparison with gGAL2 and gGAL3 which illustrates that a single direct repeat is insufficient for pre-crRNA processing by Cas12a. We also observed a decrease in eGFP fluorescence for cells transformed with crRNA controlled by *GAL10* promoter under non-inducive glucose conditions, which could be due to low levels of leaky expression (Figure 5-4A)[191].

Based on the comparison of repression levels achieved with GAL10p gRNAs with either two direct repeats or two ribozymes (gGAL2 versus gGAL3) and SNR52p gRNAs with one direct repeat or two ribozymes (gSNR1 versus gSNR2), there is no clear benefit to using self-cleaving ribozymes. The RNAP II based systems performed moderately better than the RNAP III-based systems (*eGFP* repression of 41-fold versus 26-fold for ribozyme flanked gRNAs gGAL3 and gSNR2 in galactose, respectively) (Figure 5-4A). This may be due to the known increased processivity of RNAP II over RNAP III[190]. Numerous studies have shown that the availability of the gRNA is the limiting factor for CRISPR-Cas9 mediated genome editing in eukaryotic cells[81,192,193]. Therefore, our results also suggest that despite using a multicopy plasmid in combination with a strong RNAP III promoter, the concentration of gRNA may still not be sufficient to fully saturate the available dCas12a pool.

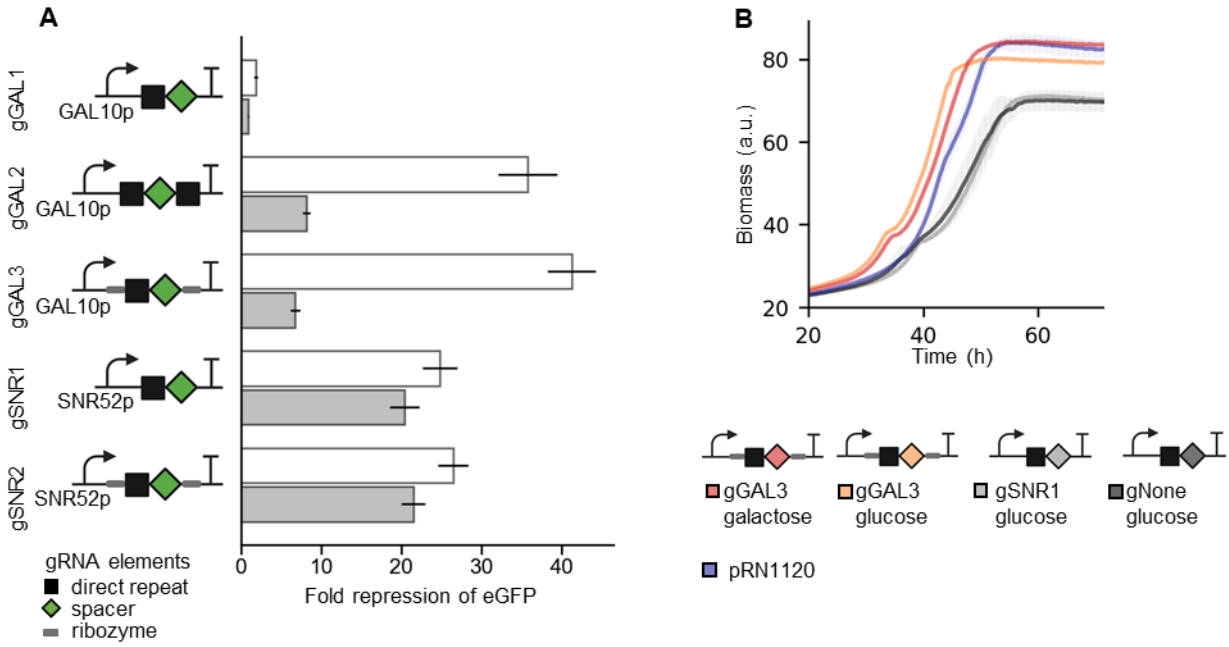


Figure 5-4. gRNA expression from RNAP III promoter and RNAP II promoter. **A**. eGFP fluorescence upon expression of dCas12a NLS Mxi1 and gRNAs targeting Kl_ENO1 promoter expressed from RNAP II promoter (GAL10p, gRNAs: gGAL1-3) and RNAP III promoter (SNR52p, gRNAs: gSNR1-2) under inducing (*i.e.* galactose, white bars) and repressing conditions (*i.e.* glucose, grey bars). Data normalized by a non-targeting control (gNone) tested under the same growth conditions as targeting gRNAs. Bars represent the mean from biological quadruplicates ± 1 standard deviation. Dashed line indicates 1-fold change (*i.e.* no repression). **B**. Biomass during a batch fermentation in microfermenter. pRN1120 indicates strain with empty gRNA recipient plasmid. Growth curves represent average of three biological replicates ± 1 standard deviation.

To better understand the dynamics of dCas12a-based CRISPRi and the influence of the expression system on cell growth, a micro-fermentation experiment was performed (Material and Methods). Expression of a CRISPRi regulated *eGFP* and a constitutively expressed *mCherry* was found to be closely linked to the cell growth with expected decreases in the eGFP fluorescence for strains harbouring a targeting gRNAs and stable of mCherry fluorescence upon reaching the stationary phase (SI Figure 5-3). Notably, antibiotic selection pressure used to maintain the empty pRN1120 plasmid caused a measurable decrease in growth rate (μ=0.056 ± 0.001 h^{-1}) when compared to the wild type strain (μ = 0.082 ± 0.006 h^{-1}, SI Figure 5-3A). In contrast, expression of the gRNA from the *GAL10* promoter did not affect growth (μ = 0.059 ± 0.003 h^{-1} under inducing and μ = 0.055 ± 0.001 h^{-1} under non-inducing conditions). Both targeting and non-targeting gRNAs expressed by the *SNR52* promoter negatively impacted cell growth (Figure 5-4B, μ = 0.044 ± 0.001 h^{-1} for gRNA gSNR1, μ = 0.043 ± 0.002 h^{-1} for gRNA gNone). We suspect that the expression of gRNAs from the RNAP III promoter leads to a competition for a shared pool of resources with native pathways leading to indirect impacts on normal cell growth[26].

Multiplexed downregulation

An advantage of dCas12a over dCas9 for CRISPRi is the automatic processing of CRISPR arrays by dCas12a, which removes the need for further accessory proteins (*e.g.* Csy4)[62,75,194]. This allows for the expression of multiple targets to be easily regulated simultaneously or enables increased repression through the expression of multiple different gRNAs targeting the same gene. In addition, Cas12a has evolved to process

long CRISPR arrays making this system robust to the production of multiple gRNAs, removing challenges faced when other approaches are used (*e.g.* expression of a gRNA array flanked by pre-tRNAs from an RNAP III promoter)[193].

To assess multiplexed regulation by dCas12a, we constructed a strain in which three fluorescence reporter proteins (BFP, mCherry and eGFP) were introduced into the genome at three integration sites and expressed by three different promoters (Sbay_TDH3, Smik_TEF1 and Kl_ENO1, respectively) (Figure 5-5A, SI Figure 5-8).

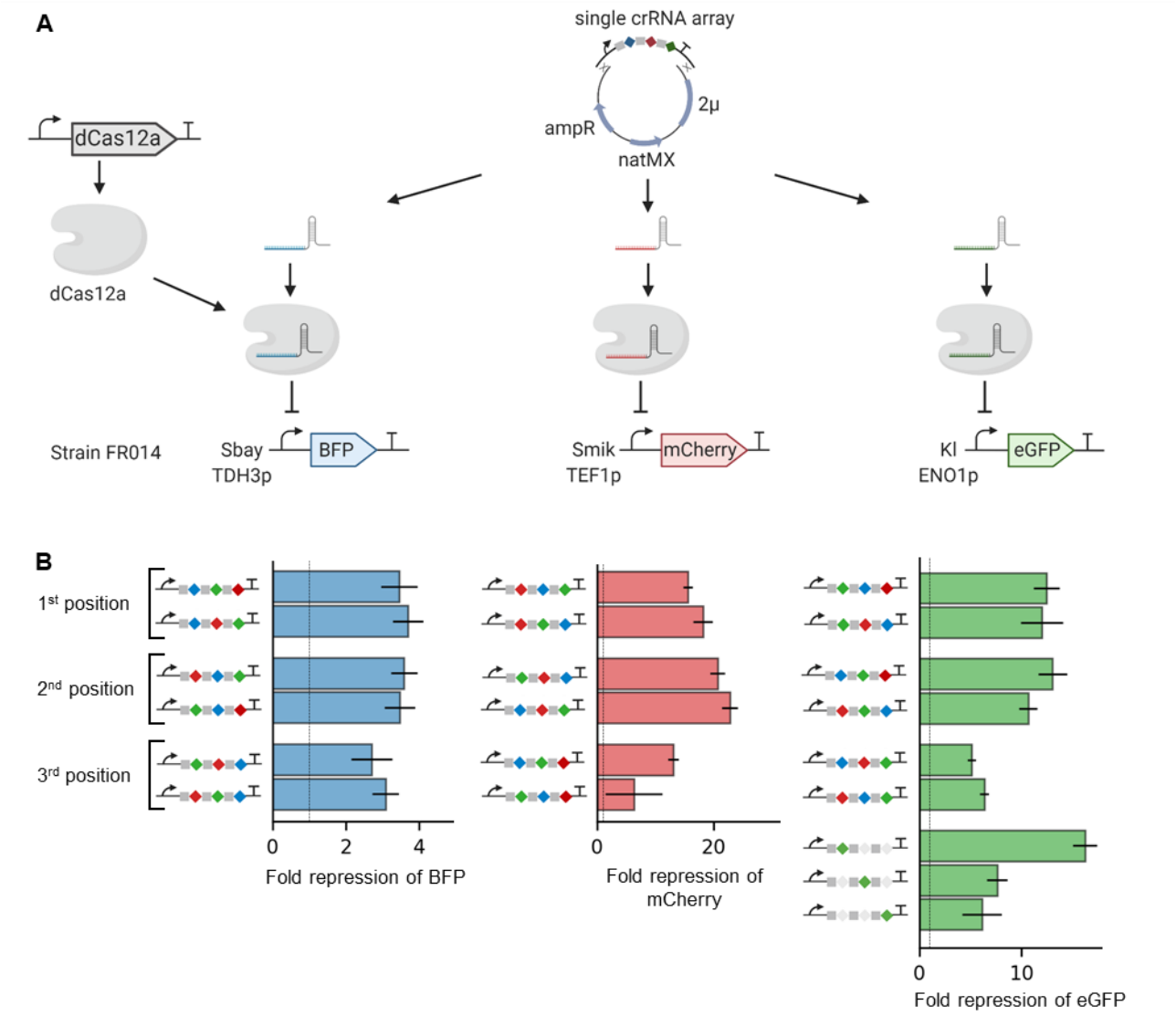


Figure 5-5. Simultaneous downregulation of fluorescent proteins with dCas12a E925A NLS Mxi1 and a single crRNA array. A. Principle of multiplex downregulation. Strain FR014 harbours three fluorescent proteins expressed from three different heterologous promoters. *dCas12a* was genome integrated to ensure stable expression. Subsequently, the resulting strain was transformed with recipient plasmid pRN1120 and crRNA expression cassette provided as linear fragment. Single crRNA array is assembled into the recipient plasmid pRN1120 via *in vivo* recombination. The crRNA array is expressed from *SNR52* promoter and *SUP4* terminator and subsequently processed by dCas12a into three individual crRNA array comprised of a spacer (blue, red, green) and direct repeat (grey). dCas12a forms a complex with crRNA and is directed to the genomic target encoded in the spacer sequence. **B**. Correlation between spacers order and repression of *eGFP*, *BFP* and *mCherry*. Bars represent fold change in BFP, mCherry and eGFP fluorescence normalized by a non-targeting array (array_10, not depicted) upon expression of dCas12a E925A NLS Mxi1 and single crRNA arrays. Six permutations in spacers order were tested and three control arrays with one spacer

targeting *eGFP* promoter and two non-targeting spacers. Bars represent fold repression between targeting and non-targeting gRNA ± 1 standard error (n = 4). Dash line indicates 1-fold change (*i.e.* no repression).

A single CRISPR array encoding three gRNAs targeting promoters driving expression of each reporter gene was then used to repress their expression simultaneously. The gRNA target sequences within each promoter were selected based on the capability to repress *eGFP* with similar strength (Figure 5-3): gRNA gTDH1 targeting Sbay_TDH3p (4.2-fold repression), gRNA gTEF4 targeting Smik_TEF1p (4.0-fold repression) and gRNA gENO6 targeting Kl_ENO1p (3.2-fold repression). To systematically assess the effect of gRNA order within an array, all possible permutations were tested with repression calculated by comparison to an array containing non-targeting gRNAs (SI Table 5-8). A strong decrease in fluorescence of all reporters upon expression of the targeting arrays was observed (Figure 5-5B). Notably, gRNAs encoded in the first or the second position within the array caused stronger repression than the same gRNA encoded at the third position (unpaired t-test, *P*-value = 0.0001 for *eGFP* and *mCherry* and *P*-value = 0.01 for *BFP*, *n* = 8, Figure 5-5B). No significant difference was observed in repression efficiency for gRNAs targeting *eGFP* and *BFP* at position 1 and 2. Dependence of the gRNA context was recently reported for CRISPRi mediated by dFnCas12a in *E. coli* with improved repression for gRNAs encoded in position 2 compared to the same gRNA encoded in position 1 or 3 within an array[195].

Stringent regulation of heterologous β-carotene production

β-carotene is a strongly coloured red-orange pigment abundant in plants, microalgae, fungi and bacteria and exhibits antioxidant properties[161,196]. While there is increasing demand for β-carotene for food additives and nutraceuticals, supply is hampered by the low-productivity and harsh conditions needed for extraction from natural sources. Microbial production of β-carotene could alleviate this difficulty and has gained significant interest in recent years[21,116,197]. To enable β-carotene production in industrially relevant microorganisms such as *S. cerevisiae*, the β-carotene biosynthesis pathway from *Xanthophyllomyces dendrorhous* can be heterologously expressed (Figure 5-6A)[21]. This pathway comprises three genes: *crtE*, *crtYB* and *crtI*, encoding geranylgeranyl diphosphate synthase, bifunctional lycopene cyclase/phytoene synthase and phytoene desaturase, respectively. Only the simultaneous expression of all three *crt* genes results in the formation of β-carotene in *S. cerevisiae* and a clear phenotypical change where the cells turn from a white into an orange colour[21].

We sought to use our dCas12a CRISPRi system to stringently regulate β-carotene production in *S. cerevisiae* as a foundation for the dynamic control of this metabolic pathway. We made use of a previously constructed strain containing the heterologous genes required for β-carotene production (CAR-034)[66]. This strain was further modified to have dCas12a E925A fused to Mxi1 and a single NLS integrated into its genome (strain CAR-041) and was transformed with a plasmid selected by dominant markers that expressed a single gRNA or crRNA array for regulation of expression of the *crt* genes. Crucially, the expression of the three genes (*crtE*, *crtYB* and *crtI*) was driven by the same heterologous promoters used for our assessment of multiplexed repression (Sbay_TDH3p, Smik_TEF1p and Kl_ENO1p, respectively), allowing the reuse of our previously tested gRNAs and crRNA arrays (Figures 5-3 and 5-5).

To begin, we used single gRNAs targeting either *crtE* (gTDH1–5), *crtYB* (gTEF3–4) or *crtI* (gENO1–6) in isolation to elucidate possible single points of control within the β-carotene pathway (Figures 5-6B,C). The

most prominent impact was seen for repression of the *crtI* gene, which lead to no detectable β-carotene production. The corresponding gRNAs (gENO2,3,4,6) have previously been shown to repress *eGFP* by 70–95% (Figure 5-3C), suggesting strong expression of *crtI* is necessary for β-carotene production. The remaining gRNAs targeting *crtI*, gENO1,5 reduced β-carotene production by 2.8- and 5-fold, although repression of *eGFP* was not observed for these gRNAs (Figure 5-3C). For repression of *crtE* and *crtYB*, β-carotene was reduced but still detectable. As expected, repression of *crtE* which catalyses the first step in carotenogenesis, was not effective likely due to the presence of a native *S. cerevisiae* GGPP synthetase encoded by *BTS1*. Production of β-carotene despite repression of *crtYB* indicates that conversion of phytoene to lycopene by CrtI limits the final lycopene cyclization to β-carotene by CrtYB. The concentration of the intermediate metabolite phytoene increased upon expression of gRNAs targeting *crtI*, which likely relates to the accumulation of this compound due to limited conversion of the reaction catalysed by CrtI into the subsequent compound lycopene. Furthermore, repression of *crtYB* decreased the phytoene production, however, only for gRNA gTEF4. Interestingly, the best performing gRNA for downregulation of the Sbay_TDH3 promoter when expressing *eGFP* (gTDH1) caused only a small reduction in phytoene concentration when used in the β-carotene pathway, whereas gRNA gTDH5 which repressed eGFP production only by 18%, halved phytoene (Figures 5-3C and 5-6B). It should be noted that *eGFP* and *crtE* were positioned on different chromosomes (11 and 10, respectively), which leads to the hypothesis that the functionality of these gRNAs might dependent on the target locus and/or genetic context. We also observed an over 8-fold drop in phytoene concentration when gRNA gTEF4 was used to repress *crtYB*, making it the limiting step in phytoene synthesis over CrtE. Although we performed lycopene quantification using the UHPLC-DAD system, none was detected for any of the tested strains. Small amounts of lycopene could have accumulated in the tested strains, but this may not have been detected due to the sensitivity threshold of the quantification method. Strains with reduced production of β-carotene exhibited cell pellet with a faded colour, in comparison with the gNone control which had unaffected carotenoids production (Figure 5-6C). The colour intensity was not fully restored to the white appearance of the wild type pellet even when no β-carotene was detected. This can be explained by the presence of other colourful carotenoids and the detection limit of the UHPLC-DAD system utilized for quantification of β-carotene.

Finally, to see if a full shutdown of carotenoids production (*i.e.* all metabolites in the β-carotene pathway) could be achieved, we exploited the multiplexed downregulation of dCas12a capability by using crRNA arrays targeting a single or all promoters expressing the *crt* genes (Figure 5-6D, SI Table 5-8). As expected, targeting solely *crtI* (arrays 7 and 9) disabled production of β-carotene but not phytoene (Figures 5-6C,D, SI Table 5-8). However, expression of the *crtI* targeting gRNA from the second position (array_8) led to growth defects and only moderately reduced β-carotene production, suggesting expression of the gRNA was hampered and had broader off-target effects. In contrast, expression of arrays_1-6 simultaneously targeting *crtE*, *crtYB* and *crtI* were found to block the production of β-carotene and decreased phytoene by 5-fold (with the exception of one array), demonstrating the ability for multiplexed CRISPRi to stringently control the entire pathway.

The production of β-carotene is known to decrease cellular fitness[33,117,118,121,125]. To evaluate the potential additional burden imposed by the CRISPRi system on the carotenogenic strain, we performed a spotting assay (Figure 5-6E). Firstly, we compared the growth of the wild type strain (CEN.PK113-7D) and

strains with genomically-integrated *Cas12a* or *dCas12a* harboring a mutation in the nuclease domain (either D832A or E925A) expressed from the strong constitutive promoter *TEF1* (strains DC001-003, respectively).

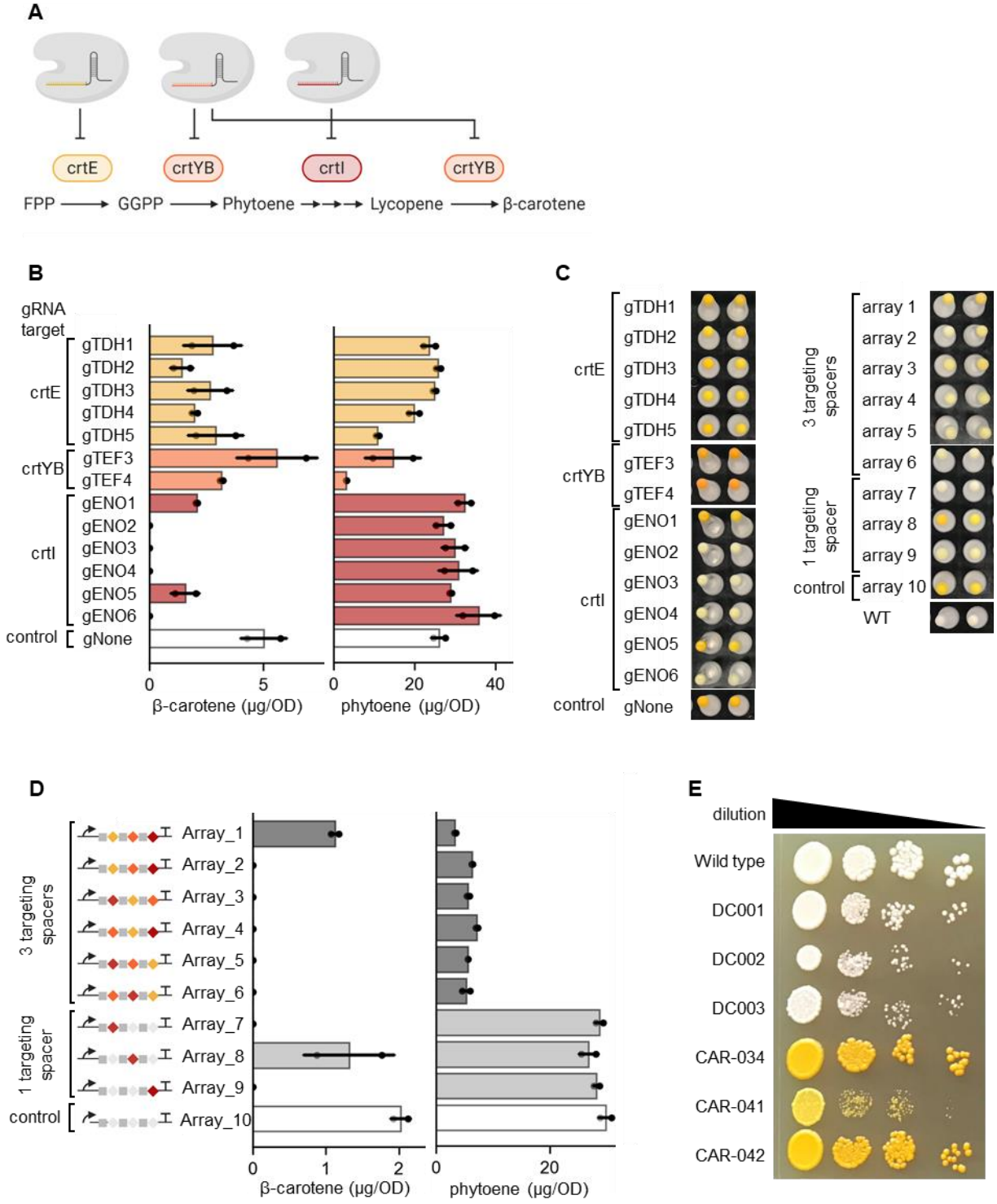


Figure 5-6. Simultaneous downregulation of β-carotene production with dCas12a E925A NLS Mxi1 and single crRNA array. A. Carotenoids production in *S. cerevisiae* is a multistep pathway catalysed by three enzymes: CrtE, CrtYB and CrtI derived from *Xanthophyllomyces dendrorhous*. The pathway can be downregulated using dCas12a in sinlgeplex manner with a gRNA or multiplex using single crRNA arrays targeting promoters in front of *crt* genes resulting in decreased enzyme levels of CrtE, CrtYB and CrtI and ultimately decreased amounts of carotenoids. Abbreviations: FPP, farnesyl pyrophosphate; GGPP, geranylgeranyl pyrophosphate. **B**. β-carotene and phytoene levels in strain CAR-041 upon singleplex downregulation with dCas12a and gRNAs harbouring a single spacer complementary to promoters in front *crt* genes. **C**. Cell pellet obtained after growing transformants of strain CAR-041 with (left panel) gRNAs used in

singleplex downregulation for targeting *crtI*, *crtE* and *crtYB*. For comparison transformants with non-targeting gRNA (gNone) and wild type strain; (right panel) single crRNA arrays targeting simultaneously *crtE*, *crtYB* and *crtI*, solely *crtI* and non-targeting control. Columns correspond to biological replicates. **D**. β-carotene and phytoene levels in strain CAR-041 upon multiplex downregulation with dCas12a and single crRNA array. Arrays with three targeting spacers (array_1-6) targeting promoters Sbay_TDH3, Smik_TEF1 and Kl_ENO1 which control expression of *crtE*, *crtYB* and *crtI*, respectively are shown in dark grey, whereas arrays with a single targeting spacer (array_7-9) targeting Kl_ENO1 promoter controlling *crtI* and two non-targeting spacer in light grey. Non-targeting control (array_10) is depicted in white. Bars represent mean ± 1 standard deviation (n=2). E. Effect of dCas12a expression in wild type and carotenogenic strains. DC001 *Cas12a*, DC002 *dCas12a* D832A, DC003 *dCas12a* E925A, CAR-034 carotenogenic strain, CAR-041 carotenogenic strain with genomically integrated *dCas12a* E925A Mxi1 controlled by the *TEF1* promoter, CAR-042 carotenogenic strain with genome integrated *dCas12a* E925A Mxi1 controlled by the *PGI1* promoter.

Strains expressing (d)Cas12a formed smaller colonies compared to the wild type strain. Next, we tested the effect of carotenoids production in combination with weak and strong expression of dCas12a E925A to understand whether the toxicity of CRISPRi system is caused by a difference in protein expression[31]. A carotenoid-producing strain with a strongly expressed dCas12a protein (CAR-041) exhibited smaller colonies sizes in comparison with a strain expressing solely dCas12a (DC002,003). However, cellular growth was restored in a carotenoids producing strain (CAR-042) when the strong *TEF1* promoter was substituted with the weaker PGI1p. These results demonstrate mild toxicity of CRISPRi dCas12a when highly expressed, which can be mitigated by using of a weaker promoter, in line with previous reports for FnCas12a[63].

5.4 Discussion

The versatility and relevance of *S. cerevisiae* in industry motivates the significant demand for new molecular tools for gene regulation in this important host organism. Although dCas12a is a suitable system for expression control of metabolic pathways due to its T-rich PAM preference and a capability to process array for regulation of multiple targets, application of CRISPRi-dCas12a has never been reported for *S. cerevisiae*. Furthermore, studies on dCas12a in other organisms typically employ design rules previously established for dCas9, which may not be optimal for dCas12a. In this work, we have addressed these issues and demonstrated the use of dCas12a for multiplexed CRISPRi in *S. cerevisiae*. Our systematic study revealed some core considerations for effective regulation when using this system. First, we found efficient downregulation was present for dCas12a with either a C-terminal NLS and Mxi1 or when fused with a C-terminal Mxi1 domain. Using fluorescence microscopy, it was observed that dCas12a lacking an NLS can localize to the nucleus, which demonstrated that dCas12a can even be effective without NLS fusion. Although our results cannot explain the mechanism of the nuclear import and accumulation of dCas12a, a possible explanation for dCas12a localization in the nucleus is a cryptic location signal in dCas12a, as speculated before for Cas9 when genomic edits were detected for the nuclease fused to a nuclear export signal (NES)[184]. Lysine is a particularly abundant amino acid in dCas12a (posing 12.5%), giving a possibility for a region with positively charged residuals to serve as a cryptic NLS undetected explicitly with available prediction tools[198–200]. An alternative mechanism of dCas12a import to nucleus is passive diffusion[201,202]. Considering recent publications[203,204] reporting slow dissociation rate of dCas9 once it is bound to a target, a low amount of dCas protein in the nucleus might be expected to mediate downregulation. To fully understand the mechanism of dCas12a import to the nucleus further research is necessary. Strong

repression was observed when using dCas12a E925A with a single C-terminal NLS and Mxi1 or MIG1 repression domain, leading to a 97% reduction in an eGFP reporter protein. Fusion to KRAB and TUP1 repression domains diminished the repressing capability of dCas12a, which has not been observed previously when used with dCas9[135]. Second, we found that dCas12a-based CRISPRi is more effective when the promoter region is targeted rather than targeting the ORF. This feature is potentially useful to downregulate many different genes simultaneously using the same gRNA if its promoter target is used for expression. Third, analysis of crRNA arrays encoding three gRNAs for multiplexed downregulation revealed a position-dependent effect on the efficiency of each gRNA. Those located in the first two positions caused significantly better repression than those in the third position. We also demonstrated gRNA expression from an inducible RNAP II-controlled *GAL10* promoter to enable inducible downregulation, opening avenues for dynamic control of gene expression using this system. Finally, as a proof-of-concept application, we showed that the inherent ability of dCas12a-based CRISPRi to preform multiplexed gene regulation can be used to stringently control a heterologous metabolic pathway producing β-carotene.

Our comparison of dCas12a-CRISPRi designs was evaluated for downregulation of eGFP and β-carotene biosynthesis, however, the presented results can serve multiple purposes. We observed a consistent performance of a dCas12a fusion with a Mxi1 domain and a C-terminal NLS regardless of a target and its chromosomal location, if the gRNA is functional. Therefore this construct is expected to be effective in regulating any metabolic pathway if working gRNAs are chosen. The modularity of the dCas12a system (NLS, repression domain, gRNA) allows to vary different elements. Among five repression domains we observed enhanced silencing activity of dCas12a for two of the domains and diminished activity of two others. Evaluation of a broader set of repression domains may lead to the discovery of even stronger repression systems. Although our work is limited to dCas12a from *Lachnospiraceae bacterium* ND2006 (LbCas12a), CRISPRi has been demonstrated for other orthologues: FnCas12a[81,95,98,173], AsCas12a[153] and EeCas12a[205]. With the recent work of Toth *et al.* this set could be further extended by the implementation of novel Cas12a variants exhibiting a preference for alternative PAM sequences such as TNTN[206] developed from LbCas12a without affecting the nuclease domain.

As the complexity of synthetic genetic systems grows, there will be an increased need for high-performance molecular tools able to regulate gene expression in a flexible way. The CRISPRi system we developed in this study exploits the inherent benefits of dCas12a for gene regulation, such as its preference for a T-rich PAM and inherent capability to process crRNA arrays to enable multiplexed regulation with minimal effort for the design. It also demonstrates the value of comprehensive studies of design parameters for such systems, which open up avenues to both refine existing systems and ensure the best possible performance is achieved; aspects that will be essential for transitioning them into industrial use.

6 Exploration of Cas13 ribonucleases for a transcript knockdown

This Chapter describes implementation of CRISPR-Cas13 system in *S. cerevisiae*. The initial work on this system was performed by Naoual Ouazzani-Chahdi (DSM, University of Wageningen), and some of the results are included in this Chapter, mainly construction of strains CC001-004 and sequencing expression cassettes of colonies with active Cas13. Nanopore sequencing data was generated during my stay at the University of Bristol hosted by Thomas E. Gorochowski (University of Bristol). Sequencing data was processed by Thomas E. Gorochowski. A potential off-target effect of selected gRNAs was evaluated using the tool developed by Jurgen Nijkamp (DSM). Thomas E. Gorochowski provided guidance for the analysis of qPCR data and change of unit for fluorescence data obtained from the flow cytometer. The live-dead staining was supported by Bianca Gielesen (DSM). Zoya Ignatova (University of Hamburg) provided valuable input for analysis of RNA integrity. The work was supervised by René Verwaal (DSM), Thomas E. Gorochowski and Johannes A. Roubos (DSM).

6.1 Introduction

CRISPR-Cas13 effector proteins function as a gRNA programmable RNases due to two Higher Eukaryotes and Prokaryotes Nucleotide-binding (HEPN) domains[207]. The presence of these domains is a common feature for all proteins classified as Cas13 (type VI-A, VI-B, VI-C and VI-D)[208], which otherwise exhibit limited sequence similarity, gRNA design preference and cleavage patern[209]. The ability to target single stranded RNA (ssRNA) in programmable manner allows for these systems to specifically regulate transcript abundance, thus controlling gene expression without altering or interacting with the underlying DNA. In eukaryotes, Cas13a nucleases provide a genetic tool for specific RNA knockdown, which to date has been applied in mammalian[70] and human cells[210], mosquitoes[211] and for interference of viral RNA in plants[90,91]. However, in prokaryotes attempts to use Cas13a have led to cellular toxicity and dormancy/death defined as *collateral effect*[68]. This process results from unspecific *trans* activity of Cas13 protein upon formation of a complex with gRNA and introduction of the first cleavage in the target encoded by gRNA[69]. The toxicity of this CRISPR system is a result of its purpose in nature, mainly preventing virial replication and spread in bacterial population infected by phages by initiating cellular dormancy[86]. Collateral effect has not been detected when the CRISPR-Cas13 system is used for transcript knockdown in higher eukaryotic cells[70,89–91] and RNA editing in unicellular fission yeast[212]. The exact explanation for collateral damage in some organisms but not others is yet to be fully understood but a potential role might be played by cellular environment or cellular response triggered by Cas13 cis-cleavage.

In this chapter, we explore the potential use of CRISPR-Cas13 for a specific target knockdown and investigate the collateral effect in *Saccharomyces cerevisiae* using four nucleases belonging to subtype VI-A (Cas13a) and VI-B (Cas13b). A previous characterization of Cas13 orthologues assayed for RNA cleavage[70,89] revealed the highest activity of Cas13a from *Leptotrichia shahii* (LshCas13a) and *Leptotrichia wadei* (LwaCas13a) and Cas13b derived from *Bergeyella zoohelcum* (BzoCas13b) and *Prevotella sp.* P5-125 (PspCas13b) which were selected for our experiments. Recently discovered accessory proteins Csx27 and Csx28 naturally control (*i.e.* repress and enhance, respectively) the activity of Cas13b, opening avenues for regulation of Cas13 based genetic circuits[92].

6.2 Materials and methods

Strains and cultivation conditions

Experiments were conducted in *S. cerevisiae* CEN.PK113-7D. Yeast cultures were grown in YEPh medium (yeast extract, 1%, Bacto™ Becton–Dickinson, Franklin Lakes, NJ, USA and phytone peptone, 2%, Difco™ BD), supplemented with either D-galactose or D-glucose (2%, Sigma Aldrich, St Louis, MO, USA). Selection of plasmids in yeast was achieved with nourseothricin (NTC, 200 µg/mL, Jena Bioscience, Germany) and geneticin (G418, 200 µg/mL, Sigma Aldrich). *E. coli* NEB 10-beta cells were used for plasmids propagation and cultivated in 2*PY medium with ampicillin (100 µg/mL, Sigma Aldrich). Yeast growth was monitored in micro-fermentation using BioLector II device (m2p-labs, Baesweiler, Germany) and MTP-R48-B FlowerPlate (m2p-labs). The chamber environment was controlled to maintain 30 °C, 800 rpm, 80% rh, O_2 20.95% and the biomass was detected at wavelengths em. 620 nm/ex. 620 nm and gain 3 in 15 min intervals for 72 h. The precultures were grown under repressive conditions (*i.e.* YEPh medium with glucose as the C-source) for one day at 30 °C and 250 rpm and subsequently cells were harvested and washed in ultrapure water. Cell suspension was used to inoculate YEPh medium supplemented with galactose or glucose to the starting OD_{600} of 0.1. Growth experiments were conducted in biological triplicates. To estimate the number of dead cells the growth experiment was conducted in 100 mL shake flasks containing 20 mL of medium and incubated at 30 °C with 250 rpm shaking. Cultures were inoculated as described for the micro-fermentation settings, with the only difference in starting OD_{600} of 0.2. Cultures subjected to the heat shock were firstly cultivated for 24 h under inducive conditions followed by 1 h incubation at 42 °C. For osmotic shock cells were grown in YEPh medium containing 1 M NaCl and the inducer (galactose).

Strain construction

Strains CC001-012 contained expression cassettes of *Cas13*, *mCherry* and *eGFP* able to integrate in non-coding region between *SOD1* (YJR104C) and *ADO1* (YJR105W) located on chromosome 10 (INT4). *eGFP* expression cassette contained *ENO1* promoter from *K. lactis* and native *TAL1* terminator, served as a target for Cas13, while *mCherry* was controlled by native *FBA1* promoter and *TDH3* terminator and was used as a reference reporter. *Cas13* nuclease gene was expressed from either a strong constitutive *TEF1* promoter (strains CC001,2,7,8) or a galactose inducible *GAL10* promoter (strains CC003-6,9-12), both derived from *S. cerevisiae* and used in combination with *GND2* terminator (Table 6-1). *Cas13* constructs were further diversified by presence (strains CC001-4,7-10) or absence of a C-terminal SV40 nuclear localization signal (NLS, strains CC005,6,11,12).

To create strains CC001-4,7-10 with integrated *Cas13* nuclease gene, *S. cerevisiae* CEN.PK113-7D genome was edited with CRISPR-Cas9 expressed from plasmid pCSN061 and gRNA gINT4 provided as linear fragment (1000 ng) along with linearized recipient plasmid pRN1120 (100 ng). Integrated expression cassettes were assembled via *in vivo* recombination from individual parts (100 ng) encoding promoter, open reading frame (ORF) coding *Cas13a* with NLS, *eGFP* and *mCherry*, terminator and 50 bp connectors. The assembly was guided by 50 bp homology flanks added to each part via PCR (primers pKC001-020, SI Table 6-1). Yeast transformation was performed using the LiAc/ssDNA/PEG method (Chapter 3)[148] .

Table 6-1. Cas13 expression cassettes in strains used in Chapter 6.

Strain name	Genotype
CC001	TEF1p LshCas13a NLS GND2t
CC002	TEF1p LwaCas13a NLS GND2t
CC003	GAL10p LshCas13a NLS GND2t
CC004	GAL10p LwaCas13a NLS GND2t
CC005	GAL10p LshCas13a GND2t
CC006	GAL10p LwaCas13a GND2t
CC007	TEF1p BzoCas13b NLS GND2t
CC008	TEF1p PspCas13b NLS GND2t
CC009	GAL10p BzoCas13b NLS GND2t
CC010	GAL10p PspCas13b NLS GND2t
CC011	GAL10p BzoCas13b GND2t
CC012	GAL10p PspCas13b GND2t

Strains CC005,6,11,12 harbouring *Cas13* constructs without NLS were obtained by removing the NLS sequence in strains CC003,4,9,10, respectively. For this purpose, strains CC003,4,9,10 were genomically engineered with CRISPR-Cas9 (plasmid pCSN061) and gRNA targeting NLS (gNLS). To repair the resulting double stranded break and create a functional Cas13 expression cassette, a DNA donor fragment was supplied. The donor DNA encoded 47 bp of Cas13 gene sequence at the 3' end, stop codon and 50 bp or 24 bp of *GND2* terminator at the 5' (SI Table 6-4).

A proper assembly and integration of expression cassettes was verified by Sanger (strains CC003-12) or Oxford Nanopore Technology sequencing (strains CC01,2). Removal of NLS was confirmed by Sanger sequencing (strains CC005,6,11,12). To generate template for sequencing genomic DNA was isolated from selected transformants (Chapter 3) and amplified by PCR using primers pKC021-022. For Sanger sequencing primers pKC023-51 were used in a reaction with BigDye™ Terminator v3.1 Cycle Sequencing Kit (ThermoFisher Scientific, USA), subsequently purified with NucleoSEQ columns for dye terminator removal (Macherey-Nagel, Düren, Germany). The analysis was performed with Genetic Analyzer 3500XL (Applied Biosystems, Waltham, MA, USA). MinION sequencing was conducted run using a FLOW-MIN106 (9.4 SpotON) flow cell and the MinION MIN-101B sequencer, as described in Chapter 5. Briefly, primers pKC052-70 were used to enrich the region of interest with unique barcodes by PCR to permit for pooling sequencing templates. Purified PCR products were prepped for sequencing using a ligation kit (SQK-LSK 109, Oxford Nanopore Technologies, UK).

gRNA design

gRNAs were designed to target either the *eGFP* or *ADE2* transcript with no impact on the expression of native *S. cerevisiae* genes, assessed with Burrows–Wheeler Aligner (BWA) and SAMtools[177,178]. The secondary structure of gRNA and potential misfolding was predicted with the ViennaRNA package[213]. Cas13 proteins exhibit preference for a protospacer flanking site (PFS), which refers to specific composition of region following the 3' end (LshCas13a), 5' end (BzoCas13b) or flanking both ends of the target

(PspCas13b). Cas13a gRNAs targets were selected taking into consideration PFS requirement and optimal length of the spacer and direct repeat (SI Table 6-2). The list of spacers used in strains construction and evaluation of Cas13 systems is enclosed in SI Table 6-3 and visualized in SI Figure 6-1. gRNA expression cassettes were comprised of the *SNR52* promoter, the *SUP4* terminator[49], the direct repeat of Cas13 or the structural component of Cas9, the spacer and homology region to the recipient plasmid pRN1120[65]. These cassettes were ordered as synthetic DNA (IDT or Twist Bioscience, San Francisco, CA, USA) and amplified in PCR reaction with primers pKC071,72 (Cas9 gRNA) or pKC073,74 (Cas13 gRNA). Plasmid pRN1120 was linearized in PCR reaction using primers pKC075,76 (for Cas9 gRNAs) or pKC077,78 (Cas13 gRNAs).

Molecular biology techniques

PCR reactions intended for cloning or sequencing were performed with Q5 polymerase (New England Biolabs, Ipswich, MA, USA) and primers obtained from IDT (Leuven, Belgium). The resulting amplicons were purified using the Wizard SV gel and PCR clean up kit (Promega, Madison, WI, USA). Template for colony PCR was prepared using SDS-LiAc solution, as detailed in Chapter 3. Reactions verifying integration of gRNA expression cassette in pRN1120 were prepared with either KAPA Robust HotStart polymerase (KAPA Biosystems, Cape Town, South Africa) or Phusion (NEB) and primers pKC079,80. RNA was isolated from yeast cells with NucleoSpin RNA kit (Macherey-Nagel, Düren, Germany) and analysed on agarose gel electrophoresis or Agilent 2100 Bioanalyzer (Santa Clara, CA, USA). Fluorescence of transformants was visualized with a colony picker Qpix 450 using FITC wavelength (Molecular Devices, San Jose, CA, USA). For quantitative PCR (qPCR), cultures of strain CC004 (LwaCas13a NLS) were grown for 5 h in YEPh medium supplemented with galactose or glucose to simulate induction or repression of Cas13a, respectively. After RNA isolation, genomic DNA was removed in enzymatic reaction using recombinant DNase I (Roche, Basel, Switzerland). cDNA library was prepared with RevertAid H Minus First Strand cDNA Synthesis Kit and oligo(dT)$_{18}$ primers. Gene transcript levels were measured in qPCR reaction using DyNAmo ColorFlash SYBR Green qPCR Kit (ThermoFisher Scientific) and primers pKC081-90. Along with quantification of heterologous transcripts *eGFP*, *mCherry* and *LwaCas13a*, the analysis included house-keeping genes *ACT1* and *TAF10*. Relative gene expression was calculated using method described by Vandesompele[214] in order to normalize abundance of heterologous transcript by two reference genes (*i.e. ACT1* and *TAF10*). qPCR was conducted in technical triplicates prepared independently for two biological replicates.

Flow cytometry and dead staining

Flow cytometry was used to evaluate whether Cas13 can mediate *eGFP* knockdown. For this purpose, yeast cultures were grown in 2 mL of selective medium in a multitier plate (well volume 10 mL) at 30 ºC, 550 rpm and 80% rh. Diluted cultures (x20 in physiological salt) were analysed with flow cytometer (BD FACSAria Fusion, Becton–Dickinson) with a bandpass filter set at 530/30 nm for eGFP detection and 610/20 nm for mCherry. Gating was applied to record events for single cells and exclude doublets and debris from the analysis. For dead staining, cells were resuspended in physiological salt to ~$5 \cdot 10^7$ cell/mL and mixed with SYTOX® Red Dead Cell Stain (Invitrogen). After 15 min incubation in dark, fluorescence was measured with flow cytometer using bandpass filter at 670/30 nm. Live cells were distinguished from dead by applying a gate on far red fluorescent channel. For dead control the cells dilution was incubated for 10 min at 80 °C.

Data analysis

Data analysis and processing was conducted in Python 3.0 and visualized with the Matplotlib package. RNA integrity was assessed with 2100 Expert Software. Flow cytometry data was recorded with BD FACSDiva 8.0.2 software and processed with the FlowCal package for Python[180].

6.3 Results

Expression of Cas13 proteins in *S. cerevisiae*

To evaluate potential toxicity of the system in *S. cerevisiae*, Cas13 was tested with or without an NLS and expression was controlled by galactose inducible *GAL10* promoter (Figure 6-1A, left panel). The stability of the construct was ensured by integration into the genome of the expression cassette harbouring *Cas13* gene. gRNA was expressed constitutively using the *SNR52* promoter from a multicopy plasmid pRN1120[65] to provide flexibility in spacer selection. Cas13 gRNAs have a different orientation of direct repeat (DR) relative to the spacer, with a DR-spacer order in Cas13a gRNA and spacer-DR for Cas13b gRNAs (Figure 6-1A, right panel)[92]. For the ease of evaluating the knockdown efficiency of Cas13a, we chose targets defining cellular phenotype, mainly *eGFP* transcript produced from genome integrated construct. Upon transformation of gRNA, cells were selected on medium supplemented with NTC for a maintenance of pRN1120 plasmid and either galactose or glucose to reflect inducive or repressive conditions for Cas13 expression. A strong inhibition in colony formation was observed for LshCas13a and LwaCas13a (SI Figure 6-2). Expression of targeting gRNAs gGFP1, gGFP2 in combination with LshCas13a NLS caused a 50% decrease in colony numbers and affected colony size under repressive growth conditions. The LshCas13a construct lacking NLS displayed growth inhibition (no formation of colonies) in combination with gGFP4. Induction of a Cas13a orthologue derived from *Leptotrichia wadei* reduced colony formation for one gRNA (gGFP3) regardless of presence or absence of the NLS. Growth impairment was not observed upon expression of any of the gRNAs and BzoCas13b NLS or BzoCas13b (SI Figure 6-3A). However, induction of the *GAL10* promoter in the construct harbouring PspCas13b diminished the capability to form colonies, also in combination with a non-targeting gRNA gNone or empty gRNA recipient plasmid which indicates toxicity of PspCas13 in *S. cerevisiae* (SI Figure 6-3B). Based on these results we sought to determine source of this toxicity and evaluate whether it is caused by the collateral effect of Cas13 nucleases. In the following experiments the focus was limited to LwaCas13a as this orthologue has been shown to be the most effective among Cas13a[70]. Fusion of NLS had no effect on performance of LwaCas13a, therefore only one construct containing NLS was used. Impact of LwaCas13a on growth was observed solely for one gRNA (gGFP3), thus four additional gRNAs targeting native *ADE2* transcript were designed. ADE2 is involved in *de novo* purine nucleotide biosynthesis and its impaired expression in cells deprived of adenine leads to accumulation of a red pigment[215]. Only one transformant was formed by strain LwaCas13a NLS with gRNA gADE2 and gADE3 (SI Figure 6-2B).

Combing data from two independent experiments, the number of LwaCas13a NLS transformants under inducing conditions (plating on galactose) dropped 10-fold in comparison with repressing conditions (plating on glucose) for gRNAs targeting the *eGFP* and *ADE2* (Figure 6-1B). In contrast, the control

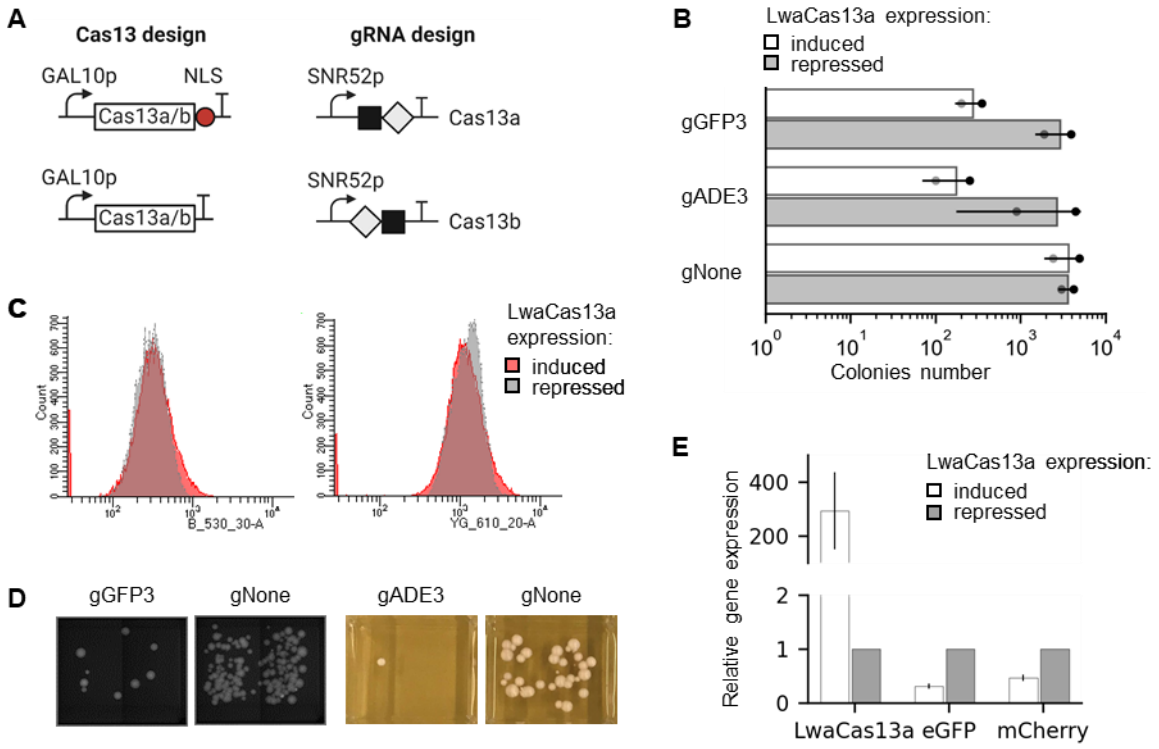


Figure 6-1. Implementation of CRISPR-Cas13 systems in *S. cerevisiae*. A. The system used for expression of Cas13 and gRNA. *Cas13a* is expressed from a galactose inducible *GAL10* promoter with or without NLS and genome-integrated. Orientation of spacer (grey diamond) and direct repeat (DR, back square) was adapted for Cas13a (DR-spacer) and Cas13b (spacer-DR). gRNA controlled by the *SNR52* promoter is provided as a linear fragment assembled via *in vivo* recombination into recipient plasmid pRN1120. **B.** Number of colonies formed upon transformation of gRNA targeting *eGFP* or *ADE2* transcripts (gGFP3 and gADE3, respectively) or non-targeting control (gNone) and under conditions inducing or repressing expression of LwaCas13a. **C.** eGFP (left panel) and mCherry fluorescence (right panel) of strain LwaCas13a NLS with gGFP3 after 24 h of growth in inducive (galactose) and repressive conditions (galactose) evaluated by flow cytometry. **D.** Transformants of LwaCas13a NLS and gRNAs gGFP3 or gNone (negative control) visualized colony picker with FITC filter used for detection of eGFP fluorescence (left panel). Photography of transformants of strain LwaCas13a NLS and gADE3 and gNone for evaluation of red colour of colonies (right panel). **E.** Relative gene expression of *LwaCas13a*, *eGFP* and *mCherry* evaluated by qPCR. Data represent mean ± standard deviation (n=2).

containing a non-targeting gRNA (gNone) produced a similar number of transformants regardless the sugar source. Transformants obtained under conditions inducing production of LwaCas13a did not display reduced eGFP fluorescence (Figure 6-1C, 6-1D left panel, when targeting *eGFP*) nor change in colonies colour from white to red (Figure 6-1D right panel, when targeting *ADE2*), despite integration of gRNA into the recipient gRNA plasmid was confirmed by colony PCR (gGFP3 and gADE3, respectively, data not shown).

The limited number of transformants appearing on agar under inducing conditions could have been caused by selective advantage of mutations which would disable CRISPR-Cas13a system if its activation was lethal to the cell. Therefore, the expression cassettes of Cas13a and gRNA were sequence-verified for the obtained transformants of strain CC001 expressing LshCas13a NLS from a *TEF1* promoter. The analysis of four transformants displayed no changed in gRNA gGFP3 and for three of the tested colonies no mutations within the LshCas13a gene. To exclude the inability to initiate trans activity of LshCas13a due to mutation in the protospacer region, which would disable structural rearrangement of LshCas13a upon the cis cleavage, *eGFP* gene was sequenced for one of the colonies, however, without detection of any mutations.

Quantification of heterologous transcripts by qPCR revealed a significant increase in LwaCas13a expression when cells were grown in galactose (inducing conditions) in comparison with glucose (repressing conditions) due to a galactose inducible promoted selected to control expression of LwaCas13a (Figure 6-1E). In this design, an *eGFP* transcript was used as a target for LwaCas13a whereas *mCherry* served as an internal control, therefore a functional knockdown would lead to decreased abundance of *eGFP* transcripts while maintenance of stable *mCherry* expression. The qPCR results showed reduced number of *eGFP* (69±5%) and *mCherry* (53±6%) transcripts when expression of LwaCas13a was induced.

Collateral effect in *S. cerevisiae*

To understand if the source of the toxicity lies in the collateral effect of LwaCas13a, total RNA was isolated from strains harbouring targeting and non-targeting gRNAs cultured for 5h under inducive or repressive conditions for Cas13a expression. Analysis of RNA by agarose gel electrophoresis revealed RNA degradation for strains carrying targeting gRNAs under Cas13a inducing conditions, which was not observed when Cas13a was repressed or for gRNAs that encoded a non-targeting spacer (**Figure 6-2A**).

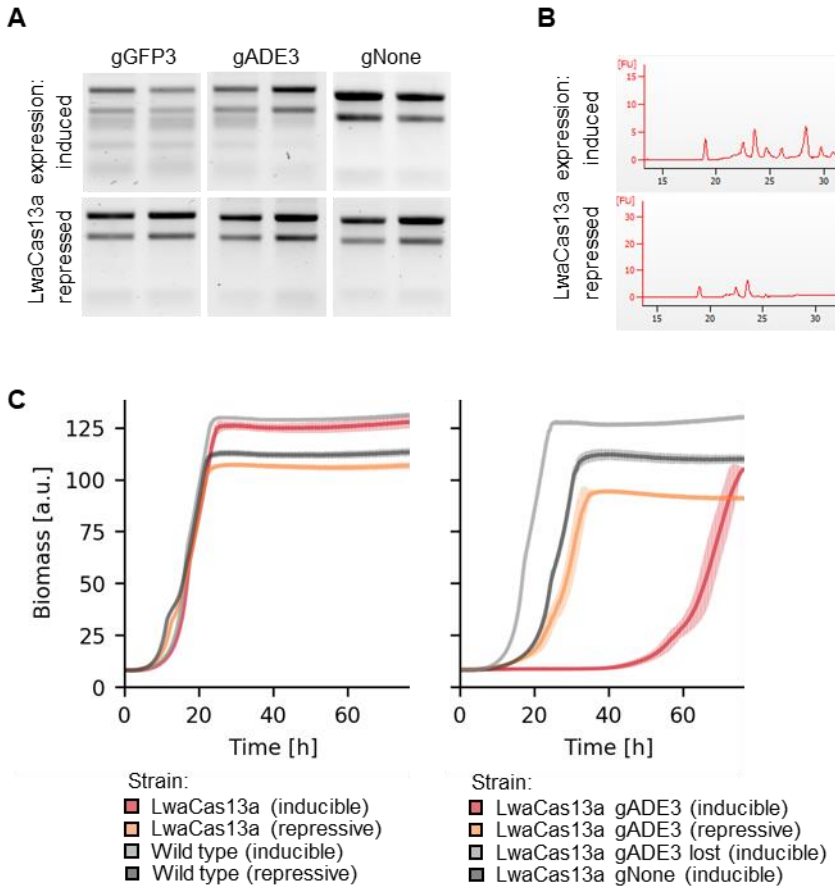


Figure 6-2. Collateral effect of CRISPR-LwaCas13a NLS in *S. cerevisiae*. A. Integrity of total RNA analysed with agarose gel electrophoresis. Total RNA was isolated after growing strains with gRNAs gGFP3, gADE3 and gNone for 5h under conditions inducing or repressing LwaCas13a NLS expression (n=2). **B.** Total RNA analysed with Agilent 2100 Analyzer from strain carrying gGFP3 and under expression (growth on galactose) or repression (growth on glucose) of LwaCas13a NLS. **C**. Growth experiment of strain LwaCas13a NLS in combination with gRNAs gADE3 and gNone. Data represents average ± 1 standard deviation (n=3).

Next, we analysed the isolated RNA with the Agilent 2100 Bionalyzer to estimate the degraded fraction by comparing abundance of ribosomal 18S and 25S RNA. Analysis of total RNA from culture grown under repressive conditions showed clear peaks corresponding to ribosomal RNA (rRNA, 5S, 5.8S, 18S and 25S, Figure 6-2B lower panel). However, induction of LwaCas13a caused a decrease in the peak area of 18S and 25S and presence of shorter RNA fractions. To quantify the degradation, abundance of ribosomal RNA was estimated from the peak area which was represented as a ration between inducible and repressive conditions for LwaCas13a NLS expression, independently for 18S and 25S. Expression of gGFP3 caused a 50% decrease in 18S and 25S abundance during cultivation in inducible medium in comparison with the medium repressing Cas13 transcription (Table 6-2). The effect was observed to a lesser extent when gADE3 gRNA was used, leading to 21-28% decrease in 18S and 25S.

Table 6-2. Abundance of ribosomal RNA (18S and 25S) under inducible and repressive conditions for LwaCas13a NLS expression

gRNA	$A_{18S\ galactose}$ / $A_{18S\ glucose}$	$A_{25S\ galactose}$ / $A_{25S\ glucose}$
gGFP3	0.49 ± 0.05	0.44 ± 0.06
gADE3	0.79 ± 0.14	0.72 ± 0.13
gNone	0.99 ± 0.01	0.99 ± 0.02

Degradation of rRNA is expected to have a severe effect on translation which may thus cause a defect in cellular growth. Micro-fermentation revealed minimal impact on biomass formation upon integration of the LwaCas13a NLS construct into the genome and induction of its expression without a gRNA (Figure 6-2C, left panel). When the strain harbouring LwaCas13a NLS was supplied with a gRNA and cultivated in presence of galactose, growth was undetectable in the first two days of incubation and entrance to the exponential phase was delayed to approximately 50 h (Figure 6-2C, right panel). Changing the C-source to glucose improved growth of this strain (LwaCas13a NLS gADE3 repressive) and caused only 13% decrease in biomass formation during the stationary phase in comparison with the parental strain lacking a gRNA (LwaCas13a NLS repressive, Figure 6-2C right and left panel). A similar decrease in biomass was observed upon expression of a non-targeting gRNA (gNone) and LwaCas13a NLS, leading to 12% drop in comparison with the parental strain LwaCas13a NLS. This growth defect might be linked to gRNA expression from the *SNR52* promoter, as observed in Chapter 5, where expression of a non-targeting gRNA caused a 17% decrease in biomass compared to the empty gRNA recipient plasmid. To test whether CRISPR-LwaCas13a toxicity is reversible, the strain expressing gADE3 was repeatedly cultivated and streaked to remove plasmid pRN1120 with the gRNA expression cassette. After three cycles, plasmid removal was confirmed by restreaking the strain on agar supplemented with NTC. Growth of the plasmid free strain (LwaCas13a gADE3 lost) was restored and during stationary phase biomass level resembled strain without gRNA (LwaCas13a NLS) and wild type cultivated in presence of galactose (Figure 6-2C, right panel).

To understand whether the delayed growth of strain LwaCas13a NLS supplied with a targeting gRNA is related to the presence of the inducer in the medium and inactivation of Cas13 system, sugar concentration was measured during a shake flask fermentation (SI Figure 6-4). After 24 h of incubation,

galactose concentration in the medium was high and oscillated around 14 g/L while glucose was depleted in cultures grown under repressive conditions. Microbial growth was suppressed in presence of galactose (OD_{600} reached 2.9) in comparison with medium containing glucose (OD_{600} of 14.1) for either of the targeting gRNAs (SI Figure 6-5). Galactose content dropped after 48 h in contrary to the increased biomass formation. A measurement conducted after 72 h revealed a complete depletion of galactose while OD_{600} of the induced cultures reached average of 21.8 in comparison with 24.6 of non-induced cultures. These results show that upon activation of Cas13a system yeast cells have limited capability to form biomass, however, galactose, which served as the only sugar source and inducer of Cas13a expression, is consumed. Depletion of the inducer inactivates Cas13a system and restores growth.

Potential lethality of the CRISPR-LwaCas13 system was tested during a growth experiment using a dead cell stain detected with a flow cytometer. Although growth was strongly suppressed in the first day of analysis (SI Figure 6-5), the population of live cells constituted more than 90% of total cells upon induction of Cas13a (Figure 6-3A, SI Figure 6-6). A slight decrease in the number of live cells was observed after 72 h from the induction (by 2.4% and 2.0% for gADE3 and gGFP3, respectively), whereas for the remaining time points significant decrease in population of live cells was not observed. Next, we examined impact of LwaCas13a system in combination with external stress factor such as heat and osmotic shock. The reasoning behind is that if the collateral effect of Cas13a is sublethal, severely damaged cells will be incapable to withstand an additional stress factor, causing lethality in cells population. In *S. cerevisiae* heat shock occurs in temperature above 36 °C triggering expression of heat shock proteins (Hsp) which act to mitigate heat impact on cellular processes and proteins[216]. Osmotic shock can be caused by high concentration of salt in the grown medium leading to imbalance of the intracellular environment and drop in turgor pressure[217]. Exposure of strain with active LwaCas13a system to the heat shock insignificantly decreased the number of live cells while osmotic shock had no impact on the population of live cells (Figure 6-3B). However, a shift in fluorescence intensity by 76-85% was observed for three out of four cultures tested under osmotic shock which would indicate injury of cells (due to the stain permeability via compromised plasma membranes, SI Figure 6-7).

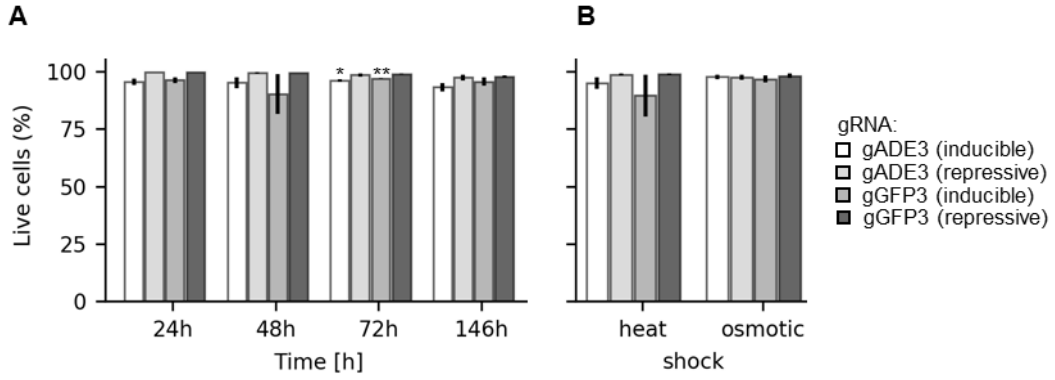


Figure 6-3. Lethality of CRISPR-LwaCas13a system in *S. cerevisiae*. **A**. Live cells during a 6 days batch fermentation under induction and repression of LwaCas13a NLS and two targeting gRNAs (gADE3, gGFP3). **B**. Live cells in cultures with active LwaCas13a system and external stress factor (*i.e.* heat or osmotic shock). Bars represent mean ± 1 standard deviation (n=2). * $p<0.05$, ** $p<0.01$, two-tailed t-test for strain with the same gRNA but with/without LwaCas13a NLS expression.

In conclusion, although Cas13a exhibit collateral effect in *S. cerevisiae* causing degradation of ribosomal RNA, the effect is not lethal to the cells which can withstand external stress factors.

6.4 Discussion

Here we have established a CRISPR-Cas13a and Cas13b system and evaluated its impact on *S. cerevisiae*. Although this system was successfully applied for transcript knockdown in eukaryotes[70,90,91,210,211], this capability was not observed in *S. cerevisiae* for any of the tested Cas13a and Cas13b orthologues. Activation of the CRISPR-Cas13 system with gRNAs targeting *eGFP* and *ADE2* transcripts did not lead to the expected change in the phenotype (decreased eGFP fluorescence and red appearance of colonies) evaluated with different detection tools (visualization of plates, qPCR and flow cytometry). Instead, expression of Cas13a nucleases in combination with the targeting gRNAs had a negative impact on cellular growth not observed for Cas13b system.

The lack of BzoCas13b function might be caused by exploration of a limited set of gRNA tested and suboptimal gRNA design. We have previously observed that in a library of gRNAs encoding different targets for CRISPRi dCas12a mediated downregulation, always yielded a number of non-functional gRNAs, despite the same design rules being applied (Chapter 5). Furthermore, new studies on gRNA architecture for Cas13 reveal new design rules which would be applicable in the future experiments[210,218,219]. Expression of PspCas13b was toxic to yeast cells as no colonies were formed even in absence of gRNA. This was not observed for PspCas13b when expressed in any other host organisms[89,218,220]. However, a similar toxic effect upon transformation with a non-targeting gRNA was noted for Cas13d in *E. coli*[221]. The reduced capability to form colonies of strains with active Cas13a system indicates cellular toxicity linked to collateral effect of Cas13a. This unspecific RNase activity of Cas13a was further confirmed by decreased integrity of RNA detected under conditions inducing the system. The RNA degradation caused a reduction in ribosomal 18S and 25S RNA which leads to the hypothesis that the growth defect might be caused by ineffective translational machinery which would unable cellular growth. Since this work was begun, the collateral damage has been reporter for more organism including mouse cerebral cortical neurones[88] and human glioblastoma cancer cells[87]. The cause of the collateral effect observed in some organism but not others is yet to elucidated.

Although the CRISPR-Cas13a system impairs growth of *S. cerevisiae*, this process is reversible when the targeting gRNA is removed. Quantification of dead cells throughout the fermentation showed that the system is not lethal to the yeast cells. Therefore, a potential application of the CRISPR-Cas13a system in *S. cerevisiae* is to control growth (Figure 6-4). A growth control device based on CRISPR-Cas13a can be built as an AND gate in order to tighten the system and prevent undesired leakage. An inducible system can be used for independent expression of Cas13a nuclease and a gRNA, *e.g.* a galactose-inducible promoter selected in this study for the control of Cas13a. Additionally, gRNA expression control can be achieved with the anhydrotetracycline-inducible *RPR1* promoter which is recognized by a RNAP III, accompanied by a stable expression of a tetracycline repressor (tetR)[96,222,223]. The advantage of the proposed system for gRNA

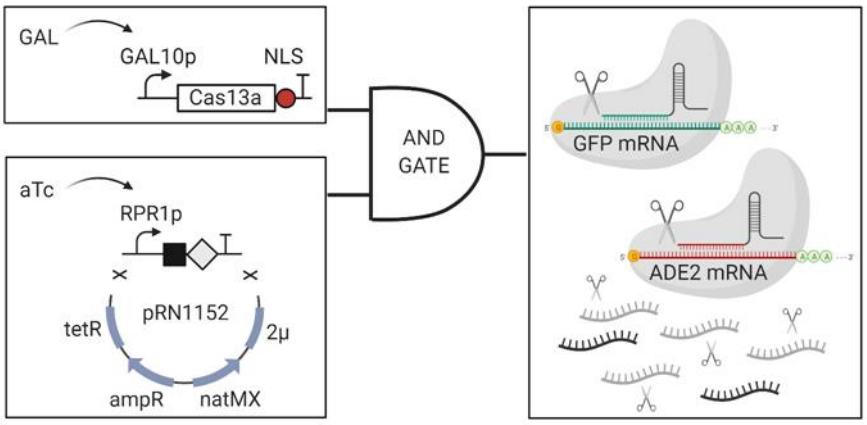


Figure 6-4. Conceptualization of CRISPR-Cas13a based growth control device. The input information defined as presence of galactose (GAL) and anhydrotetracycline (tetR) activates expression of Cas13a and gRNA, respectively. To enable aTc induction gRNA is expressed from *RPR1* promoter and the gRNA recipient plasmid pRN1152 harbours tetR expression cassette. The output of the AND gate is collateral effect which reversibly hampers cellular growth.

expression is a possibility to titrate and test whereas collateral activity is dependent on the availability of a gRNA. Presence of both inducers in the medium would activate the collateral effect of Cas13a and inhibit growth until the inducers are removed. The proposed control device could be a foundation for regulation of mixed populations during a cultivation to dynamically balance growth rate and prevent overgrowth of one of the species. For this purpose, the input in the growth control device could be substituted with signal elements participating in quorum sensing, to link the CRISPR system with cell-cell communication[224]. It should be noted that development of this system would require further analysis of the impact of Cas13a on the yeast cells.

Although the transcript knockout is not feasible in all organisms using Cas13 nucleases, a catalytically inactive CasRx Cas13d mediated translational silencing without affecting growth of *E. coli* cells[221]. In the same study a catalytically inactive Cas13a from *Leptotrichia buccalis* (Lbu) was optimized by directed evolution which revealed mutations affecting stabilization and cleaving kinetics and provided another efficient Cas13 orthologue that can be applied for transcriptional repression in *E. coli.* A further optimization of CRISPR-Cas13 systems may lead to a discovery of an effector protein suitable for a transcript knockdown in *S. cerevisiae*. This would provide a programmable post-transcriptional tool suitable for multiplex regulation, making a use of inherent Cas13 capability to process crRNA arrays[69]. Furthermore, by acquiring, sequencing and analysing genomes of microorganisms new CRISPR systems suitable for transcript knockdown in yeast and bacteria can be discovered. In fact, in November 2021, the discovery of single-effector ribonuclease Cas7-11 was reported and implemented for mRNA knockdown and editing in mammalian cells without observed toxicity[225]. Alike, *RFP* mRNA knockdown mediated by Cas7-11 in *E. coli* did not cause collateral effect observed for LwaCas13a. In conclusions, CRISPR based system for programmable mRNA knockdown in *S. cerevisiae* would expand the regulatory toolbox and provide additional level of control.

7 Biosensing stress induced by carotenoids production

This Chapter describes development of a biosensor for monitoring burden caused by the production of carotenoids. This process was supported by RNA sequencing performed by me (growth of cultures, isolation of RNA and preparation of cDNA library) and Thomas E. Gorochowski (University of Bristol, sequencing dRNA and cDNA libraries and data processing). The remaining experiments described in this Chapter were conducted by me under the supervision of René Verwaal (DSM), Thomas E. Gorochowski and Johannes A. Roubos (DSM).

7.1 Introduction

Cellular response to a change in the environment or exposure to a stress factor is mediated by transcription factors which enhance promoter activity and initiate expression of proteins to ensure cell adaptation and survival. This phenomenon can be adapted by constructing molecular devices (biosensors) capable of sensing intra- and extracellular molecules, cellular fitness and external environment[24,44,226–228]. A particular application of biosensors is to monitor well-being of a host organism during production of a compound of interest that can provide valuable information about the process such as availability of nutrients, accumulation of toxic intermediates, unfolded protein response and suboptimal conditions in the reactor such as heat and shear stress[229]. In the simplest approach, a biosensor is comprised of a native promoter sensitive to the desired stimulus and coupled with an output gene *e.g.* a fluorescent protein or a pathway to change cellular behaviour, exemplified by burden-regulated constructs. Dahl *et al.*[226] created a burden-regulated device which sensed accumulation of toxic intermediates (farnesyl pyrophosphate and HMG-CoA) during production of amorphadiene and repressed the production pathway by activating expression of an enzyme (amorphadiene synthase) which catalyses conversion of the toxic intermediate (farnesyl pyrophosphate). This was achieved by substituting promoters of the isoprenoid pathway and amorphadiene synthase with farnesyl sensitive promoters. To guide identification of native promoters responsive to the production burden, the authors exploited transcriptomics using whole-genome transcript arrays. The drawback of this system is limitation to the specific production of amorphadiene and related burden. Alternatively, a global burden sensor can be employed to mitigate any type of stress that the cell experiences during a production. In line with this approach, Ceroni *et al.*[24] determined a general burden-responsive promoter by conducting RNA sequencing for strains with various production systems (a reporter gene – green fluorescent protein, a recombinant protein or a metabolic pathway). The identified burden-responsive promoter (*htpG*) was next used in a feedback control system to dynamically optimize the production of a red fluorescent protein in a proof-of-concept case.

This chapter describes identification of a *S. cerevisiae* promoter responsive to the burden (Chapter 1, 2) caused by carotenoids production. Selection of the biosensing promoter was supported by RNA sequencing data which revealed impacts of carotenoids production on biosynthesis of ergosterol and pleiotropic drug resistance response. Next, promoters of genes strongly upregulated in carotenoids producing strain were evaluated in terms of strength and capability to express a reporter gene and a gRNA in CRISPR interference system. Finally, an attempt in understanding the tunability of the biosensing promoters with increased production of carotenoids was made.

7.2 Materials and methods

RNA sequencing

RNA sequencing was conducted for the carotenoids producing strain CAR-034 (expressing *crt* genes from strong constitutive promoters, Chapter 4) and the wild type strain CEN.PK113-7D in biological duplicates. Cultures subjected to the analysis were inoculated from the seed culture to the starting OD_{600}=0.2 in 100 mL of fresh YEPD medium. YEPD medium composition was following: 1% yeast extract (Bacto™ Becton–Dickinson, Franklin Lakes, NJ, USA), 2% phytone peptone (Difco™ BD), and 4% glucose (Sigma Aldrich, St Louis, MO, USA). Cultures were incubated at 30 °C, 250 rpm shaking and sample for RNA analysis was collected after 6 h and 24 h to gain knowledge about transcriptome during growth on glucose and ethanol. To confirm which C-source was metabolized at the time of harvest, concentration of glucose and ethanol in the medium was quantified with Dionex Ultimate 3000 HPLC system (Thermo Fisher Scientific) and Aminex HPX-87H column (300 x 7.8 mm, Bio-Rad, Hercules, CA, USA). RNA was isolated using NucleoSpin RNA kit (Macherey-Nagel, Düren, Germany) and the quality of RNA was evaluated using Agilent 2100 Bioanalyzer (Santa Clara, CA, USA). The obtained RNA fraction was split to generate the cDNA library (2500 ng of total RNA per sample) and for direct RNA analysis (500 ng of total RNA per sample). In both cases, sequencing material was prepared using Oxford Nanopore Technology kits, specifically cDNA-PCR Sequencing (SQK-PCS109) and Direct RNA Sequencing Kit (SQK-RNA002). Samples were analysed with FLOW-MIN106 (9.4 SpotON) flow cell and the MinION MIN-101B sequencer. Raw data generated by the nanopore sequencer was subjected to base calling using Guppy and the resulting reads were counted with Nanocount. The differential expression analysis was performed with DESeq2. A change in gene expression was considered if the $\log_2FC \leq -1$ (downregulation) or $\log_2FC \geq 1$ (upregulation). Expression of ergosterol biosynthesis genes was mapped to pathway using Yeast Biocyc[230].

Strain construction

To test whether promoters of upregulated genes can actively drive expression, carotenogenic strain CAR-034 (Table 7-1) was engineered with CRISPR-Cas9 to integrate an expression cassette composed of the promoter of interest, *RFP* (*mCherry*) and the *TDH3* terminator integrated into locus INT4 (non-coding region between SOD1 and ADO1 on chromosome 10). Selected promoters *ERG3*, *ERG5* and *PDR5* were amplified from genomic DNA of wild type strain CEN.PK113-7D, isolated according to the method enclosed in Chapter 3. PCR amplification of promoter regions was performed with Q5 polymerase (New England BioLabs, Ipswich, MA, USA) and primers containing BsaI recognition sites (pKC001-008, SI Table 7-1 IDT, Leuven, Belgium) to enable assembly of P-O-T via Golden Gate cloning (using BsaI and T4 ligase from NEB). Integration of the *RFP* expression cassettes into the genomic DNA was verified by colony PCR according to the protocol described in Chapter 3 using Phusion polymerase (NEB).

gRNA expression from selected promoters was conducted using carotenogenic strain CAR-047 in which dCas12a E925A NLS Mxi1 driven by the strong constitutive promoter *TEF1* was integrated. To construct strain CAR-047 (Table 7-1), expression cassette was integrated into the INT4 locus of parental strain CAR-035 (Chapter 4) using CRISPR-Cas9. The expression cassette of *dCas12a* was obtained by PCR using genomic region of strain DC010 (Chapter 5) as a template and primers pKC009-010. The correct integration of *dCas12a* was verified by Sanger sequencing using primers pKC011-024 and the BigDye™

Terminator v3.1 Cycle Sequencing Kit (ThermoFisher Scientific, USA). Sequencing reactions were purified with NucleoSEQ columns for dye terminator removal (Macherey-Nagel, Düren, Germany) and analysed with the Genetic Analyzer 3500XL (Applied Biosystems, Waltham, MA, USA).

The tunability of biosensing promoters was examined using carotenoids producing strains CAR-048, CAR-049 and FR016 (Table 7-1) which were engineered to introduce *RFP* (*mRuby2*) expression cassette into the INT4 locus. Genome editing of abovementioned strains was conducted using CRISPR-Cas9 technology as described in Chapter 5. Cells were selected on YEPD agar supplied with nourseothricin (NTC, 200 µg/mL, Jena Bioscience, Germany) and geneticin (G418, 200 µg/mL, Sigma Aldrich). Integration was confirmed by colony PCR, as described before (Chapter 3).

Table 7-1. Genotype of engineered strains used for testing biosensing promoters

Strain	Genotype	Chapter	Purpose
CAR-034	Sbay_TDH3p - crtE - Sc_TDH3t, Smik_TEF1p - crtYB - Sc_PDC1t, Kl_ENO1p - crtI - Sc_TAL1t	4	RNA-seq, test RFP expression of the selected promoters
CAR-047	Kl_PGK1p - crtE - Sc_TDH3t, Kl_ TEF2p - crtYB - Sc_PDC1t, Kl_ OLE1p - crtI - Sc_TAL1t, Sc_TEF1p - dLbCas12a E925A - NLS – Mxi1 - GND2t	7	Test expression of gRNA using selected promoters
CAR-048	Kl_PGK1p - crtE - Sc_TDH3t, Kl_ TEF2p - crtYB - Sc_PDC1t, Kl_ OLE1p - crtI - Sc_TAL1t, Kl_ENO1p - mRuby2 - Sc_TAL1t	4	Test tunability of biosensing promoters using RFP
CAR-049	Kl_TDH2p - crtE - Sc_TDH3t, Kl_ YDR1p - crtYB - Sc_PDC1t, Kl_ LEU2p - crtI - Sc_TAL1t, Kl_ENO1p - mRuby2 - Sc_TAL1t	4	Test tunability of biosensing promoters using RFP
FR016	Kl_ENO1p - mRuby2 - Sc_TAL1t	7	Test tunability of biosensing promoters using RFP

gRNA design

gRNAs used to test biosensing promoters were flanked by self-cleaving ribozymes[57] to enable expression from RNAP II promoters (ERG3 and PDR5p) and the native yeast terminator TDH3t. gRNAs contained 20 nt spacers G34, G37 targeting the ENO1p from *K. lactis* or G42 encoding a dummy sequence (SI Table 7-2), and the direct repeat sequence of LbCas12a. A non-targeting gRNA G42 controlled by the RNAP III promoter SNR52p was included in experiments as an additional control. Genome editing mediated by CRISPR-Cas9 was performed with gRNA gINT4 expressed from SNR52p.

Flow cytometry

Fluorescence of strains harbouring biosensor-reporter gene construct was quantified by flow cytometry (BD FACSAria Fusion). A gating was applied such that analysis included 20.000 events recorded for single cell while double cells and debris were excluded. RFP fluorescence was detected at wavelength 610/20 nm and recalculated to Molecules of Equivalent Phycoerythrin-TR (MEPTR) using Rainbow calibration beads with 8-peaks (BioLegend, London, UK) and the FlowCal Python package[180]. For data recording and analysis BD FACSDiva 8.0.2 software was used.

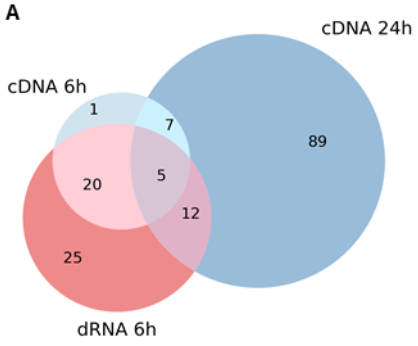
7.3 Results

Cell response to the production of β-carotene

We conducted a differential gene expression analysis to determine the transcriptional response of cells to the production of β-carotene. By gaining nann be implemented into genetic circuits as a biosensor for monitoring production related stress. For this purpose RNA sequencing (RNA-seq) was conducted by analysis of a complementary DNA (cDNA) library and direct RNA (dRNA). Standard RNA-seq methods include preparation of a cDNA library which is obtained by reverse transcription of mRNA, thus it allows to transform transcriptome information to a more stable form. With the recent developments in the RNA-seq field, direct sequencing of RNA became feasible[231]. Direct sequencing of RNA prevents the bias towards certain transcripts during preparation of a cDNA library *i.e.* reverse transcription and PCR, and allows to study full length RNA while maintaining its features lost in cDNA analysis such as base modifications.

The analysis was performed during a batch cultivation of strain CAR-034 and the wild type after 6 h and 24 h from the inoculation, to illustrate gene expression in two phases of the diauxic growth (*i.e.* during metabolism of glucose and ethanol, SI Figure 7-1). The comparison of genes upregulated according to the analysis of different sample preparations (cDNA and dRNA) and timepoints (6 h and 24 h) is presented in a Venn plot in Figure 7-1A. Sequencing dRNA led to discovery of higher number of upregulated genes in comparison with cDNA library (62 and 33 genes, respectively, Figure 7-1B) for the sample collected after 6 h fermentation. cDNA analysis conducted at different timepoints showed an increased number of upregulated genes after 24 h incubation in comparison with 6 h (113 and 33 genes, respectively), presumably as a consequence of nutrients depletion and continued production of carotenoids. Numerous genes encoding enzymes responsible for ergosterol biosynthesis (*ERG*) were upregulated regardless of the

A



B

Compared samples	Upregulated genes
dRNA 6h, cDNA 6h	*ALD6, ATF2, CHA1, DUR1, ERG1, ERG3, ERG5, ERG11, ERG25, FIT3, GDH1, NCP1, NDE1, PHO3, PHO11, PHO89, SCM4, TIR1, TIR3, YEH1*
cDNA 6h, cDNA 24h	*ERG3, ERG5, ERG25, FIT3, HXT2, PFK27, YEH1*
dRNA 6h, cDNA 24h	*ERG3, ERG5, ERG6, ERG25, FIT3, GIC2, HES1, PDR5, SNQ2, SRP40, YEH1, YOX1*
dRNA 6h, cDNA 6h, cDNA 24h	*ERG3, ERG5, ERG25, FIT3, YEH1*

Figure 7-1. Gene upregulation in strain producing β-carotene in comparison with the wild type *S. cerevisiae* strain. A. Venn diagram representing number of genes upregulated when RNA-seq for conducted for cDNA library and direct RNA (dRNA) at different time of fermentation (6 h and 24 h). **B**. List of upregulated genes common for different types of RNA-seq and timepoints.

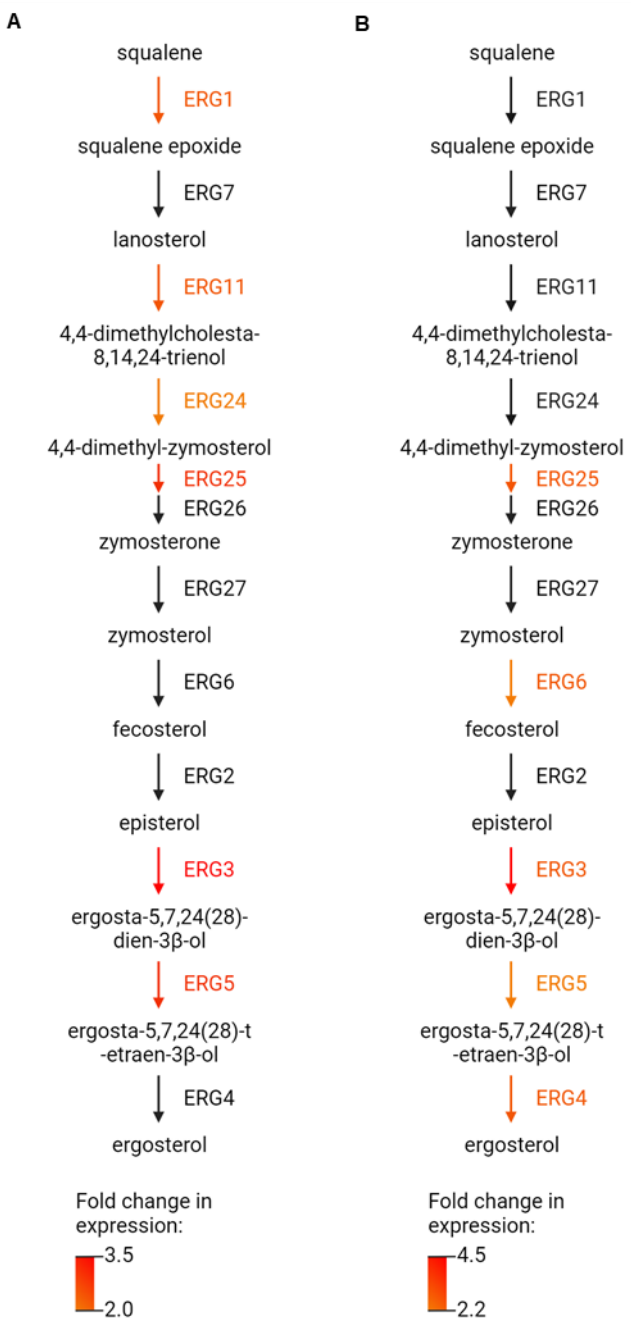


Figure 7-2. Impact of β-carotene production on ergosterol biosynthesis superpathway in *S. cerevisiae*. Enzymes involved in ergosterol formation with increased transcription are highlighted in orange (colour intensity indicates fold change in the gene expression) for cellular response after **A.** 6 h (cDNA and dRNA) and **B**. 24 h (cDNA). The first steps in the pathway are omitted (in panel A pathway is represented from farnesyl pyrophosphate conversion and in panel B – lanosterol).

applied method and time of fermentation. Ergosterol is the primary component of yeast plasma membrane which defines fluidity and permeability, and the defective biosynthesis pathway of this sterol affects cellular processes such as cell wall synthesis, endocytosis, cell polarization and fusion[120,232]. Farnesyl pyrophosphate (FPP) is the precursor for ergosterol biosynthesis which is produced in the mevalonate pathway and it is also utilized by β-carotene production. The shared resources between native ergosterol biosynthesis and carotenogenesis might cause a competition between these pathways and impose additional burden if building blocks for cellular growth are limiting. The analysis of results generated from cDNA samples identified upregulation of six *ERG* genes during exponential growth phase (*ERG1*, *ERG3*, *ERG5*, *ERG11*, *ERG24*, *ERG25* and *NCP1*, Figure 7-2A) and upregulation of five genes when cells entered stationary phase (*ERG3*, *ERG4*, *ERG5*, *ERG6*, *ERG25*, Figure 7-2B). Notably, genes *ERG1*, *ERG11* and *ERG24*

upregulated only in the initial phase of cultivation encode enzymes which catalyse first steps in ergosterol biosynthesis (Figure 7-2). In contrast, genes highly transcribed after 24 h *ERG4* and *ERG6* are responsible for the last steps of the ergosterol formation. Ergosterol biosynthesis requires iron atom for the action of diiron proteins ERG3 and ERG25, and heme which serves as a cofactor for ERG11 and ERG5[120]. Under iron deficiency transcription of *ERG1* and *ERG11* is decreased while abundance of *ERG25* and *ERG3* transcripts rises[233]. Therefore, the difference in *ERG* genes upregulated at the beginning and the end of the fermentation may be caused by depletion of iron. This hypothesis is supported by higher expression of genes induced by iron deficiency: *FRE4* (encodes ferric reductase, 4-fold increase in transcripts abundance from 6 h to 24 h determined by cDNA-seq), *FIT2* (iron regulon, 2.2-fold increase) and *HAP1* (heme-dependent transcriptional activator, 4.8-fold increase).

The RNA-seq results demonstrated increased expression of two transporters mediating export of toxic compounds: PDR5 and SNQ2. One of the first studies evaluating impact of β-carotene production on *S. cerevisiae* reported upregulated expression of the pleiotropic drug resistance transporters which facilitate detoxification of a cell[125]. In the same publication[125], PDR transporters were suggested to secrete carotenoids if a hydrophobic solvent was present in the growth medium. Using this principle Bu *et al.*[234] enhanced β-carotene secretion from yeast cells by engineering ATP-binding cassette (ABC) transporters, including SNQ2. Finally, all the tested samples consistently indicated upregulation of *FIT3* encoding a cell wall mannoprotein facilitating iron transport[235] and *YEH1* genes encoding steryl ester hydrolase involved in homeostasis of sterols[236].

Selection of promoters for biosensing production stress

With the RNA-seq data in hand, we sought to develop a biosensor suitable for monitoring production burden by selecting genes with a strong response to carotenogenesis and isolating promoters of those genes. A promoter activated by transcription factors involved in the stress response to β-carotene could be then coupled to a reporter gene for a straightforward monitoring of a burden or used for expression of a gRNA which together with the CRISPRi, system to create a feedback control system. The abovementioned results showed upregulation of genes involved in ergosterol biosynthesis and pleiotropic drug resistance, therefore promoters of genes *ERG3*, *ERG5*, *ERG25*, *PDR5* and *SNQ2* were further explored as a potential biosensors.

The highest upregulation of 6 – 8.7 fold change was noted for *ERG3* gene (Figure 7-3, SI Table 7-3). *ERG5* initially displayed 5.0 fold upregulation which decreased to 3.2 after 24 h of fermentation. Activation of ABC transporters PDR5 and SNQ2 was observed to the lesser extent in comparison with *ERG* genes with a 2.5-3.3 fold change in *PDR5* transcripts and 1.1-3.3 for *SNQ2*. To diversity the set of tested promoters, we selected genes with different transcript range based on normalized values (transcripts per million, TPM) (Table 7-2). The record TPM was noted for *ERG3*, therefore *ERG25* with a similar characteristics was excluded from the analysis. Although the number of transcript of *ERG5* was 18-27 times lower, high

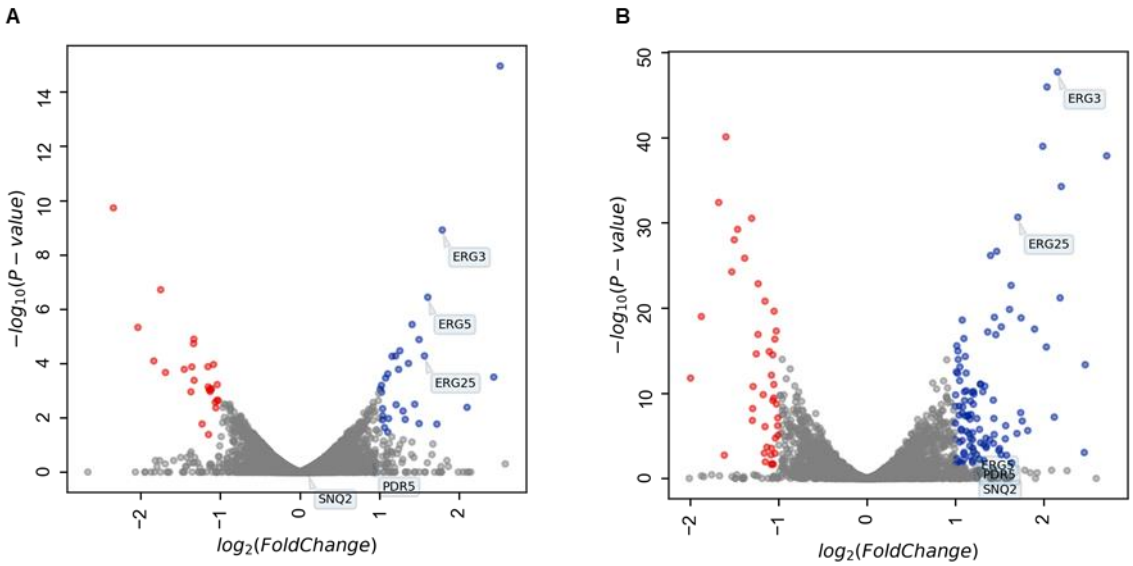


Figure 7-3. Volcano plots illustrating gene upregulation induced by production of carotenoids after A. 6 h and **B**. 24 h of cultivation. Data obtained from the analysis of cDNA library. The following criterium for applied to indicate genes with significant upregulation (blue): $log_2FC \geq 1$ and p-value<0.05 whereas for downregulation (red): $log_2FC \leq -1$ and p-value<0.05.

Table 7-2. Normalized number of transcripts (TPM) measured in carotenoids producing strain CAR-034 and wild type strain. Values indicate mean ± 1 standard deviation (n=2).

Gene	cDNA 6h		cDNA 24h	
	CAR-034	WT	CAR-034	WT
ERG3	1632±6	450±57	1185±128	254±3
ERG5	93±3	29±2	44±1	18±0
ERG25	1361±120	434±172	841±90	246±11
PDR5	3±0	2±0	31±2	13±0
SNQ2	15±1	6±1	39±3	16±2

upregulation was observed in the carotenoids producing strain compared to the wild type strain. Due to the previous correlation between β-carotene and activation of the pleiotropic drug resistance response, we decided to include *PDR5*. Altogether, evaluation of promoters potentially suitable for biosensing was limited to genes *ERG3*, *ERG5* and *PDR5*.

Activation of the *PDR5* promoter (PDR5p) enhanced by transcription factors mediating response to toxic compounds in yeast was employed for development of a biosensor for a drug diclofenac[237]. This sensing device was constructed using different lengths of region upstream *PDR5* gene (305–1551 bp) and a fluorescent protein as a reporter (mCherry). A good activation of the biosensor upon exposure to the inducer was observed when 563 bp of 5' URS of *PDR5* was used, without an improvement for a longer promoter region. Due to a known binding site of PDR1/3 transcription factors exactly at the position 563 bp of 5' URS of *PDR5*, we used a longer promoter region of 600 bp.

Development of biosensors with a similar principle based on *ERG3* or *ERG5* promoters (ERG3p, ERG5p, respectively) has not been reported, therefore two different lengths of 5' URS region upstream *ERG3* and *ERG5* genes were tested. Activity of the ERG3p was analysed using β-galactosidase enzyme assay which showed the highest effectiveness of 373 bp long promoter in comparison with the full length promoter (891 bp)[238]. The same assay was applied in a study investigating transcriptional regulation of ergosterol synthesis, using 900 bp of the ERG3p. Analysis of Hap1 binding sites in the ERG5p was performed for the full length (750 bp) non-coding region between the *SOK2* and *ERG5* genes[239], but use of a shorter ERG5p (449 bp) was also reported[240]. Based on these reports, a 900 bp and 373 bp *ERG3* promoter was tested, whereas for ERG5p a 750 bp and 449 bp upstream region was used.

To evaluate activity of selected promoters and compare their strength, a set of carotenogenic strains was created with genomically integrated expression cassette containing a promoter of interest, *RFP* (*mCherry*) and native *TDH3* terminator (Figure 7-4A). Promoter regions were amplified from genomic DNA of the wild type yeast strain CEN.PK113-7D which was used as the parental strain for carotenoids producing strains, including the strain analysed by RNA-seq. Integration of expression cassettes allowed to minimize variation in gene copy number which occurs in plasmid borne systems. Fluorescence of strains harbouring the biosensing circuit was measured with a flow cytometer after cultures were incubated 24 h starting with equivalent inoculum. ERG3p was active solely when a 900 bp upstream region was used and no RFP fluorescence was detected for 373 bp promoter (Figure 7-4B). Both of the tested *ERG5* promoters exhibited activity, however using a 449 bp promoter fluorescence was decreased by 1.9 fold which may be caused by lack of required transcription factors binding sites or another essential promoter elements. Along with the *ERG3* promoter of 900 bp, the *PDR5* promoter exhibited high activity, which was unexpected based on the RNA-seq data (Table 7-2). Furthermore, RFP fluorescence was not detected for the carotenogenic strain CAR-034 (without any fluorescent reporter) which confirms lack of interference between carotenoids and reporter system (RFP).

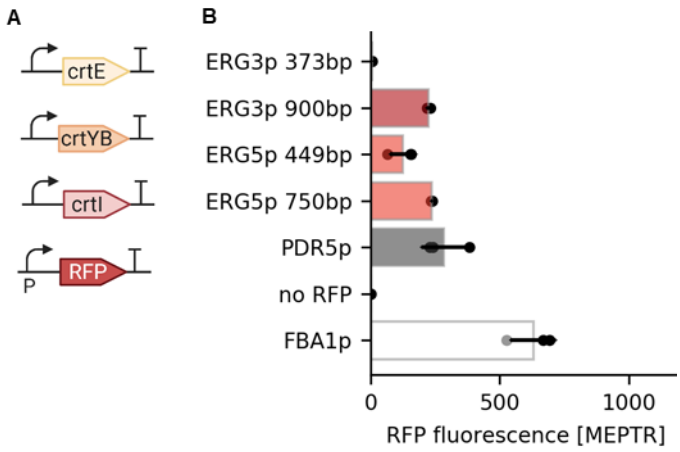


Figure 7-4. Activity test for promoters with biosensing capacity. A. A set of promoters was assayed using RFP fluorescence as a direct readout of activity. The expression cassette harbouring the promoter of interest and *RFP* was integrated into the genome of carotenogenic strain CAR-034. P indicates promoter of interest, *i.e.* ERG3p, ERG5p or PDR5p. **B**. RFP fluorescence upon expression from tested promoters. Carotenogenic strain without RFP (no RFP) and RFP driven by FBA1p were used as controls. Bars represent mean ±1 standard deviation (n=3).

gRNA expression from promoters acting as a biosensor

We sought to evaluate whether ERG3p and PDR5p can provide sufficient expression of a gRNA which would lead to effective repression of *RFP* transcription mediated by dCas12a. It was next anticipated to implement the biosensing promoter for expression of gRNA in the feedback control system dynamically regulating production of carotenoids, therefore data for RFP fluorescence repression would provide valuable insights into the gRNA expression system. The carotenogenic strain CAR-035 (Chapter 4) with moderately expressed *crt* genes was engineered to contain dCas12a E925A NLS Mxi1 and *RFP* (*mCherry*) driven by *ENO1* promoter from *K. lactis* resulting in strain CAR-047 (Figure 7-5A). Selection of strain CAR-035 over CAR-034 allowed to downregulate *RFP* by targeting ENO1p without affecting β-carotene production (in strain CAR-034 expression of *crtI* was driven by the same promoter). gRNA expression cassettes contained either the 900 bp-ERG3p or the 600 bp-PDR5p and one of the three different spacers: G34, G37 or G42. Spacer G34 and G37 encoded the ENO1p as dCas12a target and were previously assayed for repression of *eGFP* displaying 32- and 7.6-fold fold repression when expressed from the *SNR52* promoter (Chapter 5). G42 was a non-targeting control used in combination with a ERG3p, PDR5p or SNR52p. As expected, downregulation of *RFP* was observed upon expression of gRNAs targeting the *ENO1* promoter in the *RFP* construct when expressed using ERG3p and PDR5p (Figure 7-5B). A similar repression was noted for G34 achieving 4.1±0.3 and 4.7±0.6 fold decrease in RFP fluorescence (values normalized by SNR52p G42) expressed from ERG3p and PDR5p, respectively. Improved efficiency was observed for G37 driven from PDR5p (2.3±0.3 fold repression) compared to ERG3p (1.7±0.1). Overall, these results show effective gene downregulation for gRNAs flanked by self-cleaving ribozyme sites and driven by RNAP II promoters *ERG3* and *ERG5* and that the selected promoters exhibited sufficient transcription of gRNA in a carotenogenic strain.

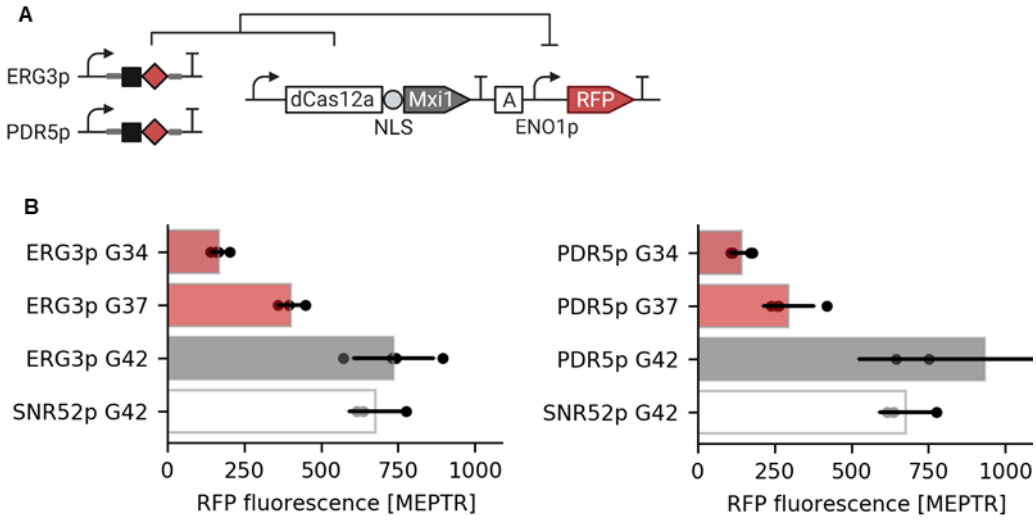


Figure 7-5. CRISPR interference of *RFP* with gRNAs driven by *ERG3* and *PDR5* promoters. A. ERG3p or PDR5p drive expression of a ribozyme-flanked gRNA encoding a target on *ENO1* promoter which controls *RFP* (*mCherry*). Targeting gRNA guides dCas12a E925A NLS Mxi1 fusion protein to the genomic locus and inhibits expression of *RFP*. Assay was conducted in carotenoids producing strain CAR-047. **B**. Repression efficiency of CRISPRi in combination with ERG3p controlled gRNAs (left panel) and PDR5p gRNAs (right panel). gRNAs G34 and G37 target ENO1p and G42 is a non-targeting control. SNR52p G42 was used as an additional control. Bars represent mean ±1 standard deviation (n=3).

Next, we sought to evaluate tunability of biosensing promoters, *i.e.* whether gene expression driven by ERG3p and PDR5p is upregulated with increased production of carotenoids and consequently, higher stress level. Such a feature would be desired for a biosensor coupled to the feedback control system and it would allow to effectively repress the production as the burden imposed on cell fitness. Three strains were selected for the experiment: carotenoids producing strains CAR-035 and CAR-036 (Chapter 4) which generate moderate and low amounts of carotenoids (respectively: 11.1 and 5.3 µg/OD after 48 incubation, Chapter 4), and wild type strain. To enable assay based on CRISPRi mediated downregulation, these strains were engineered to contain a single copy of *RFP* (*mRuby2*) in the genome resulting in strains CAR-048, CAR-049 and FR016. Subsequently these strains were provided with plasmid pCSN075 encoding dCas12a E925A NLS Mxi1 (Figure 7-6A). Quantification of RFP fluorescence revealed a high level of variability between biological replicates (Figure 7-6B). Such performance was unexpected based on the experiments conducted for strain CAR-047 with genomically integrated dCas12a as well as characterization of dCas12a which made a use of plasmid pCSN075 (Chapter 5). Furthermore, good reproducibility was achieved in previous evaluation of the sensing promoters, when *mCherry* was used as a target of CRISPRi. Due to the high variation between tested samples conclusion cannot be drawn. An optimal biosensor should be tuneable and responsive to the burden *i.e.* exhibiting minimal basal activity in the wild type strain which is boosted in the carotenoids producing strain. Therefore, it was expected to observe enhanced downregulation of *RFP* in carotenogenic strains CAR-048 and CAR-049 than in FR015, and the stronger *RFP* repression in strain CAR-048 compared to CAR-049.

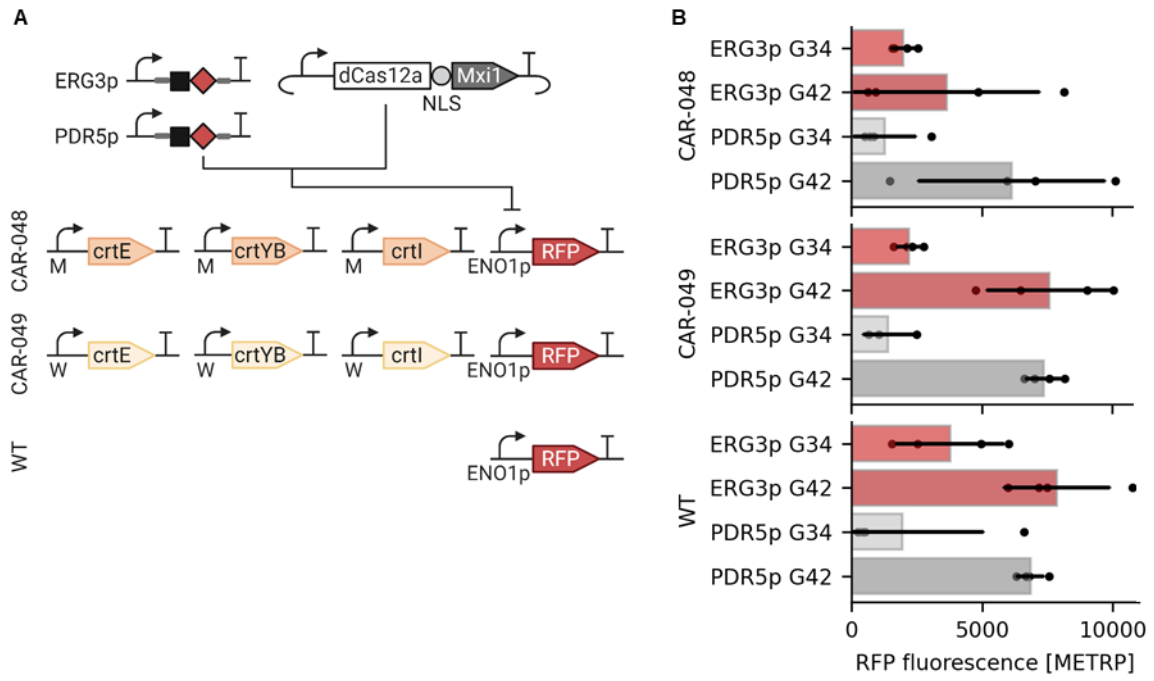


Figure 7-6. Tunability of biosensing promoters evaluated by CRISPRi mediated downregulation of *RFP*. A. gRNA expressed from ERG3p or PDR5p and dCas12a E925A NLS Mxi1 fusion protein mediates downregulation of *RFP* (*mRuby2*) in strains with different production level of β-carotene. *crt* genes in strains CAR-048 and CAR-049 were expressed from medium strength (M) or weak promoters (W) while *RFP* is expressed from ENO1p from *K. lactis*. **B**. RFP fluorescence upon downregulation with CRISPRi and ERG3p or PDR5p gRNAs in strains CAR-048, CAR-049 and FR016 with genomically integrated *RFP* expression cassette. Bars represent mean ±1 standard deviation (n=4).

7.4 Discussion

The differential expression analysis of β-carotene producing strain performed in this study revealed upregulation of genes involved in biosynthesis of ergosterol and ABC transporters involved in pleiotropic drug resistance response. Modulating expression of genes responsible for ergosterol biosynthesis and deletion of related transcriptional factors is a common method in optimization of carotenoids production due to the shared precursor between these pathways[33,94,241]. Activation of the pleiotropic drug resistance response was previously linked to the production of β-carotene[125,234] as a cell's response to accumulation of carotenoids and attempt to secrete carotenoids from the cell.

Determination of genes with increased transcription triggered by production of carotenoids allowed to select a set of promoters which could potentially sense burden related to the production. It was crucial that the sensing capacity is independent from the growth phase and available carbon source, therefore promoters were selected based on transcriptomic data obtained at different time points of the cultivation. We have identified promoters of ergosterol biosynthesis genes (*ERG3*, *ERG5*) and the pleiotropic drug resistance transporter gene *PDR5*, and showed their activity when controlling expression of a fluorescent protein.

The main objective of this thesis was to develop a feedback control system which could mitigate production burden based on a dynamically activated CRISPRi system. CRISPRi is a modular system which requires two elements, mainly nuclease-dead Cas12a protein and gRNA which gives an opportunity to control activity of the system by modulating expression of these two elements. We reasoned that the transcription of gRNA would activate CRISPRi system faster than expression of dCas12a. Therefore, we examined whether ERG3p and PDR5p can efficiently express a gRNA (and not dCas12a) using CRISPRi mediated downregulation of *RFP* as a readout. To achieve good expression for both of the tested promoters, we examined whether their activity is responsive to increased production of carotenoids, resulting in poorly reproducible data. The major differences between the system used to test possibility of expressing gRNA from ERG3p and PDR5p and experiment evaluating promoter responsiveness to the increased carotenoids production were:

1. use of *mRuby2* instead of *mCherry* as a target gene, due to a difficulty in strain construction, and
2. expression of dCas12a integrated into the genome versus expression from a low-copy plasmid.

These differences in the design could have caused poor reproducibility of the results. The maturation time of mRuby2 is longer than for mCherry (40 and 150 min, respectively), therefore the experiment should be repeated for the set of strains with *mCherry* as a target gene and possibly cell fixation prior to the measurement (none was applied here)[242]. Furthermore, a low copy number centromeric plasmid was used for expression of dCas12a, however the plasmid copy number in cell may vary, thus integration could be advantageous[243].

Development of a high throughput method for screening a larger number of promoters might lead to discovery of promoters with more attractive biosensing features. A high throughput method for screening native *E. coli* promoters as a phenylalanine biosensor was developed by Mahr *et al.*[44]. In this approach a pooled library of expression cassettes containing various promoters and a reporter gene was expressed in *E. coli* cultivated in cycles in presence or absence of phenylalanine and selected using fluorescence activated

cell sorting (FACS). Using a *S. cerevisiae* strain with tuneable production of carotenoids an analogous approach could be applied, simplifying the testing set-up and expanding the range of tested promoters.

Overall, in the presented work a promoter sensing cellular stress related to the production of β-carotene was determined based on transcriptomics data. Selected promoters ERG3p and PDR5p produced a functional gRNA which allows to couple the stress sensing circuit with a regulatory CRISPRi system to ultimately create a feedback control system in carotenoids producing strains (Chapter 8).

8 Feedback control system

This Chapter summarizes development of the feedback control system by coupling CRISPRi and burden sensor described in the previous Chapters. The growth experiment was performed with the support of René Verwaal (DSM). Sandra Pous Torres (DSM) operated UHPLC for carotenoids quantification. Remaining experiments were conducted by me under supervision of René Verwaal, Thomas E. Gorochowski (University of Bristol) and Johannes A. Roubos (DSM).

8.1 Introduction

The development of synthetic biomolecular control systems for balancing metabolic pathways is inspired by natural autoregulatory mechanisms, which allow the cell to allocate metabolic and energetic resources to produce a required metabolite only to necessary levels and prevent excessive biosynthesis. In one of the first examples of developing regulatory circuits to alleviate burden, Farmer *et al.*[124] tackled the growth defect of lycopene producing *E. coli* strain by implementing a system based on a transcription factor which modulates expression of genes involved in the production pathway depending on glucose level. With the advances in CRISPR technology and development of the gRNA programmable CRISPR systems for regulating gene expression, new types of the regulatory loop for balancing production pathways have been proposed. Ceroni *et al.*[24] linked a nuclease-defective dCas9 with a promoter monitoring cell fitness, used for control of gRNA expression. Once the production triggered a stress response, the biosensor-acting promoter was activated and produced gRNA transcripts which upon formation of complex with dCas9 bound genes involved in the production, thus blocking the transcription. Both of the aforementioned systems improved production while minimizing burden.

Our characterization of carotenoids producing strains revealed hampered growth and genetic instability of the production pathway as exemplified by a loss-of-function mutations (Chapter 2). Here, we coupled our developed dCas12-CRISPR interference system with a dynamically expressed gRNA to create a feedback control system to control β-carotene biosynthesis in *S. cerevisiae*. Implementation of the feedback control system based on dCas12a allowed us to target T-rich promoter regions due to the PAM preference (5'-TTV-3') and for a simultaneous multiplex downregulation by a simple use of a single crRNA array without a need for co-expression of additional elements. We have previously established and elucidated the design rules for a functional dCas12a CRISPRi system in *S. cerevisiae* (Chapter 5). Strong regulation was achieved using dCas12a fused to a single C-terminal NLS to ensure protein localization in the nucleus. To enhance downregulation, the Mxi1 repression domain was fused to dCas12a, leading to an improvement in *eGFP* repression (8- and 12.6-fold repression for dCas12a and dCas12a-Mxi1, respectively). Nuclease deficiency in dCas12a can be achieved by mutating the RuvC domain – either D832A or E925A. Repression of *eGFP* expression with dCas12a E925A NLS Mxi1 was slightly improved in comparison with the construct harbouring a D832A mutation (up to 2.3-fold improvement for gRNAs targeting promoter).

To provide a dynamic activation of the CRISPRi system upon carotenoids induced stress response, we performed transcriptional analysis to elucidate genes with regulation sensitive to carotenogenesis (Chapter 7). This analysis revealed a 4.4-fold upregulation (carotenogenic strain versus wild type) of the C-5 sterol desaturase gene encoded by *ERG3*, that is involved in ergosterol biosynthesis which competes with

carotenogenesis for the precursor (farnesyl diphosphate). Numerous publications[125,234,244] reported activation of the pleiotropic drug resistance (PDR) upon production of carotenoids in yeast which has also been confirmed by our analysis indicting 3-fold upregulation of *PDR5* gene encoding an ABC transporter protein responsible for cell detoxification. Therefore, promoters of *ERG3* and *PDR5* were used to enable expression of gRNA and to couple activation of the CRISPRi system upon the cell stress response. To enable gRNA expression from RNAP II promoters and production of a pre-crRNA transcript which can be processed by dCas12a, gRNAs were flanked by self-cleaving ribozymes (Chapter 5). Such design prevents gRNA from post-transcriptional modification and export from the nucleus[57]. gRNA spacers targeted the Kl_ENO1 promoter selected for expression of *crtI* encoding phytoene desaturase, which was determined as a bottleneck step in β-carotene formation. As shown in Chapter 4, selection of a weak promoter for expression of *crtI* decreased β-carotene formation by 6.7-fold in comparison with the strain with strong expression of all *crt* genes. Expression one of the remaining *crt* genes *i.e. crtE* and *crtYB* from a weak promoter did not reduce formation of β-carotene. Finally, our feedback control system for a dynamic regulation of β-carotene biosynthesis was introduced into a strain with genomically integrated production pathway (Figure 8-1A), comprised of dCas12a E925A NLS Mxi1 and gRNA expressed from *ERG3* or *PDR5* promoters, sensitive to carotenogenesis. Upon expression of a gRNA, a dCas12a-gRNA complex is formed which subsequently binds Kl_ENO1 promoter and disables *crtI* expression, thus blocking β-carotene formation. The created carotenoids producing strains with the feedback control system was evaluated in terms of carotenoids production, microbial growth and genetic stability of the production pathway during a batch fermentation.

8.2 Materials and methods

Strains

Carotenogenic strains used in this study were constructed using CRISPR-Cas9 or Cas12a to enable integration of *crtE*, *crtYB* and *crtI* expression cassettes into the genome, as described in the previous chapters (Table 8-1). All strains expressed *crt* genes from strong constitutive promoters, with a difference in the use of native *S. cerevisiae* promoters in strain CAR-002 for expression of *crtE* and *crtYB* and three different heterologous promoters in strain CAR-034. Strain CAR-034 was further engineered to contain dCas12a E925A NLS Mxi1 expression cassette in the genome, resulting in strain CAR-042. As controls in growth experiment strain CAR-034 and wild type strain (CEN.PK113-7D) were transformed with an empty gRNA recipient plasmid (pRN1120) to allow for cultivation in presence of nourseothricin (Table 8-2).

Table 8-1. Genotype of strains tested with the feedback control system

Strain	Genotype			Chapter
CAR-002	Sc_FBA1p - crtE - Sc_TDH3t,	Sc_TEF1p - crtYB - Sc_PDC1t,	Kl_ENO1p - crtI - Sc_TAL1t	2
CAR-034	Sbay_TDH3p - crtE - Sc_TDH3t,	Smik_TEF1p - crtYB - Sc_PDC1t,	Kl_ENO1p - crtI - Sc_TAL1t	4
CAR-042	Sbay_TDH3p - crtE - Sc_TDH3t, Sc_PGI1p - dLbCas12a E925A - NLS – Mxi1 - GND2t	Smik_TEF1p - crtYB - Sc_PDC1t,	Kl_ENO1p - crtI - Sc_TAL1t,	5

Cultivation

For yeast transformation rich medium composed of yeast extract (1%, Bacto™ Becton–Dickinson, Franklin Lakes, NJ, USA), phytone peptone (2% Difco™ BD) and D-glucose (2%, Sigma Aldrich, St Louis, MO, USA) was used (YEPD). The medium was supplied with geneticin (G418, 200 µg/mL) or nourseothricin (NTC, 200 µg/mL) when plasmid selection was required (pCSN075 was selected on G418 and pRN1120 – NTC). Growth experiments were conducted using rich medium or mineral medium, which contained potassium phosphate monobasic (3%, Sigma Aldrich), magnesium sulphate heptahydrate (0.5%, Merck, Darmstadt, Germany), urea (2.3%, Merck), glucose (4%, Sigma Aldrich), vitamins, and trace elements[127]. To enable plasmid selection in mineral medium, the concentration of NTC was decreased to 50 µg/mL as higher NTC concentration in combination with the mineral medium hampers growth of the yeast. Pre-cultures were set up in 100 mL shake flasks filled with 20 mL of rich medium containing NTC (200 µg/mL) and grown for 48 h at 30 ºC with 250 rpm shaking. Subsequently, pre-cultures were diluted to OD_{600} of 0.1 in 100 mL of either rich of mineral medium supplemented with NTC in 0.5 L shake flasks and cultures were incubated at 30 ºC with 250 rpm shaking. Fermentation in rich medium was conducted for 3 days, whereas in mineral medium for 5 days. Growth was evaluated by optical density measured at 600 nm and dry weight measurement. To quantify dry weight ~ 7.5 g of medium was collected in pre-dried tube (VWR, Radnor, PA, USA) and separated from the spent medium (5300 rpm, 10 min). Cell pellet was washed once in milliQ water and dried at 104 °C for 24 h. The growth experiment was conducted in biological triplicates.

Molecular biology techniques

Transformation of *S. cerevisiae* was performed according to Gietz *et al.*[148] (Chapter 3). Briefly, the yeast culture was grown to OD_{600} of 2.5-5.0 and harvested pellet was resuspended in a solution of lithium acetate dihydrate (LiAc, 0.1 M, Sigma Aldrich) and tris hydrochloride – EDTA buffer (TE, 10 mM). Subsequently, deoxyribonucleic acid sodium salt from salmon testes (ssDNA, 5 µL, Sigma Aldrich) was added and the cell suspension was mixed with all DNA parts. The following DNA parts were used: plasmid pCSN075 (100 ng), gRNA expression cassette (1000 ng) and (linearized) gRNA recipient plasmid pRN1120 (100 ng). The cell-DNA suspension (20 µL) was mixed with polyethylene glycol 4000 solution (60 µL, PEG, 50% Merck, Darmstadt, Germany) containing PEG (50%), LiAc (0.1 M) and TE (10 mM) and the mix was incubated at room temperature for 30 min. In the following incubation, DMSO (100%, 7 µL) was added and cells were subjected to heat shock at 42 ºC for 1 h. Cells were recovered in non-selective rich medium overnight at 20 ºC and plated on rich agar supplemented with G418 and/or NTC.

gRNA expression cassettes containing the *ERG3* or *PDR5* promoter, direct repeat, spacer and *GND2* terminator were obtained as synthetic DNA from IDT (Leuven, Belgium) and amplified by PCR using primers pKC001,2 (IDT, SI Table 8-1). The gRNA recipient plasmid pRN1120 was linearized by PCR with primers pKC003,4. All PCR reactions were conducted with Q5 polymerase (New England Biolabs, Ipswich, MA, USA) and PCR products were purified with the Wizard SV gel and PCR clean up kit (Promega, Madison, WI, USA).

Table 8-2. Plasmids used in Chapter 8

Plasmid	Genotype	Source	Addgene
pCSN075	*CEN6/ARS TRP1* KanMX bla TEF1p-dLbCas12a E925A-NLS-Mxi1-GND2t	Chapter 5	# 166732
pRN1120	2µm NatMX amp^R	[65]	# 101750

Carotenoids quantification

Quantification of phytoene, lycopene and β-carotene was performed with a UHPLC-DAD system according to the previously published method[181]. In line with this method, cell cultures corresponding to 1mL of OD_{600}=20 was harvested by centrifugation (13,000 rpm, 8 min) and separated from spent medium. The cell pellet was resuspended in 0.5 mL of tetrahydrofuran (Merck) and homogenized with Precellys® (Bertin, Montigny-le-Bretonneux, France). The homogenization program comprised of two 20 second cycles at 6,800 rpm and 30 second pause in between. The organic phase with carotenoids was collected after centrifugation (13,000 rpm, 8 min, 4 °C) and subjected to analysis with UHPLC. The UHPLC system contained UHPLC Ultimate 3000 system (Thermo Fisher Scientific, Waltham, MA, USA), a Waters XBridge® C18 column (3.5 µm, 2.1mm x 50 mm, Milford, MA, USA) and Ultimate 3000 photodiode array detector (Thermo Fisher Scientific). Separation on the column was achieved by a gradient of mobile phases (ethyl acetate, water and acetonitrile, SI Table 2-2). Different wavelengths were used for detection of carotenoids; for phytoene – 286 nm, lycopene – 475 nm and β-carotene 450 nm. Data was collected and processed with Chromeleon software.

8.3 Results

Construction of the feedback control system

We selected two strains to test the feedback control: CAR-002 and CAR-034. Strain CAR-002 contained *crtE* and *crtYB* expression cassettes with native *S. cerevisiae* promoters and *crtI* driven by the heterologous *ENO1* promoter from *K. lactis.* Therefore, in strain CAR-002 the feedback control system can be used to downregulate solely *crtI* without a threat of interfering with the expression of native pathways. Cultivation of CAR-002 in mineral medium improved β-carotene yield but affected pathway stability and led to loss-of-function mutations disabling the production (Chapter 2). We sought to assess if the genetic instability of strain CAR-002 can be improved upon introduction of the feedback control system. dCas12a was provided on plasmid pCSN075 (a single copy plasmid expressing dCas12a E925A fused to NLS and Mxi1 from the *TEF1* promoter, Chapter 5, Table 8-2) rather than integrated into the genome, since a potential improvement in growth was not evaluated for strain CAR-002 (the selection pressure required to maintain the plasmid pCSN075 is expected to affect growth). Due to the inability to downregulate the entire β-carotene biosynthesis pathway, it was not foreseen to further optimize the feedback control system in strain CAR-002. In strain CAR-034 genes *crtE*, *crtYB* and *crtI* were expressed from three different heterologous promoters which allows for downregulation of the full pathway. We have previously assessed toxicity of dCas12a (Chapter 5) with observed formation of small colonies when dCas12a expression was driven by a strong constitutive promoter (TEF1p). The feedback control system should mitigate burden related to a

heterologous compounds production, therefore the system itself should not impose an additional energetic or metabolic load on the host organism. To overcome mild toxicity exhibited by dCas12a, we substituted the *TEF1* promoter with the weaker constitutive promoter *PGI1*. Carotenogenic strain CAR-034 was further engineered to integrate the PGI1p dCas12a expression cassette into the genome resulting in strain CAR-042 with observed restored growth (Chapter 5).

Next, we transformed strain CAR-042 with gRNAs G34, G37 targeting *ENO1* promoter controlling *crtI* and the non-targeting control G42 expressed either from PDR5p or ERG3p, resulting in colonies with visible decrease in carotenoids production (Figure 8-1B). Downregulation mediated by PDR5p G34 caused a change in colonies colour from dark orange (PDR5p G42 non-targeting control) to pale yellow whereas expression of PDR5p G37 resulted in formation of yellow transformants. A change in colonies colour was also visible for gRNAs expressed from the ERG3p, however not to the extend observed for PDR5p gRNAs. Expression of ERG3p G34 led to appearance of yellow colonies whereas expression of ERG3p G37 decreased the intensity of the orange colour in the transformants. The difference in gRNAs efficiency may be linked to the strength and activation of the *PDR5* and *ERG3* promoters. Spacers of gRNAs G34 and G37 encoded different regions of Kl_ENO1 promoter and were previously evaluated to downregulate *eGFP* expression controlled by the Kl_ENO1 promoter. In this case gRNAs G34 and G37 were driven by RNAP III promoter (SNR52p) and resulted in a 32- and 7.6-fold decrease in eGFP fluorescence (Chapter 5), which is in agreement with a stronger repression of carotenogenesis by G34 observed here.

Finally, we tested β-carotene downregulation in strain CAR-002 using the same set of gRNAs driven by ERG3p and pCSN075 plasmid. A change in colour of the transformants was not observed which might be related to the expression system selected for dCas12a (Figure 8-1C). A comparison of dCas12a expressed genomically and from a plasmid revealed ~3-times more effective downregulation of the genomically integrated dCas12a. Integration of the dCas12a into the genome also removed the requirement of plasmid selection, which may additionally affect growth. In strain CAR-042 we optimized expression level of dCas12a by placing it under control of the *PGI1* promoter while plasmid pCSN075 applied for *crtI*

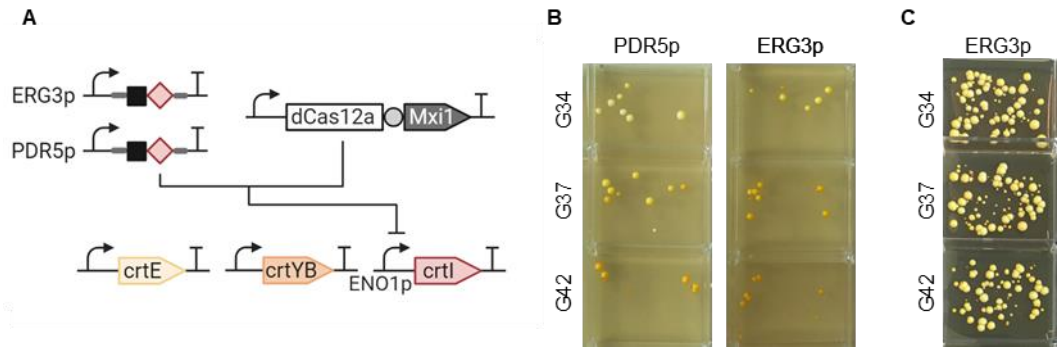


Figure 8-1. CRISPRi-based feedback control system for carotenoids downregulation in *S. cerevisiae*. A. Principle of action of the feedback control system. gRNA expression is activated by stress related to the production of β-carotene. gRNA targets dCas12a to the ENO1p in front of *crtI* and downregulates the gene expression. **B**. Downregulation of carotenoids production in strain CAR-042 with a genomically integrated dCas12a E925A NLS Mxi1 driven by PGI1p and plasmid-borne gRNAs expressed from PDR5p and ERG3p. **C**. Downregulation of carotenoids in strain CAR-002 with plasmid pCSN075 (expresses dCas12a E925A NLS Mxi1 from TEF1p) and plasmid-borne gRNAs expressed from the ERG3p. gRNAs G34 and G37 target Kl_ENO1p driving *crtI* expression; gRNA G42 in the non-targeting control.

downregulation in CAR-002 was expressed from the *TEF1* promoter. Although the *TEF1* promoter has a higher strength, repression of carotenoids production was not visually noted for CAR-002.

Influence of the feedback control system on performance of carotenogenic strains

Next, we sought to evaluate impact of the feedback control system on the growth of carotenoids producing strain CAR-042, pathway stability and confirm the visual trait of the carotenoids production with quantified production data. A batch fermentation was conducted in rich and mineral medium supplemented with NTC to select for the presence of the plasmid expressing the gRNA. For the strain harbouring the feedback control system, gRNAs expressed from PDR5p were selected due to a stronger repression evaluated based on

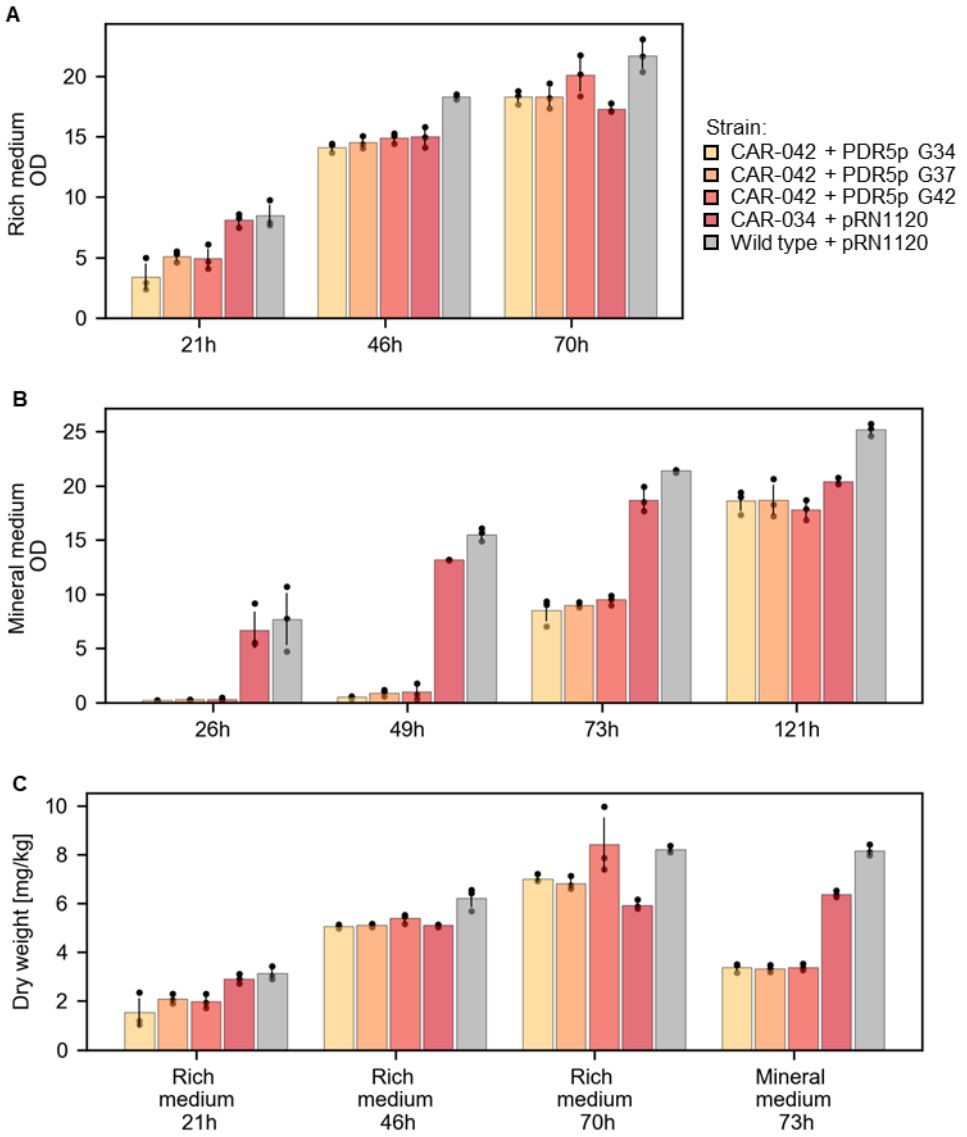


Figure 8-2. Growth experiment evaluating feedback control system. A. Optical density for cultures grown in rich medium and **B**. mineral medium. **C**. Dry weight for cultures grown in both types of medium. Strain CAR-042 – carotenogenic strain with the feedback control system; strain CAR-034 – carotenogenic strain with plasmid pRN1120 to allow cultivation in medium with NTC; WT – wild type yeast strain with pRN1120. gRNAs G34 and G37 – targeting promoter controlling *crtI* expression and gRNA G42 – non-targeting control; all gRNAs were expressed from RNAP II PDR5p. Bars represent mean ± 1 standard deviation (n=3).

transformants colour (Figure 8-1B). As a reference, carotenoids producing strain CAR-034 and a wild type carrying an empty gRNA recipient plasmid (pRN1120) were used to enable growth of all tested strains under the same conditions. Growth was evaluated by measurement of optical density and dry weight and both methods indicated hampered growth of strain with the feedback control system (CAR-042 and PDR5p G34, G37, G42) in comparison with the unmodified carotenogenic strain (CAR-034) after 21 h of cultivation in rich medium (Figure 8-2). Hampered growth of the strain with the feedback control system and dummy gRNA (PDR5p G42) may indicate additional load related to the expression of a gRNA and/or dCas12a in the carotenogenic strain CAR-042. After an additional day of incubation (46 h) the difference in growth between strains CAR-042 with gRNA and CAR-034 was diminished and all carotenogenic strains reached equal OD_{600} (14.6±0.3) and dry weight (5.2±0.1 mg/kg). The carotenogenic strain with the incorporated feedback control system (CAR-042) exhibited slightly higher OD and dry weight than unmodified strain CAR-034 when cultured in rich medium for 70 h. In the mineral medium a severe growth defect was observed for strain CAR-042 expressing any gRNA, reaching OD of only 0.3 and 0.8 after 26 and 49 h of incubation, respectively. In comparison, optical density measured for carotenogenic strain CAR-034 was 7 and 13 for the same timepoints. The growth of the feedback control strains in the mineral medium was restored after 73 h and eventually after 5 days of incubation achieved ODs similar to CAR-034.

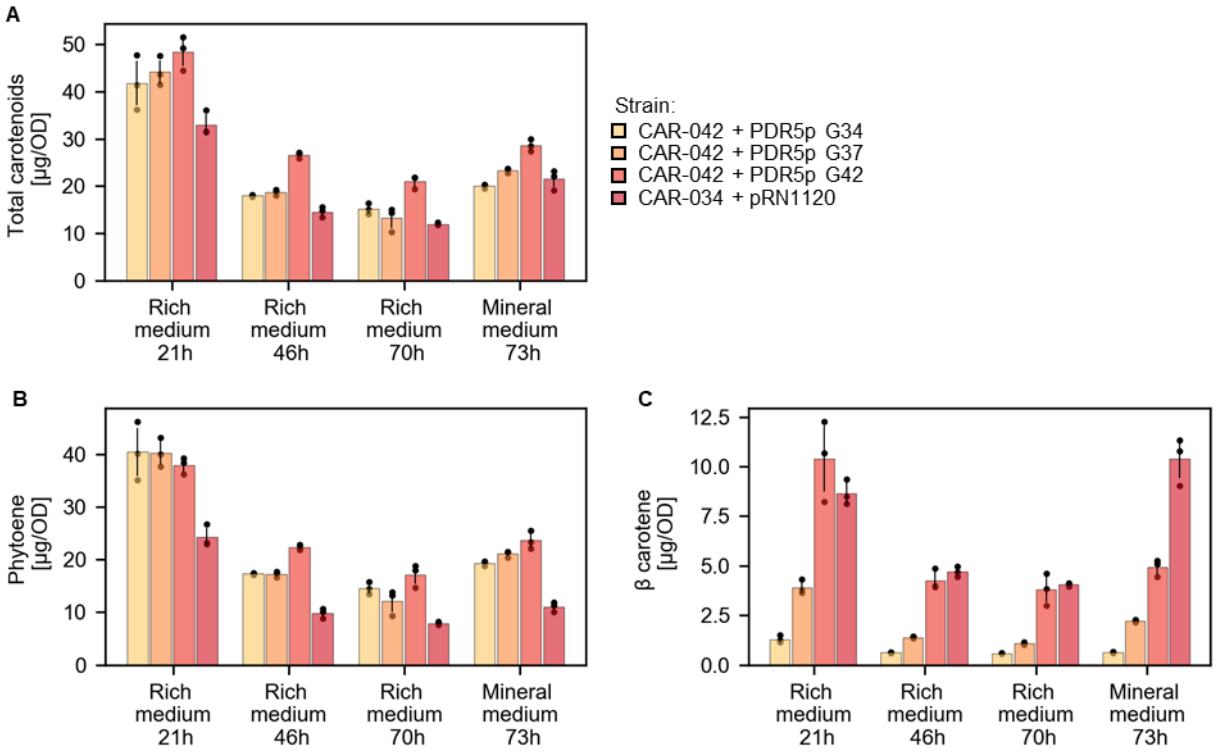


Figure 8-3. Production of carotenoids in production strain CAR-034 and strain CAR-042 with the feedback control system. A. Total carotenoids produced in fermentation on rich and mineral medium. **B**. Phytoene and **C**. β-carotene production profile. Strain CAR-042 is a carotenoids producing strain with the feedback control system comprised of dCas12a E925A NLS Mxi1 and gRNA targeting *crtI* expressed from PDR5p (G34, G37); as a control non-targeting gRNA G42 was applied. Strain CAR-034 is a control strain producing carotenoids with empty gRNA recipient plasmid pRN1120. Bars represent mean ± 1 standard deviation (n=3).

In addition to the growth experiment, carotenoids quantification was conducted for all sampling timepoints during fermentation on rich medium to confirm carotenoids production evaluated based on the colonies colour (Figure 8-3). The highest production of carotenoids was achieved after 21 h of growth in rich medium; total carotenoids levels were about two-fold higher compared to the later timepoints (46 and 70 h). In the previous analysis levels of total carotenoids were similar for timepoints between 24 and 72 h from the start of incubation (Chapter 2). However, expression of gRNA from a plasmid required supplementation of the growth medium with nourseothricin (NTC) which slowed growth of the wild type strain (OD_{600}=8.5 and 20 after one day, with and without NTC, respectively). Therefore, a difference in carotenoids production in the initial phase of the fermentation may be linked to slower growth. Total carotenoids levels produced by strain CAR-042 with the active feedback control system (gRNAs PDR5p G34, G37) resembled levels obtained for CAR-034 mostly because of higher phytoene accumulation. Downregulation of *crtI* with gRNA PDR5p G34 decreased β-carotene formation by 86±0.5% throughout the fermentation. Application of the feedback control system in combination with a less effective gRNA PDR5p G37 caused a drop in β-carotene formation by 59±3% after 21 h up to 72.1±1.3% after 70 h. gRNAs G34 and G37 controlled by SNR52p were tested before in combination with dCas12a expressed from TEF1p for downregulation of *crtI* and a full blockage formation of β-carotene was found (Chapter 5). Detection of β-carotene in strains with dCas12a expressed from a weak *PGI1* promoter in combination with gRNA controlled by the *PDR5* RNAP II promoter is presumably caused by the use of less effective expression system (lower dCas12a and gRNA expression). However, a complete cease in β-carotene production is not desirable as the feedback control system should only reduce the production level to mitigate burden. Upon cultivation in mineral medium carotenogenic strains produced 20.0–28.6 μg/OD_{600} of total carotenoids (73 h). Strain CAR-034 produced over 2-times more of β-carotene in comparison with the strain CAR-042 containing dCas12a and the non-targeting gRNA G42 (10.4±1.0 and 4.9±0.4 μg/OD_{600}), which may be caused by a significant difference in growth and prolonged exponential phase of the strains containing the feedback control system. Consistently with the results obtained for fermentation in the rich medium, a decrease in β-carotene production was observed upon expression of gRNAs PDR5p G34 and G37 in strain CAR-042. Surprisingly, phytoene accumulation in strain harbouring the feedback control system and non-functional gRNA (CAR-042 and gRNA PDR5p G42) was higher than in the strain CAR-034 despite the expression cassettes of the carotenogenic genes were the same for the both strains.

Finally genetic instability was tested by plating the production cultures on either mineral or rich agar plates (Figure 8-4, SI Figures 8-1, 8-2). Strain CAR-034 displayed better stability of β-carotene biosynthesis pathway in comparison with strain CAR-002 (Chapter 2). Cultivation of strain CAR-034 in mineral medium for 73 h followed by plating on mineral agar led to formation of uniformly orange colonies without colonies of different colour (*e.g.* white or pale pink) which would indicate disfunction of the pathway (*i.e.* loss of or mutation(s) in expression cassette(s). A possible explanation for improved pathway stability in strain CAR-034 over CAR-002 is selection of heterologous promoters (*TDH3* from *S. bayanus* and *TEF1* from *S. mikatae*) for expression of *crtE* and *crtYB* that could result in lower expression in non-native context (Table 8-1). A comparison of carotenoids production between these two strains revealed higher formation of phytoene by strain CAR-034 but lower β-carotene (4-fold in rich and 3-fold in mineral medium, Table 8-3). As observed in Chapter 2, strains with the *crt* expression cassettes containing weaker promoters

exhibited decreased carotenoids production but good growth and genetic stability. Therefore, higher production of carotenoids by strain CAR-002 could lead to genetic instability not observed for strain CAR-034 with 1.6-times lower production of total carotenoids (phytoene and β-carotene).

Table 8-3. Carotenoids production by strains CAR-002 and CAR-034 after 48 h of fermentation in rich and mineral medium. Total carotenoids value indicates sum of phytoene and β-carotene.

Strain	Medium	Phytoene (µg/OD)	β-carotene (µg/OD)	Total carotenoids (µg/OD)
CAR-002	rich	5.7±0.3	12.7±0.7	18.4±1.0
CAR-034	rich	8.0±0.6	3.2±0.0	11.2±0.6
CAR-002	mineral	6.3±0.0	20.1±0.5	26.4±0.5
CAR-034	mineral	9.4±1.1	6.6±0.1	16.0±1.2

Colonies formed by CAR-042 with gRNAs exhibited uniform colour with the same intensity as previously discussed transformants (Figures 8-1B, 8-4A). Plating mineral culture of carotenogenic strain CAR-034 after 73 h of cultivation on mineral agar did not lead to colony formation. Similarly, no colonies were obtained for strain CAR-042 and gRNA PDR5p G37. However, some colonies were produced of strain CAR-042 with other targeting gRNA PDR5p G34 and non-targeting PDR5p G42. A similar performance of tested strains was noted after 49 h of incubation but with shorter cultivation (26 h) all strains were capable of forming colonies (SI Figure 8-2B,D). Cultures that were grown in rich liquid medium for more than 46 h and subsequently plated on mineral agar did not grow, except for the wild type strain which might be caused by the change of growth conditions (*i.e.* mineral agar supplied with NTC) and carotenoids production.

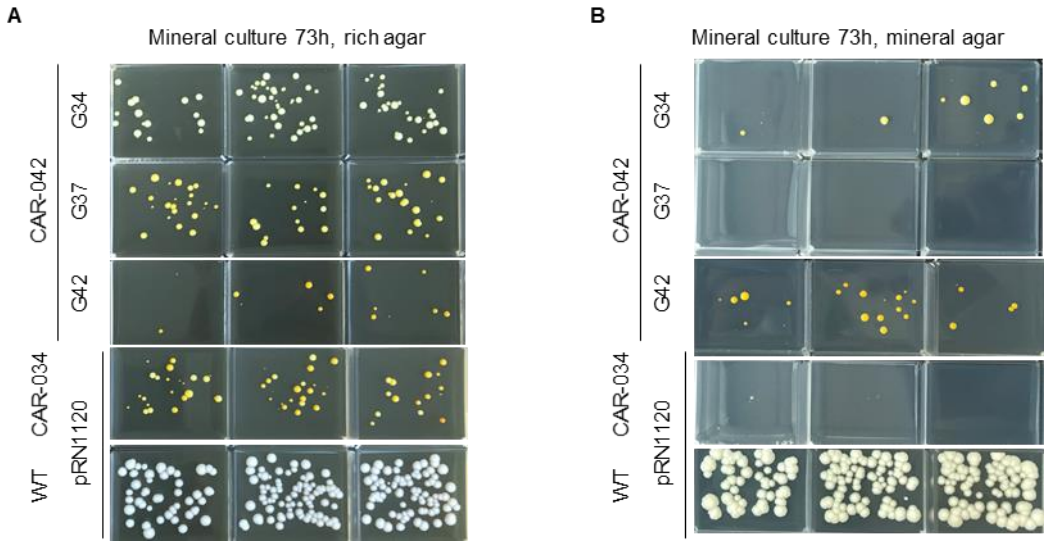


Figure 8-4. Stability of β-carotene biosynthesis pathway in strain with (CAR-042) and without the PDR5p feedback control system (CAR-034) after cultivation in mineral medium for 73 h. A. Cultures plated on rich agar and **B**. mineral agar supplemented with NTC. Each well in a row corresponds to a biological replicate (n=3).

8.4 Discussion

Here, we developed a feedback control system for a dynamic control of carotenoids formation to improve performance of the production host in terms of growth, pathway stability and productivity. In principle, the feedback control system should detect the elevated stress response induced by the production of β-carotene and temporarily reduce the production to allow the cell for restoring the energetic and metabolic equilibrium. The system was based on CRISPR interference mediated by dCas12a and coupled to the β-carotene production pathway by expression of a gRNA from the *PDR5* promoter that was identified to be upregulated upon carotenoids formation. Next, the feedback control system was implemented in the carotenoid producing strain CAR-002 which displayed genetic instability by promoting for example a loss-of function mutation in *crtYB* (Chapter 2). Expression of plasmid-borne dCas12a from a strong promoter and a targeting gRNA from the *ERG3* promoter in strain CAR-002 did not led to a visible downregulation of carotenoids formation. This system could be further optimized by integration of dCas12a which we previously shown to improve CRISPRi effectiveness in comparison with plasmid expression, and by applying gRNAs expressed from the PDR5p instead of ERG3p. To fully understand whether dCas12a expression from the TEF1p can compensate less effective plasmid system in comparison with genomically integrated PGI1p dCas12a expression cassette, further experiments are required. Optimization of the feedback control system in strain CAR-002 could provide the answer whether the genetic instability can be mitigated by careful balancing of the production level using a molecular device. Repression of carotenoids biosynthesis was achieved in strain CAR-042 with genomically integrated dCas12a controlled by a weak PGI1p and PDR5p-driven gRNAs targeting the promoter of *crtI*. Using this strain we tested if the feedback control system can restore cellular growth affected by a heterologous production. During fermentation in rich medium strain CAR-042 with the feedback control system displayed hampered growth after 21 h of incubation, but eventually achieved higher OD_{600} and dry weight than carotenogenic strain CAR-034. Optimization of the system by tuning expression levels of dCas12a and gRNA could improve growth of the strains with the feedback control system. Furthermore, expression of a gRNA from a plasmid and making use of a dominant selection marker also imposed a selection pressure which could be overcome by integrating gRNA into the genome. A severe impact on growth was seen for strain CAR-042 expressing gRNAs PDR5p G34, G37 or G42 during first two days of fermentation in mineral medium. Strain CAR-034 and the wild type strain did not display such growth defect which indicates link between hampered growth and gRNA and/or dCas12a expression. The feedback control system reduced formation of β-carotene by repressing expression of *crtI* which does not prevent cellular accumulation of phytoene. The feedback control system can be further optimized to reduce accumulation of the intermediates by multiplex downregulating expression of *crtE*, *crtYB* and *crtI*. As we showed in Chapter 5, dCas12a CRISPRi multiplex gene downregulation can be achieved with a single crRNA array encoding several genetic targets (with the current record set to 25 spacers[76]). Here, the singleplex downregulation was achieved using a gRNA expressed from RNAP II promoter and flanked by the self-cleaving ribozyme sites. To simplify architecture of a crRNA array for multiplex downregulation of *crt* genes, a spacer can be flanked solely by two direct repeats, reducing the length and complexity of the array. In Chapter 5 we have proven this gRNA design to be functional in combination with an RNAP II promoter.

Recent developments in synthetic biology explore burden-regulated circuits to alleviate production stress in microbial cell factories. Although these systems are developed to act in a dynamic manner, establishing such a system requires careful design and fine tuning. Our system was capable of repressing formation of β-carotene but did not release impaired growth in its current state. Nevertheless, with further optimization the established dCas12a-based feedback control system has the potential to improve metabolic pathways.

9 General discussion

A variety of microbial cell factories allow production of carotenoids in an effective and sustainable manner compared to the extraction from the natural sources or chemical synthesis[21,33,124]. Although the current state of genetic engineering enables expression of heterologous genes encoding the production pathway in an organism of choice, the production can negatively affect native metabolism and consequently hamper growth. To understand the stress response during the production of β-carotene in *S. cerevisiae*, we firstly characterized a set of strains producing low to high levels of carotenoids. Medium to high production of β-carotene affected cellular growth and destabilized the production pathway leading to the loss-of-function mutations in the β-carotene biosynthesis pathway gene *crtYB,* as shown in the sequencing data obtained for two transformants (**Chapter 2**). Due to the limited number of tested transformants, it is possible that the mutation disabling the production is located on a different locus than detected. To optimize the production of β-carotene we developed a feedback control system to dynamically regulate the production and prevent excessive accumulation of β-carotene and imbalanced relocation of cellular resources which can compromise metabolism and physiology. In contrast to the static control (*e.g.* promoter strength or plasmid copy number, which modulates gene expression and consequently levels of the proteins part of a metabolic pathway), implementation of a dynamic regulatory system allows to respond to a perturbation when it occurs, and to improve robustness of cells.

9.1 Feedback control system to mitigate burden in β-carotene production

Multiplex gene downregulation mediated by dCas12a-CRISPRi

To mitigate burden caused by a production of β-carotene in *S. cerevisiae* we established a feedback control system comprised of a stress-sensing circuit and regulatory circuit based on CRISPRi. To enable simultaneous downregulation of multiple genes, we developed a dCas12a system which exhibits the capability to automatically process gRNA arrays encoding multiple targets[75] (**Chapter 5**). In contrast to the selected dCas12a system, multiplex manipulation using the well-characterized dCas9 requires co-expression of RNA processing enzymes (*e.g.* Csy4) and a complex architecture of a gRNA (tracrRNA and an element enabling processing an array, *e.g.* tracrRNA or self-cleaving ribozyme)[190]. In addition, dCas12a is particularly suitable for targeting generally T-rich promoter sequences due to its PAM preference (5'-TTN-3')[74]. Next, we conducted a systematic evaluation of the key design paraments for CRISPR-dCas12a in *S. cerevisiae*, revealing important points for consideration to achieve functional downregulation for the feedback control system:

1. Fusion of a double NLS to the dCas12a reduces repression capability, therefore for the application in the feedback control system dCas12a should be equipped with a single C-terminal NLS at the 5' end of an additional repression domain.
2. Some repression domain can enhance dCas12a mediated downregulation, however some repressors abolish dCas12a activity. Our characterization of five repression domains revealed improved dCas12a performance when fused with Mxi1 and MIG1.

3. Expression of a gRNA from RNAP III promoter hampers growth of the yeast cells, which was not observed for the RNAP II promoter. As the feedback control system's function is to mitigate stress caused by production, the system itself cannot impose high energetic requirements on the cell, therefore selection of an RNAP II promoter for control of gRNA was essential.

Although in Chapter 5 we showed efficient multiplex downregulation of carotenogenesis using dCas12a and an array expressed from a *SNR52* promoter, multiplex downregulation was not tested in the feedback control system. Orchestrating expression levels of genes involved in the β-carotene biosynthesis would allow to balance the pathway and prevent accumulation of toxic intermediates. To achieve that, a gRNA array targeting *crtE*, *crtYB* and *crtI* should be expressed from a burden sensing promoter. As shown in Chapter 5, dCas12a can process a gRNA expressed from an RNAP II type promoter if a spacer is flanked by two direct repeats, reducing the length and complexity of a gRNA array (*i.e.* by omitting self-cleaving ribozymes sequences).

In our feedback control system we used gRNAs controlled by a burden-sensitive promoter, PDR5p, which strongly repressed expression of *RFP* and production of β-carotene. The repression range can be extended by selection of gRNAs containing mismatches, an approach applied in a feedback control system by Ceroni *et al.*[24]. Alternatively, gRNA with a suboptimal PAM sequence can be used, for example CTTV instead of TTTN for dLbCas12a, previously assayed by repression of a fluorescent protein gene[98]. By decreasing the degree of downregulation, production could be maintained on a minimal level without imposing burden. Furthermore, presence of a single mismatch in a gRNA was shown to reduce residency time of the dCas9-gRNA complex on the genomic DNA[204] which would provide a temporary control rather than a prolonged inhibition of the production pathway.

Monitoring burden related to the production of β-carotene

The regulatory system should be activated by a burden-sensing promoter when high level of β-carotene production compromises cell growth and metabolism. A temporary suspension of the production should allow the cell to restore its energetic and metabolic balance and re-activate the production. To provide the dynamicity of the feedback control system, promoters sensitive to the production of β-carotene were identified by RNA-seq (**Chapter 7**). An increased number of transcripts was detected for genes involved in ergosterol biosynthesis and transporters participating in pleiotropic drug resistance. Promoters of two genes from these pathways (*ERG3* – ergosterol biosynthesis and *PDR5* – pleiotropic drug resistance) were isolated and assessed for activity by expression of *RFP* and gRNA in a CRISPRi system. Although an attempt in understanding tunability of these promoters with an increasing production of carotenoids were made, obtained results were inconclusive and require further investigation.

Burden-driven control of β-carotene production

Finally, the dCas12a-CRISPRi system was coupled to the cellular burden by expression of a gRNA from a stress-sensing *PDR5* promoter to create a feedback control system (**Chapter 8**). β-carotene production and the performance of the host organism (in terms of growth and genetic stability of the production pathway) was then compared. Initially slower growth of the strain harbouring the feedback

control system was observed, however this strain achieved higher dry weight after 70 h of fermentation on rich medium compared to carotenogenic strain without the regulatory device. Improved product yield (β-carotene) was not observed when the feedback control system was applied, which may be overcome by selection of a gRNA with lower repressing capability and reduction of the production below the burdensome level instead of blocking the production[24]. To assess improvements in the genetic stability of the β-carotene biosynthesis pathway with the feedback control system, a set of strains used in this experiment should be expanded.

Post-transcriptional pathway control by Cas13-CRISPR

The developed feedback control system relies on repression of gene transcription, *i.e.* dCas12a-gRNA complex sterically blocking access of RNAP to chromosomal DNA. In addition to a dCas12a-based system, we have assessed CRISPR-Cas13 for a transcript knockdown in *S. cerevisiae* to potentially expand the feedback control system by regulation on the post-transcriptional level. A potential advantage of the CRISPR-Cas13 mediated regulation of gene expression over CRISPRi may lay in dynamics – while dCas9-gRNA complex was shown to stably bind DNA, Cas13 degrades mRNA. Further research could reveal stability of the Cas13-gRNA complex and whether the system can be titrated by gRNA abundance. Therefore, Cas13 could provide a faster and more dynamic tool to control gene expression. In addition, the PFS requirement of Cas13 is less strict than PAM sequence of dCas9 or dCas12a, expanding the number of possible targets. For example, LshCas13a favours a single nucleotide PFS (non-G at the 3' of a target) and LwaCas13a exhibits no PFS preference, in contract to the 2-4 nucleotides long PAM sequence of Cas9 and LbCas12a (5'-NGG'-3' and 5'-TTTV-3')[211]. As shown in Chapter 5, a strong downregulation cannot be achieved when dCas12a targets an ORF sequence, thus Cas13 could be applied when constrains for targeting a promoter occur.

Conducted experiments revealed a collateral effect, defined as unspecific cleavage of RNA species in the cell upon activation of Cas13 which disables a potential application of this system for specific transcript knockdown (**Chapter 6**). Instead, we show that the CRISPR-Cas13a system is not lethal to the yeast cells but rather reversibly inhibits growth, opening an avenue for a CRISPR-programmed growth control.

Designing the regulatory circuit

Here we proposed to regulate a production pathway on transcriptional and post-transcriptional level, however the feedback control system could be further extended to regulate the protein abundance by programmable proteases. Use of native or orthogonal protein degradation machinery to control a heterologous production was reported before[245–248], including a dynamic control system comprised of dCas9-CRISPRi and a protease[249]. To provide new solutions to mitigate production burden in a cell factory, we explored CRISPR technology. The advantage of the protease based system is a rapid shut down of a production pathway by degradation of involved enzymes, which cannot be achieved with CRISPR regulation. Instead, dCas12a-CRISPRi prevents allocation of the RNAP pool for transcription of a heterologous gene, whereas a functional Cas13-CRISPR silencing would allow to allocate ribosomes for translation of native proteins. A recent study examining burden in mammalian cells reveal that the limitation in the resources occurs on the transcriptional level rather than translational, contrary to

prokaryote, thus a CRISPR based feedback loop which acts on transcription might be advantageous in *S. cerevisiae*[26,29]. With the current development of this molecular toolbox, gene expression can be regulated at different stages and further research on resources allocation in *S. cerevisiae* could reveal the most optimal approach.

9.2 Expanding the CRISPR-based toolbox to optimize *S. cerevisiae* as a production chassis

To create the feedback control, CRISPR systems were explored resulting in development of three CRISPR-based tools for a potential optimization of *S. cerevisiae* as a cell factory for β-carotene production:

1. We developed a detailed protocol for multiplex genome editing using CRISPR-Cas12a and demonstrated construction of three carotenoids producing strains as a proof-of-concept (**Chapter 3**). This protocol can be transferred to create a *S. cerevisiae* strain with a multistep production pathway of choice.

2. CRISPR-dCas12a was established for repression of gene expression in multiplex manner (**Chapter 5**). This tool can be applied when repression of multiple genes is required or when target has limited number of regions containing 5'-NGG-3' PAM favourable by dCas9 but the sequence is T-rich.

3. This study revealed limitation of a CRISPR-Cas13 system for gene knockdown due to its trans (collateral) activity, however we propose use Cas13a for growth control, which is an important consideration in computerization of biological systems (**Chapter 6**).

10 References

1. Nielsen, J. & Keasling, J. D. Engineering Cellular Metabolism. *Cell* **164**, 1185-1197 (2016) doi:10.1016/j.cell.2016.02.004.
2. Nambisan, P. *Recombinant DNA Technology and Genetically Modified Organisms*. *An Introduction to Ethical, Safety and Intellectual Property Rights Issues in Biotechnology* (2017). doi:10.1016/b978-0-12-809231-6.00004-1.
3. Lee, S. Y., Mattanovich, D. & Villaverde, A. Systems metabolic engineering, industrial biotechnology and microbial cell factories. *Microb. Cell Fact.* **11**, 1-3 (2012).
4. Jeffries, T. W. Engineering yeasts for xylose metabolism. *Curr. Opin. Biotechnol.* **17**, 320–326 (2006).
5. Kavšček, M., Stražar, M., Curk, T., Natter, K. & Petrovič, U. Yeast as a cell factory: current state and perspectives. *Microb. Cell Fact.* **14**, 1-10 (2015) doi:10.1186/s12934-015-0281-x.
6. Mitsui, R. & Yamada, R. *Saccharomyces cerevisiae as a microbial cell factory*. *Microbial Cell Factories Engineering for Production of Biomolecules* vol. 1978 (Elsevier Inc., 2021).
7. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821 (2012) doi:10.1126/science.1225829.
8. Davy, A. M., Kildegaard, H. F. & Andersen, M. R. Cell Factory Engineering. *Cell Syst.* **4**, 262–275 (2017).
9. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173-1183 (2013) doi:10.1016/j.cell.2013.02.022.
10. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647-661 (2014) doi:10.1016/j.cell.2014.09.029.
11. Walker, G. M. & Walker, R. S. K. *Enhancing Yeast Alcoholic Fermentations*. *Advances in Applied Microbiology* vol. 105 (Elsevier Ltd, 2018).
12. Redden, H., Morse, N. & Alper, H. S. The synthetic biology toolbox for tuning gene expression in yeast. *FEMS Yeast Res.* **15**, 1–10 (2015).
13. Liu, L., Redden, H. & Alper, H. S. Frontiers of yeast metabolic engineering: Diversifying beyond ethanol and *Saccharomyces*. *Curr. Opin. Biotechnol.* **24**, 1023–1030 (2013).
14. Borodina, I. & Nielsen, J. Advances in metabolic engineering of yeast *Saccharomyces cerevisiae* for production of chemicals. *Biotechnology Journal* **9**, 609-620 (2014) doi:10.1002/biot.201300445.
15. Prinz, B. *et al.* Establishing a versatile fermentation and purification procedure for human proteins expressed in the yeasts *Saccharomyces cerevisiae* and Pichia pastoris for structural genomics. *J. Struct. Funct. Genomics* **5**, 29–44 (2004).
16. Curran, K. A., Leavitt, J. M., Karim, A. S. & Alper, H. S. Metabolic engineering of muconic acid production in *Saccharomyces cerevisiae*. *Metab. Eng.* **15**, 55–66 (2013).
17. Lv, X. *et al.* Enhanced isoprene biosynthesis in *Saccharomyces cerevisiae* by engineering of the native acetyl-CoA and mevalonic acid pathways with a push-pull-restrain strategy. *J. Biotechnol.* **186**, 128–136 (2014).
18. Nielsen, J. Production of biopharmaceutical proteins by yeast: Advances through metabolic engineering. *Bioengineered* **4**, 207–211 (2013).
19. Ro, D. K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast.

Nature **440**, 940–943 (2006).

20. Hansen, E. H. *et al.* De novo biosynthesis of Vanillin in fission yeast (*Schizosaccharomyces pombe*) and baker's yeast (*Saccharomyces cerevisiae*). *Appl. Environ. Microbiol.* **75**, 2765–2774 (2009).
21. Verwaal, R. *et al.* High-level production of beta-carotene in *Saccharomyces cerevisiae* by successive transformation with carotenogenic genes from *Xanthophyllomyces dendrorhous*. *Appl. Environ. Microbiol.* **73**, 4342-4350 (2007) doi:10.1128/AEM.02759-06.
22. Buijs, N. A., Siewers, V. & Nielsen, J. Advanced biofuel production by the yeast *Saccharomyces cerevisiae*. *Curr. Opin. Chem. Biol.* **17**, 480–488 (2013).
23. Borkowski, O., Ceroni, F., Stan, G. B. & Ellis, T. Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Current Opinion in Microbiology* **33**, 123-130 (2016) doi:10.1016/j.mib.2016.07.009.
24. Ceroni, F. *et al.* Burden-driven feedback control of gene expression. *Nat. Methods* **15**, 387-393 (2018) doi:10.1038/nmeth.4635.
25. Boo, A., Ellis, T. & Stan, G. B. Host-aware synthetic biology. *Curr. Opin. Syst. Biol.* **14**, 66–72 (2019).
26. Gorochowski, T. E., Avcilar-Kucukgoze, I., Bovenberg, R. A. L., Roubos, J. A. & Ignatova, Z. A Minimal Model of Ribosome Allocation Dynamics Captures Trade-offs in Expression between Endogenous and Synthetic Genes. *ACS Synth. Biol.* **5**, 710-720 (2016) doi:10.1021/acssynbio.6b00040.
27. Gyorgy, A. *et al.* Isocost Lines Describe the Cellular Economy of Genetic Circuits. *Biophys. J.* **109**, 639–646 (2015).
28. Li, G., Burkhardt, D., Gross, C. & Weissman, J. S. Quantifying Absolute Protein Synthesis Rates Reveals Principles Underlying Allocation of Cellular Resources. *Cell* **157**, 624–635 (2014).
29. Jones, R. D. *et al.* An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells. *Nat. Commun.* **11**, 1–16 (2020).
30. Mattanovich, D., Gasser, B., Hohenblum, H. & Sauer, M. Stress in recombinant protein producing yeasts. *J. Biotechnol.* **113**, 121–135 (2004).
31. Eguchi, Y. *et al.* Estimating the protein burden limit of yeast cells by measuring the expression limits of glycolytic proteins. *Elife* **7**, 1–3 (2018).
32. Moriya, H. Quantitative nature of overexpression experiments. *Mol. Biol. Cell* **26**, 3932–3939 (2015).
33. Hong, J., Park, S. H., Kim, S., Kim, S. W. & Hahn, J. S. Efficient production of lycopene in *Saccharomyces cerevisiae* by enzyme engineering and increasing membrane flexibility and NAPDH production. *Appl. Microbiol. Biotechnol.* **103**, 211-223 (2019) doi:10.1007/s00253-018-9449-8.
34. Ukibe, K., Hashida, K., Yoshida, N. & Takagi, H. Metabolic engineering of *Saccharomyces cerevisiae* for astaxanthin production and oxidative stress tolerance. *Appl. Environ. Microbiol.* **75**, 7205-7211 (2009) doi:10.1128/AEM.01249-09.
35. Liu, D., Mannan, A. A., Han, Y., Oyarzún, D. A. & Zhang, F. Dynamic metabolic control: towards precision engineering of metabolism. *J. Ind. Microbiol. Biotechnol.* **45**, 535–543 (2018).
36. Zhang, F., Carothers, J. M. & Keasling, J. D. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol.* **30**, 354-359 (2012).
37. Lalwani, M. A., Zhao, E. M. & Avalos, J. L. Current and future modalities of dynamic control in metabolic engineering. *Curr. Opin. Biotechnol.* **52**, 56–65 (2018).

38. Teixeira, P. G., Ferreira, R., Zhou, Y. J., Siewers, V. & Nielsen, J. Dynamic regulation of fatty acid pools for improved production of fatty alcohols in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **16**, 1-11 (2017).
39. Scalcinati, G. *et al.* Dynamic control of gene expression in *Saccharomyces cerevisiae* engineered for the production of plant sesquitepene α-santalene in a fed-batch mode. *Metab. Eng.* **14**, 91–103 (2012).
40. Xie, W., Ye, L., Lv, X., Xu, H. & Yu, H. Sequential control of biosynthetic pathways for balanced utilization of metabolic intermediates in *Saccharomyces cerevisiae*. *Metab. Eng.* **28**, 8–18 (2015).
41. Moser, F. *et al.* Dynamic control of endogenous metabolism with combinatorial logic circuits. *Mol. Syst. Biol.* **14**, 1–18 (2018).
42. Zhou, L. B. & Zeng, A. P. Exploring Lysine Riboswitch for Metabolic Flux Control and Improvement of L-Lysine Synthesis in *Corynebacterium glutamicum*. *ACS Synth. Biol.* **4**, 729–734 (2015).
43. Barajas, C., Gibson, J., Sandoval, L. & Del Vecchio, D. Ribosome actuator via the SpoT/ppGpp pathway to mitigate gene overexpression burden. *bioRxiv* 2021.02.11.430724 (2021).
44. Mahr, R., von Boeselager, R. F., Wiechert, J. & Frunzke, J. Screening of an *Escherichia coli* promoter library for a phenylalanine biosensor. *Appl. Microbiol. Biotechnol.* **100**, 6739–6753 (2016).
45. Barrangou, R. & Marraffini, L. A. CRISPR-cas systems: Prokaryotes upgrade to adaptive immunity. *Mol. Cell* **54**, 234–244 (2014).
46. Jiang, F., Zhou, K., Ma, L., Gressel, S. & Doudna, J. A. A Cas9-guide RNA complex preorganized for target DNA recognition. *Science* **348**, 1477–1481 (2015).
47. Mohanraju, P. *et al.* Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* **353** (2016).
48. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
49. Dicarlo, J. E. *et al.* Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* **41**, 4336-434 (2013) doi:10.1093/nar/gkt135.
50. Bandaru, S. *et al.* Structure-based design of gRNA for Cas13. *Sci. Rep.* **10**, 1–12 (2020).
51. Tang, Y. & Fu, Y. Class 2 CRISPR/Cas: An expanding biotechnology toolbox for and beyond genome editing 06 Biological Sciences 0604 Genetics. *Cell Biosci.* **8**, 1–13 (2018).
52. Schindele, P., Wolter, F. & Puchta, H. Transforming plant biology and breeding with CRISPR/Cas9, Cas12 and Cas13. *FEBS Lett.* **592**, 1954–1967 (2018).
53. Murugan, K., Babu, K., Sundaresan, R., Rajan, R. & Sashital, D. G. The Revolution Continues: Newly Discovered Systems Expand the CRISPR-Cas Toolkit. *Mol. Cell* **68**, 15–25 (2017).
54. Mans, R. *et al.* CRISPR/Cas9: A molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **15**, 1–15 (2015).
55. Horwitz, A. A. *et al.* Efficient Multiplexed Integration of Synergistic Alleles and Metabolic Pathways in Yeasts via CRISPR-Cas Identifying Determinants of Success for Multiplex. **1**, 88–96 (2015) doi:10.1016/j.cels.2015.02.001.
56. Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K. & Doudna, J. A. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **329**, 1355-1358 (2010) doi:10.1126/science.1192272.

57. Gao, Y. & Zhao, Y. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. *J. Integr. Plant Biol.* **56**, 343-349 (2014) doi:10.1111/jipb.12152.
58. Xie, K., Minkenberg, B. & Yang, Y. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 3570-3575 (2015) doi:10.1073/pnas.1420294112.
59. Walter, J. M., Chandran, S. S. & Horwitz, A. A. CRISPR-Cas-Assisted Multiplexing (CAM): Simple Same-Day Multi-Locus Engineering in Yeast. *J. Cell. Physiol.* **231**, 2563-2569 (2016) doi:10.1002/jcp.25375.
60. Tang, X. D. *et al.* Methods for enhancing clustered regularly interspaced short palindromic repeats/Cas9-mediated homology-directed repair efficiency. *Front. Genet.* **10** (2019).
61. Jang, I. S., Yu, B. J., Jang, J. Y., Jegal, J. & Lee, J. Y. Improving the efficiency of homologous recombination by chemical and biological approaches in *Yarrowia lipolytica*. *PLoS One* **13**, 1–10 (2018).
62. Zetsche, B. *et al.* Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nature Biotechnology* **35**, 31-34 (2017) doi:10.1038/nbt.3737.
63. Swiat, M. A. *et al.* FnCpf1: A novel and efficient genome editing tool for *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **45**, 12585-12598 (2017) doi:10.1093/nar/gkx1007.
64. Gier, R. A. *et al.* High-performance CRISPR-Cas12a genome editing for combinatorial genetic screening. *Nat. Commun.* **11**, 1-9 (2020) doi:10.1038/s41467-020-17209-1.
65. Verwaal, R., Buiting-Wiessenhaan, N., Dalhuijsen, S. & Roubos, J. A. CRISPR/Cpf1 enables fast and simple genome editing of *Saccharomyces cerevisiae*. *Yeast* **35**, 201-211 (2018) doi:10.1002/yea.3278.
66. Ciurkot, K., Vonk, B., Gorochowski, T. E., Roubos, J. A. & Verwaal, R. CRISPR/Cas12a Multiplex Genome Editing of *Saccharomyces cerevisiae* and the Creation of Yeast Pixel Art. *J. Vis. Exp.* **147**, e59350 (2019) doi:10.3791/59350.
67. Ciurkot, K., Gorochowski, T. E., Roubos, J. A. & Verwaal, R. Efficient multiplexed gene regulation in *Saccharomyces cerevisiae* using dCas12a. *Nucleic Acids Res.* **49**, 7775–7790 (2021) doi:10.1093/nar/gkab529.
68. Abudayyeh, O. O. *et al.* C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353** (2016) doi:10.1126/science.aaf5573.
69. East-Seletsky, A. *et al.* Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nat. Publ. Gr.* **538**, 270–273 (2016).
70. Abudayyeh, O. O. *et al.* RNA targeting with CRISPR-Cas13. *Nature* **550**, 280-284 (2017) doi:10.1038/nature24049.
71. Koonin, E. V., Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. *Current Opinion in Microbiology* **37**, 67-78 (2017) doi:10.1016/j.mib.2017.05.008.
72. Yan, W. X. *et al.* Functionally diverse type V CRISPR-Cas systems. **91**, 88–91 (2019).
73. Makarova, K. S. *et al.* An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* **13**, 722–736 (2015).
74. Zetsche, B. *et al.* Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* **163**, 759-771 (2015) doi:10.1016/j.cell.2015.09.038.

75. Fonfara, I., Richter, H., Bratovic, M., Le Rhun, A. & Charpentier, E. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* **163**, 759-771 (2016) doi:10.1038/nature17945.
76. Campa, C. C., Weisbach, N. R., Santinha, A. J., Incarnato, D. & Platt, R. J. Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. *Nat. Methods* **16**, 887-893 (2019) doi:10.1038/s41592-019-0508-6.
77. Zhang, J., Yang, F., Yang, Y., Jiang, Y. & Huo, Y. X. Optimizing a CRISPR-Cpf1-based genome engineering system for *Corynebacterium glutamicum*. *Microb. Cell Fact.* **18**, 1–13 (2019).
78. Moreno-Mateos, M. A. *et al.* CRISPR-Cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing. *Nat. Commun.* **8**, 1–9 (2017).
79. Yan, M., Yan, H., Ren, G., Zhao, J. & Guo, X. CRISPR-Cas12a-Assisted Recombineering in Bacteria. *Appl. Environ. Microbiol.* **83**, e00947-17 (2017).
80. Ramesh, A., Ong, T., Garcia, J. A., Adams, J. & Wheeldon, I. Guide RNA Engineering Enables Dual Purpose CRISPR-Cpf1 for Simultaneous Gene Editing and Gene Regulation in *Yarrowia lipolytica*. *ACS Synth. Biol.* **9**, 967–971 (2020).
81. Zhao, Y. & Boeke, J. D. CRISPR-Cas12a system in fission yeast for multiplex genomic editing and CRISPR interference. *Nucleic Acids Res.* **48**, 5788-5798 (2020) doi:10.1093/nar/gkaa329.
82. Tang, X. *et al.* A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nat. Plants* **3**, 1–5 (2017).
83. Liu, P. *et al.* Enhanced Cas12a editing in mammalian cells and zebrafish. *Nucleic Acids Res.* **47**, 4169-4180 (2019) doi:10.1093/nar/gkz184.
84. Port, F., Starostecka, M. & Boutros, M. Multiplexed conditional genome editing with Cas12a in Drosophila. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 22890–22899 (2020).
85. Connell, M. R. O. Molecular Mechanisms of RNA Targeting by Cas13-containing Type VI CRISPR – Cas Systems. *J. Mol. Biol.* **431**, 66–87 (2019).
86. Meeske, A. J., Nakandakari-Higa, S. & Marraffini, L. A. Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage. *Nature* **570**, 241–245 (2019).
87. Wang, Q. *et al.* The CRISPR-Cas13a Gene-Editing System Induces Collateral Cleavage of RNA in Glioma Cells. *Adv. Sci.* **6**, 1901299 (2019).
88. Wu, Q. W. & Kapfhammer, J. P. The Bacterial Enzyme Cas13 Interferes with Neurite Outgrowth from Cultured Cortical Neurons. *Toxins* **13**, 262 (2021).
89. Cox, D. B. T. *et al.* RNA editing with CRISPR-Cas13. *Science* **358**, 1019–1027 (2017).
90. Aman, R. *et al.* RNA virus interference via CRISPR/Cas13a system in plants. *Genome Biol.* **19**, 1–9 (2018).
91. Aman, R. *et al.* Engineering RNA virus interference via the CRISPR/Cas13 machinery in arabidopsis. *Viruses* **10**, 732 (2018).
92. Smargon, A. A., Cox, D. B. T., Pyzocha, N. K., Makarova, K. S. & Koonin, E. V. Cas13b Is a Type VI-B CRISPR-Associated RNA- Guided RNase Differentially Regulated by Accessory Article Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. *Mol. Cell* **65**, 618–630 (2017).

93. Gilbert, L. A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442 (2013).
94. Lian, J., Hamedirad, M., Hu, S. & Zhao, H. Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. *Nat. Commun.* **8**, 1-9 (2017) doi:10.1038/s41467-017-01695-x.
95. Zhang, J. lai *et al.* Gene repression via multiplex gRNA strategy in *Y. lipolytica*. *Microb. Cell Fact.* **17**, 1-13 (2018) doi:10.1186/s12934-018-0909-8.
96. Smith, J. D. *et al.* Quantitative CRISPR interference screens in yeast identify chemical-genetic interactions and new rules for guide RNA design. *Genome Biol.* **17**, 1-16 (2016) doi:10.1186/s13059-016-0900-9.
97. Didovyk, A., Borek, B., Tsimring, L. & Hasty, J. Transcriptional regulation with CRISPR-Cas9: Principles, advances, and applications. *Curr. Opin. Biotechnol.* **40**, 177–184 (2016).
98. Miao, C., Zhao, H., Qian, L. & Lou, C. Systematically investigating the key features of the DNase deactivated Cpf1 for tunable transcription regulation in prokaryotic cells. *Synth. Syst. Biotechnol.* **4**, 1-9 (2019) doi:10.1016/j.synbio.2018.11.002.
99. Zalatan, J. G. *et al.* Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* **160**, 339-350 (2015) doi:10.1016/j.cell.2014.11.052.
100. Kiani, S. *et al.* Cas9 gRNA engineering for genome editing, activation and repression. *Nat. Methods* **12**, 1051–1054 (2015).
101. Breinig, M. *et al.* Multiplexed orthogonal genome editing and transcriptional activation by Cas12a. *Nat. Methods* **16**, 51–54 (2019).
102. Wu, Y. *et al.* Design of a programmable biosensor-CRISPRi genetic circuits for dynamic and autonomous dual-control of metabolic flux in Bacillus subtilis. *Nucleic Acids Res.* **48**, 996–1009 (2020).
103. Green, A. S. & Fascetti, A. J. Meeting the Vitamin A Requirement: The Efficacy and Importance of β -Carotene in Animal Species. *Sci. World J.* (2016).
104. Goodwin, H. T. *The biochemistry of the carotenoids, vol I. Plants*. (Chapman and Hall, 1980).
105. Maoka, T. Carotenoids as natural functional pigments. *J. Nat. Med.* **74**, 1-16 (2020) doi:10.1007/s11418-019-01364-x.
106. Mathews, M. M. & W., S. S. Function of carotenoid pigments in non-phototsynthesic bacteria. *Nature* **104**, 1892–1893 (1959).
107. Rivera-Madrid, R., Carballo-Uicab, V. M., Cárdenas-Conejo, Y., Aguilar-Espinosa, M. & Siva, R. *Overview of carotenoids and beneficial effects on human health*. *Carotenoids: Properties, Processing and Applications* (Elsevier Inc., 2020). doi:10.1016/b978-0-12-817067-0.00001-4.
108. Fernandes, T. & Cordeiro, N. Microalgae as sustainable biofactories to produce high-value lipids: Biodiversity, exploitation, and biotechnological applications. *Mar. Drugs* **19**, 573 (2021).
109. Visser, H., Van Ooyen, A. J. J. & Verdoes, J. C. Metabolic engineering of the astaxanthin-biosynthetic pathway of *Xanthophyllomyces dendrorhous*. *FEMS Yeast Res.* **4**, 221–231 (2003).
110. Hong, J. *et al.* The cause of on-target point mutations generated by CRISPR-Cas9 treatment in the yeast *Xanthophyllomyces dendrorhous*. *bioRxiv* (2021).
111. Hara, K. Y. *et al.* Evaluation and screening of efficient promoters to improve astaxanthin production

in *Xanthophyllomyces dendrorhous*. *Appl. Microbiol. Biotechnol.* **98**, 6787–6793 (2014).

112. Barreiro, C. & José-Luis, B. *Microbial carotenoids. Annual review of microbiology* vol. 26 (Springer New York, 2018).
113. Borowitzka, L. & Borowitzka, M. Commercial Production of β-Carotene by Dunaliella Salina in Open Ponds. *Bull. Mar. Sci.* **47**, 244–252 (1990).
114. Schmidt, I. *et al.* Biotechnological production of astaxanthin with *Phaffia rhodozyma*/*Xanthophyllomyces dendrorhous*. *Appl. Microbiol. Biotechnol.* **89**, 555–571 (2011).
115. Zhao, J. *et al.* Engineering central metabolic modules of *Escherichia coli* for improving β-carotene production. *Metab. Eng.* **17**, 42-50 (2013) doi:10.1016/j.ymben.2013.02.002.
116. Jacobsen, I. H., Ledesma-Amaro, R. & Martinez, J. L. Recombinant β-Carotene Production by *Yarrowia lipolytica* – Assessing the Potential of Micro-Scale Fermentation Analysis in Cell Factory Design and Bioreaction Optimization. *Front. Bioeng. Biotechnol.* **8**, 29 (2020) doi:10.3389/fbioe.2020.00029.
117. Liu, P., Sun, L., Sun, Y., Shang, F. & Yan, G. Decreased fluidity of cell membranes causes a metal ion deficiency in recombinant *Saccharomyces cerevisiae* producing carotenoids. *J. Ind. Microbiol. Biotechnol.* **43**, 525-535 (2016) doi:10.1007/s10295-015-1728-0.
118. Chen, Y. *et al.* Lycopene overproduction in *Saccharomyces cerevisiae* through combining pathway engineering with host engineering. *Microb. Cell Fact.* **15**, 1-13 (2016) doi:10.1186/s12934-016-0509-4.
119. Sun, Y., Sun, L., Shang, F. & Yan, G. Enhanced production of β-carotene in recombinant *Saccharomyces cerevisiae* by inverse metabolic engineering with supplementation of unsaturated fatty acids. *Process Biochem.* **51**, 568–577 (2016).
120. Jordá, T. & Puig, S. Regulation of ergosterol biosynthesis in *Saccharomyces cerevisiae*. *Genes (Basel).* **11**, 1–18 (2020).
121. Zhou, P., Ye, L., Xie, W., Lv, X. & Yu, H. Highly efficient biosynthesis of astaxanthin in *Saccharomyces cerevisiae* by integration and tuning of algal *crtZ* and *bkt*. *Appl. Microbiol. Biotechnol.* **99**, 8419–8428 (2015).
122. Larroude, M. *et al.* A synthetic biology approach to transform *Yarrowia lipolytica* into a competitive biotechnological producer of β-carotene. Biotechnology and bioengineering, **115**, 464-472 (2018) doi:10.1002/bit.26473.
123. Xie, W. *et al.* Construction of a Controllable β-Carotene Biosynthetic Pathway by Decentralized Assembly Strategy in *Saccharomyces cerevisiae*. **111**, 125–133 (2014).
124. Farmer, W. R. & Liao, J. C. Improving lycopene production in *Escherichia coli* by engineering metabolic control. *Nat. Biotechnol.* **18**, 533–537 (2000).
125. Verwaal, R. *et al.* Heterologous carotenoid production in *Saccharomyces cerevisiae* induces the pleiotropic drug resistance stress response. *Yeast* **27**, 983-998 (2010) doi:10.1002/yea.1807.
126. Hahn-Hägerdal, B. *et al.* Role of cultivation media in the development of yeast strains for large scale industrial use. *Microbial Cell Factories* **4**, 1-16 (2005) doi:10.1186/1475-2859-4-31.
127. Verduyn, C., Postma, E., Scheffers, W. A. & Van Dijken, J. P. Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* **8**, 501–517 (1992).

128. Van Dijken, J. P. *et al.* An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. *Enzyme and Microb. Technol.* **26**, 706-714 (2000). doi:10.1016/S0141-0229(00)00162-9.
129. Rugbjerg, P. & Olsson, L. The future of self-selecting and stable fermentations. *J. Ind. Microbiol. Biotechnol.* **47**, 993–1004 (2020).
130. Meinander, N. Q. & Hahn-Hägerdal, B. Fed-batch xylitol production with two recombinant *Saccharomyces cerevisiae* strains expressing XYL1 at different levels, using glucose as a cosubstrate: A comparison of production parameters and strain stability. *Biotechnol. Bioeng.* **54**, 391–399 (1997).
131. Knott, G. J. & Doudna, J. A. CRISPR-Cas guides the future of genetic engineering. *Science* **361**, 866–869 (2018).
132. Li, Z. H., Liu, M., Lyu, X. M., Wang, F. Q. & Wei, D. Z. CRISPR/Cpf1 facilitated large fragment deletion in *Saccharomyces cerevisiae*. *J. Basic Microbiol.* **58**, 1100–1104 (2018).
133. Shao, Y., Lu, N., Qin, Z. & Xue, X. CRISPR-Cas9 Facilitated Multiple-Chromosome Fusion in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **7**, 2706–2708 (2018).
134. Stan J J Brouns, Matthijs M Jore, Magnus Lundgren, Edze R Westra, Rik J H Slijkhuis, Ambrosius P L Snijders, Mark J Dickman, Kira S Makarova, Eugene V Koonin. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–4 (2008).
135. Lian, J., HamediRad, M. & Zhao, H. Advancing Metabolic Engineering of *Saccharomyces cerevisiae* Using the CRISPR/Cas System. *Biotechnology Journal* **13**, 1700601 (2018) doi:10.1002/biot.201700601.
136. Ferreira, R., Skrekas, C., Nielsen, J. & David, F. Multiplexed CRISPR/Cas9 Genome Editing and Gene Regulation Using Csy4 in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **7**, 10–15 (2018).
137. Swarts, D. C. & Jinek, M. Cas9 versus Cas12a/Cpf1: Structure–function comparisons and implications for genome editing. *Wiley Interdisciplinary Reviews: RNA* **9**, e1481 (2018) doi:10.1002/wrna.1481.
138. Strohkendl, I., Saifuddin, F. A., Rybarski, J. R., Finkelstein, I. J. & Russell, R. Kinetic Basis for DNA Target Specificity of CRISPR-Cas12a. *Mol. Cell* **71**, 816-824.e3 (2018).
139. Jakočiunas, T., Jensen, M. K. & Keasling, J. D. CRISPR / Cas9 advances engineering of microbial cell factories. **34**, 44–59 (2016).
140. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, (2008).
141. Boeke, J. Yeast Art project. http://www.yeastart.org (2018).
142. Franklin, R. E. Molecular configuration in sodium thymonucleate. *Nature* **171**, 740-741 (1953).
143. Crick, F. & Watson, J. Molecular structure of Nucleic Acids. *Nature* **171**, 737-738 (1953).
144. Wilkins, M.H.F., Stokes, A.R., Wilson, H. R. Molecular Structure of Deoxypentose Nucleic Acids. *Nucleic Acids Res.* **171**, 738–740 (1953).
145. Young, E. M. *et al.* Iterative algorithm-guided design of massive strain libraries, applied to itaconic acid production in yeast. *Metab. Eng.* **48**, 33–43 (2018).
146. Roubos, J.A., Pel, H.J., Meijerink, B. Cloning Method. U.S. Patent US9738890B2 (2013).
147. Mandel, M. & Higa, A. Calcium-dependent bacteriophage DNA infection. *J. Mol. Biol.* **53**, 159–162 (1970).

148. Gietz, R. D., Schiestl, R. H., Willems, A. R. & Woods, R. A. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* **11**, 355-360 (1995) doi:10.1002/yea.320110408.

149. Hill, J., Donald, K. A. G. & Griffiths, D. E. DMSO-enhanced whole cell yeast transformation. *Nucleic Acids Res.* **19**, 6688 (1991).

150. Lõoke Marko, Kristjuhan Kersti, K. A. Extraction of genomic DNA from yeast for PCR-based applications. *Eur. PMC Funders Gr.* **50**, 325–328 (2011).

151. Tak, Y. E. *et al.* Inducible and multiplex gene regulation using CRISPR-Cpf1-based transcription factors. *Nat. Methods* **14**, 1163-1166 (2017) doi:10.1038/nmeth.4483.

152. Li, L. *et al.* CRISPR-Cpf1-Assisted Multiplex Genome Editing and Transcriptional Repression in *Streptomyces*. *Appl. Environ. Microbiol.* **84**, e00827-18 (2018).

153. Zhang, X. *et al.* Multiplex gene regulation by CRISPR-ddCpf1. *Cell Discov.* **3**, 1-9(2017) doi:10.1038/celldisc.2017.18.

154. Gietz, R. D. & Schiestl, R. H. Frozen competent yeast cells that can be transformed with high efficiency using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2**, 1–4 (2007).

155. Perez, A. R. *et al.* GuideScan software for improved single and paired CRISPR guide RNA design. *Nat. Biotechnol.* **35**, 347–349 (2017).

156. EauClaire, S. F., Zhang, J., Rivera, C. G. & Huang, L. L. Combinatorial metabolic pathway assembly in the yeast genome with RNA-guided Cas9. *J. Ind. Microbiol. Biotechnol.* **43**, 1001–1015 (2016).

157. Zhou, Y. *et al.* MiYA, an efficient machine-learning workflow in conjunction with the YeastFab assembly strategy for combinatorial optimization of heterologous metabolic pathways in *Saccharomyces cerevisiae*. *Metab. Eng.* **47**, 294–302 (2018).

158. Xu, N., Wei, L. & Liu, J. Recent advances in the applications of promoter engineering for the optimization of metabolite biosynthesis. *World J. Microbiol. Biotechnol.* **35**, 1–10 (2019).

159. Gao, S., Zhou, H., Zhou, J. & Chen, J. Promoter-Library-Based Pathway Optimization for Efficient (2 S)-Naringenin Production from p-Coumaric Acid in *Saccharomyces cerevisiae*. *J. Agric. Food Chem.* **68**, 6884–6891 (2020).

160. Zhou, Y. *et al.* MiYA, an efficient machine-learning work flow in conjunction with the YeastFab assembly strategy for combinatorial optimization of heterologous metabolic pathways in *Saccharomyces cerevisiae*. *Metab. Eng.* **47**, 294–302 (2018).

161. Li, C., Swofford, C. A. & Sinskey, A. J. Modular engineering for microbial production of carotenoids. *Metabolic Engineering Communications* **10**, e00118 (2020) doi:10.1016/j.mec.2019.e00118.

162. De Nadal, E., Ammerer, G. & Posas, F. Controlling gene expression in response to stress. *Nat. Rev. Genet.* **12**, 833–845 (2011).

163. Agrawal, N. *et al.* RNA Interference: Biology, Mechanism, and Applications. *Microbiol. Mol. Biol. Rev.* **67**, 657-685 (2003) doi:10.1128/mmbr.67.4.657-685.2003.

164. Beerli, R. R. & Barbas, C. F. Engineering polydactyl zinc-finger transcription factors. *Nature Biotechnology* **20**, 135-141 (2002) doi:10.1038/nbt0202-135.

165. Khalil, A. S. *et al.* A synthetic biology framework for programming eukaryotic transcription functions. *Cell* **150**, 647-658 (2012) doi:10.1016/j.cell.2012.05.045.

166. Meaker, G. A., Hair, E. J. & Gorochowski, T. E. Advances in engineering CRISPR-Cas9 as a molecular

swiss army knife. *Synth. Biol.* **5**, ysaa021 (2020).

167. Li, S. *et al.* Genome-Wide CRISPRi-Based Identification of Targets for Decoupling Growth from Production. *ACS Synth. Biol.* **9**, 1030-1040 (2020) doi:10.1021/acssynbio.9b00143.
168. Farzadfard, F., Perli, S. D. & Lu, T. K. Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. *ACS Synth. Biol.* **2**, 604-613 (2013) doi:10.1021/sb400081r.
169. Schreiber-Agus, N. *et al.* An amino-terminal domain of Mxi1 mediates anti-myc oncogenic activity and interacts with a homolog of the Yeast Transcriptional Repressor SIN3. *Cell* **80**, 777-786 (1995) doi:10.1016/0092-8674(95)90356-9.
170. Witzgall, R., O'Leary, E., Leaf, A., Onaldi, D. & Bonventre, J. V. The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 4514-4518 (1994) doi:10.1073/pnas.91.10.4514.
171. Shmakov, S. *et al.* Diversity and evolution of class 2 CRISPR-Cas systems. *Nat. Rev. Microbiol.* **15**, 169-182 (2017) doi:10.1038/nrmicro.2016.184.
172. Ottoz, D. S. M. & Rudolf, F. Constitutive and regulated promoters in yeast: How to design and make use of promoters in *S. cerevisiae*. in *Synthetic Biology: Parts, Devices and Applications* **10**, 14 (2018). doi:10.1002/9783527688104.ch6.
173. Choi, S. Y. & Woo, H. M. CRISPRi-dCas12a: A dCas12a-Mediated CRISPR Interference for Repression of Multiple Genes and Metabolic Engineering in Cyanobacteria. *ACS Synth. Biol.* **9**, 2351-2361 (2020) doi:10.1021/acssynbio.0c00091.
174. Oldenburg, K. R., Vo, K. T., Michaelis, S. & Paddon, C. Recombination-mediated PCR-directed plasmid construction in vivo in yeast. *Nucleic Acids Res.* **25**, 451–452 (1997).
175. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343-345 (2009) doi:10.1038/nmeth.1318.
176. Mircetic, J. *et al.* Purified Cas9 Fusion Proteins for Advanced Genome Manipulation. *Small Methods* **1**, 1600052 (2017).
177. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009) doi:10.1093/bioinformatics/btp324.
178. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **5**, 2078-2079 (2009) doi:10.1093/bioinformatics/btp352.
179. McMillan, J., Lu, Z., Rodriguez, J. S., Ahn, T. H. & Lin, Z. YeasTSS: An integrative web database of yeast transcription start sites. *Database* (2019) doi:10.1093/database/baz048.
180. Castillo-Hair, S. M. *et al.* FlowCal: A User-Friendly, Open Source Software Tool for Automatically Converting Flow Cytometry Data from Arbitrary to Calibrated Units. *ACS Synth. Biol.* **5**, 774-780. (2016) doi:10.1021/acssynbio.5b00284.
181. Bailey, R., Madden, K. T. & Trueheart, J. Production of carotenoids in oleaginous yeast and fungi. U.S. Patent US8288149B2 (2006).
182. Zhao, P. *et al.* One-step homozygosity in precise gene editing by an improved CRISPR/Cas9 system. *Cell Research* **26**, 633-636 (2016) doi:10.1038/cr.2016.46.
183. Beal, J. *et al.* Quantification of bacterial fluorescence using independent calibrants. *PLoS One* **13**, e0199432 (2018) doi:10.1371/journal.pone.0199432.

184. Roggenkamp, E. *et al.* Tuning CRISPR-Cas9 gene drives in *Saccharomyces cerevisiae*. *G3 Genes, Genomes, Genet.* **8**, 999-1018 (2018) doi:10.1534/g3.117.300557.

185. Liu, Y. *et al.* Engineering cell signaling using tunable CRISPR-Cpf1-based transcription factors. *Nat. Commun.* **8**, 1-8 (2017) doi:10.1038/s41467-017-02265-x.

186. Currin, A. *et al.* Highly multiplexed, fast and accurate nanopore sequencing for verification of synthetic DNA constructs and sequence libraries. *Synth. Biol.* **4**, 1–8 (2019).

187. Jensen, E. D. *et al.* Transcriptional reprogramming in yeast using dCas9 and combinatorial gRNA strategies. *Microb. Cell Fact.* **16**, 1-16 (2017) doi:10.1186/s12934-017-0664-2.

188. Kleinstiver, B. P. *et al.* Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat. Biotechnol.* **34**, 869–874 (2016).

189. Swarts, D. C., van der Oost, J. & Jinek, M. Structural Basis for Guide RNA Processing and Seed-Dependent DNA Targeting by CRISPR-Cas12a. *Mol. Cell* **66**, 221-233.e4 (2017).

190. McCarty, N. S., Graham, A. E., Studená, L. & Ledesma-Amaro, R. Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nature Communications* **11**, 1-13 (2020) doi:10.1038/s41467-020-15053-x.

191. Hofmann, A. *et al.* A tightly regulated and adjustable CRISPR-dCas9 based AND gate in yeast. *Nucleic Acids Res.* **47**, 509-520 (2019) doi:10.1093/nar/gky1191.

192. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827-832 (2013) doi:10.1038/nbt.2647.

193. Ng, H. & Dean, N. Dramatic Improvement of CRISPR/Cas9 Editing in Candida albicans by Increased Single Guide RNA Expression. *mSphere* **2**, e00385-16 (2017) doi:10.1128/msphere.00385-16.

194. McCarty, N. S., Shaw, W. M., Ellis, T. & Ledesma-Amaro, R. Rapid Assembly of gRNA Arrays via Modular Cloning in Yeast. *ACS Synth. Biol.* **8**, 906-910 (2019) doi:10.1021/acssynbio.9b00041.

195. Liao, C. *et al.* Modular one-pot assembly of CRISPR arrays enables library generation and reveals factors influencing crRNA biogenesis. *Nat. Commun.* **10**, 1-14 (2019) doi:10.1038/s41467-019-10747-3.

196. Yabuzaki, J. Carotenoids Database: Structures, chemical fingerprints and distribution among organisms. *Database* (2017) doi:10.1093/database/bax004.

197. Wu, T. *et al.* Membrane engineering - A novel strategy to enhance the production and accumulation of β-carotene in *Escherichia coli*. *Metab. Eng.* **43**, 85-91 (2017) doi:10.1016/j.ymben.2017.07.001.

198. Brameier, M., Krings, A. & MacCallum, R. M. NucPred - Predicting nuclear localization of proteins. *Bioinformatics* **23**, 1159–1160 (2007).

199. Kosugi, S., Hasebe, M., Tomita, M. & Yanagawa, H. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 10171–10176 (2009).

200. Lin, J., Mondal, A. M., Liu, R. & Hu, J. Minimalist ensemble algorithms for genome-wide protein localization prediction. **13**, 1-12 (2012).

201. Popken, P., Ghavami, A., Onck, P. R., Poolman, B. & Veenhoff, L. M. Size-dependent leak of soluble and membrane proteins through the yeast nuclear pore complex. *Mol. Biol. Cell* **26**, 1386–1394 (2015).

202. Timney, B. L. *et al.* Simple rules for passive diffusion through the nuclear pore complex. *J. Cell Biol.* **215**, 57–76 (2016).

203. Jones, D. L. *et al.* Kinetics of dCas9 target search in *Escherichia coli*. *Science* **357**, 1420–1424 (2017).
204. Ma, H. *et al.* CRISPR-Cas9 nuclear dynamics and target recognition in living cells. *J. Cell Biol.* **214**, 529–537 (2016).
205. Kim, S. K. *et al.* Efficient Transcriptional Gene Repression by Type V-A CRISPR-Cpf1 from Eubacterium eligens. *ACS Synth. Biol.* **6**, 1273–1282 (2017).
206. Toth, E. *et al.* Improved LbCas12a variants with altered PAM specificities further broaden the genome targeting range of Cas12a nucleases. *Nucleic Acids Res.* **48**, 3722–3733 (2020).
207. Shmakov, S. *et al.* Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. *Mol. Cell* **60**, 385-397 (2015) doi:10.1016/j.molcel.2015.10.008.
208. Makarova, K. S. *et al.* Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67–83 (2020).
209. East-Seletsky, A., O'Connell, M. R., Burstein, D., Knott, G. J. & Doudna, J. A. RNA Targeting by Functionally Orthogonal Type VI-A CRISPR-Cas Enzymes. *Mol. Cell* **66**, 373-383.e3 (2017).
210. Mendez-Mancilla, A. *et al.* Chemically modified guide RNAs enhance CRISPR-Cas13 knockdown in human cells. *bioRxiv* (2021).
211. Kulkarni, A. *et al.* Programmable CRISPR interference for gene silencing using Cas13a in mosquitoes. **8**, 30 (2020).
212. Jing, X. *et al.* Implementation of the CRISPR-Cas13a system in fission yeast and its repurposing for precise RNA editing. *Nucleic Acids Res.* **46**, e90-e90 (2018) doi:10.1093/nar/gky433.
213. Lorenz, R., Bernhart, S. H., Zu Siederdissen, C. H., Tafer, H., Flamm, C., Stadler, P. F., & Hofacker, I. L. ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **1**, 1–14 (6AD).
214. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, 1–12 (2002).
215. Ugolini, S. & Bruschi, C. V. The red/white colony color assay in the yeast *Saccharomyces cerevisiae*: Epistatic growth advantage of white ade8-18, ade2 cells over red ade2 cells. *Curr. Genet.* **30**, 485-492 (1996) doi:10.1007/s002940050160.
216. Morano, K. A., Grant, C. M. & Moye-Rowley, W. S. The response to heat shock and oxidative stress in *Saccharomyces cerevisiae*. *Genetics* **190**, 1157–1195 (2012).
217. Toledano, M. B., Delaunay, A., Biteau, B., Spector, D. & Azevedo, D. *Yeast stress responses*. (Springer Berlin Heidelberg, 2003). doi:10.1007/3-540-45611-2.
218. Wessels, H. H. *et al.* Massively parallel Cas13 screens reveal principles for guide RNA design. *Nat. Biotechnol.* **38**, 722–727 (2020).
219. Labun, K. *et al.* CHOPCHOP v3: Expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* **47**, W171–W174 (2019).
220. Huynh, N., Depner, N., Larson, R. & King-Jones, K. A versatile toolkit for CRISPR-Cas13-based RNA manipulation in Drosophila. *Genome Biol.* **21**, 1–29 (2020).
221. Charles, E. J. *et al.* Engineering improved Cas13 effectors for targeted post-transcriptional regulation of gene expression. *bioRxiv Prepr.* (2021).
222. Bak, G. *et al.* On-off controllable RNA hybrid expression vector for yeast three-hybrid system. *BMB Rep.* **43**, 110–114 (2010).

223. Momen-Roknabadi, A., Oikonomou, P., Zegans, M. & Tavazoie, S. An inducible CRISPR interference library for genetic interrogation of *Saccharomyces cerevisiae* biology. *Commun. Biol.* **3**, 1–12 (2020).
224. Jagtap, S. S., Bedekar, A. A. & Rao, C. V. Quorum Sensing in Yeast. *ACS Symp. Ser.* **1374**, 235–250 (2020).
225. Özcan, A. *et al.* Programmable RNA targeting with the single-protein CRISPR effector Cas7-11. *Nature* **597**, 720–725 (2021).
226. Dahl, R. H. *et al.* Engineering dynamic pathway regulation using stress-response promoters. *Nat. Biotechnol.* **31**, 1039–1046 (2013).
227. Jin, H. L., Mitchell, R. J., Byoung, C. K., Cullen, D. C. & Man, B. G. A cell array biosensor for environmental toxicity analysis. *Biosens. Bioelectron.* **21**, 500–507 (2005).
228. Carlquist, M. *et al.* Physiological heterogeneities in microbial populations and implications for physical stress tolerance. *Microb. Cell Fact.* **11**, 1–13 (2012).
229. Polizzi, K. M. & Kontoravdi, C. Genetically-encoded biosensors for monitoring cellular stress in bioprocessing. *Current Opinion in Biotechnology* **31**, 50-56 (2015) doi:10.1016/j.copbio.2014.07.011.
230. Karp, P. D. *et al.* The BioCyc collection of microbial genomes and metabolic pathways. *Brief. Bioinform.* **20**, 1085–1093 (2018).
231. Garalde, D. R. *et al.* Highly parallel direct RN A sequencing on an array of nanopores. *Nat. Methods* **15**, 201–206 (2018).
232. Joshua, I. M. & Höfken, T. From lipid homeostasis to differentiation: Old and new functions of the zinc cluster proteins Ecm22, Upc2, Sut1 and Sut2. *Int. J. Mol. Sci.* **18**, 1–17 (2017).
233. Shakoury-Elizeh, M. *et al.* Metabolic Response to Iron Deficiency in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **285**, 14823–14833 (2010).
234. Bu, X. *et al.* Engineering endogenous ABC transporter with improving ATP supply and membrane flexibility enhances the secretion of β-carotene in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **13**, 1–14 (2020).
235. Protchenko, O. *et al.* Three Cell Wall Mannoproteins Facilitate the Uptake of Iron in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **276**, 49244–49250 (2001).
236. Köffel, R. & Schneiter, R. Yeh1 constitutes the major steryl ester hydrolase under heme-deficient conditions in *Saccharomyces cerevisiae*. *Eukaryot. Cell* **5**, 1018–1025 (2006).
237. Schuller, A., Rödel, G. & Ostermann, K. Tuning the sensitivity of the PDR5 promoter-based detection of diclofenac in yeast biosensors. *Sensors* **17**, 1506 (2017).
238. Arthington-Skaggs, B. A., Crowell, D. N., Yang, H., Sturley, S. L. & Bard, M. Positive and negative regulation of a sterol biosynthetic gene (ERG3) in the post-squalene portion of the yeast ergosterol pathway. *FEBS Lett.* **392**, 161–165 (1996).
239. Hickman, M. J. & Winston, F. Heme Levels Switch the Function of Hap1 of *Saccharomyces cerevisiae* between Transcriptional Activator and Transcriptional Repressor . *Mol. Cell. Biol.* **27**, 7414–7424 (2007).
240. Cai, P. L., He, X. P., Liu, N. & Zhang, B. R. Effect of over-expression of sterol C-22 desaturase on ergosterol production in yeast strains. *Acta Microbiol. Sin.* **47**, 274–279 (2007).

241. Bröker, J. N., Müller, B., van Deenen, N., Prüfer, D. & Schulze Gronover, C. Upregulating the mevalonate pathway and repressing sterol synthesis in *Saccharomyces cerevisiae* enhances the production of triterpenes. *Appl. Microbiol. Biotechnol.* **102**, 6923–6934 (2018).

242. Lam, A. J. *et al.* Improving FRET dynamic range with bright green and red fluorescent proteins. *Nat. Methods* **9**, 1005–1012 (2012).

243. S., K. A., Curran, K. A. & Alper, H. S. Characterization of plasmid burden and copy number in *Saccharomyces cerevisiae* for optimization of metabolic engineering applications. *FEMS Yeast Res.* **13**, 107-116 (2013).

244. Reyes, L. H., Gomez, J. M. & Kao, K. C. Improving carotenoids production in yeast via adaptive laboratory evolution. *Metab. Eng.* **21**, 26–33 (2014).

245. Grilly, C., Stricker, J., Pang, W. L., Bennett, M. R. & Hasty, J. A synthetic gene network for tuning protein degradation in *Saccharomyces cerevisiae*. *Mol. Syst. Biol.* **3**, 1–5 (2007).

246. Peng, B., Nielsen, L. K., Kampranis, S. C. & Vickers, C. E. Engineered protein degradation of farnesyl pyrophosphate synthase is an effective regulatory mechanism to increase monoterpene production in *Saccharomyces cerevisiae*. *Metab. Eng.* **47**, 83–93 (2018).

247. Cameron, D. E. & Collins, J. J. Tunable protein degradation in bacteria. *Nat. Biotechnol.* **32**, 1276–1281 (2014).

248. Fernandez-Rodriguez, J. & Voigt, C. A. Post-translational control of genetic circuits using *Potyvirus* proteases. *Nucleic Acids Res.* **44**, 6493–6502 (2016).

249. Martínez, V. *et al.* CRISPR/Cas9-based genome editing for simultaneous interference with gene expression and protein stability. *Nucleic Acids Res.* **45**, 1–8 (2017).

250. Flagfeldt, D. B., Siewers, V., Huang, L. & Nielsen, J. Characterization of chromosomal integration sites for heterologous gene expression in *Saccharomyces cerevisiae*. *Yeast* **26**, 545-551 (2009) doi:10.1002/yea.1705.

251. Gootenberg, J. S. *et al.* Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. *Science* **360**, 439–444 (2018).

252. Martínez-Cámara, S., Rubio, S., del Río, H., Rodríguez-Sáiz, M. & Barredo, J.-L. Lycopene Production by Mated Fermentation of *Blakeslea trispora*. in *Microbial Carotenoids: Methods and Protocols* 257–68 (Springer New York, 2018). doi:10.1007/978-1-4939-8742-9_15.

11 List of the hazardous substances used by GHS

Substance	Pictogram	Signal word	Hazard statement	Precautionary statement
Acetonitrile		Danger	H225, H302+H312+H332, H319	P210, P280, P301+P312, P303+P361+P353, P304+P340+P312, P305+P351+P338
Ampicillin		Danger	H315, H317, H319, H334, H335	P280, P302+P352, P305+P351+P338
Butylated hydroxytoluene		Warning	H410	P273, P391, P501
Dithiothreitol		Warning	H302, H412	P273, P301+P312+P330
Ethanol		Danger	H225, H319	P210, P233, P240, P241, P242, P305+P351+P338
Ethyl acetate		Danger	H225, H319, H336	P210, P233, P240, P241, P242, P305+P351+P338
Ethylenediamine-tetraacetic acid		Warning	H373	P260, P314, P501
G418 disulfate salt		Danger	H317, H334	P261, P272, P280, P284, P302+P352, P304+P340+P312
Isopropanol		Danger	H225, H319, H336	P210, P233, P240, P241, P242, P305+P351+P338
Neomycin		Danger	H317, H334	P261, P272, P280, P284, P302+P352, P304+P340+P312
Nourseothricin		Warning	H302	P301+P312+P330
Tetrahydrofuran		Danger	H225, H302, H319, H335, H336, H351	P201, P202, P210, P301+P312, P305, P351, P338, P308, P313

12 Acknowledgments

After MSc degree I wanted to learn how to work with yeast and what CRISPR is about. Through this PhD project I became proficient with molecular manipulation of *S. cerevisiae*, worked with 5 different CRISPR systems and learned many other lab methods and instruments which resulted in the presented thesis. This could not be achieved without guidance, help and support of many people whom I want to sincerely thank.

To **René Verwaal** for all the help with arranging my stay in the Netherlands in the first weeks, the effort you put to smoothly start this project, submitting manuscripts in the late night and during holidays, and experiments run in the middle of the long weekend.

To **Hans Roubos** for shaping the project, our brainstorming sessions and countless discussions.

To **Thomas Gorochowski** for warm welcoming me in your group, always being available and open to answer any of my questions and concerns and introducing me to Nanopore sequencing (which is currently my main scientific focus). Thanks to your guidance I found my way in this project.

To **Zoya Ignatova** for critical discussions on science and sharing your expertise on RNA which helped in understanding the function of CRISPR-Cas13.

Brenda Vonk for being the most gentle and supportive teacher when I was making my first steps in molecular biology lab. I have learnt so much from you! To **Rianne van der Hoek** for endless advice on yeast protocols and sequencing and all the chat we had over pipettes. To **Bianca Gielesen** for your guidance with FACS and patience every time when I was concerned over the smallest noise FACS made. To **Priscilla Zwartjens** for critical analysis of problems I faced and fruitful discussions. To **Joep Schmitz** for always finding time to discuss my project, having a good piece of advice and enabling the most fancy experiments to happen.

Many of DSM colleagues helped me with bioinformatics – **Jurgen Nijkamp**, **Zi Di**, **Abel Folch-Fortuny** and **Irsan Kooi** thank you for that!

So many of you made me feel welcomed at DSM – **Moniek**, **Jos**, **Marco**, **Karin**, **Marcel**, **Nathalie**, **Sam** (thanks for always being available to talk about experiments) **Lisette** (thanks for moving my plates when I travelled!), **Kamesh**, **Jeffrey** (thanks for all the yeast art!), **Steven**, **Harmen**, **Laurence**, **Eric**, **Paul**, **Martine**, **Wout**, **Aad**, **Janny**, **Francine**, **Diny**, **Ingrid**, **Sandra**, **Onno**, **Arjen**, **Sander**, **Amanda**, **Wouter**, **Alex**, **Laura**, **Francisca**, **Willi** (thanks for all the chats about science!), **Segio**, **Linsey**, **Kevin**, **Leonie**, **Leonor**, **Emiel**. I would like to thank everyone for all the chit chats and showing me around in the lab.

Special thanks to **Lisa Norte** for mentoring and support in arranging PhD related matters at DSM. Also, I appreciate help from **Herman Pel in** the last months of my PhD.

To **Naoual Ouazzani Chadi** – you did a truly marvellous job on Cas13 and we could not understand this system without your dedication. It was such a great time when you stayed in the lab for the internship.

To MSCA fellows at DSM: **Federica Calderaro** for your endless support, countless times of being my lab buddy and for being there for me. **Diego Lopez Barreiro** for always listening to my PhD dramas despite of being well done with this. **Esther Volz (Hegel)** and **Stefano Grasso** for introducing me to the world of PhD.

Finally, to **my parents** and my brother **Kuba** for the support, your faith and love.

To **Andrea**, I could not make it without you.

13 Supplementary Information

List of SI Tables

SI Table 2-1 List of strains used in this study....13-6
SI Table 2-2 Gradient applied for separation of carotenoids using UHPLC....13-7
SI Table 2-3. List of primers used in the study....13-8
SI Table 2-4. Translations of *crtYB* sequence obtained in nanopore sequencing....13-9
SI Table 3-1 Table of Materials....13-12
SI Table 3-2. Single crRNA array for LbCas12a containing homology with plasmid pRN1120....13-15
SI Table 3-3. Primers sequences....13-16
SI Table 3-4. Design of constructed strains....13-18
SI Table 3-5. Sequences of donor DNA expression cassettes and flaking regions....13-19
SI Table 4-1. Genotype of strains constructed in this study....13-30
SI Table 5-1. Integration sites used for strains construction....13-31
SI Table 5-2. List of strains used in this study....13-32
SI Table 5-3. List of plasmids used in this study....13-34
SI Table 5-4. List of primers used in this study....13-35
SI Table 5-5. Features of sgRNA and crRNA design....13-42
SI Table 5-6. List of spacers used for singleplex genome editing and CRISPR interference experiments....13-43
SI Table 5-7. List of spacers in array used for multiplex genome editing and CRISPR interference experiments....13-44
SI Table 5-8. Order of spacers in single crRNA arrays used for multiplex CRISPR interference experiments....13-45
SI Table 5-9. Architecture of gRNAs expressed from RNAP II and III promoters....13-46
SI Table 5-10. Accession numbers of sequences submitted to GenBank....13-48
SI Table 6-1. List of primers used in this study....13-59
SI Table 6-2. Features of Cas13a and Cas13b gRNA design....13-64
SI Table 6-3. List of spacers used for singleplex genome editing and evaluation of CRISPR-Cas13 system....13-65
SI Table 6-4. Sequences of donor DNA used to remove NLS and construct strains CC005,6,11,12....13-66
SI Table 7-1. List of primers used in this study....13-74
SI Table 7-2. List of gRNAs used in this study....13-76
SI Table 7-3. Differential expression analysis of the carotenogenic and wild type strains....13-79
SI Table 8-1. List of primers used in this study....13-81

List of SI Figures

SI Figure 2-1. Phytoene, lycopene and β-carotene production profiles in batch fermentation by strains CAR-001, CAR-002, CAR-005, CAR-009, CAR-020, CAR-021 13-10
SI Figure 2-2. Colonies subjected to nanopore sequencing of strains CAR-002 and CAR-020. 13-11
SI Figure 4-1. Genetic instability of strains CAR-002 and CAR-034 after fermentation in mineral medium 13-29
SI Figure 5-1. Localization of dCas12a-eGFP fusion in *S. cerevisiae* cells 13-51
SI Figure 5-2. Overview on the gRNAs used for gene downregulation 13-52
SI Figure 5-3. Growth and fluorescence of strains transformed with gRNAs expressed from RNAP II and III promoters 13-53
SI Figure 5-4. Absolute fluorescence of eGFP achieved with plasmids testing NLS position and number in strain FR003 and FR009. 13-54
SI Figure 5-5. Absolute fluorescence of eGFP achieved with strains DC002-013 testing repression domains 13-55
SI Figure 5-6. Absolute fluorescence of eGFP achieved by targeting dCas12a E925A Mxi1 to promoters controlling *eGFP* or *eGFP* ORF 13-56
SI Figure 5-7. Absolute fluorescence of eGFP achieved with gRNAs expressed from RNAP II and III promoters 13-57
SI Figure 5-8. Absolute fluorescence of eGFP, BFP and mCherry achieved in multiplex downregulation using single crRNA array. 13-58
SI Figure 6-1. Visual representation of gRNAs used with Cas13a and Cas13b 13-67
SI Figure 6-2. Expression of LshCas13a or LwaCas13a and gRNAs in *S. cerevisiae*. 13-68
SI Figure 6-3. Expression of BzoCas13b or PspCas13b and gRNAs in *S. cerevisiae* 13-69
SI Figure 6-4. Sugar concentration in cultures of strain LwaCas13a NLS 13-70
SI Figure 6-5. Growth of strain LwaCas13a NLS under inducive and repressive condition 13-71
SI Figure 6-6. Live cells in culture grown for 6 h, 14 h, 18 h, 20 h, 24h and 30 h from the induction of LwaCas13a NLS 13-72
SI Figure 6-7. Dead staining of cells with active LwaCas13a NLS system exposed to heat and osmotic shock. 13-73
SI Figure 7-1. Growth curve and metabolic profile of cultures subjected to the RNA-seq 13-80
SI Figure 8-1. Instability of β-carotene biosynthesis pathway during growth on rich agar plates 13-82
SI Figure 8-2. Instability of β-carotene biosynthesis pathway during growth on mineral agar plates 13-83

2 Supplementary information to Chapter 2

SI Table 2-1 List of strains used in this study

Strain name	Genotype
CEN.PK113–7D	MAT*a MAL2–8c SUC2*
CAR-001	MAT*a MAL2-8c SUC2 ura3-52* + YIplac211 TDH3p - crtYB - CYC1t; TDH3p - crtI - CYC1t; TDH3p - crtE* - CYC1t (prototrophic strain)
CAR-002	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sc_FBA1p - Xden_crtE - Sc_TDH3t - conA - Sc_TEF1p - Xden_crtYB - Sc_PDC1t - conB - Kl_ENO1p - Xden_crtI - Sc_TAL1t - conC - INT1
CAR-005	MAT*a MAL2–8c SUC2* + INT1 - con5 - Kl_PGK1p - Xden_crtE - Sc_TDH3t - conA - Kl_TEF2p - Xden_crtYB - Sc_PDC1t - conB - Sc_ACT1p - Xden_crtI - Sc_TAL1t - conC - INT1
CAR-009	MAT*a MAL2–8c SUC2* + INT1 - con5 - Kl_TDH2p - Xden_crtE - Sc_TDH3t - conA - Kl_YDR1p - Xden_crtYB - Sc_PDC1t - conB - Sc_PRE3p - Xden_crtI - Sc_TAL1t - conC - INT1
CAR-020	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sc_FBA1p - Xden_crtE - Sc_TDH3t - conA - Sc_TEF1p - Xden_crtYB - Sc_PDC1t - conB - Kl_ENO1p - Xden_crtI - Sc_TAL1t - conC - INT1 + INT2 -con52 - Sc_FBA1p - tHMG1 - Sc_ADH2t - con2A - Kl_ENO1p - Xden_crtI - Sc_TAL1t - con23 – INT2
CAR-021	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sc_FBA1p - Xden_crtE - Sc_TDH3t - conA - Sc_TEF1p - Xden_crtYB - Sc_PDC1t - conB - Kl_ENO1p - Xden_crtI - Sc_TAL1t - conC - INT1 + INT2 -con52 - Sc_FBA1p - tHMG1 - Sc_ADH2t - con2A - Kl_ENO1p - Xden_crtI - Sc_TAL1t - con2B – Sc_TEF1p - Xden_crtYB - Sc_PDC1t - con23 - INT2

SI Table 2-2 Gradient applied for separation of carotenoids using UHPLC

Time (min)	Flow (mL/min)	Mobile phase A (ethyl acetate) %	Mobile phase B (water) %	Mobile phase C (acetonitrile) %
0	1	0	20	80
3	1	20	0	80
4.5	1	80	0	20
5	1	0	0	100
6	1	0	0	100
6.5	1	0	20	80
7	1	0	20	80

SI Table 2-3. List of primers used in the study

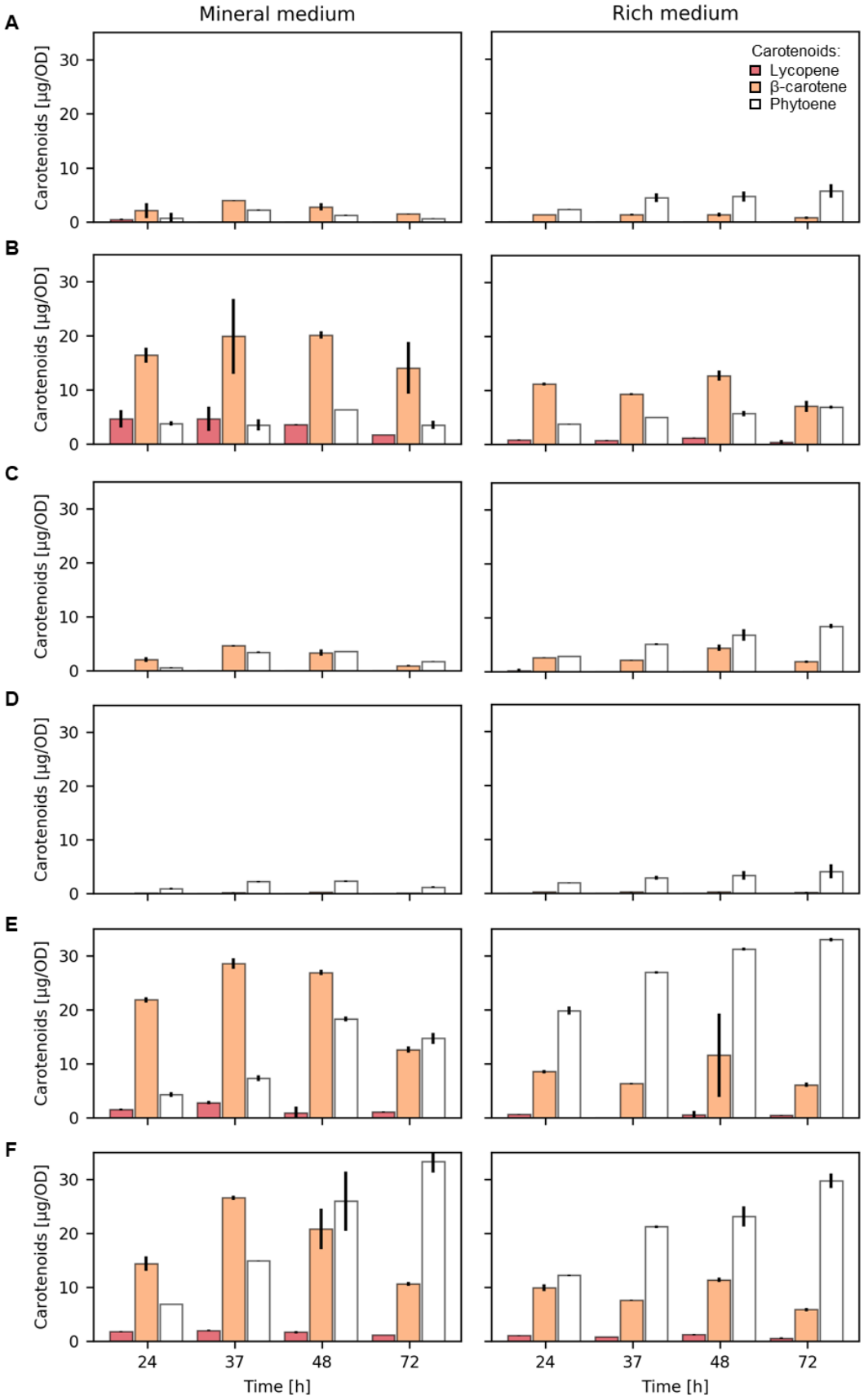
Name	Sequence	Purpose
pKC001	**GGAGTAATAC**CTATGGTGCATTTACCGACG	Amplification of expression cassette in INT1 of strains CAR-002 clone 1 and attachment of barcode (FW)
pKC002	**ACATCAGATC**CTATGGTGCATTTACCGACG	Amplification of expression cassette in INT1 of strains CAR-002 clone 2 and attachment of barcode (FW)
pKC003	**GAACTTTAAG**CTTATCAAGCCTCTCACAGAC	Amplification of expression cassette in INT1 of strains CAR-002 clone 1 and attachment of barcode (RV)
pKC004	**ATGGTACCAG**CTTATCAAGCCTCTCACAGAC	Amplification of expression cassette in INT1 of strains CAR-002 clone 2 and attachment of barcode (RV)
pKC005	**CCGAGTCAAT**CTATGGTGCATTTACCGACG	Amplification of expression cassette in INT1 of strains CAR-020 clone 1 and attachment of barcode (FW)
pKC006	**GAAGCTCGCA**CTATGGTGCATTTACCGACG	Amplification of expression cassette in INT1 of strains CAR-020 clone 2 and attachment of barcode (FW)
pKC007	**TTACAGTGTT**CTTATCAAGCCTCTCACAGAC	Amplification of expression cassette in INT1 of strains CAR-020 clone 1 and attachment of barcode (RV)
pKC008	**TACCGCATCT**CTTATCAAGCCTCTCACAGAC	Amplification of expression cassette in INT1 of strains CAR-020 clone 2 and attachment of barcode (RV)

Barcode sequence in bold, whereas binding region in normal font.

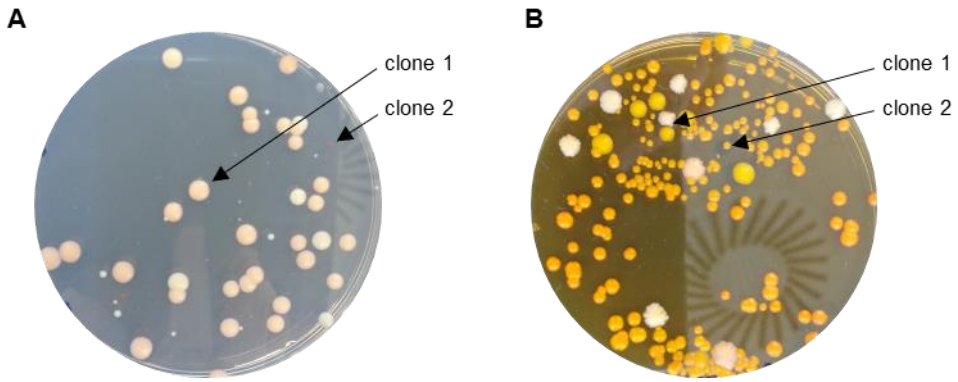
SI Table 2-4. Translations of *crtYB* sequence obtained in nanopore sequencing

Translation of *crtYB* sequence of strain CAR-002 clone #2 and CAR-020 clone #2
MTALAYYQIHLIYTLPILGLLGLLTSPILTKFDIYKISILVFIAFSATTPWDSWIIRNGAWTYPSAESGQGVF GTFLDVPYEEYAFFVIQTVITGLVYVLATRHLLPSLALPKTRSSALSLALKALIPLPIIYLFTAHPSPSPDPL VTDHYFYMRALSLLITPPTMLLAALSGEYAFDWKSGRAKSTIAAIMIPTVYLIWVDYVAVGQDSWSINDEKIV GWRLGGVLPIEEAMFFLLTNLMIVLGLSACDHTQALYLLHGRTIYGNKKMPSSFPLITPPVLSLFFSSRPYSS QPKRDLELAVKLLEKKSRSFFVASAGFPSEVRERLVGLYAFCRVTDDLIDSPEVSSNPHATIDMVSDFLTLLF GPPLHPSQPDKILSSPLLPPSHPSRPTGMYPLPPPPSLSPAELVQFLTERVPVQYHFAFRLLAKLQGLIPRYP LDELLRGYTTDLIFPLSTEAVQARKTPIETTADLLDYGLCVAGSVAELLVYVSWASAPSQVPATIEEREAVLV ASREMGTALQLVNIARDIKGDATEGRFYLPLSFFGLRDESKLAIPTDWTEPRPQDFDKLLSLSPSSTLPSSNA SESFRFEWKTYSLPLVAYAEDLAKHSYKGIDRLPTEVQAGMRAACASYLLIGREIKVVWKGDVGERRTVAGWR RVRKVLSVVMSGWEGQ
Translation of crtYB sequence of strain CAR-002 clone #1
MTALAYYQIHLIYTLPILGLLGLLTSPILTKFDIYKISILVFIAFSATTPWDSWIIRNGAWTYPSAESGQGVF GTFLDVPYEEYAFFVIQTVITGLVYVLATRHLLPSLALPKTRSSALSLALKALIPLPIIYLFTAHPSPSPDPL VTDHYFYMRALSLLITPPTMLLAALSG*YAFDWKSGRAKSTIAAIMIPTVYLIWVDYVAVGQDSWSINDEKIV GWRLGGVLPIEEAMFFLLTNLMIVLGLSACDHTQALYLLHGRTIYGNKKMPSSFPLITPPVLSLFFSSRPYSS QPKRDLELAVKLLEKKSRSFFVASAGFPSEVRERLVGLYAFCRVTDDLIDSPEVSSNPHATIDMVSDFLTLLF GPPLHPSQPDKILSSPLLPPSHPSRPTGMYPLPPPPSLSPAELVQFLTERVPVQYHFAFRLLAKLQGLIPRYP LDELLRGYTTDLIFPLSTEAVQARKTPIETTADLLDYGLCVAGSVAELLVYVSWASAPSQVPATIEEREAVLV ASREMGTALQLVNIARDIKGDATEGRFYLPLSFFGLRDESKLAIPTDWTEPRPQDFDKLLSLSPSSTLPSSNA SESFRFEWKTYSLPLVAYAEDLAKHSYKGIDRLPTEVQAGMRAACASYLLIGREIKVVWKGDVGERRTVAGWR RVRKVLSVVMSGWEGQ
Translation of *crtYB* sequence of strain CAR-020 clone #1
MTALAYYQIHLIYTLPILGLLGLLTSPILTKFDIYKISILVFIAFSATTPWDSWIIRNGAWTYPSAESGQGVF GTFLDVPYEEYAFFVIQTVITGLVYVLATRHLLPSLALPKTRSSALSLALKALIPLPIIYLFTAHPSPSPDPL VTDHYFYMRALSLLITPPTMLLAALSGEYAFDWKSGRAKSTIAAIMIPTVYLIWVDYVAVGQDSWSINDEKIV GWRLGGVLPIEEAMFFLLTNLMIVLGLSACDHTQALYLLHGRTIYGNKKMPSSFPLITPPVVLVLLLQTILLP TKERFGIGCQVVGKEVQIFLRCFCRFPI*SQRKIGWFVRFLSCHR*LD*LSRSFLQPTRYH*HGFRFLDFIIR SSIAPISTRQDFVFSIITTFPPIQTNWYVPITTTSIFVSS*IGPILD*TCPSSIPLRFQIVGQIARFDSKIPI G*IIERLHH*LDLPIVH*SRPS*KDPN*NYC*LVGLRFVCCRFCC*IVGLRFLGFRSIPSSSYY*RKRSCFGR LS*NGYRFAIGQHCQRYQG*RY*R*ILLAIVFLWFER*IQIGHSN*LD*TKTSRFRQIVVSISIFHFTIL*RF *ILQIRMEDLLFAIGCLR*RFG*ALLQGY*QITN*SPSWYESCLCFLLVDWS*NQGCLEG*CR*KKNRCWLEK SQKGFVCCHVRLGRS

Stop codons highlighted in red.



SI Figure 2-1. Phytoene, lycopene and β-carotene production profiles in batch fermentation by strains. A. CAR-001, **B**. CAR-002, **C**. CAR-005, **D**. CAR-009, **E**. CAR-020, **F**. CAR-021 in mineral (left) and rich medium (right). Data represents mean ±standard deviation (n=2).



SI Figure 2-2. Colonies subjected to nanopore sequencing of strains A. CAR-002 and **B**. CAR-020.

3 Supplementary information to Chapter 3

SI Table 3-1. Table of Materials

Name of Material or Equipment	Company	Catalog Number	Comments
Chemicals specific for the protocol			
1 Kb Plus DNA Ladder	Thermo Fisher Scientific	10787018	Electrophoresis
Ampicillin sodium salt	Sigma Aldrich	A9518	Selection of *E. coli* transformants
BsaI-HF (20 U/µl)	New England BioLabs	R353L	Golden Gate Cloning
Cell Lysis Solution (from kit Puregene Yeast/Bact. Kit B)	QIAGEN	854016	Isolation of genomic DNA from *S. cerevisiae*
CutSmart Buffer	New England BioLabs	B7204S	Linearization of pRN1120
Deoxyribonucleic acid sodium salt from salmon testes	Sigma Aldrich	D1626	Transfromation of *S. cerevisiae* (carrier DNA)
dNTPs	Invitrogen	10297018	PCRs
EcoRI-HF	New England BioLabs	R3101S	Linearization of pRN1120
Ethanol absolute for analysis	Merck	100983	Isolation of genomic DNA from *S. cerevisiae*
Ethylenediamine-tetraacetic acid	Sigma Aldrich	ED	Transformation of *S. cerevisiae*
G418 disulfate salt	Sigma Aldrich	A1720	Selection of *S. cerevisiae* transformants
Histodenz	Sigma Aldrich	D2158	Yeast pixel art
Isopropanol	Merck	100993	Isolation of genomic DNA from *S. cerevisiae*
Lithium acetate dihydrate	Sigma Aldrich	L6883	Transformation of *S. cerevisiae*
Nancy-520 DNA Gel Stain	Sigma Aldrich	1494	Electrophoresis
NEB10 competent *E. coli* cells	New England BioLabs	C3019H	Transformation of *E. coli:* dx.doi.org/10.17504/protocols.io.nkvdcw6
Nourseothricin	Jena Bioscience	AB102	Selection of *S. cerevisiae* transformants
Phusion buffer	New England BioLabs	M0530L	PCRs
Phusion High-Fidelity DNA Polymerase	New England BioLabs	M0530L	PCRs
Polyethylene glycol 4000	Merck	7490	Transformation of *S. cerevisiae*
Protein Precipitation Solution (10 M NH4AC) (from kit Puregene Yeast/Bact. Kit B)	QIAGEN	854016	Isolation of genomic DNA from *S. cerevisiae*

Purple loading dye	New England BioLabs	B7024S	Electrophoresis
QIAprep Spin Miniprep Kit	QIAGEN	27106	Purification of plasmids
RNase coctail enzyme mix	Thermo Fisher Scientific	AM2286	Isolation of genomic DNA from *S. cerevisiae*
T4 DNA ligase buffer	Invitrogen	46300-018	Golden Gate Cloning
T4 DNA Ligase (1 U/µl)	Invitrogen	1705218	Golden Gate Cloning
UltraPure Agarose	Invitrogen	16500500	Electrophoresis
Wizard SV Gel and PCR Clean-Up System Kit	Promega	A9282	Purification of PCR products and linearized pRN1120
XhoI	New England BioLabs	R0146S	Linearization of pRN1120
Zymolyase 50 mg/ml (5 units/µL)	Zymo Research	E1006	Isolation of genomic DNA from *S. cerevisiae* (yeast lysis enzyme)
Zymolyase storage buffer	Zymo Research	E1004-B	Isolation of genomic DNA from *S. cerevisiae* (necessary for the preparation of yeast lysis enzyme)
Chemicals of general use			
2*Peptone-Yeast extract (PY) agar			Plate growth of *E. coli*
2*PY medium			Cultivation of *E. coli*
Demineralized water			Transformation of *S. cerevisiae*
ELFO buffer			Electrophoresis
MQ			Multiple steps
Physiological salt solution			Transformation of *S. cerevisiae*
TE buffer			Storage of DNA, transformation of *S. cerevisiae*
Yeast extract-peptone-dextrose (YEPD; 2% glucose) medium			Cultivation of *S. cerevisiae*
YEPD (2% glucose) agar			Plate growth of *S. cerevisiae*
Consumables			
Eppendorf tubes			
Falcon tubes (50 mL)			
Microplate 96 wells			
Petri dishes			
Pipette tips 0.5 - 10 µL			
Pipette tips 10 - 200 µL			
Pipette tips 100 - 1000 µL			
Shake flasks (500 mL)			
Sterile filters			
Equipment			
Centrifuge (Falcon tubes)			
Echo 525 acoustic liquid handler			

Incubator			
NanoDrop			
Set for eletrophoresis			
Spectrophoto-meter			
Table centrifuge (Eppendorfs tubes)			
Thermocycler			
Plasmids			
pCSN067	Addgene	ID 101748	https://www.addgene.org/
pRN1120	Addgene	ID 101750	https://www.addgene.org/
Strains			
S. cerevisiae strain CEN.PK113-7D	EUROSCARF collection		http://www.euroscarf.de

SI Table 3-2. Single crRNA array for LbCas12a containing homology with plasmid pRN1120

crRNA array sequence [a,b,c,d,e,f]
CATGTTTGACAGCTTATCATCGATAATCCGGAGCTAGCATGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAG*TCT TTGAAAAGATAATGTATGATTATGCTTTCACTCATATTTATACAGAAACTTGATGTTTTCTTTCGAGTATATACAAGGTGA TTACATGTACGTTTGAAGTACAACTCTAGATTTTGTAGTGCCCTCTTGGGCTAGCGGTAAAGGTGCGCATTTTTTCACACC CTACAATGTTCTGTTCAAAAGATTTTGGTCAAACGCTGTAGAAGTGAAAGTTGGTGCGCATGTTTCGGCGTTCGAAACTTC TCCGCAGTGAAAGATAAATGATC**AATTTCTACTAAGTGTAGAT**CTGGTGGGAGAGAAAGCTTATGA**AATTTCTACTAAGTG TAGAT**GTGCCGTACGCCGGAGCCGACGG**AATTTCTACTAAGTGTAGAT**TGCCCCTCTTATACGATTATATTTTTTTTTGTT TTTTATGTCT**GGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCAT AGCTGTTTCCTGTGTG**

a. Homology to pRN1120 (bold).

b. *SNR52* promoter (italics).

c. Genomic target sequences (underlined).

d. Guide direct repeats specific for LbCas12a (italics, bold).

e. *SUP4* terminator (italics).

f. Homology to pRN1120 (bold).

SI Table 3-3. Primers sequences

Name	Sequence [a]	Description [b]	Used in point
KC-101	CATGTTTGACAGCTTATCATC	FW primer for amplification of single crRNA array	2.1.4
KC-102	CACACAGGAAACAGCTATGAC	RV primer for amplification of single crRNA array	2.1.4
KC-103	AAGCGACTTCCAATCGCTTTGC	FW primer for amplification of donor DNA with connector 5	3.6.1
KC-104	AAAGCAAAGGAAGGAGAGAAC	RV primer for amplification of donor DNA with connector A	3.6.1
KC-105	CGGATCGATGTACACAACCG	FW primer for amplification of donor DNA with connector B	3.6.1
KC-106	CAACAGGAGGCGGATGGATATAC	RV primer for amplification of donor DNA with connector C	3.6.1
KC-107	AACGTTGTCCAGGTTTGTATCC	FW primer for amplification of donor DNA with connector D	3.6.1
KC-108	AGGTACAACAAGCACGACCG	RV primer for amplification of donor DNA with connector E	3.6.1
KC-109	CACTATAGCAATCTGGCTATATG	FW primer for amplification of INT1 5' with connector 5	4,4
KC-110	**AAACGCCTGTGGGTGTGGTACTGGATATGCAAAGCGATTGGAAGTCGCTT**GACTCCTCTGCCGTCATTCC	RV primer for amplification of INT1 5' with connector 5	4,4
KC-111	**TTGCCCATCGAACGTACAAGTACTCCTCTGTTCTCTCCTTCCTTTGCTTT**AAGCGTTGAAGTTTCCTCTTTG	FW primer for amplification of INT1 3' with connector A	4,4
KC-112	TGTCAACTGGAGAGCTATCG	RV primer for amplification of INT1 3' with connector A	4,4
KC-113	AGAAGATTTCTCTTCAATCTC	FW primer for amplification of INT2 5' with connector B	4,4
KC-114	**TGCTAAGATTTGTGTTCGTTTGGGTGCAGTCGGTTGTGTACATCGATCCG**CCCTTATCAAGGATACCTGGTTG	RV primer for amplification of INT2 5' with connector B	4,4
KC-115	**ACGCTTTCCGGCATCTTCCAGACCACAGTATATCCATCCGCCTCCTGTTG**GGCGATTACACAAGCGGTGG	FW primer for amplification of INT2 3' with connector C	4,4
KC-116	TCTCCTCTTCGATGACCGGG	RV primer for amplification of INT2 3' with connector C	4,4
KC-117	GGTCGTTTTTGTGCAGCATATTG	FW primer for amplification of INT3 5' with connector D	4,4
KC-118	**GCGGAATATTGGCGGAACGGACACACGTGGATACAAACCTGGACAACGTT**TTCCAAGGAGGTGAAGAACG	RV primer for amplification of INT3 5' with connector D	4,4
KC-119	**AAATAACCACAAACATCCTTCCCATATGCTCGGTCGTGCTTGTTGTACCT**GATGGGACGTCAGCACTGTAC	FW primer for amplification of INT3 3' with connector E	4,4
KC-120	GAGCTTACTCTATATATTCATTC	RV primer for amplification of INT3 3' with connector E	4,4
KC-121	GTTACTAAACTGGAACTGTCCG	FW primer for verification of integration of con5-*crtE*-conA to INT1 5'	7.4.1
KC-122	CACTGCTAACTACGTTTACTTC	FW primer for verification of integration of con5-*crtE*-conA to INT1 3'	7.4.1

KC-123	CACTGGAACTTGAGCTTGAG	FW primer for verification of integration of conB-*crtYB*-conC to INT2 5'	7.4.1
KC-124	GTCTCCAGCTGAATTGGTCC	FW primer for verification of integration of conB-*crtYB*-conC to INT2 3'	7.4.1
KC-125	CTCTCATGAAGCAGTCAAGTC	FW primer for verification of integration of conD-*crtI*-conE to INT3 5'	7.4.1
KC-126	GATCGGTCAATTAGGTGAAG	FW primer for verification of integration of conD-*crtI*-conE to INT3 3'	7.4.1
KC-127	CCTTGTCCAAGTAGGTGTCC	RV primer for verification of integration of con5-*crtE*-conA to INT1 5'	7.4.1
KC-128	GCTGTCATGATCTGTGATAAC	RV primer for verification of integration of con5-*crtE*-conA to INT1 3'	7.4.1
KC-129	CTGGCAATGTTGACCAATTGC	RV primer for verification of integration of conB-*crtYB*-conC to INT2 5'	7.4.1
KC-130	CCAACGTGCCTTAAAGTCTG	RV primer for verification of integration of conB-*crtYB*-conC to INT2 3'	7.4.1
KC-131	CCTTACCTTCTGGAGCAGCAG	RV primer for verification of integration of conD-*crtI*-conE to INT3 5'	7.4.1
KC-132	CTGGTTACTTCCCTAAGACTG	RV primer for verification of integration of conD-*crtI*-conE to INT3 3'	7.4.1

a. Bold sequences denote connector sequences.
b. Forward and reverse primers are designated as FW and RV, respectively.

SI Table 3-4. Design of constructed strains

Expression cassette	Integration site (left)	Left connector	Promoter	ORF	Terminator	Right connector	Integration site (right)
Strain 1							
EC 1	INT1	con5	Sbay_TDH3	crtE	Sc_TDH3	conA	INT1
EC 2	INT2	conB	Smik_TEF1	crtYB	Sc_PDC1	conC	INT2
EC 3	INT3	conD	Kl_ENO1	crtI	Sc_TAL1	conE	INT3
Strain 2							
EC 4	INT1	con5	Kl_PGK1	crtE	Sc_TDH3	conA	INT1
EC 5	INT2	conB	Kl_TEF2	crtYB	Sc_PDC1	conC	INT2
EC 6	INT3	conD	Kl_OLE1	crtI	Sc_TAL1	conE	INT3
Strain 3							
EC 7	INT1	con5	Kl_TDH2	crtE	Sc_TDH3	conA	INT1
EC 8	INT2	conB	Kl_YDR1	crtYB	Sc_PDC1	conC	INT2
EC 9	INT3	conD	Kl_LEU2	crtI	Sc_TAL1	conE	INT3

SI Table 3-5. Sequences of donor DNA expression cassettes and flaking regions

Donor DNA expression cassette sequences	
Name	**Purpose**
crtE expression cassette: con5-Sbay_TDH3p-crtE-Sc_TDH3t-conA	High strength promoter for expression of *crtE*, multiplex array

AAGCGACTTCCAATCGCTTTGCATATCCAGTACCACACCCACAGGCGTTTGTGCCATTCATCTTTCACCTGCCATTAGTAA
CCCGACTTCTCATTGAGCGGGTTACGGCAGCCACAGGCCACATTCCGAATGTCTGGGTGAGCGGTCCCTTTTCCAGCATCC
ACTAAATATCTCGGATCCCGCTTTTTAATCTGGCTTCCTGAAAAAAATCAATGGAGTGATGCAAACTGACTGGAGCAAAAA
GCTGACACAAGGCAATCGACCTACGTGTCTGTCTATTTTCTCACACCTTCTATTACCTTCTAACTCTCTGGGTTGGAAAAA
ACTGAAAAAAAGGTTGTCTCCAGTTTCCACAAATCATCCCCCTGTTTGATTAATAAATATATAAAGACGACAACTATCGAT
CATAAACTCATAAAACTATAACTCCTTTACACTTCTTATTTTATAGTTATTCTATTTTAATTCTTATTGATTTTAAAACCC
CAAGAACTTAGTTTCGAAAACACACACACACAAACAATTAAAAATGGACTACGCTAACATCTTGACTGCCATTCCTTTGGA
ATTCACCCCACAAGATGACATTGTCTTGTTGGAACCATACCACTACTTAGGTAAGAACCCAGGTAAGGAAATCAGATCTCA
ATTGATTGAAGCTTTCAACTACTGGTTAGATGTCAAGAAGGAAGACTTGGAAGTTATCCAAAATGTTGTTGGTATGTTGCA
CACCGCTTCTTTGTTGATGGATGATGTTGAAGATTCTTCCGTCTTGAGAAGAGGTTCTCCAGTTGCTCATTTGATCTACGG
TATTCCACAAACCATCAACACTGCTAACTACGTTTACTTCTTGGCTTACCAAGAAATCTTCAAATTGCGTCCAACTCCAAT
TCCAATGCCAGTTATCCCACCATCTTCTGCTTCTTTGCAATCTTCTGTCTCCTCCGCCTCCTCTTCCTCTTCTGCCTCCTC
TGAAAACGGTGGTACCTCCACTCCAAACTCCCAAATCCCATTCTCCAAGGACACCTACTTGGACAAGGTTATCACTGACGA
AATGTTGTCTTTGCACCGTGGTCAAGGTTTGGAATTATTCTGGAGAGACTCTTTGACCTGTCCATCTGAAGAAGAATACGT
CAAGATGGTCTTGGGTAAGACCGGTGGTTTGTTCAGAATTGCTGTCAGATTGATGATGGCCAAGTCTGAATGTGACATTGA
CTTTGTTCAATTGGTTAACTTGATTTCCATCTACTTCCAAATCAGAGATGACTACATGAACTTGCAATCCTCTGAATACGC
TCACAACAAGAACTTCGCTGAAGACTTGACTGAAGGTAAGTTCTCCTTCCCAACCATTCACTCCATTCACGCTAACCCATC
TTCCAGATTGGTTATCAACACTTTACAAAAGAAGTCCACTTCTCCAGAAATCTTACATCACTGTGTCAACTACATGAGAAC
TGAAACCCACTCTTTCGAATACACTCAAGAAGTCTTGAACACTTTATCTGGTGCTTTGGAAAGAGAATTGGGTAGATTACA
AGGTGAATTTGCTGAAGCTAACTCCAAGATCGATTTGGGTGACGTTGAATCTGAAGGTAGAACCGGTAAGAACGTCAAATT
GGAAGCCATCTTGAAGAAATTGGCTGATATCCCTCTATAAAGTGAATTTACTTTAAATCTTGCATTTAAATAAATTTTCTT
TTTATAGCTTTATGACTTAGTTTCAATTTATATACTATTTTAATGACATTTTCGATTCATTGATTGAAAGCTTTGTGTTTT
TTCTTGATGCGCTATTGCATTGTTCTTGTCTTTTTCGCCACATGTAATATCTGTAGTAGATACCTGATACATTGTGGATGC
TGAGTGAAATTTTAGTTAATAATGGAGGCGCTCTTAATAATTTTGGGGATATTGGCTTTTTTTTTTAAAGTTTACAAATGA
ATTTTTTCCGCCAGGATCCTC**TTGCCCATCGAACGTACAAGTACTCCTCTGTTCTCTCCTTCCTTTGCTTT**

crtE expression cassette: con5-Kl_PGK1p-crtE-Sc_TDH3t-conA	Medium strength promoter for expression of *crtE*, multiplex array

AAGCGACTTCCAATCGCTTTGCATATCCAGTACCACACCCACAGGCGTTTGTGCGTTCCTCATCACTAGAAGCCGAACTGT
TGTCTTCAGTGGGGATTGGTTCGACATTTTGCCAATTGCTGTCGATGTACCCTTTCAAAGCCATGTACCTTAAATCTTCAT
CCTTGGCAAGTAGATTCATCGGGTGTGTTTGAAGTAAGAATATTTGCTTGTTTTTATGGTATCAAAGGTATATGTTGTAGA
AGACAATTTCCGGTAATCCAATTGTCTGTCTGCTCAGTTTAGCACATGTATAGTACGTTGCACATAGTCTACAATATTCAG
CATTCAGCATTCAGTATACAGCATATGGCTAAATGATCACAAATGTGATTGATGATTTGACACGACTAGAAAAGAGAACGA
AAAAGGGAAATTCCATGTCACGTGCGTTGGCACGTGACATGGAATATCGAAGAAAGAAAAAAAAAAACGATCTCGTCCTAGT
GGAAGCCCAGAGTCTGGTCCCCCCGGAGTCTTCCCAAAACAAGAAGCTGACACATGTTGACACAGAACACCCCACAGCAAA
TGCACCACGCTACGTAGATCAGGAAGCTTAACTCTAGCGACCTGTCGCTCGCCCCACAGAACCTCACCCGAGAACCACACA

TTACACGCCGCCAGCTCCCACTATACTCATCTTGCTTCCCTTAAGCGTTCTCACGATTCGTTCGCTGCCCTTCTTCAAGAG
TCTTCTGATTCTAATTCTCATTCGAAATCCTCTACAGTTAATGAATTGCTTGACATGACATTCATTGTCTCATGGTTTTGG
CTTTTTGGCTTTTGTCTTTTAAAGCTATATCAACTTTACATATAAATATACGTCAAAAGGGGATTCATTAATTAGAAAATT
CTCTTTTTCAATAGTTGCTATTCATTATCAATCTATTCAACTCAATTGGTTATTATTTTCATCTTTTTGTCATCCTAAACC
ATCAACAATATTTAAATATATCTGTTGCTACATTAAGAGTTACTTCAGAAATAACAAAAAAATCGATCAAGAATTAATAAA
AATGGACTACGCTAACATCTTGACTGCCATTCCTTTGGAATTCACCCCACAAGATGACATTGTCTTGTTGGAACCATACCA
CTACTTAGGTAAGAACCCAGGTAAGGAAATCAGATCTCAATTGATTGAAGCTTTCAACTACTGGTTAGATGTCAAGAAGGA
AGACTTGGAAGTTATCCAAAATGTTGTTGGTATGTTGCACACCGCTTCTTTGTTGATGGATGATGTTGAAGATTCTTCCGT
CTTGAGAAGAGGTTCTCCAGTTGCTCATTTGATCTACGGTATTCCACAAACCATCAACACTGCTAACTACGTTTACTTCTT
GGCTTACCAAGAAATCTTCAAATTGCGTCCAACTCCAATTCCAATGCCAGTTATCCCACCATCTTCTGCTTCTTTGCAATC
TTCTGTCTCCTCCGCCTCCTCTTCCTCTTCTGCCTCCTCTGAAAACGGTGGTACCTCCACTCCAAACTCCCAAATCCCATT
CTCCAAGGACACCTACTTGGACAAGGTTATCACTGACGAAATGTTGTCTTTGCACCGTGGTCAAGGTTTGGAATTATTCTG
GAGAGACTCTTTGACCTGTCCATCTGAAGAAGAATACGTCAAGATGGTCTTGGGTAAGACCGGTGGTTTGTTCAGAATTGC
TGTCAGATTGATGATGGCCAAGTCTGAATGTGACATTGACTTTGTTCAATTGGTTAACTTGATTTCCATCTACTTCCAAAT
CAGAGATGACTACATGAACTTGCAATCCTCTGAATACGCTCACAACAAGAACTTCGCTGAAGACTTGACTGAAGGTAAGTT
CTCCTTCCCAACCATTCACTCCATTCACGCTAACCCATCTTCCAGATTGGTTATCAACACTTTACAAAAGAAGTCCACTTC
TCCAGAAATCTTACATCACTGTGTCAACTACATGAGAACTGAAACCCACTCTTTCGAATACACTCAAGAAGTCTTGAACAC
TTTATCTGGTGCTTTGGAAAGAGAATTGGGTAGATTACAAGGTGAATTTGCTGAAGCTAACTCCAAGATCGATTTGGGTGA
CGTTGAATCTGAAGGTAGAACCGGTAAGAACGTCAAATTGGAAGCCATCTTGAAGAAATTGGCTGATATCCCTCTATAAAG
TGAATTTACTTTAAATCTTGCATTTAAATAAATTTTCTTTTTATAGCTTTATGACTTAGTTTCAATTTATATACTATTTTA
ATGACATTTTCGATTCATTGATTGAAAGCTTTGTGTTTTTTCTTGATGCGCTATTGCATTGTTCTTGTCTTTTTCGCCACA
TGTAATATCTGTAGTAGATACCTGATACATTGTGGATGCTGAGTGAAATTTTAGTTAATAATGGAGGCGCTCTTAATAATT
TTGGGGATATTGGCTTTTTTTTTTAAAGTTTACAAATGAATTTTTTCCGCCAGGATCCTC**TTGCCCATCGAACGTACAAGT**
ACTCCTCTGTTCTCTCCTTCCTTTGCTTT

crtE expression cassette: con5-Kl_TDH2p-crtE-Sc_TDH3t-conA	Low strength promoter for expression of *crtE*, multiplex array

AAGCGACTTCCAATCGCTTTGCATATCCAGTACCACACCCACAGGCGTTTGTGCCGTAAAAACTAAAACGAGCCCCCACCA
AAGAACAAAAAAGAAGGTGCTGGGCCCCCACTTTCTTCCCTTGCACGTGATAGGAAGATGGCTACAGAAACAAGAAGATGG
AAATCGAAGGAAAGAGGGAGACTGGAAGCTGTAAAAACTGAAATGAAAAAAAAAAAAAAAAAAAAAAACAAGAAGCTGAAA
ATGGAAGACTGAAATTTGAAAAATGGTAAAAAAAAAAAAGAAACACGAAGCTAAAAACCTGGATTCCATTTTGAGAAGAAG
CAAGAAAGGTAAGTATGGTAACGACCGTACAGGCAAGCGCGAAGGCAAATGGAAAAGCTGGAGTCCGGAAGATAATCATTT
CATCTTCTTTTGTTAGAACAGAACAGTGGATGTCCCTCATCTCGGTAACGTATTGTCCATGCCCTAGAACTCTCTGTCCCT
AAAAAGAGGACAAAAACCCAATGGTTTCCCCAGCTTCCAGTGGAGCCACCGATCCCACTGGAAACCACTGGACAGGAAGAG
AAAATCACGGACTTCCTCTATTGAAGGATAATTCAACACTTTCACCAGATCCCAAATGTCCCGCCCCTATTCCCGTGTTCC
ATCACGTACCATAACTTACCATTTCATCACGTTCTCTATGGCACACTGGTACTGCTTCGACTGCTTTGCTTCATCTTCTCT
ATGGGCCAATGAGCTAATGAGCACAATGTGCTGCGAAATAAAGGGATATCTAATTTATATTATTACATTATAATATGTACT
AGTGTGGTTATTGGTAATTGTACTTAATTTTGATATATAAAGGGTGGATCTTTTTCATTTTGAATCAGAATTGGAATTGCA
ACTTGTCTCTTGTCACTATTACTTAATAGTAATTATATTTCTTATTAACCTTTTTTTTAAGTCAAAACACCAAGGACAAGA
ACTACTCTTCAAAGGTATTTCAAGTTATCATACGTGTCACACACGCTTCACAGTTTCAAGTAAAAAAAAAGAATATTACAC
AATGGACTACGCTAACATCTTGACTGCCATTCCTTTGGAATTCACCCCACAAGATGACATTGTCTTGTTGGAACCATACCA
CTACTTAGGTAAGAACCCAGGTAAGGAAATCAGATCTCAATTGATTGAAGCTTTCAACTACTGGTTAGATGTCAAGAAGGA
AGACTTGGAAGTTATCCAAAATGTTGTTGGTATGTTGCACACCGCTTCTTTGTTGATGGATGATGTTGAAGATTCTTCCGT

CTTGAGAAGAGGTTCTCCAGTTGCTCATTTGATCTACGGTATTCCACAAACCATCAACACTGCTAACTACGTTTACTTCTT
GGCTTACCAAGAAATCTTCAAATTGCGTCCAACTCCAATTCCAATGCCAGTTATCCCACCATCTTCTGCTTCTTTGCAATC
TTCTGTCTCCTCCGCCTCCTCTTCCTCTTCTGCCTCCTCTGAAAACGGTGGTACCTCCACTCCAAACTCCCAAATCCCATT
CTCCAAGGACACCTACTTGGACAAGGTTATCACTGACGAAATGTTGTCTTTGCACCGTGGTCAAGGTTTGGAATTATTCTG
GAGAGACTCTTTGACCTGTCCATCTGAAGAAGAATACGTCAAGATGGTCTTGGGTAAGACCGGTGGTTTGTTCAGAATTGC
TGTCAGATTGATGATGGCCAAGTCTGAATGTGACATTGACTTTGTTCAATTGGTTAACTTGATTTCCATCTACTTCCAAAT
CAGAGATGACTACATGAACTTGCAATCCTCTGAATACGCTCACAACAAGAACTTCGCTGAAGACTTGACTGAAGGTAAGTT
CTCCTTCCCAACCATTCACTCCATTCACGCTAACCCATCTTCCAGATTGGTTATCAACACTTTACAAAAGAAGTCCACTTC
TCCAGAAATCTTACATCACTGTGTCAACTACATGAGAACTGAAACCCACTCTTTCGAATACACTCAAGAAGTCTTGAACAC
TTTATCTGGTGCTTTGGAAAGAGAATTGGGTAGATTACAAGGTGAATTTGCTGAAGCTAACTCCAAGATCGATTTGGGTGA
CGTTGAATCTGAAGGTAGAACCGGTAAGAACGTCAAATTGGAAGCCATCTTGAAGAAATTGGCTGATATCCCTCTATAAAG
TGAATTTACTTTAAATCTTGCATTTAAATAAATTTTCTTTTTATAGCTTTATGACTTAGTTTCAATTTATATACTATTTTA
ATGACATTTTCGATTCATTGATTGAAAGCTTTGTGTTTTTTCTTGATGCGCTATTGCATTGTTCTTGTCTTTTTCGCCACA
TGTAATATCTGTAGTAGATACCTGATACATTGTGGATGCTGAGTGAAATTTTAGTTAATAATGGAGGCGCTCTTAATAATT
TTGGGGATATTGGCTTTTTTTTTTAAAGTTTACAAATGAATTTTTTCCGCCAGGATCCTC**TTGCCCATCGAACGTACAAGT**
ACTCCTCTGTTCTCTCCTTCCTTTGCTTT

crtYB expression cassette: conB-Smik_TEF1p-crtYB-Sc_PDC1t-conC	High strength promoter for expression of *crtYB*, multiplex array

CGGATCGATGTACACAACCGACTGCACCCAAACGAACACAAATCTTAGCAGTGCATGCTTCAAAAACGCACTGTACTCCTT
TTTACTCTTCCGGATTTTCTCGCACTCTCCGCATCGCCGCACGAGCCAAGCCACACCCACACACCTCATACCATGTTTCCC
CTCTTTGTCTCTTTCGTGCGGCTCCATTACCCGCATGAAACTGTATAAAAGTAACAAAAGACTATTTCGTTTCTTTTTCTT
TGTCGGAAAAGGCAAAAAAAAAAATTTTTATCACATTTCTTTTTCTTGAAAATTTTTTTTGGGATTTTTTCTCTTTCGATG
ACCTCCCATTGATATTTAAGTTAATAAAAGCACTCCCGTTTTCCAAGTTTTAATTTGTTCCTCTTGTTTAGTCATTCTTCT
TCTCAGCATTGGTCAATTAGAAAGAGAGCATAGCAAACTGATCTAAGTTTTAATTACAAAATGACCGCTTTGGCTTACTAC
CAAATCCACTTGATCTACACTTTGCCAATCTTAGGTTTGCTAGGTTTGTTGACTTCTCCAATTTTGACCAAATTCGACATC
TACAAGATTTCTATCTTAGTCTTTATTGCTTTCTCTGCTACCACTCCATGGGACTCCTGGATCATCAGAAACGGTGCCTGG
ACCTACCCATCTGCTGAATCTGGTCAAGGTGTTTTCGGTACCTTTTTGGATGTCCCATACGAAGAATACGCCTTCTTTGTT
ATCCAAACCGTCATCACCGGTTTGGTTTACGTTTTGGCTACCAGACATTTGTTGCCATCTTTGGCTCTACCAAAGACCCGT
TCTTCTGCCTTGTCTCTAGCTTTGAAGGCTTTAATCCCATTGCCAATCATCTATTTGTTCACCGCTCATCCATCTCCATCC
CCAGATCCTTTGGTTACTGACCACTACTTCTACATGAGAGCTTTGTCTTTGTTGATCACCCCACCAACCATGTTGTTGGCT
GCTTTATCTGGTGAATACGCTTTCGACTGGAAATCTGGTAGAGCTAAGTCCACCATTGCTGCCATCATGATCCCAACTGTC
TACTTGATCTGGGTTGACTACGTTGCCGTTGGTCAAGACTCCTGGTCCATCAACGATGAAAAGATTGTCGGTTGGAGATTA
GGTGGTGTCTTGCCAATTGAAGAAGCTATGTTCTTCTTATTGACCAACTTGATGATCGTTTTGGGTTTGTCTGCCTGTGAC
CACACTCAAGCCTTGTACTTGTTGCACGGTAGAACTATCTACGGTAACAAGAAGATGCCATCTTCTTTCCCATTAATCACT
CCACCAGTTTTGTCCTTGTTCTTCTCCTCCAGACCATACTCCTCCCAACCAAAGAGAGATTTGGAATTGGCTGTCAAGTTG
TTGGAAAAGAAGTCCAGATCTTTCTTCGTTGCTTCTGCCGGTTTCCCATCTGAAGTCAGAGAAAGATTGGTTGGTTTGTAC
GCTTTCTGTCGTGTCACCGATGACTTGATTGACTCTCCAGAAGTTTCCTCCAACCCACACGCTACCATTGACATGGTTTCC
GATTTCTTGACTTTATTATTCGGTCCTCCATTGCACCCATCTCAACCAGACAAGATTTTGTCTTCTCCATTATTACCACCT
TCCCACCCATCCAGACCAACTGGTATGTACCCATTACCACCACCTCCATCTTTGTCTCCAGCTGAATTGGTCCAATTCTTG
ACTGAACGTGTCCCAGTTCAATACCACTTCGCTTTCAGATTGTTGGCCAAATTGCAAGGTTTGATTCCAAGATACCCATTG
GATGAATTATTGAGAGGTTACACCACTGACTTGATCTTCCCATTGTCCACTGAAGCCGTCCAAGCTAGAAAGACCCCAATT
GAAACTACTGCTGACTTGTTGGACTACGGTTTGTGTGTTGCCGGTTCTGTTGCTGAATTGTTGGTCTACGTTTCCTGGGCT

TCCGCTCCATCCCAAGTTCCAGCTACTATTGAAGAAAGAGAAGCTGTTTTGGTCGCCTCTCGTGAAATGGGTACCGCTTTG
CAATTGGTCAACATTGCCAGAGATATCAAGGGTGACGCTACTGAAGGTAGATTCTACTTGCCATTGTCTTTCTTTGGTTTG
AGAGATGAATCCAAATTGGCCATTCCAACTGACTGGACTGAACCAAGACCTCAAGATTTCGACAAATTGTTGTCTCTATCT
CCATCTTCCACTTTACCATCCTCTAACGCTTCTGAATCCTTCAGATTCGAATGGAAGACCTACTCTTTGCCATTGGTTGCT
TACGCTGAAGATTTGGCTAAGCACTCTTACAAGGGTATTGACAGATTACCAACTGAAGTCCAAGCTGGTATGAGAGCTGCT
TGTGCTTCTTACTTGTTGATTGGTCGTGAAATCAAGGTTGTCTGGAAGGGTGATGTCGGTGAAAGAAGAACCGTTGCTGGT
TGGAGAAGAGTCAGAAAGGTTTTGTCTGTTGTCATGTCCGGTTGGGAAGGTCAATAAAGCGATTTAATCTCTAATTATTAG
TTAAAGTTTTATAAGCATTTTTATGTAACGAAAAATAAATTGGTTCATATTATTACTGCACTGTCACTTACCATGGAAAGA
CCAGACAAGAAGTTGCCGACAGTCTGTTGAATTGGCCTGGTTAGGCTTAAGTCTGGGTCCGCTTCTTTACAAATTTGGAGA
ATTTCTCTTAAACGATATGTATATTCTTTTCGTTGGAAAAGATGTCTTCCAAAAAAAAAACCGATGAATTAGTGGAACCAA
GGAAAAAAAAAGAGGTATCCTTGATTAAGGAACACCTC**ACGCTTTCCGGCATCTTCCAGACCACAGTATATCCATCCGCCT**
CCTGTTG

crtYB expression cassette: conB-Kl_TEF2p-crtYB-Sc_PDC1t-conC	Medium strength promoter for expression of *crtYB*, multiplex array

CGGATCGATGTACACAACCGACTGCACCCAAACGAACACAAATCTTAGCAGTGCGAGCCTGTCCAAGCAAATGCCTTCTCA
TAAATGGTGCCAAAGACCCGCAAGCCCAAAGCAATTACCCCCCAAAAAGAAATGATATAGTGCAAGATACGTATATGACCA
TGACTTGACTAGGTGAAACAGTGCAGAAACAGCCGCACAAAAGCAGCCCTAACCCTCAGAGTCGATTTTACTCTTTCAGGT
AATAAAGCCTCGACATCAATTTTAGACAGAAGCCAGGCTGGCCTCGAGATTATAGCCATAGGCAAGCAAGAGGAGAGAAGG
GGAGGCCCCCCATGGGGGGCCTCCCCCCCGCTGTCAAGGTTTGGCAGAACCTAGCTTCATTAGGCCACTAGCCCAGCCTAA
AACGTCAACGGGCAGGAGGAACACTCCCACAAGACGGCGTAGTATTCTCGATTCATAACCATTTTCTCAATCGAATTACAC
AGAACACACCGTACAAACCTCTCTATCATAACTACTTAATAGTCACACACGTACTCGTCTAAATACACATCATCGTCCTAC
AAGTTCATCAAAGTGTTGGACAGACAACTATACCAGCATGGATCTCTTGTATCGGTTCTTTTCTCCCGCTCTCTCGCAATA
ACAATGAACACTGGGTCAATCATAGCCTACACAGGTGAACAGAGTAGCGTTTATACAGGGTTTATACGGTGATTCCTACGG
CAAAAATTTTTCATTTCTAAAAAAAAAAAGAAAAATTTTTCTTTCCAACGCTAGAAGGAAAAGAAAAATCTAATTAAATTG
ATTTGGTGATTTTCTGAGAGTTCCCTTTTTCATATATCGAATTTTGAATATAAAAGGAGATCGAAAAAATTTTTCTATTCA
ATCTGTTTTCTGGTTTTATTTGATAGTTTTTTTGTGTATTATTATTATGGATTAGTACTGGTTTATATGGGTTTTTCTGTA
TAACTTCTTTTTATTTTAGTTTGTTTAATCTTATTTTGAGTTACATTATAGTTCCCTAACTGCAAGAGAAGTAACATTAAA
AATGACCGCTTTGGCTTACTACCAAATCCACTTGATCTACACTTTGCCAATCTTAGGTTTGCTAGGTTTGTTGACTTCTCC
AATTTTGACCAAATTCGACATCTACAAGATTTCTATCTTAGTCTTTATTGCTTTCTCTGCTACCACTCCATGGGACTCCTG
GATCATCAGAAACGGTGCCTGGACCTACCCATCTGCTGAATCTGGTCAAGGTGTTTTCGGTACCTTTTTGGATGTCCCATA
CGAAGAATACGCCTTCTTTGTTATCCAAACCGTCATCACCGGTTTGGTTTACGTTTTGGCTACCAGACATTTGTTGCCATC
TTTGGCTCTACCAAAGACCCGTTCTTCTGCCTTGTCTCTAGCTTTGAAGGCTTTAATCCCATTGCCAATCATCTATTTGTT
CACCGCTCATCCATCTCCATCCCCAGATCCTTTGGTTACTGACCACTACTTCTACATGAGAGCTTTGTCTTTGTTGATCAC
CCCACCAACCATGTTGTTGGCTGCTTTATCTGGTGAATACGCTTTCGACTGGAAATCTGGTAGAGCTAAGTCCACCATTGC
TGCCATCATGATCCCAACTGTCTACTTGATCTGGGTTGACTACGTTGCCGTTGGTCAAGACTCCTGGTCCATCAACGATGA
AAAGATTGTCGGTTGGAGATTAGGTGGTGTCTTGCCAATTGAAGAAGCTATGTTCTTCTTATTGACCAACTTGATGATCGT
TTTGGGTTTGTCTGCCTGTGACCACACTCAAGCCTTGTACTTGTTGCACGGTAGAACTATCTACGGTAACAAGAAGATGCC
ATCTTCTTTCCCATTAATCACTCCACCAGTTTTGTCCTTGTTCTTCTCCTCCAGACCATACTCCTCCCAACCAAAGAGAGA
TTTGGAATTGGCTGTCAAGTTGTTGGAAAAGAAGTCCAGATCTTTCTTCGTTGCTTCTGCCGGTTTCCCATCTGAAGTCAG
AGAAAGATTGGTTGGTTTGTACGCTTTCTGTCGTGTCACCGATGACTTGATTGACTCTCCAGAAGTTTCCTCCAACCCACA
CGCTACCATTGACATGGTTTCCGATTTCTTGACTTTATTATTCGGTCCTCCATTGCACCCATCTCAACCAGACAAGATTTT
GTCTTCTCCATTATTACCACCTTCCCACCCATCCAGACCAACTGGTATGTACCCATTACCACCACCTCCATCTTTGTCTCC

AGCTGAATTGGTCCAATTCTTGACTGAACGTGTCCCAGTTCAATACCACTTCGCTTTCAGATTGTTGGCCAAATTGCAAGG
TTTGATTCCAAGATACCCATTGGATGAATTATTGAGAGGTTACACCACTGACTTGATCTTCCCATTGTCCACTGAAGCCGT
CCAAGCTAGAAAGACCCCAATTGAAACTACTGCTGACTTGTTGGACTACGGTTTGTGTGTTGCCGGTTCTGTTGCTGAATT
GTTGGTCTACGTTTCCTGGGCTTCCGCTCCATCCCAAGTTCCAGCTACTATTGAAGAAAGAGAAGCTGTTTTGGTCGCCTC
TCGTGAAATGGGTACCGCTTTGCAATTGGTCAACATTGCCAGAGATATCAAGGGTGACGCTACTGAAGGTAGATTCTACTT
GCCATTGTCTTTCTTTGGTTTGAGAGATGAATCCAAATTGGCCATTCCAACTGACTGGACTGAACCAAGACCTCAAGATTT
CGACAAATTGTTGTCTCTATCTCCATCTTCCACTTTACCATCCTCTAACGCTTCTGAATCCTTCAGATTCGAATGGAAGAC
CTACTCTTTGCCATTGGTTGCTTACGCTGAAGATTTGGCTAAGCACTCTTACAAGGGTATTGACAGATTACCAACTGAAGT
CCAAGCTGGTATGAGAGCTGCTTGTGCTTCTTACTTGTTGATTGGTCGTGAAATCAAGGTTGTCTGGAAGGGTGATGTCGG
TGAAAGAAGAACCGTTGCTGGTTGGAGAAGAGTCAGAAAGGTTTTGTCTGTTGTCATGTCCGGTTGGGAAGGTCAATAAAG
CGATTTAATCTCTAATTATTAGTTAAAGTTTTATAAGCATTTTTATGTAACGAAAAATAAATTGGTTCATATTATTACTGC
ACTGTCACTTACCATGGAAAGACCAGACAAGAAGTTGCCGACAGTCTGTTGAATTGGCCTGGTTAGGCTTAAGTCTGGGTC
CGCTTCTTTACAAATTTGGAGAATTTCTCTTAAACGATATGTATATTCTTTTCGTTGGAAAAGATGTCTTCCAAAAAAAAA
ACCGATGAATTAGTGGAACCAAGGAAAAAAAAAGAGGTATCCTTGATTAAGGAACACCTC**ACGCTTTCCGGCATCTTCCAG**
ACCACAGTATATCCATCCGCCTCCTGTTG

crtYB expression cassette: conB-Kl_YDR1p-crtYB-Sc_PDC1t-conC	Low strength promoter for expression of *crtYB*, multiplex array

CGGATCGATGTACACAACCGACTGCACCCAAACGAACACAAATCTTAGCAGTGCTTTTCTTTTTTTGCGGTCACCCCCATG
TGGCGGGGAGGCAGAGGAGTAGGTAGAGCAACGAATCCTACTATTTATCCAAATTAGTCTAGGAACTCTTTTTCTAGATTT
TTTAGATTTGAGGGCAAGCGCTGTTAACGACTCAGAAATGTAAGCACTACGGAGTAGAACGAGAAATCCGCCATAGGTGGA
AATCCTAGCAAAATCTTGCTTACCCTAGCTAGCCTCAGGTAAGCTAGCCTTAGCCTGTCAAATTTTTTTCAAAATTTGGTA
AGTTTCTACTAGCAAAGCAAACACGGTTCAACAAACCGAAAACTCCACTCATTATACGTGGAAACCGAAACAAAAAAACAA
AAACCAAAATACTCGCCAATGAGAAAGTTGCTGCGTTTCTACTTTCGAGGAAGAGGAACTGAGAGGATTGACTACGAAAGG
GGCAAAAACGAGTCGTATTCTCCCATTATTGTCTGCTACCACGCGGTCTAGTAGAATAAGCAACCAGTCAACGCTAAGACA
GGTAATCAAAATACCAGTCTGCTGGCTACGGGCTAGTTTTTACCTCTTTTAGAACCCACTGTAAAAGTCCGTTGTAAAGCC
CGTTCTCACTGTTGGCGTTTTTTTTTTTTTGGTTTAGTTTCTTATTTTTCATTTTTTTCTTTCATGACCAAAAACAAACAA
ATCTCGCGATTTGTACTGCGGCCACTGGGGCGTGGCCAAAAAAATGACAAATTTAGAAACCTTAGTTTCTGATTTTTCCTG
TTATGAGGAGATATGATAAAAAATATTACTGCTTTATTGTTTTTTTTTTATCTACTGAAATAGAGAAACTTACCCAAGGAG
GAGGCAAAAAAAAGAGTATATATACAGCAGGTACCATTCAGATTTTAATATATTCTTTTCTCTTCTTCTACACTATTATTA
TAATAATTTTACTATATTCATTTTTAGCTTAAAACCTCATAGAATATTATTCTTCAGTCACTCGCTTAAATACTTATCAAA
AATGACCGCTTTGGCTTACTACCAAATCCACTTGATCTACACTTTGCCAATCTTAGGTTTGCTAGGTTTGTTGACTTCTCC
AATTTTGACCAAATTCGACATCTACAAGATTTCTATCTTAGTCTTTATTGCTTTCTCTGCTACCACTCCATGGGACTCCTG
GATCATCAGAAACGGTGCCTGGACCTACCCATCTGCTGAATCTGGTCAAGGTGTTTTCGGTACCTTTTTGGATGTCCCATA
CGAAGAATACGCCTTCTTTGTTATCCAAACCGTCATCACCGGTTTGGTTTACGTTTTGGCTACCAGACATTTGTTGCCATC
TTTGGCTCTACCAAAGACCCGTTCTTCTGCCTTGTCTCTAGCTTTGAAGGCTTTAATCCCATTGCCAATCATCTATTTGTT
CACCGCTCATCCATCTCCATCCCCAGATCCTTTGGTTACTGACCACTACTTCTACATGAGAGCTTTGTCTTTGTTGATCAC
CCCACCAACCATGTTGTTGGCTGCTTTATCTGGTGAATACGCTTTCGACTGGAAATCTGGTAGAGCTAAGTCCACCATTGC
TGCCATCATGATCCCAACTGTCTACTTGATCTGGGTTGACTACGTTGCCGTTGGTCAAGACTCCTGGTCCATCAACGATGA
AAAGATTGTCGGTTGGAGATTAGGTGGTGTCTTGCCAATTGAAGAAGCTATGTTCTTCTTATTGACCAACTTGATGATCGT
TTTGGGTTTGTCTGCCTGTGACCACACTCAAGCCTTGTACTTGTTGCACGGTAGAACTATCTACGGTAACAAGAAGATGCC
ATCTTCTTTCCCATTAATCACTCCACCAGTTTTGTCCTTGTTCTTCTCCTCCAGACCATACTCCTCCCAACCAAAGAGAGA
TTTGGAATTGGCTGTCAAGTTGTTGGAAAAGAAGTCCAGATCTTTCTTCGTTGCTTCTGCCGGTTTCCCATCTGAAGTCAG

AGAAAGATTGGTTGGTTTGTACGCTTTCTGTCGTGTCACCGATGACTTGATTGACTCTCCAGAAGTTTCCTCCAACCCACA
CGCTACCATTGACATGGTTTCCGATTTCTTGACTTTATTATTCGGTCCTCCATTGCACCCATCTCAACCAGACAAGATTTT
GTCTTCTCCATTATTACCACCTTCCCACCCATCCAGACCAACTGGTATGTACCCATTACCACCACCTCCATCTTTGTCTCC
AGCTGAATTGGTCCAATTCTTGACTGAACGTGTCCCAGTTCAATACCACTTCGCTTTCAGATTGTTGGCCAAATTGCAAGG
TTTGATTCCAAGATACCCATTGGATGAATTATTGAGAGGTTACACCACTGACTTGATCTTCCCATTGTCCACTGAAGCCGT
CCAAGCTAGAAAGACCCCAATTGAAACTACTGCTGACTTGTTGGACTACGGTTTGTGTGTTGCCGGTTCTGTTGCTGAATT
GTTGGTCTACGTTTCCTGGGCTTCCGCTCCATCCCAAGTTCCAGCTACTATTGAAGAAAGAGAAGCTGTTTTGGTCGCCTC
TCGTGAAATGGGTACCGCTTTGCAATTGGTCAACATTGCCAGAGATATCAAGGGTGACGCTACTGAAGGTAGATTCTACTT
GCCATTGTCTTTCTTTGGTTTGAGAGATGAATCCAAATTGGCCATTCCAACTGACTGGACTGAACCAAGACCTCAAGATTT
CGACAAATTGTTGTCTCTATCTCCATCTTCCACTTTACCATCCTCTAACGCTTCTGAATCCTTCAGATTCGAATGGAAGAC
CTACTCTTTGCCATTGGTTGCTTACGCTGAAGATTTGGCTAAGCACTCTTACAAGGGTATTGACAGATTACCAACTGAAGT
CCAAGCTGGTATGAGAGCTGCTTGTGCTTCTTACTTGTTGATTGGTCGTGAAATCAAGGTTGTCTGGAAGGGTGATGTCGG
TGAAAGAAGAACCGTTGCTGGTTGGAGAAGAGTCAGAAAGGTTTTGTCTGTTGTCATGTCCGGTTGGGAAGGTCAATAAAG
CGATTTAATCTCTAATTATTAGTTAAAGTTTTATAAGCATTTTTATGTAACGAAAAATAAATTGGTTCATATTATTACTGC
ACTGTCACTTACCATGGAAAGACCAGACAAGAAGTTGCCGACAGTCTGTTGAATTGGCCTGGTTAGGCTTAAGTCTGGGTC
CGCTTCTTTACAAATTTGGAGAATTTCTCTTAAACGATATGTATATTCTTTTCGTTGGAAAAGATGTCTTCCAAAAAAAAA
ACCGATGAATTAGTGGAACCAAGGAAAAAAAAAGAGGTATCCTTGATTAAGGAACACCTC**ACGCTTTCCGGCATCTTCCAG**
ACCACAGTATATCCATCCGCCTCCTGTTG

crtI expression cassette: conD-Kl_ENO1p-crtI-Sc_TAL1t-conE	High strength promoter for expression of *crtI*, multiplex array

AACGTTGTCCAGGTTTGTATCCACGTGTGTCCGTTCCGCCAATATTCCGCGTGCCGTATCCCTATCTGGATTAACATCACT
GCCACAGATCGAATTGCAAGAAGCCACACTTCACGTGATCCACTCGTTCATCAGGTTTGTAGCTTCATGGCGCAGGACTTC
TGATGATGAACTATCTGGCTCATCCGGATGGATCACAAGGATGGATACCCTCAGACAGTACGTTTCCGTATGGAGCGATCT
TACAAGAACCAACAGTTGATCCTATTACTTTTTTTTTATTTTTTGTCCCTCCGGGATGGCAAGAGGGACAAAGAAGAATCT
TCGTTCTTCTTTCTTGTTCTCAACTTCCCAGCTTCCGTGTGATTACCCTCCGGGACAACAGAAAAACTGGCATTCGGTATC
CCGGGAATCTGCTGAGAAGGAAAGAAAACGAAAAAAAAATTGTACATTTGTGTCACATTATGAATTACAGGAAGTCAGAAA
ACAGGCAGCACATGTCTCGCACATGCATGTCCATCAGACGAGACATTATGAGACATGCACGCGTGTGAGAGACATAGCAAA
AGTCTCTCCAGTACACACAGAAAGACACGTTCACAATCCAGGCACCCCACAGAGAAAAAAAAAAGAAGAAGCCCGGAAGCT
GGCACGCCATCATCAACCACCGCTCGGTTTACACGCATCCCAACTGTCTTTTTTTTCTGGAATCCTATAATAACTGGCATC
TGGAAATCACGTTGTATGTTGCACCATAGTGACTGGCTGTCTGACTAGCAAACATTGATTCCCTGATTCCCATTTGGCTCA
ATTTTGATGAGAAACAGTTGATTGATTCTTGTCAATTTTTTTTTCTTTGGACCACCACCAACCAATTGACATTGAAGTACT
TTCCCATGATTTGAGGTTATATAAAAGGACGTTCAAATCACTTTCAAGGTTAATTCAGTTTTGTCAATTGATTTAAGTTCA
ATTGTTAACAAATTTAATTTAATTCGAAACAAACCAAACCAATTCATTTGAATTAACAAACCAACCCACAAAACAAAAAAA
AATGGGTAAGGAACAAGACCAAGACAAGCCAACTGCCATCATCGTTGGTTGTGGTATCGGTGGTATTGCTACCGCTGCCAG
ATTAGCTAAGGAAGGTTTCCAAGTTACCGTCTTTGAAAAGAACGACTACTCCGGTGGTAGATGTTCTTTGATTGAAAGAGA
TGGTTACAGATTCGACCAAGGTCCATCTTTGTTGCTATTACCAGACTTGTTCAAGCAAACCTTCGAAGATTTGGGTGAAAA
GATGGAAGACTGGGTTGATTTGATCAAGTGTGAACCAAACTACGTTTGTCACTTCCATGATGAAGAAACTTTCACCTTCTC
CACTGACATGGCTTTATTGAAGAGAGAAGTCGAAAGATTTGAAGGTAAAGATGGTTTCGACAGATTCTTGTCTTTCATCCA
AGAAGCTCACAGACATTACGAATTGGCTGTTGTCCACGTCTTGCAAAAGAACTTCCCAGGTTTCGCTGCTTTCTTGAGATT
ACAATTCATCGGTCAAATCTTAGCTTTGCACCCATTTGAATCCATCTGGACCAGAGTTTGTCGTTACTTCAAGACTGACAG
ATTGAGAAGAGTCTTCTCCTTTGCCGTTATGTACATGGGTCAATCTCCATACTCTGCTCCAGGTACCTACTCCTTGTTGCA
ATACACTGAATTGACTGAAGGTATCTGGTACCCAAGAGGTGGTTTCTGGCAAGTTCCAAACACTTTGTTGCAAATCGTCAA

GAGAAACAACCCATCTGCTAAGTTCAACTTCAACGCTCCAGTTTCTCAAGTTTTGTTGTCTCCAGCTAAGGACAGAGCTAC
CGGTGTCAGATTAGAATCTGGTGAAGAACACCACGCTGATGTTGTCATTGTCAATGCTGACTTGGTCTACGCTTCTGAACA
TTTGATTCCAGATGATGCTAGAAACAAGATCGGTCAATTAGGTGAAGTTAAGCGTTCCTGGTGGGCTGATTTGGTTGGTGG
TAAGAAGTTGAAGGGTTCTTGTTCTTCTTTGTCTTTCTACTGGTCTATGGACAGAATCGTTGACGGTTTGGGTGGTCACAA
CATCTTCTTGGCTGAAGACTTCAAGGGTTCCTTCGACACCATTTTCGAAGAATTGGGTTTGCCAGCTGACCCATCTTTCTA
TGTTAACGTTCCATCCAGAATTGACCCTTCTGCTGCTCCAGAAGGTAAGGATGCCATTGTCATCTTAGTCCCATGTGGTCA
CATCGATGCTTCCAACCCTCAAGACTACAACAAATTGGTTGCCAGAGCCAGAAAGTTCGTCATCCAAACCTTGTCTGCCAA
GTTGGGTCTACCAGATTTCGAAAAGATGATTGTTGCTGAAAAGGTTCACGATGCTCCATCCTGGGAAAAGGAATTCAACTT
GAAGGACGGTTCCATTTTGGGTTTGGCTCACAACTTCATGCAAGTCTTGGGTTTCAGACCATCCACCAGACACCCAAAGTA
CGACAAATTGTTCTTTGTCGGTGCTTCTACCCACCCAGGTACTGGTGTTCCAATTGTCTTGGCTGGTGCCAAATTGACTGC
TAACCAAGTTTTGGAATCCTTCGATCGTTCTCCAGCTCCAGATCCTAACATGTCTTTGTCTGTTCCATACGGTAAGCCATT
GAAATCCAACGGTACTGGTATTGACTCTCAAGTCCAATTGAAATTCATGGACTTGGAACGTTGGGTTTACCTATTAGTCTT
GTTGATTGGTGCTGTTATCGCCAGATCCGTCGGTGTCTTGGCCTTTTAAAGGAAGTATCTCGGAAATATTAATTTAGGCCA
TGTCCTTATGCACGTTTCTTTTGATACTTACGGGTACATGTACACAAGTATATCTATATATATAAATTAATGAAAATCCCC
TATTTATATATATGACTTTAACGAGACAGAACAGTTTTTTATTTTTTATCCTATTTGATGAATGATACAGTTTCTTATTCA
CGTGTTATACCCACACCAAATCCAATAGCAATACCGGCCATCACAATCACTGTTTCGGCAGCCCCTAAGATCAGACAAAAC
ATCCGGAACCACCTTAAATCAACGTCCCTC**AAATAACCACAAACATCCTTCCCATATGCTCGGTCGTGCTTGTTGTACCT**

crtI expression cassette: conD-Kl_OLE1p-crtI-Sc_TAL1t-conE	Medium strength promoter for expression of *crtI*, multiplex array

AACGTTGTCCAGGTTTGTATCCACGTGTGTCCGTTCCGCCAATATTCCGCGTGCCAAAGGGGGGGCAGGGACAGGGATACG
ACAAGGGCTGGGGAAAAAAAAAAAGATAGATACGATTGGCCGGGTAAGCCTGGGGAAATGTAGCAAGTGCGGGTAAGTTAA
AAGGTAACCACGTGACTCCGGAAGAGTCACGTGGTTACGGACTTTTTTCTCTAGATCTCAGCTTTTTATCGGTCTTACCCT
GCCCTCCTGCCCCCTGCCCCTTCCCTTTGCCCCAAAAAGAAAGGAAATCTGTTGGATTTCGCTCAGGCCATCCCTTTCGTT
AATATCGGTTATCGCTTTACACACTGCACATCCTTCTGTCCAAAAGGAATCCAGAAGTTTAGCTTTTCCTTCCTTTCCCAC
AGACATTAGCCTAGGCCCTCTCTCATCATTTGCATGCCTCAGCCAATGTACCAAGAATAACGCAACGAGGTTGGGAAATTT
TAACCCAACAATCGATGCAGATGTGACAAGAGATTAGACACGTTCCAGATACCAGATTACACAGCTTGTGCTAGCAGAGTG
ACATATGGTGGTGTTGTGTCTCGTTTAGTACCTGTAATCGAGAGTGTTCAAATCAGTCGATTTGAACACCCTTACTGCCAC
TGAATATTGATTGAATACCGTTTATTGAAGGTTTTATGAGTGATCTTCTTTCGGTCCAGGACAATTTGTTGAGCTTTTTCT
ATGTAGAGTTCCGTCCCTTTTTTTTTTTTTTTGCTTTCTCGCACTTACTAGCACTATTTTTTTTTCACACACTAAAACAC
TTTATTTTAATCTATATATATATATATATATATATATGTAGGAATGGAATCACAGACATTTGATACTCATCCTCATCCTTATT
AATTCTTGTTTTAATTTGTTTGACTTAGCCAAACCACCAATCTCAACCCATCGTATTTCAGGTATTGTGTGTCTAGTGTGT
CTCTGGTATACGGAAATAAGTGCCAGAAGTAAGGAAGAAACAAAGAACAAGTGTCTGAATACTACTAGCCTCTCTTTTCAT
AATGGGTAAGGAACAAGACCAAGACAAGCCAACTGCCATCATCGTTGGTTGTGGTATCGGTGGTATTGCTACCGCTGCCAG
ATTAGCTAAGGAAGGTTTCCAAGTTACCGTCTTTGAAAAGAACGACTACTCCGGTGGTAGATGTTCTTTGATTGAAAGAGA
TGGTTACAGATTCGACCAAGGTCCATCTTTGTTGCTATTACCAGACTTGTTCAAGCAAACCTTCGAAGATTTGGGTGAAAA
GATGGAAGACTGGGTTGATTTGATCAAGTGTGAACCAAACTACGTTTGTCACTTCCATGATGAAGAAACTTTCACCTTCTC
CACTGACATGGCTTTATTGAAGAGAGAAGTCGAAAGATTTGAAGGTAAAGATGGTTTCGACAGATTCTTGTCTTTCATCCA
AGAAGCTCACAGACATTACGAATTGGCTGTTGTCCACGTCTTGCAAAAGAACTTCCCAGGTTTCGCTGCTTTCTTGAGATT
ACAATTCATCGGTCAAATCTTAGCTTTGCACCCATTTGAATCCATCTGGACCAGAGTTTGTCGTTACTTCAAGACTGACAG
ATTGAGAAGAGTCTTCTCCTTTGCCGTTATGTACATGGGTCAATCTCCATACTCTGCTCCAGGTACCTACTCCTTGTTGCA
ATACACTGAATTGACTGAAGGTATCTGGTACCCAAGAGGTGGTTTCTGGCAAGTTCCAAACACTTTGTTGCAAATCGTCAA
GAGAAACAACCCATCTGCTAAGTTCAACTTCAACGCTCCAGTTTCTCAAGTTTTGTTGTCTCCAGCTAAGGACAGAGCTAC

CGGTGTCAGATTAGAATCTGGTGAAGAACACCACGCTGATGTTGTCATTGTCAATGCTGACTTGGTCTACGCTTCTGAACA
TTTGATTCCAGATGATGCTAGAAACAAGATCGGTCAATTAGGTGAAGTTAAGCGTTCCTGGTGGGCTGATTTGGTTGGTGG
TAAGAAGTTGAAGGGTTCTTGTTCTTCTTTGTCTTTCTACTGGTCTATGGACAGAATCGTTGACGGTTTGGGTGGTCACAA
CATCTTCTTGGCTGAAGACTTCAAGGGTTCCTTCGACACCATTTTCGAAGAATTGGGTTTGCCAGCTGACCCATCTTTCTA
TGTTAACGTTCCATCCAGAATTGACCCTTCTGCTGCTCCAGAAGGTAAGGATGCCATTGTCATCTTAGTCCCATGTGGTCA
CATCGATGCTTCCAACCCTCAAGACTACAACAAATTGGTTGCCAGAGCCAGAAAGTTCGTCATCCAAACCTTGTCTGCCAA
GTTGGGTCTACCAGATTTCGAAAAGATGATTGTTGCTGAAAAGGTTCACGATGCTCCATCCTGGGAAAAGGAATTCAACTT
GAAGGACGGTTCCATTTTGGGTTTGGCTCACAACTTCATGCAAGTCTTGGGTTTCAGACCATCCACCAGACACCCAAAGTA
CGACAAATTGTTCTTTGTCGGTGCTTCTACCCACCCAGGTACTGGTGTTCCAATTGTCTTGGCTGGTGCCAAATTGACTGC
TAACCAAGTTTTGGAATCCTTCGATCGTTCTCCAGCTCCAGATCCTAACATGTCTTTGTCTGTTCCATACGGTAAGCCATT
GAAATCCAACGGTACTGGTATTGACTCTCAAGTCCAATTGAAATTCATGGACTTGGAACGTTGGGTTTACCTATTAGTCTT
GTTGATTGGTGCTGTTATCGCCAGATCCGTCGGTGTCTTGGCCTTTTAAAGGAAGTATCTCGGAAATATTAATTTAGGCCA
TGTCCTTATGCACGTTTCTTTTGATACTTACGGGTACATGTACACAAGTATATCTATATATATAAATTAATGAAAATCCCC
TATTTATATATATGACTTTAACGAGACAGAACAGTTTTTTATTTTTTATCCTATTTGATGAATGATACAGTTTCTTATTCA
CGTGTTATACCCACACCAAATCCAATAGCAATACCGGCCATCACAATCACTGTTTCGGCAGCCCCTAAGATCAGACAAAAC
ATCCGGAACCACCTTAAATCAACGTCCCTC**AAATAACCACAAACATCCTTCCCATATGCTCGGTCGTGCTTGTTGTACCT**

crtI expression cassette: conD-Kl_LEU2p-crtI-Sc_TAL1t-conE	Low strength promoter for expression of *crtI*, multiplex array

AACGTTGTCCAGGTTTGTATCCACGTGTGTCCGTTCCGCCAATATTCCGCGTGCGCTGTGAAGATCCCAGCAAAGGCTTAC
AAAGTGTTATCTCTTTTGAGACTTGTTGAGTTGAACACTGGTGTTTTCATCAAACTTACCAAGGACGTGTACCCATTGTTG
AAACTTGTATCACCATATATTGTTATCGGACAACCTTCACTTGCATCTATCCGTTCTTTAATCCAAAAGAGATCTAGAATA
ATGTGGCAAAGGCCAGAAGATAAAGAACCAAAAGAGATAATCTTGAATGACAACAATATCGTTGAAGAGAAATTAGGTGAT
GAAGGTGTCATTTGTATCGAGGATATCATCCATGAGATTTCGACGTTGGGCGAAAATTTCTCGAAATGTACTTTCTTCCTA
TTACCATTCAAATTGAACAGAGAAGTCAGTGGATTCGGTGCCATCTCCCGTTTGAATAAACTGAAAATGCGCGAACAAAAC
AACAAGACACGTCAAATTTCAAACGCTGCCACGGCTCCAGTTATCCAAGTAGATATCGACACAATGATTTCCAAGTTGAAT
TGATTAACTATAAAAGGAAAATATCTGTACAATAGACATCGGGCTCCCATTGGCCCTACCCACATATGTAGAAATACATTA
CTCTATTCACTACTGCATTTAGTTATGTTTAACATTTGATATAGCAGACTACCGCCAGGCACAATATATTCCCCTTCCCTC
TTGCCATTCGCTGTACTTGTGGTGGATTCCAATTCAGCGCAGTCACGTGCTAGTAATCACCGCATTTTTTTCTTTTCCTTT
CAGGCTAAAACCGGTTCCGGGCCTGATCCCTGCACTCATTTTCTAACGGAAAACCTTCAGAAGCATAACTACCCATTCCAG
TTTAGAGACATGACAGGTTCAACATCAGATGCTTCATATACTTTTATATATTGAATTATATAAATATATCTATGTACTCTA
AGTAAGTACATCTGCTTTAACGCATTCCTACATTTGCTTCGATTTATTTTTATTGTTGATACCTATTTGAAGAAGTAAAAA
GTATCCCACACTACACAGATTATACAATGGGTAAGGAACAAGACCAAGACAAGCCAACTGCCATCATCGTTGGTTGTGGTA
TCGGTGGTATTGCTACCGCTGCCAGATTAGCTAAGGAAGGTTTCCAAGTTACCGTCTTTGAAAAGAACGACTACTCCGGTG
GTAGATGTTCTTTGATTGAAAGAGATGGTTACAGATTCGACCAAGGTCCATCTTTGTTGCTATTACCAGACTTGTTCAAGC
AAACCTTCGAAGATTTGGGTGAAAAGATGGAAGACTGGGTTGATTTGATCAAGTGTGAACCAAACTACGTTTGTCACTTCC
ATGATGAAGAAACTTTCACCTTCTCCACTGACATGGCTTTATTGAAGAGAGAAGTCGAAAGATTTGAAGGTAAAGATGGTT
TCGACAGATTCTTGTCTTTCATCCAAGAAGCTCACAGACATTACGAATTGGCTGTTGTCCACGTCTTGCAAAAGAACTTCC
CAGGTTTCGCTGCTTTCTTGAGATTACAATTCATCGGTCAAATCTTAGCTTTGCACCCATTTGAATCCATCTGGACCAGAG
TTTGTCGTTACTTCAAGACTGACAGATTGAGAAGAGTCTTCTCCTTTGCCGTTATGTACATGGGTCAATCTCCATACTCTG
CTCCAGGTACCTACTCCTTGTTGCAATACACTGAATTGACTGAAGGTATCTGGTACCCAAGAGGTGGTTTCTGGCAAGTTC
CAAACACTTTGTTGCAAATCGTCAAGAGAAACAACCCATCTGCTAAGTTCAACTTCAACGCTCCAGTTTCTCAAGTTTTGT
TGTCTCCAGCTAAGGACAGAGCTACCGGTGTCAGATTAGAATCTGGTGAAGAACACCACGCTGATGTTGTCATTGTCAATG

CTGACTTGGTCTACGCTTCTGAACATTTGATTCCAGATGATGCTAGAAACAAGATCGGTCAATTAGGTGAAGTTAAGCGTT
CCTGGTGGGCTGATTTGGTTGGTGGTAAGAAGTTGAAGGGTTCTTGTTCTTCTTTGTCTTTCTACTGGTCTATGGACAGAA
TCGTTGACGGTTTGGGTGGTCACAACATCTTCTTGGCTGAAGACTTCAAGGGTTCCTTCGACACCATTTTCGAAGAATTGG
GTTTGCCAGCTGACCCATCTTTCTATGTTAACGTTCCATCCAGAATTGACCCTTCTGCTGCTCCAGAAGGTAAGGATGCCA
TTGTCATCTTAGTCCCATGTGGTCACATCGATGCTTCCAACCCTCAAGACTACAACAAATTGGTTGCCAGAGCCAGAAAGT
TCGTCATCCAAACCTTGTCTGCCAAGTTGGGTCTACCAGATTTCGAAAAGATGATTGTTGCTGAAAAGGTTCACGATGCTC
CATCCTGGGAAAAGGAATTCAACTTGAAGGACGGTTCCATTTTGGGTTTGGCTCACAACTTCATGCAAGTCTTGGGTTTCA
GACCATCCACCAGACACCCAAAGTACGACAAATTGTTCTTTGTCGGTGCTTCTACCCACCCAGGTACTGGTGTTCCAATTG
TCTTGGCTGGTGCCAAATTGACTGCTAACCAAGTTTTGGAATCCTTCGATCGTTCTCCAGCTCCAGATCCTAACATGTCTT
TGTCTGTTCCATACGGTAAGCCATTGAAATCCAACGGTACTGGTATTGACTCTCAAGTCCAATTGAAATTCATGGACTTGG
AACGTTGGGTTTACCTATTAGTCTTGTTGATTGGTGCTGTTATCGCCAGATCCGTCGGTGTCTTGGCCTTTTAAAGGAAGT
ATCTCGGAAATATTAATTTAGGCCATGTCCTTATGCACGTTTCTTTTGATACTTACGGGTACATGTACACAAGTATATCTA
TATATATAAATTAATGAAAATCCCCTATTTATATATATGACTTTAACGAGACAGAACAGTTTTTTATTTTTTATCCTATTT
GATGAATGATACAGTTTCTTATTCACGTGTTATACCCACACCAAATCCAATAGCAATACCGGCCATCACAATCACTGTTTC
GGCAGCCCCTAAGATCAGACAAAACATCCGGAACCACCTTAAATCAACGTCCCTC**AAATAACCACAAACATCCTTCCCATA**
TGCTCGGTCGTGCTTGTTGTACCT

Integration flank DNA sequences

Name	**Purpose**
INT1 5’-con5	Multiplex array

CACTATAGCAATCTGGCTATATGATATGCAGAAAATAGTGATTACTGTGCTCTTTGCTGCATCATGCATCATATGTTTGGC
ATTCTGTCCTGTCAGATGAATGGGCGAAACAATTCGGGGAGCTTTGTTGCGACTTGGTACCCGGCAAGCCGCGACCTATAT
TTTGCATTAAACGAGTTCATCGTGGATTTATCGCCATGAGCCTTAGCTTATCAAGCCTCTCACAGACAATTAAGCAATGAA
AAAGGTATACCATCGGCGCAGAATGGTTAACTAGTGGGTTCATACTGCTGTGTTATAGATTGTTACCTAAGTGATCACCAA
AAAAAAGTGCAAAAAGGAAAAAAAAATAAGAGACAGGTAACTTCCACAAGCTTATTCTTCCAAAAATCAATCTTATCTTCA
TGCCAGCAATAGTTGCGTGCTGAGCTCAACAGTGCCCAACCCTTGTGCACCGTAGAATTGTAGAATACAAATACATAAATA
AGTGTGTTCCCGAAGGACTAAGGAATGACGGCAGAGGAGTC**AAGCGACTTCCAATCGCTTTGCATATCCAGTACCACACCC**
ACAGGCGTTT

conA-INT1 3’	Multiplex array

TTGCCCATCGAACGTACAAGTACTCCTCTGTTCTCTCCTTCCTTTGCTTTAAGCGTTGAAGTTTCCTCTTTGTATATTTGA
GATCTTCATTTTATCGGATTCTTTGTCATCAGACAACTTGTTGAGTGGTACTAAAGGAGTGCTTTTCATCATCCTTTTGGT
GAACGATTTCAAATACGTTAGTGTTTTCTGAGCTAGTTTTGATCAATTCAGGTGATTCGTTATCAGAACTCTCAGGTTTGT
ATTCGTGTCCAGTTGTGTAGCATTCGCCTAACGTGTAAGCACGGATTTCTTCCTCAGAAATTTCACTGTATGGAATCATGC
CCTTCTTTCTCGCTTCTTCGTCGGTAAATGCACCATAGTAATCTTTGTCATCATGTCTAACAGTAATTTTGAATGGGAAGA
AGACACATAGCCCCCAGTAAACGAAAAAAGAAATCAAAAAGGAGAAGAAAGAATCACCATAAAAGAATTTAACAATACCTG
AGTCGTGGAAATAGTTATTGTTGACTTCCCAAGCGATACCAGGTAGACCGGGAGCCATACCACACACCCAGGCAACGATAG
CTCTCCAGTTGACA

INT2 5’-conB	Multiplex array

AGAAGATTTCTCTTCAATCTCCTTCTCCTTAACACTTAATTTTGGCACTTCGTCAACTTTAATTTTCTTGGAAGCCATTGC
TATGTATTCTTAGTGTACTGCGACTTGGGTTTCTTTCTTGCTTTGTATAAAAGTGCGATGAGTTCAATCATCAAGTGAATT
TTTTTTTTTCAGCCTGCATTGAAAAAAAAAAAATTGTATGCTCGAGCGATGGGAAGGTTTTTTCCTCTTCCCGTTCAAAGC
TTCCTCGGACCAACTCCACTTTTTCTCGGCCCGTTGTAAAGAAATGGGTAAGAGAACTAAATTGCAATAAGCTTCGAATTA
TCGTTTTGATCCAGACATAAAGAGATGTACAAACTACAAAACAAGAGTAAAATAACTCGAACGAAGGGTACACCTTCTGTA
ATAAAACAAATTCACTATAACTGCCTTATTGGAGTGAAAAAACAAGATTATGTAGACCCCCTCTCGCGATAAAGAACATTA
ATGACTGCAACACCCGCACACCACAAGGTGAAGGGCAAAGGTGAATTAACCAAAGTGAAGAGGACGACGTAGAATCAACCA
GGTATCCTTGATAAGGG**CGGATCGATGTACACAACCGACTGCACCCAAACGAACACAAATCTTAGCA**

conC-INT2 3’	Multiplex array

ACGCTTTCCGGCATCTTCCAGACCACAGTATATCCATCCGCCTCCTGTTGGGCGATTACACAAGCGGTGGCGTCTTGAAAA
AAGTTGGGTTTCTTATCTGTACGGCGATCACTCCCCTCATATGATAGCTGTGTAGAGTGACTTGCTCTTTGATTGGCCTTA
CTTCTTGCATCTATCATCCTTTCTTTTACTCTCTCCAATCACCACAACTCATATCATTCTACATTTCATTTTATTGTTCTG
TACAATTTTTAGGTTTGTAAAATATATAAGTGATGTACATATCTTCTATATCCTTTCCGTTTGTGTCTATTTTAGGGTTGG
CTTATTACGCATTATTCCTACTTTATTTTTACTTGCTTACTTGAAAAAGAAGGTATACCACAGTGCTGCTCCACTGCAAAC
AATATAAAATATTATCAATTATTTTTTTTTTTCTTCTTTTGAATTAGATCACTAATAAACTCTTATAATGCTCAACCATCC
CAGCCAAGGCTCAGATGATGCTCAGGACGAAAAGCAAGGCGACTTCCCGGTCATCGAAGAGGAGA

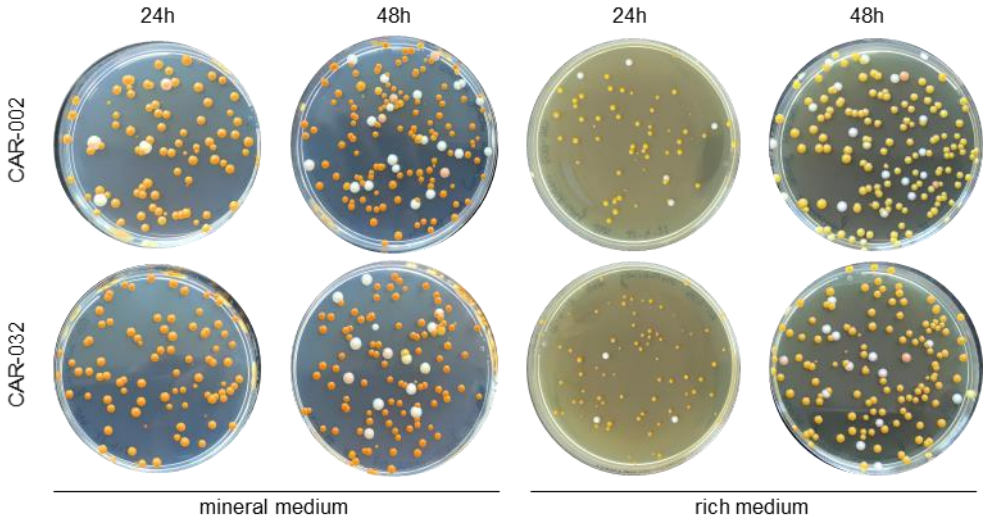
INT3 5’-conD	Multiplex array

GGTCGTTTTTGTGCAGCATATTGTCCTCTAGATGCAAACTCTGCAGGTCCATTTGCAGTAAAGTGAGTTGCCTCTCGAAGA
ATCATTAATTTCGTATAACCGTCACTATTAAAGTCAGAAAATAAATTCTGTCGTAGACAATGTTACCATAATGTTCTTGTC
CATTTTGCATACACTTTAAATATTCATTTGATTTCTCAGGGTTCATGATCATAATAAATTGCGCATTCGCAAGGCGGTAGT
ATTATAATGGGGTCCATCATTCTGTAGCAAGAAGTTACAGTACGCTGTTCAAGCGTTAAACAAGATAAGTAATCTCGAATG
AAACATTCATATTTCGCATGAGCCAACATACAGTTGCTGAGTAATCTTCATTGCGCTTATTTATCGGCATTGAGATTGTAA
AGGAAGTAAAACGCATTTTTGCAGATCTGTTCTCTTATGTATTTTTAATCGTCCTTGTATGGAAGTATCAAAGGGGACGTT
CTTCACCTCCTTGGAA**AACGTTGTCCAGGTTTGTATCCACGTGTGTCCGTTCCGCCAATATTCCGC**

conE-INT3 3’	Multiplex array

AAATAACCACAAACATCCTTCCCATATGCTCGGTCGTGCTTGTTGTACCTGATGGGACGTCAGCACTGTACTTGTTTTTGC
GACTAGATTGTAAATCATTCTTTATTTAATCTCTTTCTTTAACTACTGCTTAAAGTATAATTTGGTCCGTAGTTTAATAAC
TATACTAAGCGTAACAATGCATACTGACATTATAAGCCTGAACATTACGAGTTTAAGTTGTATGTAGGCGTTCTGTAAGAG
GTTACTGCGTAAATTATCAACGAATGCATTGGTGTATTTGCGAAAGCTACTTCTTTTAACAAGTATTTACATAAGAATAAT
GGTGATCTGCTCAACTGATTTGGTGATAACTCTAACTTTTTTAGCAACAATTTAAAAGATAATTCGAACATATATAACAGT
AGGAAGAATTTGTGTACGTCAAATTAAGATAATTTAGCATTACCAAAGTTATTAACCTAAACATAAAATATATATGAGACA
CATGTGGAAATCGTATGAAACAACTGTTATGAAACTGACAAGAATGAATATATAGAGTAAGCTC

4 Supplementary information to Chapter 4



SI Figure 4-1. Genetic instability of strains CAR-002 and CAR-034 after fermentation in mineral medium.

Preventing genetic instability during carotenoids quantification: A prolonged cultivation in mineral medium lead to mutations in the production pathway and incapability to form β-carotene (Chapter 2). To prevent this behaviour during characterization of carotenogenic strains in mineral medium, cultivation conditions were adapted. Pre-cultures were started from strains restreaked on rich agar instead of directly inoculation from the glycerol stocks. Precultures were grown for 24h and production cultures for 48h (instead of 72h for both steps respectively). Harvest after 24h and 48h was followed by testing instability of strains CAR-002 and CAR-032 (Figure 4.1). The number of white colonies increased after 48h in comparison to 24h on both mineral and rich agar. However, a good growth was observed under all tested conditions and in none of the cases population of white cells overgrown carotenoids producing orange colonies.

SI Table 4-1. Genotype of strains constructed in this study

Strain name	Genotype
CEN.PK113–7D	MAT*a MAL2–8c SUC2*
CAR-002	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sc_FBA1p - Xden_crtE - Sc_TDH3t - conA - Sc_TEF1p - Xden_crtYB - Sc_PDC1t - conB - Kl_ENO1p - Xden_crtI - Sc_TAL1t - conC - INT1
CAR-022 (SSM)	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sbay_TDH3p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Smik_TEF1p – Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_OLE1p - Xden_crtI - Sc_TAL1t - conE - INT3
CAR-023 (SSW)	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sbay_TDH3p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Smik_TEF1p – Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_LEU2p - Xden_crtI - Sc_TAL1t - conE - INT3
CAR-024 (SMS)	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sbay_TDH3p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Kl_TEF2p – Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_ENO1p - Xden_crtI - Sc_TAL1t - conE - INT3
CAR-025 (SWS)	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sbay_TDH3p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Kl_YDR1p – Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_ENO1p - Xden_crtI - Sc_TAL1t - conE - INT3
CAR-026 (SMM)	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sbay_TDH3p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Kl_TEF2p – Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_OLE1p - Xden_crtI - Sc_TAL1t - conE - INT3
CAR-027 (SWW)	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sbay_TDH3p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Kl_YDR1p – Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_LEU2p - Xden_crtI - Sc_TAL1t - conE - INT3
CAR-028 (MSS)	MAT*a MAL2–8c SUC2* + INT1 - con5 - Kl_PGK1p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Smik_TEF1p – Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_ENO1p - Xden_crtI - Sc_TAL1t - conE - INT3
CAR-029 (MWW)	MAT*a MAL2–8c SUC2* + INT1 - con5 - Kl_PGK1p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Kl_YDR1p – Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_LEU2p - Xden_crtI - Sc_TAL1t - conE - INT3
CAR-030 (WSS)	MAT*a MAL2–8c SUC2* + INT1 - con5 - Kl_TDH2p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Smik_TEF1p – Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_ENO1p - Xden_crtI - Sc_TAL1t - conE - INT3
CAR-031 (WMM)	MAT*a MAL2–8c SUC2* + INT1 - con5 - Kl_TDH2p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Kl_TEF2p – Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_OLE1p - Xden_crtI - Sc_TAL1t - conE - INT3
CAR-032	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sc_FBA1p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Sc_TEF1p - Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_ENO1p - Xden_crtI - Sc_TAL1t - conE – INT3
CAR-033 (-SS)	MAT*a MAL2–8c SUC2* + INT1 - con5 - conA - INT1; INT2 - conB - Smik_TEF1p – Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_ENO1p - Xden_crtI -Sc_TAL1t - conE - INT3
CAR-034 (SSS)	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sbay_TDH3p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Smik_TEF1p – Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_ENO1p - Xden_crtI - Sc_TAL1t - conE - INT3
CAR-035 (MMM)	MAT*a MAL2–8c SUC2* + INT1 - con5 - Kl_PGK1p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Kl_TEF2p – Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_OLE1p - Xden_crtI - Sc_TAL1t - conE - INT3
CAR-036 (WWW)	MAT*a MAL2–8c SUC2* + INT1 - con5 - Kl_TDH2p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Kl_YDR1p – Xden_crtYB - Sc_PDC1t - conC - INT2; INT3 - conD - Kl_LEU2p - Xden_crtI - Sc_TAL1t - conE - INT3

5 Supplementary information to Chapter 5

SI Table 5-1. Integration sites used for strains construction

Integration site	Chromosome	Location
INT1	15	Non-coding region between *NTR1* (YOR071c) and *GYP1* (YOR070c)
INT2	11	Non-coding region between *SRP40* (YKR092C) and *PTR2* (YKR093W)
INT3	16	Ty4 long terminal repeat (YPRCtau3) [250]
INT4	10	Non-coding region between *SOD1* (YJR104C) and *ADO1* (YJR105W)

SI Table 5-2. List of strains used in this study

Strain name	Genotype	Origin
CEN.PK113–7D	MAT*a MAL2–8c SUC2*	[128]
CEN.PK113–9D	MAT*a ura3-52 trp1-289 MAL2-8c SUC2*	[128]
CSN0001	MAT*a MAL2–8c SUC2* + pCSN061	[65]
CSN0004	MAT*a MAL2–8c SUC2* + pCSN067	[65]
FR003	MAT*a MAL2–8c SUC2* + INT4 - conA - Sc_FBA1p - mCherry - Sc_TDH3t - conB - Kl_ENO1p - eGFP - Sc_TAL1t - conC – INT4	This study
FR007	MAT*a MAL2–8c SUC2* + INT1 - con5 - conA - INT1; INT2 - conB - Sbay_TDH3p - eGFP - Sc_TDH3t - conC - INT2; INT3 - conD - Sc_FBA1p - mCherry - Sc_TDH3t - conE - INT3	This study
FR008	MAT*a MAL2–8c SUC2* + INT1 - con5 - conA - INT1; INT2 - conB - Smik_TEF1p - eGFP - Sc_PDCt - conC - INT2; INT3 - conD - Sc_FBA1p - mCherry - Sc_TDH3t - conE - INT3	This study
FR009	MAT*a MAL2–8c SUC2* + INT1 - con5 - conA - INT1; INT2 - conB - Kl_ENO1p - eGFP - Sc_TAL1t - conC - INT2; INT3 - conD - Sc_FBA1p - mCherry - Sc_TDH3t - conE - INT3	This study
FR013	MATa *ura3-52 trp1-289 MAL2-8c SUC2* + INT1 - con5 - Sbay_TDH3p - BFP - Sc_TDH3t - conA - INT1; INT2 - conB - Smik_TEF1p - mCherry - Sc_PDCt - conC - INT2; INT3 - conD - Kl_ENO1p - eGFP - Sc_TAL1t - conE - INT3	This study
FR014	MATa *ura3-52 trp1-289 MAL2-8c SUC2* + INT1 - con5 - Sbay_TDH3p - BFP - Sc_TDH3t - conA - INT1; INT2 - conB - Smik_TEF1p - mCherry - Sc_PDCt - conC - INT2; INT3 - conD - Kl_ENO1p - eGFP - Sc_TAL1t - conE - INT3; INT4 - Sc_TEF1p - dLbCas12a E925A - NLS – Mxi1 - GND2t - INT4	This study
DC001	MAT*a MAL2–8c SUC2* + INT4 - Sc_TEF1p - LbCas12a - NLS - GND2t - conA - Sc_FBA1p - mCherry - Sc_TDH3t - conB - Kl_ENO1p - eGFP - Sc_TAL1t - INT4	This study
DC002	MAT*a MAL2–8c SUC2* + INT4 - Sc_TEF1p - dLbCas12a D832A - NLS - GND2t - conA - Sc_FBA1p - mCherry - Sc_TDH3t - conB - Kl_ENO1p - eGFP - Sc_TAL1t - INT4	This study
DC003	MAT*a MAL2–8c SUC2* + INT4 - Sc_TEF1p - dLbCas12a E925A - NLS - GND2t - conA - Sc_FBA1p - mCherry - Sc_TDH3t - conB - Kl_ENO1p - eGFP - Sc_TAL1t - INT4	This study
DC004	MAT*a MAL2–8c SUC2* + INT4 - Sc_TEF1p - dLbCas12a D832A - NLS - KRAB - GND2t - conA - Sc_FBA1p - mCherry - Sc_TDH3t - conB - Kl_ENO1p - eGFP - Sc_TAL1t - INT4	This study
DC005	MAT*a MAL2–8c SUC2* + INT4 - Sc_TEF1p - dLbCas12a D832A - NLS - Mxi1 - GND2t - conA - Sc_FBA1p - mCherry - Sc_TDH3t - conB - Kl_ENO1p - eGFP - Sc_TAL1t - INT4	This study
DC006	MAT*a MAL2–8c SUC2* + INT4 - Sc_TEF1p - dLbCas12a D832A - NLS - MIG1 - GND2t - conA - Sc_FBA1p - mCherry - Sc_TDH3t - conB - Kl_ENO1p - eGFP - Sc_TAL1t - INT4	This study
DC007	MAT*a MAL2–8c SUC2* + INT4 - Sc_TEF1p - dLbCas12a D832A - NLS - TUP1 - GND2t - conA - Sc_FBA1p - mCherry - Sc_TDH3t - conB - Kl_ENO1p - eGFP - Sc_TAL1t - INT4	This study
DC008	MAT*a MAL2–8c SUC2* + INT4 - Sc_TEF1p - dLbCas12a D832A - NLS – UME6 - GND2t - conA - Sc_FBA1p - mCherry - Sc_TDH3t - conB - Kl_ENO1p - eGFP - Sc_TAL1t - INT4	This study
DC009	MAT*a MAL2–8c SUC2* + INT4 - Sc_TEF1p - dLbCas12a E925A - NLS - KRAB - GND2t - conA - Sc_FBA1p - mCherry - Sc_TDH3t - conB - Kl_ENO1p - eGFP - Sc_TAL1t - INT4	This study
DC010	MAT*a MAL2–8c SUC2* + INT4 - Sc_TEF1p - dLbCas12a E925A - NLS - Mxi1 - GND2t - conA - Sc_FBA1p - mCherry - Sc_TDH3t - conB - Kl_ENO1p - eGFP - Sc_TAL1t - INT4	This study
DC011	MAT*a MAL2–8c SUC2* + INT4 - Sc_TEF1p - dLbCas12a E925A - NLS - MIG1 - GND2t - conA - Sc_FBA1p - mCherry - Sc_TDH3t - conB - Kl_ENO1p - eGFP - Sc_TAL1t - INT4	This study
DC012	MAT*a MAL2–8c SUC2* + INT4 - Sc_TEF1p - dLbCas12a E925A - NLS - TUP1 - GND2t - conA - Sc_FBA1p - mCherry - Sc_TDH3t - conB - Kl_ENO1p - eGFP - Sc_TAL1t - INT4	This study
DC013	MAT*a MAL2–8c SUC2* + INT4 - Sc_TEF1p - dLbCas12a E925A - NLS - UME6 - GND2t - conA - Sc_FBA1p - mCherry - Sc_TDH3t - conB - Kl_ENO1p - eGFP - Sc_TAL1t - INT4	This study

CAR-034	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sbay_TDH3p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Smik_TEF1p - Xden_crtYB - Sc_PDC1t - conC – INT2; INT3- conD - Kl_ENO1p - Xden_crtI - Sc_TAL1t - conE - INT3	[66]
CAR-041	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sbay_TDH3p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Smik_TEF1p - Xden_crtYB - Sc_PDC1t - conC – INT2; INT3- conD - Kl_ENO1p - Xden_crtI - Sc_TAL1t - conE - INT3; INT4 - Sc_TEF1p - dLbCas12a E925A - NLS – Mxi1 - GND2t - INT4	This study
CAR-042	MAT*a MAL2–8c SUC2* + INT1 - con5 - Sbay_TDH3p - Xden_crtE - Sc_TDH3t - conA - INT1; INT2 - conB - Smik_TEF1p - Xden_crtYB - Sc_PDC1t - conC – INT2; INT3- conD - Kl_ENO1p - Xden_crtI - Sc_TAL1t - conE - INT3; INT4 - Sc_PGI1p - dLbCas12a E925A - NLS – Mxi1 - GND2t - INT4	This study

The constructed strains were name according to the following rule: FR - reporter strains with fluorescent protein genes integrated; DC - dCas12a variants integrated into genomic DNA in combination with fluorescent protein genes; CAR - carotenoids producing strains. All strains construction were done in the CEN.PK113-7D background, except for FR013 and FR014 which were performed in the CEN.PK113-9D background.

SI Table 5-3. List of plasmids used in this study

Plasmid	Genotype	Reference	Addgene
pCSN061	*CEN6/ARS TRP1* KanMX bla Kl11p-SpCas9-GND2t	[65]	# 101725
pCSN067	*CEN6/ARS TRP1* KanMX bla Kl11p-LbCas12a-GND2t	[65]	# 101748
pC-Mxi1 (pCSN069)	*CEN6/ARS TRP1* KanMX bla TEF1p-dLbCas12a E925A-Mxi1-GND2t	This study	# 166728
pC-Mxi1-NLS (pCSN071)	*CEN6/ARS TRP1* KanMX bla TEF1p-dLbCas12a E925A-Mxi1-NLS-GND2t	This study	# 166729
pC-NLS-Mxi1-NLS (pCSN072)	*CEN6/ARS TRP1* KanMX bla TEF1p-dLbCas12a E925A-NLS-Mxi1-NLS-GND2t	This study	# 166730
pC-NLS (pCSN074)	*CEN6/ARS TRP1* KanMX bla TEF1p-dLbCas12a E925A-NLS-GND2t	This study	# 166731
pC-NLS-Mxi1 (pCSN075)	*CEN6/ARS TRP1* KanMX bla TEF1p-dLbCas12a E925A-NLS-Mxi1-GND2t	This study	# 166732
dCas12a-eGFP (pCSN078)	*CEN6/ARS TRP1* KanMX bla TEF1p-dLbCas12a E925A-NLS-Mxi1-eGFP-GND2t	This study	# 171629
pRN1120	2µm NatMX amp^R	[65]	# 101750

SI Table 5-4. List of primers used in this study

Name	Sequence	Purpose
pKC001	CTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCGCGGTCTTTGAAAAGATAATG	Amplification of Cas9 sgRNA expression cassette (RV)
pKC002	AAAATACAACAAATAAAAAACACTCAATGACCTGACCATTTGATGGAGTTCCGCGGAGACATAAAAAAC	Amplification of Cas9 sgRNA expression cassette (FW)
pKC003	AACTCCATCAAATGGTCAGG	Amplification of Cas9 sgRNA recipient plasmid pRN1120 (FW)
pKC004	AACAAAAGCTGGGTACCGGG	Amplification of Cas9 sgRNA recipient plasmid pRN1120 (RV)
pKC005	CATGTTTGACAGCTTATCATC	Amplification of Cas12a single crRNA array expression cassette (FW)
pKC006	CACACAGGAAACAGCTATGAC	Amplification of Cas12a single crRNA array expression cassette (RV)
pKC007	TACCGTCGACCTCGAGGGG	Amplification of Cas12a crRNA recipient plasmid pRN1120 (FW)
pKC008	ATATCGAATTCCTGCAGCCCG	Amplification of Cas12a crRNA recipient plasmid pRN1120 (RV)
pKC009	GGTCTCGAATGTCTAAAGGTGAAGAATTATTCA	Amplification of GFP with BsaI recognition site for GGC (FW)
pKC010	GGTCTCCTTTATTTGTACAATTCATCCATACCA	Amplification of GFP with BsaI recognition site for GGC (RV)
pKC011	GGTCTCGAATGGTAAGTAAGGGTGAAGAAGA	Amplification of mCherry with BsaI recognition site for GGC (FW)
pKC012	GGTCTCCTTTATTTGTACAATTCGTCCATTCC	Amplification of mCherry with BsaI recognition site for GGC (RV)
pKC013	GGTCTCGAATGTCCGAATTGATCAAGGAAAAC	Amplification of BFP with BsaI recognition site for GGC (FW)
pKC014	GGTCTCCTTTAGTTTAATTTGTGACCCAACTTG	Amplification of BFP with BsaI recognition site for GGC (RV)
pKC015	CACTATAGCAATCTGGCTATATG	Amplification of INT1 left flank (FW)
pKC016	AAACGCCTGTGGGTGTGGTACTGGATATGCAAAGCGATTGGAAGTCGCTTGACTCCTCTGCCGTCATTCC	Amplification of INT1 left flank (RV) with homology to con5

pKC017	TTGCCCATCGAACGTACAAGTACTCCTCTGTTCTCTCCTTCCTTTGCTTTAAGCGTTGAAGTTTCCTCTTTG	Amplification of INT1 right flank (FW) with homology to conA
pKC018	TGTCAACTGGAGAGCTATCG	Amplification of INT1 right flank (RV)
pKC019	AGAAGATTTCTCTTCAATCTC	Amplification of INT2 left flank (FW)
pKC020	TGCTAAGATTTGTGTTCGTTTGGGTGCAGTCGGTTGTGTACATCGATCCGCCCTTATCAAGGATACCTGGTTG	Amplification of INT2 left flank (RV) with homology to conB
pKC021	ACGCTTTCCGGCATCTTCCAGACCACAGTATATCCATCCGCCTCCTGTTGGGCGATTACACAAGCGGTGG	Amplification of INT2 right flank (FW) with homology to conC
pKC022	TCTCCTCTTCGATGACCGGG	Amplification of INT2 right flank (RV)
pKC023	GGTCGTTTTTGTGCAGCATATTG	Amplification of INT3 left flank (FW)
pKC024	GCGGAATATTGGCGGAACGGACACACGTGGATACAAACCTGGACAACGTTTTCCAAGGAGGTGAAGAACG	Amplification of INT3 left flank (RV) with homology to conD
pKC025	AAATAACCACAAACATCCTTCCCATATGCTCGGTCGTGCTTGTTGTACCTGATGGGACGTCAGCACTGTAC	Amplification of INT3 right flank (FW) with homology to conE
pKC026	GAGCTTACTCTATATATTCATTC	Amplification of INT3 right flank (RV)
pKC027	GATAACGACGCTTCTGCCTAC	Amplification of INT4 left flank (FW)
pKC028	AAAGCAAAGGAAGGAGAGAACAGAGGAGTACTTGTACGTTCGATGGGCAAGTGTGTCGGAATTAGTAAGCGG	Amplification of INT4 left flank (RV) with homology to conA
pKC029	ACGCTTTCCGGCATCTTCCAGACCACAGTATATCCATCCGCCTCCTGTTGGATGCCGATAATTTTCACATGGA	Amplification of INT4 right flank (FW) with homology to conC
pKC030	GTTACTAAACTGGAACTGTCCG	Amplification of construct integrated into INT1 (FW)
pKC031	GCTGTCATGATCTGTGATAAC	Amplification of construct integrated into INT1 (RV)
pKC032	CACTGGAACTTGAGCTTGAG	Amplification of construct integrated into INT2 (FW)
pKC033	CCAACGTGCCTTAAAGTCTG	Amplification of construct integrated into INT2 (RV)
pKC034	CTCTCATGAAGCAGTCAAGTC	Amplification of construct integrated into INT3 (FW)
pKC035	CTGGTTACTTCCCTAAGACTG	Amplification of construct integrated into INT3 (RV)
pKC036	GGAACCCACGTTCTGCGTTAG	Amplification of construct integrated into INT4 (FW)

pKC037	CTTGGTACATAGACTTGACACC	Amplification of construct integrated into INT4 (RV)
pKC038	GCGATGGCGAGCCCAGCGGAAGGGATGTCCG CTTACTAATTCCGACACACTTGGCTGATAAT AGCGTATAAACAA	Amplification of Sc_TEF1p with 50 bp homology to integration site (INT4) (FW)
pKC039	AAAGTCTTAGACAAAGAGTAACAGTTGGTGA ATTTTTCCAACTTAGACATTTTGTAATTAAA ACTTAGATTAGATTG	Amplification of Sc_TEF1p with 50 bp homology to integration site (INT4) (RV)
pKC040	ATGTCTAAGTTGGAAAAATTCACCA	Amplification of Cas12a ORF or dCas12a ORF with SRAD linker and SV40 NLS (FW)
pKC041	TACCTTTCTCTTCTTCTTTGGATC	Amplification of Cas12a ORF or dCas12a ORF with SRAD linker and SV40 NLS (RV)
pKC042	TCCAGAGCTGATCCAAAGAAGAAGAGAAAGG TATAAAGGAGTTAAAGGCAAAGTTTTCTT	Amplification of GND2t with 36 bp homology to SRAD linker and SV40 NLS (FW)
pKC043	AGGAGTTAAAGGCAAAGTTTTCTT	Amplification of GND2t (used for constracts with repression domains) (FW)
pKC044	AAAGCAAAGGAAGGAGAGAACAGAGGAGTAC TTGTACGTTCGATGGGCAAAACAATAGCGAT CCGAAAGGCG	Amplification of GND2t with 50 bp homology to connector A (RV)
pKC045	AAACCTCCGTCAAGCACTCCA	Amplification of constructs carrying repression domains (FW)
pKC046	AATTATTTGTGGGAACGGCTCTA	Amplification of constructs carrying repression domains (RV)
pKC047	TTGCCCATCGAACGTACAAGTACTCCTCTGT TCTCTCCTTCCTTTGCTTTCTACTTGGCTTC ACATACGTTG	Amplification of Sc_FBA1p with 50 bp homology to connector A (FW)
pKC048	ATGAATTCCTTAATGATCGCCATATTGTCTT CTTCACCCTTACTTACCATTTTGAATATGTA TTACTTGGTTATGG	Amplification of Sc_FBA1p with 50 bp homology to mCherry ORF (RV)
pKC049	ATGGTAAGTAAGGGTGAAGAAGA	Amplification of mCherry ORF (FW)
pKC050	TTATTTGTACAATTCGTCCATTCC	Amplification of mCherry ORF (RV)
pKC051	GAGCTGAAGGTAGGCACTCAACGGGTGGAAT GGACGAATTGTACAAATAAAGTGAATTTACT TTAAATCTTGCATT	Amplification of Sc_TDH3t with 50 bp homology to mCherry ORF (FW)
pKC052	TGCTAAGATTTGTGTTCGTTTGGGTGCAGTC GGTTGTGTACATCGATCCGATCCTGGCGGAA AAAATTCATTTG	Amplification of Sc_TDH3t with 50 bp homology to connector A (FW)

pKC053	CGGATCGATGTACACAACCGACTGCACCCAAACGAACACAAATCTTAGCACGTATCCCTATCTGGATTAACAT	Amplification of Kl_ENO1p with 50 bp homology to connector B (FW)
pKC054	TCAACCAAAATTGGGACAACACCAGTGAATAATTCTTCACCTTTAGACATTTTTTTTTGTTTTGTGGGTTGGTTT	Amplification of Sc_FBA1p with 50 bp homology to eGFP ORF (RV)
pKC055	ATGTCTAAAGGTGAAGAATTATTCA	Amplification of eGFP ORF (FW)
pKC056	TTATTTGTACAATTCATCCATACCA	Amplification of GFP ORF (RV)
pKC057	TTGTTACTGCTGCTGGTATTACCCATGGTATGGATGAATTGTACAAATAAAGGAAGTATCTCGGAAATATTAATT	Amplification of Sc_TAL1t with 50 bp homology to GFP ORF (FW)
pKC058	CATCACGTACGACGACACTGCGACTCATCCATGTGAAAATTATCGGCATCGACGTTGATTTAAGGTGGTTCC	Amplification of Sc_TAL1t with 50 bp homology to integration site (INT4) (FW)
pKC059	CGTGTTCGTTTACTGTTAATTGGTGCTTGGTACATAGACTTGACACC	Amplification of template from INT4 for MinION sequencing (RV)
pKC060	GCACATAAGCAATATCGTAGTCCGTCTTGGTACATAGACTTGACACC	Amplification of template from INT4 for MinION sequencing (RV)
pKC061	AGATGATGGGTACGGGCCTCTAATACTTGGTACATAGACTTGACACC	Amplification of template from INT4 for MinION sequencing (RV)
pKC062	CATCCAACACTCTACGCCCTCTTCACTTGGTACATAGACTTGACACC	Amplification of template from INT4 for MinION sequencing (RV)
pKC063	AGAGCTAGAAGGGCACCCTGCAGTTCTTGGTACATAGACTTGACACC	Amplification of template from INT4 for MinION sequencing (RV)
pKC064	GGAAAGGGAATTATTTCGTAAGGCGGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC065	TAGTTTCGAACCACGGTTACTAATCGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC066	ATGAGTGTCAGCGAGTGTAACTCGAGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC067	CAATCCCGAGGCCTGACGCGACATAGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC068	CCACAGCGAGGAAGTAAACTGTTATGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)

pKC069	TTGTACTAATCGGCTTCAACGTGCCGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC070	CGCCTGCTACAACAGGAGTATCAAAGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC071	AAGCCCATCGTGGTCCTTAGACTTGGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC072	TCCGCAAGGCGCAGCAGTGCACAAGGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC073	CAAATGACAATTAACCACCGTGTATGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC074	ATGCCATTGGCTCTTAGACAGCCCGGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC075	CAGTTTACCGCATCTTGACCTAACTGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC076	CATTTGTATGTTCACCTATCTACTAGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC077	TGGCAGTAAGTACGCCTTCTGAATTGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC078	ACCTCTATCACCGCTTCATGCTAAGGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC079	TACGCTCATCTAATACATCCAACTCGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC080	TAATGAGAACAACCACACCATAGCGGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC081	AGCAATCACTTGCATAGCTGCGTATGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC082	CAAAAGATGATAAGCTCCGGCAAGCGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC083	AATATTGAACAACGCAAGGATCGGCGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC084	AGTCGTCCGCGGAGCACTCTGGTAAGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC085	GGACAAGCATAGCATAGCCATTTATGGAACC CACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)

pKC086	CAGAAATCCGCGTGATTACGAGTCGGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC087	CCCTGCGGTCACGTCTATAGAAATTGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC088	ACACCTTTAACCCTCCTGAGAACCGGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC089	AGAACCCGGCCCCACGCAATGGAACGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC090	GCCGAACGTAAGTATAGCGCATATTGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC091	CCATAAACGAGCTATTAGTTATGAGGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC092	ATGGTACATTTGGACAATATTGAATGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC093	GCCACGGCGAATGTCGGGGAGACAGGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC094	TATACCAGGTGTCCTGTGAGCAGCGGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC095	TAATAGCACACACGGGGCAATACCAGGAACCCACGTTCTGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC096	AGGAGTTAAAGGCAAAGTTTTC	Amplification of pCN061 backbone without Cas9 (FW)
pKC097	CACTATAGGGCGAATTGGGTACC	Amplification of pCN061 backbone without Cas9 (RV)
pKC098	TTGGCTGATAATAGCGTATAAACA	Amplification of dCas12a with homology to pRN backbone (FW)
pKC099	AACAATAGCGATCCGAAAGGCG	Amplification of dCas12a with homology to pRN backbone (RV)
pKC100	GTGCTTGACGGAGGTTTGAGC	Amplification of dCas12a without C-terminal NLS (RV) and pCSN backbone with dCas12a (RV) to create pC-Mxi1, pC-Mxi1-NLS, pC-NLS-Mxi1-NLS
pKC101	AGGAGTTAAAGGCAAAGTTTTCTT	Amplification of pCSN backbone with dCas12a (FW)
pKC102	AGAACCAGCAGAACCTGGAGCACCACCAGCAGCACCACCAGAAGATCTTGGAGATGGCATGGATGG	Amplification of dCas12a, backbone and linker from pCSN069 (RV)

pKC103	GCTGGTGGTGCTCCAGGTTCTGCTGGTTCTGCTGCTGGTTCTGGTATGTCTAAAGGTGAAGAATTATTCA	Amplification of eGFP and 45bp of linker for dCas12a-eGFP fusion (FW)
pKC104	GGGAACGGCTCTAGAAAAGAAAACTTTGCCTTTAACTCCTTTATTTGTACAATTCATCCATACCA	Amplification of eGFP and 45bp of GND2t for dCas12a-eGFP fusion (FW)

SI Table 5-5. Features of sgRNA and crRNA design

Cas protein	Purpose	PAM	Spacer length	Direct repeat sequence
Cas9	Singleplex genome editing	NGG	20 nt	GTTTTAGAGCTAGAAATAGCAAGT TAAAATAAGGCTAGTCCGTTATCA ACTTGAAAAAGTGGCACCGAGTCG GTGGTGC
Cas12a	Multiplex genome editing	TTTV	23 nt	AATTTCTACTAAGTGTAGAT
dCas12a	Singleplex CRISPRi	TTTV	20 nt	TAATTTCTACTAAGTGTAGAT
dCas12a	Multiplex CRISPRi	TTTV	23 nt	AATTTCTACTAAGTGTAGAT

SI Table 5-6. List of spacers used for singleplex genome editing and CRISPR interference experiments

Name	Target	TSS distance	Strand	Spacer sequence	PAM	Vendor
Spacers used for singleplex genome editing using Cas9						
gCAS9	INT4	-	T	TCAGGACATGATCTCAGTAG	CGG	IDT
Spacers used for singleplex CRISPRi using dCas12a						
gENO1	Kl_ENO1p	-455	T	CTATGTCTCTCACACGCGTG	TTTG	Twist
gENO2	Kl_ENO1p	-319	NT	CACGCATCCCAACTGTCTTT	TTTA	Twist
gENO3	Kl_ENO1p	-139	NT	GACCACCACCAACCAATTGA	TTTG	Twist
gENO4	Kl_ENO1p	-106	T	TATAACCTCAAATCATGGGA	TTTA	Twist
gENO5	Kl_ENO1p	-5	T	TTTTGTGGGTTGGTTTGTTA	TTTG	Twist
gENO6	Kl_ENO1p	+27	NT	ATTCGAAACAAACCAAACCA	TTTA	Twist
gGFP1	eGFP	+94 to +110	T	TGACCATTAACATCACCATC	TTTG	Twist
gGFP2	eGFP	+139 to +155	T	CCGTAAGTAGCATCACCTTC	TTTA	Twist
gGFP3	eGFP	+302 to +318	T	TTGAACATAACCTTCTGGCA	TTTC	Twist
gGFP4	eGFP	+436 to +452	T	TGACCTAAAATGTTACCATC	TTTG	Twist
gTDH1	Sbay_TDH3p	-297	T	AGGAAGCCAGATTAAAAAGC	TTTC	Twist
gTDH2	Sbay_TDH3p	-216	T	CTCCAGTCAGTTTGCATCAC	TTTG	Twist
gTDH3	Sbay_TDH3p	-171	T	AGTTTTTTCCAACCCAGAGA	TTTC	Twist
gTDH4	Sbay_TDH3p	-121	T	TTAATCAAACAGGGGGATGA	TTTA	Twist
gTDH5	Sbay_TDH3p	+18	T	TGTGTGTGTGTTTTCGAAAC	TTTG	Twist
gTEF1	Smik_TEF1p	-265	NT	CCCTCTTTGTCTCTTTCGTG	TTTC	IDT
gTEF2	Smik_TEF1p	-121	NT	GGATTTTTTCTCTTTCGATG	TTTG	IDT
gTEF3	Smik_TEF1p	-44	NT	TTCCTCTTGTTTAGTCATTC	TTTG	IDT
gTEF4	Smik_TEF1p	-31	NT	GTCATTCTTCTTCTCAGCAT	TTTA	IDT
gNone	Dummy	-	-	ACTAACCCATCAACCTGTAC	–	Twist

Strand indicates whereas gRNA encodes template (T) or non-template (NT) strand. Position indicates proximity of gRNA's midpoint from transcription start site (estimated according to [96]).

SI Table 5-7. List of spacers in array used for multiplex genome editing and CRISPR interference experiments

Name	Target	Position	Strand	Spacer sequence	PAM
Spacers in single crRNA array used for simultaneous genome editing using Cas12a					
gINT1	INT1	-	NT	CTGGTGGGAGAGAAAGCTTATGA	TTTG
gINT2	INT2	-	NT	GTGCCGTACGCCGGAGCCGACGG	TTTG
gINT3	INT3	-	NT	TGCCCCTCTTATACGATTATATT	TTTA
Spacers in single crRNA array used for simultaneous CRISPRi using dCas12a					
gENO6	Kl_ENO1p	959	NT	ATTCGAAACAAACCAAACCAATT	TTTA
gTDH1	Sbay_TDH3p	129	T	AGGAAGCCAGATTAAAAAGCGGG	TTTC
gTEF4	Smik_TEF1p	360	NT	GTCATTCTTCTTCTCAGCATTGG	TTTA
gNone	dummy	-	-	ACTAACCCATCAACCTGTACTCG	–
gNone2	dummy	-	-	AGTCTACGCCTCCTACGGTGTGC	–
gNone3	dummy	-	-	TTACACGCATTGGCTAGAGCAGC	–

Strand indicates whereas gRNA encodes template (T) or non-template (NT) strand.

SI Table 5-8. Order of spacers in single crRNA arrays used for multiplex CRISPR interference experiments

Name	Graphical representation	Spacer in position 1	Spacer in position 2	Spacer in position 3
Array_1		gTDH1	gTEF4	gENO6
Array_2		gTDH1	gENO6	gTEF4
Array_3		gTEF4	gTDH1	gENO6
Array_4		gENO6	gTDH1	gTEF4
Array_5		gTEF4	gENO6	gTDH1
Array_6		gENO6	gTEF4	gTDH1
Array_7		gENO6	gNone2	gNone3
Array_8		gNone	gENO6	gNone3
Array_9		gNone	gNone2	gENO6
Array_10		gNone	gNone2	gNone3

SI Table 5-9. Architecture of gRNAs expressed from RNAP II and III promoters

gGAL1

CATGTTTGACAGCTTATCATCGATAATCCGGAGCTAGCATGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGC
TGCAGCATTTGAATAAGAAGTAATACAAACCGAAAATGTTGAAAGTATTAGTTAAAGTGGTTATGCAGTTTTT
GCATTTATATATCTGTTAATAGATCAAAAATCATCGCTTCGCTGATTAATTACCCCAGAAATAAGGCTAAAAA
ACTAATCGCATTATCATCCTATGGTTGTTAATTTGATTCGTTCATTTGAAGGTTTGTGGGGCCAGGTTACTGC
CAATTTTTCCTCTTCATAACCATAAAAGCTAGTATTGTAGAATCTTTATTGTTCGGAGCAGTGCGGCGCGAGG
CACATCTGCGTTTCAGGAACGCGACCGGTGAAGACGAGGACGCACGGAGGAGAGTCTTCCTTCGGAGGGCTGT
CACCCGCTCGGCGGCTTCTAATCCGTACTTCAATATAGCAATGAGCAGTTAAGCGTATTACTGAAAGTTCCAA
AGAGAAGGTTTTTTTAGGCTAAGATAATGGGGCTCTTTACATTTCCACAACATATAAGTAAGATTAGATATGG
ATATGTATATGGATATGTATATGGTGGTAATGCCATGTAATATGATTATTAAACTTCTTTGCGTCCATCCAAA
AAAAAAGTAAGAATTTTTGAAAATTC***TAATTTCTACTAAGTGTAGAT***TATAACCTCAAATCATGGGA*AGGAGT*
TAAAGGCAAAGTTTTCTTTTCTAGAGCCGTTCCCACAAATAATTATACGTATATGCTTCTTTTCGTTTACTAT
ATATCTATATTTACAAGCCTTTATTCACTGATGCAATTTGTTTCCAAATACTTTTTTGGAGATCTCATAACTA
GATATCATGATGGCGCAACTTGGCGCTATCTTAATTACTCTGGCTGCCAGGCCCGTGTAGAGGGCCGCAAGAC
CTTCTGTACGCCATATAGTCTCTAAGAACTTGAACAAGTTTCTAGACCTATTGCCGCCTTTCGGATCGCTATT
*GTT*GGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCA
TAGCTGTTTCCTGTGTG

gGAL2

CATGTTTGACAGCTTATCATCGATAATCCGGAGCTAGCATGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGC
TGCAGCATTTGAATAAGAAGTAATACAAACCGAAAATGTTGAAAGTATTAGTTAAAGTGGTTATGCAGTTTTT
GCATTTATATATCTGTTAATAGATCAAAAATCATCGCTTCGCTGATTAATTACCCCAGAAATAAGGCTAAAAA
ACTAATCGCATTATCATCCTATGGTTGTTAATTTGATTCGTTCATTTGAAGGTTTGTGGGGCCAGGTTACTGC
CAATTTTTCCTCTTCATAACCATAAAAGCTAGTATTGTAGAATCTTTATTGTTCGGAGCAGTGCGGCGCGAGG
CACATCTGCGTTTCAGGAACGCGACCGGTGAAGACGAGGACGCACGGAGGAGAGTCTTCCTTCGGAGGGCTGT
CACCCGCTCGGCGGCTTCTAATCCGTACTTCAATATAGCAATGAGCAGTTAAGCGTATTACTGAAAGTTCCAA
AGAGAAGGTTTTTTTAGGCTAAGATAATGGGGCTCTTTACATTTCCACAACATATAAGTAAGATTAGATATGG
ATATGTATATGGATATGTATATGGTGGTAATGCCATGTAATATGATTATTAAACTTCTTTGCGTCCATCCAAA
AAAAAAGTAAGAATTTTTGAAAATTC***TAATTTCTACTAAGTGTAGAT***TATAACCTCAAATCATGGGA***TAATTT***
CTACTAAGTGTAGAT*AGGAGTTAAAGGCAAAGTTTTCTTTTCTAGAGCCGTTCCCACAAATAATTATACGTAT*
ATGCTTCTTTTCGTTTACTATATATCTATATTTACAAGCCTTTATTCACTGATGCAATTTGTTTCCAAATACT
TTTTTGGAGATCTCATAACTAGATATCATGATGGCGCAACTTGGCGCTATCTTAATTACTCTGGCTGCCAGGC
CCGTGTAGAGGGCCGCAAGACCTTCTGTACGCCATATAGTCTCTAAGAACTTGAACAAGTTTCTAGACCTATT
*GCCGCCTTTCGGATCGCTATTGTT*GGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTCCG
AGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG

gGAL3

CATGTTTGACAGCTTATCATCGATAATCCGGAGCTAGCATGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGC
TGCAGCATTTGAATAAGAAGTAATACAAACCGAAAATGTTGAAAGTATTAGTTAAAGTGGTTATGCAGTTTTT
GCATTTATATATCTGTTAATAGATCAAAAATCATCGCTTCGCTGATTAATTACCCCAGAAATAAGGCTAAAAA
ACTAATCGCATTATCATCCTATGGTTGTTAATTTGATTCGTTCATTTGAAGGTTTGTGGGGCCAGGTTACTGC
CAATTTTTCCTCTTCATAACCATAAAAGCTAGTATTGTAGAATCTTTATTGTTCGGAGCAGTGCGGCGCGAGG
CACATCTGCGTTTCAGGAACGCGACCGGTGAAGACGAGGACGCACGGAGGAGAGTCTTCCTTCGGAGGGCTGT
CACCCGCTCGGCGGCTTCTAATCCGTACTTCAATATAGCAATGAGCAGTTAAGCGTATTACTGAAAGTTCCAA
AGAGAAGGTTTTTTTAGGCTAAGATAATGGGGCTCTTTACATTTCCACAACATATAAGTAAGATTAGATATGG
ATATGTATATGGATATGTATATGGTGGTAATGCCATGTAATATGATTATTAAACTTCTTTGCGTCCATCCAAA
AAAAAAGTAAGAATTTTTGAAAATTC*AAATTACTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTC****TAAT***
TTCTACTAAGTGTAGAT**TATAACCTCAAATCATGGGA**GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGC
TGGGCAACATGCTTCGGCATGGCGAATGGGAC*AGGAGTTAAAGGCAAAGTTTTCTTTTCTAGAGCCGTTCCCA*
CAAATAATTATACGTATATGCTTCTTTTCGTTTACTATATATCTATATTTACAAGCCTTTATTCACTGATGCA
ATTTGTTTCCAAATACTTTTTTGGAGATCTCATAACTAGATATCATGATGGCGCAACTTGGCGCTATCTTAAT
TACTCTGGCTGCCAGGCCCGTGTAGAGGGCCGCAAGACCTTCTGTACGCCATATAGTCTCTAAGAACTTGAAC
*AAGTTTCTAGACCTATTGCCGCCTTTCGGATCGCTATTGTT*GGGGGGCCCGGTACCCAGCTTTTGTTCCCTTT
AGTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG

gSNR1

CATGTTTGACAGCTTATCATCGATAATCCGGAGCTAGCATGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGC
TGCAGTCTTTGAAAAGATAATGTATGATTATGCTTTCACTCATATTTATACAGAAACTTGATGTTTTCTTTCG
AGTATATACAAGGTGATTACATGTACGTTTGAAGTACAACTCTAGATTTTGTAGTGCCCTCTTGGGCTAGCGG
TAAAGGTGCGCATTTTTTCACACCCTACAATGTTCTGTTCAAAAGATTTTGGTCAAACGCTGTAGAAGTGAAA
GTTGGTGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATC***TAATTTCTACTAAGTGTA***
GAT**TATAACCTCAAATCATGGGA***TTTTTTTGTTTTTTATGTCTGGGGGGCCCGGTACCCAGCTTTTGTTCCCT
TTAGTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG

gSNR2

CATGTTTGACAGCTTATCATCGATAATCCGGAGCTAGCATGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGC
TGCAGTCTTTGAAAAGATAATGTATGATTATGCTTTCACTCATATTTATACAGAAACTTGATGTTTTCTTTCG
AGTATATACAAGGTGATTACATGTACGTTTGAAGTACAACTCTAGATTTTGTAGTGCCCTCTTGGGCTAGCGG
TAAAGGTGCGCATTTTTTCACACCCTACAATGTTCTGTTCAAAAGATTTTGGTCAAACGCTGTAGAAGTGAAA
GTTGGTGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATC*AAATTACTGATGAGTCCG*
*TGAGGACGAAACGAGTAAGCTCGTC****TAATTTCTACTAAGTGTAGAT*****TATAACCTCAAATCATGGGA*GGCCGGC
ATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAATGGGAC*TTTTTTTGTTTT*
*TTATGTCT*GGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTCCGAGCTTGGCGTAATCAT
GGTCATAGCTGTTTCCTGTGTG

Legend:

GAL10 promoter (RNAP II)

SNR52 promoter (RNAP III)

LbCas12a direct repeat

Spacer gENO4

Hammer head ribozyme

HDV ribozyme

GND2 terminator

SUP4 terminator

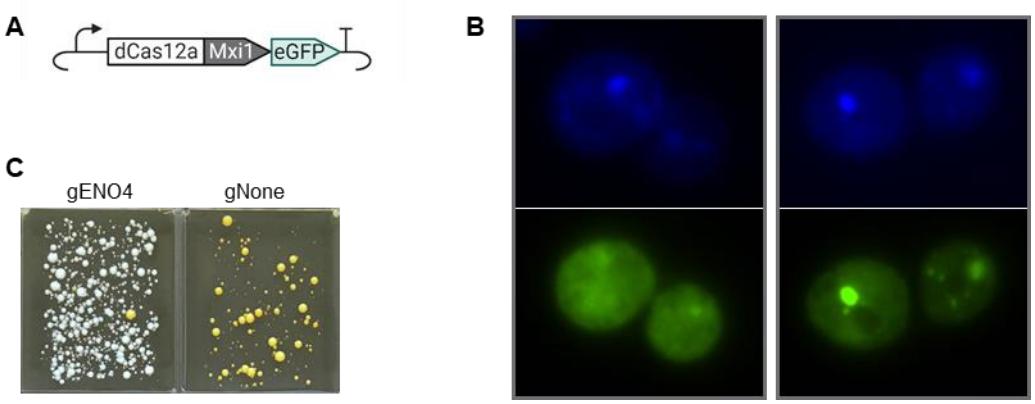
Homology to pRN1120 backbone

SI Table 5-10. Accession numbers of sequences submitted to GenBank

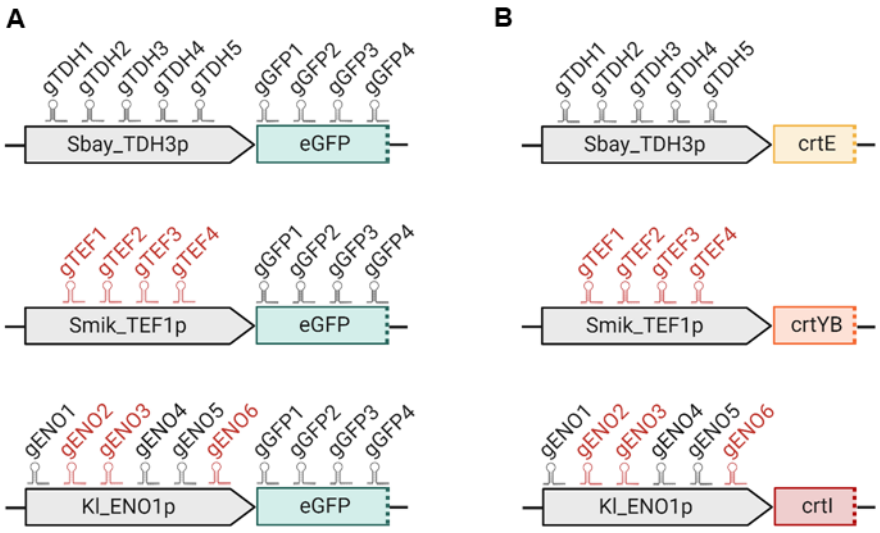
Construct name	GenBank accession number	Content
DC002	MW584243	Integration site INT4 of strain DC002 with expression cassette containing dLbCas12a_D832A
DC003	MW584244	Integration site INT4 of strain DC003 with expression cassette containing dLbCas12a_E925A
DC004	MW766349	Integration site INT4 of strain DC004 with expression cassette containing dLbCas12a_D832A fused to SV40 and KRAB domain
DC005	MW584245	Integration site INT4 of strain DC005 with expression cassette containing dLbCas12a_D832A fused to SV40 and Mxi1 domain
DC006	MW766350	Integration site INT4 of strain DC005 with expression cassette containing dLbCas12a_D832A fused to SV40 and MIG1 domain
DC007	MW766351	Integration site INT4 of strain DC005 with expression cassette containing dLbCas12a_D832A fused to SV40 and TUP1 domain
DC008	MW766352	Integration site INT4 of strain DC005 with expression cassette containing dLbCas12a_D832A fused to SV40 and UME6 domain
DC009	MW584246	Integration site INT4 of strain DC009 with expression cassette containing dLbCas12a_E925A fused to SV40 and KRAB domain
DC010	MW584247	Integration site INT4 of strain DC010 with expression cassette containing dLbCas12a_E925A fused to SV40 and Mxi1 domain
DC011	MW584248	Integration site INT4 of strain DC011 with expression cassette containing dLbCas12a_E925A fused to SV40 and MIG1 domain
DC012	MW584249	Integration site INT4 of strain DC012 with expression cassette containing dLbCas12a_E925A fused to SV40 and TUP1 domain
DC013	MW584250	Integration site INT4 of strain DC013 with expression cassette containing dLbCas12a_E925A fused to SV40 and UME6 domain
FR003_INT4	MW584251	Integration site INT4 of strain FR003 with cassette containing mCherry and eGFP expressed from Kl_ENO1p
FR007,_FR008,_ FR009_INT1	MW584252	Integration site INT1 of strains FR007, FR008, FR009 with cassette containing con5 and conA

FR007,_FR008,_ FR009_INT3	MW584253	Integration site INT3 of strains FR007, FR008, FR009 with cassette containing mCherry
FR007_INT2	MW584254	Integration site INT2 of strains FR007 with cassette containing eGFP expressed from Sbay_TDH3p
FR008_INT2	MW584255	Integration site INT2 of strains FR008 with cassette containing eGFP expressed from Smik_TEF1p
FR009_INT2	MW584256	Integration site INT2 of strains FR009 with cassette containing eGFP expressed from Kl_ENO1p
FR013_INT1	MW584257	Integration site INT1 of strains FR013 with cassette containing BFP expressed from Sbay_TDH3p
FR013_INT2	MW584258	Integration site INT2 of strains FR013 with cassette containing mCherry expressed from Smik_TEF1p
FR013_INT3	MW584259	Integration site INT3 of strains FR013 with cassette containing eGFP expressed Kl_ENO1p
gSNR1 (gENO4, G12034)	MW584260	LbCas12a gRNA expressed from SNR2p (RNAP III) with one direct repeat
gSNR2	MZ222245	LbCas12a gRNA expressed from SNR2p (RNAP III) with two ribozymes
gGAL1	MZ222243	LbCas12a gRNA expressed from GAL10p (RNAP II) with one direct repeat
gGAL2	MZ222244	LbCas12a gRNA expressed from GAL10p (RNAP II) with two direct repeats
gGAL3 (G12053)	MW584261	LbCas12a gRNA expressed from GAL10p (RNAP II) with two ribozymes
pC-Mxi1 (pCSN069)	MW584262	Single copy yeast plasmid expressing DNase-dead dLbCas12a E925A fused to Mxi1, codon optimized for *Saccharomyces cerevisiae*
pC-Mxi1-NLS (pCSN071)	MW584263	Single copy yeast plasmid expressing DNase-dead dLbCas12a E925A fused to Mxi1 and NLS, codon optimized for *Saccharomyces cerevisiae*

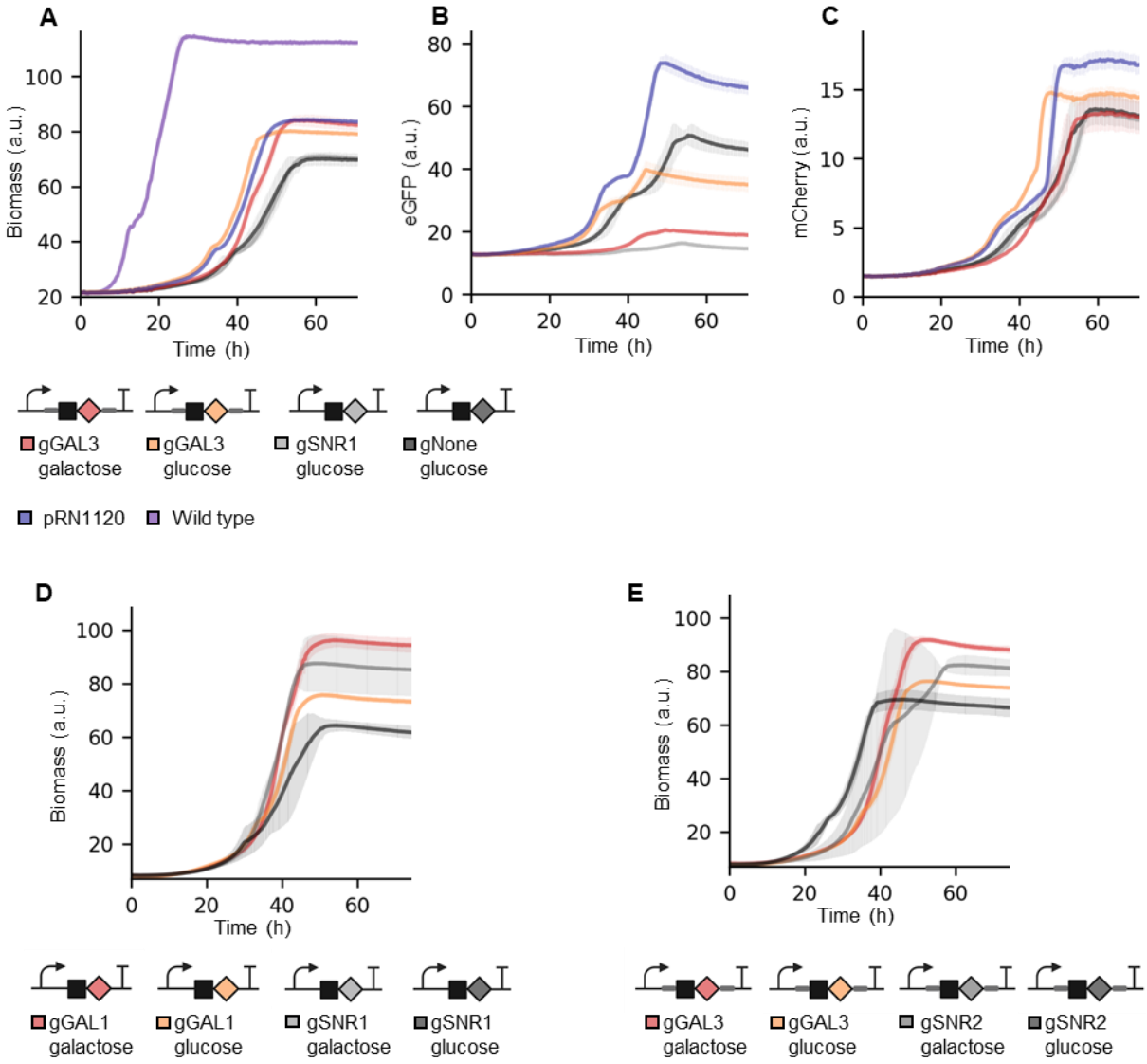
pC-NLS-Mxi1-NLS (pCSN072)	MW584264	Single copy yeast plasmid expressing DNase-dead dLbCas12a E925A fused to NLS, Mxi1 and NLS, codon optimized for *Saccharomyces cerevisiae*
pC-NLS (pCSN074)	MW584265	Single copy yeast plasmid expressing DNase-dead dLbCas12a E925A fused to NLS, codon optimized for *Saccharomyces cerevisiae*
pC-NLS-Mxi1 (pCSN075)	MW584266	Single copy yeast plasmid expressing DNase-dead dLbCas12a E925A fused to NLS and Mxi, codon optimized for *Saccharomyces cerevisiae*
dCas12a-eGFP (pCSN078)	MZ222246	Single copy yeast plasmid expressing DNase-dead dLbCas12a E925A fused to Mxi1 and eGFP, codon optimized for *Saccharomyces cerevisiae*



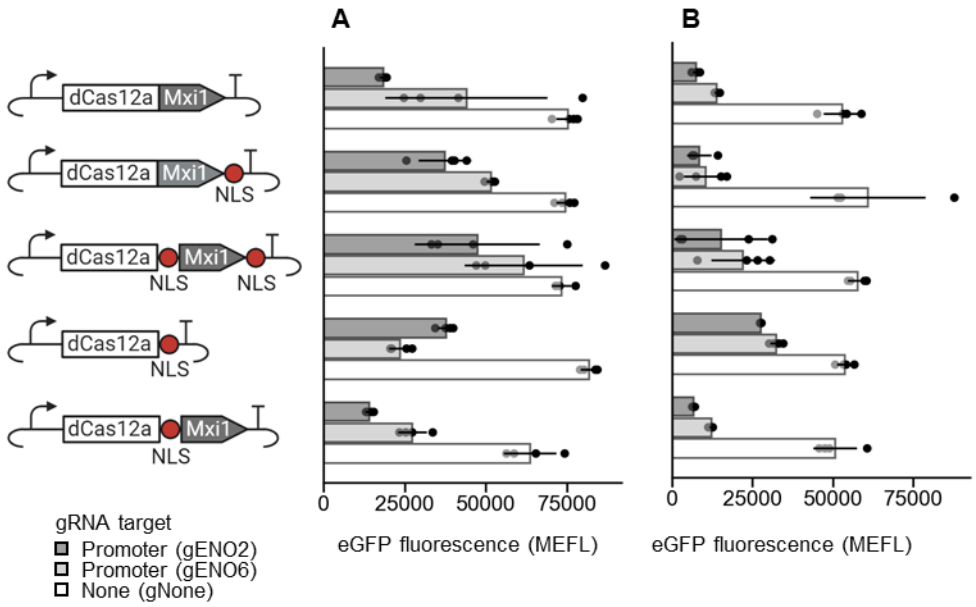
SI Figure 5-1. Localization of dCas12a-eGFP fusion in *S. cerevisiae* cells. A. Map of plasmid pCSN078, **B**. Fluorescent microscopy of yeast cells transformed with pCSN078 and stained with Vybrant® DyeCycle™ Violet dye to visualize nucleus (top) and dCas12a-eGFP fusion (bottom). **C**. Downregulation of carotenoids production, exemplified by colour loss of the transformants (appearance of white colonies is caused by suppression of β-carotene biosynthesis), in strain CAR-034 using pCSN078 and the gRNA targeting the Kl_ENO promoter controlling crtI (gENO4) or the non-targeting control (gNone).



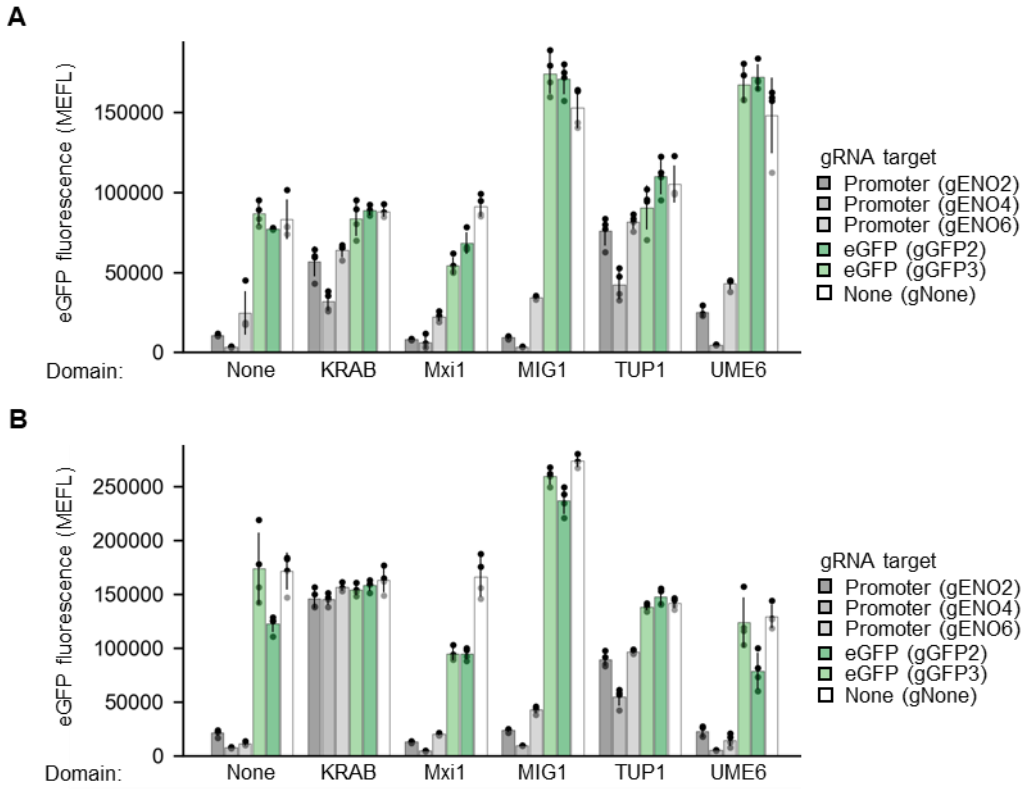
SI Figure 5-2. Overview on the gRNAs used for gene downregulation. A. gRNAs used to downregulate *eGFP*. gTDH1-5, gTEF1-4 and gENO1-6 target three different promoters in front of *eGFP* and gGFP1-4 target ORF of *eGFP*. gNone served as a non-targeting control (not depicted). **B**. gRNAs used to downregulate β-carotene biogenesis. gRNAs target promoters in front of crt genes, as described for *eGFP*. gRNAs encoding target on the template strand are depicted in black and non-template strand in red.



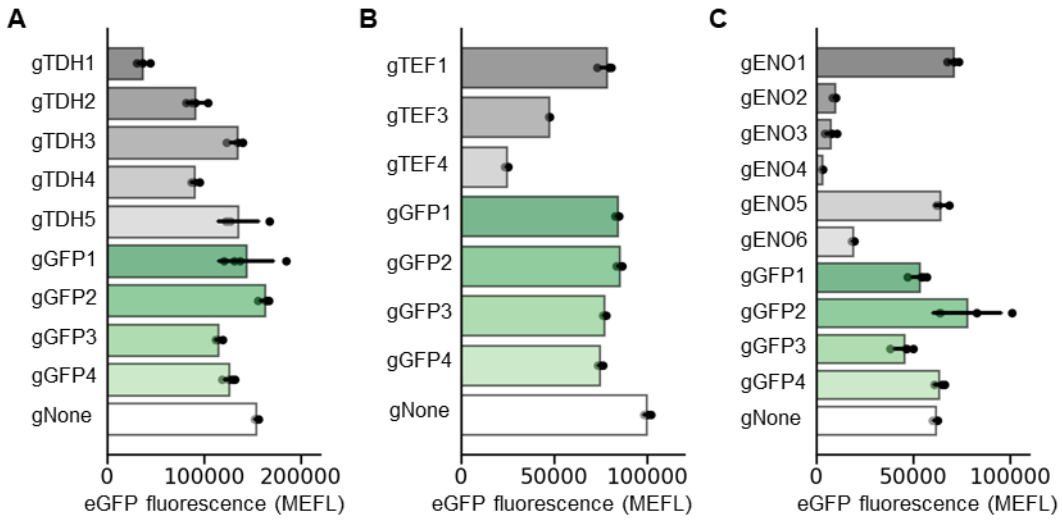
SI Figure 5-3. Growth and fluorescence of strains transformed with gRNAs expressed from RNAP II and III promoters. A. Biomass, **B**. GFP and **C**. mCherry fluorescence for the strains set comparing RNAP II gRNA with ribozymes, RNAP III gRNA with a single direct repeat (DR), empty gRNA recipient plasmid (pRN1120) and wild type strain; **D**. biomass for the strains set comparing RNAP II and RNAP III gRNA with a single DR; **E**. biomass for the strains set comparing RNAP II and RNAP III gRNA with ribozymes. Growth curves represent mean ± 1 standard deviation (n=3).



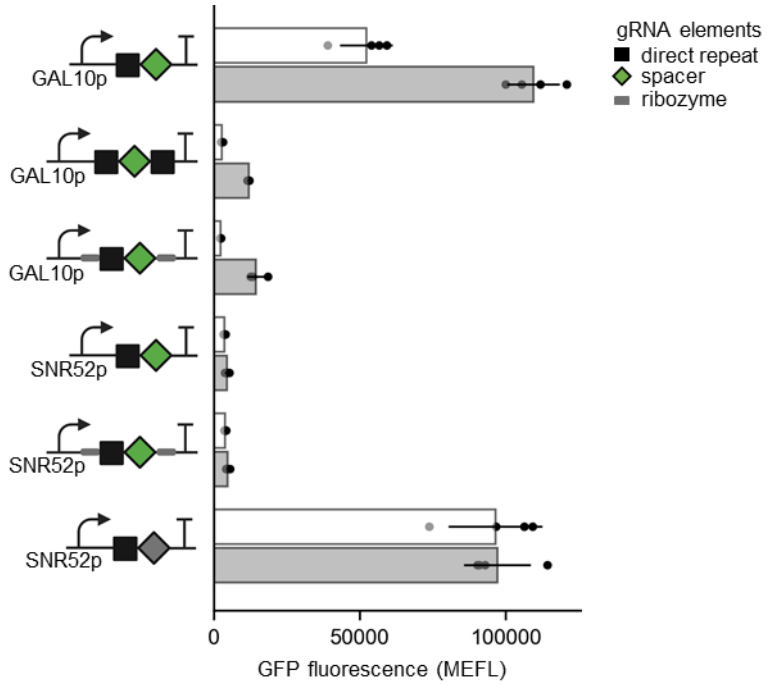
SI Figure 5-4. Absolute fluorescence of eGFP achieved with plasmids testing NLS position and number A. in strain FR003, and **B**. FR009.



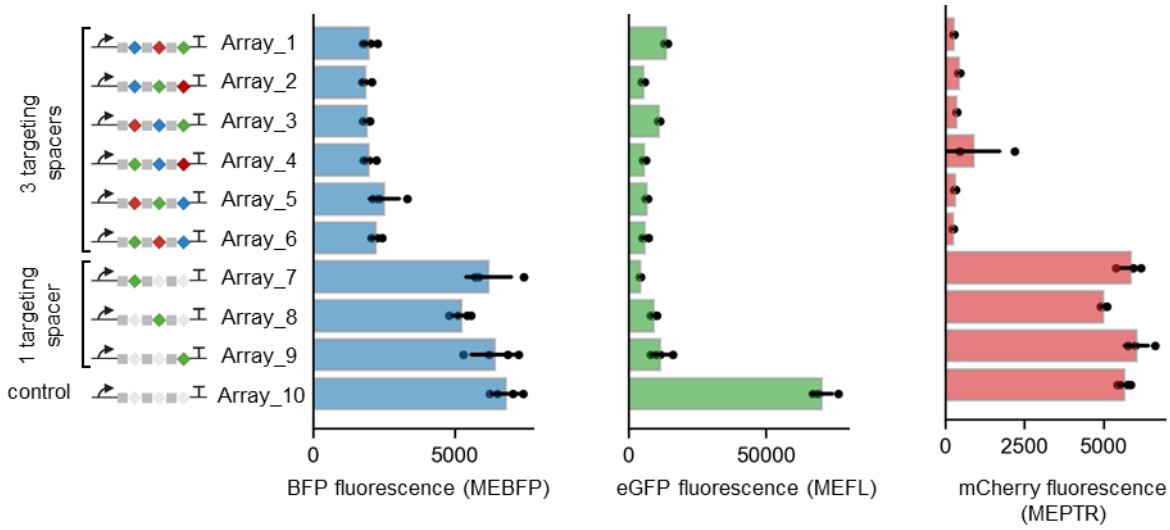
SI Figure 5-5. Absolute fluorescence of eGFP achieved with strains DC002-013 testing repression domains. A. in combination with dCas12a D832A and **B**. dCas12a E925A.



SI Figure 5-6. Absolute fluorescence of eGFP achieved by targeting dCas12a E925A Mxi1 to promoters controlling *eGFP* or *eGFP* ORF. A. gRNAs targeting eGFP and *S. bayanus TDH3* promoter in strain FR007, **B**. gRNAs targeting *eGFP* and *S. mikatae TEF1* promoter in strain FR008, **C**. gRNAs targeting *eGFP* and *K. lactis ENO1* promoter in strain FR009.



SI Figure 5-7. Absolute fluorescence of eGFP achieved with gRNAs expressed from RNAP II and III promoters. RNAP II promoter – *GAL10* or RNAP III promoter – *SNR52*.



SI Figure 5-8. Absolute fluorescence of eGFP, BFP and mCherry achieved in multiplex downregulation using single crRNA array.

6 Supplementary information to Chapter 6

SI Table 6-1. List of primers used in this study

Name	Sequence	Purpose
pKC001	GCGATGGCGAGCCCAGCGGAAGGGATGTCCGCTTACTA ATTCCGACACACTTGGCTGATAATAGCGTATAAACAA	Amplification of Sc_TEF1p with 50 bp homology to integration site (INT4) (FW)
pKC002	TGTTTCTTGTAGTGGGAGACACCGTCAATCTTGGTAAT CTTCATGTACATTTTGTAATTAAAACTTAGATTAGATT G	Amplification of Sc_TEF1p with 50 bp homology to LwaCas13a (RV)
pKC003	GCGATGGCGAGCCCAGCGGAAGGGATGTCCGCTTACTA ATTCCGACACACCATTTGAATAAGAAGTAATACAAAC	Amplification of Sc_GAL10p with 50 bp homology to integration site (INT4) (FW)
pKC004	TGTTTCTTGTAGTGGGAGACACCGTCAATCTTGGTAAT CTTCATGTACATGAATTTTCAAAAATTCTTACTTTTTT TT	Amplification of Sc_GAL10p with 50 bp homology to LwaCas13a (RV)
pKC005	ATGTACATGAAGATTACCAAGATTG	Amplification of LwaCas13a ORF with SV40 NLS (FW)
pKC006	TTATACCTTTCTCTTCTTCTTTGG	Amplification of LwaCas13a ORF with SV40 NLS (RV)
pKC007	TCCAGAGCTGATCCAAAGAAGAAGAGAAAGGTATAAAG GAGTTAAAGGCAAAGTTTTCTT	Amplification of GND2t with 36 bp homology to SRAD linker and SV40 NLS (FW)
pKC008	AAAGCAAAGGAAGGAGAGAACAGAGGAGTACTTGTACG TTCGATGGGCAAAACAATAGCGATCCGAAAGGCG	Amplification of GND2t with 50 bp homology to connector A (RV)
pKC009	TTGCCCATCGAACGTACAAGTACTCCTCTGTTCTCTCC TTCCTTTGCTTTCTACTTGGCTTCACATACGTTG	Amplification of Sc_FBA1p with 50 bp homology to connector A (FW)
pKC010	ATGAATTCCTTAATGATCGCCATATTGTCTTCTTCACC CTTACTTACCATTTTGAATATGTATTACTTGGTTATGG	Amplification of Sc_FBA1p with 50 bp homology to mCherry ORF (RV)
pKC011	ATGGTAAGTAAGGGTGAAGAAGA	Amplification of mCherry ORF (FW)
pKC012	TTATTTGTACAATTCGTCCATTCC	Amplification of mCherry ORF (RV)
pKC013	GAGCTGAAGGTAGGCACTCAACGGGTGGAATGGACGAA TTGTACAAATAAAGTGAATTTACTTTAAATCTTGCATT	Amplification of Sc_TDH3t with 50 bp homology to mCherry ORF (FW)
pKC014	TGCTAAGATTTGTGTTCGTTTGGGTGCAGTCGGTTGTG TACATCGATCCGATCCTGGCGGAAAAAATTCATTTG	Amplification of Sc_TDH3t with 50 bp homology to connector A (FW)
pKC015	CGGATCGATGTACACAACCGACTGCACCCAAACGAACA CAAATCTTAGCACGTATCCCTATCTGGATTAACAT	Amplification of Kl_ENO1p with 50 bp homology to connector B (FW)
pKC016	TCAACCAAAATTGGGACAACACCAGTGAATAATTCTTC ACCTTTAGACATTTTTTTTTGTTTTGTGGGTTGGTTT	Amplification of Kl_ENO1p with 50 bp homology to GFP ORF (RV)
pKC017	ATGTCTAAAGGTGAAGAATTATTCA	Amplification of GFP ORF (FW)
pKC018	TTATTTGTACAATTCATCCATACCA	Amplification of GFP ORF (RV)
pKC019	TTGTTACTGCTGCTGGTATTACCCATGGTATGGATGAA TTGTACAAATAAAGGAAGTATCTCGGAAATATTAATT	Amplification of Sc_TAL1t with 50 bp homology to GFP ORF (FW)

pKC020	CATCACGTACGACGACACTGCGACTCATCCATGTGAAA ATTATCGGCATCGACGTTGATTTAAGGTGGTTCC	Amplification of Sc_TAL1t with 50 bp homology to integration site (INT4) (FW)
pKC021	GGAACCCACGTTCTGCGTTAG	Amplification of construct integrated into INT4 (FW)
pKC022	CTTGGTACATAGACTTGACACC	Amplification of construct integrated into INT4 (RV)
pKC023	CCTTTAACACTGCGACTGCTTG	Sanger sequencing of INT4 (FW)
pKC024	CAGTTTCATTTTTCTTGTTCTATTAC	Sanger sequencing of TEF1p (FW)
pKC025	ATGGATATGTATATGGTGGTAATG	Sanger sequencing of GAL10p (FW)
pKC026	TCTACAATCAACAGATCGCTTCA	Sanger sequencing of FBA1p (FW)
pKC027	CGTATGTATCCAGAAGATGGCG	Sanger sequencing of mCherry ORF (FW)
pKC028	GCAGGACTTCTGATGATGAACT	Sanger sequencing of ENO1p (FW)
pKC029	TCTTTGGACCACCACCAACCA	Sanger sequencing of ENO1p (FW)
pKC030	AGACACAACATTGAAGATGGTTC	Sanger sequencing of eGFP ORF (FW)
pKC031	CTAGAGCCGTTCCCACAAATAA	Sanger sequencing of GND2t (FW)
pKC032	ATTATAGAAAACGAAACCGAGAAG	Sanger sequencing of LshCas13a ORF (FW)
pKC033	GATCTTAGTCAACGAACAAAAGG	Sanger sequencing of LshCas13a ORF (FW)
pKC034	CATCCTTCTCCAAAGTCTTACC	Sanger sequencing of LshCas13a ORF (FW)
pKC035	TCCAGAAGATGAAGGAAATTGAG	Sanger sequencing of LshCas13a ORF (FW)
pKC036	TTCGCTAAGAACATACAAAACAAG	Sanger sequencing of LshCas13a ORF (FW)
pKC037	AGTTCAAGCTAATTGGGAACAAC	Sanger sequencing of LshCas13a ORF (FW)
pKC038	GAAATGACTTTCTACGAATTGCTG	Sanger sequencing of LwaCas13a ORF (FW)
pKC039	GTTGAACTATTTCTTCGACTTCG	Sanger sequencing of LwaCas13a ORF (FW)
pKC040	AGAAGGAGATGGAACGTGTCG	Sanger sequencing of LwaCas13a ORF (FW)
pKC041	GAGAAAGAGAACATTGAAGACTAC	Sanger sequencing of LwaCas13a ORF (FW)

pKC042	GAATTTATTGATTAAGCTCTTCTCC	Sanger sequencing of LwaCas13a ORF (FW)
pKC043	CAAAGAGAGAGAGGTGAAAAGG	Sanger sequencing of BzoCas13b ORF (FW)
pKC044	ATTCGCTCAATTCCCAACTTTGA	Sanger sequencing of BzoCas13b ORF (FW)
pKC045	CATATCAAGATGGTAACTCCACC	Sanger sequencing of BzoCas13b ORF (FW)
pKC046	TAAGGGTAAGACTTTCAAGGGTA	Sanger sequencing of BzoCas13b ORF (FW)
pKC047	TGAAGATGTTGAATCCTACAAGG	Sanger sequencing of BzoCas13b ORF (FW)
pKC048	TAACACTGGTTTCTTCTTATCCTT	Sanger sequencing of PspCas13b ORF (FW)
pKC049	CCCACTACATTTTGGAAAACAAC	Sanger sequencing of PspCas13b ORF (FW)
pKC050	TTCTTATACAAGGTTTTCGCTCG	Sanger sequencing of PspCas13b ORF (FW)
pKC051	GCCGACTTCGATGGAGAAAGA	Sanger sequencing of PspCas13b ORF (FW)
pKC052	GGAAAGGGAATTATTTCGTAAGGCGGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC053	TAGTTTCGAACCACGGTTACTAATCGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC054	ATGAGTGTCAGCGAGTGTAACTCGAGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC055	CAATCCCGAGGCCTGACGCGACATAGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC056	CCACAGCGAGGAAGTAAACTGTTATGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC057	TTGTACTAATCGGCTTCAACGTGCCGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC058	CGCCTGCTACAACAGGAGTATCAAAGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC059	AAGCCCATCGTGGTCCTTAGACTTGGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC060	TCCGCAAGGCGCAGCAGTGCACAAGGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC061	CAAATGACAATTAACCACCGTGTATGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC062	ATGCCATTGGCTCTTAGACAGCCCGGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC063	CAGTTTACCGCATCTTGACCTAACTGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)

pKC064	CATTTGTATGTTCACCTATCTACTAGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC065	TGGCAGTAAGTACGCCTTCTGAATTGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC066	ACCTCTATCACCGCTTCATGCTAAGGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC067	TACGCTCATCTAATACATCCAACTCGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC068	TAATGAGAACAACCACACCATAGCGGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC069	AGCAATCACTTGCATAGCTGCGTATGGAACCCACGTTC TGCGTTAG	Amplification of template from INT4 for MinION sequencing (FW)
pKC070	AGAGCTAGAAGGGCACCCTGCAGTTCTTGGTACATAGA CTTGACACC	Amplification of template from INT4 for MinION sequencing (RV)
pKC071	CTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTAC CCAGCTTTTGTTCCGCGGTCTTTGAAAAGATAATG	Amplification of Cas9 sgRNA expression cassette (RV)
pKC072	AAAATACAACAAATAAAAAACACTCAATGACCTGACCA TTTGATGGAGTTCCGCGGAGACATAAAAAAC	Amplification of Cas9 sgRNA expression cassette (FW)
pKC073	AACTCCATCAAATGGTCAGG	Amplification of Cas9 sgRNA recipient plasmid pRN1120 (FW)
pKC074	AACAAAAGCTGGGTACCGGG	Amplification of Cas9 sgRNA recipient plasmid pRN1120 (RV)
pKC075	CATGTTTGACAGCTTATCATC	Amplification of Cas13 gRNA expression cassette (FW)
pKC076	CACACAGGAAACAGCTATGAC	Amplification of Cas13 gRNA expression cassette (RV)
pKC077	TACCGTCGACCTCGAGGGG	Amplification of Cas13 gRNA recipient plasmid pRN1120 (FW)
pKC078	ATATCGAATTCCTGCAGCCCG	Amplification of Cas13 gRNA recipient plasmid pRN1120 (RV)
pKC079	GCTCACATGTTCTTTCCTGCGT	Verification of gRNA expression cassette integration in pRN1120 (FW)
pKC080	GCTGAGCTGGATCAATAGAGTA	Verification of gRNA expression cassette integration in pRN1120 (RV)
pKC081	GTCTCCGGTGAAGGTGAAGGT	Quantification of eGFP transcript in qPCR reaction (FW)
pKC082	AGGTTGGCCATGGAACTGGC	Quantification of eGFP transcript in qPCR reaction (RV)
pKC083	TCCAGAAGATGGCGCGCTTA	Quantification of mCherry transcript in qPCR reaction (FW)
pKC084	TGGTCTTAACTTCCGCGTCGT	Quantification of mCherry transcript in qPCR reaction (RV)
pKC085	GTTCATTGTCATCGCTATCAAGCA	Quantification of LwaCas13a transcript in qPCR reaction (FW)

pKC086	CCTTTGCTGTTGTCGGGTGG	Quantification of LwaCas13a transcript in qPCR reaction (RV)
pKC087	CACCGCTTTGGCTCCATCTT	Quantification of ACT1 transcript in qPCR reaction (FW)
pKC088	AGAACCACCAATCCAGACGG	Quantification of ACT1 transcript in qPCR reaction (RV)
pKC089	GGCGTGCAGCAGATTTCACAA	Quantification of TAF10 transcript in qPCR reaction (FW)
pKC090	AGCCCGTATTCAGCAACAGC	Quantification of TAF10 transcript in qPCR reaction (RV)

SI Table 6-2. Features of Cas13a and Cas13b gRNA design

Donor organism	PFS	Spacer (nt)	DR sequence	DR (nt)	Reference
Leptotrichia shahii	Single site: 3' H	28	CCACCCCAATATCGAAGGGGACTAAAAC	28	[68,70,212,251]
Leptotrichia wadei	No PFS	28	GATTTAGACTACCCCAAAAACGAAGGGGACTAAAAC	36	[70,251]
Bergeyella zoohelcum	Single site: 5' D	30	GTTGGAACTGCTCTCATTTTGGAGGGTAATCACAAC	36	[89,92]
Prevotella sp. P5-125	Double site: 5' DD and 3' NAN or NNA	30	GTTGTGGAAGGTCCAGTTTTGAGGGGCTATTACAAC	36	[89]

Where: PFS - protospacer flanking site, DR – direct repeat, H – adenine, cytosine or thymine, and D – guanine, adenine or thymine.

SI Table 6-3. List of spacers used for singleplex genome editing and evaluation of CRISPR-Cas13 system

Name	Target	Sequence	PAM/PFS	Vendor
Spacers used for CRISPR-Cas9 genome editing				
g83	INT4	TCAGGACATGATCTCAGTAG	CGG	IDT
gNLS2	NLS	TGATCCAAAGAAGAAGAGAA	AGG	IDT
Spacers used for CRISPR-LshCas13a and LwaCas13a				
gGFP1	eGFP	AGCATCACCTTCACCTTCACCGGAGACA	A	Twist
gGFP2	eGFP	CACCATAAGTTAAAGTAGTGACTAAGGT	T	Twist
gGFP3	eGFP	TATCACCTTCAAACTTGACTTCAGCTCT	C	Twist
gGFP4	eGFP	GGATAAGTAATGGTTGTCTGGTAACAAG	A	Twist
gADE1	ADE2	ATCAATCTGATTGTTTCTGGAGAAGGGT	A	Twist
gADE2	ADE2	TAATGGTGGAGAAAGATGTGAAATTCTT	C	Twist
gADE3	ADE2	TTAGCATAATGGCGTTCGTTGTAATGGT	A	Twist
gADE4	ADE2	ACAAGAACCTTTTACGGGCACACCGATG	C	Twist
NT	-	ACTAACCCATCAACCTGTACGGGAACAT	–	Twist
Spacers used for CRISPR-BzoCas13b				
gGFP1	eGFP	ATTTACCGTAAGTAGCATCACCTTCACCTT	5′G	Twist
gGFP2	eGFP	GAACACCATAAGTTAAAGTAGTGACTAAGG	5′A	Twist
gGFP3	eGFP	TCACCTTCAAACTTGACTTCAGCTCTGGTC	5′A	Twist
gGFP4	eGFP	GTGGATAAGTAATGGTTGTCTGGTAACAAG	5′T	Twist
NT	-	ACTAACCCATCAACCTGTACGGGAACATTC	–	Twist
Spacers used for CRISPR-PspCas13b				
gGFP1	eGFP	GCATCACCTTCACCTTCACCGGAGACAGAA	5′TT, 3′TAC	Twist
gGFP2	eGFP	ATTTACCGTAAGTAGCATCACCTTCACCTT	5′TG, 3′TGA	Twist
gGFP3	eGFP	GAACACCATAAGTTAAAGTAGTGACTAAGG	5′AA, 3′AAT	Twist
gGFP4	eGFP	GTGGATAAGTAATGGTTGTCTGGTAACAAG	5′GT, 3′TCA	Twist
NT	-	ACTAACCCATCAACCTGTACGGGAACATTC	–	Twist

SI Table 6-4. Sequences of donor DNA used to remove NLS and construct strains CC005,6,11,12

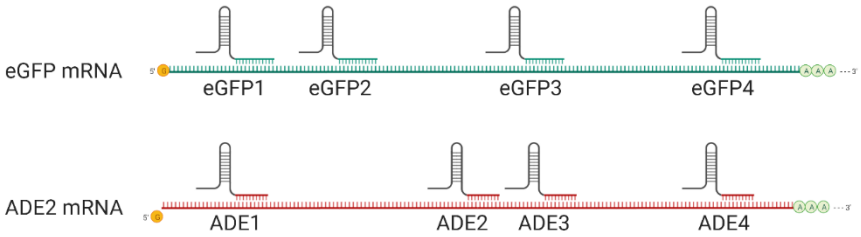
Strain	Sequence	Vendor
CC005	TGATCATTGAATTGCTGACAAAGATTGAGAACACCAACGACACTTTG**TAA***AGGAGTTAAAGGCAA AGTTTTCTTTTCTAGAGCCGTTCCCACAAATAATT*	IDT
CC006	AGAATGAATTTTTGGAAAACGTCAAGGCCATGTTGGAATATTCGGAG**TAA***AGGAGTTAAAGGCAA AGTTTTCTTTTCTAGAGCCGTTCCCACAAATAATT*	IDT
CC011	AATATTTCGCTGAAGTCTTTAAGAAGGAAAAGGAAGCTTTGATCAAG**TAA***AGGAGTTAAAGGCAA AGTTTTCTT*	IDT
CC012	TTGCTATGTCCATTAAGAAGGCTTTCGGTGAATATGCCATCATGAAG**TAA***AGGAGTTAAAGGCAA AGTTTTCTT*	IDT

Legend:

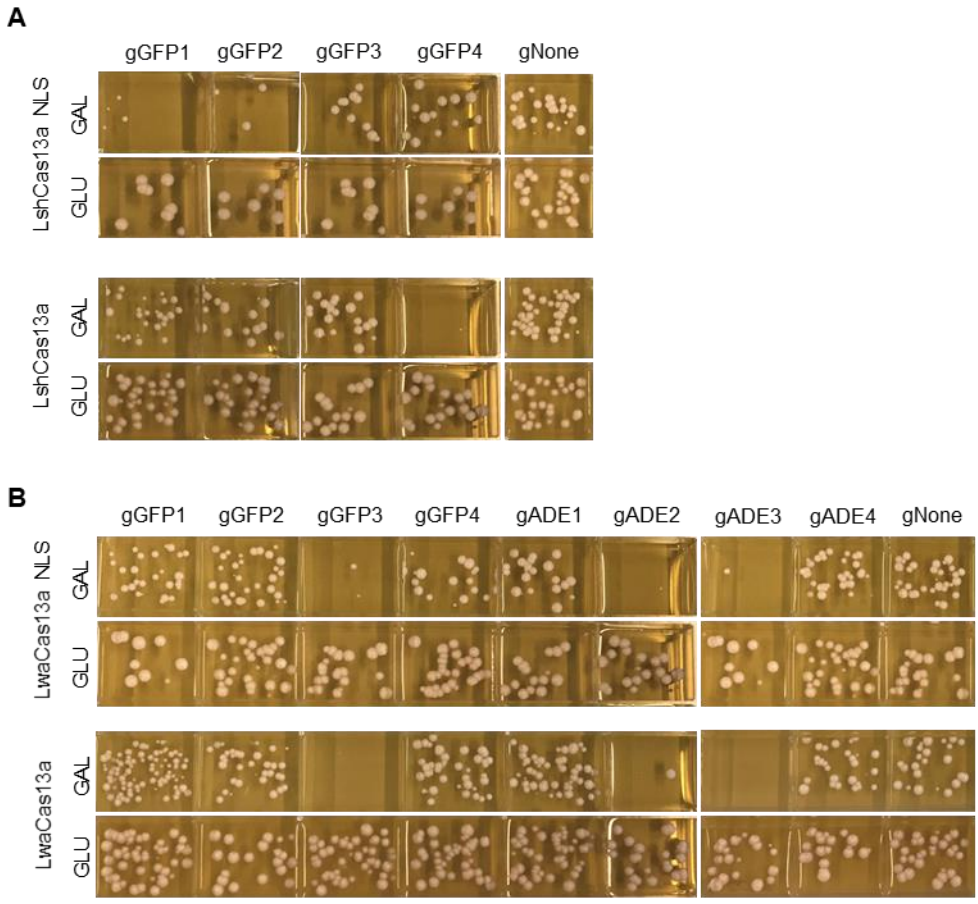
Cas13 sequence homology

Stop codon TAA

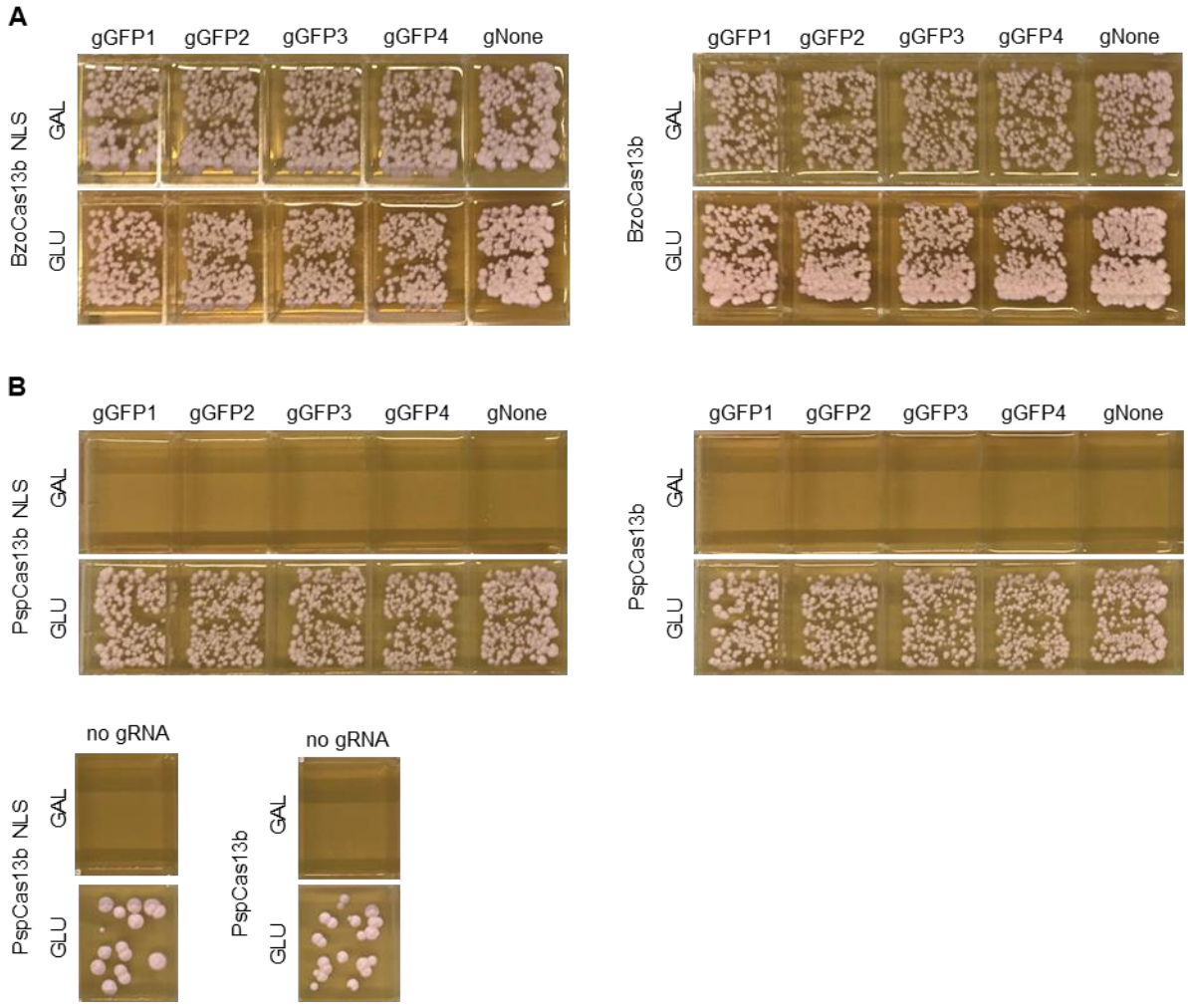
GND2 terminator homology



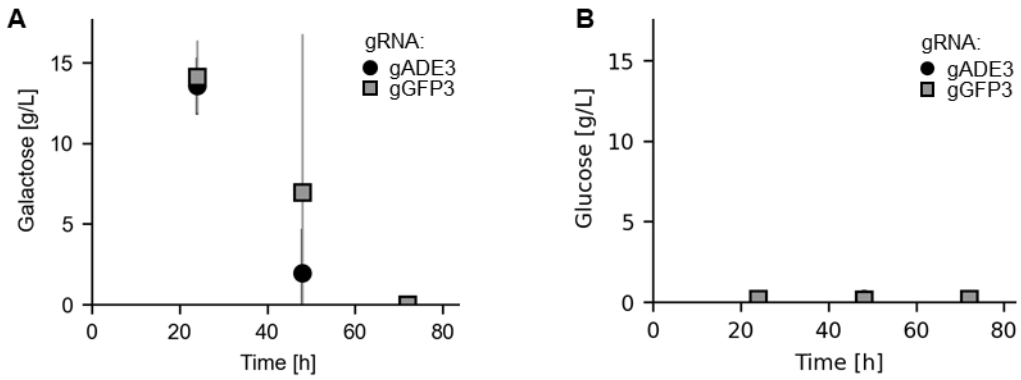
SI Figure 6-1. Visual representation of gRNAs used with Cas13a and Cas13b. Four gRNAs were used to target *eGFP* and *ADE2* transcripts. G indicates 5'-cap and AAA – poly(A) tail of mRNA. gRNAs are exemplified for Cas13a with 5' direct repeat (black) and 3' spacer (green or red), for Cas13b orientation of spacer and direct repeat is reversed.



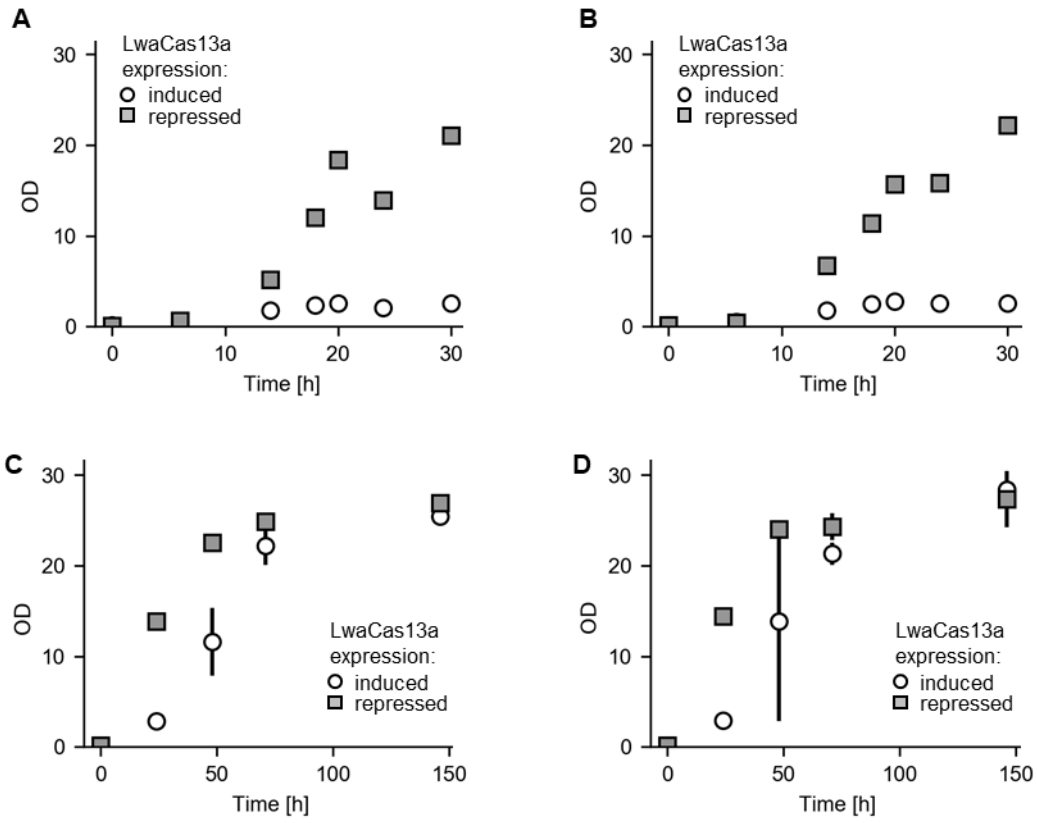
SI Figure 6-2. Expression of A. LshCas13a or B. LwaCas13a and gRNAs in *S. cerevisiae*. Strain CC003 (LwaCas13a NLS), CC004 (LwaCas13a), CC005 (LshCas13a NLS) and CC006 (LshCas13a) were used. In all strains Cas13a gene was expressed from the galactose inducible promoter *GAL10*. Transformants were plated on YEPh agar with galactose (2% GAL, inducing conditions) or glucose (2% GLU, repressing conditions). Selected gRNAs were targeting *eGFP* mRNA (gGFP1-4) or *ADE2* mRNA (gADE1-4). As a control gNone with a non-targeting spacer was used. The experiment was conducted twice on independent days.



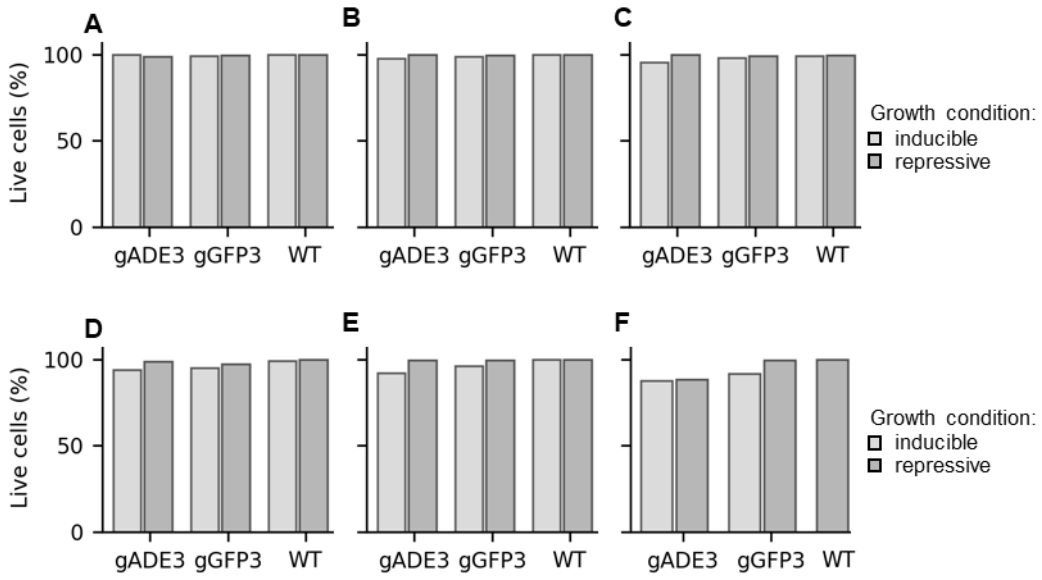
SI Figure 6-3. Expression of A. BzoCas13b or B. PspCas13b and gRNAs in *S. cerevisiae*. Strain CC009 (BzoCas13b NLS), CC010 (BzoCas13b), CC011 (PspCas13b NLS) and CC012 (PspCas13b) were used. In all strains Cas13b gene was expressed from galactose inducible promoter *GAL10*. Transformants were plated on YEPh agar with galactose (2% GAL, inducing conditions) or glucose (2% GLU, repressing conditions). Selected gRNAs were targeting eGFP mRNA (gGFP1-4). As a control gNone with a non-targeting spacer was used. Additionally, transformants of PspCas13b (±NLS) are shown with the empty gRNA recipient plasmid pRN1120 (no gRNA). The experiment was conducted twice on independent days.



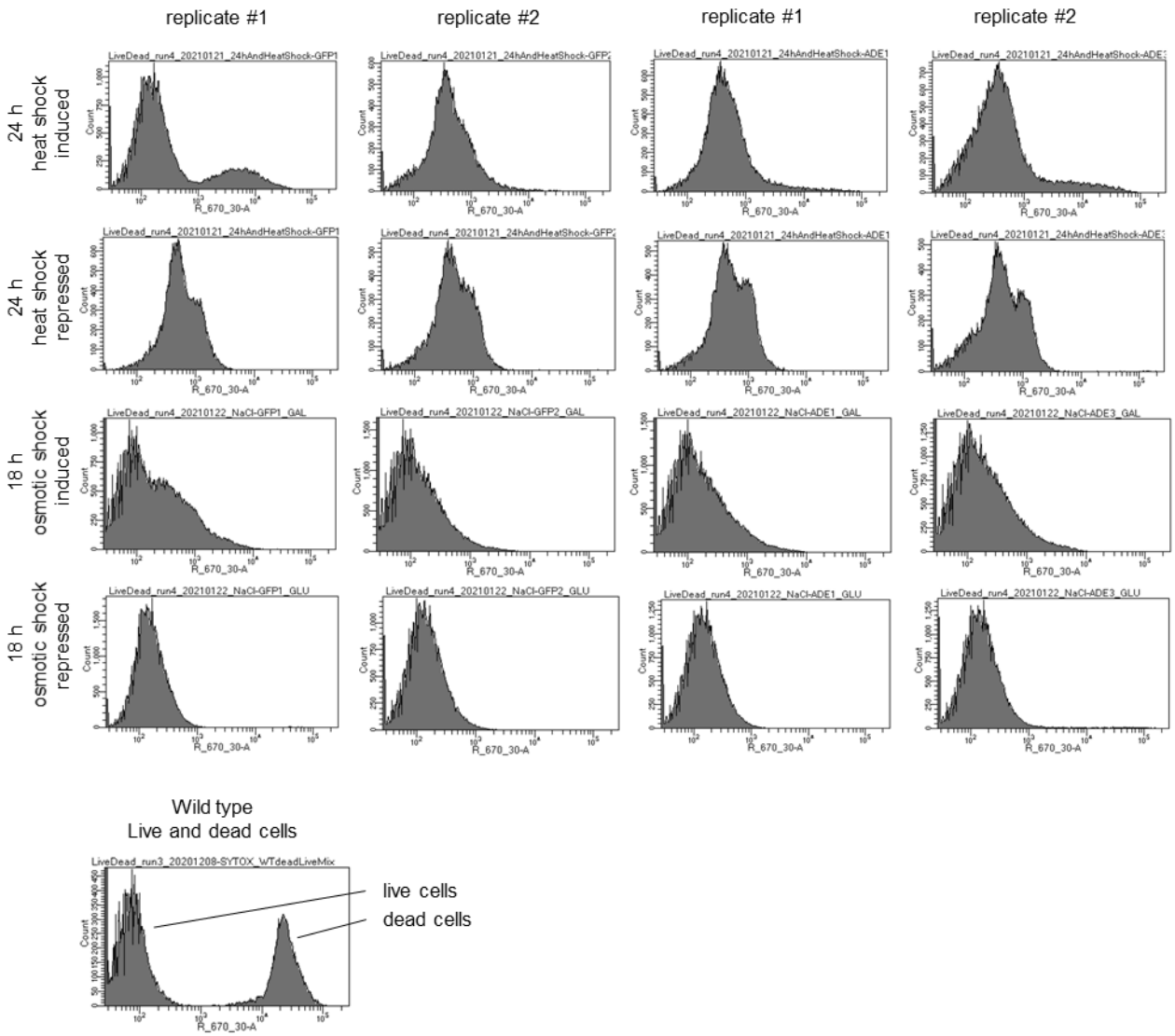
SI Figure 6-4. Sugar concentration in cultures of strain LwaCas13a NLS under A. inducing conditions (galactose) and **B**. repressive conditions (glucose). Plot represents mean ± 1 standard deviation (n=2).



SI Figure 6-5. Growth of strain LwaCas13a NLS under inducive and repressive condition. A. In combination with gRNA gADE3 and **B**. gGFP3 during 30 h batch fermentation. **C**. In combination with gRNA gADE3 and **D**. gGFP3 during 6 days batch fermentation. The 30 h fermentation was conducted without replicates while the 5 days fermentation in biological duplicates. In the 30 h fermentation cultures were prepared in two sets: 6h, 24, 30 h and 14 h, 18, 20 h and results were represented cumulatively. Panels C and D represent average ±1 standard deviation (n=2).



SI Figure 6-6. Live cells in culture grown for A. 6 h, **B**. 14 h, **C**. 18 h, **D**. 20 h, **E**. 24h and **F**. 30 h from the induction of LwaCas13a NLS. Experiment was conducted without replicates.



SI Figure 6-7. Dead staining of cells with active LwaCas13a NLS system exposed to heat and osmotic shock. Far red fluorescence in range approximately 0-10^3 a.u. indicated live cells while approximately 10^4-10^5 - dead cells. Shift in the peak of live cells for cells under inducive conditions and osmotic stress indicates injured cells.

7 Supplementary information to Chapter 7

SI Table 7-1. List of primers used in this study

Name	Sequence	Purpose
pKC001	GGTCTCGGTGCCGTCATCATTCCCCTCGAGGT	Amplification of ERG3p 900 bp (FW)
pKC002	GGTCTCGGTGCCGGGCATTATTTCGGTCGTTTA	Amplification of ERG3p 373 bp (FW)
pKC003	GGTCTCCCATTATCTCAAATCTAGACGAATATTTTTC	Amplification of ERG3p (RV)
pKC004	GGTCTCGGTGCAACAACTATGAATTCCTTGGATAATT	Amplification of ERG5p 750 bp (FW)
pKC005	GGTCTCGGTGCGTAAAGAGACAGTGTATTTGTTCC	Amplification of ERG5p 449 bp (FW)
pKC006	GGTCTCCCATTTTTGTTAAAAGGTATTTATTGTCTATTG	Amplification of ERG5p (RV)
pKC007	GGTCTCGGTGCACGATTATCACGACACAACCTTG	Amplification of PDR5p (FW)
pKC008	GGTCTCCCATTTTTTGTCTAAAGTCTTTCGAACGA	Amplification of PDR5p (RV)
pKC009	ATAGTTACAGCTAAACATTTGCCC	Amplification of dCas12a expression cassette (FW)
pKC010	CATCACGTACGACGACACTGCGACTCATCCATGTGAAAATTATCGGCATCAACAATAGCGATCCGAAAGGCG	Amplification of dCas12a expression cassette (RV)
pKC011	GGAACCCACGTTCTGCGTTAG	Amplification of sequencing template from INT4 (FW)
pKC012	CTTGGTACATAGACTTGACACC	Amplification of sequencing template from INT4 (RV)
pKC013	CCTTTAACACTGCGACTGCTTG	Sanger sequencing of INT4
pKC014	CAGTTTCATTTTTCTTGTTCTATTAC	Sanger sequencing of TEF1p
pKC015	CACGAAGTCCAAGAAATCAAGG	Sanger sequencing of dCas12a
pKC016	ACGGTTCCTCTGAAAAGTTATTC	Sanger sequencing of dCas12a
pKC017	ACAAGGATATTGCTGGTTTCTAC	Sanger sequencing of dCas12a
pKC018	CATCAAAGAATTGAAGGCTGGTT	Sanger sequencing of dCas12a
pKC019	GCTTTGATGTCCTTGATGTTGC	Sanger sequencing of dCas12a

pKC020	CGTATGTATCCAGAAGATGGCG	Sanger sequencing of mCherry
pKC021	GGTTACACTCACATGGCTTTGA	Sanger sequencing of mRuby2
pKC022	GCAGGACTTCTGATGATGAACT	Sanger sequencing of ENO1p
pKC023	TCTTTGGACCACCACCAACCA	Sanger sequencing of ENO1p
pKC024	CTAGAGCCGTTCCCACAAATAA	Sanger sequencing of GND2t

SI Table 7-2. List of gRNAs used in this study

ERG3 G34

CATGTTTGACAGCTTATCATCGATAATCCGGAGCTAGCATGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGC
TGCAGATCTCAAATCTAGACGAATATTTTTCTTATTATCTTTTTTTACAAATGCAACAAACTGACCTCCTCTG
CCACCACTTCAAAGAACAAGACAAATGTAAAGAATGAATGAACAAGTGAATGAAAAAACGAATTGATAAGAAA
GATAGAAATCCGCCTCTAATATATACCGCGATGTTTCTTTCGACCGTTACATAATAGCAGCAACAACAATACC
CGATCGCTGCTGAACCTCTTGTTTCAATGCACCTGCTGTTCTTCTTCACCTTCTATTGTGCCACTTATACGAG
CCTGAATTTTTCGTTTATCGCAGCCATCGGCCTCGTATACACCCCGGTTGCCCTGGGGCGCAACTAAACGACC
GAAATAATGCCCGGGAGTTTTCGGGTGGTCTGCCCCGGGCACCGATCTCGGTCGTTTAGGCCGCAGTGGGTCG
TGCCACACGGCAGCGATGCAAGGGACAACGCGTCGTAACTTGCGATGCCCGACAATCGCAACAAAAGGCGTCG
TTTCGCGGTTTTTTCGTATCATTTTTTTTTTCAGTTTTTCTTCTTTCGGCTATGAAAAAAATGGCAGAGCTGC
CCTAATCCATTTTTATGGATGCCCTAATATATTGTCACTCTGTACGTAGGGCTTCCAGAATGCCCTTTCAGGA
AGGATTCTAGCATAATAGACGCATCATATATTATAGCGTAGTGGTGCCAGATGCTGGGCTGTAAAAGGCTGTA
GCAGTAACTGTAGTGTACTAAAGGCTCAGTTTTTTTTTTTTCTTCTTTTACGTAATGGACGAGGTTGACGATA
TTCTAATTAACAACCAGGTGGTGGATGACGAGGAGGATGACGAAGAGATGCTGAGTGGGCTGGAAAACGACTC
AAAGCAGGACCTCGAGGGGAATGATGACG*AAATTACTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTC**T*
AATTTCTACTAAGTGTAGATTATAACCTCAAATCATGGGAGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCC
GGCTGGGCAACATGCTTCGGCATGGCGAATGGGAC*AGGAGTTAAAGGCAAAGTTTTCTTTTCTAGAGCCGTTC*
CCACAAATAATTATACGTATATGCTTCTTTTCGTTTACTATATATCTATATTTACAAGCCTTTATTCACTGAT
GCAATTTGTTTCCAAATACTTTTTTGGAGATCTCATAACTAGATATCATGATGGCGCAACTTGGCGCTATCTT
AATTACTCTGGCTGCCAGGCCCGTGTAGAGGGCCGCAAGACCTTCTGTACGCCATATAGTCTCTAAGAACTTG
*AACAAGTTTCTAGACCTATTGCCGCCTTTCGGATCGCTATTGTT*GGGGGGCCCGGTACCCAGCTTTTGTTCCC
TTTAGTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG

ERG3 G37

CATGTTTGACAGCTTATCATCGATAATCCGGAGCTAGCATGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGC
TGCAGATCTCAAATCTAGACGAATATTTTTCTTATTATCTTTTTTTACAAATGCAACAAACTGACCTCCTCTG
CCACCACTTCAAAGAACAAGACAAATGTAAAGAATGAATGAACAAGTGAATGAAAAAACGAATTGATAAGAAA
GATAGAAATCCGCCTCTAATATATACCGCGATGTTTCTTTCGACCGTTACATAATAGCAGCAACAACAATACC
CGATCGCTGCTGAACCTCTTGTTTCAATGCACCTGCTGTTCTTCTTCACCTTCTATTGTGCCACTTATACGAG
CCTGAATTTTTCGTTTATCGCAGCCATCGGCCTCGTATACACCCCGGTTGCCCTGGGGCGCAACTAAACGACC
GAAATAATGCCCGGGAGTTTTCGGGTGGTCTGCCCCGGGCACCGATCTCGGTCGTTTAGGCCGCAGTGGGTCG
TGCCACACGGCAGCGATGCAAGGGACAACGCGTCGTAACTTGCGATGCCCGACAATCGCAACAAAAGGCGTCG
TTTCGCGGTTTTTTCGTATCATTTTTTTTTTCAGTTTTTCTTCTTTCGGCTATGAAAAAAATGGCAGAGCTGC
CCTAATCCATTTTTATGGATGCCCTAATATATTGTCACTCTGTACGTAGGGCTTCCAGAATGCCCTTTCAGGA
AGGATTCTAGCATAATAGACGCATCATATATTATAGCGTAGTGGTGCCAGATGCTGGGCTGTAAAAGGCTGTA
GCAGTAACTGTAGTGTACTAAAGGCTCAGTTTTTTTTTTTTCTTCTTTTACGTAATGGACGAGGTTGACGATA
TTCTAATTAACAACCAGGTGGTGGATGACGAGGAGGATGACGAAGAGATGCTGAGTGGGCTGGAAAACGACTC
AAAGCAGGACCTCGAGGGGAATGATGACG*AAATTACTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTC**T*
AATTTCTACTAAGTGTAGATATTCGAAACAAACCAAACCAGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCC
GGCTGGGCAACATGCTTCGGCATGGCGAATGGGAC*AGGAGTTAAAGGCAAAGTTTTCTTTTCTAGAGCCGTTC*
CCACAAATAATTATACGTATATGCTTCTTTTCGTTTACTATATATCTATATTTACAAGCCTTTATTCACTGAT
GCAATTTGTTTCCAAATACTTTTTTGGAGATCTCATAACTAGATATCATGATGGCGCAACTTGGCGCTATCTT
AATTACTCTGGCTGCCAGGCCCGTGTAGAGGGCCGCAAGACCTTCTGTACGCCATATAGTCTCTAAGAACTTG
*AACAAGTTTCTAGACCTATTGCCGCCTTTCGGATCGCTATTGTT*GGGGGGCCCGGTACCCAGCTTTTGTTCCC
TTTAGTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG

ERG3 G42

CATGTTTGACAGCTTATCATCGATAATCCGGAGCTAGCATGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGC
TGCAGATCTCAAATCTAGACGAATATTTTTCTTATTATCTTTTTTTACAAATGCAACAAACTGACCTCCTCTG
CCACCACTTCAAAGAACAAGACAAATGTAAAGAATGAATGAACAAGTGAATGAAAAAACGAATTGATAAGAAA
GATAGAAATCCGCCTCTAATATATACCGCGATGTTTCTTTCGACCGTTACATAATAGCAGCAACAACAATACC
CGATCGCTGCTGAACCTCTTGTTTCAATGCACCTGCTGTTCTTCTTCACCTTCTATTGTGCCACTTATACGAG
CCTGAATTTTTCGTTTATCGCAGCCATCGGCCTCGTATACACCCCGGTTGCCCTGGGGCGCAACTAAACGACC
GAAATAATGCCCGGGAGTTTTCGGGTGGTCTGCCCCGGGCACCGATCTCGGTCGTTTAGGCCGCAGTGGGTCG
TGCCACACGGCAGCGATGCAAGGGACAACGCGTCGTAACTTGCGATGCCCGACAATCGCAACAAAAGGCGTCG
TTTCGCGGTTTTTTCGTATCATTTTTTTTTTCAGTTTTTCTTCTTTCGGCTATGAAAAAAATGGCAGAGCTGC
CCTAATCCATTTTTATGGATGCCCTAATATATTGTCACTCTGTACGTAGGGCTTCCAGAATGCCCTTTCAGGA
AGGATTCTAGCATAATAGACGCATCATATATTATAGCGTAGTGGTGCCAGATGCTGGGCTGTAAAAGGCTGTA
GCAGTAACTGTAGTGTACTAAAGGCTCAGTTTTTTTTTTTTCTTCTTTTACGTAATGGACGAGGTTGACGATA
TTCTAATTAACAACCAGGTGGTGGATGACGAGGAGGATGACGAAGAGATGCTGAGTGGGCTGGAAAACGACTC

AAAGCAGGACCTCGAGGGGAATGATGACG*AAATTACTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTC**T**
AATTTCTACTAAGTGTAGATACTAACCCATCAACCTGTACGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCC
GGCTGGGCAACATGCTTCGGCATGGCGAATGGGAC*AGGAGTTAAAGGCAAAGTTTTCTTTTCTAGAGCCGTTC*
CCACAAATAATTATACGTATATGCTTCTTTTCGTTTACTATATATCTATATTTACAAGCCTTTATTCACTGAT
GCAATTTGTTTCCAAATACTTTTTTGGAGATCTCATAACTAGATATCATGATGGCGCAACTTGGCGCTATCTT
AATTACTCTGGCTGCCAGGCCCGTGTAGAGGGCCGCAAGACCTTCTGTACGCCATATAGTCTCTAAGAACTTG
*AACAAGTTTCTAGACCTATTGCCGCCTTTCGGATCGCTATTGTT*GGGGGGCCCGGTACCCAGCTTTTGTTCCC
TTTAGTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG

PDR5 G34

CATGTTTGACAGCTTATCATCGATAATCCGGAGCTAGCATGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGC
TGCAGACGATTATCACGACACAACCTTGCCGCCGAGAAAACGTCCGTGGAGAACCATTCGGTCGATTGCTTCC
CACGGAACGAGTGGACTGAAACTTAAGACTGCCCCTCTCTTTCCGCGGAATCGCTCATGCCGCGGTGCCACAA
CATTTTCAGATTTACTAAGACTCCGGTGAGTGTGGGCTCACCCGCGGGTCGTGATCACGATTCAGCACCCTTT
GGACTCGTGATTCCGTGGAAAGGTCAGATCTGTATTCCTACTTATGGTAATGTGCTAAAAAAAGAGAAATGTC
TCCGCGGAACTCTTCTACGCCGTGGTACGATATCTGTTGAACGTAATCTGAGCAATACAAACAAGGCCTCTCC
TATACATATATAATTGTGATGTGCATAACCTTATGGCTGTTCGCTTTTATTATCATACCTTAGAATGAAATCC
AAAAGAAAAAAGTCACGCAAAGTTGCAAACATATAACAACTGTGTTAGTTATCACTCGACTTTGTTATTCTAA
TTATAAATAAATTGGCAACTAGGAACTTTCGAAAAAGAAATTAAAGACCCTTTTAAGTTTTCGTATCCGCTCG
TTCGAAAGACTTTAGACAAAA*AAATTACTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTC****TAATTTCTA***
***CTAAGTGTAGAT*TATAACCTCAAATCATGGGA**GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGC
AACATGCTTCGGCATGGCGAATGGGAC*AGGAGTTAAAGGCAAAGTTTTCTTTTCTAGAGCCGTTCCCACAAAT*
AATTATACGTATATGCTTCTTTTCGTTTACTATATATCTATATTTACAAGCCTTTATTCACTGATGCAATTTG
TTTCCAAATACTTTTTTGGAGATCTCATAACTAGATATCATGATGGCGCAACTTGGCGCTATCTTAATTACTC
TGGCTGCCAGGCCCGTGTAGAGGGCCGCAAGACCTTCTGTACGCCATATAGTCTCTAAGAACTTGAACAAGTT
*TCTAGACCTATTGCCGCCTTTCGGATCGCTATTGTT*GGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGA
GGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG

PDR5 G37

CATGTTTGACAGCTTATCATCGATAATCCGGAGCTAGCATGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGC
TGCAGACGATTATCACGACACAACCTTGCCGCCGAGAAAACGTCCGTGGAGAACCATTCGGTCGATTGCTTCC
CACGGAACGAGTGGACTGAAACTTAAGACTGCCCCTCTCTTTCCGCGGAATCGCTCATGCCGCGGTGCCACAA
CATTTTCAGATTTACTAAGACTCCGGTGAGTGTGGGCTCACCCGCGGGTCGTGATCACGATTCAGCACCCTTT
GGACTCGTGATTCCGTGGAAAGGTCAGATCTGTATTCCTACTTATGGTAATGTGCTAAAAAAAGAGAAATGTC
TCCGCGGAACTCTTCTACGCCGTGGTACGATATCTGTTGAACGTAATCTGAGCAATACAAACAAGGCCTCTCC
TATACATATATAATTGTGATGTGCATAACCTTATGGCTGTTCGCTTTTATTATCATACCTTAGAATGAAATCC
AAAAGAAAAAAGTCACGCAAAGTTGCAAACATATAACAACTGTGTTAGTTATCACTCGACTTTGTTATTCTAA
TTATAAATAAATTGGCAACTAGGAACTTTCGAAAAAGAAATTAAAGACCCTTTTAAGTTTTCGTATCCGCTCG
TTCGAAAGACTTTAGACAAAA*AAATTACTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTC****TAATTTCTA***
***CTAAGTGTAGAT*ATTCGAAACAAACCAAACCA**GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGC
AACATGCTTCGGCATGGCGAATGGGAC*AGGAGTTAAAGGCAAAGTTTTCTTTTCTAGAGCCGTTCCCACAAAT*
AATTATACGTATATGCTTCTTTTCGTTTACTATATATCTATATTTACAAGCCTTTATTCACTGATGCAATTTG
TTTCCAAATACTTTTTTGGAGATCTCATAACTAGATATCATGATGGCGCAACTTGGCGCTATCTTAATTACTC
TGGCTGCCAGGCCCGTGTAGAGGGCCGCAAGACCTTCTGTACGCCATATAGTCTCTAAGAACTTGAACAAGTT
*TCTAGACCTATTGCCGCCTTTCGGATCGCTATTGTT*GGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGA
GGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG

PDR5 G42

CATGTTTGACAGCTTATCATCGATAATCCGGAGCTAGCATGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGC
TGCAGACGATTATCACGACACAACCTTGCCGCCGAGAAAACGTCCGTGGAGAACCATTCGGTCGATTGCTTCC
CACGGAACGAGTGGACTGAAACTTAAGACTGCCCCTCTCTTTCCGCGGAATCGCTCATGCCGCGGTGCCACAA
CATTTTCAGATTTACTAAGACTCCGGTGAGTGTGGGCTCACCCGCGGGTCGTGATCACGATTCAGCACCCTTT
GGACTCGTGATTCCGTGGAAAGGTCAGATCTGTATTCCTACTTATGGTAATGTGCTAAAAAAAGAGAAATGTC
TCCGCGGAACTCTTCTACGCCGTGGTACGATATCTGTTGAACGTAATCTGAGCAATACAAACAAGGCCTCTCC
TATACATATATAATTGTGATGTGCATAACCTTATGGCTGTTCGCTTTTATTATCATACCTTAGAATGAAATCC
AAAAGAAAAAAGTCACGCAAAGTTGCAAACATATAACAACTGTGTTAGTTATCACTCGACTTTGTTATTCTAA
TTATAAATAAATTGGCAACTAGGAACTTTCGAAAAAGAAATTAAAGACCCTTTTAAGTTTTCGTATCCGCTCG
TTCGAAAGACTTTAGACAAAA*AAATTACTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTC****TAATTTCTA***
***CTAAGTGTAGAT*ACTAACCCATCAACCTGTAC**GGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGC
AACATGCTTCGGCATGGCGAATGGGAC*AGGAGTTAAAGGCAAAGTTTTCTTTTCTAGAGCCGTTCCCACAAAT*
AATTATACGTATATGCTTCTTTTCGTTTACTATATATCTATATTTACAAGCCTTTATTCACTGATGCAATTTG
TTTCCAAATACTTTTTTGGAGATCTCATAACTAGATATCATGATGGCGCAACTTGGCGCTATCTTAATTACTC
TGGCTGCCAGGCCCGTGTAGAGGGCCGCAAGACCTTCTGTACGCCATATAGTCTCTAAGAACTTGAACAAGTT

*TCTAGACCTATTGCCGCCTTTCGGATCGCTATTGTT*GGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGA
GGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG

Legend:

ERG3 promoter

PDR5 promoter

LbCas12a direct repeat

Spacer

Hammer head ribozyme

HDV ribozyme

GND2 terminator

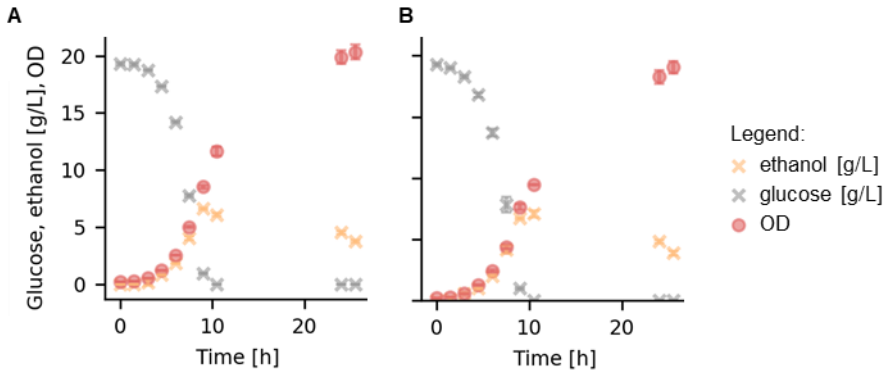
Homology to pRN1120 backbone

SI Table 7-3. Differential expression analysis of the carotenogenic and wild type strains. Analysis was performed for cDNA library at timepoints 6 h and 24 h and direct RNA for timepoint 6 h. logFC – logarithm of fold change; logCPM - logarithm of counts per million reads; FDR - False Discovery rate.

Gene	cDNA 6h			
	logFC	logCPM	PValue	FDR
ERG3	1.79	10.01	1.2E-09	7.5E-07
ERG5	1.61	5.97	3.6E-07	1.8E-04
ERG25	1.56	9.79	5.3E-05	1.5E-02
PDR5	0.91	2.19	3.4E-01	1.0E+00
SNQ2	0.10	2.63	1.0E+00	1.0E+00

Gene	cDNA 24h			
	logFC	logCPM	PValue	FDR
ERG3	2.16	9.48	1.9E-48	1.1E-45
ERG5	1.17	5.04	2.0E-05	3.8E-04
ERG25	1.71	9.08	2.2E-31	7.1E-29
PDR5	1.20	4.60	2.7E-04	3.5E-03
SNQ2	1.20	3.68	1.4E-02	8.2E-02

Gene	dRNA 6h			
	logFC	logCPM	PValue	FDR
ERG3	1.71	8.57	3.8E-21	2.3E-18
ERG5	1.12	7.78	9.0E-09	2.5E-06
ERG25	1.19	8.49	9.8E-11	5.0E-08
PDR5	1.43	4.74	2.4E-05	2.3E-03
SNQ2	1.10	4.86	2.1E-04	1.2E-02

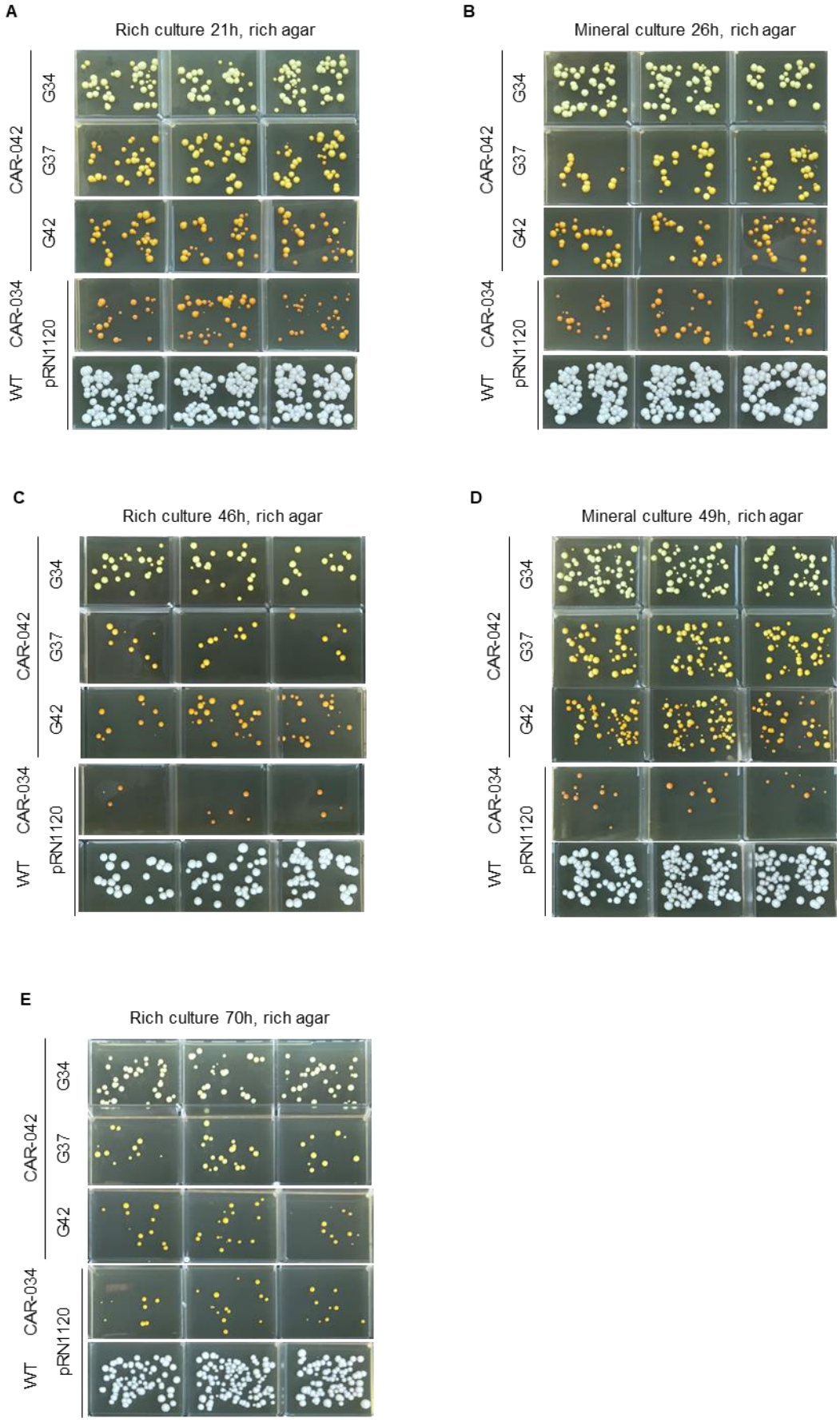


SI Figure 7-1. Growth curve and metabolic profile of cultures subjected to the RNA-seq. A. Carotenoids producing strain CAR-034 and **B**. wild type strain. Plots represent mean ± standard deviation (n=3).

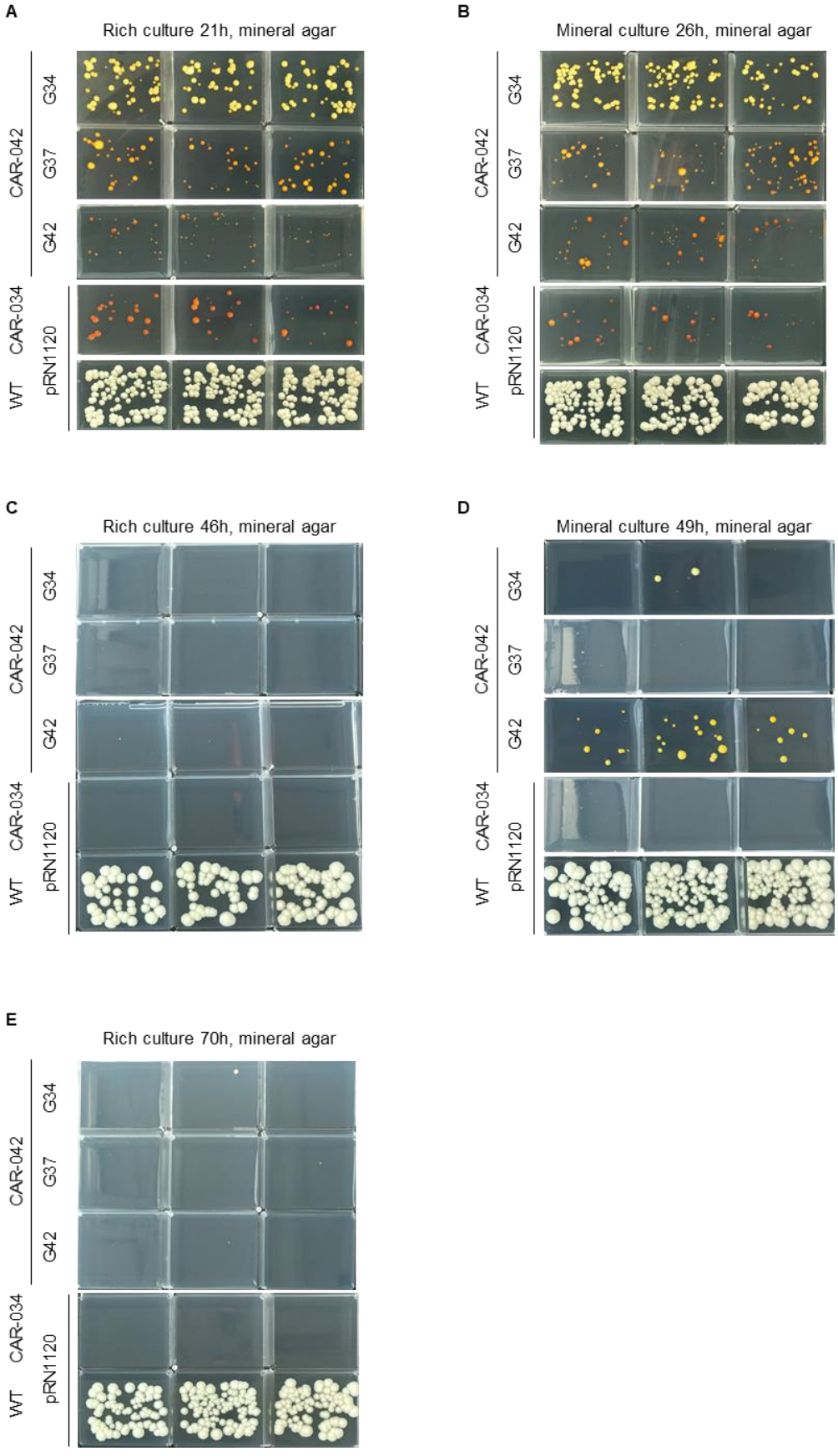
8 Supplementary information to Chapter 8

SI Table 8-1. List of primers used in this study

Name	Sequence	Purpose
pKC001	CATGTTTGACAGCTTATCATC	Amplification of dCas12a crRNA expression cassette (FW)
pKC002	CACACAGGAAACAGCTATGAC	Amplification of dCas12a crRNA expression cassette (RV)
pKC003	TACCGTCGACCTCGAGGGG	Amplification of dCas12a crRNA recipient plasmid pRN1120 (FW)
pKC004	ATATCGAATTCCTGCAGCCCG	Amplification of dCas12a crRNA recipient plasmid pRN1120 (RV)



SI Figure 8-1. Instability of β-carotene biosynthesis pathway during growth on rich agar plates



SI Figure 8-2. Instability of β-carotene biosynthesis pathway during growth on mineral agar plates