

# Synthetic noise control in eukaryotic gene expression and signal transduction



## Dissertation

zur

Erlangung des Doktorgrades  
der Naturwissenschaften  
(*Doctor rerum naturalium*)

Dem Fachbereich Biologie  
der Philipps-Universität Marburg  
vorgelegt von

**Max Mundt**  
aus Darmstadt

Marburg, 2018



Die Untersuchungen der vorliegenden Arbeit wurden von März 2014 bis November 2017 unter Betreuung von Prof. Dr. Victor Sourjik in Marburg am LOEWE-Zentrum für Synthetische Mikrobiologie (SYNMIKRO) der Philipps-Universität Marburg und dem Max-Planck-Institut für terrestrische Mikrobiologie in der Abteilung *Systems and Synthetic Microbiology* durchgeführt.

Vom Fachbereich Biologie der Philipps-Universität Marburg  
als Dissertation angenommen am 24.05.2018

Erstgutachter: Prof. Dr. Victor Sourjik

Zweitgutachter: Dr. Tobias Erb

Tag der mündlichen Prüfung: 08.06.2018



## **Erklärung**

Ich versichere, dass ich meine Dissertation mit dem Titel „Synthetic Noise Control in Eukaryotic Gene Expression and Signal Transduction“ selbstständig, ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfsmittel bedient habe.

Diese Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Max Mundt

Marburg, den \_\_\_\_\_



Zum Zeitpunkt der Einreichung dieser Dissertation wird die folgende Originalpublikation vorbereitet um die erzielten Ergebnisse zu veröffentlichen:

**Mundt, M.;** Anders, A.; Murray, S. & Sourjik, V. (2018). A system for gene expression noise control in yeast (in preparation)



*Meiner Familie*  
*Meinen Freunden*



## Table of contents

Erklärung .....	IV
Table of contents.....	X
Zusammenfassung.....	1
Summary .....	3
Abbreviations .....	5
1 Introduction.....	6
1.1 Gene expression noise .....	6
1.2 Determining the noise in a system.....	7
1.2.1 Fluorescent proteins.....	7
1.2.2 Flow cytometry.....	7
1.2.3 Noise calculation .....	8
1.3 Sources of biological noise .....	9
1.3.1 Intrinsic and extrinsic noise.....	9
1.3.2 Discriminating intrinsic from extrinsic noise.....	10
1.4 Noise control .....	12
1.5 Synthetic control of mRNA levels.....	14
1.5.1 Inducible promoters .....	14
1.5.2 Inducible ribozymes.....	15
1.6 Noise in signal transduction pathways.....	16
1.6.1 Signal transduction pathways .....	16
1.6.2 The yeast life cycle.....	17
1.6.3 The yeast mating pathway .....	17
1.6.4 Pathway noise versus gene expression noise .....	19
1.7 Aim of this thesis .....	20
2 Results .....	21
2.1 Development of a low-noise measurement platform.....	21
2.1.1 Optimizing biological parameters .....	22
2.1.2 Optimizing measurement parameters .....	24
2.1.3 Optimizing analysis.....	31
2.2 Proof-of-concept of a noise tuner .....	34
2.2.1 Characterization of ribozyme induction.....	34
2.2.2 Simultaneous control of transcription rate and mRNA degradation rate.....	35
2.2.3 Benchmarking of the noise tuner.....	41
2.2.4 Stochastic simulation of the noise tuner.....	42

2.3	Noise tuning in a signaling pathway.....	45
2.3.1	Ste2.....	47
2.3.2	Ste5.....	48
2.3.3	Fus3 .....	49
2.3.4	Msg5 .....	50
2.3.5	Sst2 .....	52
2.4	Noise control via expression rates in native genes .....	57
3	Discussion .....	61
3.1	Noise-optimized measurement systems.....	61
3.2	Synthetic control of gene expression noise .....	64
3.3	Synthetic control of pathway noise.....	67
3.4	Noise tuning via expression rates: A strategy employed in nature?.....	70
3.5	Conclusion and outlook.....	71
4	Materials and Methods .....	73
4.1	Chemicals and consumables .....	73
4.1.1	Growth media.....	73
4.1.2	Buffers and stock solutions .....	74
4.1.3	Reaction kits .....	75
4.2	Yeast strains.....	75
4.3	Plasmids.....	77
4.4	Molecular cloning.....	78
4.4.1	Primers .....	78
4.4.2	Polymerase Chain Reaction (PCR) .....	81
4.4.3	Gibson Assembly .....	82
4.4.4	Restriction-Ligation Cloning .....	82
4.4.5	Electrocompetent <i>E. coli</i> cells.....	83
4.4.6	<i>E. coli</i> transformation .....	83
4.5	Yeast strain development.....	83
4.5.1	Competent yeast cells .....	83
4.5.2	Genomic integration.....	84
4.6	Growth conditions.....	84
4.7	Fluorescent proteins.....	84
4.8	Flow cytometry.....	85
4.9	Microscopy .....	86
4.10	Calculation of intrinsic and extrinsic noise.....	86

4.10.1	Dual reporter assay .....	87
4.10.2	Reduced gate size analysis .....	87
4.11	Stochastic simulation.....	87
4.12	Calculation of mutual information .....	88
4.13	Morphology analysis .....	88
4.14	Software .....	89
5	References.....	90
6	Supplementary information .....	96
6.1	Supplementary table .....	96
6.2	Sequence alignments .....	96
6.3	Supplementary figures .....	101
6.4	Chemicals and consumables .....	104
7	Acknowledgements .....	105
8	Curriculum vitae .....	106

### Zusammenfassung

Jedem biologischen Prozess wohnt eine Zufallskomponente inne. Während der Genexpression entsteht „Rauschen“ durch stochastische Fluktuationen und Schwankungen in der Anzahl und Aktivität der involvierten Expressionsmaschinerie. Daraus resultiert, dass sich Zellen einer genetisch identischen Population in der Zahl ihrer Proteine unterscheiden, was sich negativ auf biologische Prozesse auswirken kann. Evolutionär hat die Natur auf Mechanismen selektiert, die Rauschen vermindern und der zellulären Fitness zugutekommen. Die meisten synthetisch-biologischen genetischen Schaltkreise und Signaltransduktionswege verfügen allerdings nicht über Systeme, die Genexpressionsrauschen regulieren, wodurch ihre Funktionalität limitiert werden kann.

In der vorliegenden Studie beschreiben wir ein synthetisches System zur Kontrolle von Genexpressionsrauschen in der Hefe *Saccharomyces cerevisiae*. Die präsentierten Daten wurden zum Großteil mittels Durchflusszytometrie aufgenommen mit einer Messkonfiguration, welche wir zuvor auf minimales unspezifisches biologisches und technisches Rauschen optimiert haben. Das von uns entwickelte System ermöglicht es die Variation in der Expression eines Zielgens einzustellen, indem dessen Transkriptionsrate und mRNA-Degradationsrate gesteuert werden. Dafür verwenden wir induzierbare Promotor- und Ribozym-Sequenzen. Messungen in denen wir einen Unterschied im Genexpressionsrauschen eines Fluoreszenzreporters um bis zu 300 % einstellen, demonstrieren die Funktionalität unseres Systems. Wir bewerten die Leistungsfähigkeit unseres „Noise Tuners“ indem wir ihn mit semi-synthetischen Systemen vergleichen, deren mRNA-Degradationsrate durch native Hefe-Terminatoren gesteuert wird.

Wir verwenden ein analytisches Modell, welches Genexpression mit der Verteilung von Proteinkonzentrationen auf Populationsebene verbindet. Auf der Basis dieses Modells betreiben wir stochastische Simulationen, die die experimentellen Ergebnisse reproduzieren und einen tieferen Einblick in die zugrundeliegenden Mechanismen erlauben: Im gegebenen Parameterbereich wird der Grad des Rauschens von der Transkriptionsrate definiert, wohingegen die Expressionsniveaus sowohl von der Transkriptionsrate wie auch der mRNA-Degradationsrate gesteuert werden.

Mit der Entwicklung des Noise Tuners verfolgen wir zwei Ziele: Zum einen die Reduktion von Genexpressionsrauschen in Systemen, in denen es sich als schädlich erweist; zum anderen wollen wir ihn als Werkzeug etablieren um den Einfluss von Rauschen in komplexeren Systemen zu untersuchen. Dazu testen wir den Noise Tuner für verschiedene Gene des Signaltransduktionswegs für die sexuelle Paarung in *S. cerevisiae*, einem Modellsignalweg für synthetisch-biologische Studien. Angewandt auf verschiedene Gene zeigen wir, dass der Noise Tuner das Rauschen in der Aktivität des Signaltransduktionswegs kontrollieren kann.

Eine detaillierte Analyse zum negativen Regulator *SST2* zeigt Unterschiede in der Variabilität des Signaltransduktionswegs von bis zu 50 %. Wir zeigen die physiologische Relevanz dieser Veränderung durch verbesserte Informationsweiterleitung, wenn *SST2*-Rauschen herabreguliert ist. Darüber hinaus untersuchen wir die Morphologien von ca. 10.000 Hefezellen und zeigen, dass Zellen mit niedrigem *SST2*-Rauschen bei Stimulation des Signalwegs präziser mit dem Paarungs-Phänotyp reagieren als Zellen mit verrauschter *SST2*-Genexpression.

Des Weiteren untersuchten wir, ob die hier beschriebene und erfolgreich getestete Strategie zur Kontrolle von Variabilität in der Genexpression auch in natürlichen Genen gezeigt werden kann. Dazu

## Zusammenfassung

wählten wir fünf Gene der Hefe aus, für die wir auf Basis publizierter Daten extrem hohe bzw. niedrige mRNA-Degradationsraten berechnet hatten. Wir nutzen die entsprechenden Promotoren und Terminatoren dieser Gene um das Genexpressionsrauschen in einem Reporter-gen zu messen und konnten einen generellen Trend feststellen für niedriges Rauschen bei hoher Transkriptions- und mRNA-Degradationsrate und genauso den umgekehrten Trend für hohes Rauschen bei niedrigen Raten. Für die zwei Gene mit der extremsten (hohen bzw. niedrigen) Variabilität konnten wir zudem Hinweise finden, die der Hypothese entsprechen, das Genexpressionsrauschen mit bestimmten Genfunktionen verknüpft ist.

In der vorliegenden Arbeit zeigen wir die Entwicklung, Konstruktion und erfolgreiche Anwendung eines synthetischen Werkzeugs zur Kontrolle des Expressionsrauschens von Genen. Die vorgestellten Ergebnisse illustrieren den Einfluss den das Genexpressionsrauschen einer individuellen Komponente eines Signaltransduktionswegs auf dessen Leistungsfähigkeit haben kann; wir zeigen aber auch, wie dieser Einfluss kontrolliert werden kann. Wir empfehlen, dass generelle Design-Prinzipien für niedrige Variabilität in der Genexpression - wie die hier vorgestellte Methode - verwendet werden um die Leistungsfähigkeit synthetischer Netzwerke zu verbessern.

## Summary

### Summary

A certain level of randomness is inherent to every biological process, causing individual cells in a clonally identical population to vary in the number of protein molecules. This variation that was termed gene expression noise arises from stochastic fluctuations and the variability of numbers and states of the involved expression machinery. Noise causes suboptimal protein concentrations, which can negatively affect biological processes. Nature has selected for noise-reducing mechanisms when they benefit cellular fitness. Most synthetic genetic circuits and signaling pathways, however, lack systems that control gene expression noise, which can reduce their functionality.

Here, we report the construction of a synthetic noise tuning system in *Saccharomyces cerevisiae*. We present data acquired by flow cytometry, using a measurement setup that we optimized for minimal nonspecific biological and technical variations. The system we developed allows the tuning of expression noise of a target gene using externally added small molecules to control the transcription rate via inducible promoters and the mRNA degradation rate via inducible ribozyme sequences. We demonstrate the functionality of the noise tuner by achieving up to 3-fold noise differences in the expression of a fluorescence reporter gene. We benchmarked the performance of the noise tuner by comparing it to semi-synthetic systems with fixed mRNA degradation rates, mediated by native yeast terminators.

Stochastic simulations of an analytical model that links gene expression to population-level distributions of protein numbers were used to reproduce the experimental findings and revealed the mechanisms underlying the observations: In the given parameter space, noise was mainly affected by the transcription rate, whereas the mean expression was governed by both, the transcription rate and the translational burst size defined by the mRNA degradation rate.

The objective of the development of the noise tuner was twofold: the first goal was to reduce gene expression noise in contexts where it proves to be detrimental. The second goal was to establish the noise tuner as a tool to investigate the influence of noise in complex networks. We applied the noise tuner to different genes in the yeast mating pathway, a model signal transduction pathway and the basis for numerous studies in the field of synthetic biology. We determined that the noise tuner, when applied to different genes of the pathway resulted in detectable changes in pathway noise.

Detailed analysis of the negative pathway regulator gene *SST2* set to either high or low noise resulted in up to 50 % difference in pathway noise between the two settings. We demonstrated that the low noise setting of *SST2* expression lead to improved information transmission through the pathway. Categorization of cell morphologies during stimulation with mating pheromone suggested a more precise, switch-like response in the low-noise *SST2* cells.

To investigate whether noise tuning principle we describe here was also applicable to native genes, we selected five yeast genes with reportedly extreme mRNA production and degradation rates. We used the corresponding promoters and terminators to drive the expression of a reporter gene to observe a general trend towards low noise for genes with high transcription and mRNA degradation rates and *vice versa*. The results of a gene ontology analysis of the two most extreme cases supported a hypothesis that noise levels are linked to protein function.

In this thesis we report the design, construction and successful application of a synthetic noise tuner. Our results illustrate the impact of gene expression noise of individual components on pathway

## Summary

performance – but we also show that this can be controlled. We suggest that design principles for low-noise gene expression, such as those presented in this thesis, should be taken into account for the synthetic modification and *de novo* design of signal transduction pathways and other networks.

## Abbreviations

### Abbreviations

a.u.	Arbitrary units
ATP	Adenosine triphosphate
bp	Base pair
ddH <sub>2</sub> O	Double-distilled water
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
FP	Fluorescence protein
GTP/GDP	Guanosine tri-/ diphosphate
mNG	mNeongreen
mRNA	Messenger RNA
nt	Nucleotide
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PRG	Pheromone response gene
RNA	Ribonucleic acid
rpm	Revolutions per minute
SD	Standard deviation
UTR	Untranslated region

### 1 Introduction

This work focuses on a molecular tool that we have developed to control gene expression noise in yeast. The following pages explain how noise can be measured using single-cell techniques, which underlying mechanisms cause noise, and the known existing mechanisms by which noise can be controlled. A strategy to control gene expression noise will be described which employs the regulation of the transcription rate and the mRNA degradation rate. Lastly, noise will be discussed in the context of biological networks, using the yeast mating pathway as a model for signal transduction.

#### 1.1 Gene expression noise

Living organisms share the remarkable capability to act deterministically at the macro level despite fluctuations at the micro level. From microorganisms to mammals, this characteristic is found throughout all domains of life: At any given time, thousands of random molecular processes occur within a cell. Still, the cell is capable of spatial and temporal organization, signaling and response to stimuli, homeostasis, and the other characteristics of life. How this randomness can be controlled and its sources and consequences are subjects of ongoing research.

For a long time, this randomness (we refer to it henceforth as noise) was ignored because measurements were mostly done “in bulk” and individual fluctuations were averaged out using the statistics of large numbers. The development of techniques to measure properties of single cells, such as flow cytometry, enabled the analysis of cell-to-cell variations in large populations of cells. Furthermore, the emergence of Synthetic Biology has advanced research on gene expression noise, based on the observation that synthetic genetic circuits show large fluctuations in their behavior (1).

One central notion of noise research is that on the level of individual cells, small numbers of molecules can play a crucial role for cellular function. Perhaps the most prominent example of such a molecule is DNA, which is present in only one to a few copies per cell and is the very basis for the process of gene expression that gives rise to a certain phenotype. Early noise research has shown that even in a clonally identical population of cells, the expression of individual genes differs from one cell to another (2). This gene expression noise results in diverse phenotypes that can be observed in microorganisms (see e.g. 2) as well as higher organisms (3), as illustrated in Figure 1.



Figure 1: Phenotypic consequences of gene expression noise. The fingerprints of identical twins exhibit different patterns (adapted from 3).

This thesis will focus on the characterization and control of gene expression noise in the yeast *Saccharomyces cerevisiae*, one of the most extensively studied unicellular eukaryotes. *S. cerevisiae* has been a model organism for the analysis of eukaryotic gene expression and its regulation, e.g. by heterochromatin remodeling (4, 5) and mRNA processing (6, 7, 8, 9). Moreover, yeast strains represent capable hosts for the recombinant production of value chemicals (e.g. 10) and are employed as biosensors (e.g. 11). The deep understanding of the regulation of genes and the wide variety of

## Introduction

biotechnological applications makes *S. cerevisiae* an excellent organism to study gene expression noise.

### 1.2 Determining the noise in a system

Noise can be determined by measuring a range of cellular variables, using single-cell level techniques and the extent of variability can be expressed in different numbers. Here, we focus on flow cytometry measurements of populations expressing fluorescent reporter genes which are analyzed by calculating the moments of the corresponding distributions (e.g. mean and variance).

#### 1.2.1 Fluorescent proteins

Noise is typically measured on the protein level. Proteins are the main effector molecules that mediate cellular functions. Thus, measuring protein heterogeneity gives the most insight into the expression capacity of a cell and its physiology. Their ease of use and the development of fast single-cell fluorescence measurement techniques (see section 1.2.2) have made fluorescent proteins (FPs) the main objects for gene expression noise analysis.

Independent of the color, all FPs share a similar structure with the chromophore being formed by three amino acid residues that are shielded by a barrel-like shape formed by beta sheets (12). FPs with the desired properties have to be chosen based on the aim of the experiment and the experimental and biological setup. The main properties that have to be considered are the excitation and emission spectra and the brightness.

The chromophore of a FP is only excitable with light in a certain range of wavelengths. A photon within this range can be absorbed to excite the chromophore which then can emit a photon with lower energy (i.e. higher wavelength) than the absorbed photon. This difference in wavelengths is referred to as Stokes shift and causes the difference between excitation and emission spectra in fluorescence measurements (see Figure 55 in the methods section).

Bright FPs increase the signal-to-noise ratio and allow high sensitivity. Brightness is defined by two parameters: The extinction coefficient and the quantum yield. The extinction coefficient is a metric for how well the fluorescent protein attenuates light of a certain wavelength, i.e. how much light it absorbs. The quantum yield describes how many photons of fluorescence light are emitted from the protein per photon from the excitation light source (12).

To allow the expression and measurement of more than one fluorescence reporter, the excitation and emission spectra of the FPs have to be well interspaced and separated. To measure in a heterologous host such as yeast, fluorescent reporter genes are usually codon optimized. In this process, some codons in the original sequence are changed so that the codons of the gene correspond to the tRNA abundance of the heterologous host organism.

#### 1.2.2 Flow cytometry

Noise measurements are challenged by the fact that all involved processes are subject to randomness. That does not only involve the cellular processes but also the selected measurement technique, in this case, flow cytometry. The resulting technical noise becomes part of the total noise that is measured and cannot readily be distinguished from the biological noise unless it has been estimated on the basis of experimental data (see e.g. Figure 23 in the results section).

## Introduction

Flow cytometers consist of a fluidics, optics and electronics part. Technical parameters at the optics and electronics level can be chosen to optimize the acquisition and to minimize technical noise. This is general best practice to improve signal-to-noise in measurements. For noise analysis it is even more crucial because here, not only is the signal supposed to be distinguishable from the noise but also one source of noise is supposed to be distinguishable from a separate source of noise.

In flow cytometry, laser beams hit the cells and create two types of signals. In the first, the laser is either diffracted in the forward or the side direction, creating forward (FSC) or side scatter signals (SSC), which are affected by the size and the inner granularity of the cells. Photons from the laser can also be absorbed by the cells which are then emitted as photons of lower energy (see section 1.2.1). These fluorescence signals scale with the concentration of fluorophores in the cell. Adjusting the laser power can be a potential parameter to optimize the measurement (see Figure 21 in the results section).

Unlike the diffracted signal, the fluorescence light emitted from the sample is not emitted in a single wavelength but in a spectrum of wavelengths. In order to analyze this complex signal, it is split into different frequency bands to be collected by the appropriate detector that translates signals into an electronic current. For this purpose, the filtered parts of the spectrum are transmitted to photomultiplier tubes (PMTs). Voltages are applied to regulate the strength of the electronic current. These PMT voltages are important technical parameters that can be adjusted to optimize flow cytometry measurements (see Figure 22 in the results section).

### 1.2.3 Noise calculation

Single-cell measurement techniques such as flow cytometry are a prerequisite for the analysis of cellular variation. This is due to the fact that in order to calculate the variation, apart from the first moment (mean,  $\mu$ ) also the second moment of the distribution (variance,  $\sigma^2$ ) has to be known. The square root of the variance is the standard deviation  $\sigma$  which, divided by the mean of a distribution gives the coefficient of variation (CV). The CV is commonly used to describe gene expression noise or the amount of cell-to-cell heterogeneity in a population. Normalization of the standard deviation to the mean is necessary to take into account that the variation scales with the mean (see section 1.3.2).

The mean and CV are common statistical estimators. In measurements of biological parameters, however, data often contain outliers. For that reason, it can be useful to employ alternative metrics for noise that are more robust towards extreme values. The median of a distribution is the middle value, separating the lower from the higher half of the distribution. By this definition, outliers have a smaller impact on the median as compared to the mean. The absolute deviation of the median can be calculated for every data point in the distribution, creating a second set of values. The median of this second dataset is the median absolute deviation (MAD), which again is impacted only marginally by outliers. To use the MAD as a consistent estimator of the standard deviation of normal distributions, it is scaled with the factor 0.6745 (because for normal distributions  $MAD \approx 0.6745\sigma$ ). The MAD normalized by the median gives a robust version of the coefficient of variation (rCV).

## Introduction

### 1.3 Sources of biological noise

The biological sources of noise can be discriminated into an intrinsic and an extrinsic component which will be discussed in the following sections.

#### 1.3.1 Intrinsic and extrinsic noise

Gene-specific or intrinsic noise describes the stochasticity in the expression of individual genes (Figure 2, left side). This accounts for the inherent randomness of the chemical reactions, binding events, and micro-fluctuations that occur e.g. during promoter activation, transcription and translation of an individual gene. RNA polymerase for instance can be recruited to a specific promoter by transcription factors that bind to specific cis-regulatory elements in the DNA. This process depends on the grade of condensation of the DNA in that locus, transcription factor-DNA binding affinity, and how well RNA polymerase is recruited by the transcription factor. Further randomness during transcription can be introduced by RNA polymerase stalling (6). Stochastic processes on the transcriptional level include mRNA maturation and degradation. Translational efficiency is also subject to fluctuations, e.g. due to stochastic tRNA binding and availability.

These and further factors lead to transcription and translation occurring in random bursts: Once a promoter is accessible and the transcription machinery is recruited, many transcripts can be synthesized in a short period of time, creating temporal peaks in mRNA concentration. Transcriptional bursts are in turn passed on to the translational level because the produced mRNA molecules are the template for many rapid rounds of translation, leading to many proteins being produced from few transcripts. The measured noise partly depends on how many molecules are produced per burst (burst size) and how often these bursts occur in a given timeframe (burst frequency). Big, infrequent bursts create more noise than small, frequent bursts (13).

## Introduction

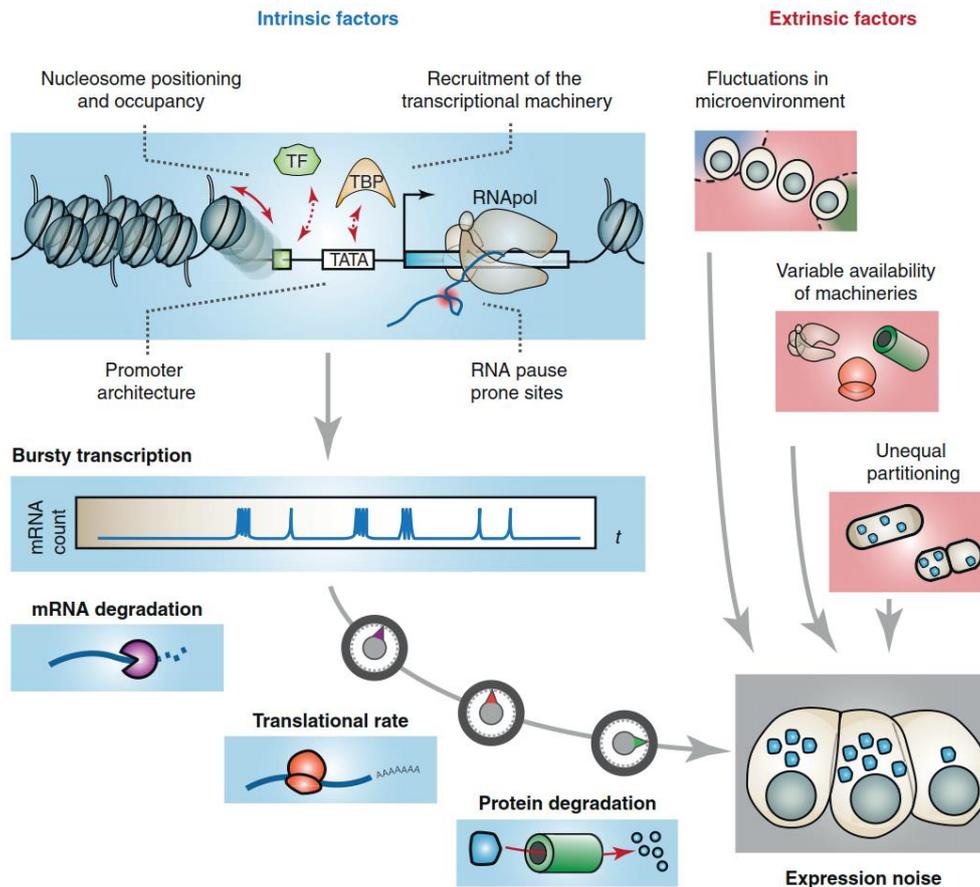


Figure 2: Factors contributing to intrinsic and extrinsic noise. Intrinsic noise (left) is gene-specific and includes the stochasticity of all rates involved in gene expression. Extrinsic noise (right) comprises global differences between one cell and another. Both categories of noise add up to the total observed noise in the concentration of a protein (adapted from 14).

Extrinsic noise refers to differences between clonal cells which cause variation in gene expression capacity that affects all genes of a cell globally (Figure 2, right side). Especially on solid medium, cells in a population can be affected by differential availability of nutrients. For instance, cells on the edge of a colony have a higher local concentration of nutrients than cells in the center. As a result, cells with more nutrients available have a more active gene expression. Unequal partitioning can also play a role for cell-to-cell differences in expression, especially in the case of *Saccharomyces cerevisiae*, which divides by budding, creating a much smaller daughter cell with the same set of genes but lower numbers of polymerases and ribosomes to express them. Generally, the cell size scales with the overall transcription and translation output (15). Extrinsic noise has also been explained by changes in transcriptional activity in different stages of the cell cycle (16).

### 1.3.2 Discriminating intrinsic from extrinsic noise

Extrinsic and intrinsic noise can be determined and distinguished experimentally. The first method to describe intrinsic and extrinsic sources as orthogonal contributors to the total observed noise was published by Elowitz and colleagues in 2002. Two identically regulated fluorescence reporter genes were introduced into *Escherichia coli* and the fluorescence intensities were measured by flow cytometry (2). The intensities of the two fluorescent reporters are plotted against each other. Extrinsic noise is defined as the correlated variation of those reporters (Figure 3 B and C, top panel and middle panel, respectively). A hypothetical population exhibiting only extrinsic noise would differ only in absolute, but not relative fluorescence intensity. Intrinsic noise is defined as the uncorrelated variation of the two reporters that introduces differences in the expression of the reporters in individual cells

## Introduction

(Figure 3 B and C, bottom panel and left panel, respectively). Thus, the cells do not only differ in absolute amounts of protein molecules but also in relative amounts, so that at a given time a cell can contain more of one reporter than of the other.

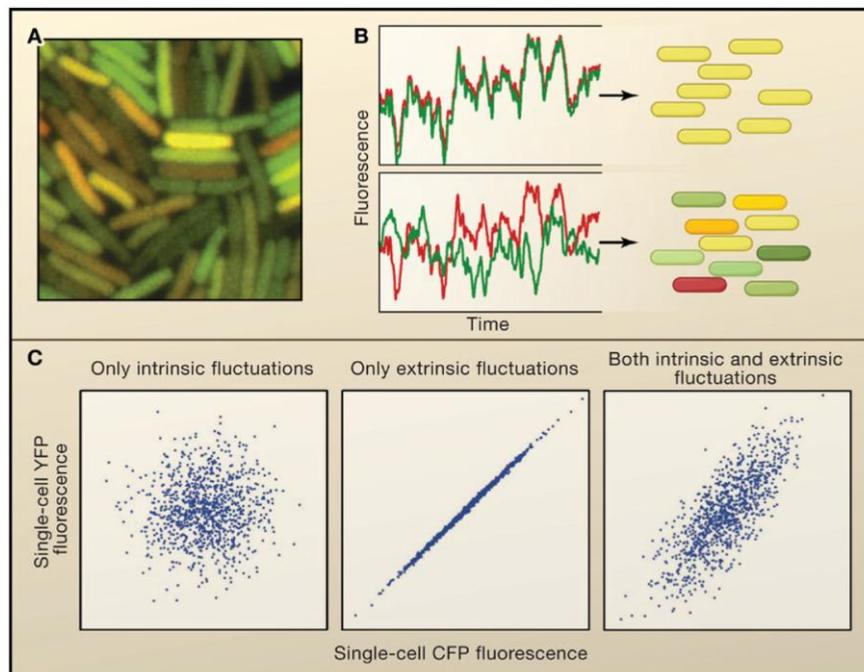


Figure 3: Intrinsic and extrinsic noise measured in a dual-reporter assay. (A) Fluorescence microscopic image of *E. coli* cells harboring two different but identically controlled fluorescent reporter genes. (B) Differences in brightness and color are caused by fluctuations that are correlated (top panel) or uncorrelated (bottom panel). (C) Hypothetical scatter plots of cells exhibiting only intrinsic (left panel) or extrinsic noise (middle panel), and the combination of both as it can be measured in a dual-reporter assay (right panel, adapted from 17)

For yeast experiments, another method can be used to discriminate intrinsic from extrinsic noise, which takes advantage of the fact that FSC and SSC signals from flow cytometry experiments are good predictors for the gene expression capacity. For this method, termed reduced gate size analysis (18), FSC and SSC are plotted for a population of cells and a subset of the population is defined by drawing a circular gate around it. Cells within the gate share more similar FSC and SSC values with each other than with cells outside of the gate. The phenotypic differences, defined by the variation of FSC and SSC values in the subset become smaller as the radius of the gate is reduced. When the noise of the fluorescence intensities is plotted for subsets of the population with decreasing gate radii, the values converge towards a noise level that corresponds to the contributions of intrinsic and technical sources of noise (Figure 4). The difference in noise between the reduced gate and the full gate corresponds to the extrinsic noise evoked by phenotypic differences of the cells.

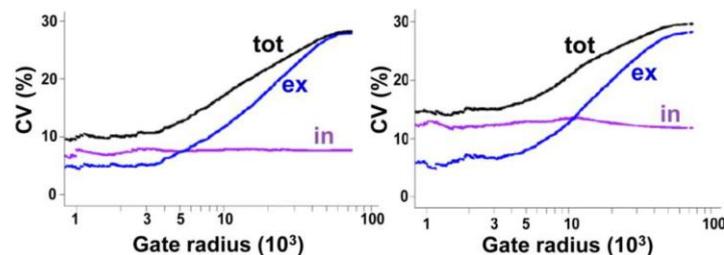


Figure 4: Reduced gate size analysis. Two examples of strains with differentially regulated fluorescence reporter genes. Whereas intrinsic noise (purple) does not change with gate radius, the extrinsic noise (blue) converges to a minimum when cells have sufficient phenotypic similarity (adapted from 19).

## Introduction

The necessary number of acquired cells has to be sufficiently high for the reduced gate size analysis. Otherwise, if the gate size is reduced beyond a critical point, the sample size is too small to yield meaningful values.

The contributions of intrinsic and extrinsic noise to the total noise depend on the mean expression level. It was found that at low mean expression, intrinsic noise makes a bigger contribution to total noise, whereas at high means, extrinsic noise is dominant (20). Generally, when a process such as gene expression occurs with a constant probability in time (Poissonian probability distribution), the coefficient of variation scales with the inverse of the square root of the mean protein abundance (Figure 5).

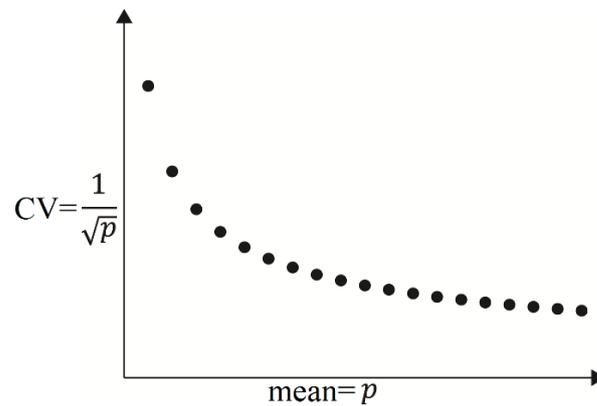


Figure 5: General mean/ noise relationship. In Poissonian distributions, the mean  $\mu$  equals the variance  $\sigma^2$ , so that  $CV = \frac{\sigma}{\mu} = \frac{1}{\sqrt{\mu}}$  meaning that the noise (Y-axis) scales with the inverse square root of the mean protein abundance  $p$  (X-axis).

This can also be observed experimentally and is a major reason why the noise of two genes/ states/ timepoints etc. have to be compared with regard to similar mean expression.

### 1.4 Noise control

The correlation of mean and noise suggests that high expression rates are beneficial as low noise would be achieved. However, many genes have to be expressed at an evolutionary conserved optimal level with fitness deficits for the cell if the protein concentration deviates too far from this optimum (21). The same is true for synthetic circuits: metabolic pathways, for instance, employ a set of enzymes that, in a multi-step process, produce secondary metabolites. Overexpression of these enzymes has been shown to result in suboptimal yields. This can be explained by the high expression burden on the cells as well as, depending on the pathway, accumulation of toxic intermediates (22, 23). Low noise at the optimal mean is vital for the performance of metabolic, genetic, and signaling circuits. This section will cover strategies by which mean expression can be decoupled from expression variation to allow noise control.

Even at the same mean, noise levels of native genes can differ largely from one another. It is an accepted view, that the variation of native proteins reflects a tradeoff of costs and benefits of noise control in that particular gene. High variability in gene expression has been shown to be beneficial for the cells in some contexts or for certain genes. Noisy stress response genes are suggested to increase adaptability of a population of cells when environmental conditions change e.g. a subpopulation that deviated a lot from the mean expression might have an advantage because it can facilitate a faster stress response. Noisiness in stress-response genes can thus be a trait that is selected for to favor heterogeneous populations. (24) For most genes, however, expression noise reduces cellular fitness

## Introduction

(25) and is generally counter-selected (26). Low noise is beneficial for the cell in the majority of genes, including those involved in homeostasis, cell cycle control, and protein synthesis (18). The different strategies that are observed in nature and that are employed by researchers to control variation in gene expression will be discussed in the following section.

A strategy to decouple mean from noise has been the subject of several studies. Their aim was to independently control the expression and activity of an upstream inducible regulator of the target gene (27, 28, 29, 30). In these cases, low expression of the regulator increases its noise and thus, the noise in the transcription of the target gene. To adjust the expression of the target gene for different concentrations of regulator proteins, different levels of inducer molecules are used, that bind to the regulator and change its affinity to the promoter of the target gene (31).

Other studies suggest that instead of regulating an inducer protein, noise can be effectively controlled via the basic rate constants of gene expression. In the simplest model of gene expression, the steady state level of a protein is governed by the rate constants of four processes: transcription, mRNA degradation, translation, and protein degradation. Thus, to achieve a desired concentration of proteins, two strategies can be used: The gene is transcribed at a high rate but transcripts have a low translation rate. This leads to many mRNA molecules that are translated rarely or not at all. The second strategy is that genes are transcribed at a low rate but the translation rate is high. In this case, only few transcripts are made but they are translated multiple times (Figure 6). Stochastic modeling of intrinsic/ gene-specific noise in these two cases show that the first case creates lower variation of protein concentration over time than the second case because stochasticity at the transcription level is averaged out due to a higher number of transcripts (32). This concept is universal but has been shown to be specifically pronounced in eukaryotes, where transcription can be particularly noisy due to stronger transcriptional bursts that occur caused by heterochromatin remodeling (3).

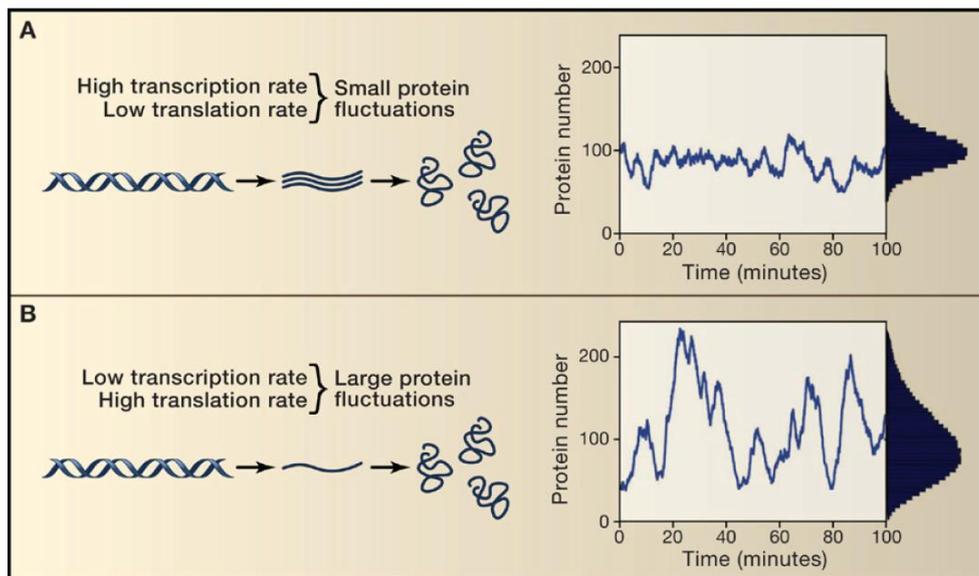


Figure 6: Two strategies by which expression rates of individual genes change noise at the same mean. (A) High transcription and low translation rates filter out stochasticity of transcription and lead to small variation of the protein concentration over time. (B) Low transcription and high translation rates amplify transcriptional stochasticity and lead to big fluctuations of the protein concentration over time (adapted from 17).

This mechanism was used to artificially control gene expression noise in yeast, using an inducible promoter to set the transcription rate combined with GFP in two different codon variants that

## Introduction

conferred different translation efficiencies. High translation efficiencies were shown to amplify transcriptional noise (27). Another study showed that the mRNA stability conferred by different native yeast terminator regions scaled with the transcriptional burst size of fluorescence reporter gene expression (33).

Fraser and colleagues used available whole genome data for yeast mRNA transcription and degradation rates, as well as translation rates to estimate the noise of genes based on the above model. They hypothesized that noise should be especially low in essential genes and those genes that encode subunits of multiprotein complexes. They found that the calculated noise in those genes was low due to high transcription rates and low mRNA stability and/ or translation rates (34).

Noise control that acts on mRNA level has been suggested to be the most effective strategy (35) and also produces low wastage rates, meaning low metabolic burden on the cells (36). The following section will elaborate on the means to control steady state mRNA levels using inducible genetic elements, namely promoters and terminators.

### 1.5 Synthetic control of mRNA levels

#### 1.5.1 Inducible promoters

In the native context, inducible promoters allow cells to react to environmental signals (toxins or nutrients) by direct interaction with a mediator (activator or repressor) that in turn changes its binding affinity to an operator site to regulate the expression of a target gene. One example of this is the regulation of tetracycline resistance by tetracycline-controlled transcriptional activation in *Escherichia coli*. Molecules of the antibiotic tetracycline bind to the repressor TetR, which in turn is released from Tet operator sites (TetO) to allow transcription of the exporter TetA. This system, thus, mediates resistance by export of tetracycline only on demand when tetracycline is present.

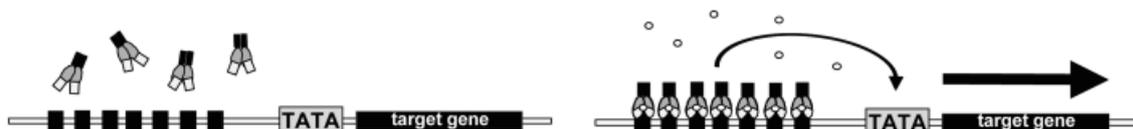


Figure 7: The Tet-ON expression system. Left: Reverse tetracycline trans-activators are constitutively expressed. Proteins consist of a DNA binding domain (white), an inducer binding domain (light gray), and an activation domain (black). Right: Upon binding to doxycycline (white circles), the reverse tetracycline trans-activator (rtTA) binds to the TetO sites in the proximal promoter (black squares) and recruits components of the transcription machinery to the core promoter (adapted from 37).

The tetracycline-controlled gene expression system has been adapted and modified for yeast and higher eukaryotes (38). Fusing activator domains to TetR in yeast created a tetracycline trans-activator (tTA, 39) and mutagenesis of tTA yielded a variant with reversed characteristics: rtTA binds to the TetO sites to activate transcription (40). Added tetracycline, or its derivate doxycycline, binds to the rtTA and triggers a conformational change that increases the binding affinity of rtTA to the TetO sites. Binding of rtTA to the DNA initiates transcription. Different versions of tetracycline-responsive promoters have been developed. Variation of the number of TetO sites has been shown to influence absolute expression and leakiness (41).

## Introduction

### 1.5.2 Inducible ribozymes

Downstream adjacent to the open reading frame, the 3' untranslated region (3' UTR) and the terminator are located. The corresponding sequences can overlap partly, which makes it useful to broadly define the combinations of both elements as "terminator region".

Terminator regions affect gene expression through transcriptional termination and post-transcriptional regulation. During transcriptional termination poly(A) signals and GU-rich sequences in the pre-mRNA determine the site for mRNA cleavage and subsequent polyadenylation. On the post-transcriptional level, the part of the terminator region encoding the 3' UTR plays a role in regulating mRNA degradation rate and translation rate. Regulatory proteins can bind to motifs in the 3' UTR to repress translation or flag the mRNA for degradation (42) or to stabilize the transcript (43).

Native terminator regions can be employed to increase mRNA and protein yields (44). A comparison of over 5000 sequences inserted downstream of a fluorescent reporter gene showed that native terminator regions confer a 70-fold range of expression levels in yeast (45). A study on expression-enhancing terminators found that they primarily act by increasing mRNA stability (46).

Similar to the approaches to develop artificial, well defined promoter systems, different efforts are being made to standardize and functionalize terminators in yeast for improved controllability and performance. Synthetic yeast terminators of less than 100 nt have been developed that contain a single poly(A) signal and the minimum number of necessary elements for efficient termination that confer an expression level similar to or stronger than a commonly used native terminator (47).

Riboswitches are a class of genetic regulatory elements which consist of RNA sequences that are typically found within the UTRs of genes. Riboswitches contain an aptamer domain that binds a metabolite, triggering a conformational change that affects gene expression by attenuating transcriptional termination or changing translation efficiency (48). Inducible ribozymes are a class of riboswitches that exhibit a nuclease activity which, upon metabolite binding to the aptamer, can be activated or disrupted. Such a construct embedded into the 3' UTR of a mature eukaryotic mRNA can be used to tune transcript stability: through its autocatalytic activity, part of the mRNA containing the ribozyme and the poly(A) tail is cleaved off from the part containing the ORF. Without a poly(A) tail that protects the transcript from exonucleolytic decay, the ORF is rapidly digested by the mRNA degradation machinery. As a result, the steady state of protein concentration is reduced (49).

## Introduction

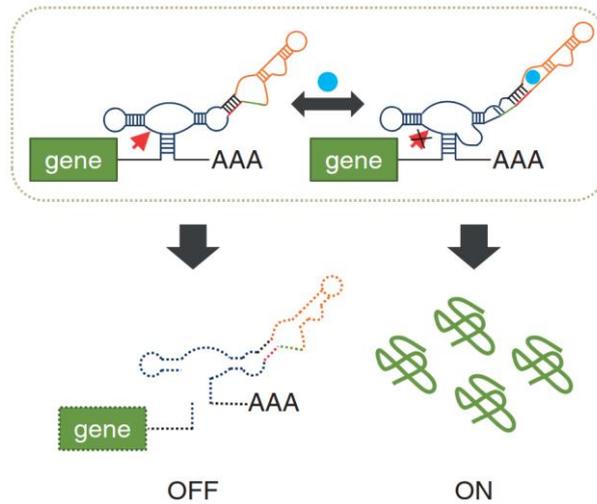


Figure 8: Design of an inducible ribozyme. After transcription, the ribozyme sequence folds into a secondary structure with an autocatalytic cleavage function, resulting in a strand break within the sequence (red arrow), dissociation of the poly(A) tail and subsequent rapid decay of the mRNA. Inducible ribozymes have been developed to bind to the secondary structure and induce a shift in the stem loop that disrupts the autocatalytic function. The mRNA is stabilized and has a higher probability of being translated (image adapted from 49).

The above examples of inducible genetic elements to control mRNA levels are the basis for the gene expression noise tuning systems presented in results section 2.2 of this thesis.

### 1.6 Noise in signal transduction pathways

Gene expression noise can impact the variability of the network or physiological process that the gene product is part of. In that respect, it is particularly crucial to consider noise in processes that involve the conveyance of information. Information is typically transmitted either in genetic circuits or signal transduction pathways (50). As opposed to genetic circuits, signal transduction pathways act on the level of protein activity, allowing information to be passed on faster as compared to transcriptional activation. Ultimately, the output of a signal transduction pathway can act on gene expression and thus, can modulate genetic circuits. The following sections describe general principles of signal transduction pathways, using the example of the *Saccharomyces cerevisiae* mating pathway. Furthermore, it will be elucidated which role noise plays in such pathways.

#### 1.6.1 Signal transduction pathways

Signal transduction pathways allow cells to rapidly process environmental signals and initiate an appropriate response. These biochemical networks relay the information to downstream effectors which can modulate gene expression to ensure fitness. Signal transduction pathways are studied in a wide variety of organisms to understand information flow from external stimuli upstream to the corresponding downstream effects, such as the modulation of gene expression. The signals are typically being transduced by phosphorylation. Kinases phosphorylate a substrate protein by transferring a phosphate group from a donor (usually adenosine triphosphate, ATP). Phosphorylation triggers a conformational change in the target protein's structure, activating or deactivating it, or otherwise changing its function. Phosphatases can release phosphate groups from the substrate protein to bring the protein back to its original state.

Mitogen-activated protein kinase (MAPK) cascades are sequences of phosphorylation events. In this process, a kinase activates another kinase by phosphorylation, which in turn activates another kinase. MAPK cascades are conserved in all eukaryotes from yeast to human as a central integration module

## Introduction

for cellular information processing. The cascade topology allows it to amplify and insulate signals and to improve the signal-to-noise ratio (51). The mating pathway of the yeast *Saccharomyces cerevisiae* is a model MAPK signal transduction pathway that processes an external stimulus to make a cell fate decision crucial for cellular fitness. The next section describes the *S. cerevisiae* life cycle and the topology and function of the mating pathway.

### 1.6.2 The yeast life cycle

*Saccharomyces cerevisiae* cells can exist in a haploid or a diploid form. Haploid cells can have either of two mating types, MATa or MAT $\alpha$ , that differ in one locus (mating type locus). MATa and MAT $\alpha$  cells, as well as diploids, undergo a mitotic life cycle. Haploid cells sense pheromones secreted by the opposite mating type and initiate a mating response. Depending on the pheromone concentration, cells undergo drastic morphological changes to prepare for cell fusion. In the process, cells adopt a unipolar budding pattern (52) along the gradient towards the highest pheromone concentration (53) before the cell cycle is synchronized by cell cycle arrest in G1 phase. At higher pheromone concentrations, the cells undergo morphological changes, become elongated and increase in volume by chemotropic growth towards the pheromone source. Cells that sense pheromone concentrations that indicate a mating partner is in close proximity develop a protrusion (shmoo) to achieve physical contact and ultimately fuse with a cell of the opposite mating type to form a diploid (Figure 9).

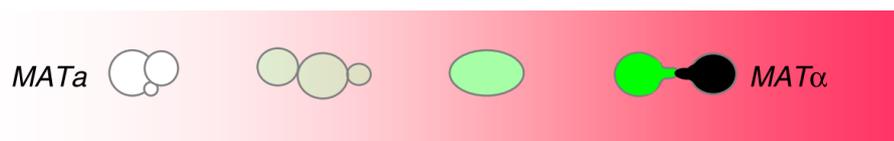


Figure 9: MATa cells (white and green) sense pheromone (red) secreted by MAT $\alpha$  cell (black). Cells respond to different pheromone concentrations with different morphogenesis: Axial budding when no pheromone is present, pseudohyphal and chemotropic growth at intermediate pheromone concentrations, and the shmooing phenotype at high levels of pheromone. Shade of green indicates mating pathway activity (adapted from 54).

The decision to mate poses a risk to the facultatively sexual *S. cerevisiae* cells. Successful sexual reproduction is an evolutionary advantage but it comes at a cost for the cells: During the mating response the cell wall integrity is reduced and cells lose time to proliferate due to the cell cycle arrest. In addition to that, expressing the components of the mating pathway and the pheromone response genes is a metabolic burden to the cell. This indicates a fitness trade-off between growth rate and mating efficiency (55). Accurate sensing of sexual partners is crucial and explains why yeast has evolved to sense the ratio of potential mating partners to competitors (54) and to track even very shallow gradients of pheromone to increase mating efficiency (56).

### 1.6.3 The yeast mating pathway

The molecules activating the mating pathway are the small peptide pheromones  $\alpha$  factor (secreted by MAT $\alpha$  cells) and a factor (secreted by MATa cells). Pheromones bind to G protein-coupled receptors (GPCRs) on the surface of yeast cells of the opposite mating type (Ste2 and Ste3 for MATa and MAT $\alpha$ , respectively). Upon ligand binding, the receptor Ste2 undergoes a conformational change that allows it to replace guanosine diphosphate (GDP) with guanosine triphosphate (GTP) in Gpa1p, the  $\alpha$ -subunit of the G protein. This causes a conformational change that dissociates the G $\beta\gamma$  complex (Ste4-Ste18).

The G $\beta\gamma$  complex associates with the scaffolding protein Ste5 and serves as an adapter for the kinase Ste20 to phosphorylate Ste11. This initiates the MAPK cascade during which Ste11 (a MAPKK kinase) phosphorylates Ste7 (a MAPK kinase), which in turn phosphorylates the MAP kinase Fus3. The cascade

## Introduction

is acted out while all involved kinases are attached to the Ste5 scaffold. Binding to Ste5 increases the phosphorylation rate of Fus3 by Ste7 (57) but scaffold proteins also prevent signal leakage: Ste11, for instance, is a component in both, the mating pathway and the high osmolarity / glycerol (HOG) pathway. Ste5 channels Ste11 activity towards a mating response, isolating the pathways from one another (58).

When phosphorylated, the MAPK Fus3 dissociates from Ste5 and translocates into the nucleus where it activates the transcription factor Ste12, which in turn activates expression of ca. 200 pathway response genes (PRGs). One of these encodes the cell cycle inhibitor Far1, which induces polarized growth and cell cycle arrest in G1 phase (59).

Among the activated genes are also components of the pathway itself, such as *FUS3*, resulting in an autocatalytic positive reinforcement of pathway activity. Positive feedback loops can convert the graded pathway response into a switch-like cell cycle response to allow an “all or nothing” decision once a threshold pheromone concentration has been reached.

Ste12 also upregulates expression of negative feedback regulators such as the phosphatase Msg5 and the GTPase-activating protein Sst2. Negative feedback can increase the dynamic range of the pathway and prevents random activation as well as prolonged activity, which can lead to cell death (60). Msg5 targets Fus3 to allow adaptation of the pathway response (61). *SST2* was the first gene identified to regulate G protein signaling (62). Sst2 accelerates hydrolysis of GTP to GDP to downregulate activity of the G $\alpha$  subunit Gpa1p. Typical for negative feedback regulators, Sst2 also reduces pathway noise (63, 64).

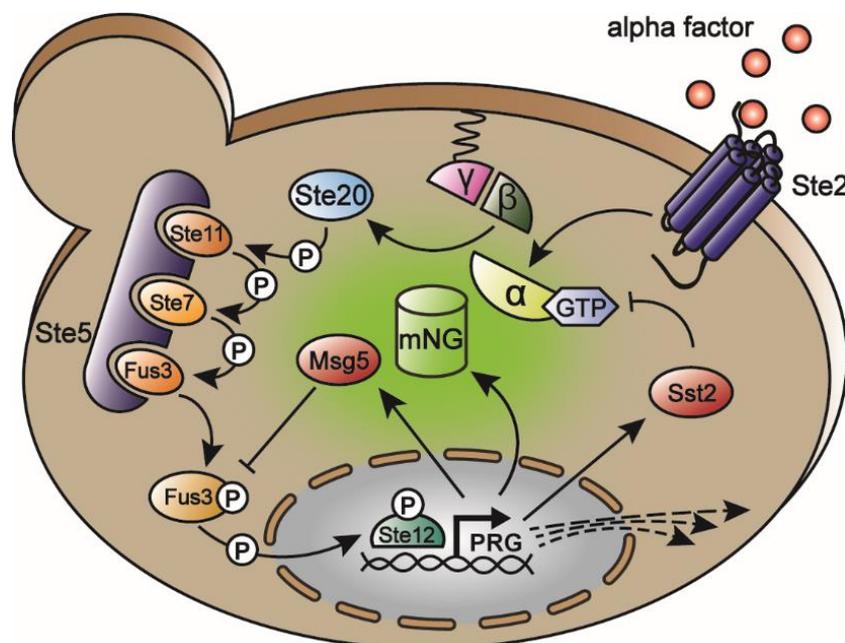


Figure 10: Important components of the yeast mating pathway. Alpha factor binds to Ste2 on a MATa cell, leading to the release of the  $\alpha$  subunit of the G protein from the  $\beta/\gamma$  subunits ( $G\beta\gamma$ ).  $G\beta\gamma$  then recruits the kinase Ste20. A primary substrate of Ste20 is the MAPKKK Ste11, which upon phosphorylation by Ste20 phosphorylates Ste7, which in turn phosphorylates Fus3. The MAPK cascade is facilitated by the scaffolding protein Ste5. Active Fus3 phosphorylates the transcription factor Ste12 which induces ca. 200 pheromone response genes (PRG). Among those are the negative feedback regulators Msg5 and Sst2 that act downstream and upstream in the pathway, respectively. Fluorescence reporter genes (here mNeogreen, mNG), driven by a PRG promoter can be used as reporter of pathway activity.

## Introduction

### 1.6.4 Pathway noise versus gene expression noise

The concept of intrinsic and extrinsic noise described in section 1.3 can be applied to the expression of pathway output (i.e. pheromone response genes) and to the signal transduction itself (Table 1). Pathway variability caused by intrinsic factors are termed transmission noise. These are the stochastic fluctuations that determine whether or not a signal is transmitted from one pathway component to another; for example, whether or not a phosphate group is transferred from ATP to Ste7, catalyzed by Ste11. Correspondingly, stochastic fluctuations in expression cause intrinsic gene expression noise.

Extrinsic factors of pathway noise are governed by the number, location and activity of pathway components. These determine whether or not a pathway component is available to help transmit the signal. For instance, if the number of Ste5 scaffolding proteins is too low, the signal cannot be transmitted efficiently through the MAPK cascade. This extrinsic pathway noise limits the capacity with which cells can transmit signals. Correspondingly, extrinsic gene expression noise limits the expression capacity (65).

*Table 1: A signaling pathway response can be subdivided into a pathway subsystem and a gene expression subsystem. The measured noise in a pathway reporter is the combination of intrinsic and extrinsic noise from the two subsystems (definitions based on 65).*

<b>Subsystems \ sources of variation</b>	<b>Intrinsic factors (stochastic fluctuations)</b>	<b>Extrinsic factors (Cell-to-cell differences in number, location and activity of involved molecules)</b>
Pathway	Transmission noise (Intrinsic pathway noise)	Extrinsic pathway noise (→ pathway capacity)
Gene expression	Intrinsic gene expression noise	Extrinsic gene expression noise (→ expression capacity)

In this noise model, which takes the pathway variability into account, the measured noise of a pathway output is a combination of the intrinsic and extrinsic factors of pathway noise and gene expression noise. Genes that are regulated similarly share substantial parts of their variation (66), indicating that the capacity of the upstream regulatory pathway is a major contributor to the noise in the observed output.

The yeast mating pathway has been the basis for a growing number of synthetic signaling circuits. Pathway components have been altered in several ways to change pathway response dynamics (67, 68). Cell fate control was achieved using synthetic control systems (69). Rewiring of pathway components can be used to change the pathway output altogether (70).

Engineered mating pathways have been used to develop artificial quorum sensing (71) and altered social behaviors (72). Synthetic pathway variants have been developed for unicellular (73) and multicellular (74) memory devices and for biological computation in multicellular networks (75). Additionally, different approaches highlight the use of the mating pathway for biosensing, e.g. by coupling it to heterologously expressed mammalian GPCRs to identify ligands (11, 76, 77).

Synthetic alterations of the mating pathway come along with disturbances of the evolutionarily optimized pathway topology. Controlling the noise in critical pathway components can improve the performance of the desired pathway function.

## Introduction

### 1.7 Aim of this thesis

Synthetic biology has brought about a large variety of synthetic networks, including genetic circuits and signal transduction pathways with artificial inputs and outputs. The unicellular eukaryote *Saccharomyces cerevisiae* has been a model platform for the development of such networks, due to the immense knowledge of its genetics and cell biology and the availability of well-characterized genetic control elements. Unlike native systems, synthetic networks lack parts or all of the control mechanisms that allow functionality despite the noise that is inherent in all involved processes. External tuning of the expression noise of a network gene provides a good approach to ensure accurate concentrations of components that are critical for network performance. Tuning gene expression noise means that the variation of protein concentration has to be decoupled from the mean protein concentration. Studies suggest that noise control is most effective and metabolically cheap when executed on the transcript level (36, 35).

The first aim of this thesis was to develop a robust measurement platform with minimal nonspecific biological and technical noise (results, section 2.1). This is a prerequisite for the analysis: Distinguishing a signal from noise is not sufficient in this context, in fact, the measurement system has to enable the distinction of one source of noise from another source of noise.

The second aim was to develop a system in yeast that allowed external tuning of gene expression noise by orthogonally controlling the transcription rate via inducible promoters and the mRNA degradation rate via inducible ribozymes (results, section 2.2). This proof-of-concept system should be capable of decoupling mean and noise of a fluorescence reporter gene in the same expression range as native genetic regulatory elements. It should allow noise tuning at a desired mean expression to either reduce noise when necessary, or as an analysis tool to investigate the impact of gene expression noise on entire pathways.

Next, we wanted to test the developed system in a complex network, namely the yeast mating pathway (results, section 2.3) to find out if our noise tuner, applied to individual pathway components, was capable of altering pathway noise and correspondingly, pathway precision.

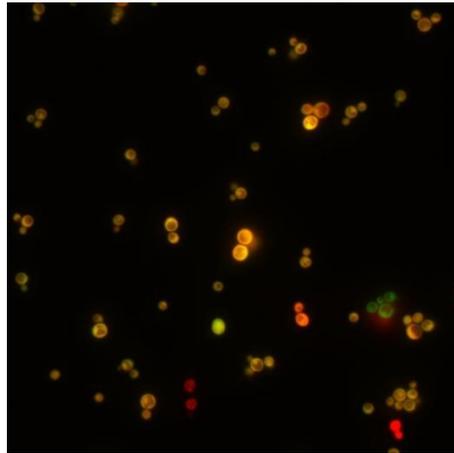
Finally, we wanted to assess experimentally if the described strategy for gene expression noise control was employed by native yeast promoters and terminators (results, section 2.4) and if this was in accordance with existing hypotheses that established a link of the noise level of a gene to its function.

## Results

### 2 Results

#### 2.1 Development of a low-noise measurement platform

In order to measure noise with high sensitivity, the measurement system has to be optimized to prevent nonspecific noise from masking the sources of variation that are intended to be detected. We chose multiple noise-reducing approaches covering different aspects of the measurement. Biological noise was reduced by changing genetic elements or the genetic context and technical noise was estimated by screening parameters of the flow cytometer. Measurements were done with single-, dual-, or triple-reporter strains. An example of noisiness in raw data of a dual-reporter strain measurement is shown in Figure 11.



*Figure 11: Example fluorescence microscopy image of cells from a dual-reporter strain harboring identically regulated mCherry and mNeogreen genes. Differences in brightness and color of the cells indicates different sources of noise.*

We chose mTurquoise2, mNeogreen and mCherry as fluorescence reporters because they have low sequence similarity (see supplemental figures). This is an important feature to consider to reduce the risk of homologous recombination, which *Saccharomyces cerevisiae* does very efficiently. Low sequence homology should thus improve genomic integrity and reduce artifacts that otherwise might affect noise measurements.

## Results

### 2.1.1 Optimizing biological parameters

To increase expression from the Tet promoters even at low promoter activity, we replaced the original *CYC1* 5' UTR with the strong *ACT1* 5' UTR (78). Single-reporter yeast strains harboring either one of the two promoter variants driving mCherry or mNeogreen expression were measured with the fully induced Tet-ON system (Figure 12, for details on the Tet-ON system see Figure 7).

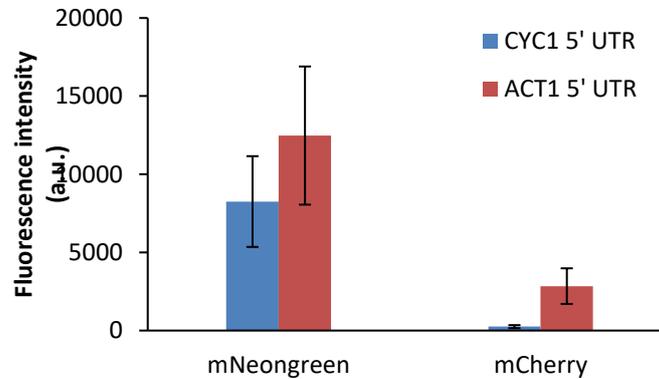


Figure 12: Impact of 5' UTRs on reporter expression. Flow cytometry analysis of four single-reporter strains harboring either mNeogreen (left) or mCherry (right) driven by fully induced (100 ng/ $\mu$ L) Tet promoters. Fluorescence intensities (raw data) are given in arbitrary units (a.u.). Error bars indicate standard deviation of ca. 10,000 measured cells.

The *ACT1* 5' UTR showed improved expression by ca. 50% for mNeogreen and a more than 10-fold increase in mCherry expression. The remarkably large difference in fold-change increase is possibly caused by interactions between the 5' UTR with the open reading frame. It has been shown before that the effect of 5' UTR sequences on expression levels is context-specific: The expression strength conferred by a 5' UTR for a fluorescence reporter gene showed only weak correlation with the expression obtained for the same 5' UTR with a different fluorescence reporter gene (79).

The fluorescence reporter genes in this study are all genomically integrated. To test the influence of the integration site and the relative genomic positioning of two fluorescence reporter genes on gene expression noise, we constructed dual-reporter strains harboring the reporter genes either in the same locus or in two different loci on separate chromosomes (Figure 13). For the single-locus approach we integrated a cassette of converging mNeogreen and mCherry genes into the *his3* locus, which is commonly used for integration. For the two-loci approach we used the same parental yeast strain but made two subsequent integrations. This time only mCherry was integrated into the *his3* locus while mNeogreen was integrated into another commonly used integration site, the *ura3* locus.

First, we assessed whether the two different loci conferred different gene expression noise by comparing the robust CVs of mNeogreen integrated into the two different loci. We found that both loci conferred to similar noise levels in the reporter gene expression. mCherry, expressed from the same locus in both strains, served as a control and as expected exhibited similar noise in both strains (Figure 13, bottom left panel).

## Results

Next, we calculated the intrinsic noise  $\eta_{int}^2$  and the extrinsic noise  $\eta_{ext}^2$  (see materials and methods, section 4.10.1) to estimate how much of the variation in the two strains was correlated and how much was uncorrelated between the reporters. The strains showed similar extrinsic noise but varied largely in their intrinsic noise with the single locus approach leading to a 4-fold reduction in uncorrelated variation (Figure 13, bottom right panel).

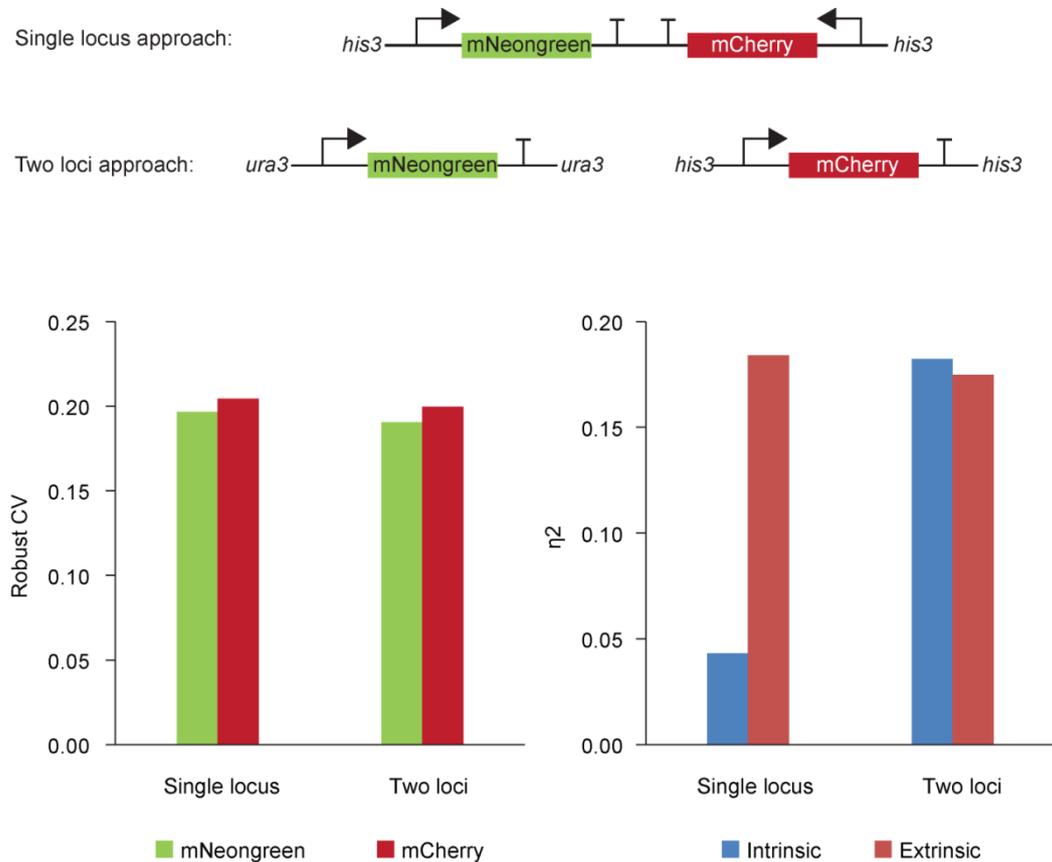


Figure 13: Effects of the integration loci on gene expression noise: Top: Two strategies to build a dual-reporter strain. Two fluorescence reporter genes, driven by identical TetO7 promoters and ADH1 terminators were either integrated into a single locus or into two different loci on separate chromosomes. Bottom left panel: Isolated analysis of the two reporter genes by plotting the robust CV of mNeogreen and mCherry for both strains. *ura3* and *his3* loci exhibit similar noise levels. Bottom right panel: Calculation of intrinsic and extrinsic noise for both strains. Single integration into *his3* locus reduces uncorrelated (intrinsic) noise 4-fold while correlated (extrinsic) noise remains unaffected. Noise calculation according to (2). Tet promoter induction with 100 ng/ $\mu$ L doxycycline. Calculations were done with raw, not normalized fluorescence intensity data.

These findings emphasize the impact of the relative genomic distance on uncorrelated noise between those genes. Our results hint that it is in fact the physical separation of the genes which creates intrinsic noise. It appears that the rare probabilistic events of mRNA production are correlated for genes in close proximity, possibly because the transcription of one gene increases the likelihood of the other gene being transcribed, since RNA polymerase is already present in the locus and/ or the chromatin structure is open.

## Results

We established a third fluorescence reporter to be used as an internal control for the overall gene expression capacity of a cell. As the originally obtained mTurquoise2 gene exhibited fluorescence levels close to auto-fluorescence, a codon-optimized mTurquoise2 sequence, supplied by Alexander Anders, was tested instead (Figure 14). Both strains - one harboring the original, one harboring the codon-optimized gene - were driven by identical GPD promoter and terminator sequences, integrated into the same locus of the same parental strain.

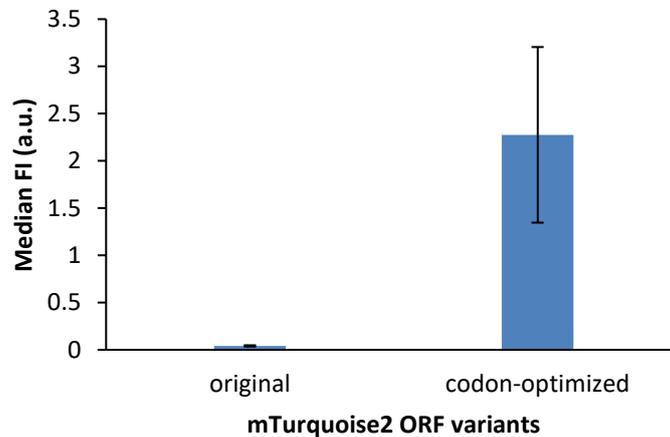


Figure 14: Codon optimization of mTurquoise2 open reading frame. Bars show mean of medians of fluorescence intensity normalized to the FSC-A signal. Error bars show mean of MADs of fluorescence intensity for three biological replicates.

The codon-optimized gene variant exhibited a more than 100-fold increased fluorescence compared to the original sequence. Codon optimization algorithms typically match the codon usage of the DNA sequence to the codon bias of the host organism's genome and exclude rare codons. To retain amino acid identity, usually only the third base in a triplet is changed ("wobble base"). Alignment of both sequences showed 69.4 % sequence identity, indicating that almost all wobble bases had been replaced. A sequence alignment is shown in section 6.2. The new sequence had a GC content of 34 % compared to 62 % in the original sequence. It is possible that in the case of the original mTurquoise2 sequence, one or several extremely rare codons prevented translation almost entirely. Only the codon-optimized version was used for subsequent experiments with mTurquoise2.

### 2.1.2 Optimizing measurement parameters

To ensure reproducibility, noise measurements (as most other experiments) should be performed with cells growing in exponential phase where the growth rate is constant. Given that yeast cultures are inoculated with the same number of cells, spectrophotometric measurements can be used to estimate the growth rate based on the optical density (OD). For high-throughput studies, such as those described in this thesis, OD measurements of every sample become experimentally challenging. We wanted to develop a high-throughput method to estimate growth rate based solely on flow cytometric data that would make additional time- and resource-consuming OD measurements unnecessary.

## Results

We prepared dilutions of a dense yeast culture and measured them by spectrophotometry and flow cytometry to assess whether the events per microliter of culture volume recorded by the flow cytometer correlated with the OD measurements (Figure 15). The number of cells per volume measured by flow cytometry correlated well with the OD over the entire linear range of the spectrophotometer. This shows that no additional OD measurements are necessary: If cultures are inoculated to the same initial concentration of cells, the growth rate at a given timepoint can be estimated from the concentration of cells calculated from the event per second and the flow rate of the cytometer.

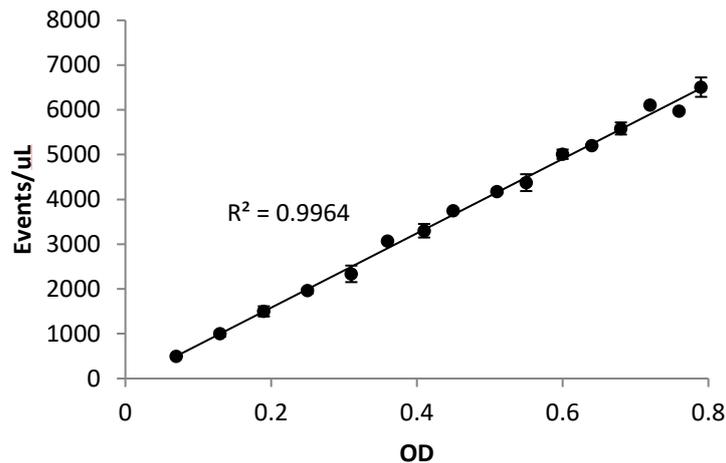


Figure 15: Correlation of events per microliter measured by flow cytometry with optical density measured with a spectrophotometer. Events/μL were calculated from the recorded time points of acquired events and the sample flow rate. A coefficient of determination  $R^2$  close to 1 indicates high correlation between OD and events/μL. Error bars indicate standard deviation from the mean of three individual flow cytometry measurements.

Using those findings, we performed time course flow cytometry experiments with cells inoculated from stationary overnight cultures. The *Saccharomyces cerevisiae* strains we used in our experiments typically produce two distinct populations that can be distinguished in the forward scatter area (FSC-A) and width (FSC-W) channels as shown in Figure 16. We call the cells producing a peak at lower FSC-W values the singlet population and the cells producing a peak at higher FSC-W values the doublet population. We observed that the ratio of singlets to doublets changed over the time course of the experiment.

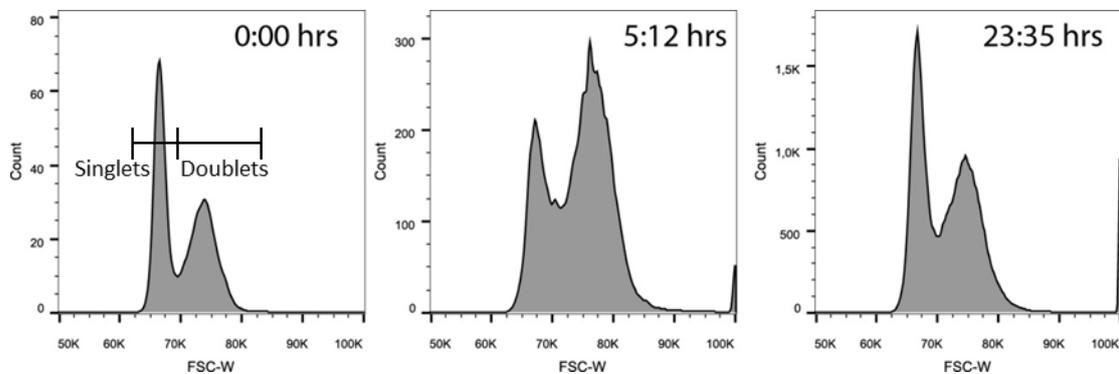


Figure 16: Two morphological distinct populations of yeast cells change their ratio during time course experiments. Example histograms of the FSC-W signal acquired at three different time points of the experiment. At inoculation and at the end point measurement the following day, the low-FSC-W population produces a bigger peak than 5:12 hours after inoculation from stationary phase.

## Results

To test the efficiency of the flow cytometer for the measurement of yeast growth kinetics we acquired the events per second and the FSC signals of a culture over a time course of 24 hours. The events per microliter calculated for different time points after inoculation yielded a highly reproducible growth curve and showed that growth kinetics can be measured by flow cytometry. Plotting the ratio of the doublets to all cells shows a maximum in mid-exponential phase, suggesting the doublet ratio is growth rate dependent.

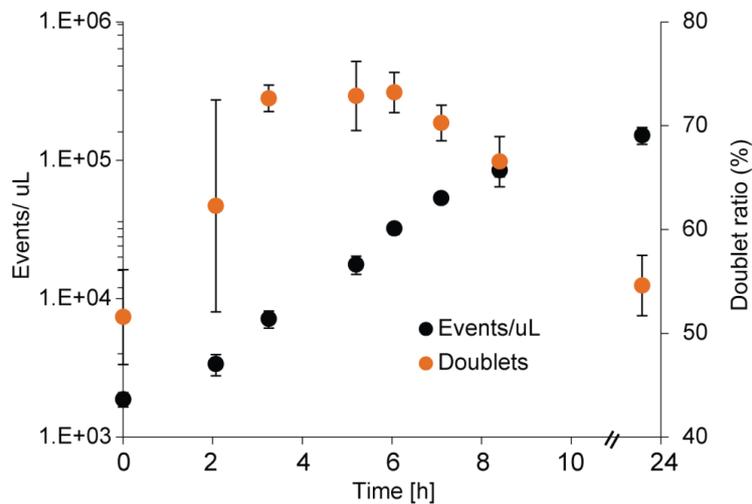


Figure 17: Growth kinetics measured by flow cytometry. Events per microliter measured at different time points after inoculation (black, left Y-axis). Ratio of cells in doublet gate is given as the percentage of total cells (orange, right Y-axis). Error bars indicate standard deviation of three biological replicates.

The high ratio of doublets in mid-exponential phase suggests that the two distinct peaks observed could be due to morphological changes during the cell cycle that occur when a daughter cell is formed during budding.

To test whether the doublet population indeed consists of cells that are attached to each other we exposed exponentially growing yeast cells to sonication and measured the ratio of doublets. The assumption was that late-stage buds could be detached from the mother cells by the ultrasound waves. Indeed, we observed a more than 50% drop of the doublet population (Figure 18) and a corresponding rise of the singlet population.

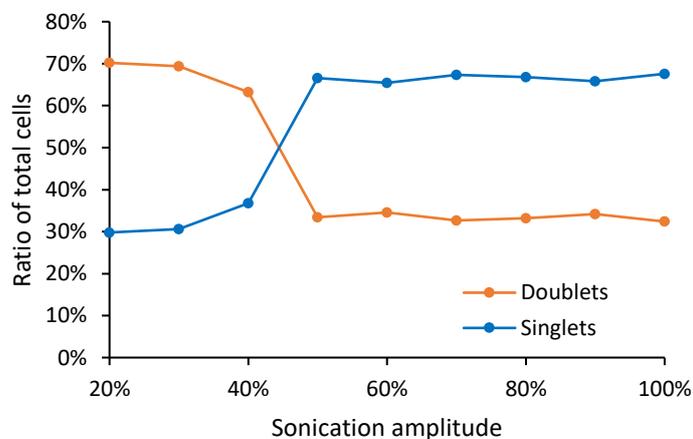


Figure 18: Dispersion of cell doublets. The ratio of doublets dropped sharply when cells were sonicated with increasing amplitude. Percentage of amplitude refers to an arbitrary scale from the manufacturer of the sonicator.

## Results

We assume that the doublet population is in fact a mixture of aggregated cells and budding cells. Aggregation would explain why even in stationary phase we observed a doublet population. Aggregated cells and late-stage buds can be dispersed whereas early buds are not yet fully separated from the mother cells, which is why they cannot be dispersed, which explains why we do not observe the doublet ratio dropping to zero.

To assess whether cells would recover the doublet characteristic, we dispersed an exponentially growing yeast culture by sonication with 60 % amplitude and measured the doublet ratio over a period of 60 minutes in 5-minute-intervals. Within the measurement time, the ratio of doublets tripled (Figure 19), indicating that sonicated cells continued to aggregate and/ or bud.

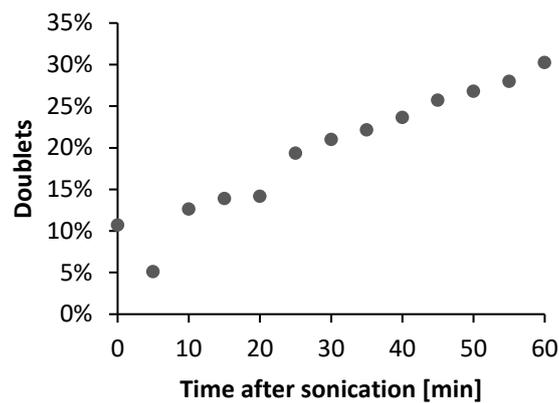


Figure 19: Recovery of doublets after sonication. Cells from a single tube were measured at different time points after sonication. During the course of the experiment, cells were kept at room temperature without shaking.

Sonication of yeast cultures is common practice to disperse cells prior to flow cytometric measurements. Here we show that, although it efficiently reduces the doublet population, sonication is not well suited to create more homogeneous populations for high-throughput experiments with up to 96 samples per measurement and runtimes of up to ca. one hour.

## Results

We further investigated whether gating of *Saccharomyces cerevisiae* populations according to their forward scatter signal could reduce noise. This would be the case e.g. if appropriate gating isolated cells of similar morphologies that corresponded to the same cell cycle stages, which correlate with gene expression rates, and therefore also noise (16). To assess to which extent gating of morphologically similar cells can be used to reduce variation, we analyzed a dual-reporter strain with fully induced mCherry and mNeogreen expression. Comparison of intrinsic and extrinsic noise of the total population with the gated population showed no difference in intrinsic noise but reduction of extrinsic noise by more than 80% for the gated singlets (Figure 20).

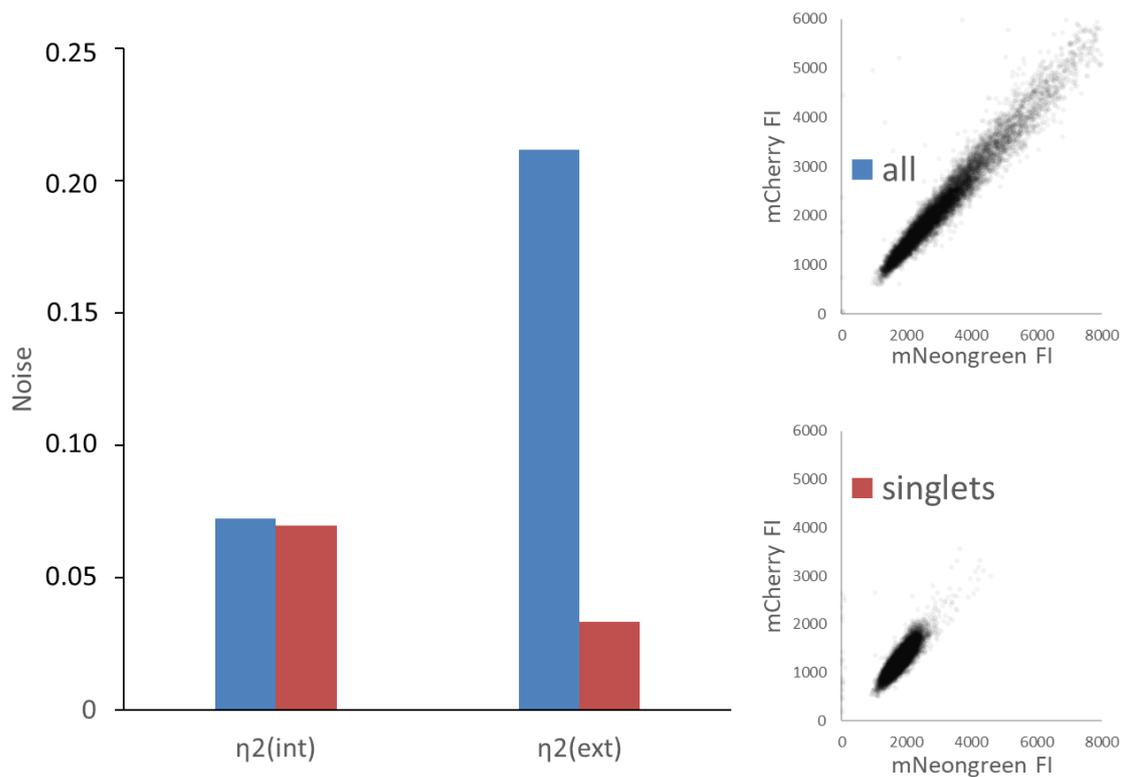


Figure 20: Gating of subpopulations reduces extrinsic noise. Left: Intrinsic and extrinsic noise for ungated (blue) and gated (red) population. Noise values were calculated according to (2). Right: Scatter plots of mCherry (X-axis) and mNeogreen (Y-axis) for ungated (top panel) and gated cells (bottom panel). Calculations based on raw, not normalized fluorescence intensities.

Gating of singlets shows a similar effect as the reduced gate size analysis introduced in section 1.3.2 and determined in subsequent results. Selecting a homogeneous subpopulation based on phenotypical parameters reduces extrinsic noise but has only marginal effect on intrinsic noise.

## Results

We further assessed whether nonspecific noise could be reduced by adjusting parameters of the flow cytometer. Fully induced mNeogreen fluorescence compared at 50 mW and 100 mW laser power, adjusted to the same median via the PMT voltages, resulted in similar CV values. In the tested parameter space, noise thus does not change with the laser power. The default laser power for the subsequent experiments was set to the default setting (100 mW to detect mNeogreen and mCherry, 75 mW to detect mTurquoise2).

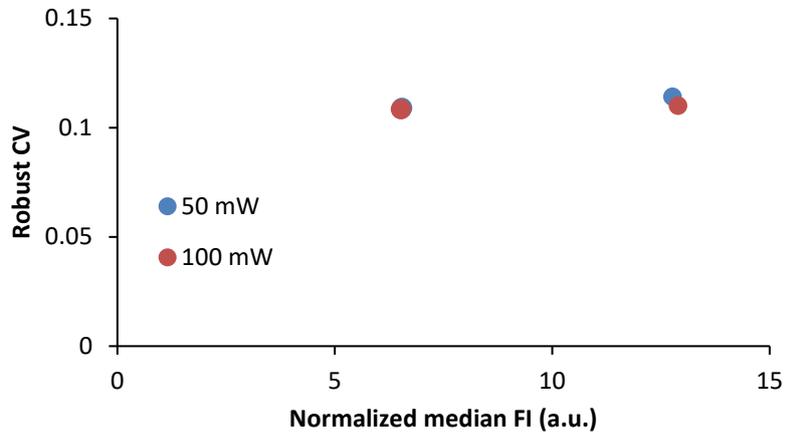


Figure 21: Noise analysis at 50 mW and 100 mW laser power for the measurement of mNeogreen. In a median expression range between 6 and 12 arbitrary units, laser powers of 50 or 100 mW do not change noise. PMT voltages were adjusted to allow similar median at low and high laser power. The robust CV was calculated from FSC-normalized mNeogreen fluorescence intensities.

To test the influence of the PMT voltages of the flow cytometer on the measured noise, we measured the fluorescence in a yeast strain with fully induced mNeogreen expression (100 ng/ $\mu$ L doxycycline). Prior to the measurement, we adjusted the voltages in the PMTs such that an uninduced control strain would give a low, medium, or high median fluorescence signal. The voltages used for the PMTs can be found in the appendix (section 6.1, Table 8). We found overall only small differences in measured noise, with the medium settings resulting in the lowest values, 10 % less than the low settings (Figure 22). Subsequent experiments were performed with the medium PMT settings.

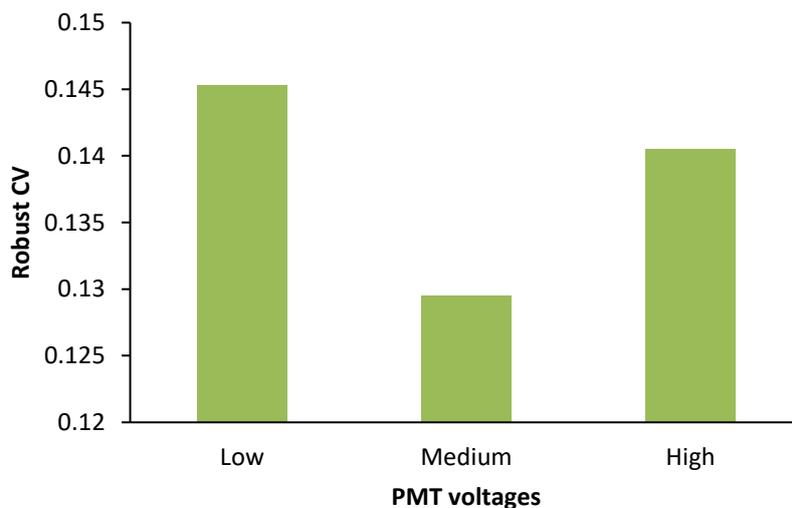


Figure 22: Noise-dependence on PMT voltages. Robust coefficients of variation are shown for mNeogreen fluorescence intensities normalized to the respective forward scatter signal. Medium PMT voltages result in slightly lower noise than high and low voltages.

## Results

The laser power and PMT tests indicated that changing instrument parameters did not affect the noise measurements to a big part, which suggests that technical noise introduced by the instrument was generally low. To estimate the lower limit of noise that could be achieved with flow cytometric measurements, we measured fluorescent beads of a defined, uniform size but different surface densities of a fluorophore, creating different fluorescence intensities. Figure 23 compares the calculated noise for a given median fluorescence to the theoretical minimum according to the noise/mean relationship discussed in section 1.3.2. The measurements were done for the blue laser (488 nm) that we used in other experiments to acquire mNeongreen fluorescence.

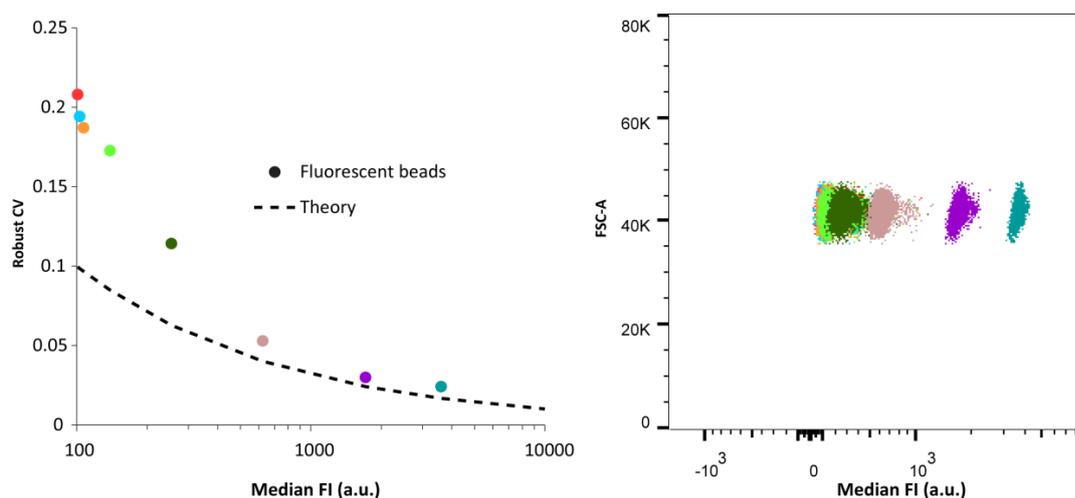


Figure 23: Left panel: Robust coefficient of variation of beads with different, distinct fluorescence intensities. Theoretically expected noise was calculated as the inverse of the square root of the median fluorescence (dashed line). Right panel: Scatterplot of the FSC signal from the beads against their fluorescence intensities. The colors from the dots in the left panel correspond to the colors of the populations in the scatterplot.

Uniform bead sizes at different fluorescence intensities were confirmed for excitation with the cytometer's blue laser. Above a median intensity of ca. 600 arbitrary units, the experimental rCV converges with the theoretical rCV, indicating that technical noise of the flow cytometer is low. At the highest intensity, the fluorescent beads exhibited a robust coefficient of variation around 0.025. The lowest rCV of a yeast sample presented in this thesis was about four times higher, illustrating that biological noise is by far the main contributor in the experiments shown. Below ca. 600 arbitrary units, the rCV values deviate from the theoretical calculation, indicating either the increased influence of technical noise or low, and thus, more stochastic coating of the beads with the fluorophore that we measured.

The scatterplot in Figure 23 shows high overlap of beads with supposedly different but low fluorescence intensities. This points towards another consideration for noise measurements, which is accurate subtraction of auto-fluorescence. Auto-fluorescence, is the part of the signal that does not correspond to the fluorescence reporter but other molecules in the cell. The recorded signal is the sum of the reporter signal and the auto-fluorescence. If the background is not subtracted, rCV calculations can yield artificially low values due to normalization of the MAD to an artificially high median. The left panel in Figure 24 shows an example of overlapping fluorescence intensity distributions for a population without mNeongreen gene (red) and a population with lowly expressed mNeongreen (green).

## Results

It is common practice to subtract the mean value of the auto-fluorescence  $\langle a \rangle$  from all values  $i$  of the signal  $s_i$ . For  $s_i < \langle a \rangle$ , this results in negative fluorescence intensities that have to be excluded or set to zero (Figure 24, middle panel). Instead, for each signal value  $s_i$ , we subtract a value  $a_i$  that was picked randomly from the auto-fluorescence probability distribution given that  $a_i \leq s_i$ . Thus, the smallest possible value after subtraction is 0 (Figure 24, right panel).

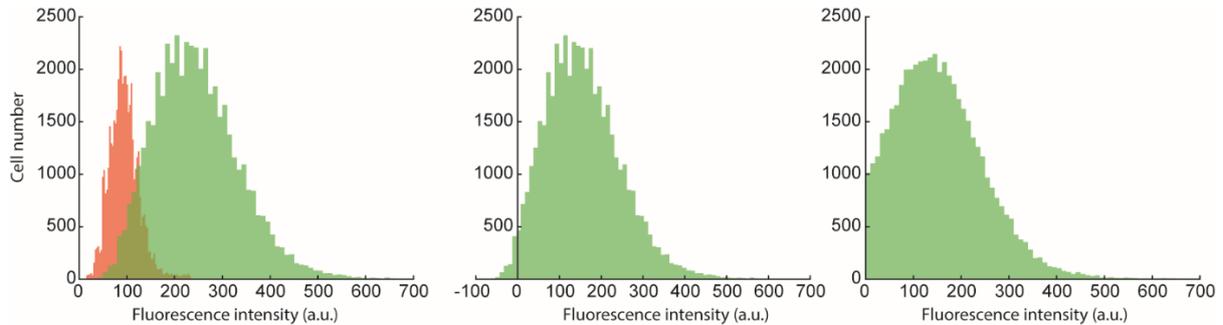


Figure 24: Comparison of two background subtraction strategies. Left panel: Fluorescence distribution of a yeast strain not harboring mNeogreen in red and a yeast strain with low mNeogreen induction in green. Middle panel: Same data subtracted by the median auto-fluorescence value. Right panel: Distribution after applying the subtraction method developed by Seán Murray.

This approach, developed in our lab by Seán Murray, takes into account that the auto-fluorescence of cells differs and that a cell's auto-fluorescence can never be bigger than the measured signal. We note the limitation of this method for very low signals that have high overlap with the auto-fluorescence distribution. Otherwise, however, this method results in a rapid and accurate auto-fluorescence subtraction. Here, accurate refers to agreement between the original signal and the reconstituted original signal obtained by convolving the corrected distribution with the auto-fluorescence distribution.

### 2.1.3 Optimizing analysis

Normalization of acquired fluorescence data is crucial because it can filter out nonspecific sources of variation, e.g. sources that are not directly related to gene expression. We tested two approaches for normalization of raw fluorescence data (Figure 25).

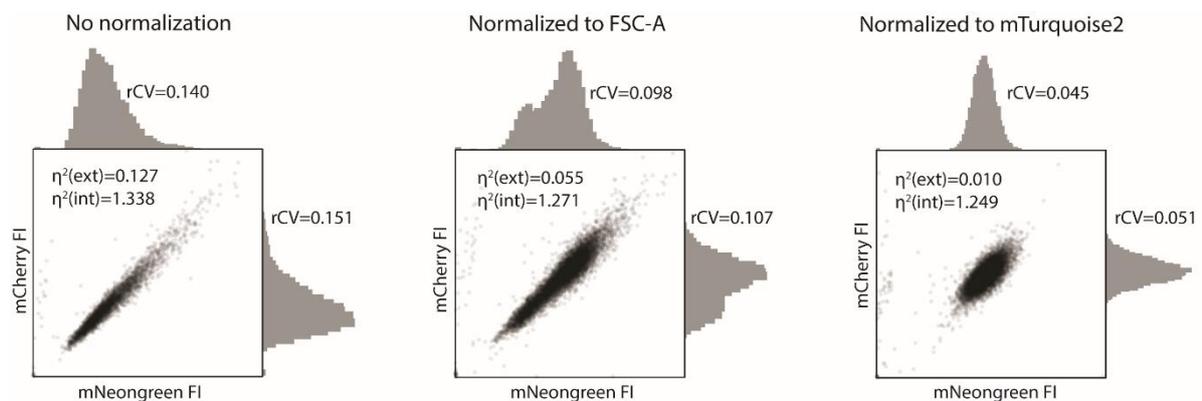


Figure 25: Comparison of normalization techniques for a triple-reporter strain. Scatter plot of fluorescence signals of mCherry and mNeogreen either as raw data (left), fluorescence normalized to the forward scatter signal (middle) or normalized to the constitutive mTurquoise2 expression control module (right). Values for intrinsic and extrinsic noise are given in the scatterplots. Adjacent histograms show the distributions of the individual reporters with their respective robust CVs. Axes in the middle and right scatterplot are scaled to approximately match the size of the left distribution. The ratio of the axes is identical in all three scatterplots and all are in linear scale. Fluorescence intensities in arbitrary units.

## Results

Dividing the raw data by the FSC signal (see section 1.2.2) reduced extrinsic noise by 57%. However, normalizing to an internal reporter for the expression capacity of the cell (i.e. the constitutively expressed mTurquoise2 gene) reduced extrinsic noise by 92%. The high absolute values for intrinsic noise cannot be explained conclusively, but the relative change is rather small (5 % and 7 % reduction for FSC-normalization and mTurquoise2-normalization, respectively), so that we conclude that normalization has only a negligible effect on intrinsic noise. Regarding the FSC-normalized distributions for mNeongreen and mCherry, we find that despite the overall reduced noise compared to the raw data, an additional source of heterogeneity seems to be introduced that creates two peaks. That might be explained by the distinct singlet and doublet distributions that we observe in the FSC channels but was not further investigated.

We performed a reduced gate size analysis on the raw and normalized datasets to assess how it compared to the dual reporter assay. As described in the introduction, the reduced gate size analysis takes advantage of the observation that phenotypic differences in the FSC/ SSC channels correspond to extrinsic noise. Cells in the full gate (radius 1) display the total noise. Reducing the gate size (here: up to radius 0.1) reduces the phenotypic heterogeneity and thus, the extrinsic noise. The difference in rCV between the full gate and the reduced gate radius corresponds to the extrinsic noise. The results are well in accordance with the results above: The raw data (blue lines in Figure 26) show a difference in rCV of 0.14 and 0.12 for mNeongreen and mCherry, respectively. The FSC-normalized data (yellow lines) show a difference of 0.05 and 0.07. The mTurquoise2-normalized data show virtually no extrinsic noise. The robust coefficients of variation converge to similar values for all datasets, which shows that the different normalization techniques only filter out extrinsic noise and lead to the only minor changes in intrinsic noise observed in Figure 25.

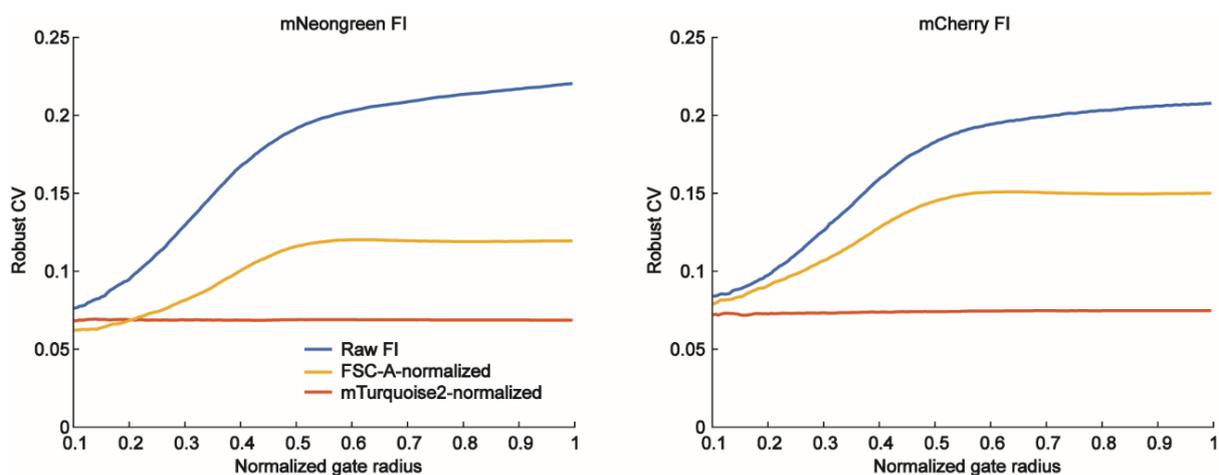


Figure 26: Reduced gate size analysis of differently normalized fluorescence data. X-axis indicates normalized gate radius with 1 corresponding to a gate that contains all measured cells. Y-axis shows rCVs for mNeongreen and mCherry (left and right panel, respectively) at different gate radii. Extrinsic noise is filtered out by reducing the gate radius. Raw (not normalized) data (blue) exhibits more extrinsic noise than cells normalized to FSC-A (yellow). Cells normalized to mTurquoise2 show no extrinsic noise.

The reduced gate size analysis allowed an analysis of intrinsic and extrinsic noise that was qualitatively similar to the analysis by the dual reporter assay. Since it requires only one reporter, reducing the effort for cloning and transformation, we constructed strains with only a mNeongreen reporter (plus the mTurquoise2 expression control) for subsequent experiments.

## Results

The results presented in this section illustrate that extrinsic noise in part depends on biological differences between cells but is also largely influenced by sample handling and data analysis. To measure effects of different expression modes on the noise in individual genes (see following sections), we aimed for minimal influence of variability that was not specific to those genes. Normalization to the general expression capacity of each individual cell proved to be very effective in reducing extrinsic noise and made gating of singlets or any other preselection of subsets of cells unnecessary.

## Results

### 2.2 Proof-of-concept of a noise tuner

To decouple gene expression noise from the mean expression of a reporter gene, we characterized the behavior of different inducible promoters and terminators. For transcription rate control we used the well-established Tet-ON system (Figure 7) with TetO7 promoters (seven Tet operator sites, hereafter called “Tet promoter”). Some of the experiments were performed using TetO2 promoters (two Tet operator sites), as stated specifically in the following text. Tet promoters were induced with doxycycline (dox). Transcript stability was controlled using inducible ribozymes that can be either activated or deactivated by the small molecule theophylline (theo). We used different concentrations of doxycycline and theophylline to change the modes of expression and calculated the gene expression noise.

#### 2.2.1 Characterization of ribozyme induction

We used previously published inducible ribozymes as tools to regulate gene expression via the mRNA degradation rate. Using these sequences, we wanted to enable expression control on the post-transcriptional level in addition to the transcriptional control mediated by the Tet promoter. By exerting control on both levels of gene expression, we aimed to build a synthetic system for noise tuning that employed the strategies introduced in section 1.4.

Inducible ribozyme sequences with different characteristics (49) were tested in the 3' UTRs of mNeogreen genes driven by Tet promoters. We inserted the synthetic minimal terminator T(Synth27) (47) 3' of each ribozyme sequence. Two ribozymes (L2b8-t47: “Ribo 1” and L2b8-a1-t41: “Ribo 2”) have their autocatalytic cleavage function inhibited when binding theophylline. One ribozyme (L2bOFF1-a14: “Ribo 3”) activates cleavage upon theophylline binding. The Ribo 1 strain showed 13-fold, Ribo 2 showed 14-fold induction; Ribo 3 showed a 2.7-fold decrease in expression (Figure 27).

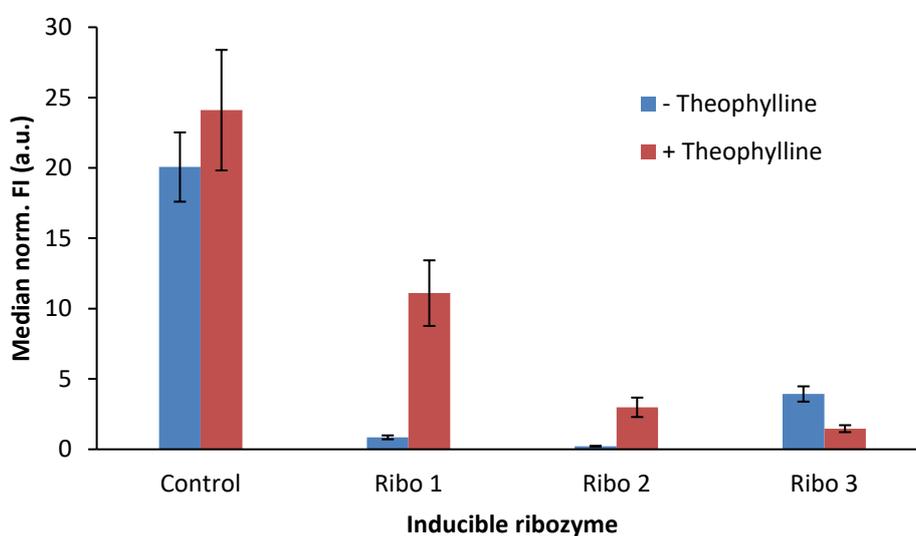


Figure 27: mNeogreen intensities for genes with different stabilized and destabilized ribozymes. Strains were induced with 64 ng/ $\mu$ L doxycycline and 0 mM (-) or 25.6 mM (+) theophylline overnight and in a day culture grown to mid-exponential phase. Fluorescence data was normalized to mTurquoise2 expression. Error bars indicate the MAD of ca. 10,000 cells.

The control construct contains a ribozyme with a scrambled core sequence that has no autocatalytic cleavage function. The reduced expression of the active ribozymes compared to the control indicates some cleavage activity even if the construct is turned off. Due to the low inducibility and high leakiness in the off-state, we did not continue working with Ribo 3 and focused on the two Theo-ON systems.

## Results

Histograms of the induction of those three strains and a strain harboring only the synthetic terminator sequence T(Synth27) are shown in Figure 28.

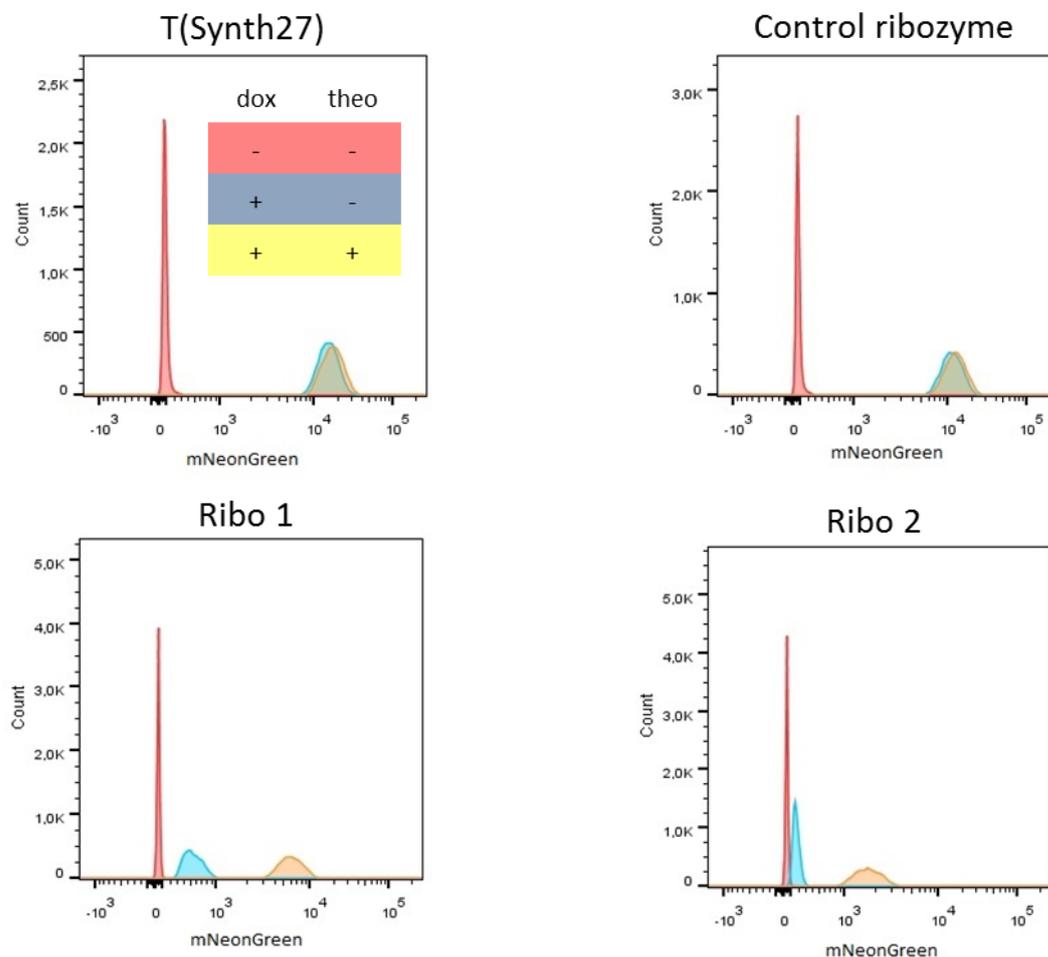


Figure 28: Fluorescence distributions of yeast strains harboring mNeongreen driven by a Tet promoter and different terminator regions. T(Synth27) is a synthetic minimal terminator that was cloned behind all ribozyme constructs. Constructs were induced with 0 (-) or 64 ng/ $\mu$ L (+) doxycycline and 0 mM (-) or 25.6 mM (+) theophylline overnight and in a day culture grown to exponential phase. Colors in the top left insert correspond to histogram colors.

The distributions of mNeongreen fluorescence intensities show, as above, similar fold changes for the two Theo-ON systems but differences in the strains with destabilized (- theo) mRNA. Whereas the distribution of the destabilized Ribo 1 construct was shaped like a low-mean on-state, the distribution of the destabilized Ribo 2 construct resembled the off-state (Figure 28, bottom left and right panels, respectively). This unexpected finding might be explained by the absolute activity ranges of the two ribozymes. While the stabilizing effect of theophylline is similar in both, Ribo 1 and Ribo 2, the lower absolute expression conferred by Ribo 2 results practically in an almost complete shutdown of expression.

### 2.2.2 Simultaneous control of transcription rate and mRNA degradation rate

The main goal of this project was to decouple gene expression noise from the mean expression by orthogonally tuning the transcription rate via doxycycline and the mRNA degradation rate via theophylline. To evaluate which combinations of doxycycline and theophylline concentrations resulted in similar expression rates, we induced mNeongreen expression over a range of different concentrations of both inducers and calculated the median and robust CV for each combination. Each

## Results

of the experiments described in the following section contains data of ca. 3 million cells and up to 64 different combinations of inducer concentrations.

To get the most informative view of our results, we integrated the information for median and noise as functions of the inducer concentrations into a single plot (see e.g. Figure 29 and Figure 31). Each square in those heatmaps represents an individual flow cytometric measurement of 50,000 yeast cells. The color of the squares indicates the robust CV that corresponds to the fluorescence distribution of that measurement. The lines on the heatmap are interpolations of the corresponding median expression levels of the distributions. Changes in color along a line (i.e. increase or decrease of the robust CV along a fixed median) would indicate decoupling between absolute expression and noise.

We did this experiment first for the control ribozyme to confirm that different theophylline concentrations did not markedly change expression levels at a given doxycycline concentration if the ribozyme was inactive (Figure 29).

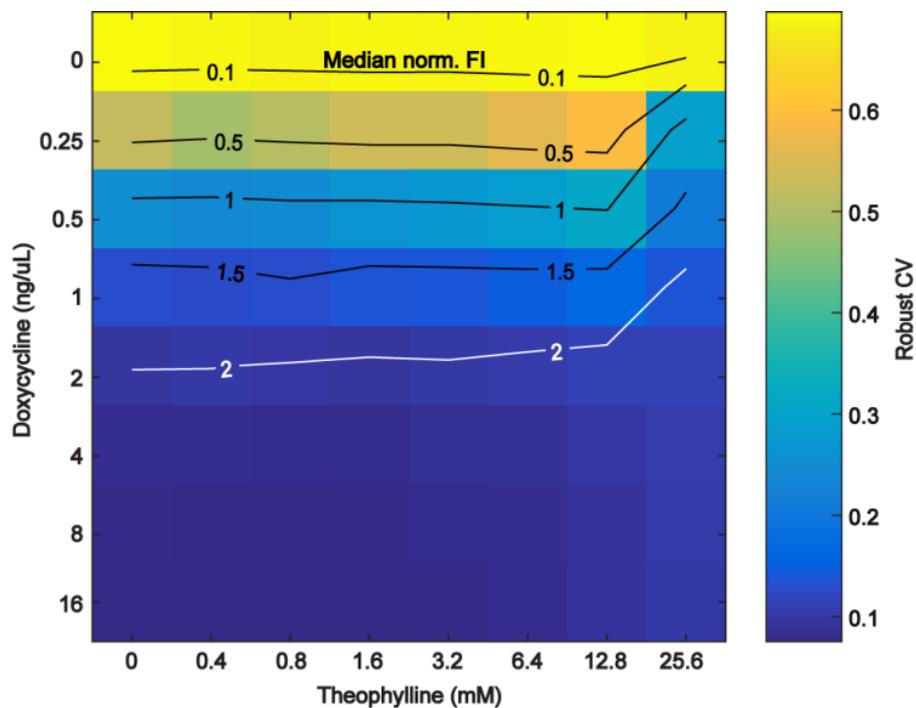


Figure 29: Noise heatmap for the ribozyme control strain induced with combinations of different dox/ theo concentrations. Colors indicate noise as the robust CV, given by the color scale on the right. The numbered lines indicate median mNeogreen expression normalized to the internal mTurquoise2 expression control. As expected for the theo-unresponsive control, both the noise and the median expression scale only with the dox concentration. High (25.6 mM) concentration of theo resulted in a reduced median expression.

Apart from the highest concentration used (25.6 mM), theophylline showed no effect on the mNeogreen expression. The fluorescence intensities scaled exclusively with the doxycycline levels. As expected, the noise scaled inversely with the gene expression, with the lowest noise at the highest fluorescence intensity. The fluorescence signal of the internal expression control, the constitutively expressed mTurquoise2 gene, was unaffected by the inducer concentration in terms of both, median expression (Supplementary Figure 5) and noise (Supplementary Figure 6).

To assess whether the decreased median expression at high theophylline concentrations was an artifact caused by a general impairment of the cellular physiology, we measured the growth of yeast cultures at different theophylline concentrations. For this we calculated the events recorded by the

## Results

flow cytometer per microliter of sample for all samples of a given theophylline concentration (Figure 30). We found that the events/  $\mu\text{L}$  were in a stable range between 4000 and 5000, which corresponds to an optical density between 0.5 and 0.6 (see Figure 15). The only exception was the highest theophylline concentration, where the cell density dropped below 3500 events/  $\mu\text{L}$ . This reduction in cell growth at 25.6 mM theophylline is significant ( $p < 0.02$  in two-sided t-tests for pairwise comparison with all other theophylline concentrations).

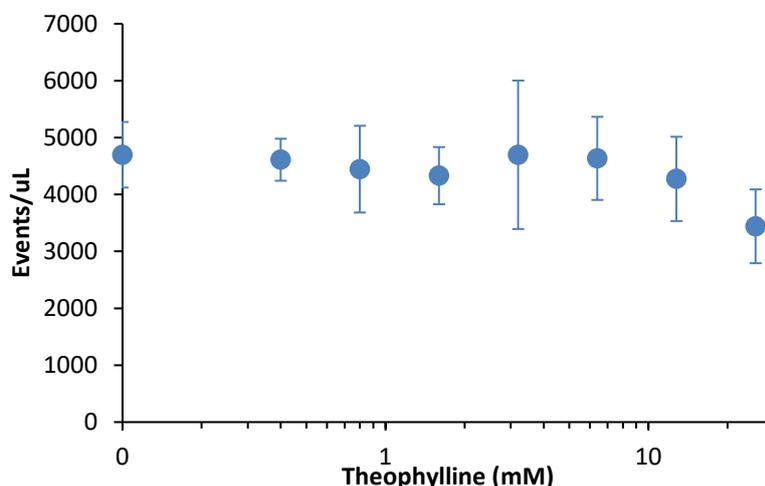


Figure 30: Effect of theophylline concentration on cell growth. LD medium with different concentrations of theophylline was inoculated with a yeast overnight culture to an initial OD of 0.05 and grown for 6 hours at 30 °C with shaking. Concentration of cells (events/  $\mu\text{L}$ ) of theophylline-induced cells was comparable to the uninduced sample (4000 to 5000) for all theophylline concentrations except the highest (25.6 mM). Error bars indicate the standard deviation of eight yeast cultures that were induced with different doxycycline concentrations from 0 to 16 ng/  $\mu\text{L}$  as shown in Figure 29.

We conclude that at 25.6 mM, where the concentration is close to its limit of solubility in aqueous solutions, theophylline appears to have a detrimental effect on cell growth. This concentration was therefore omitted from subsequent experiments.

After establishing appropriate experimental parameters and a comprehensive protocol for data visualization to measure expression levels and noise, we performed experiments similar to that in Figure 29 for the two active ribozymes Ribo 1 and Ribo 2 (Figure 31 A and B, respectively). As expected, the median fluorescence intensity was maximized at the highest concentrations of both inducer molecules where the transcription rate is high and the mRNA degradation rate is low, leading to high steady-state mRNA concentrations and consequently high protein concentrations.

Surprisingly, the robust CV measured with both systems scaled with the doxycycline concentration but the theophylline concentration had almost no effect on it. As a result, the rCV heatmaps in Figure 31 are qualitatively similar to the control heatmap in Figure 30. We note that the observed noise is reduced drastically over a small range of doxycycline induction. We observed this switching from a high-noise regime to a low-noise regime in the two constructs with active ribozymes as well as the control construct.

Accordingly, we found that when we follow a fixed median expression line in Figure 31, we can cross from the high noise regime to the low noise regime, indicating a decoupling of the mean and the noise. Taken together, these results show that we could successfully develop a desired noise tuning system.

## Results

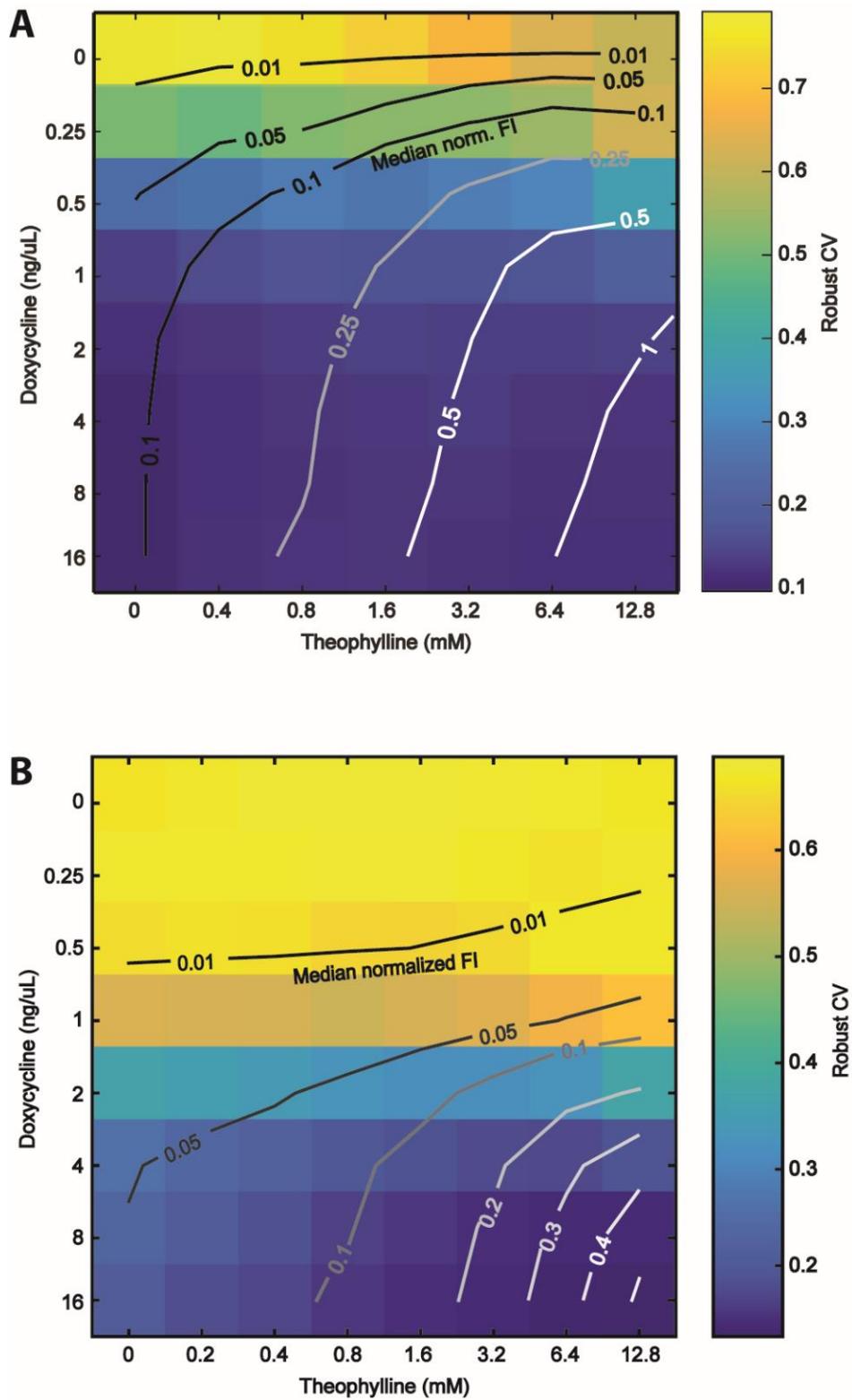


Figure 31: Noise heatmaps of mNeogreen expression with the Ribo 1 construct (A) and the Ribo 2 construct (B). Each square corresponds to a unique combination of theophylline (X-axis) and doxycycline concentrations (Y-axis). The colors of the squares correspond to the noise of the respective measurement, given as robust coefficient of variation of the mNeogreen expression. The rCV range is given by a color scale next to the heatmaps on the right. The lines on the heatmaps correspond to the median expression of mNeogreen normalized to the mTurquoise2 expression capacity control. Both heatmaps show that at an appropriate median expression, the noise measured from the  $P_{TetO7}$ -mNeogreen-Ribozyme constructs can be tuned up to 3-fold range.

## Results

It is noticeable that the low noise regime is reached at different doxycycline concentrations and thus, different transcription rates for the two noise tuning systems shown above. This shows that the mRNA degradation rate does have an impact on noise inasmuch as the different ribozymes confer different absolute expression levels (as shown in Figure 27) and the general mean-noise relationship discussed in Figure 5 does still apply. At higher absolute expression levels conferred by Ribo 1 we observe fewer noise as compared to Ribo 2. However, if similar median expression levels are compared, the noise is comparable in both systems.

From the results in Figure 31 we derive two strategies by which a particular median fluorescence intensity can be achieved: Strategy 1 employs low transcription rate (low doxycycline concentration) and low mRNA degradation rate (high theophylline concentration). Few mRNA molecules are present but they have a long half-life, so that each molecule is being translated very often. At low transcription rates, however, production of an mRNA molecule is a stochastic process and this noisiness on the transcript level is passed on to the protein level. Strategy 2 employs high transcription and mRNA degradation rates. Many mRNA molecules are produced but they have a short half-life, so that individual transcripts are rarely translated before being degraded. At high transcription rate, mRNA production is less stochastic so that less noise is being passed on to the protein level. The two strategies differ in their mRNA turnover and thus, in the energy required. Strategy 1 is noisier but energetically cheap, whereas strategy 2 is less noisy but requires more energy.

Both Theo-ON systems (Ribo 1 and Ribo 2) showed gene expression noise tuning capabilities for a fluorescence reporter gene in combination with an inducible TetO7 promoter. Due to the tighter off-state (see Figure 27), we proceeded with the Ribo 2 construct for all subsequent ribozyme experiments. In the following sections, constructs that regulate expression noise of a gene of interest via TetO7 promoter and Ribo 2 will be referred to as “noise tuner”.

To gain more insight from the expression responses of individual populations at different combinations of inducer concentrations, we selected distributions from the noise tuner measurements described in Figure 31 B with similar median expression but different noise for pairwise comparison. Figure 32 shows two examples; one at low and one at high expression. Populations that reach a given mean expression with higher doxycycline and lower theophylline concentration express the reporter gene more homogeneously with higher peaks and less spread towards lower or higher extreme values.

## Results

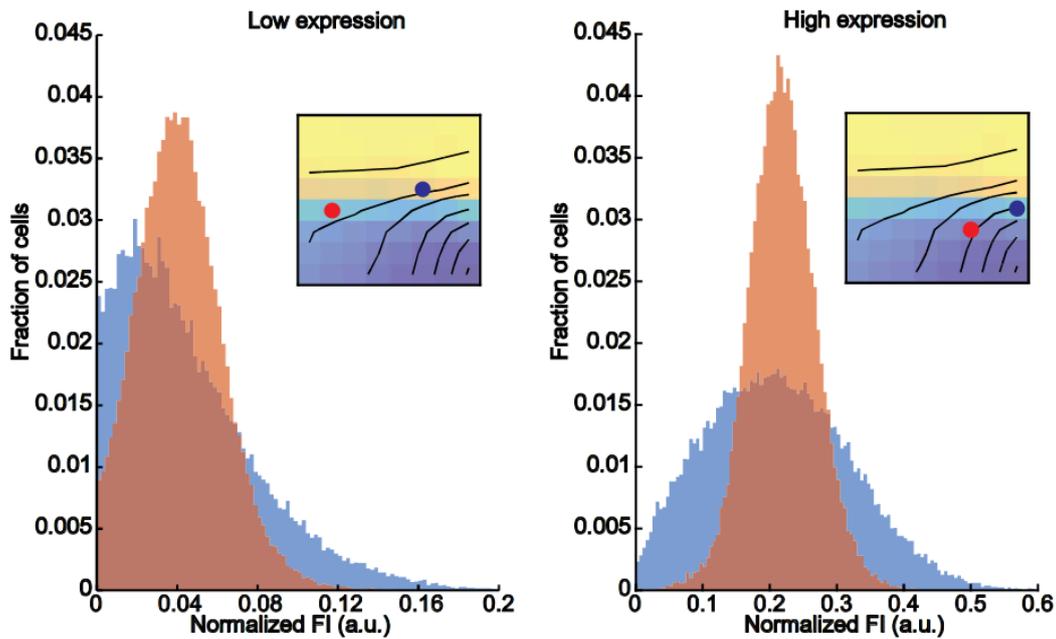


Figure 32: Example distributions of the mNeogreen noise tuner with similar median expression but different noise settings. Populations with higher dox and lower theo (red) exhibit lower expression heterogeneity than populations with lower dox and higher theo (blue). Insets indicate the compared populations from Figure 31 B. Colors of the dots in insets correspond to histogram colors.

To discern which sources of noise are affected by the two different expression strategies (high dox/ low theo and low dox/ high theo) we performed a reduced gate size analysis to estimate the contribution of extrinsic noise to the total observed noise (see section 1.3.2 for a description of the method). Figure 33 shows the reduced gate size analysis for the two examples in Figure 32. The extrinsic noise corresponds to the difference in robust CV between the full gate radius and the reduced gate radius (here: 0.1).

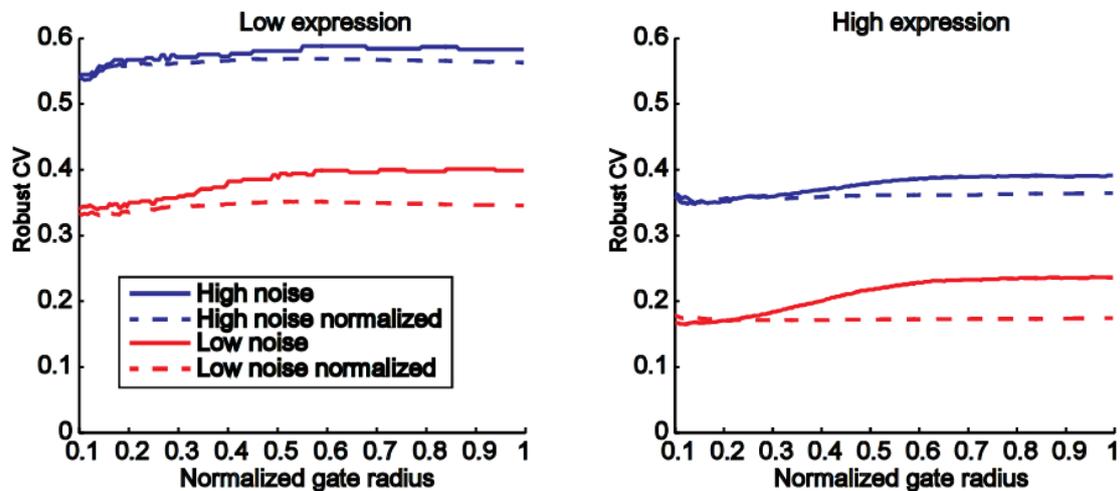


Figure 33: The noise tuner changes intrinsic variability. The radius of a circular gate around the median forward (FSC-A) and side scatter (SSC-A) is decreased to comprise only cells with similar FSC-A/ SSC-A values to filter out the extrinsic contribution of the observed noise. Background-subtracted raw data is shown with full lines; background-subtracted data normalized to the expression capacity control is shown with dashed lines. Samples shown here correspond to the examples in Figure 32.

The amount of extrinsic noise in the raw data was generally low. Data normalized to the expression capacity control (constitutive mTurquoise2 gene) showed no reduction of noise for reduced gate size, indicating again that extrinsic noise is filtered out efficiently. Assuming that technical noise is negligible

## Results

(see Figure 23), we note that the main factors that are changed between high and low noise settings of the noise tuner are intrinsic, as we would expect for gene-specific manipulation of expression.

To test the influence of the promoter strength on the noise tuning capabilities, we exchanged the TetO7 promoter for a TetO2 promoter with two instead of seven Tet operator sites for rtTA to bind. Noise did not change significantly when a given median expression was either achieved with high dox and low theo or *vice versa* (Figure 34).

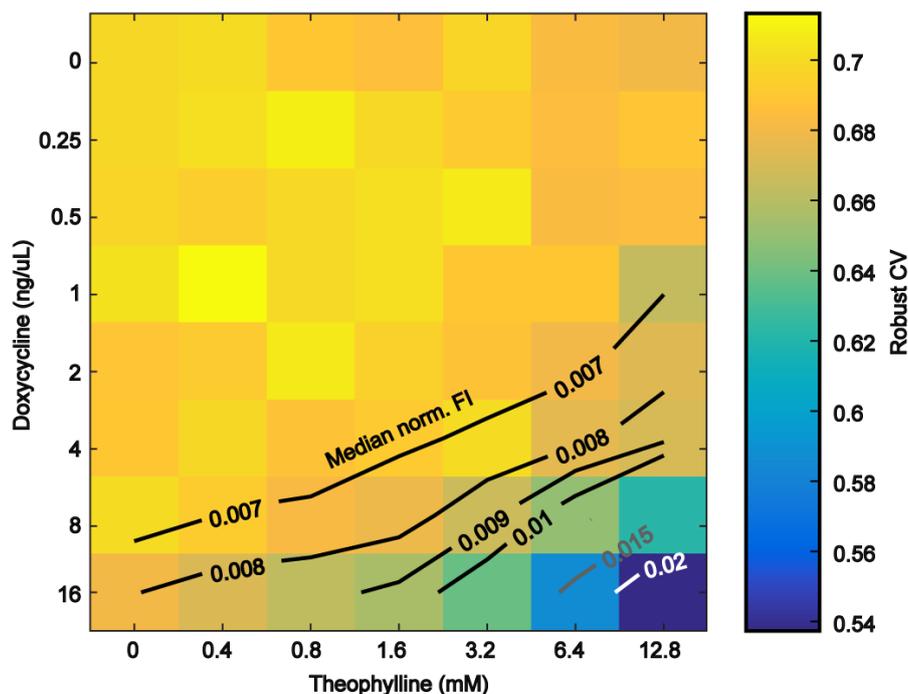


Figure 34: Median and rCV for 64 different combinations of theophylline and doxycycline concentrations of a yeast strain with *mNeongreen* driven by  $P_{TetO2}$  and Ribo 2. Colors indicate noise, given as robust coefficient of variation (color scale on the right) for individual populations measured by flow cytometry. Lines indicate identical median normalized fluorescence signal, interpolated from the measured data.

The  $P_{TetO2}$  variant exhibited overall much lower expression and did not leave the noisy regime of gene expression that we had observed previously at rCV values below ca. 0.4. This indicates that noise tuning by this strategy requires a sufficiently strong promoter.

### 2.2.3 Benchmarking of the noise tuner

The experiments described in Figure 31 and Figure 34 indicate that the noise tuning capabilities depend on the parameter space that is defined by the properties of the promoter and the terminator region used. The Tet-ON system is commonly used to induce expression, but less so the synthetic sequence comprising Ribo 2 and the T(synth27) terminator. To employ the noise tuner not only for the regulation of a fluorescence reporter gene, but also to tune noise in genes with a variety of functions and expression levels, we wanted to estimate whether the noise tuner operated in a similar mean and noise range as was conferred by native terminator regions.

We benchmarked the range of the noise tuner's expression and variability against other strains harboring the same  $P_{TetO7}$ -*mNeongreen* construct but different native yeast terminator sequences. For this comparison we selected terminators from the literature that had been reported to confer extremely high or low expression of a fluorescence reporter gene (45).

## Results

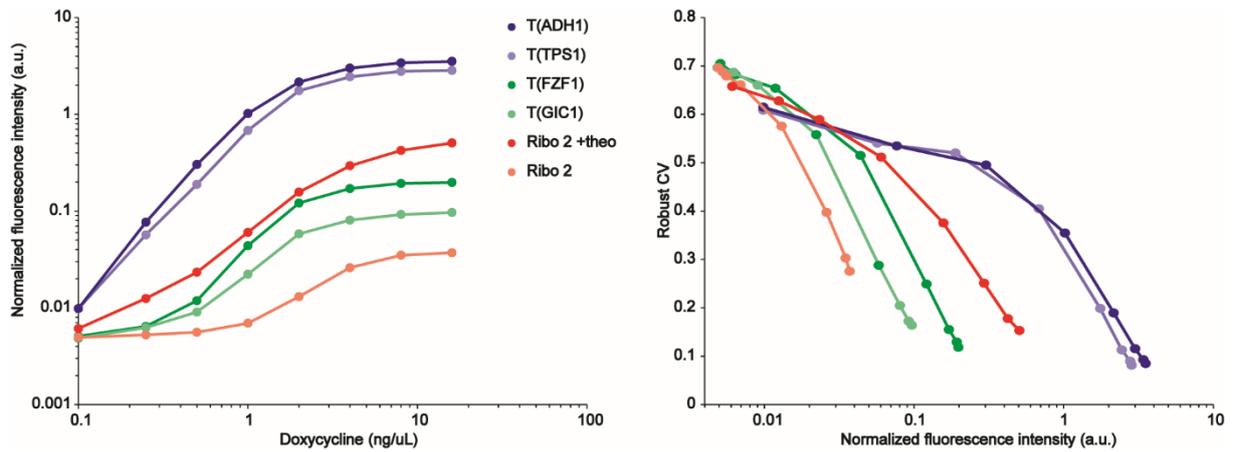


Figure 35: Comparison of the noise tuner with similar constructs harboring native terminators. The stabilized or destabilized Ribo 2 construct and strains with either strong native terminators (T(TPS1), T(ADH1)) or weak native terminators (T(FZF1), T(GIC1)) according to (45) were induced with different doxycycline concentrations. Dose-response curves of mNeogreen are shown in the left panel. Corresponding noise, given as the robust CV, is plotted against the median mNeogreen intensity in the right panel. Ribo 2 was induced with either 0 mM or 12.8 mM theophylline (“+ theo”).

The noise tuner showed expression in the lower range of expression achieved using native 3’ UTRs (Figure 35, left). The mean-noise relationship was comparable to the native constructs (Figure 35, right). The results show that changing 3’ UTRs of native sequences can have a similar effect on gene expression and noise as stabilizing or destabilizing the synthetic construct. At high expression, the native terminator regions conferred even less noise than the noise tuner. We further explored noise control via native control elements in section 3.4.

### 2.2.4 Stochastic simulation of the noise tuner

To further examine the mechanisms behind the observation that doxycycline and theophylline regulate mean expression and expression noise differently, we turned to the theoretical results of Shahrezaei and Swain on the distribution of protein copy numbers (80). Assuming that the protein lifetime is much larger than the mRNA lifetime, they found that the steady state number of molecules in the cell follows a negative binomial distribution (see materials and methods, section 4.11). Furthermore, they found that this distribution depends on only two parameters:  $a = \frac{v_0}{d_1}$ , where  $v_0$  is the mRNA transcription rate and  $d_1$  is the protein degradation rate; and  $b = \frac{v_1}{d_0}$ , where  $v_1$  is the translation rate and  $d_0$  is the mRNA degradation rate (see Figure 36).

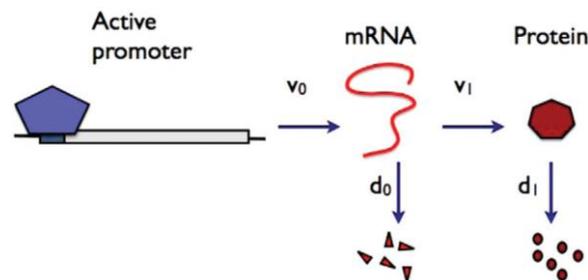


Figure 36: The basic probability rates of gene expression. Steady state mRNA concentration depends on transcription probability  $v_0$  and mRNA degradation probability  $d_0$  per unit time. Protein concentration depends on translation probability  $v_1$  and protein degradation probability  $d_1$  per unit time (adapted from 80).

The parameter  $b$  is known as the translational burst size – the average number of proteins translated per mRNA, while  $a$  can be interpreted in a similar way as the average number mRNA transcribed during

## Results

the lifetime of a single protein. These theoretical findings assume a two-stage model of gene expression, wherein the protein steady-state is governed by transcription and translation only, and does not factor in heterochromatin remodeling.

The mean of the distribution is

$$\mu = a \cdot b$$

while its coefficient of variation is

$$CV = \frac{1}{\sqrt{a}} \left( \sqrt{1 + \frac{1}{b}} \right).$$

We note that if the burst size  $b \gg 1$  then the term in brackets in above equation converges towards 1. In that case, the CV depends only on the parameter  $a$ , i.e. it does depend on the translation rate or the mRNA degradation rate. Large  $b$  implies that many proteins are produced from few mRNAs which is a likely assumption (see e.g. 81 and 82).

As the independence of the CV of the burst size was not investigated by the authors in (80), we decided to perform stochastic simulations to confirm this result and determine if we could reproduce the experimental observations in a biologically reasonable parameter space. We found that simulations of the two-stage model were in well accordance with the experimental results. While the mean protein number was maximized when both,  $a$  and  $b$ , were high, the CV scaled with  $a$  and only weakly with  $b$  (Figure 37).

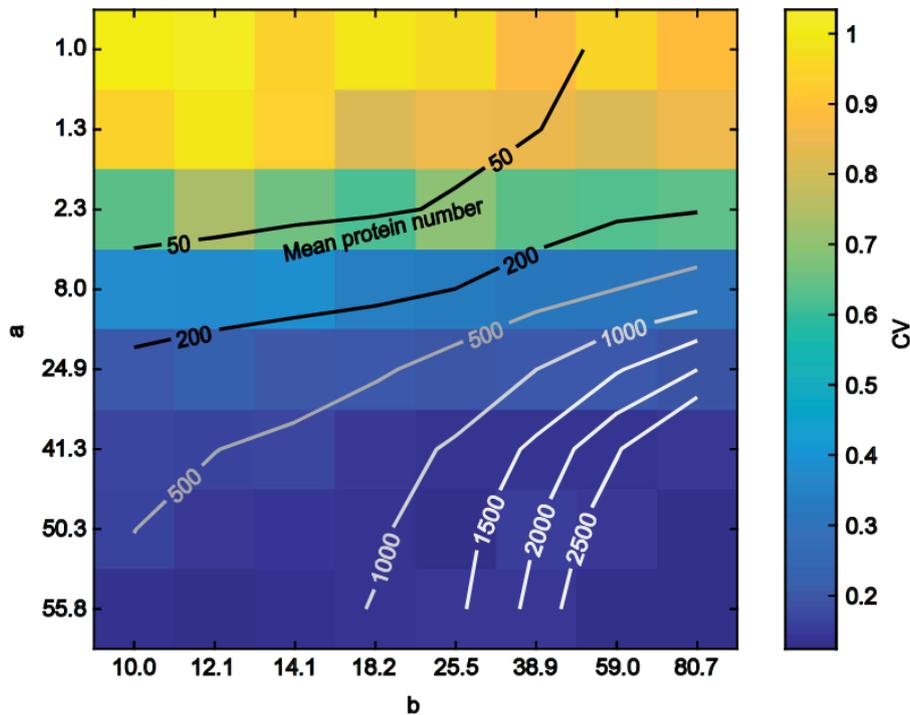


Figure 37: Stochastic modeling of mean and CV of steady state protein number as function of  $a$  and  $b$ . Lines indicate same protein number, colors indicate coefficient of variation (color scale on the right). Scaling of the  $a$  and  $b$  axes is chosen to correspond to the dose-response of doxycycline and theophylline, respectively. See methods for parameter values.

This observation is consistent with the experimental data (Figure 31 B): The doxycycline-induced promoter activity enters into  $a$  via the mRNA transcription rate  $\nu_0$ , while theophylline-induced mRNA

## Results

stability enters into  $b$  via the mRNA degradation rate  $d_0$ . Thus, increasing the transcription rate increases  $a$  while decreasing the mRNA degradation rate increases  $b$ .

Summing up this results section, we have shown experimental evidence for the successful construction of a gene expression noise tuner that employs the orthogonal control of the transcription rate and the mRNA degradation rate to decouple the median expression from the expression variation. The noise tuner operates at the lower bound of expression conferred by native terminator regions, in a parameter space where noise mainly stems from stochastic promoter activity and the transcript stability is changed to adjust the median expression level.

## Results

### 2.3 Noise tuning in a signaling pathway

Genes and the proteins they encode do not act in isolation but interact in complex networks. Depending on the network topology, the gene function, protein number and other factors, gene expression noise of individual components can have a smaller or bigger impact on the network's robustness and its reliable function. We used the noise tuner to investigate effects of noise for five different genes of the yeast mating pathway, as shown schematically in Figure 38. This analysis effectively probes noise sensitivity / robustness of the mating pathway.

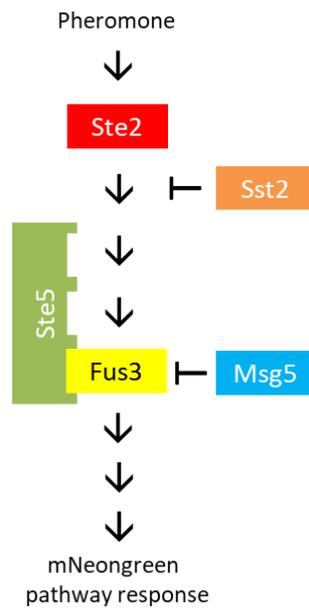


Figure 38: Overview of the pathway genes whose expression were controlled by the noise tuning system. A detailed description of the yeast mating pathway is given in section 1.6.3.

Similar to the proof-of-concept noise tuner described in the previous section, the experimental procedures included the orthogonal expression control of pathway genes with doxycycline and theophylline. Again, we aimed to set the gene expression noise at a given expression by using different transcription rates and mRNA degradation rates. We constructed yeast strains that had the native promoter and terminator of a pathway gene replaced with our inducible elements in the native gene locus and induced them with a range of doxycycline concentrations with either destabilized (- theo) or stabilized (+ theo) mRNAs. The promoter replacements had been done by Alexander Anders for previous experiments.

The following results show flow cytometric analyses of normalized mNeogreen measurements, similar to the results presented in the previous section. A notable experimental and conceptual difference, however, is that the mNeogreen reporter gene was placed under the control of the pathway response promoter P(Fus1). The measured mNeogreen fluorescence intensity, thus, reflects pathway activity ("pathway reporter", see Figure 10 in the introduction). The strains also contained the constitutively expressed mTurquoise2 gene to filter out extrinsic noise. We expected that induced changes in our strains which resulted in differences in the noise of the pathway reporter reflected the intrinsic gene expression noise plus the intrinsic pathway noise.

In order to measure fluorescence from the pathway reporter, we had to stimulate the mating pathway. We induced a mating response by incubating haploid MATa cells with purified  $\alpha$  pheromone for three

## Results

hours, choosing pheromone concentrations in the dynamic response range. We included a wildtype (parental) strain in the measurements to estimate the mean expression and noise under native control for a given pheromone concentration (Figure 39).

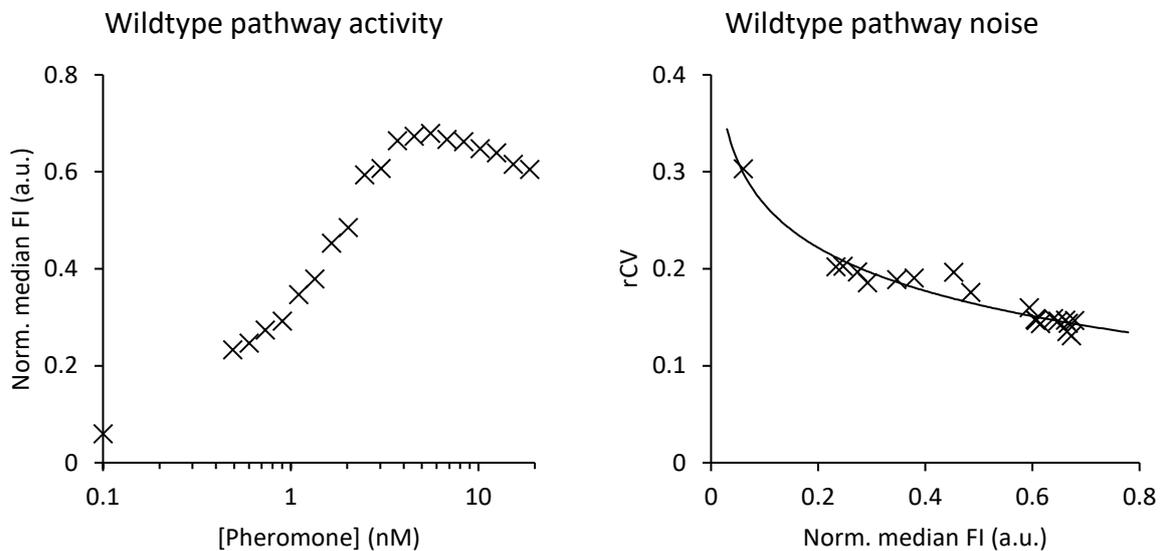


Figure 39: Pathway activity (left panel) and pathway noise (right panel) three hours after stimulation of a wildtype yeast strain with different pheromone concentrations. We define the pathway activity as the fluorescence measured from an mNeongreen reporter gene that is controlled by a pathway response promoter (see main text). We define pathway noise as the robust coefficient of variation measured from the pathway reporter. Fluorescence values are normalized to an internal constitutively expressed mTurquoise2 expression capacity control.

The wildtype measurements are intended to serve as a general orientation for the expected ranges of the pathway activity and the noise at a given pheromone concentration. The primary focus was, however, the comparison of different doxycycline and theophylline levels in the noise tuning strains. From the results we could make primarily two types of observations for each tested gene:

- 1) We could assess for which genes the noise tuner can be employed to change the pathway activity. If changes in gene expression via doxycycline and theophylline are reflected by changes in the pathway activity, that shows that the noise tuner works in this context.
- 2) We could observe whether the noise tuner, when applied to individual genes, can change pathway noise. If changes in pathway noise occur that correspond to the expected changes in expression noise of the targeted gene, that shows that the induced noise propagates through the pathway.

The following sections sum up these two kinds of observations for the five mating pathway genes shown in Figure 38. First, the results will be shown for the pathway receptor Ste2, scaffolding protein Ste5 and the MAPK Fus3. Then more detailed results will be shown for the negative pathway regulators Msg5 and Sst2.

## Results

### 2.3.1 Ste2

Ste2 is the seven-transmembrane domain G protein coupled receptor of MATa cells that binds  $\alpha$  pheromone. Ste2 is itself upregulated by the pathway (83). Signaling via Ste2 is terminated by ligand-induced receptor endocytosis (84). We induced the expression of *STE2* via doxycycline and theophylline and stimulated the pathway with 1.39 nM pheromone. As expected, the pathway showed increasing activity at higher concentrations of doxycycline. Depending on the doxycycline concentration, adding theophylline increased pathway activity between 11 % and 25 % (Figure 40, left panel). At a given median fluorescence, yeast cells without theophylline exhibited slightly lower pathway noise (Figure 40, right panel).

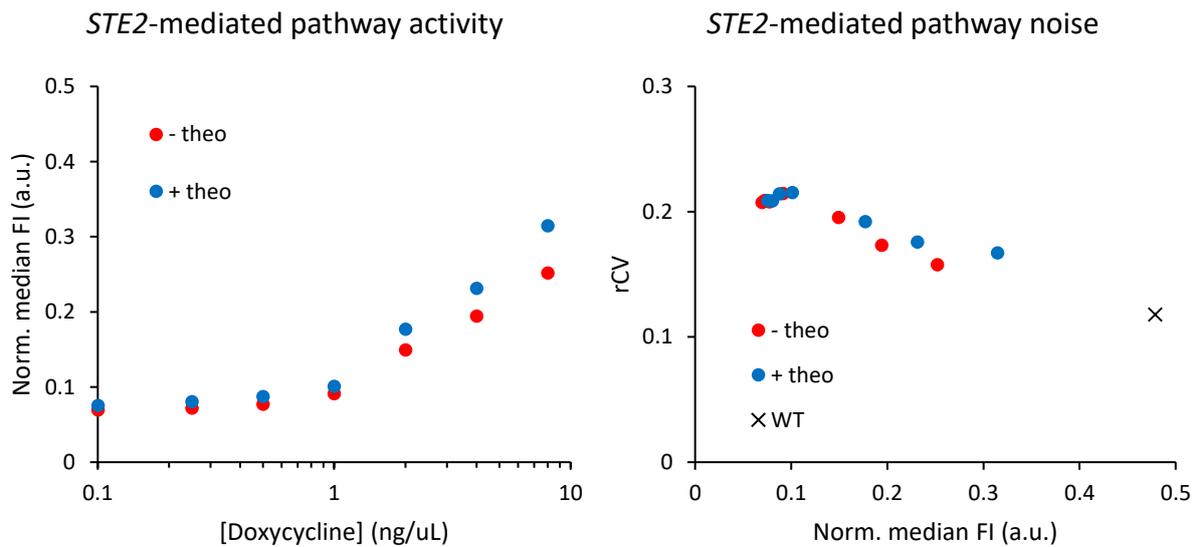


Figure 40: Pathway activity for different *STE2* regulation conditions. *STE2* was induced over a range from 0 to 8 ng/ $\mu$ L doxycycline and either without theophylline (red dots) or with 10 mM theophylline (blue dots). Left panel: Doxycycline dose and corresponding mNeogreen pathway reporter response. A slightly higher pathway response can be achieved when theophylline is added. Right panel: Pathway noise, given as rCV, at different median expression levels. Median and rCV of the wildtype are indicated with a black X.

The curve progression in the left panel in Figure 40 indicates that higher dox concentrations could further increase pathway activity. As Supplementary Figure 8 shows, *STE2* is expressed at very high transcription rates.

## Results

### 2.3.2 Ste5

*STE5* encodes a scaffold protein for the MAPK cascade to facilitate efficient signaling. We stimulated a yeast strain with *STE5* controlled by our noise tuner with 1.39 nM pheromone and measured the pathway activity for different doxycycline concentrations, with or without theophylline. The left panel in Figure 41 shows pathway induction over the entire dynamic range, mediated by the induced expression of *STE5*. Adding theophylline resulted in an increase in pathway activity to up to 2.5-fold over the activity with destabilized *STE5* transcripts. Absolute differences in fluorescence output were highest in the dynamic range of the pathway response.

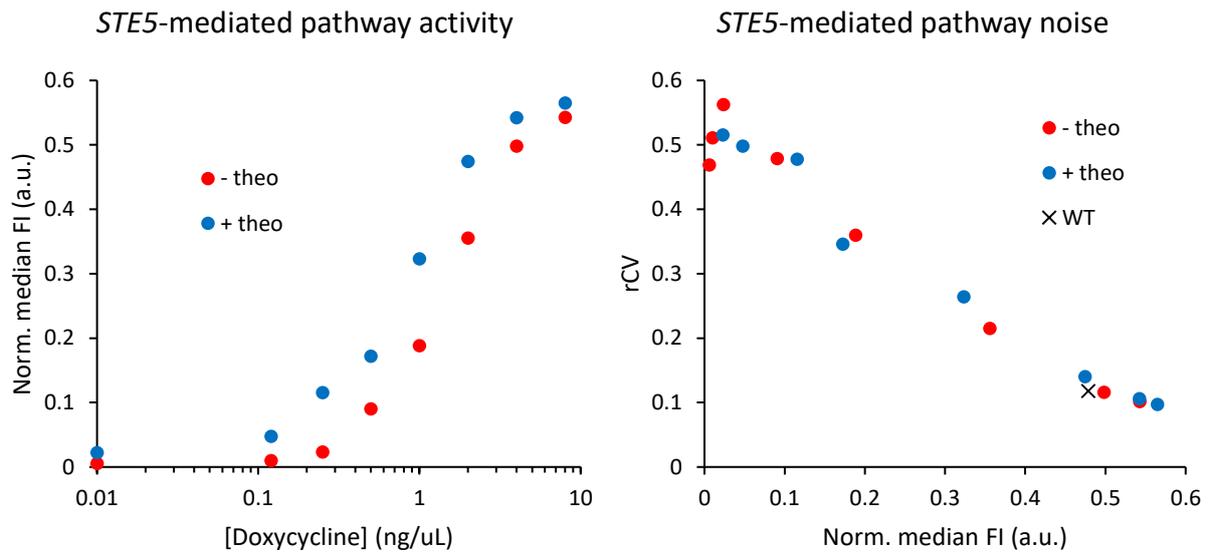


Figure 41: Pathway activity and pathway noise for the *STE5*-noise tuner. Measurement parameters and description as in Figure 40.

Based on the observations made for the proof-of-concept noise tuner (Figure 31 B), the critical transcription rates for the transition between the high and the low noise regime are achieved with doxycycline concentrations between approximately 1 and 4 ng/μL. This appears to be within the dynamic range of the pathway activity when *STE5* expression is controlled. In this range, however, the pathway noise was similar for yeast cells treated with theophylline and untreated cells (Figure 41, right panel), indicating that noise differences were not propagated through the pathway. At wildtype expression, *STE5*-regulation resulted in similar noise as the wildtype.

## Results

### 2.3.3 Fus3

Fus3 is the MAPK of the yeast mating pathway. As for *STE5*, *FUS3* expression was induced by increasing the transcription rate via doxycycline and this was mirrored by a corresponding increase in pathway activity in cells stimulated with 1.39 nM pheromone. Theophylline regulation of *FUS3* showed only a marginal increase of pathway activity in the dynamic range. We suspect that this could be due to a limited ability of theophylline to stabilize the transcript. This could be the case if e.g. the *FUS3* coding sequence contains elements that destabilize the mRNA. Doxycycline concentrations above 2 ng/  $\mu$ L led to reduced pathway activity of the theophylline-treated cells but not the untreated cells (Figure 42, left panel).

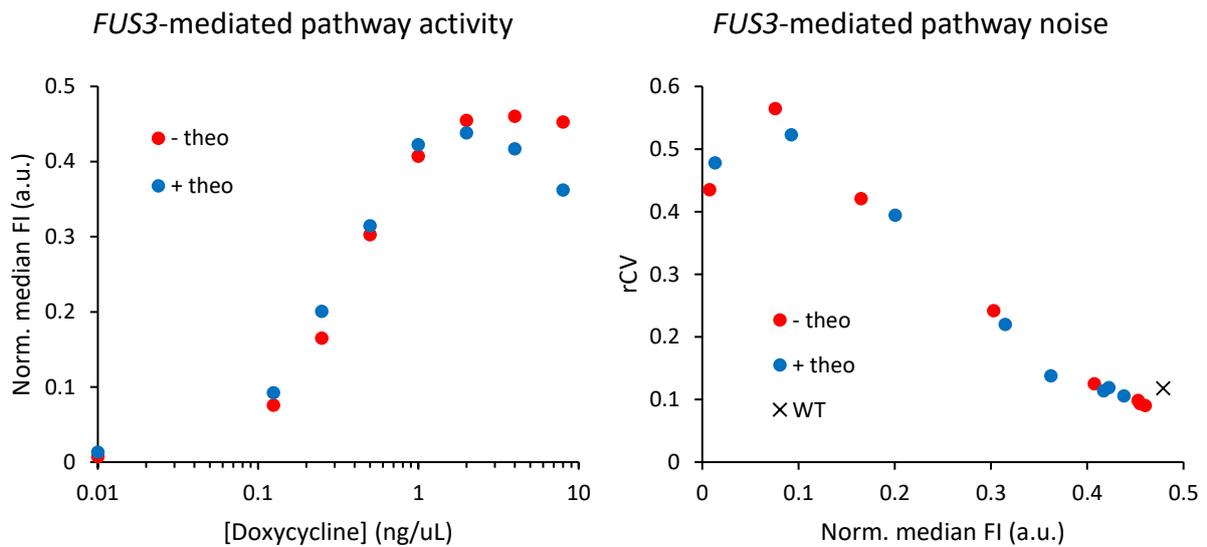


Figure 42: Pathway activity and pathway noise for a yeast strain with *FUS3* noise tuner. Measurement parameters and description as in Figure 40.

Theophylline also showed no significant influence on the mean-noise relationship of the fluorescence signals from the pathway reporter (Figure 42, right panel). The strain with our noise tuner reaches rCV values similar to the wildtype at comparable median expression.

## Results

### 2.3.4 *Msg5*

*MSG5* encodes a downstream negative feedback regulator of the pathway. *Msg5* directly interacts with the MAPK *Fus3* and dephosphorylates it, resulting in lower pathway activity at increased *MSG5* expression. The pathway activity was further reduced by up to 50 % when the growth medium contained theophylline, indicating that *MSG5* transcripts were successfully stabilized and *Msg5* protein levels were increased (Figure 43, left panel).

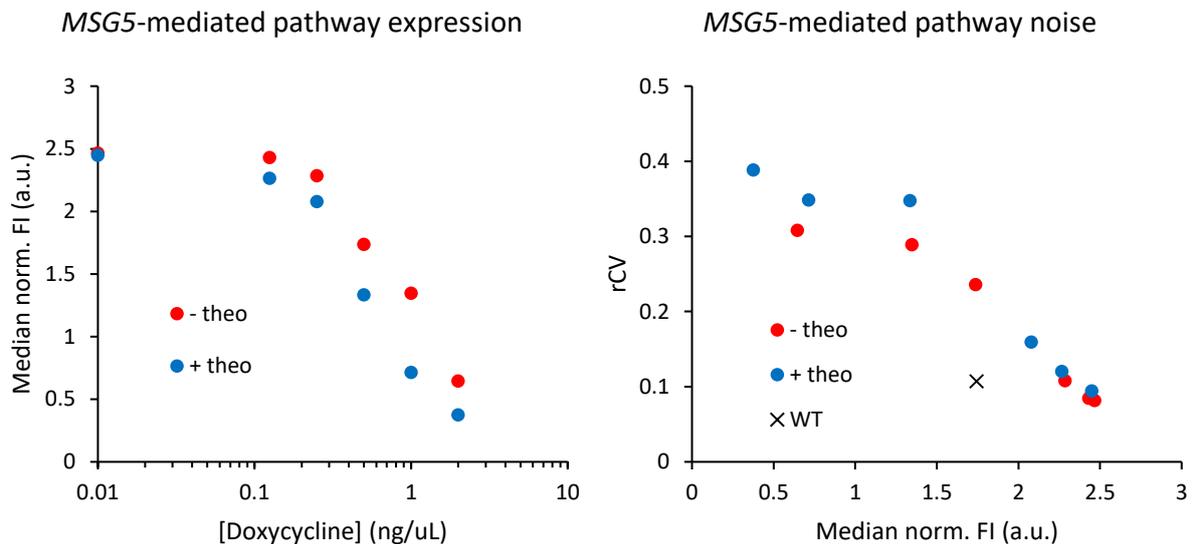


Figure 43: Pathway activity and pathway noise for the *MSG5*-noise tuner. *MSG5* was induced over a range from 0 to 2 ng/ $\mu$ L doxycycline and either without theophylline (red dots) or with 12.8 mM theophylline (blue dots). Cells were stimulated with 1.39 nM pheromone. Left panel: Doxycycline dose and corresponding mNeogreen pathway reporter response. Right panel: Pathway noise, given as robust CV, for corresponding median mNeogreen expression. Black X denotes wildtype expression and noise. These and the upcoming measurements in this section were made with reduced PMT voltage for mTurquoise2 detection, hence the increased normalized fluorescence compared to the previous figures.

The *MSG5* noise tuner strain showed significantly higher noise than the wildtype (Figure 43, right panel). Analyzing the transcription rate and mRNA degradation rate of *MSG5* from published data, we found that both rates were at the upper limit that was observed for yeast genes in the respective range of steady-state mRNA levels (Supplementary Figure 8). This indicates that the native *MSG5* might already employ the low noise expression strategy that we described for our noise tuner in section 2.2.

The right panel in Figure 43 shows pathway noise differences between cells with and without theophylline at high *MSG5* expression/ low pathway activity. These results prompted us to perform further experiments to assess whether noise in *MSG5* expression would impact the pathway response towards pheromone. *Saccharomyces cerevisiae* cells adjust their mating behavior (mating pathway activity and cell morphology) in accordance to the pheromone concentration that the cells sense (see Figure 9), which might be affected by noisy *MSG5* expression.

To evaluate this, we first had to find doxycycline/ theophylline combinations that produced similar median expression levels in the *MSG5* noise tuner strain. From further experiments (data not shown) we concluded that 4 ng/ $\mu$ L dox without theo and 2 ng/ $\mu$ L dox with 12.8 mM theo showed appropriate median fluorescence intensities. We refer to them as low- and high-noise *MSG5* settings, respectively.

## Results

Cells set to low or high *MSG5* noise were stimulated with a range of pheromone concentrations to induce the full range of pheromone responses from sub-shmooing to full stimulation.

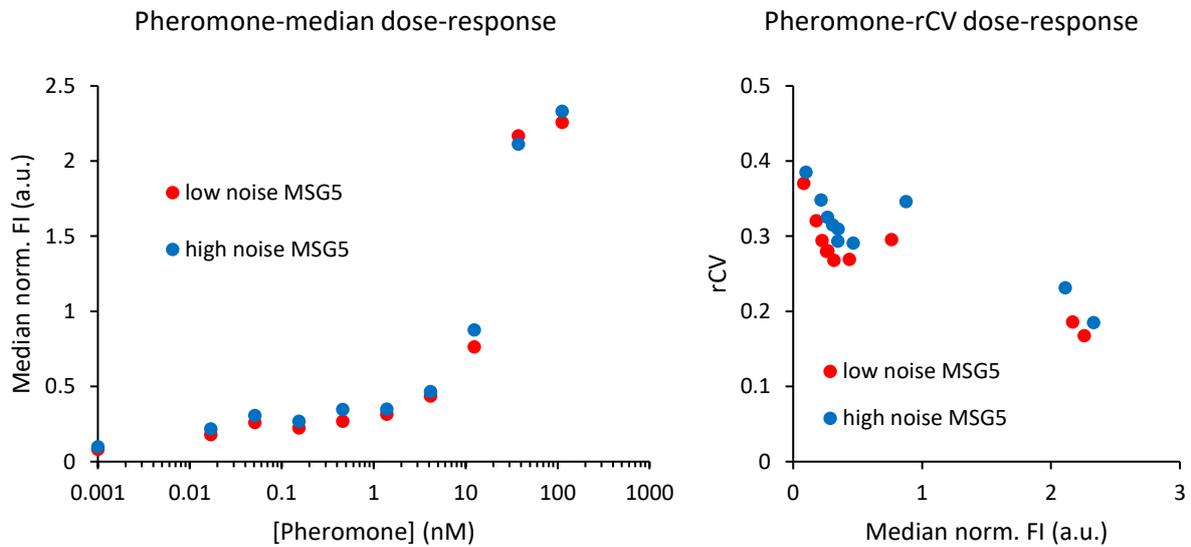


Figure 44: Pheromone response for cells with *MSG5* expression either set to low noise (4 ng/ $\mu$ L doxycycline, red) or high noise (2 ng/ $\mu$ L doxycycline and 12.8 mM theophylline, blue). Left panel: Pheromone dose response shows similar dynamics for both settings. Right panel: Low noise *MSG5* setting results in lower pathway noise over the entire range of pheromone induction. All values derived from *mTurquoise2*-normalized fluorescence intensities.

For both noise settings, cells showed very similar pheromone responses and the typical switch-like pathway activation at high pheromone concentrations (left panel in Figure 44). We calculated the robust coefficient of variation for all data points and found that the cells set to low *MSG5* expression noise indeed exhibited slightly less variability over the entire range of pheromone stimulation (right panel in Figure 44), indicating that *MSG5* expression noise did indeed affect pathway noise, albeit only to a small extent.

## Results

### 2.3.5 Sst2

We tested the capabilities of the noise tuning system to control *SST2* and in turn, change pathway expression and noise. *Sst2* is a GTPase accelerating protein that acts upstream in the mating pathway and attenuates signaling. *Sst2* has been described as pathway noise suppressor, making it a promising candidate for our noise modulation studies (64).

Regulation of *SST2* expression by doxycycline and theophylline had a strong effect on the pathway response. Because *Sst2* is a negative regulator, pathway activity is reduced with increased *SST2* expression (Figure 45, left panel). When *SST2* transcription was induced, the pathway showed more than 7-fold repression when *SST2* mRNA was stabilized, as compared to 1.3-fold repression when no theophylline was added.

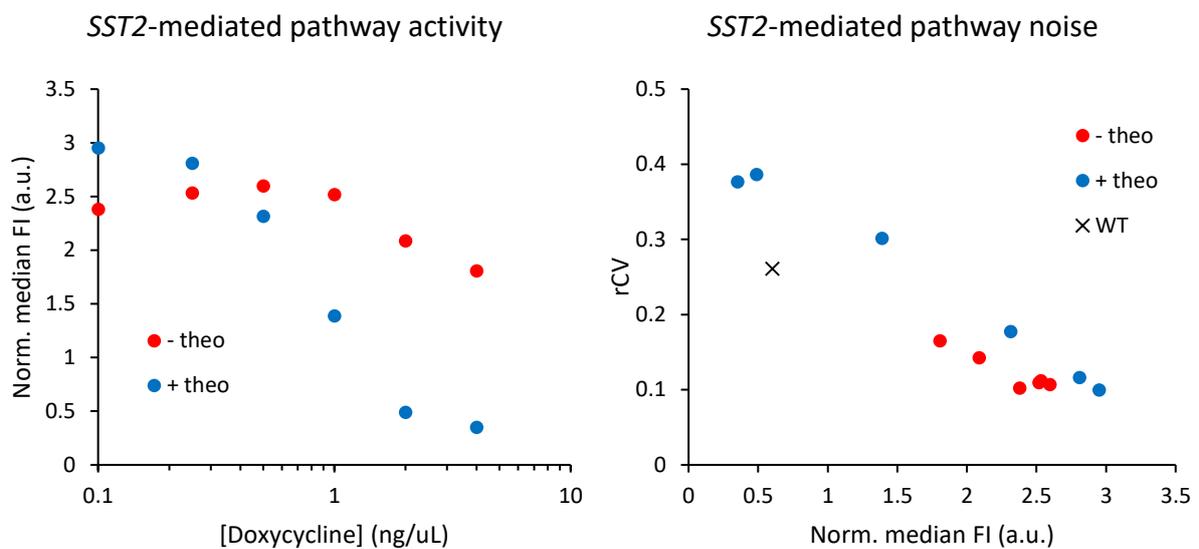


Figure 45: Pathway activity for different *SST2* regulation conditions. The *SST2* noise tuner was induced over a range from 0 to 4 ng/ $\mu$ L doxycycline and either without theophylline (red dots) or with 12.8 mM theophylline (blue dots). Cells were stimulated with 2.47 nM pheromone. Left panel: Doxycycline dose and corresponding *mNeogreen* pathway reporter response. Right panel: Pathway noise, given as rCV, for corresponding median expression. Black X denotes wildtype expression and noise. These measurements were made with reduced PMT voltage for *mTurquoise2* detection, hence the increased normalized fluorescence.

Over the entire expression range cells with destabilized *SST2* transcripts exhibited lower pathway noise than those with stabilized *SST2* transcripts (Figure 45, right panel). Following up on these initial results for the *SST2* noise tuner, we screened a wider range of doxycycline and theophylline concentrations to regulate *SST2* expression and analyzed the pathway activity and the pathway noise. Figure 46 shows the results plotted as noise heatmap, similar as for the proof-of-concept noise control constructs in Figure 31. We observed noise tunability at intermediate pathway activity, around 2.0 to 2.5 arbitrary units.

## Results

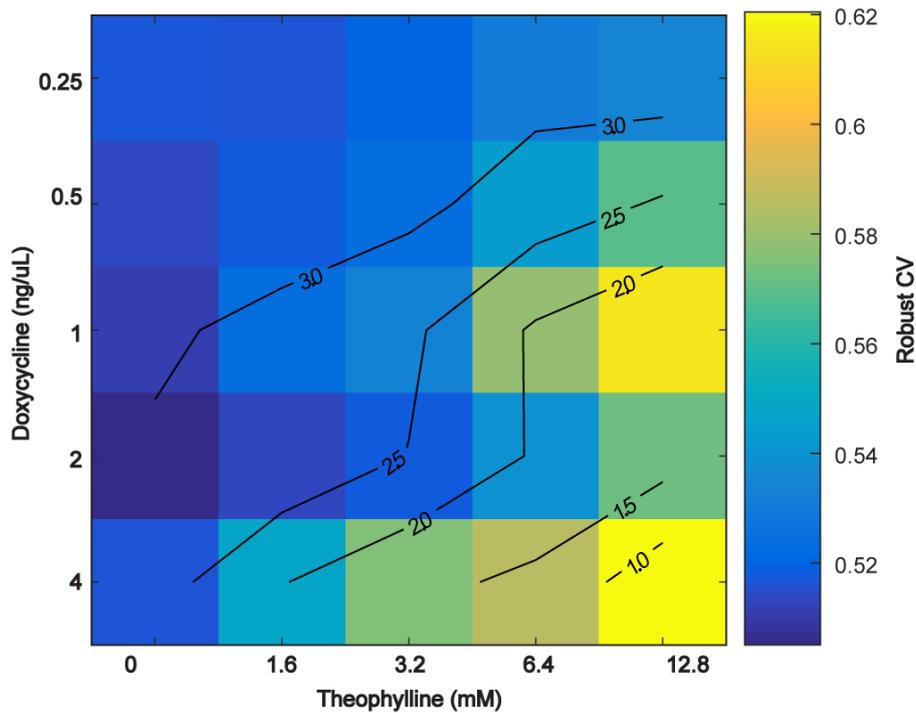


Figure 46: Noise heatmap for mating pathway activity when *SST2* is regulated with different combinations of dox/ theo concentrations. Pathway noise is tunable via *SST2* at intermediate pathway activity. Colors indicate robust coefficient of variation of the pathway reporter. Lines indicate median pathway reporter fluorescence intensity in arbitrary units. Cells were stimulated with 2.47 nM pheromone.

The results above show that the *SST2* noise tuner is capable of accurate control of mating pathway activity and the corresponding pathway variation. Based on the good inducibility of the pathway by controlling *SST2* expression with both inducers, we focused our efforts on *SST2* as a model to demonstrate the impact of gene expression noise of a single pathway component on the overall pathway performance, using the noise tuner.

As for *MSG5*, we screened dox/ theo combinations that resulted in a similar pheromone dose-response. The *SST2* noise tuner driven by the inducible Tet promoter and the inducible ribozyme terminator was grown with destabilized (0 mM theophylline) and stabilized (10 mM theophylline) 3' UTR with varying concentrations of doxycycline. Cells were stimulated over the entire pheromone response range and the pathway response was measured by flow cytometry. We found that cultures induced with 1 ng/  $\mu$ l dox and 10 mM theo produced similar dose response curves as cultures induced with 8.5 ng/  $\mu$ l and 0 mM theo (Figure 47). We will refer to these combinations of inducer concentrations as high-noise and low-noise setting, respectively.

## Results

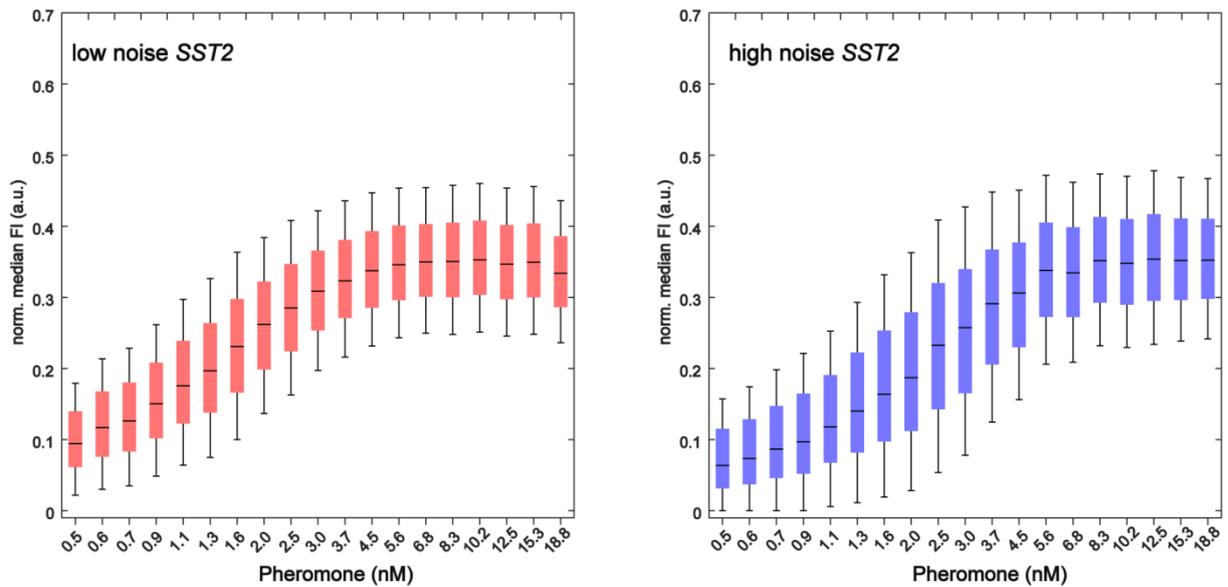


Figure 47: Dose-response curves of low and high noise SST2 strains for different levels of pheromone stimulation. Left panel: Low noise setting with SST2 expression induced with 8.5 ng/  $\mu$ l and 0 mM theophylline. Right panel: High noise setting with SST2 induced with 1 ng/  $\mu$ l dox and 10 mM theophylline. Black lines in boxes indicate median normalized fluorescence intensity of the population. Bottom and top of boxes indicate 25<sup>th</sup> and 75<sup>th</sup> percentiles, respectively. Whisker length is set to 50% of box length. Outliers not shown.

We calculated the robust coefficient of variation for all data points and plotted the values as a function of the pathway activity (Figure 48). At low pathway activity we observed ca. 50 % noise difference between strains with SST2 set to high expression noise and strains with SST2 set to low expression noise. The low noise setting results in less pathway variation over a wide range of pheromone concentrations. Pathway noise values for high and low SST2 expression noise converge when the pathway activity is saturated (normalized median fluorescence intensity of ca. 0.35).

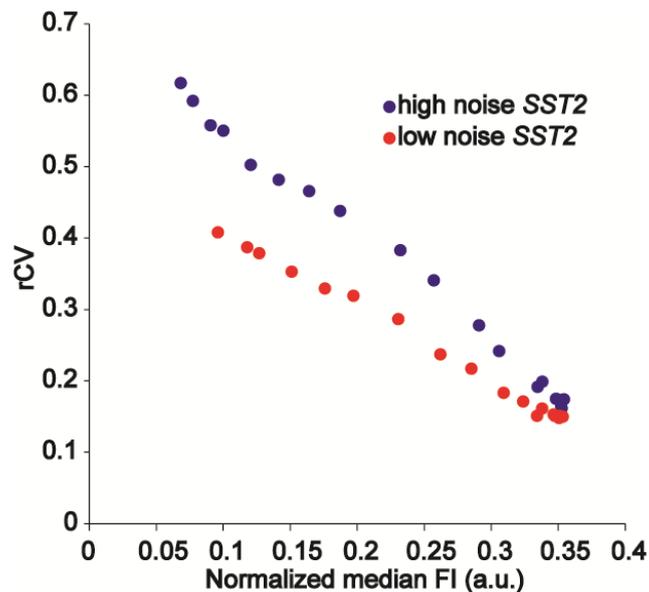


Figure 48: Mean-noise relationship of the pathway reporter output for SST2 set to either high noise (blue) or low noise (red). Low noise SST2 results in lower pathway noise over the entire range of pathway activity. Dots indicate different stimulation with pheromone.

We selected individual data points from low and high noise settings at similar median expression and compared the distributions. The SST2 noise tuner set to low noise resulted in slightly sharper and less

## Results

skewed peaks over the entire range of pheromone concentrations. Examples of distributions are shown in Supplementary Figure 9.

Next, we assessed whether the differences in pathway noise affected the cells' ability to integrate the input signal and compute an appropriate output. We calculated the mutual information between windows of four adjacent pheromone concentrations and the fluorescence intensities of individual cells that were stimulated with either of those concentrations (for details see materials and methods, section 4.10). We found that when the *SST2*-noise tuner was set to low noise, the signal transduction was improved by up to 50 % compared to the high-noise setting (Figure 49). The most precise information transmission was observed in the dynamic range of the pathway (ca. 0.15 to 0.3 a.u.). Once the pathway is saturated (0.35 a.u.), no information about the pheromone concentration can be inferred from the pathway activity and the mutual information drops to 0.

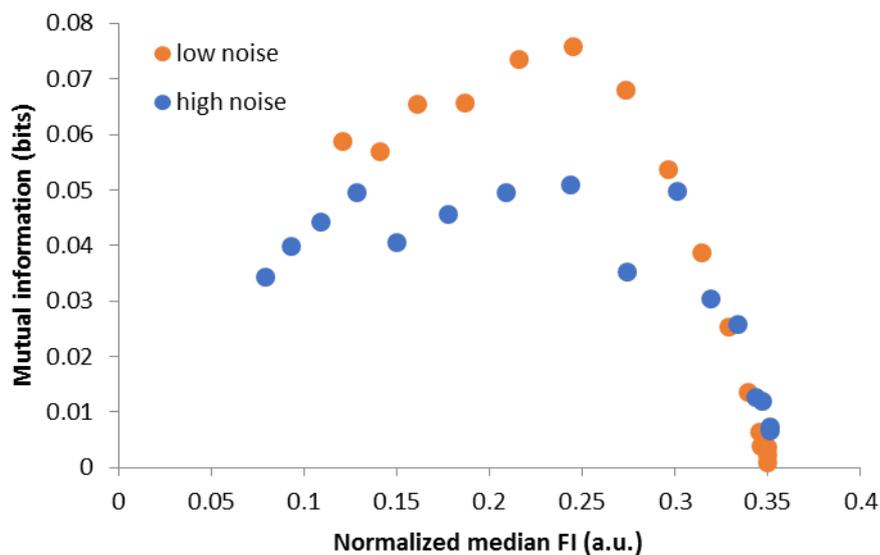


Figure 49: Mutual information of input (pheromone) and output (fluorescence intensity) from the datasets in Figure 47. Higher mutual information indicates a more precise signal transduction. When the *SST2*-noise tuner is set to low noise, the pathway shows up to 50 % increased information transmission.

In *Saccharomyces cerevisiae* the mating pathway response correlates with distinct phenotypes (Figure 9). To find out whether the change in pathway noise also had an impact on the morphology of the cells at different pheromone concentrations, we acquired microscopy images of low and high *SST2* expression noise at different pheromone concentrations. Figure 50 shows example microscopy images of predominantly unstimulated, chemotropic, and shmooing cells.

## Results

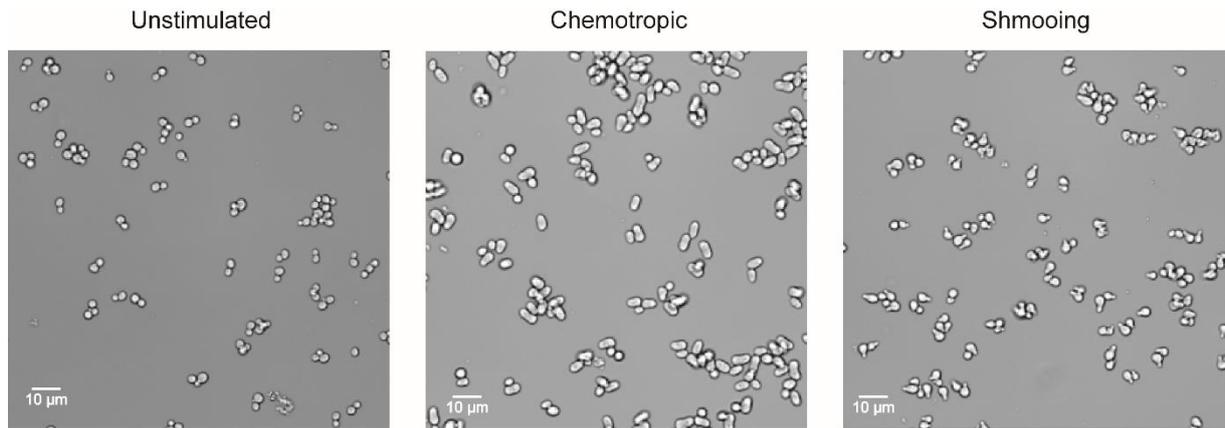


Figure 50: Three distinct phenotypes during mating response. Microscopy images of yeast cells after stimulation with different pheromone concentrations. At 0 nM pheromone, the cells retain the unstimulated phenotype (left). Cells respond with chemotropic growth at intermediate pheromone concentrations (middle). At high pheromone concentration, cells develop mating protrusions (shmooring, right).

In a blinded experiment, individual cells were classified into three categories: Unstimulated, chemotropic, and shmooring. The classification was done based on time-lapse movies for a total of ca. 9800 cells (see Materials and Methods, section 4.13). We counted the number of cells per category and microscopy image and calculated the ratio of the three phenotypes for different pheromone concentrations (Figure 51).

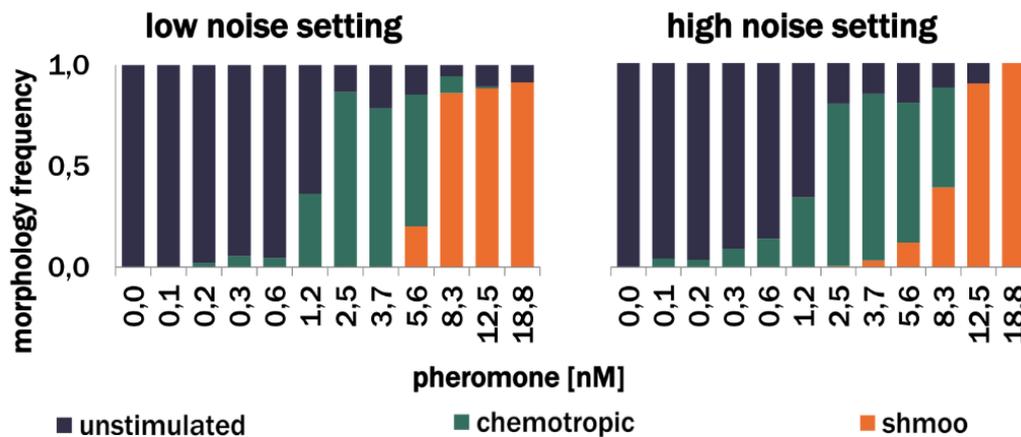


Figure 51: Ratio of unstimulated, chemotropic, and shmooring phenotypes at different pheromone concentrations. When the *SST2* noise tuner was set to low noise, cells respond switch-like with the shmooring phenotype (orange bars, left panel). For the high-noise setting, the shmooring phenotype is developed more gradually (right panel).

We found differences in the ratios of phenotypes dependent on *SST2* expression noise. When the *SST2* noise tuner was set to low noise, cells developed the shmooring phenotype more switch-like than cells with the high noise setting.

Taken together, the results obtained for *SST2* show that the noise tuner developed in this study can have a strong impact on pathway activity and pathway noise. Moreover, we have shown here that different expression settings mediated by our noise tuning system have a physiologically relevant effect on the pathway precision, which we demonstrated by changes in signal transmission and the phenotypical response to the pathway input.

## Results

### 2.4 Noise control via expression rates in native genes

We screened the literature for native yeast genes with extreme mRNA production or degradation rates to find out whether they could be employed to adjust expression variability similar to our noise tuner. We used data of mRNA synthesis rates and mRNA half-lives of yeast genes (85) to calculate the mRNA decay rate and the steady-state mRNA copy number per cell. With this we identified genes that reach a given steady-state mRNA level by extremely high or extremely low mRNA decay rates. Figure 52 shows the result of the analysis for 3788 yeast genes.

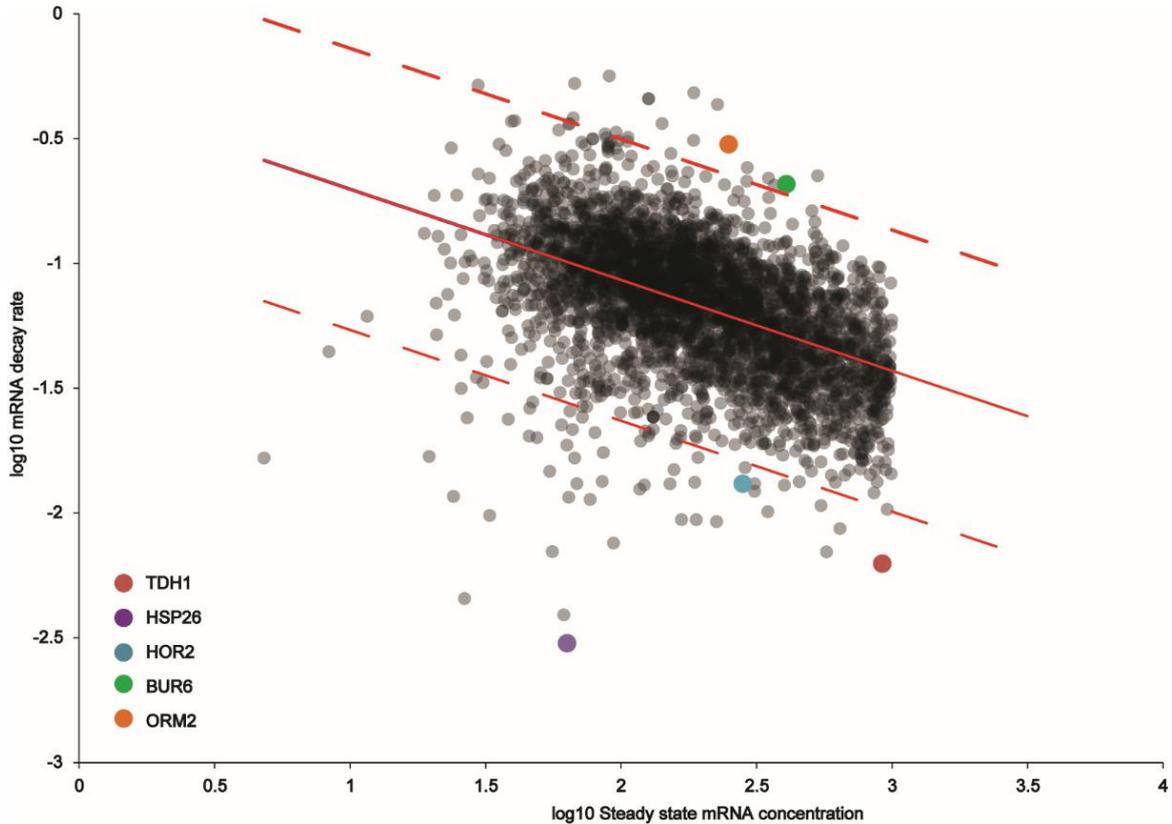


Figure 52: Analysis of extreme mRNA decay rates at a given steady state mRNA concentration. Calculations based on mRNA half-lives and mRNA synthesis rates from (85). Each of the 3788 dots represents a yeast gene. The full red line is an extrapolation of the mean decay rate at a given mRNA concentration. Genes separated from the mean by a dashed line were defined as outliers with either extremely high (e.g. BUR6 and ORM2) or extremely low mRNA decay rate (e.g. TDH1, HSP26, and HOR2) at a given steady state mRNA concentration. Only genes with a log10 steady state mRNA concentration below 3 are plotted.

From the data we selected the five genes with extreme mRNA decay rates highlighted in Figure 52 for further experimental analysis. The mRNA synthesis rate  $v_0$ , mRNA decay rate  $d_0$  and the steady state mRNA number  $\mu = \frac{v_0}{d_0}$  of those genes are given in Table 2.

## Results

Table 2: mRNA synthesis rate  $v_0$ , mRNA decay rate  $d_0$  and the steady state mRNA number  $\mu$  of five yeast genes with extreme  $d_0$ . *BUR6* and *ORM2* reach a given mRNA steady-state by high synthesis and decay rates, *HSP26*, *TDH1* and *HOR2* by low synthesis and decay rates. Calculations based on reported values from (85).

Gene	$v_0$ ( $\text{min}^{-1}$ )	$d_0$ ( $\text{min}^{-1}$ )	$\mu$
<i>BUR6</i>	84.49	0.208	407
<i>ORM2</i>	74.82	0.300	249
<i>HSP26</i>	0.19	0.003	63
<i>TDH1</i>	5.76	0.006	919
<i>HOR2</i>	3.68	0.013	281

According to these calculations and the observations made for our noise tuner construct, which can employ a high-noise and a low-noise strategy to reach a desired expression level, we hypothesized that *BUR6* and *ORM2* should have rather low gene expression noise, while *TDH1*, *HOR2* and especially *HSP26* should have high gene expression noise.

We constructed yeast strains harboring mNeogreen driven by the promoter/ terminator pairs of the five gene candidates in Table 2. Promoters and terminator regions were defined as the 500 bp upstream and downstream of the ORF, respectively. Median and noise of mNeogreen intensities of these strains were measured by flow cytometry and compared to the median-noise space that could be set by the noise tuner (Figure 53).

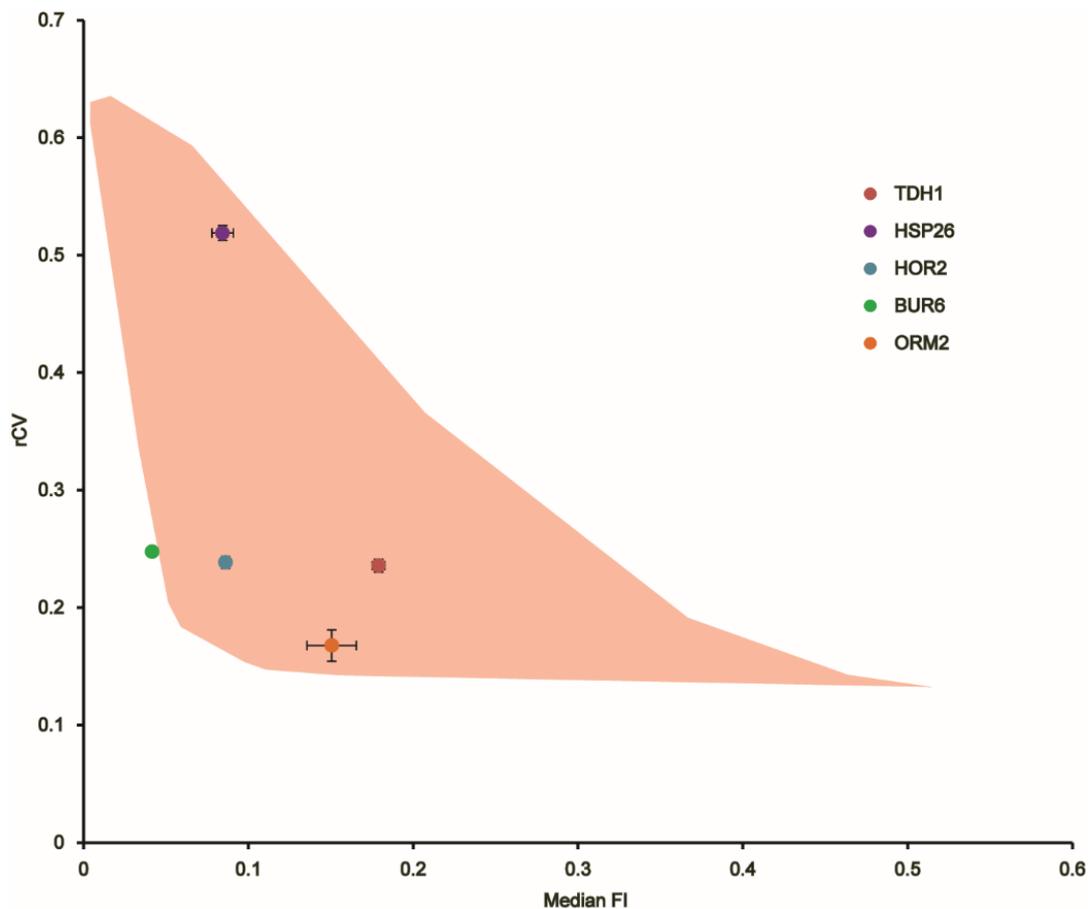


Figure 53: Median-noise space of the mNeogreen noise tuner compared to native regulation. Dots indicate the median and robust CV of mNeogreen driven by native yeast promoter/ terminator pairs from five different genes. The light-red area indicates the median/ rCV space that can be set by the noise tuner. Error bars indicate standard deviation of three biological replicates. The median fluorescence intensity was normalized to the constitutively expressed mTurquoise2 module.

## Results

We found that the noise tuner covered the median-noise values of four of the five genes. The *BUR6* promoter/ terminator pair was an exception because at its mean it conferred slightly lower noise than could be achieved by the noise tuner. The mRNA production and degradation rates in Table 2 correlate for the most part with the measured protein noise in Figure 53. High rates (*BUR6* and *ORM2*) lead to low variation at a given median protein concentration. Low rates (*TDH1* and *HSP26*) result in high protein noise at a given median. The *HOR2* promoter and terminator pair is an exception, conferring relatively low noise despite low mRNA production and degradation rates. In terms of median expression, the measured fluorescence intensities only partly reflect the mRNA steady state numbers, e.g. for *TDH1* which has the highest calculated steady state mRNA number and also the highest fluorescence intensity in this experiment. Notably, the noise tuner showed potential to reduce the noise for three of the five promoter-terminator combinations tested (*TDH1*, *HSP26*, *HOR2*).

To evaluate how much of the noise could be attributed to gene-specific variation, we again performed a reduced gate size analysis (Figure 54). For all five genes we found only minor contribution of extrinsic noise to the total observed variation.

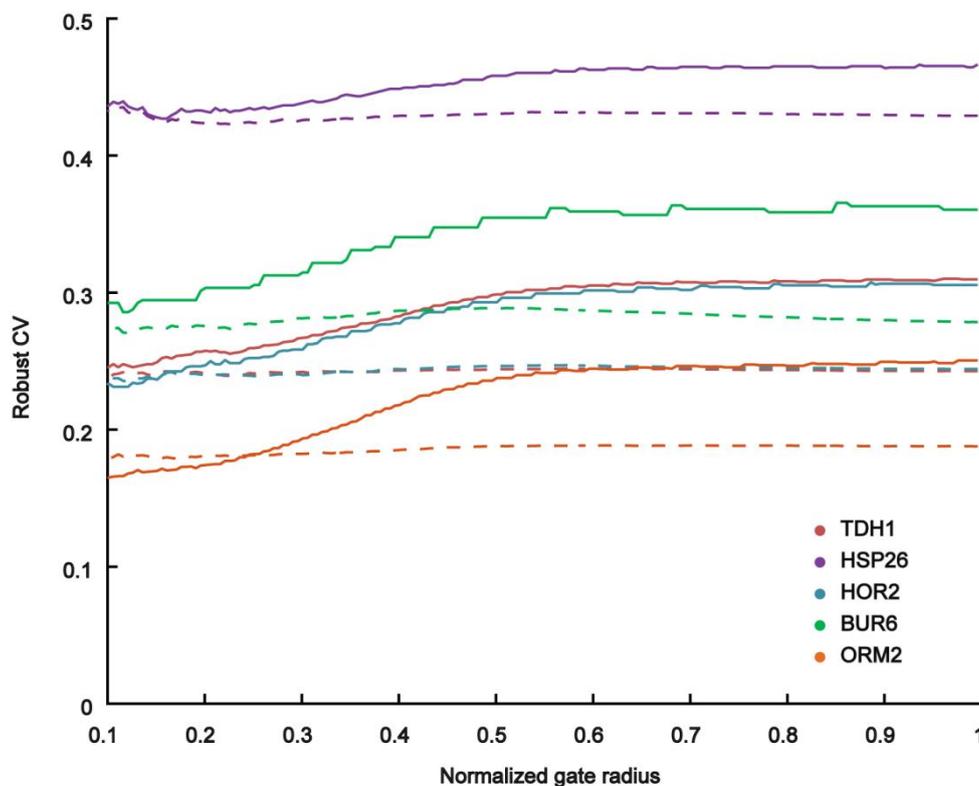


Figure 54: Reduced gate size analysis to measure extrinsic contribution to noise for mNeogreen driven by five different native promoter/ terminator pairs. Raw data shown in full lines, normalized fluorescence normalized to internal mTurquoise2 control shown in dashed lines. Extrinsic noise can be estimated from the rCV values of the full gate radius to the reduced gate radius.

As for the noise tuner, this suggests that most of the observed noise is intrinsic and thus, is created by gene-specific regulation, such as regulation of transcription rate and mRNA degradation rate.

We wanted to find out whether our calculations and experimental observations of the native promoter/ terminator pairs could be linked to protein function. For this we studied the literature for the biological functions of the two extreme cases on the noise spectrum, *BUR6* and *HSP26*.

## Results

The *BUR6* promoter and terminator conferred extremely low mNeogreen expression noise, slightly lower even than what we could achieve with our synthetic mNeogreen noise tuner. *BUR6* is an essential gene and its gene product is a subunit of a heterodimeric transcription factor. The heterodimer is a transcriptional regulator that represses the basal transcription machinery (86). Noise has been hypothesized to be detrimental in genes whose deletion is lethal and those that encode subunits of multiprotein complexes (34). Corresponding to this theory, we found that *BUR6* expression exhibits low noise. Moreover, our experiments indicate that *BUR6* expression follows the low-noise strategy described in this thesis, mediated by its promoter and terminator sequence.

*HSP26* expression is at the other extreme of the noise spectrum, with very low transcription and mRNA degradation rates. Hsp26, as the name suggests, is a heat-shock protein with a chaperone function (87). Expression of the non-essential *HSP26* gene is induced under different stressors including heat shock and salt shock (88). Stress response genes have been suggested to exhibit noisy expression as a bet-hedging strategy which allows clonally identical cells to cover a large dynamic range of protein concentrations, resulting in subpopulations with increased fitness when environmental conditions change (18). Here, we reported experiments for *HSP26* which indicate that this bet-hedging strategy can be naturally utilized by setting the transcription and mRNA degradation rates of the stress-response gene; the same mechanisms that we employ for the high-noise setting of our noise tuner.

To sum up, we have shown experimentally that yeast genes can exhibit extremely high or low variability in their expression, which is mediated by the promoter and terminator regions and that the resulting noise is in line with existing hypotheses that make a connection between gene expression noise and certain gene functions.

### 3 Discussion

#### 3.1 Noise-optimized measurement systems

Each part and process involved in experimental measurements is subject to noise. We call these sources of noise “nonspecific” to distinguish them from the variation that we can specifically control via the noise tuner. In order to reliably distinguish even small differences in controllable noise, we characterized the sources of nonspecific noise and minimized them by changing biological parameters, measurement parameters or data processing.

##### **Optimizing biological parameters**

We exchanged the 5' UTR in our reporter genes to increase expression and thus achieve a better separation from the auto-fluorescence signal, resulting in a 50 % increase in mNeogreen expression and a ca. 10-fold increase in mCherry expression (Figure 12). We explain the large difference in expression gain by context-dependent effects, mediated by the different open reading frames of mNeogreen and mCherry. This is a common observation which hampers attempts to standardize genetic parts in the field of yeast synthetic biology. Modularity of genetic elements in eukaryotes is only maintained to a certain extent and is very much context-specific. Approaches to predict expression strength for different (native or synthetic) 5' UTRs have shown to be accurate for a given ORF but showed only limited success when the 5' UTRs were combined with a different ORF that the algorithm was not trained for (79, 89).

We employed mTurquoise2 as additional fluorescence reporter for the expression capacity of individual cells. Initially, however, the fluorescence was almost undetectable which is why we used codon-optimization to increase expression (Figure 14). This strongly increased fluorescence intensities and allowed us to make reliable measurements with triple-reporter strains. For the codon optimization, over 30 % of the original primary sequence was changed (see supplementary figures).

We observe that the two loci we used for integration of fluorescent reporter genes mCherry and mNeogreen are qualitatively similar in terms of their noise (Figure 13). This is not necessarily what one would expect, since considerable variation in expression noise from different chromosomal locations has been reported (90). This makes the large difference in intrinsic noise measured with the one-locus and the two-loci approach even more striking: Integration of mCherry and mNeogreen into the same locus resulted in 4-fold reduced intrinsic noise compared to when mNeogreen was integrated into a different locus. One reason for this might be heterochromatin remodeling, which is likely to be uncorrelated in different loci. Another possible reason which has been hypothesized is spatially inhomogeneous distribution of regulators which can create uncorrelated variations when the two genes are situated in different positions in the nucleus (91). If genes are located in close physical proximity to each other, noise that arises from the recruitment of transcription factors should be mitigated. If this is the underlying reason for reduced intrinsic noise in the same locus, we would expect the Tet operator sites to act as a sink for the activator protein rtTA. Correspondingly, we would expect a higher local concentration of rtTA when both Tet promoters are situated in the same locus, which should result in higher gene expression. Indeed, we find that mCherry fluorescence is increased by 37% when mNeogreen is expressed from the same locus (Supplementary Figure 10). This finding represents a nice example for context-dependent expression and might be a potential strategy for synthetic expression increase, by placing additional *cis*-elements in close proximity to a gene of interest to serve as a sink for transcription factors. The influence of relative genomic localization has practical

## Discussion

implications: Genes whose products are required in a fixed stoichiometry (e.g. subunits of the same protein complex or enzymes in the same production pathway) should be integrated into the same locus to reduce intrinsic noise.

### Optimizing measurement parameters

Growth kinetics measured by flow cytometry allowed us to observe an increase in the doublet population that peaked at maximum growth rate. Doublets are a common observation in flow cytometry and are explained by aggregated cells that pass the laser beams at the same time, creating artificially high fluorescence values. For that reason, doublets are usually gated out and excluded from further analysis. Here we observe a correlation between doublets and growth rate for *Saccharomyces cerevisiae* that indicates that part of the doublets are not random aggregates but budding cells (Figure 17). This could be confirmed e.g. by cell sorting of singlets and doublets and subsequent microscopy to observe budding.

Sonication lead to a decrease of the doublet population, but only by 50 percent (Figure 18). We assume that the successfully dispersed cells are mainly aggregates with weak cell surface interaction and that the cells that remain in the doublet population are mother cells that are still covalently attached to their daughter cells. For the purpose of our high-throughput experiments we found that sonication was not applicable because of the fast regeneration of doublets after dispersion (Figure 19). Normalization to the expression capacity, as discussed below, turned out to be a more efficient way to control for cell-to-cell variation and rendered gating of singlets unnecessary.

Subtraction of auto-fluorescence is crucial for the noise calculation, especially at low expression levels. Auto-fluorescence adds to the reporter signal, leading to a higher median fluorescence intensity, which in turn leads to calculation of artificially low rCV values. A commonly used method for auto-fluorescence removal is to simply subtract the mean or median fluorescence of a control strain without the reporter gene. While this might be accurate enough when the measured samples exhibit fluorescence that is well separated from auto-fluorescence, it has major limitations at intensities that overlap with auto-fluorescence, creating artificial negative values. Gene expression noise, as shown in this thesis and in numerous other studies cited, is more pronounced at low steady-state protein levels where protein production occurs by rare stochastic expression events. This makes noise analysis at low expression levels particularly interesting but unfortunately it also makes auto-fluorescence a particular nuisance. We report here an auto-fluorescence subtraction method that alleviates the limitations of noise measurements at low expression levels (Figure 24). The algorithm, devised in our lab by Seán Murray, makes two very logical assumptions: Firstly, we assume that the auto-fluorescence probability distributions in the negative control and the tested strains are the same. This is likely to be the case, since the negative control we use is usually the parental strain. Secondly, we assume that the auto-fluorescence can never be bigger than the measured signal. This assumption simply states that fluorescence attributed to a reporter gene cannot be negative. The described algorithm is a statistical method that cannot determine the actual auto-fluorescence of a cell, meaning that while it is unfeasible on the level of individual cells, it produces good results for sufficiently large datasets.

### Optimizing analysis

To discriminate between intrinsic and extrinsic noise, we employ two different methods: The dual reporter assay, as introduced by Elowitz and colleagues (2), and the reduced gate size analysis used by Newman and colleagues (18). To the knowledge of the author of this thesis, these two methods have

## Discussion

never been compared directly, which is however important because both methods have benefits and drawbacks. The dual reporter assay requires a more sophisticated biological setup because two fluorescence genes have to be genomically integrated, plus ideally a third one should be included as control. This requires additional time and effort to construct strains. Moreover, the assumption is made that the two reporters are regulated identically, which might not always be true, because the genetic context, such as the controlled open reading frame, can influence expression. The reduced gate size analysis on the other hand requires only one reporter and thus, can be readily applied to already existing fluorescence reporter strains. If the number of available cells is limited, however, this is not the method of choice. By reducing the gate size, the number of analyzed cells is further reduced and beyond a critical limit it will generate unreliable values for the coefficient of variation. Moreover, the reduced gate size analysis can only be used to estimate the contribution of the extrinsic noise to the total noise; more specifically, only the fraction of extrinsic noise that can be accounted for by the forward and side scatter signals. The remaining noise contains intrinsic as well as technical noise factors, which cannot be readily distinguished. This limitation can be circumvented when the technical noise can be estimated, which we did in this thesis by measuring fluorescent beads and testing the laser power and PMT voltages of the flow cytometer (Figure 21, Figure 22, and Figure 23). The reduced gate size method proves useful for the analysis of noise in budding yeast because most of the noise is caused by parameters that are described by the forward and side scatter signals, with the cell size being the main factor, presumably. For other organisms with a more homogeneous size distribution the amount of extrinsic noise that can be filtered out by this method might be smaller.

We use  $\eta^2$  and the (robust) CV as parameters for the noise in dual- or single reporter assays, respectively. Mathematically,  $CV^2$  and  $\eta^2$  both describe the normalized variance, meaning the size of protein fluctuations compared to the mean protein concentration. In noise research  $\eta^2$  is used conventionally when the signals are further deconvoluted mathematically into an extrinsic ( $\eta_{ext}^2$ ) and an intrinsic term ( $\eta_{int}^2$ ), which is described in more detail in (92). We find that intrinsic noise is largely of biological nature, changing only with biological parameters, such as the integration locus or (as we show in this thesis) the transcription and mRNA degradation rates of a gene. Extrinsic noise on the other hand is a more diffuse concept which is affected by biological parameters as well as data processing. Due to its “catchall nature” (17) the mechanisms of extrinsic noise are difficult to analyze. However, for the purpose of the experiments presented in this thesis the focus lay on the control of gene-specific, intrinsic factors of noise. For that reason, in our analysis pipeline we decided to filter out extrinsic noise as efficiently as possible, which made deconvolution of signals into their intrinsic and extrinsic components largely unnecessary.

Normalization to an internal control of the expression capacity, in this thesis the constitutively expressed mTurquoise2 gene, proved to be remarkably efficient for filtering out extrinsic noise. This was most apparent in a dual-reporter assay in which extrinsic noise was compared in raw data and mTurquoise2-normalized data, the latter showing almost no extrinsic noise (Figure 25 and Figure 26). Efficient noise filtering improves the quality of the results presented here, as it reduces the risk that other sources of noise mask the gene expression noise that we want to measure. There is a caveat, however, regarding the mating pathway experiments that we performed later on: Pheromone stimulation triggered an increased expression of the GPD promoter-driven mTurquoise2 gene. Supplementary Figure 7 shows that the expression of the control gene is increased twofold over the range of the pheromone stimulation. This illustrates that a promoter which is considered constitutive under one condition might still exhibit altered expression levels under a different condition.

## Discussion

### 3.2 Synthetic control of gene expression noise

The central achievement reported in this thesis is the development of a system that is capable of tuning gene expression noise. To the knowledge of the author, the noise tuner presented here is currently the only system that allows external direct control of the noise associated with principal steps of the central dogma of molecular biology. The noise tuner is a device consisting of two parts: An inducible promoter for transcription rate control and an inducible ribozyme for mRNA degradation rate control. It employs two inputs, doxycycline and theophylline, to generate a single output with a new property: The capability to decouple mean expression from its noise. In that respect, what we report here is an example for a phenomenon termed emergence. Emergent properties such as the mean-noise decoupling arise out of the properties of the parts that make up the device and are qualitatively novel and irreducible with respect to them. The parts of the noise tuner (promoter and ribozyme) have properties that, when combined, create a device that can decouple mean from noise, although the individual parts are incapable of mean-noise decoupling.

In the introduction we briefly described another strategy for mean-noise decoupling via external inducers, published in (30) and based on (27). Their strategy utilizes a two-input one-output approach to tune the abundance and the activity of a transcriptional regulator controlling expression of a reporter gene. The authors report different noise levels but similar mean for two settings: Low-abundant, highly active transcriptional regulator proteins created high noise and highly abundant transcriptional regulators that had low activity created low noise. While this system might prove useful to analyze the consequences of additional noise in a cellular process, it is not suitable to reduce noise in that process. We propose our noise tuner as a complementary approach that is more suitable when the goal is to reduce gene expression noise. While a direct comparison of the noise values was not possible because the raw data was unavailable, we argue that their low-noise setting corresponds to our high-noise setting: Both involve a highly expressed regulator that controls the transcription rate of a reporter gene. In our case this is the reverse tetracycline trans-activator (rtTA), driven by a strong constitutive promoter which produces low variation in the target gene (93). In our noise tuning system, as opposed to (30), we can further decrease noise also by upregulating rtTA activity, thus, reducing variability by stochastic binding of rarely activated rtTA proteins. When we choose high rtTA activity (high dox), we adjust the expression downstream by destabilizing the mRNA (low theo) to allow comparable mean expression levels.

#### **mRNA stability control with inducible ribozymes**

To control the mRNA stability, we built rationally designed terminator regions consisting of two parts: an inducible ribozyme and a synthetic terminator sequence that contains motifs that are sufficient for efficient termination (47). The combined sequences have a length of about 200 nt, which lies within the typical length of *S. cerevisiae* terminator regions (including the length of the 3' UTR, 94). The relative ease of combining those two elements to create a functionalized terminator region indicates the possibility for further functionalized terminator regions with desired properties, such as adjusted translational efficiency (42) or spatial control of expression (95) by including the corresponding sequence motifs into the terminator region.

To evaluate the performance of the inducible ribozymes, we focused on two key parameters of genetic control elements: the basal expression and the dynamic range of the controlled gene. Both of the tested Theo-ON systems (i.e. ribozymes that activate gene expression when theophylline is added) showed a comparable fold change (13- and 14-fold for Ribo 1 and Ribo 2, respectively; Figure 27). Due

## Discussion

to differences in the basal expression, the two ribozymes differ in their absolute expression range. Which one to choose for expression control depends on the individual application. Ribozymes that confer low basal expression could prove useful in combination with e.g. toxic genes that require a complete shutoff of expression under certain conditions. A limitation of both ribozymes is clearly the high concentration of theophylline (mM range), that is required to deactivate the endonuclease activity of the ribozymes tested here. We tested theophylline concentrations of up to 25.6 mM, while the solubility limit of theophylline in aqueous solutions is 40 mM. As a result, theophylline powder had to be dissolved directly into the medium for each experiment, instead of using highly concentrated stock solutions, with the potential risk of reduced reproducibility. Additionally, high theophylline concentrations resulted in growth defects (Figure 30) so that in later experiments the maximum concentration used was confined to 12.8 mM. The dose response curves suggest that higher concentrations of theophylline could stabilize the ribozymes even further but due to the drawbacks of poor solubility and growth defects we cannot reach saturating concentrations.

### Decoupling mean from noise

By combining an inducible promoter with an inducible ribozyme, we essentially built an AND-gate that requires doxycycline and theophylline to produce a signal. We harnessed the dynamic ranges that we could achieve for the transcription rate and the mRNA degradation rate to select concentrations of the two inducers that result in similar mean expression. As predicted from theoretical considerations (32), we were successful in decoupling absolute expression levels from the noise in gene expression. Somewhat surprisingly, however, we found that the mRNA degradation rate had almost no effect at all on the noise (Figure 31). This shows that, while stochasticity in transcription is the main contributor to noise, the process of mRNA degradation is not stochastic. One possible explanation for this observation is that mRNA degradation is a general process governed by the ubiquitous mRNA degradation machinery that breaks down transcript molecules efficiently and fast once the transcript is deadenylated (96).

We show that the noise tuner predominantly controls the intrinsic component of the observed noise, which is what is to be expected for gene-specific modulation of expression (Figure 33). Perhaps more noteworthy is the observation that in our experiments intrinsic and not extrinsic noise is clearly the dominant component. We note that this observation is in contrast to some reports that show dominant extrinsic noise in yeast (e.g. 3, 97). One explanation could be the differences in data processing. We show here that extrinsic (but not intrinsic) noise measured by flow cytometry is sensitive to normalization and gating. No standard procedures exist in the field of noise research so that results from different publications can rarely be compared directly. Moreover, as opposed to most other studies that investigate noise, our noise tuner operates at the lower bound of protein abundance. At lower abundance, intrinsic noise becomes more dominant (20). Entire classes of functionally important proteins are expressed at low abundances, including transcription factors (98). This shows the large potential of the noise tuner for noise reduction in synthetic genetic circuits.

We show the limits of the noise tuner by exchanging the promoter. In contrast to the TetO7 promoter, a promoter variant with only two Tet operator sites does not show the mean-noise decoupling characteristic (Figure 34). We suspect that the TetO2 promoter recruits RNA polymerase more stochastically and is simply less efficient, so that expression never reaches the levels necessary to transition into the low-noise regime.

## Discussion

### Model and stochastic simulation of the noise tuner

To further explore the experimental finding that noise depends on doxycycline but not theophylline, we employed an analytical model that allows us to describe mean and noise by only two parameters  $a$  and  $b$ . In simplified terms, the doxycycline concentration, which controls the transcription rate  $v_0$ , is a proxy for  $a$  and the theophylline concentration, which controls the mRNA lifetime  $d_0^{-1}$ , is a proxy for  $b$ . Mathematically, the mean protein abundance is the product of  $a$  and  $b$ , so that e.g. reducing  $a$  by half and doubling the value of  $b$  results in the same mean. Correspondingly, if we do the same exercise for the experimental data within the linear dose-response ranges of doxycycline and theophylline, we reach approximately the same median fluorescence intensity.

The most notable finding we make based on these theoretical considerations is that the CV as a measure of noise scales with the inverse square root of  $a$  and is independent of  $b$  if the latter is large. The assumption that  $b \gg 1$  implies that many protein molecules are being produced from one mRNA molecule. This is supported by previous studies that estimate values for  $b$  between 20 and several hundred molecules (99 and 100, respectively). In the respective parameter space, noise is caused by large amounts of proteins created from few, stochastically produced mRNA molecules at low promoter activity (or: low doxycycline concentrations). The stochastic simulation of the model is in good agreement with the experimental results (Figure 31 B and Figure 37).

The model implies that our experimental approach is only one of four different possible strategies to decouple mean from noise via expression rate control. The essential point is to control the parameters  $a$  and  $b$ , which in this study is done via  $v_0$  and  $d_0$ , respectively. The theory predicts that the same could be achieved by simultaneous control of  $a$  via the protein degradation rate  $d_1$  and  $b$  via the translation rate  $v_1$ . The other two possibilities are the simultaneous control of  $v_0$  and  $v_1$  and the simultaneous control of  $d_0$  with  $d_1$ . We argue that of those four potential strategies our system embodies the preferred approach for mainly two reasons: Firstly, the good availability and documentation of inducible promoters and ribozymes makes the control of the transcription rate and mRNA degradation rate the preferred choice for experimenters. Secondly, we argue that keeping the control on (post)transcriptional level is energetically cheaper for the cells than mean and noise tuning on protein level, because the pool of molecules that have to be produced and/ or degraded is much smaller. mRNA abundances in yeast range typically from close to zero to 100 copies per cell (101), while proteins are present in up to 1 million copies (98).

### Applicability of the noise tuner

As mentioned in the results section, we note that the different noise settings that can be achieved with our system come along with different energy requirements. Low noise through high transcription and mRNA degradation rates results in a high mRNA turnover which requires more energy than the high-noise setting that has a low mRNA turnover. Apart from detrimental growth at extreme theophylline concentrations we did not observe a correlation between low noise and reduced growth.

For the use in synthetic pathways one will have to test on a case-by-case basis if the benefits of low noise outweigh the increased energy requirements of maintaining a high mRNA turnover. This will depend on the pathway topology and the involved network motifs, the abundance of the protein of interest and other factors. Of course, it also depends on how the benefits are defined: For the development of a synthetic biosensory pathway, high specificity and sensitivity are the overarching goals. To achieve that, the increased energy consumption for the low-noise expression of critical

## Discussion

components is a small price to pay. For the example of the yeast mating pathway we have shown that the noise tuner is capable of improving sensory performance when it is applied to critical pathway components, which will be discussed in detail in the next section.

For other synthetic applications, e.g. metabolic pathways, benefits might be defined as increased product yield. Here, higher mRNA turnovers of enzymes involved in the pathway might be more crucial as that would divert energy away from the biosynthesis of the product. To the knowledge of the author, there are currently no published studies that specifically investigate the effect of expression noise of genes encoding metabolic enzymes on product yields in synthetic metabolic pathways. There are indications, however, that gene expression noise of flux-limiting metabolic enzymes is propagated into the metabolic fluxes in native pathways (recently reviewed in 102).

### 3.3 Synthetic control of pathway noise

We built the noise tuner with two applications in mind. On the one hand, we aimed to achieve noise reduction to improve homogeneous expression in a population for a variety of applications, e.g. in the field of metabolic engineering. On the other hand, we wanted to use it as a device to test the robustness of pathways towards noisiness of individual components. This could be instrumental for the identification of noise-critical components and noise propagation through a pathway.

There are various reasons why we chose the yeast mating pathway as our model to test the potential applications of the noise tuner. Firstly, it is one of the signaling pathways that have been described most thoroughly. Detailed knowledge about individual pathway components, interactions, and in some cases even expression noise is available. Secondly, as described in the introduction, the mating pathway has developed into a model pathway for synthetic biology, which we see as a likely field of application for the noise tuner. Moreover, the pathway has two well-defined readouts: a fluorescence reporter driven by a widely used pathway response promoter (see e.g. 54, 67, 73) and the distinct morphologies (unstimulated, chemotropic, shmooing) that correlate with the level of pheromone stimulation. Lastly, unlike many other sensory systems the mating pathway induces radical changes that affect the cell as a whole, including cell cycle arrest and partial cell wall disassembly at the mating protrusion. The high costs and high benefits of the pathway activation suggest that the accurate regulation of the pathway activity and the pathway noise is crucial. Hence, we expected that in such a pathway, changing the expression noise of a critical component could create a detectable change in the noise of the pathway output.

Regarding the setup of the described experiments, there are two further remarks. Firstly, the pathway reporter signal (mNeogreen controlled by pathway response promoter P(Fus1)) does not reflect purely the pathway noise but also the gene expression noise of the reporter itself. This noise, however, should be identical for all tested strains and conditions. Secondly, it is an indirect reporter. As opposed to the proof-of-concept experiments described before, we did not acquire direct expression data of the genes we intended to regulate. This would have been possible in principle by genetic fusion of the gene of interest to a fluorescence reporter. Genetic fusions, however, can have the disadvantage of obstructing protein function and might not be detectable for genes with low expression.

We tested the noise tuner for five selected pathway components that represent the main functions of the pathway from signal sensing (Ste2) and signal transmission (Ste5, Fus3) to response regulation (Msg5, Sst2). We analyzed strains that harbored the corresponding genes in the native locus but were controlled by a Tet promoter and an inducible ribozyme. The pathway genes were expressed with

## Discussion

different transcription rates and with their mRNAs either stabilized with theophylline or destabilized. For each component, we expected to observe one of three scenarios:

- 1) Theophylline has no effect on pathway activity and pathway noise. This indicates that no transcript stabilization can be achieved, possibly because the coding sequence of the mRNA contains elements that confer instability. Correspondingly we would expect no changes in pathway noise.
- 2) Adding theophylline changes pathway activity but no changes in pathway noise are observed. This indicates that the transcript can be stabilized, resulting in increased expression and higher or lower pathway activity for positive and negative regulators, respectively. No change in pathway noise can have two reasons: **(A)** The critical transcription rates to pass the threshold between the high and the low noise regime (approx. 1 to 4 ng/ $\mu$ L doxycycline) lie outside of the dynamic range of the pathway response. If for example the pathway activity is already saturated at doxycycline concentrations that are below the threshold, we would not expect differences in pathway noise when stabilizing the transcript. The same is true if pathway activity can only be achieved with doxycycline concentrations well above the threshold. In both cases the noise tuner would have to be changed to a device with a more suitable dynamic range. **(B)** If the noise tuning range of the pathway component matches the dynamic range of the pathway activity but still no change in pathway noise can be observed, then this would mean that expression noise of the target gene is filtered out in the subsequent steps of the signal transduction.
- 3) The noise tuner changes pathway activity as well as pathway noise. As for scenario 2, this shows that gene expression noise control via the noise tuner is possible. It also shows that fluctuations in the target gene expression are propagated to the measured output of the pathway, indicating the role of the gene for phenotypic robustness. The term 'phenotypic stabilizer' has been proposed for genes that are critical for robustness (103).

We observe all three scenarios with the five genes we tested. Applying the noise tuner to the mating pathway's MAP kinase Fus3 shows no impact on pathway activity (Figure 42), which we can explain with scenario 1. The noise tuner is not able to increase the stability of the transcript. The scaffold protein Ste5 shows the characteristics of scenario 2 with increased expression but no change in pathway noise when the *STE5* mRNA is stabilized (Figure 41). The dynamic range of the pathway activity lies within a doxycycline range where we would expect noise tuning to be possible, so it seems likely that the expression noise of *STE5* is filtered out by the pathway. One possible explanation for this is post-translational regulation of the Ste5 proteins via translocation. Ste5 in naïve cells resides primarily in the nucleus and upon pheromone stimulation are translocated to the mating protrusion (104). Compartmentalization of Ste5 might lead to localized abundances of proteins that are so high that individual translational bursts are efficiently buffered.

In scenario 3 the pathway output is susceptible to gene expression noise in individual genes. We observe this to a small extent for the pathway receptor Ste2 and the phosphatase Msg5, and to a bigger extent for the GTPase-activating protein Sst2. While for *STE2* the limited experimental data do not allow a final conclusion, reduced expression noise in *MSG5* and *SST2* showed reduced pathway noise over a wide range of pheromone stimulation levels (Figure 44 and Figure 47). Notably, while for the other three tested genes we reach wildtype noise levels, for *MSG5* and *SST2* controlled by the noise tuner we generally observe higher pathway noise. We explain this by the nature of the network motifs that these two genes represent in the native pathway: As negative feedback regulators their

## Discussion

expression is upregulated with pheromone leading to noise suppression in the wildtype. In our noise tuning strains, however, the native promoter is replaced by a Tet promoter which disrupts the feedback and thus, the noise suppression. In the case of *SST2* this might also be the reason for a less switch-like response to pheromone compared to the wildtype response.

The GTPase-activating protein Sst2 is critical to suppresses noise in the mating pathway response (64) and accordingly showed strong relative differences in pathway noise (Figure 48). These differences correlate with the amount of information that can be conveyed through the pathway. The mutual information between the input and output signals is a key metric for the accuracy of a signaling pathway. A 50 % difference, as we observe for low- and high-noise *SST2* settings (Figure 49), could make the difference between an accurate shmooing response and an inaccurate response, either by shmooing prematurely or shmooing too late. In the dynamic range of the pathway, the increased ability to resolve inputs enables low-noise *SST2* cells to make a switch-like mating decision, indicating relatively high pathway precision.

Taken together, we show here that the noise tuner, when applied to phenotypic stabilizers such as *SST2*, is indeed capable of adjusting pathway noise with a substantial effect on pathway precision. Thus, regarding the broader applicability of the noise tuner, an important question is: How common are phenotypic stabilizers? In terms of network motifs, negative autoregulation and negative feedback loops have been reported to act as stabilizers (105). Thus, pathway noise might be affected when genes that comprise these network motifs are noisier. This is in accordance with our findings that show that noise in *MSG5* and *SST2* expression affects pathway noise.

It has also been reported that components that are further upstream in the network can have a stronger effect on output noise. Noise in upstream regulators of synthetic genetic circuits showed a larger impact on the output noise than noise in downstream elements. Moreover, output noise was shown to scale with the length of the cascade (106). In signaling pathways, upstream elements determined the limit of signal information transduction capacity (107). This is also reflected by our finding that from all five tested genes, *SST2* noise had the strongest effect on pathway noise. These observations indicate that upstream regulators would be good candidates for noise tuning if no other relevant information on pathway components is available.

More mating pathway genes would have to be tested to allow a general statement on pathway topology and the gene expression noise of individual components. However, one general observation that we made is that with none of the five components tested did we achieve the strong noise control that we reached for the proof-of-concept strains (up to 3-fold noise tuning). Either this means that we have been unfortunate when choosing the candidate genes to test the noise tuner, or, more likely, the pathway has inherent mechanisms to buffer noise from individual components.

These findings suggest that noise could become an emerging hurdle for synthetic pathways which are increasing in size and complexity. Optimizing synthetic pathways for low noise has been largely overlooked so far, indicating a great potential for noise tuning. A first step to alert researchers in metabolic engineering and synthetic biology to the issue of gene expression noise would be to characterize large libraries of genetic control elements not only according to absolute gene expression but also to investigate the gene expression noise that arises from different combinations of promoters and terminators.

## Discussion

### 3.4 Noise tuning via expression rates: A strategy employed in nature?

The previous parts describe the design, construction and testing of a synthetic noise tuner. In the final set of experiments that we performed within the scope of this thesis, we investigated whether similar noise differences can be observed in native genetic control elements that confer different transcription and mRNA degradation rates.

The transcription and mRNA degradation rates can vary largely among yeast genes, even at the same steady-state mRNA concentration, and it has been shown that these rates can be dynamically regulated (85). The literature suggests that there is indeed selective pressure for some genes to be more or less noisy (e.g. 25, 108). The question remains if nature employs expression rate control to adjust noise levels in a similar fashion as we present it here for the noise tuner. Theoretical considerations have suggested this by showing that high transcription and mRNA degradation rates correlate with genes that were hypothesized to require low-noise expression (34). To the knowledge of the author, this has previously not been investigated experimentally.

In our experimental setup we defined native promoter and terminator regions from genes with reportedly extreme transcription and mRNA degradation rates and used them to control the expression of a fluorescence reporter gene outside of the original genomic locus. We assume that the transcription rate is mainly governed by the promoter and the mRNA degradation rate is mainly governed by the 5' and 3' untranslated regions, encoded in what we defined as the promoter and terminator regions, respectively. As mentioned before in this thesis, such an experimental approach is limited by potential context-dependent effects in the expression of the native gene or (post)transcriptional regulatory properties of the native coding sequence. Despite these limitations we could observe the expected trend towards high or low noise for four out of our five candidates (Figure 53). For the two extreme cases, *BUR6* on the low-noise end and *HSP26* on the high-noise end, these observations supported the hypothesis that links noise to gene function.

*BUR6* encodes a transcription factor participating in the formation of the preinitiation complex. It is an essential gene, making it likely that strong deviations from the optimal Bur6 abundance could have detrimental effects on a cell. Moreover, as a subunit of a heterodimer Bur6 levels should preferably be in a 1:1 ratio with its binding partner Ncb2. Noisy expression would result in one of the two subunits being in excess, which essentially would mean a waste of the resources that went into producing it. Noteworthy, based on the results in (85) we calculated that *NCB2* is expressed with above average transcription and mRNA degradation rates as well, indicating that *NCB2* noise is controlled in a similar fashion as *BUR6* noise (data not shown).

On the opposite extreme we found that the *HSP26* promoter and terminator conferred high expression noise of the reporter gene. *HSP26* encodes a chaperone that prevents aggregation of unfolded proteins. Our observation is in accordance with the hypothesis that stress response genes are selected for high noise, creating a phenotypic heterogeneity that serves for bet-hedging in a clonal population of cells. For such genes it could be assumed that noise control is dynamic, with high noise for bet-hedging when environmental stress is absent and low noise for an efficient response when a stressor is present. Preliminary results show that in the construct with  $P_{HSP26}$  and  $T_{HSP26}$  driving mNeogreen expression, noise is reduced upon addition of salt (data not shown). This hints to a mechanism where cells can transiently adjust gene expression noise in response to the environment, by tuning transcription and mRNA degradation rates according to react to stress.

## Discussion

This is an interesting notion to consider in the context of any sort of single-cell experiment that compares protein abundances under different treated and untreated conditions. Potentially, a number of physiologically relevant changes remain undetected because they occur on the level of protein variance without significant effects on mean protein abundance. In that respect it might be worthwhile to (re)analyze such data.

### 3.5 Conclusion and outlook

The noise tuner was developed using a measurement platform tested for minimal technical variation and an analysis pipeline that can filter out extrinsic noise almost entirely. We report the successful design, construction and testing of two versions of the noise tuner, capable of changing noise up to 3-fold. The mechanism is based on the differential contributions of the transcription rate and the mRNA degradation rate to the mean expression and the expression noise. The noise tuner was built with the intention to 1) reduce noise if necessary and 2) to allow noise control over a wide range to use it as a tool to investigate pathway robustness. Applied to a phenotypic stabilizer, the noise tuner allowed alterations of the precision of the mating pathway response, which we show through changes in the information transmission and the phenotypic response to pathway stimulation.

Our small study on noise from native yeast promoters and terminators should be expanded. It would be useful to characterize entire libraries of promoters and terminators according to the mean and noise they create in a reporter gene. The results could be compared with the mean and noise that would be expected from transcription and mRNA degradation rates of the corresponding genes. The results from such a study would inform about what properties to look for when selecting promoters and terminator regions for the expression of a gene of interest with a desired noise profile.

Looking further into the regulation of *HSP26* will give insights into whether yeast cells are able to reshape noise dynamically using the proposed strategy of transcription and mRNA degradation rate control via the promoter and terminator region. For improved detection of dynamic changes, a destabilized fluorescence reporter should be used in conjunction with  $P_{HSP26}$  and  $T_{HSP26}$ . It should also be tested if a strain with a low-noise promoter/ terminator pair for *HSP26* exhibits a different growth rate under stress conditions in a competition assay against wildtype *HSP26*.

Regarding gene expression control more generally, we assume that further functionalized synthetic terminator regions will improve expression systems. Functionalization does not stop at tuning of the degradation rate. There are binding motifs that change translation efficiency or facilitate translocation of the mRNA. Various applications for such terminator regions can be imagined, e.g. yeast production strains that have the synthesis of a value chemical downregulated by an inhibitor that has its mRNA translocated to the bud tip, so that inhibitor concentration is high in daughter cells but low in mother cells. This could allow mother cells to spend more energy on production whereas daughter cells first increase their cell volume and only start production when they first bud off a new daughter cell.

Furthermore, it would be interesting to see how the noise tuner performs in a synthetic pathway. We assume that synthetic pathways could benefit most from our system, since they typically lack the (known and unknown) noise control mechanisms of native pathways. The yeast *Saccharomyces cerevisiae* is a main chassis for synthetic metabolic pathways to produce numerous value chemicals (109). Engineered metabolic pathways that produce value chemicals have been reported to suffer from flux imbalances, leading to excessive intermediates and/ or the accumulation of pathway components (22). Different strategies have been developed for pathway balancing, including a range of methods to

## Discussion

optimize expression levels of metabolic enzymes, but – to the knowledge of the author – only in the context of absolute expression and not expression noise. The noise tuner could be applied to the expression of the enzymes in such a synthetic metabolic pathway to answer the question whether reducing expression noise can reduce the accumulation of intermediates and increase the product yield.

Generally, we expect the focus on noise optimization to increase along with the number and complexity of synthetic networks that are being developed for numerous applications. This work highlights the importance of noise control in individual genes as well as entire pathways and offers a tool to externally tune the variation in gene expression in such networks.

## 4 Materials and Methods

### 4.1 Chemicals and consumables

All chemicals and consumables used in this thesis are listed in the appendix (section 6.4).

#### 4.1.1 Growth media

##### **Luria broth (LB) medium**

- 10 g Bacto tryptone
- 5 g Bacto yeast extract
- 5 g NaCl

ddH<sub>2</sub>O added to a total volume of 1L. Adjusted with NaOH to pH 7. 100 µg/mL ampicillin was added for antibiotic selection.

##### **LB plates**

- 1 L LB medium
- 15 g Agar

100 µg/mL ampicillin was added for antibiotic selection.

##### **SOC medium (super optimal broth with catabolite repression)**

- 20 g Bacto tryptone
- 5 g Bacto yeast extract
- 4.8 g Magnesium sulfate
- 3.6 g Glucose
- 0.5 g NaCl
- 0.186 g KCl

ddH<sub>2</sub>O added to a total volume of 1L.

##### **Yeast peptone dextrose (YPD) medium**

- 50 g YPD powder

ddH<sub>2</sub>O added to a total volume of 1L.

##### **YPD plates**

- 1 L YPD liquid medium
- 15 g Agar

For antibiotic selection, one of the following was added: 500 µg/mL G418 (geneticin, KanMX marker), 200 µg/mL Hygromycin B (HPH marker), 100 µg/mL clonNAT (nourseothricin, NAT marker)

##### **Synthetic defined (SD) medium**

- 6.9 g YNB
- 790 mg Appropriate amino acid dropout mix
- 20 g Glucose

ddH<sub>2</sub>O added to a total volume of 1L.

##### **Low fluorescent SD medium (LD)**

- 6.9 g YNB LoFlo
- 790 mg Appropriate amino acid dropout mix
- 20 g Glucose

ddH<sub>2</sub>O added to a total volume of 1L.

## Materials and Methods

### SD dropout plates

1 L SD medium  
15 g Agar

SD dropout medium was used, lacking either histidine, uracil, leucine or tryptophan.

### 4.1.2 Buffers and stock solutions

#### LiAc solution 10x (1M, pH 7.5)

20.4 g LiAc dihydrate  
180 mL ddH<sub>2</sub>O

Adjust to pH 7.5 with 10 % acetic acid. Add ddH<sub>2</sub>O to 200 mL.

#### Tris-HCl solution 10x (100 mM, pH 8)

12.1 g Tris HCl  
900 mL ddH<sub>2</sub>O

Adjust to pH 8 with concentrated HCl. Add ddH<sub>2</sub>O to 1 L.

#### EDTA solution 10x (10 mM, pH 8)

2.92 g EDTA  
900 mL ddH<sub>2</sub>O

Adjust to pH 8 with 1 M NaOH. Add ddH<sub>2</sub>O to 1 L.

#### PLAG solution for competent yeast cells

4 mL PEG4000  
1 mL LiAc solution 1 M  
1 mL Tris-HCl 100 mM  
1 mL EDTA 10 mM  
1.5 mL Glycerol  
1.5 mL ddH<sub>2</sub>O

#### TAE buffer (50x) for gel electrophoresis

242 g Tris base  
57.1 g Glacial acetic acid  
100 mL EDTA (0.5M, pH 8)

ddH<sub>2</sub>O added to a total volume of 1L.

#### Antibiotic stock solutions

Ampicillin 1000x	100 mg/mL	in ddH <sub>2</sub> O
Hygromycin B 250x	50 mg/mL	in ddH <sub>2</sub> O
G418 200x	100 mg/mL	in ddH <sub>2</sub> O
clonNAT 1000x	100 mg/mL	in ddH <sub>2</sub> O

#### Other stock solutions

Doxycycline 100x	10 mg/mL	in ddH <sub>2</sub> O
α pheromone 100x	1875 nM	in LD
Casein 100x	200 mM	in LD

## Materials and Methods

### 4.1.3 Reaction kits

GeneJET DNA Purification Kit, ThermoFisher Scientific, Dreieich

GeneJET Plasmid Miniprep Kit, ThermoFisher Scientific, Dreieich

GeneJET Gel Extraction Kit, ThermoFisher Scientific, Dreieich

Gibson Assembly Master Mix, New England Biolabs, Frankfurt am Main

The kits were used according to the guidelines given by the manufacturers.

### 4.2 Yeast strains

*Saccharomyces cerevisiae* strains used in this thesis are listed in Table 3. All strains are derived from the haploid MATa strain SEY6210a (*MATa leu2-3,112 ura3-52 his3Δ200 trp1Δ901 lys2-801 suc2Δ9*). Mating pathway reporter strains are derived from yAA328 (Alexander Anders). Strains used for optimization of the measurements and those used to characterize the gene expression noise tuners contain reporter cassettes in well-defined loci (*URA3, HIS3, LEU2, TRP1, LYS2*). Strains that were used to tune expression noise of mating pathway genes were modified in the respective native loci.

Table 3: Yeast strains used in this work.

Strain	Genotype	Reference
yAA321	<i>LYS2::rtTA mfa2Δ::hphNT1 mfa1Δ::kITRP1 bar1Δ::natNT</i>	A. Anders
yAA322	<i>LYS2::rtTA mfa2Δ::[P<sub>Tet07-5'</sub> ACT1-mCherry- T<sub>ADH1</sub>:HIS3] mfa1Δ::kITRP1 bar1Δ::natNT</i>	A. Anders
yAA328	<i>LYS2::rtTA mfa2Δ::[P<sub>Tet07-5'</sub> ACT1-mCherry-T<sub>ADH1</sub>:HIS3] mfa1Δ::kITRP1 leu2Δ::[P<sub>GPD</sub>-mTurquoise2-T<sub>GPD</sub>:CgLEU2] bar1Δ::[P<sub>FUS1</sub>-mNeogreen-T<sub>FUS1</sub>:URA3]</i>	A. Anders
yMFM001	<i>lys2-801Δ::HPH</i>	This study
yMFM003	<i>his3-Δ200::[P<sub>Tet</sub>-mNeogreen-T<sub>ADH1</sub> HIS3]</i>	This study
yMFM006	<i>trp1-Δ901::[P<sub>Tet</sub>-mCherry-T<sub>ADH1</sub> TRP1]</i>	This study
yMFM008	<i>ura3-52::[P<sub>Tet</sub>-mNeogreen-T<sub>ADH1</sub> URA3]</i>	This study
yMFM010	<i>ura3-52::[P<sub>Tet</sub>-mCherry-T<sub>ADH1</sub> HPH]</i>	This study
yMFM011	<i>his3-Δ200::[P<sub>Tet</sub>-mNeogreen-T<sub>ADH1</sub> P<sub>Tet</sub>-mCherry-T<sub>ADH1</sub> URA3]</i>	This study
yMFM012	<i>leu2-3,112::[P<sub>GPD</sub>-mTurquoise2-T<sub>ADH1</sub> LEU2]</i>	This study
yMFM023	<i>his3-Δ200::[P<sub>Tet</sub>-mNeogreen-T<sub>ADH1</sub> P<sub>Tet</sub>-mCherry-T<sub>ADH1</sub> TRP1]</i>	This study
yMFM025	<i>his3-Δ200::[P<sub>Tet</sub>-mNeogreen-T<sub>GIC1</sub> P<sub>Tet</sub>-mCherry-T<sub>ADH1</sub> URA3]</i>	This study
yMFM027	<i>his3-Δ200::[P<sub>Tet</sub>-mNeogreen-T<sub>FZF1</sub> P<sub>Tet</sub>-mCherry-T<sub>ADH1</sub> URA3]</i>	This study

## Materials and Methods

yMFM028	<i>his3-Δ200::[P<sub>Tet</sub>-mNeongreen-T<sub>TPS1</sub> P<sub>Tet</sub>-mCherry-T<sub>ADH1</sub> URA3]</i>	This study
yMFM029	<i>his3-Δ200::[P<sub>Tet</sub>-mNeongreen-T<sub>STRSV</sub> P<sub>Tet</sub>-mCherry-T<sub>ADH1</sub> URA3]</i>	This study
yMFM030	<i>his3-Δ200::[P<sub>Tet</sub>-mNeongreen-T<sub>L2b8-t47</sub> P<sub>Tet</sub>-mCherry-T<sub>ADH1</sub> URA3]</i>	This study
yMFM031	<i>his3-Δ200::[P<sub>Tet</sub>-mNeongreen-T<sub>L2b8-a1-t47</sub> P<sub>Tet</sub>-mCherry-T<sub>ADH1</sub> URA3]</i>	This study
yMFM032	<i>his3-Δ200::[P<sub>Tet</sub>-mNeongreen-T<sub>L2bOFF1-a141-a14</sub> P<sub>Tet</sub>-mCherry-T<sub>ADH1</sub> URA3]</i>	This study
yMFM034	<i>leu2-3,112::[P<sub>GPD</sub>-yomTurquoise2-T<sub>GPD</sub> LEU2]</i>	This study
yMFM034	<i>leu2-3,112::[P<sub>GPD</sub>-yomTurquoise2-T<sub>GPD</sub> LEU2]</i>	This study
yMFM035	<i>leu2-3,112::[P<sub>GPD</sub>-yomTurquoise2-T<sub>GPD</sub> LEU2]</i>	This study
yMFM036	<i>his3-Δ200::[P<sub>Tet-5'UTR(ACT1)</sub>-mNeongreen-T<sub>ADH1</sub> HIS3]</i>	This study
yMFM037	<i>trp1-Δ901::[P<sub>Tet-5'UTR(ACT1)</sub>-mNeongreen-T<sub>ADH1</sub> TRP1]</i>	This study
yMFM038	<i>his3-Δ200::[P<sub>Tet-5'UTR(ACT1)</sub>-mCherry-T<sub>ADH1</sub> HIS3]</i>	This study
yMFM039	<i>trp1-Δ901::[P<sub>Tet-5'UTR(ACT1)</sub>-mCherry-T<sub>ADH1</sub> TRP1]</i>	This study
yMFM042	<i>his3-Δ200::[P<sub>Tet-5'UTR(ACT1)</sub>-mNeongreen-T<sub>ADH1</sub> P<sub>Tet-5'UTR(ACT1)</sub>-mCherry-T<sub>ADH1</sub> URA3]</i>	This study
yMFM043	<i>his3-Δ200::[P<sub>Tet-5'UTR(ACT1)</sub>-mNeongreen-T<sub>ADH1</sub> P<sub>Tet-5'UTR(ACT1)</sub>-mCherry-T<sub>ADH1</sub> URA3]</i>	This study
YMFM046	<i>his3-Δ200::[P<sub>Tet-5'UTR(ACT1)</sub>-mNeongreen-T<sub>L2b-OFF1-a14-T(Synth27)</sub> HIS3]</i>	This study
YMFM047	<i>his3-Δ200::[P<sub>Tet-5'UTR(ACT1)</sub>-mNeongreen-T<sub>Synth27</sub> HIS3]</i>	This study
YMFM048	<i>his3-Δ200::[P<sub>Tet-5'UTR(ACT1)</sub>-mNeongreen-T<sub>STRSV</sub> HIS3]</i>	This study
YMFM049	<i>his3-Δ200::[P<sub>Tet-5'UTR(ACT1)</sub>-mNeongreen-T<sub>L2b8-t47-T(Synth27)</sub> HIS3]</i>	This study
YMFM050	<i>his3-Δ200::[P<sub>Tet-5'UTR(ACT1)</sub>-mNeongreen-T<sub>L2b8-a1-t41-T(Synth27)</sub> HIS3]</i>	This study
YMFM054	<i>T<sub>MSG5</sub>::[T<sub>L2b8-a1-t41-T(Synth27)</sub> NAT]</i>	This study
YMFM058	<i>T<sub>FUS3</sub>::[T<sub>L2b8-a1-t41-T(Synth27)</sub> NAT]</i>	This study
YMFM060	<i>T<sub>STE2</sub>::[T<sub>L2b8-a1-t41-T(Synth27)</sub> NAT]</i>	This study
YMFM070	<i>P<sub>STE5</sub>::[P<sub>TetO7-5'ACT1</sub> KAN]</i>	This study
YMFM074	<i>T<sub>SST2</sub>::[T<sub>L2b8-a1-t41-T(Synth27)</sub> NAT]</i>	This study
YMFM079	<i>P<sub>TetO7-ACT1</sub>-mNeongreen::[P<sub>TetO2-CYC1</sub> KAN]</i>	This study

## Materials and Methods

YMFM101	$T_{L2b8-a1-t41-T(Synth27)}::[T_{FZF1} KAN]$	This study
YMFM102	$T_{L2b8-a1-t41-T(Synth27)}::[T_{GIC1} KAN]$	This study
YMFM103	$T_{L2b8-a1-t41-T(Synth27)}::[T_{TPS1} KAN]$	This study
YMFM104	$T_{L2b8-a1-t41-T(Synth27)}::[T_{ADH1} KAN]$	This study
YMFM106	$his3-\Delta200::[P_{TDH1}-mNeogreen-T_{TDH1} HIS3]$	This study
YMFM107	$his3-\Delta200::[P_{HSP26}-mNeogreen-T_{HSP26} HIS3]$	This study
YMFM109	$his3-\Delta200::[P_{HOR2}-mNeogreen-T_{HOR2} HIS3]$	This study
YMFM111	$his3-\Delta200::[P_{BUR6}-mNeogreen-T_{BUR6} HIS3]$	This study
YMFM115	$his3-\Delta200::[P_{ORM2}-mNeogreen-T_{ORM2} HIS3]$	This study

### 4.3 Plasmids

Two sorts of plasmids were used in this thesis to construct new yeast strains: Reporter plasmids contain entire fluorescence reporter cassettes, flanked by homology regions for genomic integration. Tagging plasmids contain promoter or terminator sequences, flanked by standard motifs (110) to attach homology arms via overlap PCR. All plasmids used in this thesis are described in Table 4. All plasmids contain an ori sequence that enables replication in *Escherichia coli* cells and an ampicillin resistance cassette as selection marker.

Table 4: Plasmids used in this work.

Plasmid	Description	Source
pAA151	<i>In vivo</i> tagging plasmid with TetO2-CYC1 promoter	A. Anders
pAA207	<i>In vivo</i> tagging plasmid with FZF1 terminator	A. Anders
pAA208	<i>In vivo</i> tagging plasmid with GIC1 terminator	A. Anders
pAA209	<i>In vivo</i> tagging plasmid with TPS1 terminator	A. Anders
pAA263	<i>In vivo</i> tagging plasmid with TetO7-ACT1 promoter	A. Anders
pMFM001	<i>HIS3</i> -integrative plasmid with mNeogreen driven by TetO7 <i>CYC1</i> promoter	This study
pMFM007	<i>TRP1</i> -integrative plasmid with mCherry driven by <i>CYC1</i> TetO7 promoter	This study
pMFM009	<i>HIS3</i> -integrative plasmid with mNeogreen driven by TetO7 promoter for dual-reporter assay in two loci	This study
pMFM010	<i>URA3</i> -integrative plasmid with mCherry driven by TetO7 promoter for dual reporter assay in two loci	This study
pMFM011	<i>HIS3</i> -integrative plasmid with converging mNeogreen and mCherry genes driven by TetO7 promoters for dual reporter assay in one locus	This study
pMFM017	<i>LEU2</i> -integrative plasmid with mTurquoise2 driven by <i>GPD</i> promoter	This study
pMFM022	<i>URA3</i> -integrative plasmid with mNeogreen driven by TetO7 promoter	This study
pMFM027	<i>LEU2</i> -integrative plasmid with codon-optimized mTurquoise2 driven by <i>GPD</i> promoter	This study
pMFM034	<i>HIS3</i> -integrative plasmid with mNeogreen driven by TetO7 promoter	This study
pMFM035	<i>HIS3</i> -integrative plasmid with mNeogreen driven by TetO7 <i>ACT1</i> promoter	This study

## Materials and Methods

pMFM036	<i>HIS3</i> -integrative plasmid with mCherry driven by TetO7 <i>ACT1</i> promoter	This study
pMFM037	<i>TRP1</i> -integrative plasmid with mCherry driven by TetO7 <i>ACT1</i> promoter	This study
pMFM038	<i>HIS3</i> -integrative plasmid with converging mNeogreen and mCherry genes driven by TetO7 <i>ACT1</i> promoters for dual reporter assays	This study
pMFM042	<i>HIS3</i> -integrative plasmid with mNeogreen driven with TetO7 <i>ACT1</i> promoter and T(Synth27) terminator	This study
pMFM044	<i>HIS3</i> -integrative plasmid with mNeogreen driven with TetO7 <i>ACT1</i> promoter and L2bOFF1-a14 ribozyme and T(Synth27) terminator	This study
pMFM046	<i>HIS3</i> -integrative plasmid with mNeogreen driven with TetO7 <i>ACT1</i> promoter and sTRSV control ribozyme and T(Synth27) terminator	This study
pMFM047	<i>HIS3</i> -integrative plasmid with mNeogreen driven with TetO7 <i>ACT1</i> promoter and L2b8-t47 ribozyme and T(Synth27) terminator	This study
pMFM048	<i>HIS3</i> -integrative plasmid with mNeogreen driven with TetO7 <i>ACT1</i> promoter and L2b8-a1-t41 ribozyme and T(Synth27) terminator	This study
pMFM058	<i>In vivo</i> tagging plasmid with L2b8-a1-t41 ribozyme and T(Synth27) terminator	This study
pMFM065	<i>In vivo</i> tagging plasmid with ADH1 terminator	This study
pMFM071	<i>HIS3</i> -integrative plasmid with mNeogreen driven by <i>TDH1</i> promoter and terminator	This study
pMFM072	<i>HIS3</i> -integrative plasmid with mNeogreen driven by <i>HSP26</i> promoter and terminator	This study
pMFM074	<i>HIS3</i> -integrative plasmid with mNeogreen driven by <i>HOR2</i> promoter and terminator	This study
pMFM076	<i>HIS3</i> -integrative plasmid with mNeogreen driven by <i>BUR6</i> promoter and terminator	This study
pMFM080	<i>HIS3</i> -integrative plasmid with mNeogreen driven by <i>ORM2</i> promoter and terminator	This study

## 4.4 Molecular cloning

### 4.4.1 Primers

Primers used for sequencing, attachment of flanking homology sequences and to create cloning fragments are listed in Table 5.

Table 5: Primers used in this work.

Primer name	Sequence (5' to 3')	Source
AA159	CGAGCTCGAATTCATCG	A. Anders
AA231	CACTAACGACGAAAGACCTCAAGAACTCATTGGAACGAAATATTTAGT CGTACGCTGCAGGTCGAC	A. Anders
AA249	TTATACCGAAGGTCACGAAATTACTTTTTCAAAGCCGTAAATTTTGATCAA TCGATGAATTCGAGCTCG	A. Anders
AA301	GTTTGAACTACAAGGAAATAAGGCAGAGAAAAAGAAAGGAAAATAATAT GCGTACGCTGCAGGTCGAC	A. Anders
AA302	ATATGTATACATTGTTCTTCGGGTTGATATTTAATGATAATGATGGCTAAT CGATGAATTCGAGCTCG	A. Anders
AA307	GCCCGCATTTTTAATTCTTGTATCATAAATTCAAAAATTATATTATATCAAT CGATGAATTCGAGCTCG	A. Anders
AA382	ATAGCAAAGAAACCGTAAATTTGGACGTAAGATGATAAACAAAA GCGTACGCTGCAGGTCGAC	A. Anders

**Materials and Methods**

AA383	CTGTGTATATACTTTGTCGATAATAAGGTGTACCCTGCTTGTCTACGTTAAT CGATGAATTCGAGCTCG	A. Anders
AA385	GTCTTATAGAGAGCGGTAATAACAACCTGTCCCCTCCATATGGATTATATAC GTACGCTGCAGGTCGAC	A. Anders
AA386	TGGCGGGATGCTTTCTTTTATTATTGCATAAAATTTAGTGTATACTCTAAT CGATGAATTCGAGCTCG	A. Anders
AA393	TGGACAACGAATCAATATCTACTGCCCGGAACAGATGATGTTTCTTCCTC GTACGCTGCAGGTCGAC	A. Anders
AA394	CCTGCAGATCTGATAGTCATAAAGATTCTTTGGGAGATGCTGTCCCTTAA TCGATGAATTCGAGCTCG	A. Anders
AA395	TTGAGGTAGATACCAATCGAAGGTCCGATAAAAACCTTCCAGATGCAACC CGTACGCTGCAGGTCGAC	A. Anders
AA481	AATAGTAGCGGAAGCGATAGTAATACATTTGGTAGCGATAAGTGCACATG CGTACGCTGCAGGTCGAC	A. Anders
AA491	CTGAGGCGTTATAGTTCAATTTGGTAATTAAGATAGAGTTGTAAGATG CGTACGCTGCAGGTCGAC	A. Anders
AA493	CCGCGCTAAAAAGGAAGATACAGGATACAGCGGAAACAACITTTTAAAT GCGTACGCTGCAGGTCGAC	A. Anders
AA495	TTTCCTTTAAGAGCAGGATATAAGCCATCAAGTTTCTGAAAATCAAATGC GTACGCTGCAGGTCGAC	A. Anders
AA553	AGGACTGTTTGTGCAATTGTACCTGAAGATGAGTAAGACTCTCAATGAAA ATCGATGAATTCGAGCTCG	A. Anders
MFM009	CCATCTACGCCATAGCATCAATGGGCCCTAATTCGCGCCACTTC	This study
MFM010	GGGCCCATTGATGCTATGGCGTAGATGGCTTTCCTTTCGTTTATCTTGCC	This study
MFM011	CGTATCCGATACCTACGACTGAGGATCCAGATCTGTTTAGCTTGCCCTC	This study
MFM012	GGATCCTCAGTCGTAGGTATCGGATACGCCACACAATTATAAGCAAAGGG	This study
MFM013	GGATTGCTACCAACGAACCTCCGCTAGCCACACAATTATAAGCAAAGGG	This study
MFM014	GCTAGCGGAGTTCGTTGGTAGCAATCCCATCGATGAATTCGAGCTCG	This study
MFM015	CTATGCGATACTAGCTCAGTCTAGAGACACCGATTATTTAAAGCTGC	This study
MFM016	TCTAGACTGAGCTAGTATCGCATAGTAATTCGCGCCACTTCTAAATAAGC	This study
MFM017	GGATTGCTACCAACGAACCTCCGCTAGCTTGACTGCGCTGGCGAAGCATG	This study
MFM018	TCTAGACTGAGCTAGTATCGCATAGCAAAGCCGAATCCACCACGGTC	This study
MFM020	CGTATCCGATACCTACGACTGAGGATCCCTTCGTACGCTGCAGGTCGAC	This study
MFM107	TTTGTTTGTTTATGTGTGTTTATTTCG	This study
MFM108	GAGCTCTTATTTGTATAACTCATCC	This study
MFM109	GGATGAGTTATACAATAAGAGCTCGTGAATTTACTTTAAATCTTGC	This study
MFM110	CGAATAAACACACATAAACAAACAAAATGGTTAGAGTCTCTAAAGGCGAA GAATTG	This study
MFM113	CAAAGCCGAATCCACCAC	This study
MFM115	GTTTCGTTTCCCAGATACATGC	This study
MFM116	CCACCGAAGTCGGTGATG	This study
MFM117	GCATGTATCGGAAACGAACAGTTCGAGTTTATCATTATCAATAC	This study
MFM118	CAGCATCACCAGCTTCGGTGGTATATGTTATCTTATCTTGCGCGTAC	This study
MFM123	TGTTAATTCAGTAAATTTTCGATCTTGG	This study
MFM125	GATCGAAAATTTACTGAATTAACAATGGTTTCTAAGGGTGAAGAAGAC	This study
MFM126	AATTCGTATATAATTTAGCTATTTGCTTAACCATAACGTCAGTGAAAGC	This study
MFM127	GATCGAAAATTTACTGAATTAACAATGGTGAGCAAGGGCGAGG	This study
MFM128	AATTCGTATATAATTTAGCTATTTGCTTAGCCGCCGGTGGAGTGGCG	This study
MFM129	GCAAATAGCTAAATTATATACGAATTAATATTATG	This study
MFM149	CTAGAAATAAGAGTATCATCTTTCAAAGACACCGATTATTTAAAGCTGC	This study

**Materials and Methods**

MFM150	ACAGTTATATATATATATATATATATATATATATATATACCACCCACTCGAGCGCGCCTA GGTTAACCCATAACGTCAGTGAAAG	This study
MFM161	GCCTTAATTAACCCGGGATCCTTTGAAAGATGATACTCTTTATTTCTAGA CAG	This study
MFM166	GGATCCCCGGGTTAATTAAGGCGTGAGTAAGGAAAGAGTGAGGAACTAT C	This study
MFM167	GTCGTCAAGAGTGGTACCCATGCTAGCTGTTTTATATTTGTTGTA AAAAGT AGATAATTACTTCC	This study
MFM168	GCATGCCCTGCCCTAACCCATGGAGCTTTTTGATTAAGCCTTCTAGTCC	This study
MFM169	GATATCATCGATGAATTCGAGCTCGATACATGGGTGACCAAAGAGCG	This study
MFM170	ATGGGTACCACTCTTGACGAC	This study
MFM171	TTAGGGGCAGGGCATGC	This study
MFM172	CACAGCAACTTGATTACAC	This study
MFM173	GAACATCGATGCCTTAC	This study
MFM174	GATGGCGATCTGTAGAAG	This study
MFM184	CAGACGTATACAAAGATGCTAGCGCTTTAATAGAAATCCAAGAAAAGTGC CGTACGCTGCAGGTCGAC	This study
MFM185	ACTGTTTGTGCAATTGTACCTGAAGATGAGTAAGACTCTCAATGAAATTA TCGATGAATTCGAGCTCG	This study
MFM186	TGCTACAGGAGAGGTAAATG	This study
MFM187	CTTCGAAGAATATACTAAAAAATGAGCAGGCAAGATAAACGAAGGCAAA GCGTACGCTGCAGGTCGAC	This study
MFM189	GCTTACTGCTTTTTTCTCCCAAGATCGAAAATTTACTGAATTAACAATGCG TACGCTGCAGGTCGAC	This study
MFM193	ATATACACATGTATATATATCGTATGCTGCAGCTTTAAATAATCGGTGTCA TCGATGAATTCGAGCTCG	This study
MFM216	AATTGAACTTCAAGGAATGGCAAAGGCTTTCCTACTGACGTTATGGGTAA CGTACGCTGCAGGTCGAC	This study
MFM223	CGTGGTGGATTCCGGCTTTGGGAATAGGATATGCGACGAAGAC	This study
MFM224	GTCTTCTTCACCCTTAGAAACCATTTTGTGTTGTGTAAATTTAGTGAAGT AC	This study
MFM225	GCTTTCCTGACGTTATGGGTAAATAAAGCAATCTTGATGAGGATAATG	This study
MFM226	GCAGCTTTAAATAATCGGTGTCTTCTTAGTATATATACTGCTCAAGGGC	This study
MFM227	CGTGGTGGATTCCGGCTTTGGCAGCAGCAACTCCGTGTG	This study
MFM228	GTCTTCTTCACCCTTAGAAACCATGTTAATTTGTTTAGTTTGTGTTTGTGCTT TTTTG	This study
MFM229	GCTTTCCTGACGTTATGGGTAAAGTGACCTGGCTCTATAGTGTTG	This study
MFM230	GCAGCTTTAAATAATCGGTGTCTTCAATACAAAACCTCGATTCTAATCGG	This study
MFM235	CGTGGTGGATTCCGGCTTTGTCTCAAGTATTTTGGCACCTCGC	This study
MFM236	GTCTTCTTCACCCTTAGAAACCATTCCGAATATTGTTTTATTGTTTTATGTT TTTCC	This study
MFM237	GCTTTCCTGACGTTATGGGTAAATCCTCTAAAATCGAACATATTTGAGTA ATAATTC	This study
MFM238	GCAGCTTTAAATAATCGGTGTCTTCTTACTAGAAATTAATTTTAATCCGT TGTG	This study
MFM243	CGTGGTGGATTCCGGCTTTGGATTATTTTTCTTTAATCAACTTTTTCGGATA TCG	This study
MFM244	GTCTTCTTCACCCTTAGAAACCATTAAGAATGTAGTGATGAGTACTGTAAC TGG	This study
MFM245	GCTTTCCTGACGTTATGGGTAAAGCAGAACGGGGCGATGTAAC	This study
MFM246	GCAGCTTTAAATAATCGGTGTGAGAATCCTCGAGCTAGCCCTTC	This study

## Materials and Methods

MFM259	CGTGGTGGATTCCGGCTTTGGATTAATTTAGGGTCCCCGGC	This study
MFM260	GTCTTCTTCACCCTTAGAAACCATGATGAAACTGTTTCTTATGCGTTTTTAT AATG	This study
MFM261	GCTTTCACCTGACGTTATGGGTAAATTGGCTCTGCCTATACGCATATATG	This study
MFM262	GCAGCTTTAAATAATCGGTGTGAGAGAAATATAAAGAAAGATGATGATGTA AGGTATAC	This study
MFM271	TGGTTTCTTCAGAAACAACGGTG	This study

### 4.4.2 Polymerase Chain Reaction (PCR)

Three PCR protocols and polymerases were used in this study for different purposes.

#### PCR to create DNA fragments for cloning

Reaction mix:

10 $\mu$ L	5X Phusion HF buffer (NEB)
2.5 $\mu$ L	Forward primer (10 $\mu$ M)
2.5 $\mu$ L	Reverse primer (10 $\mu$ M)
1 $\mu$ L	dNTPs(10 mM)
1 $\mu$ L	Template DNA (1 to 10 ng/ $\mu$ L)
0.5 $\mu$ L	Phusion DNA polymerase (NEB)
32.5 $\mu$ L	ddH <sub>2</sub> O

Thermocycler settings:

30 cycles	[	30 s	98 °C
		10 s	98 °C
		30 s	T <sub>m</sub> + 3 °C
		30 s / kb	72 °C
		5 min	72 °C
		hold	8 °C

#### Colony PCR to identify positive transformants

Reaction mix:

10 $\mu$ L	MangoMix (Taq polymerase master mix, Bioline)
1 $\mu$ L	Forward primer (10 $\mu$ M)
1 $\mu$ L	Reverse primer (10 $\mu$ M)
8 $\mu$ L	ddH <sub>2</sub> O

Thermocycler settings:

30 cycles	[	5 min	95 °C
		30 s	95 °C
		30 s	T <sub>m</sub> - 5 °C
		1 min / kb	72 °C
		5 min	72 °C
		hold	8 °C

#### GXL PCR to create fragments for *in vivo* tagging

Reaction mix:

10 $\mu$ L	PrimeSTAR GXL Buffer (Takara)
1 $\mu$ L	dNTPs (10 mM)
1 $\mu$ L	Forward primer (10 $\mu$ M)
1 $\mu$ L	Reverse primer (10 $\mu$ M)
1 $\mu$ L	Template DNA (1 to 10 ng/ $\mu$ L)

## Materials and Methods

1  $\mu$ L PrimeSTAR GXL DNA Polymerase (Takara)  
35  $\mu$ L ddH<sub>2</sub>O

Thermocycler settings:

30 cycles [ 10 s 98 °C  
15 s 55 °C  
1 min / kb 68 °C  
hold 8 °C

PCRs were performed in peqSTAR thermocyclers (PEQLAB). Produced fragments were analyzed by gel electrophoresis with TAE gels containing 1 % agarose. DNA fragments were purified either with the GeneJET PCR Purification Kit or from gel slices with the GeneJET Gel Extraction Kit.

### 4.4.3 Gibson Assembly

Gibson assembly was used to create plasmids from up to four DNA fragments created by Phusion PCR. Fragments had flanking sequences of 15 to 25 bp homologous to each other. The reaction was performed according to the manufacturer's protocol. Essentially, we calculated the molar mass of 100 ng of the backbone DNA fragment and the masses of insert DNA fragments in a molar ratio of 2:1 or 3:1 to the backbone, depending on insert length. The corresponding volumes of PCR-purified DNA fragments were transferred to a 1.5 mL tube on ice and ddH<sub>2</sub>O was added to a total volume of 10  $\mu$ L. The total molar mass of fragments was between 0.02 pmol and 1 pmol. 10  $\mu$ L of Gibson Assembly Master Mix (NEB) was added and reaction mix was transferred to a heat block preheated to 50 °C and incubated for 1 hour. 1  $\mu$ L of a 1:4 dilution of the reaction mixture in ddH<sub>2</sub>O was used for transformation.

### 4.4.4 Restriction-Ligation Cloning

Classic cloning based on restriction and ligation was used when appropriate cut-sites were present or when sequences were cloned that created stable secondary structures, such as the ribozyme sequences. Restriction digestion was also used to linearize plasmids for genomic integration (PmeI) and to remove methylated template plasmid DNA from PCRs (DpnI). The restriction enzymes used in this study are AvrII, DpnI, NotI, PmeI, PspOMI, and XhoI and were all purchased from New England Biolabs in their high-fidelity (HF) version, wherever available. All enzymes had 100 % activity in CutSmart buffer (NEB) and were heat-inactivatable at 65 °C with the exception of AvrII.

#### Restriction digestion

1  $\mu$ g DNA  
2  $\mu$ L CutSmart buffer  
0.5  $\mu$ L Restriction enzyme 1  
0.5  $\mu$ L Restriction enzyme 2  
To 20  $\mu$ L ddH<sub>2</sub>O

Reagents were mixed in a 1.5 mL tube and incubated on a heat block at 37 °C for 1 hour, followed by heat-inactivation at 65 °C for 20 min.

To prevent re-ligation of backbone fragments with their inserts, digested backbone DNA was either gel-extracted or treated with phosphatase to dephosphorylate the restriction sites.

## Materials and Methods

### Phosphatase treatment of backbone digest

17 $\mu$ L	Digestion reaction
2 $\mu$ L	Antarctic phosphatase reaction buffer (NEB)
1 $\mu$ L	Antarctic phosphatase (NEB)

The reaction mixture was incubated at 37 °C for 30 min.

Assembly of restricted DNA fragments was done with T4 DNA Ligase purchased from ThermoFisher Scientific.

### Ligation

0.02 pmol	Digested backbone DNA
0.06 pmol	Digested insert DNA
2 $\mu$ L	T4 DNA Ligase Buffer (ThermoFisher Scientific)
1 $\mu$ L	5 % PEG4000
0.5 $\mu$ L	T4 DNA Ligase (ThermoFisher Scientific)
To 20 $\mu$ L	ddH <sub>2</sub> O

Reaction mixture was incubated for 1 hour at room temperature. 1  $\mu$ L of the ligation mixture was used for transformation.

#### 4.4.5 Electrocompetent *E. coli* cells

Highly efficient electrocompetent *E. coli* cells were prepared according to (111). Briefly, 200  $\mu$ L of *E. coli* NEB Turbo cell suspension was grown to an OD between 0.4 and 0.6 in LB medium, chilled on ice water and pelleted by centrifugation. The supernatant was discarded and the pellet resuspended in 40 mL ice-cold ddH<sub>2</sub>O. The cell suspension was transferred into two 50 mL tubes. 10 mL of ice-cold high-density solution containing 20% glycerol and 1.5 % mannitol were added to each of the two tubes to form a second layer below the cell suspension. Tubes were centrifuged with 2000 g at 4 °C for 15 min. The layer of cells that formed during centrifugation was removed, resuspended in 200  $\mu$ L of the ice-cold glycerol/ mannitol solution and aliquoted into 1.5 mL tubes with 50  $\mu$ L per tube. Cells were then either used directly for electroporation or frozen in liquid nitrogen prior to storage at -80 °C.

#### 4.4.6 *E. coli* transformation

50  $\mu$ L of electrocompetent cells were mixed with 1 to 5  $\mu$ L DNA and transferred into a pre-chilled electroporation cuvette on ice. The cuvette was placed into a MicroPulser electroporator (BIO-RAD). After the electric pulse, cells were recovered in 950  $\mu$ L SOC medium pre-heated to 37 °C and incubated for one hour prior to plating on appropriate medium and incubation over night at 37 °C.

### 4.5 Yeast strain development

All yeast strains described in this thesis were built by genomic integration of linear DNA sequences derived from plasmids. Plasmids contained no yeast ori to prevent them from being proliferated.

#### 4.5.1 Competent yeast cells

Heat shock competent yeast cells were prepared by growing an overnight culture in YPD and inoculating 50 mL YPD medium in an Erlenmeyer flask to an initial OD of 0.05. The day culture was grown at 30 °C and 200 rpm shaking to an OD between 0.8 and 1. Cell suspension was transferred into a 50 mL plastic tube and centrifuged at 4000 rpm for 2 minutes at room temperature. The supernatant was discarded and the cell pellet was resuspended in 2 mL PLAG solution. 200  $\mu$ L denatured salmon

## Materials and Methods

sperm DNA were added and 200  $\mu\text{L}$  aliquots of the cell suspension were transferred into chilled 1.5 mL tubes. Aliquots were either immediately used for transformation or frozen at  $-80\text{ }^{\circ}\text{C}$ .

### 4.5.2 Genomic integration

All genes and genetic elements were integrated genomically by transforming linear DNA with flanking sequences that are homologous to the desired integration locus in the genome. For integration of reporter genes, linear DNA derived from PmeI-digested plasmids containing the desired genes and homology regions (ca. 200 bp). For promoter and terminator exchanges, linear DNA derived from PCR fragments created with primers that had overhangs (50 nt) homologous to the desired locus.

20  $\mu\text{L}$  of linear DNA were incubated with an aliquot of competent yeast cells at  $35\text{ }^{\circ}\text{C}$  with shaking for 1 to 2 hours. Cells were shifted to  $42\text{ }^{\circ}\text{C}$  without shaking for 15 minutes and then plated on the appropriate agar plates to select for prototrophic cells. For antibiotic resistance selection, cells were incubated in 5 mL YPD for at least 2 hours prior to plating on the appropriate antibiotic agar plates. After two days, colonies were re-streaked onto selective agar and single colonies from those plates were streaked onto YPD agar.

Colonies from the YPD plate were used to inoculate a YPD overnight culture (200 rpm,  $30\text{ }^{\circ}\text{C}$ ). For long-term storage of strains, 1 mL of the overnight culture was mixed with 500  $\mu\text{L}$  60 % glycerol in a cryo tube and stored at  $-80\text{ }^{\circ}\text{C}$ .

### 4.6 Growth conditions

All experiments were performed with cells grown in low fluorescent synthetic defined media (LD). Cells were inoculated from single colonies and grown at  $30\text{ }^{\circ}\text{C}$  with shaking over night for at least 16 hours. Day cultures were inoculated to an initial OD of 0.05 and grown at  $30\text{ }^{\circ}\text{C}$  with shaking for 5 to 6 hours to an OD between 0.4 and 0.8 prior to the measurement. For mating pathway stimulation experiments, media was supplied with 2  $\mu\text{M}$  casein (Sigma-Aldrich) and cells were grown to at least OD 0.2 prior to stimulation with pheromone. Single timepoint measurements were conducted 3 hours after addition of  $\alpha$ -pheromone (Sigma-Aldrich). For time course experiments, data was acquired immediately after pheromone stimulation to up to 12 hours after stimulation.

For strains harboring genes controlled by Tet promoters and inducible ribozymes, doxycycline and theophylline were added to the final concentrations as indicated in the results section. Due to the low solubility in aqueous solutions, LD containing theophylline was prepared by adding the appropriate amount of theophylline powder directly to the media.

### 4.7 Fluorescent proteins

Table 6 lists the properties of the three fluorescent proteins that have been used to acquire the data presented in this thesis. mTurquoise2 is a third-generation cyan variant of the original GFP from *Aequorea victoria*, with extremely high quantum yield (112). mNeongreen emits light in the green to yellow spectrum and is derived from LanYFP, a yellow fluorescent protein found in the fish *Branchiostoma lanceolatum*. To create the monomeric mNeongreen, 21 substitutions were introduced into LanYFP through a directed evolution approach (113). mCherry is a monomeric red-fluorescent variant of DsRed, originally isolated from the anemone *Discosoma*. mCherry was modified for increased photostability and excitation coefficient (114).

## Materials and Methods

Table 6: Properties of fluorescent proteins used in this thesis.  $Ex_{max}$ : Maximum of excitation spectrum;  $Em_{max}$ : Maximum of emission spectrum; EC: Extinction coefficient; QY: Quantum yield. Maturation times (time to half-maximum fluorescence) are not indicated because values were either not reported or measured in other organisms than yeast.

Fluorescence protein	Source organism	$Ex_{max}$ (nm)	$Em_{max}$ (nm)	EC ( $mM^{-1}\cdot cm^{-1}$ )	QY	Reference
<b>mTurquoise2</b>	<i>Aequorea victoria</i>	434	474	30	0.93	Goedhardt et al. 2012
<b>mNeongreen</b>	<i>Branchiostoma lanceolatum</i>	506	517	116	0.80	Shaner et al. 2013
<b>mCherry</b>	<i>Discosoma sp.</i>	587	610	72	0.22	Shaner et al. 2004

### 4.8 Flow cytometry

The main part of the data presented in this thesis has been acquired using a Fortessa Special Order flow cytometer (BD Biosciences), which allows rapid measurements (up to 10,000 cells per second) of multiple parameters with minimal sample preparation. Through the sample injection port, a cell suspension is either aspirated from a tube or a 96-well plate from a high-throughput sampler. From the sample injection port, the cells pass into the flow cell. Here, sheath fluid is added through a second input and the suspension is pressed through a nozzle. In this process termed hydrodynamic focusing, cells are aligned in single file before they pass the laser beam(s). Fluorescence and refracted light are channeled towards the detectors and split into ranges of wavelengths. This is done using a system of filters with different properties. Band pass filters only pass a certain window of the spectrum. Short pass filters block light above a certain wavelength so that only shorter wavelengths can pass. The opposite is true for long pass filters. Dichroic filters split the light according to wavelength: A dichroic long pass filter will e.g. pass light above 600 nm but will reflect light below 600 nm. These filters are organized in a setup that allows isolation of fluorescent signals that are expected from e.g. different FPs in a sample.

Figure 55 shows a setup of lasers and filters that has been used in the experiments described in this thesis.

## Materials and Methods

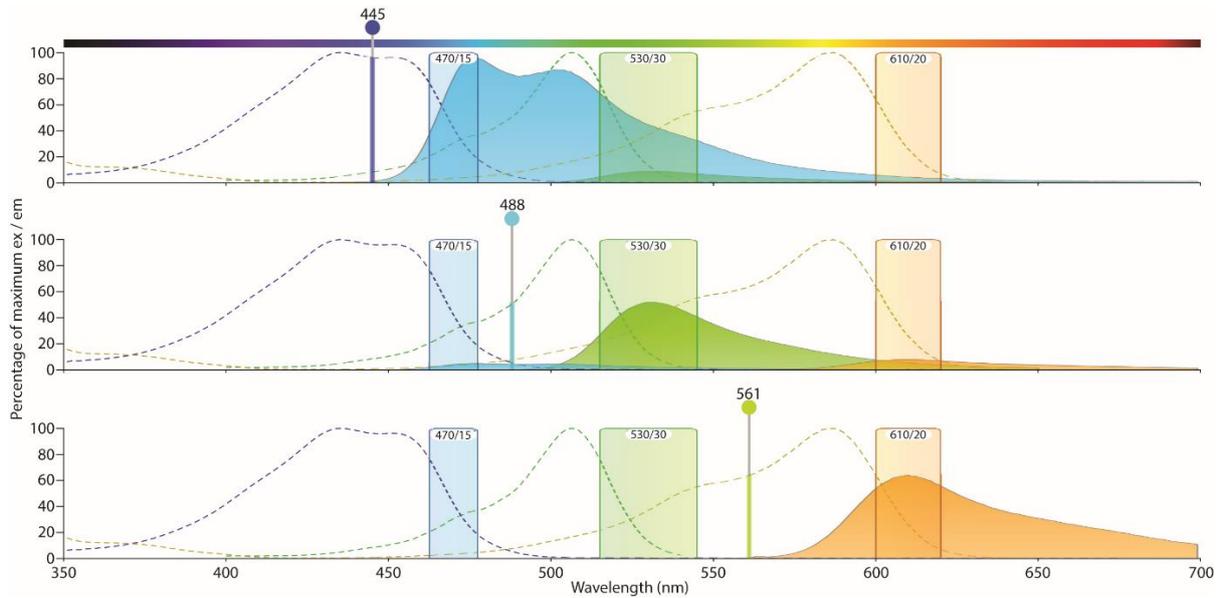


Figure 55: Excitation spectra (dashed lines) and emission spectra (full peaks) for fluorescent proteins. Laser wavelengths are indicated with vertical lines; boxes indicate acquisition filters. Top: mTurquoise2 (blue) is excited at 445 nm. The mTurquoise2 fluorescence signal is acquired in a 15-nm window around 470 nm. Middle: mNeogreen is excited at 488 nm and fluorescence is acquired in a 30-nm window around 530 nm. Bottom: mCherry is excited at 561 nm and fluorescence signals are acquired in a 20-nm band around 610 nm. Adapted from the BD Biosciences spectrum viewer ([www.bdbiosciences.com/us/s/spectrumviewer](http://www.bdbiosciences.com/us/s/spectrumviewer), accessed 14.01.2018)

A photon inside a PMT hits a photocathode which then ejects a primary electron. This electron in turn is accelerated towards another electrode, termed dynode. When it hits the dynode, more electrons are ejected, which are accelerated towards the next dynode. The process continues until enough secondary electrons are ejected to create a measurable current that is picked up at the anode at the end of the tube.

Gating of the cells can be done in real-time. If not otherwise stated, one gate was applied that excluded events that hit the upper and lower detector limits. 50,000 cells were analyzed unless otherwise stated.

### 4.9 Microscopy

Microscopic images were acquired with a wide-field microscope (Olympus MT20) equipped with a 150W mercury-xenon lamp, a motorized stage, a 40x dry objective (Olympus UPLSAPO N/A=0.95) and a EM-CCD camera (Hamamatsu C9100). mNeogreen fluorescence was acquired using a 474/23 excitation filter and a 525/45 emission filter, mCherry signal with 562/40 and 641/75 filters, and mTurquoise2 with 434/17 and 480/40 filters. Cell suspensions were transferred to a 96-well glass-bottom plate (Greiner Bio-One) coated with concanavalin A (Sigma-Aldrich) and centrifuged for 1 min at 4000 rpm to attach the cells to the bottom of the wells. Image acquisition was started immediately after alpha-factor addition and repeated periodically at defined time intervals over the course of several hours.

### 4.10 Calculation of intrinsic and extrinsic noise

For most experiments presented in this thesis, we calculated the robust CV as a measure of noise (see main text). To distinguish between intrinsic and extrinsic noise in dual-reporter strains, we used the mathematical method originally proposed by Elowitz and colleagues (2). For experiments with strains that harbored only one fluorescent noise reporter, we used the reduced gate size method as described by Newman and colleagues (18).

#### 4.10.1 Dual reporter assay

For a measurement of two fluorescence intensities  $x$  and  $y$  in a cell, the total normalized variance is given as the sum of extrinsic and intrinsic normalized variance, or

$$\eta_{tot}^2 = \eta_{ext}^2 + \eta_{int}^2$$

where the intrinsic component is

$$\eta_{int}^2 = \frac{\langle (x - y)^2 \rangle}{2\langle x \rangle \langle y \rangle}$$

and the extrinsic component is

$$\eta_{ext}^2 = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\langle x \rangle \langle y \rangle}$$

Taking the square root gives the corresponding noise values for  $\eta_{ext}$  and  $\eta_{int}$ .

#### 4.10.2 Reduced gate size analysis

The analysis was performed as described in the main text. The center of the gate was defined as the centroid of the population. Robust CVs were calculated for different gate radii using a script written in MATLAB (MathWorks).

#### 4.11 Stochastic simulation

The simulation is based on the analytical model for two-stage gene expression developed by Sharezaei and Swain (80). The model describes the probability distribution of protein numbers in a population of cells, based on the underlying probabilities of mRNA synthesis  $v_0$ , mRNA degradation  $d_0$ , protein synthesis  $v_1$  and protein degradation ( $d_1$ , see Figure 36 in the results section). Given the assumption that the protein lifetime is much longer than the lifetime of a transcript ( $\frac{d_0}{d_1} = \gamma \gg 1$ ), a probability distribution can be derived for the protein number  $n$  at a given time:

$$P_n = \frac{\Gamma(a + n)}{\Gamma(n + 1)\Gamma(a)} \left(\frac{b}{1 + b}\right)^n \left(1 - \frac{b}{1 + b}\right)^a$$

where the production and degradation probabilities are reformulated as mRNA production per protein lifetime  $a = \frac{v_0}{d_1}$  and the translational burst size  $b = \frac{v_1}{d_0}$ .  $\Gamma()$  denotes the gamma function to take non-integer values into account.

To confirm that the CV becomes independent of the translational burst size  $b$  as long as  $b \gg 1$ , we performed stochastic simulations of the four processes above using the Gillespie formalism (115) as was also done by Sharezaei and Swain. We fixed the protein half-life to be two hours, which is a conservative estimate based on of GFP data (116), resulting in the protein degradation rate  $d_1 = \frac{\log 2}{7200} \text{ s}^{-1}$ . We varied the parameter  $a$ , thereby achieving a corresponding range in values of the mRNA synthesis rate  $v_0$ . Up to a normalization, the set of values for  $a$  and  $b$  were chosen by non-negative matrix factorization (using the function *nmmf* in MATLAB) of the matrix of experimental median fluorescent values (Figure 31 B). Since the mean is given theoretically by  $ab$ , this provides a conversion from the doxycycline and theophylline concentrations used in the experiment to (up to a factor) the corresponding  $a$  and  $b$  values. The normalization factors were chosen to match the experimental range

## Materials and Methods

of rCV and to give reasonable values for the mean number of molecules and the burst size. Finally, the protein synthesis rate was fixed  $v_1 = 0.11 \text{ s}^{-1}$  such that the range of mRNA half-lives fell within the biologically expected range (minutes).

The resultant values were

$$\begin{aligned} a &= (1 \quad 1.32 \quad 2.33 \quad 7.23 \quad 18.53 \quad 30.06 \quad 36.82 \quad 40.80), \\ b &= (10 \quad 12 \quad 13.98 \quad 18.09 \quad 25.32 \quad 38.43 \quad 58.18 \quad 79.48), \end{aligned}$$

which corresponded to mRNA synthesis rates (in units of  $\text{hour}^{-1}$ ) of

$$v_0 = (0.35 \quad 0.46 \quad 0.81 \quad 2.5 \quad 6.42 \quad 10.42 \quad 12.76 \quad 14.14)$$

and mRNA half-lives (in units of mins) of

$$\frac{\log 2}{d_0} = (1.05 \quad 1.26 \quad 1.47 \quad 1.90 \quad 2.66 \quad 4.04 \quad 6.11 \quad 8.35).$$

Simulations were performed using these values and the properties of the steady-state protein distribution were compared to the experimental result (Figure 31 B and Figure 37).

### 4.12 Calculation of mutual information

Mutual information describes how much information (in bits) from an input distribution (here: pheromone concentrations) can be inferred from an output distribution (here: fluorescence intensities). In the experiments described in this thesis only the output distribution is known (continuous data sets of fluorescence intensities), whereas the input can only be described by discrete values (20 pheromone concentrations ranging from 0 nM to 18.75 nM). We used the formula derived in (117) to calculate mutual information  $I$  between a discrete dataset  $X$  and a continuous dataset  $Y$ :

$$I(X, Y) = \sum_x \int \log \frac{\mu(x, y)}{p(x)\mu(y)} dy$$

with the discrete probability function  $p(x)$  and the continuous densities  $\mu(x, y)$  and  $\mu(y)$ . We calculated  $I(X, Y)$  for a moving window of four experimental subsets of subsequent pheromone concentrations, considering pheromone/ fluorescence data of 45,000 cells for each subset. For the 20 pheromone concentrations tested, we end up with 17 values for mutual information (first window: 1 to 4; last window: 17 to 20). We note that the absolute values of mutual information calculated is somewhat arbitrary as it depends heavily on the number of different inputs measured: A high-resolution dose-response with many input pheromone concentrations contains only little mutual information when adjacent concentrations are compared. In contrast, when only off- and on-state of the pathway are compared (0 nM and 18.75 nM pheromone) the mutual information would be very high but would not allow qualitative distinctions between the high-noise and the low-noise setting.

### 4.13 Morphology analysis

Between 100 and 300 cells per noise setting and pheromone concentration were categorized. Categorization was not based on morphologies at a fixed timepoint but on *changes* in morphology over time (typically 8 hours) after pheromone stimulation. Including this information allowed classification of cells not only by size and shape but also by behavior. Table 7 lists the categories and predictors that were defined prior to the analysis.

## Materials and Methods

Table 7: Parameters for classification of yeast mating phenotypes.

	<b>Unstimulated cells</b>	<b>Chemotropic cells</b>	<b>Shmooring cells</b>
<b>Size</b>	Small	Big	Intermediate
<b>Shape</b>	Round	Elongated	Pear-like
<b>Behavior</b>	Bud and divide	Unidirectional increase in size. Do not bud and divide.	Develop one or more pointy protrusions. Do not increase in size. Do not bud and divide

To reduce bias, the analysis was done blinded with original filenames of the time-lapse movies replaced by random numbers.

### 4.14 Software

Acquisition of flow cytometric data and initial gating was done using the FACSDIVA software by BD Biosciences. For visualization of raw data and export of .fcs files to .csv files we used the FlowJo software (BD Biosciences). Crude flow cytometric data processing was done with MS Excel (Microsoft).

Data normalization, reduced gate size analysis, stochastic simulations, and calculation of mutual information were done in MATLAB (Mathworks).

Acquisition of microscopic images was done with the NIS-Elements Advanced Research Software (Nikon). Analysis of time-lapse movies was done with R (R Foundation) and ImageJ (NIH).

Adobe Illustrator was used to design figures and cartoons.

Sequence alignments were done using MAFFT multiple sequence alignment software version 7. (118)

## 5 References

1. M. B. Elowitz, S. Leibler, A synthetic oscillatory network of transcriptional regulators. *Nature*. **403**, 335–338 (2000), doi:10.1038/35002125.
2. M. B. Elowitz, A. J. Levine, E. D. Siggia, P. S. Swain, Stochastic Gene Expression in a Single Cell. *Science*. **297**, 1183–1186 (2002), doi:10.1126/science.1070919.
3. J. M. Raser, E. K. O'Shea, Control of Stochasticity in Eukaryotic Gene Expression. *Science*. **304**, 1811–1814 (2004), doi:10.1126/science.1098641.
4. T. G. Fazio *et al.*, Widespread Collaboration of Isw2 and Sin3-Rpd3 Chromatin Remodeling Complexes in Transcriptional Repression. *Molecular and Cellular Biology*. **21**, 6450–6460 (2001), doi:10.1128/MCB.21.19.6450-6460.2001.
5. G. Mizuguchi *et al.*, ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1 Chromatin Remodeling Complex. *Science*. **303**, 343–348 (2004), doi:10.1126/science.1090701.
6. H. E. Mischo, N. J. Proudfoot, Disengaging polymerase: terminating RNA polymerase II transcription in budding yeast. *Biochimica et biophysica acta*. **1829**, 174–185 (2013), doi:10.1016/j.bbagr.2012.10.003.
7. Z. Moqtaderi, J. V. Geisberg, Y. Jin, X. Fan, K. Struhl, Species-specific factors mediate extensive heterogeneity of mRNA 3' ends in yeasts. *Proceedings of the National Academy of Sciences of the United States of America*. **110**, 11073–11078 (2013), doi:10.1073/pnas.1309384110.
8. J. V. Geisberg, Z. Moqtaderi, X. Fan, F. Ozsolak, K. Struhl, Global Analysis of mRNA Isoform Half-Lives Reveals Stabilizing and Destabilizing Elements in Yeast. *Cell*. **156**, 812–824 (2014), doi:10.1016/j.cell.2013.12.026.
9. I. Gupta *et al.*, Alternative polyadenylation diversifies post-transcriptional regulation by selective RNA-protein interactions. *Molecular systems biology*. **10**, 1–11 (2014), doi:10.1002/msb.135068.
10. D.-K. Ro *et al.*, Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*. **440**, 940–943 (2006), doi:10.1038/nature04640.
11. J. Minic *et al.*, Functional expression of olfactory receptors in yeast and development of a bioassay for odorant screening. *FEBS Journal*. **272**, 524–537 (2005), doi:10.1111/j.1742-4658.2004.04494.x.
12. D. M. Chudakov, M. V. Matz, S. Lukyanov, K. A. Lukyanov, Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiological Reviews*. **90**, 1103–1163 (2010), doi:10.1152/physrev.00038.2009.
13. A. Sanchez, I. Golding, Genetic Determinants and Cellular Constraints in Noisy Gene Expression. *Science*. **342**, 1188–1193 (2013).
14. G. Chalancon *et al.*, Interplay between gene expression noise and regulatory network architecture. *Trends in genetics : TIG*. **28**, 221–232 (2012), doi:10.1016/j.tig.2012.01.006.
15. S. Marguerat, J. Bähler, Coordinating genome expression with cell size. *Trends in Genetics*. **28**, 560–565 (2012), doi:10.1016/j.tig.2012.07.003.
16. C. J. Zopf, K. Quinn, J. Zeidman, N. Maheshri, Cell-cycle dependence of transcription dominates noise in gene expression. *PLoS computational biology*. **9**, e1003161 (2013), doi:10.1371/journal.pcbi.1003161.
17. A. Raj, A. van Oudenaarden, Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*. **135**, 216–226 (2008), doi:10.1016/j.cell.2008.09.050.
18. Newman, John R. S. *et al.*, Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature*. **441**, 840–846 (2006), doi:10.1038/nature04785.
19. E. Dacheux *et al.*, Translation initiation events on structured eukaryotic mRNAs generate gene expression noise. *Nucleic Acids Research*. **45**, 6981–6992 (2017), doi:10.1093/nar/gkx430.
20. A. Bar-Even *et al.*, Noise in protein expression scales with natural protein abundance. *Nat Genet*. **38**, 636–643 (2006), doi:10.1038/ng1807.
21. L. Keren *et al.*, Massively Parallel Interrogation of the Effects of Gene Expression Levels on Fitness. *Cell*. **166**, 1282–1294.e18 (2016), doi:10.1016/j.cell.2016.07.024.

## References

22. J. A. Jones, Ö. D. Toparlak, Koffas, Mattheos A G, Metabolic pathway balancing and its role in the production of biofuels and chemicals. *Current opinion in biotechnology*. **33**, 52–59 (2015), doi:10.1016/j.copbio.2014.11.013.
23. M. Jeschek, D. Gerngross, S. Panke, Rationally reduced libraries for combinatorial pathway optimization minimizing experimental effort. *Nature communications*. **7**, 11163 (2016), doi:10.1038/ncomms11163.
24. S. L. Holland, T. Reader, P. S. Dyer, S. V. Avery, Phenotypic heterogeneity is a selected trait in natural yeast populations subject to environmental stress. *Environmental microbiology*. **16**, 1729–1740 (2014), doi:10.1111/1462-2920.12243.
25. Z. Wang, J. Zhang, Impact of gene expression noise on organismal fitness and the efficacy of natural selection. *Proceedings of the National Academy of Sciences*. **108**, E67-E76 (2011), doi:10.1073/pnas.1100059108.
26. B. Lehner, Selection to minimise noise in living systems and its implications for the evolution of gene expression. *Mol Syst Biol*. **4**, 1389 (2008), doi:10.1038/msb.2008.11.
27. W. J. Blake, M. Kaern, C. R. Cantor, J. J. Collins, Noise in eukaryotic gene expression. *Nature*. **422**, 633–637 (2003).
28. W. J. Blake *et al.*, Phenotypic consequences of promoter-mediated transcriptional noise. *Molecular cell*. **24**, 853–865 (2006), doi:10.1016/j.molcel.2006.11.003.
29. K. F. Murphy, G. Balazsi, J. J. Collins, Combinatorial promoter design for engineering noisy gene expression. *PNAS*. **104**, 12726–12731 (2007).
30. A. Aranda-Díaz, K. Mace, I. Zuleta, P. Harrigan, H. El-Samad, Robust Synthetic Circuits for Two-Dimensional Control of Gene Expression in Yeast. *ACS synthetic biology*. **6**, 545–554 (2017), doi:10.1021/acssynbio.6b00251.
31. K. F. Murphy, R. M. Adams, X. Wang, G. Balázs, J. J. Collins, Tuning and controlling gene expression noise in synthetic gene networks. *Nucleic Acids Research*. **38**, 2712–2726 (2010), doi:10.1093/nar/gkq091.
32. M. Thattai, A. van Oudenaarden, Intrinsic noise in gene regulatory networks. *PNAS*. **98**, 8614–8619 (2001).
33. O. Shalem *et al.*, Measurements of the Impact of 3' End Sequences on Gene Expression Reveal Wide Range and Sequence Dependent Effects. *PLoS Comput Biol*. **9**, 1–12 (2013), doi:10.1371/journal.pcbi.1002934.
34. H. B. Fraser, A. E. Hirsh, G. Giaever, J. Kumm, M. B. Eisen, Noise Minimization in Eukaryotic Gene Expression. *Plos Biol*. **2**, 834–838 (2004), doi:10.1371/journal.pbio.0020137.
35. A. Singh, Negative feedback through mRNA provides the best control of gene-expression noise. *IEEE transactions on nanobioscience*. **10**, 194–200 (2011).
36. P. S. Swain, Efficient Attenuation of Stochasticity in Gene Expression Through Post-transcriptional Control. *Journal of molecular biology*. **344**, 965–976 (2004), doi:10.1016/j.jmb.2004.09.073.
37. C. Berens, W. Hillen, Gene regulation by tetracyclines. Constraints of resistance regulation in bacteria shape TetR for application in eukaryotes. *Eur J Biochem*. **270**, 3109–3121 (2003), doi:10.1046/j.1432-1033.2003.03694.x.
38. M. Gossen, A. L. Bonin, S. Freundlieb, H. Bujard, Inducible gene expression systems for higher eukaryotic cells. *Current opinion in biotechnology*. **5**, 516–520 (1994).
39. U. Baron, M. Gossen, H. Bujard, Tetracycline-controlled transcription in eukaryotes: novel transactivators with graded transactivation potential. *Nucleic Acids Research*. **25**, 2723–2729 (1997).
40. S. Urlinger *et al.*, Exploring the sequence space for tetracycline-dependent transcriptional activators: Novel mutations yield expanded range and sensitivity. *PNAS*. **97**, 7963–7968 (2000).
41. G. Belli, E. Gari, L. Piedrafita, M. Aldea, E. Herrero, An activator/repressor dual system allows tight tetracycline-regulated gene expression in budding yeast. *Nucleic Acids Research*. **26**, 942–947 (1998).
42. M. A. Miller, W. M. Olivas, Roles of Puf proteins in mRNA degradation and translation. *Wiley interdisciplinary reviews. RNA*. **2**, 471–492 (2011), doi:10.1002/wrna.69.

## References

43. R. Shalgi, M. Lapidot, R. Shamir, Y. Pilpel, A catalog of stability-associated sequence elements in 3' UTRs of yeast mRNAs. *Genome Biology*. **6**, R86.1 - R86.15 (2005).
44. M. Yamanishi, S. Katahira, T. Matsuyama, TPS1 Terminator Increases mRNA and Protein Yield in a *Saccharomyces cerevisiae* Expression System. *Biosci. Biotechnol. Biochem.* **75**, 2234–2236 (2011).
45. M. Yamanishi *et al.*, A Genome-Wide Activity Assessment of Terminator Regions in *Saccharomyces cerevisiae* Provides a "Terminatome" Toolbox. *ACS Synth. Biol.* **2**, 337–347 (2013), doi:10.1021/sb300116y.
46. K. A. Curran, A. S. Karim, A. Gupta, H. S. Alper, Use of expression-enhancing terminators in *Saccharomyces cerevisiae* to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metabolic Engineering*. **19**, 88–97 (2013), doi:10.1016/j.ymben.2013.07.001.
47. K. A. Curran *et al.*, Short Synthetic Terminators for Improved Heterologous Gene Expression in Yeast. *ACS synthetic biology*. **4**, 824–832 (2015), doi:10.1021/sb5003357.
48. A. Wittmann, B. Suess, Engineered riboswitches: Expanding researchers' toolbox with synthetic RNA regulators. *FEBS Letters*. **586**, 2076–2083 (2012), doi:10.1016/j.febslet.2012.02.038.
49. J. C. Liang, A. L. Chang, A. B. Kennedy, C. D. Smolke, A high-throughput, quantitative cell-based screen for efficient tailoring of RNA device activity. *Nucleic Acids Research*. **40**, e154 (2012), doi:10.1093/nar/gks636.
50. C. Kiel, E. Yus, L. Serrano, Engineering Signal Transduction Pathways. *Cell*. **140**, 33–47 (2010), doi:10.1016/j.cell.2009.12.028.
51. H. M. Sauro, B. N. Kholodenko, Quantitative analysis of signaling networks. *Progress in Biophysics and Molecular Biology*. **86**, 5–43 (2004), doi:10.1016/j.pbiomolbio.2004.03.002.
52. S. Erdmann, M. Snyder, A Filamentous Growth Response Mediated by the Yeast Mating Pathway. *Genetics*. **159**, 919–928 (2001).
53. J. E. Segall, Polarization of yeast cells in spatial gradients of alpha mating factor. *Proceedings of the National Academy of Sciences*. **90**, 8332–8336 (1993), doi:10.1073/pnas.90.18.8332.
54. A. Banderas, M. Koltai, A. Anders, V. Sourjik, Sensory input attenuation allows predictive sexual response in yeast. *Nat Comms*. **7**, 12590 (2016), doi:10.1038/ncomms12590.
55. G. I. Lang, A. W. Murray, D. Botstein, The cost of gene expression underlies a fitness trade-off in yeast. *Proceedings of the National Academy of Sciences*. **106**, 5755–5760 (2009).
56. T. I. Moore *et al.*, Robust Spatial Sensing of Mating Pheromone Gradients by Yeast Cells. *PLoS ONE*. **3**, e3865 (2008), doi:10.1371/journal.pone.0003865.
57. M. Good, G. Tang, J. Singleton, A. Reményi, W. A. Lim, The Ste5 Scaffold Directs Mating Signaling by Catalytically Unlocking the Fus3 MAP Kinase for Activation. *Cell*. **136**, 1085–1097 (2009), doi:10.1016/j.cell.2009.01.049.
58. R. E. Chen, J. Thorner, Function and regulation in MAPK signaling pathways. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*. **1773**, 1311–1340 (2007), doi:10.1016/j.bbamcr.2007.05.003.
59. O. Atay, J. M. Skotheim, Spatial and temporal signal processing and decision making by MAPK pathways. *J. Cell Biol.* **216**, 317–330 (2017), doi:10.1083/jcb.201609124.
60. N.-N. Zhang, Multiple Signaling Pathways Regulate Yeast Cell Death during the Response to Mating Pheromones. *Molecular Biology of the Cell*. **17**, 3409–3422 (2006), doi:10.1091/mbc.E06-03-0177.
61. K. Doi *et al.*, MSG5, a novel protein phosphatase promotes adaptation to pheromone response in *S. cerevisiae*. *EMBO*. **13**, 61–70 (1994).
62. H. G. Dohlman, J. Song, D. Ma, W. E. Courchesne, J. Thorner, Sst2, a negative regulator of pheromone signaling in the yeast *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*. **16**, 5194–5209 (1996).
63. D. Nevozhay, R. M. Adams, K. F. Murphy, K. Josic, G. Balazsi, Negative autoregulation linearizes the dose–response and suppresses the heterogeneity of gene expression. *PNAS*. **106**, 5123–5128 (2009).

## References

64. G. Dixit, J. B. Kelley, J. R. Houser, T. C. Elston, H. G. Dohlman, Cellular noise suppression by the regulator of G protein signaling Sst2. *Molecular cell*. **55**, 85–96 (2014), doi:10.1016/j.molcel.2014.05.019.
65. A. Colman-Lerner *et al.*, Regulated cell-to-cell variation in a cell-fate decision system. *Nature*. **437**, 699–706 (2005), doi:10.1038/nature03998.
66. J. Stewart-Ornstein, J. S. Weissman, H. El-Samad, Cellular Noise Regulons Underlie Fluctuations in *Saccharomyces cerevisiae*. *Molecular cell*. **45**, 483–493 (2012), doi:10.1016/j.molcel.2011.11.035.
67. S. G. Peisajovich, J. E. Garbarino, P. Wei, W. A. Lim, Rapid Diversification of Cell Signaling Phenotypes by Modular Domain Recombination. *Science*. **328**, 368–372 (2010), doi:10.1126/science.1182376.
68. C. J. Bashor, N. C. Helman, S. Yan, W. A. Lim, Using Engineered Scaffold Interactions to Reshape MAP Kinase Pathway Signaling Dynamics. *Science*. **319**, 1539–1543 (2008), doi:10.1126/science.1151153.
69. K. E. Galloway, E. Franco, C. D. Smolke, Dynamically Reshaping Signaling Networks to Program Cell Fate via Genetic Controllers. *Science*. **341**, 12350051–123500511 (2013), doi:10.1126/science.1235005.
70. S.-H. Park, A. Zarrinpar, W. A. Lim, Control of Regulatory T Cell Development by the Transcription Factor Foxp3. *Science*. **299**, 1057–1061 (2003), doi:10.1126/science.1079490.
71. T. C. Williams, L. K. Nielsen, C. E. Vickers, Engineered Quorum Sensing Using Pheromone-Mediated Cell-to-Cell Communication in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **2**, 136–149 (2012), doi:10.1021/sb300110b.
72. H. Youk, W. A. Lim, Secreting and Sensing the Same Molecule Allows Cells to Achieve Versatile Social Behaviors. *Science*. **343**, 1242782 (2014), doi:10.1126/science.1242782.
73. N. T. Ingolia, S. Ghaemmaghami, Newman, John R. S., J. S. Weissman, Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science*. **324**, 218–223 (2009).
74. A. Urrios *et al.*, A Synthetic Multicellular Memory Device. *ACS Synth. Biol.* **5**, 862–873 (2016), doi:10.1021/acssynbio.5b00252.
75. S. Regot *et al.*, Distributed biological computation with multicellular engineered networks. *Nature*. **469**, 207–211 (2010), doi:10.1038/nature09679.
76. N. Fukuda, J. Ishii, M. Kaishima, A. Kondo, Amplification of agonist stimulation of human G-protein-coupled receptor signaling in yeast. *Analytical Biochemistry*. **417**, 182–187 (2011), doi:10.1016/j.ab.2011.06.006.
77. J. Ishii *et al.*, Improved identification of agonist-mediated G $\alpha$ i-specific human G-protein-coupled receptor signaling in yeast cells by flow cytometry. *Analytical Biochemistry*. **426**, 129–133 (2012), doi:10.1016/j.ab.2012.04.012.
78. C. Springer, O. Valerius, A. Strittmatter, G. H. Braus, The Adjacent Yeast Genes ARO4 and HIS7 Carry No Intergenic Region. *J. Biol. Chem.* **272**, 26318–26324 (1997), doi:10.1074/jbc.272.42.26318.
79. T. Decoene, G. Peters, S. L. de Maeseneire, M. de Mey, Toward Predictable 5'UTRs in *Saccharomyces cerevisiae*. *ACS synthetic biology* (2018), doi:10.1021/acssynbio.7b00366.
80. V. Shahrezaei, P. S. Swain, Analytical distributions for stochastic gene expression. *PNAS*. **105**, 17256–17261 (2008).
81. A. Raj, A. van Oudenaarden, Single-molecule approaches to stochastic gene expression. *Annual review of biophysics*. **38**, 255–270 (2009), doi:10.1146/annurev.biophys.37.032807.125928.
82. D. R. Larson, R. H. Singer, D. Zenklusen, A single molecule view of gene expression. *Trends in Cell Biology*. **19**, 630–637 (2009), doi:10.1016/j.tcb.2009.08.008.
83. C. J. Roberts *et al.*, Signaling and Circuitry of Multiple MAPK Pathways Revealed by a Matrix of Global Gene Expression Profiles. *Science*. **287**, 873–880 (2000), doi:10.1126/science.287.5454.873.
84. K. A. Schandel, D. D. Jenness, Direct Evidence for Ligand-Induced Internalization of the Yeast  $\alpha$ -Factor Pheromone Receptor. *Molecular and Cellular Biology*. **14**, 7245–7255 (1994).

## References

85. C. Miller *et al.*, Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. *Mol Syst Biol.* **7**, 1–13 (2011), doi:10.1038/msb.2010.112.
86. A. Goppelt, M. Meisterernst, Characterization of the Basal Inhibitor of Class II Transcription NC2 from *Saccharomyces cerevisiae*. *Nucleic Acids Research.* **24**, 4450–4455 (1996), doi:10.1093/nar/24.22.4450.
87. L. Petko, S. Lindquist, Hsp26 is not required for growth at high temperatures, nor for thermotolerance, spore development, or germination. *Cell.* **45**, 885–894 (1986), doi:10.1016/0092-8674(86)90563-5.
88. M. Amoros, F. Estruch, Hsf1p and Msn2/4p cooperate in the expression of *Saccharomyces cerevisiae* genes HSP26 and HSP104 in a gene- and stress type-dependent manner. *Mol Microbiol.* **39**, 1523–1532 (2001), doi:10.1046/j.1365-2958.2001.02339.x.
89. R. M. C. Portela *et al.*, Synthetic Core Promoters as Universal Parts for Fine-Tuning Expression in Different Yeast Species. *ACS synthetic biology.* **6**, 471–484 (2017), doi:10.1021/acssynbio.6b00178.
90. C. Batenchuk *et al.*, Chromosomal position effects are linked to sir2-mediated variation in transcriptional burst size. *Biophysical Journal.* **100**, L56–8 (2011), doi:10.1016/j.bpj.2011.04.021.
91. A. Becskei, B. B. Kaufmann, A. van Oudenaarden, Contributions of low molecule number and chromosomal positioning to stochastic gene expression. *Nature Genetics.* **37**, 937–944 (2005), doi:10.1038/ng1616.
92. P. S. Swain, M. B. Elowitz, E. D. Siggia, Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences of the United States of America.* **99**, 12795–12800 (2002), doi:10.1073/pnas.162041399.
93. A. Becskei, B. Seraphin, L. Serrano, Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO.* **20**, 2528–2535 (2001).
94. V. Pelechano, J. García-Martínez, J. E. Pérez-Ortín, A genomic study of the inter-ORF distances in *Saccharomyces cerevisiae*. *Yeast (Chichester, England).* **23**, 689–699 (2006), doi:10.1002/yea.1390.
95. N. Bobola, R.-P. Jansen, T. H. Shin, K. Nasmyth, Asymmetric accumulation of Ash1p in postanaphase nuclei depends on a myosin and restricts yeast mating-type switching to mother cells. *Trends in Genetics.* **12**, 252 (1996), doi:10.1016/0168-9525(96)90034-7.
96. J. S. Anderson, R. P. Parker, The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *The EMBO journal.* **17**, 1497–1506 (1998), doi:10.1093/emboj/17.5.1497.
97. D. Volfson *et al.*, Origins of extrinsic variability in eukaryotic gene expression. *Nature.* **439**, 861–864 (2006), doi:10.1038/nature04281.
98. S. Ghaemmaghami *et al.*, Global analysis of protein expression in yeast. *Nature.* **425**, 737–741 (2003).
99. L. Cai, N. Friedman, X. S. Xie, Stochastic protein expression in individual cells at the single molecule level. *Nature.* **440**, 358–362 (2006), doi:10.1038/nature04599.
100. L. B. Carey, D. van Dijk, P. M. A. Slood, J. A. Kaandorp, E. Segal, Promoter sequence determines the relationship between expression level and noise. *PLoS biology.* **11**, e1001528 (2013), doi:10.1371/journal.pbio.1001528.
101. M. J. Holland, Transcript abundance in yeast varies over six orders of magnitude. *The Journal of biological chemistry.* **277**, 14363–14366 (2002), doi:10.1074/jbc.C200101200.
102. V. Takhaveev, M. Heinemann, Metabolic heterogeneity in clonal microbial populations. *Current opinion in microbiology.* **45**, 30–38 (2018), doi:10.1016/j.mib.2018.02.004.
103. J. Masel, M. L. Siegal, Robustness. *Trends in genetics : TIG.* **25**, 395–403 (2009), doi:10.1016/j.tig.2009.07.005.
104. L. S. Garrenton *et al.*, Nucleus-specific and cell cycle-regulated degradation of mitogen-activated protein kinase scaffold protein Ste5 contributes to the control of signaling competence. *Molecular and Cellular Biology.* **29**, 582–601 (2009), doi:10.1128/MCB.01019-08.
105. U. Alon, Network motifs. *Nature reviews. Genetics.* **8**, 450–461 (2007), doi:10.1038/nrg2102.

## References

106. S. Hooshangi, S. Thiberge, R. Weiss, Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proceedings of the National Academy of Sciences of the United States of America*. **102**, 3581–3586 (2005), doi:10.1073/pnas.0408507102.
107. R. Cheong, A. Rhee, J. Wang, I. Nemenman, A. Levchenko, Information Transduction Capacity of Noisy Biochemical Signaling Networks. *Science*. **334**, 354–358 (2011).
108. Z. Zhang, W. Qian, J. Zhang, Positive selection for elevated gene expression noise in yeast. *Molecular systems biology*. **5**, 299 (2009), doi:10.1038/msb.2009.58.
109. M. S. Siddiqui, K. Thodey, I. Trenchard, C. D. Smolke, Advancing secondary metabolite biosynthesis in yeast with synthetic biology tools. *FEMS yeast research*. **12**, 144–170 (2012), doi:10.1111/j.1567-1364.2011.00774.x.
110. C. Janke *et al.*, A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast*. **21**, 947–962 (2004), doi:10.1002/yea.1142.
111. D. J. Warren, Preparation of highly efficient electrocompetent *Escherichia coli* using glycerol/mannitol density step centrifugation. *Analytical Biochemistry*. **413**, 206–207 (2011), doi:10.1016/j.ab.2011.02.036.
112. J. Goedhart *et al.*, Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. *Nature communications*. **3**, 751 (2012), doi:10.1038/ncomms1738.
113. N. C. Shaner *et al.*, A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum*. *Nature methods*. **10**, 407–409 (2013), doi:10.1038/nmeth.2413.
114. N. C. Shaner *et al.*, Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nature Biotechnology*. **22**, 1567–1572 (2004), doi:10.1038/nbt1037.
115. D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361 (1977), doi:10.1021/j100540a008.
116. A. Natarajan, S. Subramanian, F. Srienc, Comparison of mutant forms of the green fluorescent protein as expression markers in Chinese hamster ovary (CHO) and *Saccharomyces cerevisiae* cells. *Journal of Biotechnology*. **62**, 29–45 (1998), doi:10.1016/S0168-1656(98)00040-6.
117. B. C. Ross, Mutual information between discrete and continuous data sets. *PLoS ONE*. **9**, e87357 (2014), doi:10.1371/journal.pone.0087357.
118. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7. *Molecular biology and evolution*. **30**, 772–780 (2013), doi:10.1093/molbev/mst010.

## 6 Supplementary information

### 6.1 Supplementary table

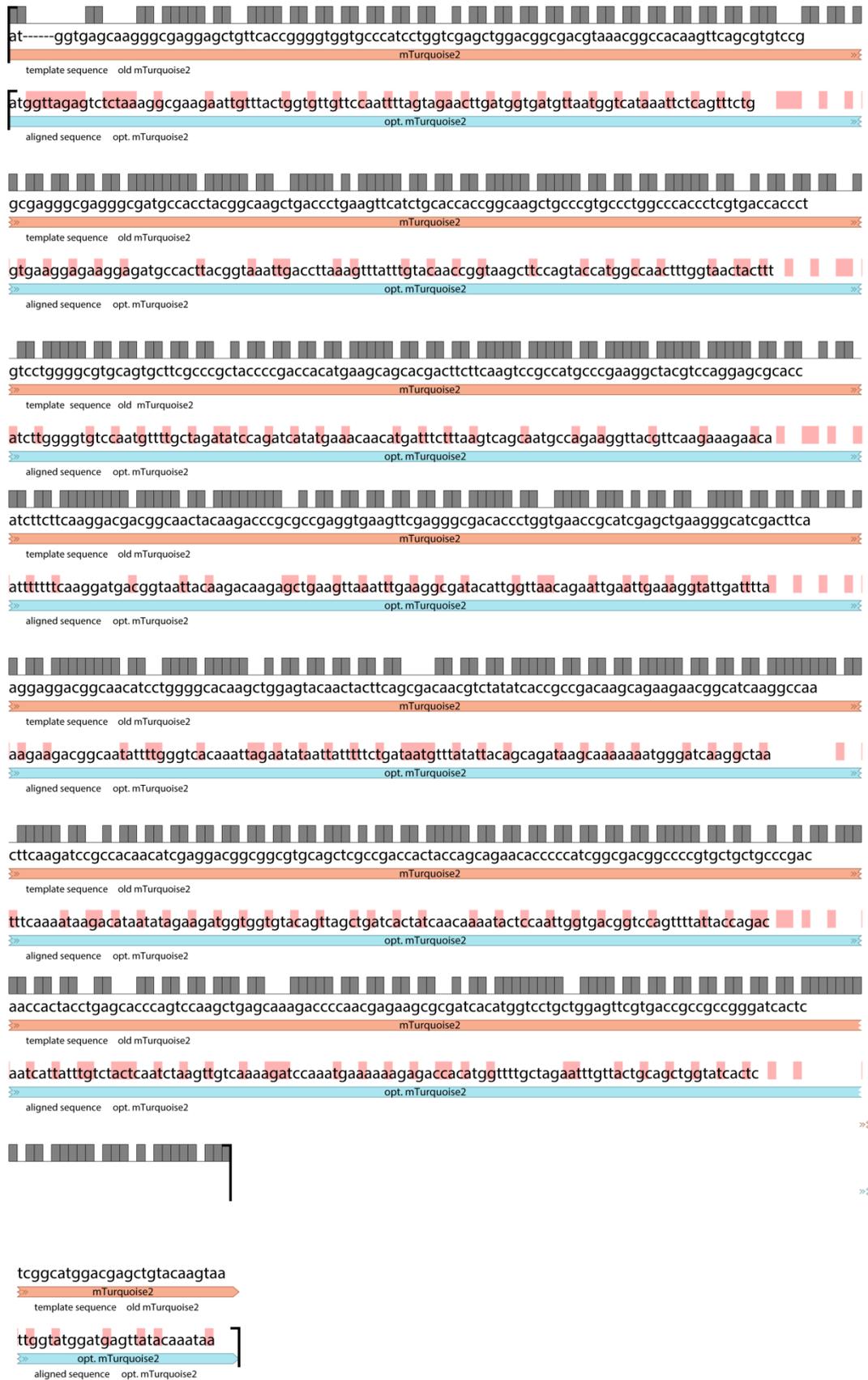
Table 8: Voltages used to test PMT noise as described in section 2.1.2. 488 nm laser wavelength and the 530/30 filter correspond to mNeogreen settings. 561 nm laser with 610/20 filter correspond to mCherry settings.

	488 nm laser with 530/30 filter		561 nm laser with 610/20 filter	
	Voltage	Median FI	Voltage	Median FI
“low”	215	3	365	3
“medium”	245	10	436	10
“high”	340	100	590	102

### 6.2 Sequence alignments

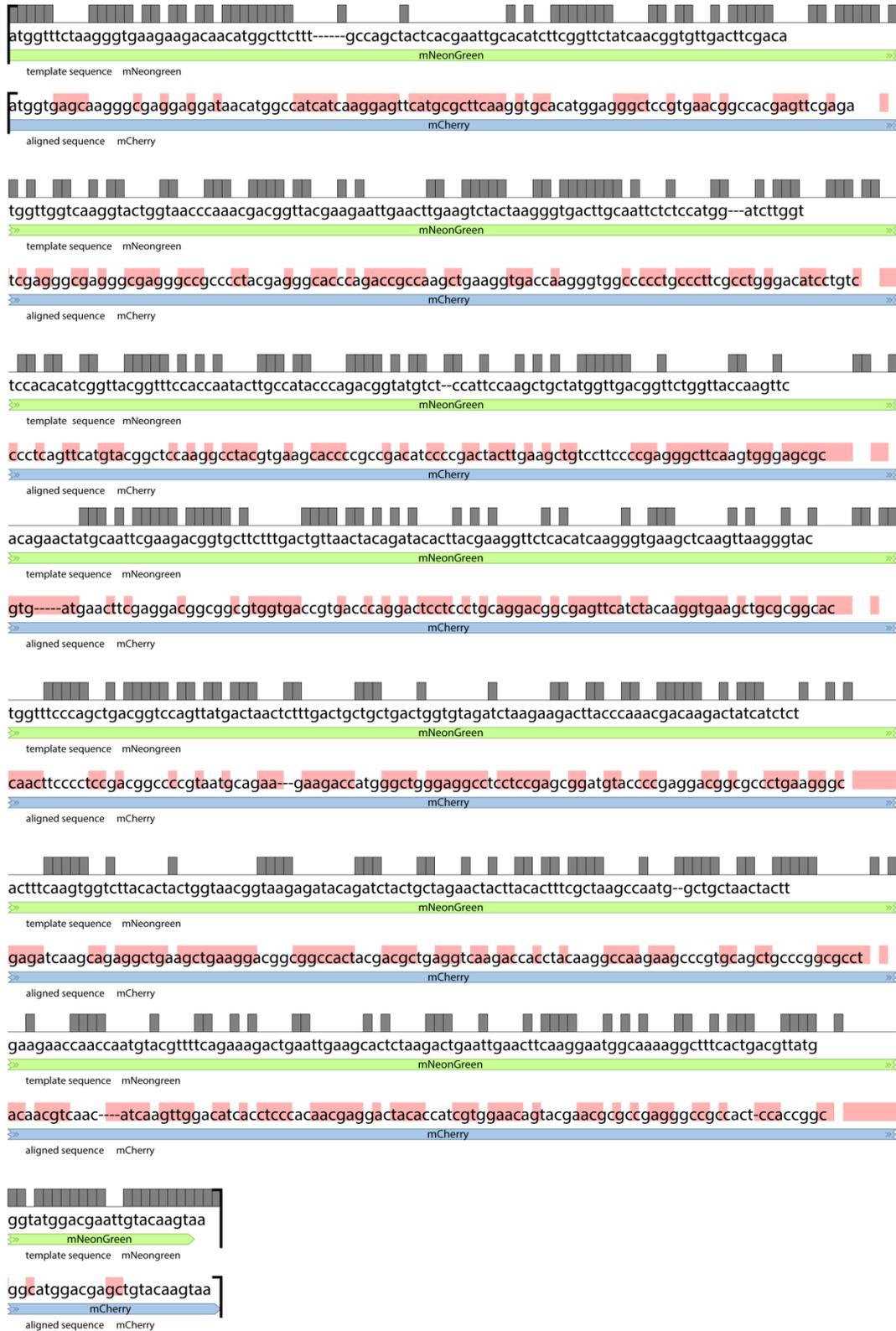
Due to *S. cerevisiae*'s capability to do efficient homologous recombination, genomic integrity can be reduced if genes with similar sequences are genomically integrated. This could be the case e.g. when fluorescence reporters are used that all derive from GFP and differ in only few base pairs to change the excitation and emission spectra of the protein. In this study, we used fluorescent reporter genes from three different original hosts. The following sequence alignments illustrate the low sequence identity between those genes. Supplementary Figure 1 shows the sequences of the original mTurquoise2 gene aligned to our codon-optimized version.

## Supplementary information



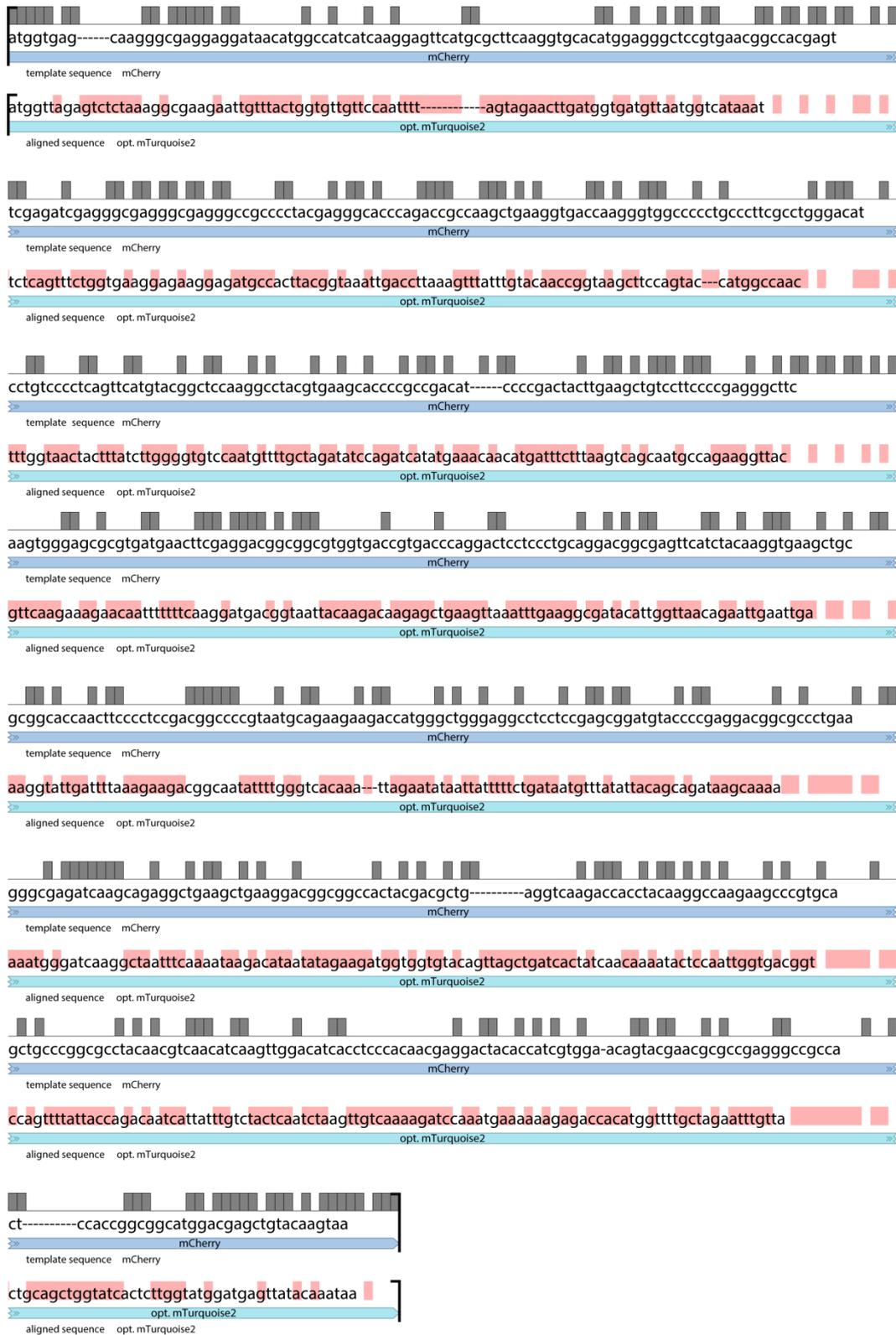
Supplementary Figure 1: Sequence alignment of codon-optimized mTurquoise2 open reading frame (blue) and original mTurquoise2 sequence (red). Red background behind bases indicates mismatch. Sequence alignment was done using MAFFT multiple sequence alignment software (118).

## Supplementary information



Supplementary Figure 2: Sequence alignment of mNeonGreen (green) and mCherry (blue) open reading frames. Red background behind bases indicates mismatch. Sequence alignment was done using MAFFT multiple sequence alignment software (118).

## Supplementary information



Supplementary Figure 3: Sequence alignment of mCherry (blue) and the codon-optimized mTurquoise2 version (turquoise). Red background behind bases indicates mismatch. Sequence alignment was done using MAFFT multiple sequence alignment software (118).

## Supplementary information

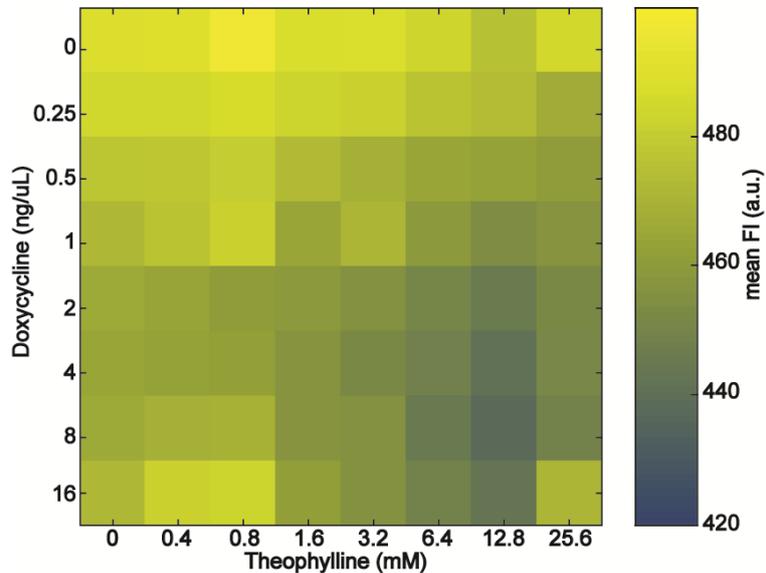


Supplementary Figure 4: Sequence alignment of codon-optimized mTurquoise2 (turquoise) and mNeonGreen (green) open reading frames. Red background behind bases indicates mismatch. Sequence alignment was done using MAFFT multiple sequence alignment software (118).

## Supplementary information

### 6.3 Supplementary figures

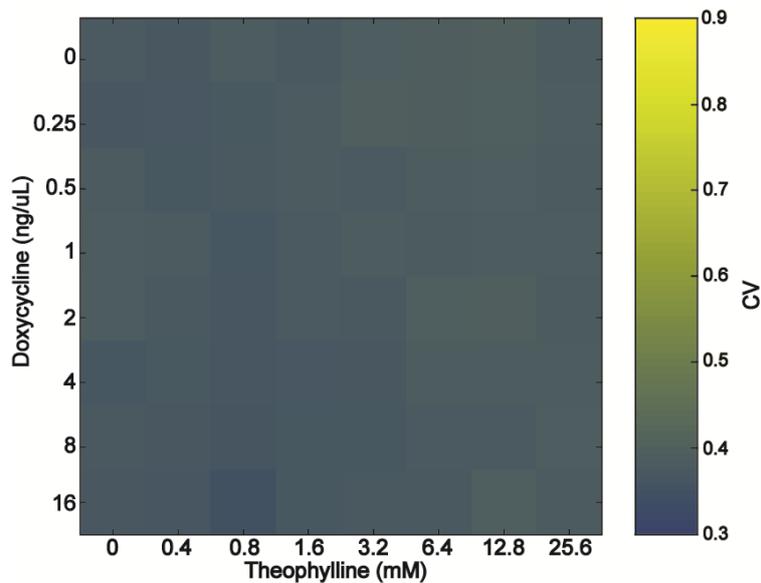
The fluorescence and noise of the constitutive mTurquoise2 reporter was analyzed to assure that the expression capacity control was not affected by doxycycline and theophylline.



Supplementary Figure 5: Mean fluorescence intensities of mTurquoise2 control module for 8-by-8 matrix of doxycycline and theophylline concentrations. Mean fluorescence intensities in arbitrary units are given on the color scale on the right.

We found that even at high theophylline levels, mTurquoise2 expression showed no change. The observed growth defect that resulted in reduced signal from the blue laser (mNeongreen excitation) showed no effect on the signal from the blue-violet laser (mTurquoise2 excitation). Generally, the expression is in a very narrow range.

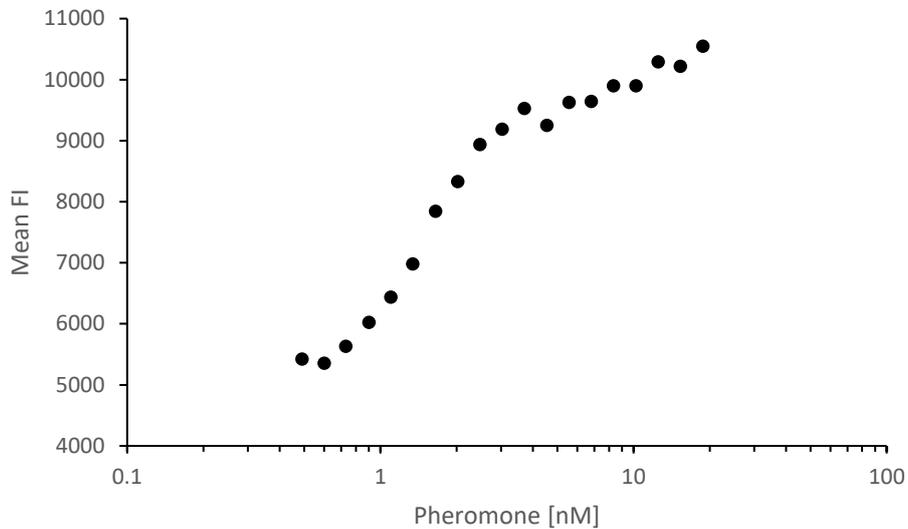
The corresponding expression noise of mTurquoise2 showed constant low noise (CV of ca. 0.3) with no changes when doxycycline or theophylline levels were altered.



Supplementary Figure 6: Noise, given as CV of mTurquoise2 control in STRSV strain. Color scale corresponds to the typical range of CVs measured in this study

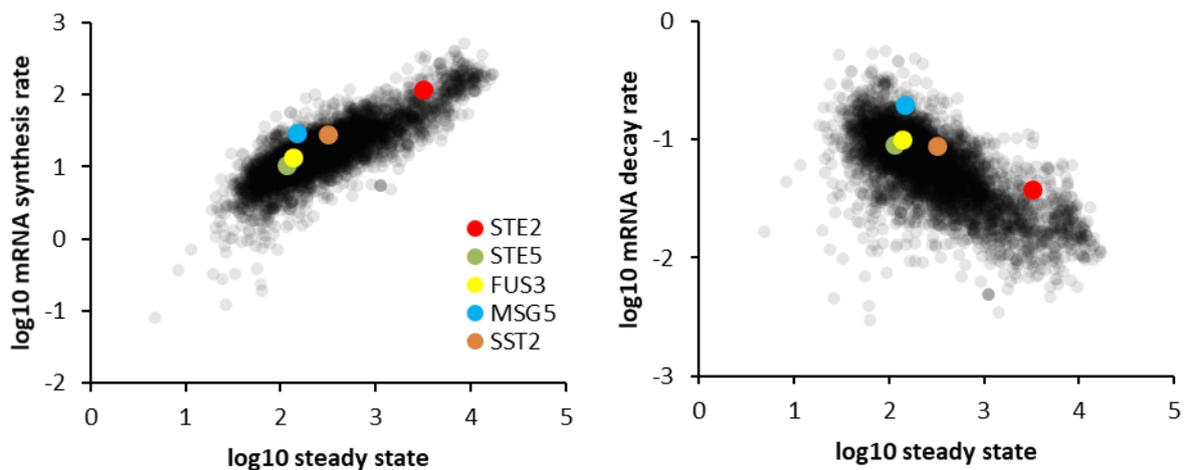
## Supplementary information

Expression of the mTurquoise2 control was also analyzed for mating pathway noise experiments. We found that expression increased by 2-fold with increasing pheromone concentration, indicating that the GPD promoter and/ or terminator is slightly responsive to pheromone



Supplementary Figure 7: Expression of constitutive control mTurquoise2 increases over a 2-fold range with increasing pheromone concentration.

The mRNA synthesis and degradation rates as a function of the steady state mRNA level were calculated from data acquired by Miller and colleagues. The rates follow the expected correlation for higher steady state levels when the transcription rate is high and/ or the degradation rate low. The data was used to analyze how the rates of pathway genes at a given steady state compare to the average at that steady state.

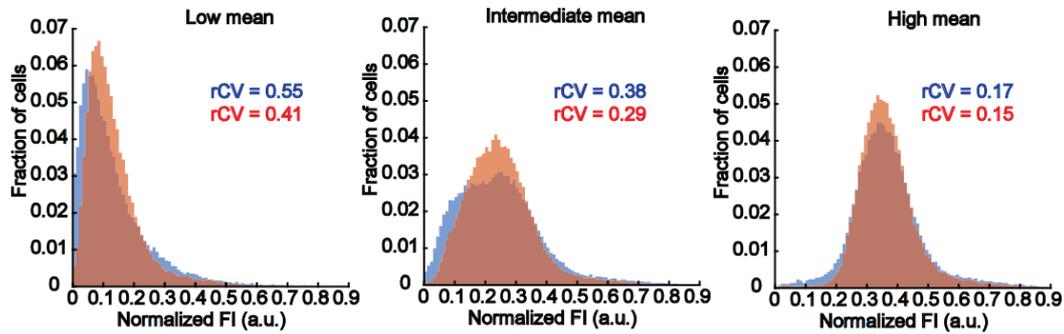


Supplementary Figure 8: Synthesis and decay rates of yeast transcripts. Mating pathway genes discussed in this thesis are highlighted. The pathway receptor STE2 exhibits high mRNA steady-state levels mediated by high transcription rates. The negative feedback regulator Msg5 shows mRNA synthesis and decay rates at the upper limits for the given steady-state mRNA levels. Values were calculated from data obtained in (85).

We compared histograms of normalized pathway output for the SST2 noise tuning strain, either set to high or low SST2 expression noise. The low-noise distribution exhibits less spread and a higher peak

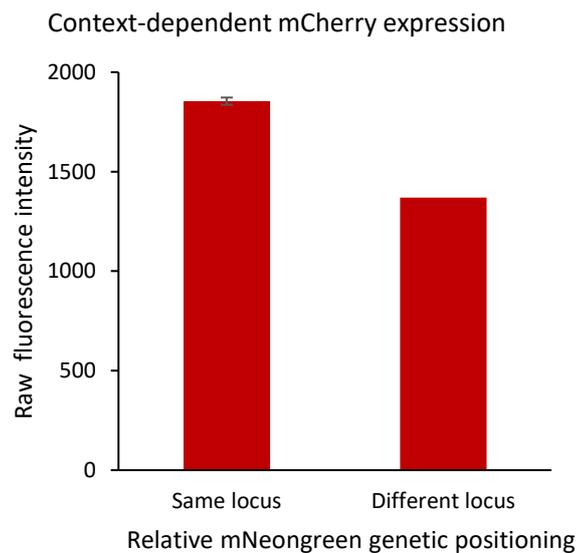
## Supplementary information

over the entire mean expression range, but the differences to the high-noise distribution are smaller as compared to the proof-of-concept strain.



Supplementary Figure 9: Pathway reporter histograms at different pheromone stimulation levels. Red and blue distributions correspond to low- and high-noise setting, respectively. Robust CVs of the distributions are given in the respective colors.

Actively transcribed genes in close physical proximity seem to increase expression of a target gene. Additionally to exhibiting 4-fold lower noise when expressed in the same locus, we also found that mCherry was expressed 37 % stronger when mNeogreen was expressed from the same locus, compared to being expressed from a different locus (Supplementary Figure 10).



Supplementary Figure 10: mCherry expression governed by a TetO7 promoter increases when mNeogreen driven by the same TetO7 promoter is expressed from the same locus (*his3*) as compared to mNeogreen expression from a different locus (*ura3*). Error bar indicates standard deviation of four biological replicates. "Different locus" measurement was measured with only one biological replicate.

## 6.4 Chemicals and consumables

<b>Chemical</b>	<b>Supplier</b>
Agarose	Roth
Alpha pheromone	Sigma-Aldrich
Amino acid mix	Formedium
Ampicillin sodium salt	Roth
Bacto agar	BD Biosciences
Bacto tryptone	BD Biosciences
Bacto yeast extract	BD Biosciences
Casein sodium salt	Sigma-Aldrich
ClonNAT	Sigma-Aldrich
Concanavalin A	Sigma-Aldrich
Doxycycline	Sigma-Aldrich
EDTA	Roth
Ethanol	Roth
G418	Sigma-Aldrich
Glacial acetic acid	Roth
Glucose	Sigma-Aldrich
Glycerol	Roth
HCl	Roth
Hygromycin B	Formedium
KCl	Roth
Lithium acetate dihydrate	Sigma-Aldrich
Magnesium sulfate	Sigma-Aldrich
Mannitol	Sigma-Aldrich
NaCl	Roth
NaOH	Roth
PEG4000	Sigma-Aldrich
Salmon sperm DNA	ThermoFisher
Theophylline	Sigma-Aldrich
Tris base	Roth
Tris HCl	Roth
YNB (yeast nitrogen base)	Formedium
YNB LoFlo	Formedium
YPD powder	Roth

## Acknowledgements

### 7 Acknowledgements

I would like to express my sincere gratitude to Prof. Dr. Victor Sourjik for providing a great working environment and for giving me the trust and liberty to explore different paths to develop as a scientist, inside and outside the lab.

I also would like to thank the other members of my Thesis Advisory Committee, Prof. Dr. Hans-Ulrich Mösch and Dr. Tobias Erb for valuable advice to advance the projects described in this thesis.

I am especially grateful to Dr. Alexander Anders for sharing his knowledge and resources with me and for his constant guidance over the past four years. Many thanks also to Dr. Seán Murray for his help with data analysis and many fruitful discussions. I am especially indebted to these two colleagues for taking over the majority of the reviewing process of my manuscript, making this a much better thesis than it would have been otherwise.

Furthermore, I would like to thank Dr. Gabriele Malengo and Silvia González Sierra who taught me to operate a flow cytometer and for their advice on the topic.

I wish to thank all current and former colleagues in the Sourjik lab, especially Anne Löchner, Nico Krink, Daniel Hürtgen and Joana Lopes for the great experiences working in a team together and organizing so many different scientific events in the past years.

Schließlich danke ich meiner Mutter für jahrzehntelange bedingungslose Unterstützung, sowie meiner ganzen Familie, allen voran meinen Großeltern und meinem Bruder für den Zuspruch und Rückhalt.

## 8 Curriculum vitae