Universität Bremen

Development of an online reporter system to follow plasmid stability in *Saccharomyces cerevisiae*

Johannes Öhrlein

Dissertation submitted in partial fulfilment
of the requirements for the degree

Dr. rer. nat.

Bremen, Feburary 2014

The experiments of the presented study have been carried out from August 2009 until July 2013 under the guidance of Prof. Dr Tilman Achstetter in the Laboratory of Industrial Microbiology at the University of Applied Sciences Bremen, Germany in cooperation with Prof. Dr. Barbara Reinhold-Hurek from the Department 2 Biology/Chemistry of Bremen University, Germany.

First reviewer: Prof. Dr. Barbara Reinhold-Hurek

Second reviewer: Prof. Dr. Tilman Achstetter

Defense of the doctor's thesis was at 4th of April 2014

TABLE OF CONTENT

ABBREVIATIONS

APS	ammonium persulfate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
ER	endoplasmatic reticulum
FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide
GPI	glycosylphosphatidylinositol
HTH	helix turn helix
LB	Luria-Bertani
LTR	long terminal repeat
NAD	nicotinamide adenine dinucleotide
OD	optical density
PCR	polymerase chain reaction
RNase	ribonuclease
RT	room temperature
S.c.	*Saccharomyces cerevisiae*
SD	synthetic defined
SDS	sodium dodecyl sulfate
TEMED	tetramethylethylenediamine
TFIID	transcription factor II D
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet
YPD(AU)	yeast extract, peptone, dextrose, (adenine sulphate, uracil)

SUMMARY

Control over a process is important to ensure the correct outcome. However, this control is often missing in certain situations such as the monitoring of the stability of process relevant genetic information in producing organisms, especially if this information is passed over in form of plasmid DNA.

Therefore, this work aims to develop a reporter system that makes it possible to follow plasmid stability online in *Saccharomyces cerevisiae,* which is widely used in industrial settings. The idea behind this is that a chromosomally integrated signaling protein is controlled by a repressor encoded on the plasmid DNA. As soon as the plasmid is absent, the concentration of the repressor is reduced and the signaling protein can be expressed.

Several foundations are needed to develop such a system, like a stable expression of the signaling protein with an optimized location for signal strength. This was found out to be the fluorescent protein yEGFP3 under the control of the *PGK1* promoter and localized into the peroxisome.

Following these results, the *PGK1* promoter was improved by insertion of lambda phage based operator sites in specific locations in order to control the reporter expression. Here, it could be recognized that a singular insertion of up to 48 bp in most cases does not interfere with *PGK1* promoter activity to any significant extent.

Two lambda phage repressor variants were tested at different expression levels in combination with the genome integrated expression cassettes variants. In these experiments, the combination with a highly expressed wild type repressor in most promoter variants lead to a decreased expression of yEGFP3 of about 100%, whereas in case of the combination with repressors containing a nuclear localization sequence site, either no effect on the expression or even an increase of the expression has been observed.

The developed combinations with the highest repression of the reporter protein expression were able to detect a plasmid loss of 50% or 30% of the total cells, depending on the detection method.

SUMMARY GERMAN

Die Kontrolle über einen Prozess ist entscheidend um dessen Ergebnis zu gewährleisten. In vielen Situationen ist eine deratige Kontrolle aber nicht möglich, da ein effizienter Kontrollmechanismus fehlt. Ein Beispiel hierfür sind prozessrelevante genetische Informationen in Organismen, insbesondere dann, wenn es sich dabei um Plasmid basierte Informationen handelt.

Daher wird in dieser Arbeit ein Reportersystem entwickelt, welches online die Plasmidstabilität in dem oft industriell verwendeten Organismus *Saccharomyces cerevisiae* überprüfen kann. Dies basiert auf der Idee, dass ein ins Genom integriertes Reportergen durch einen Repressor kontrolliert wird, welcher auf dem Plasmid codiert ist. Sobald das Plasmid verloren geht, verringert sich die Konzentration des Repressor und ein Signal wird erzeugt.

Verschiedenste Grundlagen müssen zur Erstellung eines solchen Systems entwickelt werden, wie z.B. die konstante und unabhängige Produktion des Signalmoleküls. Dafür hat sich das Fluoreszenzprotein yEGFP3 unter Kontrolle des *PGK1* Promoters und Peroxisomenlokalisation als geeignet gezeigt.

Des Weiteren wurden im Promoter an verschiedenen Positionen Operatorsequenzen integriert, die auf dem Lambda Phagen basieren, um eine Kontrolle des Promoters zu ermöglichen. Dabei hinterließ die Integration von bis zu 48 bp in den meisten Fällen einen noch aktiven *PGK1* Promoter.

In Kombination mit diesen Promotervarianten wurden jeweils zwei Versionen des Lambda Phagen Repressors getestet In diesen Tests zeigte sich, dass der natürliche Repressor mit fast allen Promoter Variationen eine Reduktion der Expression von bis zu 100% bewirkt. Die Kombination mit einem Repressor, welcher ein Kernlokalisierungssignal trägt, führte hingegen entweder zur keiner Reaktion oder sogar zu einer Steigerung der Expression.

Die Kombinationen mit der höchsten Repression ermöglichten es, den Plasmidverlust in Zellen zu beobachten, sobald - abhängig von der Detektionsmethode - 50% oder 30% der Zellen das Plasmid verloren hatten.

1 INTRODUCTION

Control over a process is important to ensure the correct outcome. However, this control is often not available in certain situations such as the monitoring of the stability of process relevant genetic information in producing organisms, especially if this information is passed over faithfully in form of plasmid DNA. Therefore, this work aims to develop a reporter system that makes it possible to follow plasmid stability online in *Saccharomyces cerevisiae*. Contrary to the statement of Kozak (2007) that "a frequent problem is deciding the answer first and then constructing a test system to show what one wants to see", the answer or goal is predetermined and has to be established by different approaches; this will be explained in detail later on. Next, an explicit introduction about the background is following, starting with the description of the particularities of the organism used in this work.

1.1 SACCHAROMYCES CEREVISIAE

The unicellular organism *Saccharomyces cerevisiae* has already been accompanying humanity for a long time in different processes like brewing - as its name indicates - as well as baking and is equivalently called yeast, which is accurately covering several species. This long connection is also reflected in the scientific work, where it is used widely, for example is "enzyme" Greek for "in yeast".

Especially its numerous positive features make *S. cerevisiae* to one of the important eukaryotic model organisms. It is for example fast growing on different media, laboratory strains show a stable haploid and diploid phase and its genome was the first eukaryotic one to be sequenced. Furthermore, a wide range of genomic manipulations is established (Sherman 2002), many studies about protein interaction detectable in yeast are available (Goll and Uetz 2006), as well as knock out strains from 96% of the annotated open reading frames (Giaver *et al.* 2002). Even the subcellular localization of

known proteins is identified and thereby complementing the genome wide data - and as expected most of them are concentrated in the cytoplasm or nucleus, but still 44% of the total observed proteins are localized in other specific subcellular regions (Huh *et al.* 2003). However, it has to be kept in mind that there are no really wild type *Saccharomyces* strains commonly used in genetic studies and even the domesticated strains for brewing and baking are not genetically comparable with laboratory ones, which derived from few strains (Sherman 2002). This makes the comparison of different results sometimes difficult.

The reference strain for the genomic analysis presented in this study is *S. cerevisiae* S288C, which is haploid containing 16 chromosomes with a total of around 12,000 kpb chromosomal DNA and around 6100 open reading frames (Sherman 2002). Exceptional in comparison to other single cell eukaryotes is thereby the high genome density with 72% of the total sequence representing genes and that only 3.8% of them possess introns, despite the well-established splicing mechanism apparatus (Dujon 1996). One tentative explanation for this could be found in the multiplicity of intracellular RNA viruses (Wickner 1996), which are present in almost all *S. cerevisiae* strains and the homologous recombination ability (Kass and Jasin 2010). Nevertheless, genes for splicing as well as genes for transcription, ribosome biosynthesis, translation, cell wall and membrane biogenesis, DNA replication, nuclear transport, and basic cytoskeletal functions belong to the important 20% of yeast genes required for viability and are therefore not part of most knockout analyses (Mnaimneh *et al.* 2004). In the year 2007, more than 1000 genes of *Sac*charomyces were called uncharacterized and in the midst of 2013, the function of 759 still remain unclear. A reason for that could be their connection to the wild type lifestyle of yeast, where the cells are most of the time dormant and stick to insects like the fruit fly *Drosphila melanogaster* thereby being transported from fruit to fruit. The popular large-scale experiments by contrast use most of the time only artificial conditions, where an effect of these wild type lifestyle genes is rarely seen (Peňa-Castillo and Hughes 2007).

1.1.1 THE ORGANISM AND ITS INDUSTRIAL USE

Under laboratory conditions, prototrophic strains have a doubling time between 90 (rich medium) and 140 min (synthetic medium) during exponential phase (Sherman 2002). In any case yeast cells are aging over time and are able to produce around 30 daughter cells in average with a maximum of 40 (Falcon *et al.* 2005). The doubling time increases as a matter of course further with the quantity of mutations accumulated and introduced into the strains. Besides that, the growth behavior of auxotrophic strains is getting more and more attention again. This leads to a discussion about the optimal amount of complementing substrate in synthetic media and on the other hand, the observation that even in rich media like YPD supplements like adenine and uracil might be added in order to reach similar growth behavior as prototrophic strains under the same conditions (Corbacho *et al.* 2011).

Large-scale productions are especially influenced by the price, market size, purity and solubility of the respective product. In some circumstances, it is useful to apply defined media to save for example purification steps, but most of the time undefined complex media like molasses and a reduced growth rate is used. This leads to a constricted application of expression systems, because most of them react differently in terms of activity and strenght compared to the laboratory experiments. Furthermore, several other limitations and problems occur on a larger scale. An important limitation of growth rate is the exchange of heat and oxygen in a culture. Overall, local gradients in temperature, concentrations or pressure cannot be avoided. Another important fact is that around 35 to 37 generations are required for an efficient large-scale process and most of the product is created in the final phase. The consequence is a high impact on the productivity by the expression construct stability, which is leading for example to an 8.4% decrease of product at a 1% loss rate per generation. Besides that, any kind of maintenance energy affects the productivity and has to be kept in mind by the choice of the fermentation conditions and host organism (Hensing *et al.* 1995).

As soon as the production of heterologous proteins starts, various other stress reactions can be observed in cells. One of them is the unfolded protein response, which activates the endoplasmatic reticulum (ER) associated protein degradation among other processes. All of this leads to a decrease of protein secretion or even protein production.

At high cell densities, yeast cells can even affect each other through aromatic alcohols like phenylethanol and tryptophol or ammonia, which act autostimulatory in a quorum sensing fashion (Granek and Magwene 2010, Meunier and Choder 1999). Also on cell-to-cell contact a mutual influence is observed (Mattanovich *et al.* 2004), which becomes more obvious in yeast colonies. There for example a two phasic growth is observed, which is independent from the shift of fermentation to aerobic metabolism and also an active feeding mechanism between the cells cannot be ruled out (Meunier and Choder 1999).

The asymmetrical cell division of *S. cerevisiae* enforces the impression that yeast cultures are most of the time strongly heterogenic in cell appearance and biochemical activity. This heterogenic activity in a similar environment, whereby the normal slight genetic variability of a culture is already left out, is based on so-called “noise” in gene expression, which can be separated into intrinsic and extrinsic fluctuations (Longo and Hasty 2006). The intrinsic noise is based on the stochastic nature of biochemical reactions, which is influenced by the system size and thermal condition (Hasty 2000). In contrast to this is the extrinsic noise the global and local variation of components concentration like transcription factors or polymerases (Longo and Hasty 2006). Both effects are connected as described in the example of seemingly Ca^{2+} dependent stochastic bursts of transcription factors into the nucleus in yeast (Crabtree and Graef 2009). The majority of large-scale studies approves intrinsic noise as more influencing (Longo and Hasty 2006), whereby in eukaryotes the transcription and in prokaryotes the translation step generates the most noise (Blake *et al.* 2003).

1.1.2 PLASMIDS IN YEAST

The prior described heterogeneity is additionally reinforced by the so-called selfish elements observed in yeast. Next to the widespread introns, retrons, transposons, non-LTR and LTR retrotransposons (Velmurugan *et al.* 2003), and already mentioned RNA viruses most of the genera *Saccharomyces*, *Zygosaccharomyces*, *Kluyveromyces*, and *Torulaspora* possess an additional selfish DNA plasmid (cells with plasmid are called cir^+). Best analyzed from the 2µm family is the *Saccharomyces* plasmid with around 6.4 kpb (Frank and Wolfe 2009). It is a multi-copy plasmid but the exact number seems to vary in cells between 60-100 (Ahn *et al.* 1997), 40-60 (Ghosh *et al.* 2006), 14–34 (Karim *et al.* 2012) and 24-88 copies (Harrison *et al.* 2012). Assured is that with a loss rate of 10^{-5}–10^{-4} per generation it is nearly as stable as a chromosome (10^{-5}-10^{-6} per generation) (Jayaram *et al.* 2004). This is achieved by an active segregation and copy number control system through four plasmid-encoded proteins among them a DNA recombinase (Cui *et al.* 2009).

There seems to be a constant trade-off between the high plasmid number as well as stability and the fitness of the host. Not only is there the danger of mutation by the plasmid-encoded recombinase, but also every additional kpb of DNA was calculated to cost yeast around 0.025% of growth rate or fitness (Harrison *et al.* 2012), which is explaining why 2µm bearing cells have a 1.5% to 3% slightly longer generation time than plasmid free cells (Volkert *et al.* 1989). For the first time and in contradiction to the former results, it has recently been observed that in galactose rich media 2µm containing cells seem to have an around 12% higher fitness than plasmid free cells (Harrison *et al.* 2012). This indicates that the four encoded proteins on the 2µm plasmid could have more functions than only ensuring plasmid segregation and copy number.

The knowledge about the yeast molecular functions is used to establish different artificial forms of plasmids in yeast. Most of them are shuttle vectors bearing also DNA sequences for the transformation into *Escherichia*

coli and markers for both organisms. A simple form of a vector is an integrating plasmid, which has a selective marker and additionally a homologous DNA region to the chromosome or both in one (Gunge 1983). After transformation, such a vector integrates by homologous recombination into the chromosome, whereby a higher efficiency can be observed with previous linearized plasmids (Simon and Moore 1986). Depending on the homologous region, also one or multiple integrations can be achieved.

An autonomously replicating plasmid is constructed by the addition of autonomous replicating sequences (*ARS*), which are fairly different in sequence and can be found around every 32-40 kpb in the yeast genome (Gunge 1983). The copy number of these plasmids can be in quite high ranges of around 200–300 in several generation old cells (Falcon *et al.* 2005). However, their stability is low, because there is a barrier between the mother cell and the daughter cell by the bud respectively mother cell compartments as well as a too short interval for equal distribution (Liu *et al.* 2013). To overcome these stability problems a centromere region might be added giving rise to one of the more often used plasmid forms, which mimic a mini-chromosome. The so-called centromere plasmids (*CEN*/*ARS*) remain with a high stability over generations in low copy numbers of around 1-3 in the cell.

Another often-used plasmid is the artificial 2µm plasmid, which is carrying the replication origin and the *STB* (stability) locus of the wild type 2µm plasmid (Gunge 1983). These artificial 2µm are thereby dependent on the wild type 2µm plasmid-encoded proteins Rep 1 (replication) and Rep2 for stability and therefore only useable in strains bearing the wild type 2µm (Ahn *et al.* 1997). These artificial plasmids are not as stable as centromere plasmids, but can reach higher copy numbers. Thereby numerous aspects of copy number regulation and segregation of artificial 2µm are not known (Falcon *et al.* 2005) but it is assumed that they undergo homologous recombination with the resident 2µm. With long lasting high selection pressure the artificial 2µm plasmid can in certain circumstances outgrow the wild type 2µm plasmid by its replicative advantage, which is leading over

time to complete plasmid free cells (called cir^0), because of the dependency of the artificial 2µm plasmid on Rep1 and Rep2 (Gunge 1983).

In contrary, any kind of excessive high plasmid copy number leads to a yeast phenotype with nibbled colony morphology (Volkert *et al.* 1989) and faster aging (Falcon and Aris 2003). Interestingly the wild type 2µm has no such effect, but the artificial 2µm plasmids, which are also not as stable (Falcon *et al.* 2005). This is that way, because many other factors play a role in plasmid stability, which will be mentioned in the next paragraph.

1.2 PLASMID STABILITY

The term plasmid stability addresses events dealing with plasmid loss or its structural change, which is influenced by numerous factors, like the used organism, medium, culture conditions, plasmid size, selection marker, copy number, the produced heterologous protein, promoter of the heterologous expression block, the utilized replication origin, segregation origin and the DNA sequence itself.

All of these factors are connected to each other, whereby a strong effect is visible at the copy number of a plasmid. Normally it is assumed that a high copy number makes a plasmid more stable, because of the pure higher chance of being segregated during cell division if plasmids are segregated randomly between both cells. However, different observations are contrary to this (Zhang *et al.* 1996). Despite its high plasmid number and efficient replication, around 10-50% of *ARS* plasmids are lost per cell division (Scott-Drew and Murray 1998) through the earlier mentioned retention inside the mother cell. An active segregation seems to be important in yeast for plasmid stability making 2µm-based plasmids apparently the only useful multi-copy plasmid in yeast.

The copy number of the 2µm-based plasmid is highly influenced by the selection marker. Depending on the used addiction system, the copy number of 2µm-based plasmids is fluctuating around 14-34 copies per haploid cell in

the work of Karim *et al.* (2012). Remarkably, no such fluctuation was observed in the same study for the *CEN/ARS* plasmids (2-3 per cell) and by using antibiotic`s resistance as a marker both plasmid types show around 5-6 copies per cell (Karim *et al.* 2012). Besides that, it is possible to enhance the copy number of 2µm-based plasmids extremely by the usage of auxotrophic markers with truncated or even deleted promoters (Verma and Singh 2012). A truncated *LEU2* (YCL018W) marker promoter for example can rise the plasmid copy number up to 122-150 (Fujimura *et al.* 1996), which presents itself in a higher gene dosis after that and therefore may lead to a higher productivity (Verma and Singh 2012). Interestingly, it was observed that exactly this *LEU2* marker with its wild type promoter leads to a decreased fitness of cells and plasmid copy number in comparison to other markers (Karim *et al.* 2012). This indicates that there can be yeast cells created with an even higher plasmid copy number coming close to the amount of foreign DNA, which certainly cannot be maintained stable and reasonable in the cell because of the DNA fitness burden (Harrison *et al.* 2012) or sometimes toxic overproduction of cell foreign components (Zhang *et al.* 1996). Nevertheless, it has to be kept in mind that the increase in copy numbers is reached through a high selective pressure and it is not clear what other changes are needed in order to cope with or are accompanied by such a high plasmid number.

Also connected to the selection marker is a growth rate decline between 5 and 25%, especially in haploid cells, while in diploid cells the factor of DNA burden seems to be more relevant in that case (Karim *et al.* 2012). In this context, the influence of the medium is uncertain since the different mixtures of supplements are not comparable (Corbacho *et al.* 2011); yet other studies found negligible growth differences (Zhang *et al.* 1997, Kilonzo *et al.* 2009). Nonetheless, in both cell types, haploids and diploids, it is observed that the growth rate of cells with 2µm-based plasmids is lower than with *CEN/ARS* plasmids (Karim *et al.* 2012).

Additionally, the number of plasmid free cell seems to increase during the exponential growth phase but decreased during the stationary phase (Chen *et al.* 1997). Interestingly, in some cells with auxotrophic marker plasmids it

was observed that even the promoter strength of the gene of interest had a reducing effect on the plasmid copy number (Karim *et al.* 2012).

Another factor of plasmid stability is the DNA sequence itself. Especially if the sequence has several repeating motifs, structural instability can occur and the plasmid can form instable higher-order or non-Watson–Crick conformations, whereby the Z-form seems to be the most unstable one. Alternatively, the repeats are the basis for mutations and genetic rearrangements leading to a considerable percentage of structurally and sequentially changed plasmids in a population (Oliveira *et al.* 2009).

Particularly the chimeric part of a plasmid and its insert arrangement can bear for example cryptic promoters leading to instability (Romanos *et al.* 1992) in the same way as insufficient terminator sequences especially if their transcript is covering the *CEN* or *STB* site of the plasmid (Russo and Sherman 1989). The size and some kind of incompatibility between plasmids seem to have a similar effect on the stability (Futcher and Cox 1984). Moreover, a higher stability for *ARS* as well as for 2µm-based plasmids - although not that distinct - was observed in polyploid strains (Kilonzo *et al.* 2009). The influence of the host is clearly seen in a study of 266 gene knock-outs where already 12 showed higher plasmid instability (Hegemann *et al.* 1999). The pH (Kilonzo *et al.* 2009) and temperature finish the list of factors influencing plasmid stability and like most factors, the actual mechanism behind it is not clear, especially in the case of 2µm-based plasmids (Zhang *et al.* 1996).

Over time, plasmid free cells overgrow plasmid-bearing cells through a kinetic advantage (Lú-Chau *et al.* 2004) also called competitive instability (Kilonzo *et al.* 2009) and can therefore have enormous effects on the productivity. The quality as well as the quantity of the plasmid is important for a controlled and confirmed process, whereby non- or maybe abnormal product producers are not desired (Fujimura *et al.* 1996).

That is why there are several ways undertaken to improve plasmid stability in yeast cultures. One involves the development of new addiction systems

like the exploitation of the mevalonate pathway, which was established in *E. coli* (Kroll *et al.* 2009), and also introduced into yeast with plasmid stability in non-selective medium of up to 100% in specific mutants. Remarkably, this also seemed to improve significantly the structural stability of the plasmid (Lange and Steinbüchel 2011). Another technique is the periodical fermentation involving oscillation of culturing conditions (Impoolsup *et al.* 1989) like cyclic oxygen tension (Lú-Chau *et al.* 2004), yeast extract concentration (Gupta and Mukherjee 2001), dilution rate (Gupta and Mukherjee 2002), cyclic glucose starvation, growth and production phase separation, cell immobilization or recycling of plasmid carrying cells (Chen *et al.* 1997). However, several of these methods only work at low cell densities (Gupta and Mukherjee 2002) and are explained through a faster response of plasmid free cells to environment factors than plasmid bearing cells (Chen *et al.* 1997). Most of these techniques result in a significant loss of productivity of the culture and therefore a loss of the desired product. To overcome this problem it would be useful to have an online reporter system to follow plasmid stability, which makes it possible to fine-tune counteracting mechanisms against plasmid loss and as a result improve the productivity.

1.3 IMPORTANCE OF A PLASMID DEDECTION SYSTEM

In academia, *E. coli* is widely used because its characteristics make a simpler manipulation possible in comparison to yeasts, making it even easier to produce a proof of principle strain (Hong and Nielsen 2012). Nevertheless the same problems with plasmid stability emerge and this is why more and more techniques for genome integration are developed (Ying *et al.* 2010), like the chemically inducible chromosomal evolution using the RecA site. This method not only produces highly stable constructs, but also overcomes the typical copy number drawback of integration via high antibiotic concentration selection (Kachroo *et al.* 2009). Several established methods are transferred to yeast, because contrary to *E. coli* it has several advantages like being more pH, or osmo-tolerant, and phage resistant (Hong and Nielsen

2012), whereby somehow the expression level of 2µm-based plasmids are not reached by multiple integrations (Romanos *et al.* 1992).

One of these methods is for example the usage of the *LoxP* sequence, which is the target site for the bacterial *Cre* recombinase, thus allowing selective recombination (Leite *et al.* 2012). However, in general it seems that chromosomal insertions are not as structurally stable as plasmid based constructions (Lange and Steinbüchel 2011) and that they inhibit an easy alteration of engineered pathways later on (Dahm and Jennewein 2010). But it has to be kept in mind that every modification in a plasmid leads to a change in behavior in growth and plasmid copy number, which has to be tested (Karim *et al.* 2012).

This points out why both ways of influencing yeasts are likewise important in industrial fermentations and often even used at the same time, like in the work of Ignea *et al.* (2011) for the production of monoterpene, which can be used as a basis for flavors, fragrances, drugs and alternative fuels. In that case, yeast cells selected for their sterol levels were improved with a combination of recyclable integration cassettes and a plasmid-driven co-expression of process essential enzymes thus producing a maximized amount of monoterpene (Ignea *et al.* 2011). In a similar way, a precursor for an artemisinin based malaria medicament is produced by fermentation from engineered and improved *S. cerevisiae*. Around 2.5 g/L of amorpha-4,11-diene could be generated with the help of a single expression plasmid and the complete synthesis of artemisinin acid failed only because of high plasmid instability (Westfall *et al.* 2012). Nevertheless, some industrial processes depend purely on plasmid-based expression, like the production of Hepatitis B vaccines (WHO 1985), the taxoid biosynthesis (Dahm and Jennewein 2010) or the R/S-hydroxy ketones production (Calam *et al.* 2013). All these examples point out how important plasmid based expression systems are and why a plasmid stability reporter system is reasonable, especially if plasmid instability is the only factor inhibiting new products. Noteworthy is also that a patent search on “www.depatisnet.dpma.de” as well as on “www.google.de/?tbm=pts” could not detect such a system.

1.4 ONLINE REPORTER SYSTEM

For construction of a successful online reporter system for plasmid stability, the plasmid loss should be connected to an easily detectable signal. One way of achieving this is to integrate a constitutive signal expression block into the plasmid like described in the work of Hegemann *et al.* (1999) and so as soon as the signal level is 5% lower than the wild type level a plasmid loss could be assumed. However, on the one hand, only 80% of the *S. cerevisiae* cells produced a signal at the outset and on the other hand, a signal increase is easier and more confident to detect.

As an example for a different approach, the work of Bahl *et al.* (2004) about the quantification of plasmid loss in *E. coli* can be consulted. The signal is integrated into the chromosome and controlled by a repressor encoded on the plasmid DNA. As soon as the plasmid is absent, the repressor is lost and a signal is expressed, whereby a delay is possible because of the repressor degradation and signal production. Following the reporter gene this way, its localization and its control have to be investigated. The theoretical basis therefore will be described in the next paragraphs always with the focus on the usage in industrial processes.

1.4.1 REPORTER

The reporter for an online detection system for plasmid loss has to fulfill several characteristics like non-toxicity, or sensitive detection in high cell densities.

A useful reporter gene could be therefore the green fluorescent protein (GFP), which was discovered in the jellyfish *Aequorea victoria* (Ormö *et al.* 1996). The protein structure of GFP has the form of a barrel made out of eleven ß-sheets, with one α-helix inside, and a short helical segment at the end (Yang *et al.* 1996). Through this arrangement, it is possible that the amino acids (aa) Ser^{65}, Tyr^{66} and Gly^{67} form the 4-(*p-hydroxybenzylidene*)-*imidazolidin-5-one* chromophore through an oxidation and dehydration which emits light at

509 nm as soon as it is excited with 395 nm (Cubitt *et al.* 1995). Since only Gly^{67} seems to be important for the functional chromophore forming (Remington 2006), a series of mutation experiments were undertaken to improve the wild type form (Snaith *et al.* 2010). Thereby a change of Ser^{65} to Gly and Ser^{72} to Ala leads to 7.5 faster chromophore assembly, lower background (Welsh and Kay 1997), a 40 times brighter fluorescence, higher solubility and an excitation shift from 395 nm to 488 nm compared to the wild type (Cormack *et al.* 1997). Interestingly, in all variants - as well as in the wild type - cyclizes the protein backbone under anaerobic conditions (Remington 2006) and some variants even form a red fluorescent species of not yet fully understood nature, if they have been completely maturated before (Tsien 1998).

These can only be partially compared to the wild typely found red fluorescent proteins, which undergo an additional oxidation step to form their chromophore even with having a green emitting intermediary, too (Remington 2006). From the mushroom coral *Discosoma striata* the obligatory tetrameric dsRed was isolated and formed through mutations into the dimeric tdTomato as well as via the monomer mRFP1 into the monomer mCherry (Shaner *et al.* 2004). As tdTomato is brighter but mCherry smaller, both are not always comparable to some variants of GFP in regard to photostability, brightness, chromophore maturation time, effects on fused protein and phototoxicity (Snaith *et al.* 2010). This phototoxicity seems to occur in the generation of reactive oxygen at the chromophore leading to a certain toxicity of fluorescent proteins on high expression levels (Tsien 1998).

Nevertheless, red fluorescent proteins have an important advantage over GFP variants regarding the fluorescence background of a cell. Particular yeast cells possess many chemicals, which show diverse forms of fluorescence, too. When some amino acids, reduced nicotinamides, oxidized flavins, as well as lipofuscins are excited at 488 nm, this leads to a background emission peak at around 520 nm. Thereby this autofluorescence is getting stronger as soon as the cells reach the stationary phase (Billinton *et al.* 1998).

The structures of the proteins (Yang *et al.* 1996), but not the fluorescence (Doherty *et al.* 2010) are quite resistant against pH change, whereby the red variants are slightly less affected (Baird *et al.* 2000). In addition the cellular localization in peroxisomes or mitochondria can increase the folding of both (Tsien 1998), but it has to be kept in mind that the fluorescence is strongly influenced by intracellular physiological changes of e.g. the pH (Doherty *et al.* 2010). Another positive aspect of a subcompartimentation is that around 300-3000 fluorescent protein molecules are sufficient for detection by microscopy, whereby the influence on actual non-imaging detection methods seems to be neglectable (Tsien 1998).

When fluorescent proteins are compared to another reporter gene like for example luciferase, which enables a bioluminescent reaction (Baumik and Gambhir 2002), both have most of the time opposing positive and negative features. Whereas fluorescent proteins are stable over several hours (Grilly *et al.* 2007), luciferases have only a half-life of around two hours (Robertson *et al.* 2008), but the degradation process of GFP can be influenced in both directions (Grilly *et al.* 2007). Of course, luciferases avoid autofluorescence, but they bear the deciding negative feature by being dependent on luciferin for a reaction (Greer and Szalay 2002), while fluorescent proteins only need the correct wavelength for activation (Robertson *et al.* 2008). For this reasons the focus of this work is determined to the usage of a fluorescent reporter gene.

1.4.2 Reporter localization

The localization of the reporter can be understood in two manners: One of these is the subcompartimentation of the reporter protein, both constant physiological conditions and signal strength having to be reflected. The other one includes considerations about the localization of the reporter gene on the chromosome, which is important to guarantee high expression and access of control elements.

1.4.2.1Intracellular subcompartimentation of the reporter protein

As mentioned earlier, the volume of yeast cells diverges widely within a culture, but the ratio between cell volume and total major organelle volume stays constant (Uchida *et al.* 2011). In the next paragraphs, the possible effects of the cell compartments cytoplasm, endoplasmic reticulum, cell wall and peroxisomes are discussed briefly regarding their requirements, properties and influence on the signal of fluorescent proteins.

Cytoplasm

With around 70% of the total cell volume the cytoplasm is the biggest compartment in a yeast cell (Uchida *et al.* 2011), and so a large part of the metabolic processes like the glycolysis take place there, too, influencing among other things the pH of this compartment (Sigler *et al.* 1980). This substrate dependent fluctuation effect as well as the growth phase (Valli *et al.* 2005) could be the reason for the different declarations of the cytoplasmic pH alternating around 5 (Valli *et al.* 2006), 6 (Peña *et al.* 2005), and 7, which is normally maintained (Orij *et al.* 2009). Such pH oscillation and the signal dilution through the volume could affect the reporter signal negatively, on the positive side compared to other compartments no additional modification in the sense of a targeting signal added to reporter protein is needed.

Endoplasmatic reticulum

For the co-translational insertion of proteins into the ER lumen following the secretory pathway, special signal sequences on the primary polypeptide chain are essential. These 15-30 aa long sequences have no strong sequence homology, but with the N-terminal region containing often positively charged residues, the hydrophobic region with at least six hydrophobic residues and the C-terminal region with polar uncharged residues three segments can typically be observed (Emanuelsson *et al.* 2007). To remain inside of the soluble ER proteins need ER lumen retention signals involving in most of the cases the amino sequences KDEL or HDEL at the C- terminus (Pagny *et al.* 1999).

However, proteins are not always secreted or structurally stably retained in the ER in yeast. This is depending on the right combination of secretion signal and protein, whereby the signal sequences of various yeast proteins like Bgl2 (YGR282C) (Achstetter *et al.* 1992), Inu1, Pho5 (YBR093C), Mel1 (Li *et al.* 2002), Suc2 (YIL162W) or Pho1 (Q0085) (Hashimoto *et al.* 1998), are already often used, all differing in length. Despite this functional signal sequences combination, some heterologous proteins can fail the ER quality control especially at high expression levels, which leads to an ER associated degradation (Ahner and Brodsky 2004). The pH of the ER is around neutrality (Kim *et al.* 1998) and the volume is quite small (Perktold *et al.* 2007), which offers good conditions for fluorescent proteins if successfully transported and remained in the ER.

Cell wall

The incorporation of proteins into the plasma membrane or cell wall, which represents around 25% of the dry weight (Papanikou and Glick 2009), has some constraints and specific requirements, making specific localization experiments difficult (Smits *et al.* 2006). One way is the attachment of a glycosylphosphatidylinositol (GPI) anchor, which adds the property to a protein to associate with lipid rafts (Castillion *et al.* 2009). Only few GPI proteins stay attached, because most are released from the plasma membrane through a phospholipase C in the periplasm and are anchored in the cell wall by a covalent connection to glucans (Pittet and Conzelmann 2006).

Most of the time heterologous proteins are fused to host cell walls or periplasm membrane proteins with a GPI domain if an attachment to the cell is desired (Ueda and Tanaka 2000). Often used are the C-terminal halfs containing a GPI anchor attachment signal of the proteins α agglutinin (Sag1, YJR004C) or the a-agglutinin anchorage subunit (Aga1, YNR044W), which binds a-agglutinin (Aga2, YGL032C) via disulfide bridges (Cappelaro *et al.* 1994).

Some studies show a connection between GPI anchors and prion diseases, leading to a possible misfolding of the fused protein (Speare *et al.* 2010), which nevertheless has to bypass the ER quality control as well. Therefore, outside of the cell no compartments and compounds of the yeast cell disturb emission and excitation. However, the pH is depending on the environment as well as on the acidification activity of the cell (Sigler *et al.* 1980), and furthermore the reporter could be detached from the cell wall leading to a continuously increasing background noise.

Peroxisome

Most proteins assigned for the peroxisomes carry a signal sequence at the C-terminus with the amino acid combination SKL and are transported through the membrane in a folded state (Brocard *et al.* 1994). The main function of peroxisomes is the β-oxidation of fatty acids and therefore only few or one of them can be observed in full media, but the number increases with upturning metabolism (van der Klei and Veenhuis 1997). The pH of the peroxisomes is about 8.2 (van Roermund *et al.* 2004) and as mentioned earlier fluorescent proteins seem to fold better in this small volume compartment, but in some studies a decrease of fitness was detected by targeting heterologous proteins into the peroxisome (Leskinen *et al.* 2003).

Since no distinct location seems to be optimal, the cytoplasm, ER, cell wall and peroxisome will be tested for their potentials in this study.

1.4.2.2Genome integration of the reporter gene

Having discussed the impact of the reporter protein localization, the focus is now on the influence of the localization respectively integration of the reporter gene in the yeast genome. Identical expression cassettes in the genome of *Saccharomyces cerevisiae* showed an up to 8.7 fold difference in their expression levels in varying positions, indicating the high influence of chromosomal localization or rather epigenetic effects (Flagfeldt *et al.* 2009).

This derives among other factors from the special way eukaryotic chromosomal DNA is organized in nucleosomes, DNA-protein complexes, which strung together with an intervening linker DNA, form the so-called chromatin. In nucleosomes 147 bp of DNA is wrapped 1.7 times around an octamer of the core histones, H2A, H2B, H3, and H4, whereby each occurs twice (Rando and Chang 2009). On this structural level, a high variation is already possible, as the flexible location of the histone complex can disturb the accessibility of regulator sequences. Isoforms of histones as well as a wide range of modifications can alter either the DNA–histone binding affinity, or the stability of nucleosome interactions, which both change the chromatin condensation level. A high condensation, also called heterochromatin, hinders the basal transcriptional machinery and so chromatin modifying protein complexes are required to form the more open euchromatin, since these complexes or related factors are often recruited to more accessible sequences in promoter regions (Uffenbeck and Krebs 2006).

In yeast, roughly two types of promoters can be distinguished from their structure and functionality. Around 80% of yeast promoters depend on the transcription factor IID (TFIID) and possess - compared to the remaining - no typical TATA box DNA sequences as binding site for other transcription factors. Especially TFIID depending promoters appear to be in a nearly nucleosome-free region, whereas TATA box promoters are continuously occupied by randomly delocalized nucleosomes. Thereby the DNA sequence itself can influence histone binding, because dinucleotide AT rich sites are thermodynamically favored to bend and bind histones, whereas regions with long runs of A or T are more stiff (Rando and Chang 2009). This illustrates that dynamic processes and localization play a critical role for the accessibility of transcription factors or regulatory elements in the genome (Buhler and Gasser 2009). A suitable spot was identified in the yeast genome on chromosome 16 represented by a long terminal repeat of a Ty element (YPRCΔ15, Flagfeldt *et al.* 2009).

1.4.3 REPORTER CONTROL

After having discussed the type of reporter and its localization, the control of it comes into focus, especially its expression control. This can on the one hand be conducted by the promoter of the reporter gene, which should allow constituive expression under any environmental condition after initiation and produce thus a strong reporter signal, while it should also not be the strongest promoter to limit complications for industrial applications. On the other hand, the most important part of the online detection system is the direct control of the promoter via a repressor. This repressor should be cell based, highly specific with strong repression and should not disturb the normal cell function. Moreover, its expression cassette should be plasmid based with a balanced intensity to enable repression without decreasing cell fitness and its concentration connected to the plasmid loss.

1.4.3.1Promoter of reporter gene

As already mentioned, yeast promoters can be distinguished via the absence or presence of a typical TATA box, and yet A/T-richness of intergenic regions and missing consensus sequences make this determination not easy. Nevertheless, a multiplicity of TATA box containing promoters seem to be more stress responsive, noisier in gene expression, but more controlled and able to cover higher as well as lower expression intensities, thus separating the totality of yeast promoters into housekeeping promoters, being mostly constitutive and TATA box bearing promoters, being responsible for stress response (Basehoar *et al.* 2004). Next to the TATA box, as part of the core promoter region, promoters possess an upstream region where most of the transcription activators, enhancers, or *trans*-regulatory elements bind to the so-called upstream activating sequences (UAS) (Blount *et al.* 2012).

Especially these elements have an enormous influence on the expression of a promoter leading to the possibility to improve even the strongest constitutive promoters by additional UAS elements or an overcoming of the enhancer limitation (Blazeck *et al.* 2012). The finding that hybrid transcription factors made of bacterially DNA-binding domains and eukaryotic transcription

activation domains are working through the recruitment ability of the factor, underlines the importance of these activators (Ptashne 2011). In addition, these factors seem to have a stronger influence on the alluded extrinsic noise by generating mRNA fluctuations, than the potential effects of the chromatin structure and rearrangement (Bar-Even *et al.* 2006).

However, the ability of high mRNA production aroused through a wild type or modified promoter does not automatically lead to a high protein production. On the one hand, the promoter could alter the overall fitness or expression of other genes through exhaustion of essential and rare transcription activators, while on the other hand translational and posttranslational feedback mechanisms play an important role (Lul and Jeffries 2007). This seems to affect especially heterologous genes, on which compared to the wild type 3-phosphor glycerate kinase 1 (YCR012W, $PGK1_{prom}$) promoter product a reduction of protein production from 50% to 2% of total soluble protein has been observed (Romanos *et al.* 1992). Only through integration of around 120 gene copies restored to 37% (van der Aar *et al.* 1989). This is attributed for example to a missing internal or downstream activating sequence, decreasing mRNA or protein stability and improper codon usage (Romanos *et al.* 1992). Moreover, this change might be based on a less efficient Kozak sequence, preventing start codon recognition by the scanning mechanism of the ribosome (Kozak 1999). Furthermore, the influence of DNA copy number to a certain degree in most cases correlates with protein abundance, but sometimes yeast already seems to express genes more than needed and favors direct regulation over feedback regulation (Springer *et al.* 2010).

In summary, this leads to the importance of a previous characterization of yeast homologous or derivate promoters in connection with copy number and chosen heterologous protein before experimenting. As a support for the optimal promoter decision a multiplicity of promoter analyses are already conducted and information is available for many promoters. Described are for example different environments (Monfort *et al.* 1999), time (Partow *et al.* 2010), mutability (Nevoigt *et al.* 2006), transcription factor dependency

(Görgens *et al.* 2000), copy number (Nacken *et al.* 1996), and relative strength in comparison to others (Mumberg *et al.* 1995).

Based on all depicted information, the relatively strong constitutive promoters of the gene triose-phosphate dehydrogenase (YGR192C, $TDH3_{prom}$), 3-phospho glycerate kinase ($PGK1_{prom}$), alcohol dehydrogenase (YOL086C, $ADH1_{prom}$) and actin (YFL039C, $ACT1_{prom}$) were analysed. Moreover, the weaker constitutive promoters of a subtilisin-like processing protease (YNL238W, $KEX2_{prom}$), a subunit histone deacetylase (YJR025C, $HAD1_{prom}$) and a tRNA ligase (YJL087C, $TRL1_{prom}$) were tested in this study, thus covering a wide spectrum of expression intensities.

1.4.3.2 Repressor system

The multi-step process from DNA to enzyme activity enables multiple ways of interfering and influence, whereby next to cell based, also chemical molecules like polyamides can be applied (Gottesfeld *et al.* 1997) as well as even physical signals like electric fields (Hasty *et al.* 2000). However, the control particularly of the transcription in eukaryotes is essentially based on activation, while in prokaryotic cells repression is widely spread, especially through their organization of multiple metabolic-interrelated genes in operons (Kærn *et al.* 2003). In contrast to the eukaryotic activators, the relative position of repressors on the promoter region is important for their function particularly because some can even act in a different location as activators (Ptashne 2011).

There are examples of prokaryotic repressor systems which have been transferred to eukaryotic cells and working there, given that only a minor if any interference with the host cell's metabolism must be expected (Kærn *et al.* 2003), and basal expression could be eliminated, thus improving the desired method outcome (Ptashne 2011). A strong bacterial based repressor system can be established from phages, which are viruses specialized on bacteria and archaea. Especially the group of temperate phages, for which it

is characteristic that the genome of the virus is transitionally integrated into the genome of the host and nearly complete silenced, seems reasonable to investigate, because it can shift from this lysogenic into a lytic state through a protein-based switch. Nearly all temperate phages control this switch roughly in a similar way, like the mycobacteriophage L1 (Bandhu *et al.* 2009), the lambda, p22, or 434 phage, differing just in DNA and protein sequence (Bushmann 1992). Best studied is the lambda phage and will be therefore described next.

Briefly, the proteins Cro (activating lytic state) and CI (maintaining lysogenic state) regulate the switch by acting both timely staggered as activators and repressors enhancing or preventing RNA polymerase binding by competing for the same operator sites, however CI has the higher affinity (Bakk *et al.* 2004). In the lambda genome two of such operator sites exist, which are separated around 2.4 kpb and called O_L and O_R site (Flashman 1978). Thereby each site contains three 17 bp long, 3 to 7 bp spaced, binding sequences named O_R or O_L 1, 2, and 3, each of them differing slightly in their sequence leading to varying affinity of Cro and CI to them (Gedeon *et al.* 2008). For the control of the switch in particular the O_R sites are important, located between the coding sequences for Cro and CI (see figure 1 A). In the lysogenic state dimeric CI binds the O_R1 site followed by O_R2, forming a tetramer afterwards. Through this, the expression of Cro is blocked and the expression of CI is enhanced, until a sufficient high concentration of CI is reached to bind O_R3, inhibiting in turn the expression of CI (Hochschild and Lewis 2009). The same binding occurs at the operator site O_L whereby a loop and a CI octamer are formed (see figure 1 B). This leads on the one hand at the beginning to a higher expression activity for CI (Anderson and Yang 2008) and on the other hand to a more stable maintenance of the lysogenic state (Zurla *et al.* 2009).

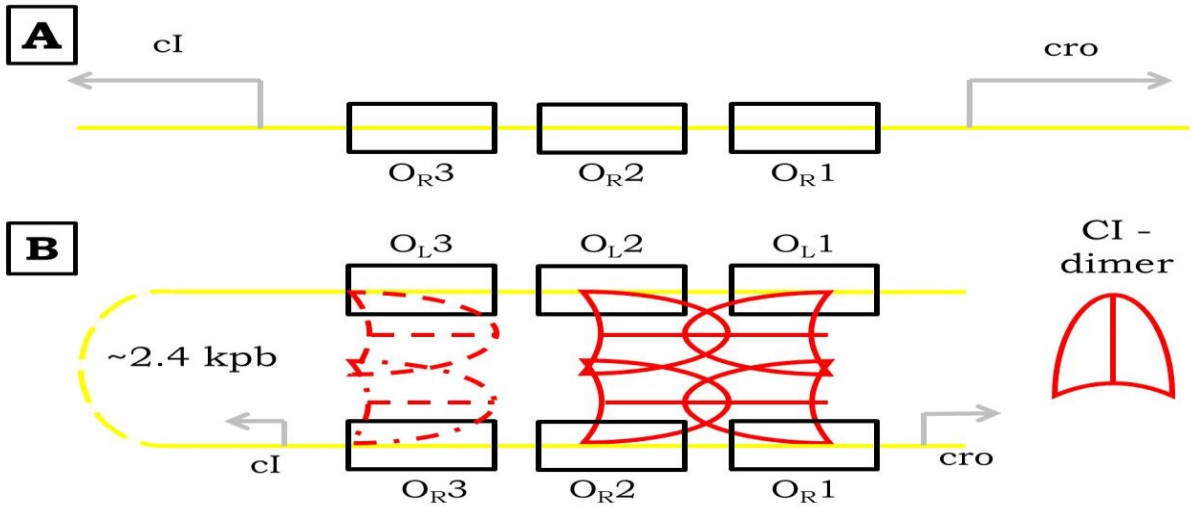


Figure 1: **Promoter organization of phage lambda bidirectional promoter region at cI and cro**. In (A) organization of operator sequences O_R1, O_R2, and O_R3 is shown not occupied by CI. In (B) operator sequences O_R1, O_R2, and O_R3, as well as operator sequences O_L1, O_L2, and O_L3 are shown with the CI octamer occupying O_R1 and O_R2 on one hand, and O_L1 and O_L2 on the other hand. The loop formed encompasses coding sequences of CI (all sequences not drawn to scale).

The condition is so stable that less than 2 x 10^{-9} per cell per generation of spontaneous switch activations are observed making sequence mutations of the involved lambda genes or the activation of the host SOS response the only factors abolishing the lysogenic state (Morelli *et al.* 2009). The SOS response is activated through UV light irradiation (Gedeon *et al.* 2008), or other DNA damage, as well as factors like temperature or starvation and initiates the expression of *e.g.* RecA. This among others functions triggers the cleavage and thus the inactivation of the lambda CI repressor (Avlund *et al.* 2009), regardless of its dimerization (Pal and Chattopadhyaya 2009) or binding to DNA (Pal and Chattopadhyaya 2012). In this situation, Cro can be expressed, occupying the site O_R3 also as a dimer but with a higher affinity than CI for it, thus leading to an inhibition of CI expression. Gradually all operator sites are bound by Cro, which only forms dimers, and the genes for the initiation of the lytic state are activated (Dodd *et al.* 2005).

Especially the repressor CI is used in some construction to alter promoter expression in yeast. For instance, Wedler and Wambutt (1995) exchanged the wild type TATA box of $TDH3_{prom}$ with an artificial TATA box flanked by O_L1 and O_L2. This construct did not show the wild type activity of the promoter, but allowed an up to 85% repression with the 266 aa CI present (Wedler and Wambutt 1995). Another example is the use of the 434-phage

operator site in different location of $PGK1_{prom}$, showing most of the time despite insertion wild type activity and a repression between 16 and 61% in presence of the CI repressor (Webster and Brammer 1995).

Several other repressor systems like the usage of RNA, the TET system or the Lac system were considered, but none of the found repressor systems completely fulfilled the requested properties for an online reporter system for plasmid stability. The RNA systems are connected with a high effort and preliminary work, because the introduction of the RNAi system to yeast (Suk *et al.* 2011), as well as the development of suitable antisense or riboswitch RNA (Bayer and Smolke 2005) is time consuming and after that not flexible applicable. Furthermore, the regulating functions of riboswitches depend on ligands (Waters and Storz 2009), whose concentration is most of the time very complex to connect to plasmid loss. This problem is especially obvious with the Tet (Kærn *et al.* 2003) and the Lac system, where the ligands are either an in industrial scales costly antibiotic (Gari *et al.* 1997), or part of a metabolism (Wheatley *et al.* 2013), which is not present in yeast, respectively its chemical analog. Nevertheless, a usage of these repressors is still possible given that such repressors are assumingly removed from their repressor site by degradation. However, the tight repression via the Tet system is connected to a constricted promoter (Maya *et al.* 2008) and costly patents for industrial usage (Loew *et al.* 2011). Such constraints are not connected with the *lac* system, but its repression is not very stable leading to higher background in the system (Lewis 2011).

In contrast to that, the phage system is based on a single small repressor protein with no necessity of an additional ligand and the degradation of it is not only supported by RecA, but also in a smaller extent non-enzymatically through the amino acids histidine and lysine (Pal and Chattopadhyaya 2012). However, phage repressors are – in contrast to Tet or Lac repressors - seemingly kicked off by RNA polymerase, prohibiting the insertion of operator sites inside of the open reading frame (Webster and Brammer 1995). Nevertheless, through its regulation of an all-or-nothing decision, the phage system in nature is one of the best and tightest controlled

mechanisms, functioning under different conditions also in yeast (Wedler and Wambutt 1995). It is therefore best suited to be the basis for the repression control of the reporter gene. Unfortunately, in yeast the same degree of repression has not yet been accomplished.

1.5 OBJECTIVES

In summary, several objectives arose to develop an online reporter system to follow plasmid stability in *Saccharomyces cerevisiae*. First of all a stable expression of fluorescent proteins is needed with an optimized subcellular location for signal strength. Tested will be red as well as green variants of fluorescent proteins and different constitutive yeast promoters, which all are still not the strongest available. An important point will be hereby also the localization, though is the genome location already choosen, the localization inside the different cell compartments is still unknown.

After these experiments, the most important point is the control of the promoter by the lambda phage based repressor system. Again, localization and number of operator sites in the promoter of the signal protein, expression intensities as well as repressor-to-operator affinity improvments have to be reflected.

Finally, investigations about plasmid stability under selective and nonselective conditions have to be considered to connect the plasmid loss to a correlating fluorescent signal.

2 MATERIALS AND METHODS

2.1 CHEMICALS, BUFFERS AND SOLUTIONS

Unless otherwise stated all chemicals, reagents, and components for media and agar were purchased from the following companies: VWR (Darmstadt, Germany), Sigma Aldrich (Seelze, Germany), Merck (Darmstadt, Germany), Applichem (Darmstadt, Germany), Serva (Heidelberg, Germany), or Roth (Karlsruhe, Germany). The chemicals had at least analytical grade p.A. or were tested for their molecular biological usage. In all reactions water from the water clearance machine Ultra clear UV plus Basic (SG Water clearance and regeneration, Hamburg, Germany) was used.

2.2 OLIGONUCLEOTIDES

Primers for PCR and oligonucleotides for cloning experiments are listed in the following table 1 concerning name, sequence, origin, application, melting temperature (Tm), and additional sites, which are marked in bold or colors. The oligonucleotides were purchased from Biomers (Ulm, Germany) and are separated in reporter gene, repressor gene, promoter mutation, and genome insertion for a better overview.

Table 1: Description of primers used in this study. Additional sites are marked respectively in bold or colors in the sequence of the primer. The first temperature (Tm) is connected to the binding of primers without additional sites and the second to the binding of the complete primer.

Name	Sequence 5´ - 3´	Application	Tm [°C]	Additional sites
Reporter gene				
forJOE0001	AT**GAGCTC**TTTCAGCAGCTCTGATGTAG	*KEX2*$_{prom}$ sequence from *S.c.*	53/57	*Sac*I
revJOE0002	AT**GTCGAC**TAGTAGTTTATAATTATGTGACGAGGCC	*KEX2*$_{prom}$ sequence from *S.c.*	55/61	*Sal*I
forJOE0024	CT**GAGCTC**TTTCAGCAGCTCTGATGTAG	*KEX2*$_{prom}$ sequence from *S.c.*	58/60	*Sac*I

revJOE0027	CT**TCTAGA**TGCAGAT ATTTTGGCTGCAAA ACCAGCCAAAAGGA AAAGGAAAGCTTGC AAAAGCATCTGATAA TGGGTTAGTAGTTTA TAATTATGTG	*KEX2*$_{prom}$ sequence from *S.c.*	56/ 84	*Xba*I *SUC2*$_{secr}$
forJOE0052	CT**GCGCGC**TTTCAGC AGCTCTGATGTAG	*KEX2*$_{prom}$ sequence from *S.c.*	58/ 71	*Pau*I
revJOE0053	CT**TCTAGA**GGGCATC TGATAATGGGTTAGT AGTTTATAATTATGTG	*KEX2*$_{prom}$ sequence from *S.c.*	56/ 69	*Xba*I
revJOE0067	CT**GCATGC**CCATCTG ATAATGGGTTAGTAG TTTATAATTATGTG	*KEX2*$_{prom}$ sequence from *S.c.*	62/ 69	*Sph*I
revJOE0040	GT**TCTAGA**TGCAGA TATTTTGGCTGCAAA ACCAGCCAAAAGGA AAAGGAAAGCTTGC AAAAGCATTGTTTTA TATTTGTTGTAAAAAG TAGATAATTAC	*PGK1*$_{prom}$ sequence from *S.c.*	58/ 78	*Xba*I *SUC2*$_{secr}$
revJOE0041	CATTGTTTTATATTTG TTGTAAAAAGTAGAT AATTAC	*PGK1*$_{prom}$ sequence from *S.c.*	58	-
forJOE0042	TA**GCGCGC**AGATTCC TGACTTCAACTCAAG AC	*PGK1*$_{prom}$ sequence from *S.c.*	61/ 72	*Pau*I
revJOE0050	CT**TCTAGA**GGGCATT GTTTTATATTTGTTGT AAAAAGTAGATAATTA C	*PGK1*$_{prom}$ sequence from *S.c.*	56/ 67	*Xba*I
revJOE0066	CT**GCATGC**CCATTGT TTTATATTTGTTGTAA AAAGTAGATAATTAC	*PGK1*$_{prom}$ sequence from *S.c.*	60/ 67	*Sph*I
revJOE0071	ATCGT**GCATGC**C**GA GATTTGAAACACTTA ATAGTGCTAGAACT AACGATTTATAGTTA ACCATGATTGTTGA AAAAATGTTAAATAA TGAGGAAACGCTCT CCAT**TGTTTTATATTT GTTGTAAAAAGTAGA T	*PGK1*$_{prom}$ sequence from *S.c.*	58/ 78	*Sph*I virus *K28* secretion signal

revJOE0043	GT**TCTAGA**TGCAGA TATTTTGGCTGCAAA ACCAGCCAAAAGGA AAAGGAAAGCTTGC AAAAGCATTGTATAT GAGATAGTTGATTGT ATGCTTGGTATAG	*ADH1*$_{prom}$ sequence from *S.c.*	58/ 80	*Xba*I *SUC2*$_{secr}$
revJOE0044	CATTGTATATGAGAT AGTTGATTGTATGCT TGGTATAG	*ADH1*$_{prom}$ sequence from *S.c.*	58	-
forJOE0045	TA**GCGCGC**GTAATAA TAGGCGCATGCAACT TCTTTTC	*ADH1*$_{prom}$ sequence from *S.c.*	60/ 74	*Pau*I
revJOE0064	CT**GCATGC**CCATTGT ATATGAGATAGTTGA TTGTATGCTTGGTAT AG	*ADH1*$_{prom}$ sequence from *S.c.*	65/ 71	*Sph*I
forJOE0065	TC**GCGCGC**TGCAACT TCTTTTCTTTTTTTTTC TTTTC	*ADH1*$_{prom}$ without *Sph*I from *S.c.*	61/ 70	*Pau*I
revJOE0049	CT**TCTAGA**GGGCATT GTATATGAGATAGTT GATTGTATGCTTGGT ATAG	*ADH1*$_{prom}$ sequence from *S.c.*	58/ 71	*Xba*I
revJOE0046	GT**TCTAGA**TGCAGA TATTTTGGCTGCAAA ACCAGCCAAAAGGA AAAGGAAAGCTTGC AAAAGCATTGTTAAT TCAGTAAATTTTCGAT CTTGGGAAGAA	*ACT1*$_{prom}$ sequence from *S.c.*	56/ 80	*Xba*I *SUC2*$_{secr}$
revJOE0047	CATTGTTAATTCAGTA AATTTTCGATCTTGG GAAGAA	*ACT1*$_{prom}$ sequence from *S.c.*	56	-
forJOE0048	TA**GCGCGC**TTAATGC ACAACATTTAACCTA CATTCTTCCTTATC	*ACT1*$_{prom}$ sequence from *S.c.*	59/ 73	*Pau*I
revJOE0051	CT**TCTAGA**GGGCATT GTTAATTCAGTAAATT TTCGATCTTGGGAAG AA	*ACT1*$_{prom}$ sequence from *S.c.*	59/ 72	*Xba*I
forJOE0063	CT**GCATGC**CCATTGT TAATTCAGTAAATTTT CGATCTTGGGAAGAA	*ACT1*$_{prom}$ sequence from *S.c.*	66/ 72	*Sph*I

forJOE0068	CT**GCATGC**TTGCAGATATTTTGGCTGCAAAAC	*Suc2*$_{secr}$ for *ACT1*$_{prom}$/ *ADH1*$_{prom}$/ *PGK1*$_{prom}$ from JOE0005/6/7	62/ 70	*Sph*I
forJOE0069	AGTT**GCATGC**T**TGCAGATATTTTGGCTGCAAACCAGCCAAAAGGAAAAGGAAAGCTTGCAAAAGCAT**GAATTCTTTGTTTGTTTATGTGTG	*TDH3*$_{prom}$ from pRS316- mCherry	60/ 81	*Sph*I *SUC2*$_{secr}$
revJOE0070	TA**GCGCGC**AATAAAAACACGCTTTTTCAGTTC	*TDH3*$_{prom}$ from pRS316- mCherry	59/ 70	*Pau*I
forJOE0028	CT**GAATTC**TTTGTACAATTCATCCATACCATGGGTAATAC	yEGFP3 sequence from pUG35	63/ 66	*Eco*RI
revJOE0029	CT**TCTAGA**TCTAAAGGTGAAGAATTATTCACTGGTGTTGTC	yEGFP3 sequence from pUG35	61/ 65	*Xba*I
revJOE0062	CT**GCATGC**AATCTAAAGGTGAAGAATTATTCACTGGTGTTGTC	yEGFP3 Sequence from pUG35	66/ 71	*Sph*I
forJOE0036	CT**TCTAGA**GTTTCAAAAGGTGAAGAAGATAATATGG	mCherry sequence from pRS316-mCherry	61/ 65	*Xba*I
revJOE0037	CT**GAATTC**TTTATATAATTCATCCATACCACCAGTTG	mCherry sequence from pRS316-mCherry	61/ 66	*Eco*RI
forJOE0060	CT**GCATGC**AAGTTTCAAAAGGTGAAGAAGATAATATGG	mCherry sequence from pRS316-mCherry	61/ 68	*Sph*I
forJOE0038	CT**TCTAGA**GTGAGCAAGGGCGAGGAGGTCATC	tdTomato sequence from pks392	70/ 73	*Xba*I
revJOE0039	CT**GAATTC**CTTGTACAGCTCGTCCATGCCGTAC	tdTomato sequence from pks392	69/ 73	*Eco*RI

forJOE0061	CT**GCATGC**AAGTGAG CAAGGGCGAGGAGG TCATC	tdTomato sequence from pks392	71/ 76	*Sph*I
revJOE0054	TC**GCCGGC**TCACTAT AGGGCGAATTGG	*CYC1*$_{term}$ sequence from pUG35	58/ 71	*Nae*I
forJOE0055	TC**GAATTC**CATCATC ATCATCATCATCATC ATCTCGAG**AAAGATG AACTA**TAA**CCGCGG**C TGGTCGAGTCATGTA ATTAGTTATG	*CYC1*$_{term}$ sequence from pUG35 with KDEL and His(8)tag	61/ 78	*Eco*RI His(8)tag *Xho*I KDEL *Sac*II
forJOE0056	CT**GAATTC**CATCATC ATCATCATCATCATC ATCTCGAG**AGTAAAC TA**TAA**CCGCGG**CTGG TCGAGTCATGTAATTA GTTATG	*CYC1*$_{term}$ sequence from pUG35 with SKL and His(8)tag	61/ 78	*Eco*RI His(8)tag *Xho*I SKL *Sac*II
forJOE0057	CTGGTCGAGTCATGT AATTAGTTATG	*CYC1*$_{term}$ sequence from pUG35	61	-
forJOE0030	CT**GAATTC**AGCGCCA AAAGCTCTTTTATCT	GPI anchor region *SAG1* fragment from *S.c.*	54/ 64	*Eco*RI
revJOE0031	CTTT**GCCGGC**TACTT AATCAGTGGTCTACT GAAAC	GPI anchor region *SAG1* fragment from *S.c.*	56/ 67	*Nae*I
forJOE0058	TC**CTCGAG**GTGGTGA CTACCAGCACAAAAC	Short GPI anchor region *SAG1* fragment from *S.c.*	60/ 70	*Xho*I
revJOE0059	CT**CCGCGG**TTAGAAT AGCAGGTACGACAAA AG	Short GPI anchor region *SAG1* fragment from *S.c.*	63/ 70	*Sac*II

Repressor gene

forJOE0072	TA**GCATGC**AAAGCAC AAAAAAGAAACCATTA AC	CI from λ phage DNA	58/ 67	*Sph*I
revJOE0073	AT**CTCGAG**TTAGCCA AACGTCTCTTCAGGC CACTGACT	CI from λ phage DNA	72/ 75	*Xho*I
forJOE0074	GCATGCAA**CCCAAGA AAAAGCGCAAGGTA**A GCACAAAAAAGAAAC CATTAAC	CI from λ phage DNA with NLS	58/ 77	*Sph*I NLS
revJOE0075	AT**CTCGAG**CTTA**TAC CTTGCGCTTTTTCTT GGG**GCCAAACGTCTC TTCAGGCCACTGAC	CI from λ phage DNA with NLS	71/ 82	*Xho*I NLS
forJOE0076	TA**GCATGC**AA**CATCA TCATCATCATCAT**AG CACAAAAAAGAAACC ATTAAC	CI from λ phage DNA with His(8)tag	58/ 72	*Sph*I His(8)tag
revJOE0077	TA**CTCGAG**TTA**ATGA TGATGATGATGATG** GCCAAACGTCTCTTC AGGCCACTGACTA	CI from λ phage DNA with His(8)tag	71/ 78	*Xho*I His(8)tag
forJOE0078	ATGTTCTCACCTGAG CTTAGA	Intern1 CI from λ phage DNA	60	-
revJOE0079	CTAAGCTCAGGTGAG AACAT**TCTAGT**CTGA ACATGAGAAAAAACA GGGTACTCAT	Intern2 CI from λ phage DNA	65/ 75	RecA site mutation
revJOE0080	CTAAGCTCAGGTGAG AACAT**TCTAGTTACC TTGCGCTTTTTCTTA GG**GTACTCATACTCA CTTCTAA	Intern 2 CI from λ phage DNA with NLS	59/ 77	RecA site mutation NLS
forJOE0081	TT**GCATGC**TTAATCT GTTCGGCGTAT	*HDA1*$_{prom}$ sequence from *S.c.*	63/ 67	*Sph*I
revJOE0082	TT**GAGCTC**CAGTACC AGCGTCTGCATAG	*HDA1*$_{prom}$ sequence from *S.c.*	62/ 71	*Sac*I
forJOE0083	TT**GAGCTC**ACTGCCA GCACTTTAAGG	*TRL1*$_{prom}$ sequence from *S.c.*	59/ 69	*Sac*I

revJOE0084	TT**GCATGC**TTCTTCG TATGAATACTT	*TRL1*$_{prom}$ sequence from *S.c.*	54/ 63	*Sph*I
forJOE0103	TT**GAGCTC**GCTTGTC CAGACTGAATG	*TRL1*$_{prom}$ long from *S.c.*	57/ 66	*Sac*I
forJOE0085	AT**GAGCTC**TTTCAGC AGCTCTGATGTAG	*KEX2*$_{prom}$ sequence from *S.c.*	59/ 66	*Sac*I
revJOE0086	AT**GCATGC**CCATCTG ATAATGGGTTAGTAG	*KEX2*$_{prom}$ sequence from *S.c.*	54/ 68	*Sph*I
revJOE0115	CTCTTGTTACCCATCA TTG	2µm Ori from yEpGAP-mCherry	55	-
forJOE0116	ATCAGAGCAGATTGT ACTG	2µm Ori from yEpGAP-mCherry	55	-
revJOE0117	**GCAGTGACTCCTAG CGCTCACCAAGCTCT TAAAACGGGAATTTA TG**ATCAGAGCAGATT GTACTG	2µm Ori from yEpGAP-mCherry, Ω PCR	55/ 80	pUG35
forJOE0118	**CATTTATCAGGGTTA TTGTCTCATGAGCG GATACATATTTGAAT GTA**CTCTTGTTACCC ATCATTG	2µm Ori form yEpGAP-mCherry, Ω PCR	55/ 75	pUG35

Promoter modifications

forJOE0087	**AGGTAAAATAGTCAA CACGCACGGTGTTA GATATTTATCCCTTG CGGTGATAGATTTAA CGTA**TATTTGGTCTTT TCTAATTCGTAGT	Pos1 from JOE0031	58/ 78	O_R
revJOE0088	**TGCGTGTTGACTATT TTACCTCTGGCGGT GATAATGGTTGCA**GT ATTTCTTGCATTGACC AATTT	Pos1 from JOE0031	58/ 79	O_R

forJOE0089	**AGGTAAAATAGTCAA CACGCACGGTGTTA GATATTTATCCCTTG CGGTGATAGATTTAA CGTA**CATATATATAAA CTTGCATAAATTG	Pos2 from JOE0031	54/ 76	O_R
revJOE0090	**TGCGTGTTGACTATT TTACCTCTGGCGGT GATAATGGTTGCA**AA ATGTAAGTTTCACGA GGTTCTA	Pos2 from JOE0031	60/ 79	O_R
revJOE0093	**TCTAACACCGTGCGT GTTGACTATTTTACC TCTGGCGGTGATAAT** ATATTTGGTCTTTTCT AATTCGTAGT	nPos1+/- from JOE0031	58/ 77	O_R
forJOE0094	**AGTCAACACGCACG GTGTTAGA**GTATTTC TTGCATTGACCAATTT	nPos1+ from JOE0031	58/ 75	O_R
forJOE0095	**AGTCAACACGCACG GTGTTAGA**ATGTAAG TTTCACGAGGTTCTA C	nPos1- from JOE0031	60/ 75	O_R
revJOE0096	**TCTAACACCGTGCGT GTTGACTATTTTACC TCTGGCGGTGATAAT** ACATATATATAAACTT GCATAAATTG	nPos2+/- from JOE0031	54/ 76	O_R
forJOE0097	**AGTCAACACGCACG GTGTTAGA**AAATGTA AGTTTCACGAGGTTC TA	nPos2+ from JOE0031	60/ 75	O_R
forJOE0098	**AGTCAACACGCACG GTGTTAGA**GTTAGAA GAAAAGAGTGTGTGA GAAC	nPos2- from JOE0031	61/ 76	O_R
revJOE0099	**TCTAACACCGTGCGT GTTGACTATTTTACC TCTGGCGGTGATAAT** TCTGGAATGGCGGGA AAGG	nPos3+/- from JOE0031	64/ 81	O_R
forJOE0100	**AGTCAACACGCACG GTGTTAGA**TGTTGTG TGACGAAATTG	nPos3+ from JOE0031	55/ 75	O_R
forJOE0101	**AGTCAACACGCACG GTGTTAGA**ACCTTCG ATTGCTTGTTACA	nPos3- from JOE0031	59/ 76	O_R

revJOE0112	TT**ATCACCGCCGGTG ATAT**ATAAATT**ATCAC CGCCGGTGATAT**GC AAGAAATACATATTTG GTCTTTTC	TAT from JOE0031	58/7 8	O_R
forJOE0113	**TATATCACCGGCGGT GATAA**ATGTAAATGT AAGTTTCACGAGGTT CTAC	TAT from JOE0031	57/7 8	O_R
forJOE0102	AT**GCGCGC**TAGAACC TCGTGAAACTTAC	*PGK1*$_{prom}$ short from *S.c.*	56/7 0	*Pau*I
forJOE0091	AT**GCCGGC**CCCGGC CTCAGCGCCGGGTTT TCTTTG	OL from λ phage DNA	77/ 84	*Nae*I
revJOE0092	CG**ATGCAT**GTGCTCA GTATCACCGCCAGTG GTATTTATG	OL from λ phage DNA	70/7 3	*Nsi*I
forJOE0114	ATCTGCAGTATCACC GCCGGTGATATATTT ATTATCACCGCCGGT GATACTGCAGTG	OL from λ phage DNA		-

Genome insertions and controls

Name	Sequence 5´ - 3´	Application	Tm [°C]	Additional sites
forJOE0003	CTTCCACGCCAGTCT CATTG	Insertions verification of pUG35 in *S.c.*	53	-
revJOE0004	TTCACCTTCACCGGA GACAG	Insertions verification of pUG35 in *S.c.*	52	-
forJOE0005	GGAACGTGCTGCTAC TCATCC	Verification of *URA3* Sequence from pUG35	54	-
revJOE0006	TTGCTGGCCGCATCT TCTC	Verification of *URA3* Sequence from pUG35	56	-

revJOE0007	TTCCTTTCTCGCCAC GTTCG	Verification of vector and *URA3* sequence from pUG35	58	-
forJOE0008	GGGTGGAAGAGATGA AGGTTAC	Verification of vector and *URA3* sequence from pUG35	51	-
revJOE0009	TGCTTCCGGCTCCTA TGTTG	Insertions verification of pUG35 in *S.c.*	55	-
revJOE0104	**GTAGTTTTAAAATTT CAAATCCGAACAACA GAGCATAGGGTTTC GCAAAC**CGCAGATTC CTGACTTCAAC	*Pau*I Insertion from JOE0060-70, 73-78	60/ 78	*YPRCΔ15*
forJOE0105	**GTTCTGTATTGTTCT TCTTAGTGCTTGTAT ATGCTCATCCCGACC TTCCATT**CGGCATCA GAGCAGATTG	*URA3* insertion from JOE0060-65, 73-78	60/ 79	*YPRCΔ15*
forJOE0106	**GTTCTGTATTGTTCT TCTTAGTGCTTGTAT ATGCTCATCCCGACC TTCCATT**CACTATAG GGCGAATTGG	*Nae*I insertion from JOE0060-65	56/ 78	*YPRCΔ15*
forJOE0107	**GTTCTGTATTGTTCT TCTTAGTGCTTGTAT ATGCTCATCCCGACC TTCCATT**AAATGCAT GTGCTCAGTATC	O_L insertion from JOE0066-70	58/ 78	*YPRCΔ15*
forJOE0108	CAGATGAATGGACGC GAATGC	Insertion test *YPRCΔ15*	64	-
revJOE0109	TTCACCTTCACCGGA GACAG	Insertion test yEGFP3	63	-
forJOE0110	**CACGTCAAGACAAAG AAAGAAAGAAAGAAA AACTAACACATTAAT** GTAGTTTTAAAATTTC AAATCCGAACAAC	Extension insertion *Pau*I	58/ 74	*YPRCΔ15*

revJOE0111	**TCAAAACACTCGGTT TTACTCGAGCTTGTA GCACAATAATACCGT GTAGA**GTTCTGTATT GTTCTTCTTAG	Extension insertion O_L, *Nae*I, *URA3*	60/ 76	*YPRCΔ15*

2.3 STRAINS AND PLASMIDS

The strains and plasmids used in this study as well as their construction ways are listed in table 2 and table 3. All plasmids established in this work, were controlled by either restriction digestion (see 2.5.4) or sequencing (see 2.5.9) respectivly both. Some selected plasmids are shown in figure 2 and are colored (yellow) accordingly to this in table 3 as well as plasmids, which were partially sequenced (green, see for details Appendix 7.4).

Table 2: Description of strains used in this study.

Organism	Description	Reference
***Escherichia coli* DH5α**	F', *rec*A1, *end*A1, *hsdR17,* (r_k^-,m_k^+) *pho*A, *supE44, thi-1, relA1,* λ$^-$, Φ 80*lacZΔM15,* (*lacZYA-argF*), U169, *deo*R	Life Technologies, Darmstadt, Germany
Saccharomyces cerevisiae	wild type, Meyen ex E.C. Hansen, DSMZ70468	DSMZ, Braunschweig, Germany
DBY947	*MATa, ade2-101, ura3-52*	Neff *et al. 1983*
SY992	*MATa, ura3Δ0, his3Δ1, leu2Δ0, trp1Δ63, ade2Δ0, lys2Δ0, ADE8*	Tomlin *et al.* 2001
YJOE0001	Homologous integration of *BmgB*I cut JOE0001 into $KEX2_{prom}$ region of *SY992*	This study
YJOE0002	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0105 and forJOE0110 / revJOE0111 from JOE0060 into *YPRCΔ15* region of SY992	This study

YJOE0003	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0105 and forJOE0110/ revJOE0111 from JOE0061 into *YPRCΔ15* region of SY992	This study
YJOE0004	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0105 and forJOE0110/ revJOE0111 from JOE0062 into *YPRCΔ15* region of SY992	This study
YJOE0005	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0105 and forJOE0110/ revJOE0111 from JOE0063 into *YPRCΔ15* region of SY992	This study
YJOE0006	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0105 and forJOE0110/ revJOE0111 from JOE0064 into *YPRCΔ15* region of SY992	This study
YJOE0007	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0105 and forJOE0110/ revJOE0111 from JOE0065 into *YPRCΔ15* region of SY992	This study
YJOE0008	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0105 and forJOE0110/ revJOE0111 from JOE0073 into *YPRCΔ15* region of SY992	This study
YJOE0009	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0105 and forJOE0110/ revJOE0111 from JOE0074 into *YPRCΔ15* region of SY992	This study
YJOE0010	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0105 and forJOE0110/ revJOE0111 from JOE0075 into *YPRCΔ15* region of SY992	This study

YJOE0011	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0105 and forJOE0110/ revJOE0111 from JOE0076 into *YPRCΔ15* region of SY992	This study
YJOE0012	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0105 and forJOE0110/ revJOE0111 from JOE0077 into *YPRCΔ15* region of SY992	This study
YJOE0013	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0105 and forJOE0110/ revJOE0111 from JOE0078 into *YPRCΔ15* region of SY992	This study
YJOE0014	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0107 and forJOE0110/ revJOE0111 from JOE0066 into *YPRCΔ15* region of SY992	This study
YJOE0015	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0107 and forJOE0110/ revJOE0111 from JOE0067 into *YPRCΔ15* region of SY992	This study
YJOE0016	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0107 and forJOE0110/ revJOE0111 from JOE0068 into *YPRCΔ15* region of SY992	This study
YJOE0017	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0107 and forJOE0110/ revJOE0111 from JOE0069 into *YPRCΔ15* region of SY992	This study
YJOE0018	Homologous integration of PCR fragment with primers revJOE0104/ forJOE0107 and forJOE0110/ revJOE0111 from JOE0070 into *YPRCΔ15* region of SY992	This study

Table 3: Description of plasmids used in this study.

Plasmid	Description	Reference
pUG23	*MET25*$_{prom}$, yEGFP3 (C-fus), *CYC*$_{term}$, *HIS3*, *CEN6/ARSH4*, Ap^r, *ColE1* origin	Hegemann *et al.* 1999
pUG34	*MET25*$_{prom}$, yEGFP3 (N-fus), *CYC*$_{term}$, *HIS3*, *CEN6/ARSH4*, Ap^r, *ColE1* origin	Hegemann *et al.* 1999
pUG35	*MET25*$_{prom}$, yEGFP3 (C-fus), *CYC*$_{term}$, *URA3*, *CEN6/ARSH4*, Ap^r, *ColE1* origin	Hegemann *et al.* 1999
pUG36	*MET25*$_{prom}$, yEGFP3 (N-fus), *CYC*$_{term}$, *URA3*, *CEN6/ARSH4*, Ap^r, *ColE1* origin	Hegemann *et al.* 1999
pks392	tdTomato, kanMX6, Ap^r, *pBR322* origin, T7P	Snaith *et al.* 2005
pks397	P3nmt1, tdTomato, kanMX6, Ap^r, *pBR322* origin, T7P	Snaith *et al.* 2005
pRS316-mCherry	*TDH3*$_{prom}$, mCherry, *URA3*, *CEN6/ARSH4*, *ColE1* origin, Ap^r	Keppler Ross *et al.* 2008
yEpGAP-mCherry	*TDH3*$_{prom}$, mCherry, *URA3*, 2µ Ori, *ColE1* origin, Ap^r	Keppler Ross *et al.* 2008
pJET1.2/blunt	T7P, LacUV5P, eco47IR, *pMB1* origin, Ap^r	Thermo Fisher Scientific, St.LeonRot, Germany
JOE0001	*Sac*I/*Sal*I *KEX2*$_{prom}$ fragment (forJOE0001/ revJOE0002, 590 bp) in *Sac*I/*Sal*I site of pUG35	This study
JOE0002	*Sac*I/*Xba*I *KEX2*$_{prom}$ with *SUC2*$_{secr}$ fragment (forJOE0024/revJOE0027, 647 bp) in *Sac*I/*Xba*I site of pUG35	This study
JOE0003	*Xba*I/*Eco*RI yEGFP3 fragment (forJOE0028/ revJOE0029, 717 bp) in *Xba*I/*Eco*RI site of JOE0002	This study
JOE0004	*Eco*RI/*Nae*I *SAG1* fragment (forJOE0030/ revJOE0031, 1242 bp) in *Eco*RI/*Nae*I site of JOE0003	This study
JOE0005	*Pau*I/*Xba*I *ACT1*$_{prom}$ with *SUC2*$_{secr}$ fragment (forJOE0048/ revJOE0046, 635 bp) in *Pau*I/*Xba*I site of JOE0004	This study
JOE0006	*Pau*I/*Xba*I *ADH1*$_{prom}$ with *SUC2*$_{secr}$ fragment (forJOE0045/ revJOE0043, 546 bp) in *Pau*I/*Xba*I site of JOE0004	This study
JOE0007	*Pau*I/*Xba*I *PGK1*$_{prom}$ with *SUC2*$_{secr}$ fragment (forJOE0042/ revJOE0040, 841 bp) in *Pau*I/*Xba*I site of JOE0004	This study

JOE0008	*Xba*I/*Eco*RI tdTomato fragment (forJOE0038/ revJOE0039, 1431 bp) in *Xba*I/*Eco*RI site of JOE0005	This study
JOE0009	*Xba*I/*Eco*RI tdTomato fragment (forJOE0038/ revJOE0039, 1431 bp) in *Xba*I/*Eco*RI site of JOE0006	This study
JOE0010	*Xba*I/*Eco*RI tdTomato fragment (forJOE0038/ revJOE0039, 1431 bp) in *Xba*I/*Eco*RI site of JOE0007	This study
JOE0011	*Xba*I/*Eco*RI mCherry fragment (forJOE0036/ revJOE0037, 711 bp) in *Xba*I/*Eco*RI site of JOE0005	This study
JOE0012	*Xba*I/*Eco*RI mCherry fragment (forJOE0036/ revJOE0037, 711 bp) in *Xba*I/*Eco*RI site of JOE0006	This study
JOE0013	*Xba*I/*Eco*RI mCherry fragment (forJOE0036/ revJOE0037, 711 bp) in *Xba*I/*Eco*RI site of JOE0007	This study
JOE0014	*Pau*I/*Xba*I *ADH1*$_{prom}$ fragment (forJOE0045/ revJOE0049, 489 bp) in *Pau*I/*Xba*I site of JOE0003	This study
JOE0015	*Pau*I/*Xba*I *ACT1*$_{prom}$ fragment (forJOE0048/ revJOE0051, 578 bp) in *Pau*I/*Xba*I site of JOE0003	This study
JOE0016	*Pau*I/*Xba*I *PGK1*$_{prom}$ fragment (forJOE0042/ revJOE0050, 784, bp) in *Pau*I/*Xba*I site of JOE0003	This study
JOE0017	*Pau*I/*Xba*I *KEX2*$_{prom}$ fragment (forJOE0052/ revJOE0053, 590 bp) in *Pau*I/*Xba*I site of JOE0007	This study
JOE0018	*Eco*RI/*Nae*I *CYC1*$_{term}$ with His(8)tag and KDEL fragment (forJOE0055/revJOE0054, 340 bp) in *Eco*RI/*Nae*I site of JOE0005	This study
JOE0019	*Eco*RI/*Nae*I *CYC1*$_{term}$ with His(8)tag and KDEL fragment (forJOE0055/revJOE0054, 340 bp) in *Eco*RI/*Nae*I site of JOE0006	This study
JOE0020	*Eco*RI/*Nae*I *CYC1*$_{term}$ with His(8)tag and KDEL fragment (forJOE0055/revJOE0054, 340 bp) in *Eco*RI/*Nae*I site of JOE0007	This study
JOE0021	*Eco*RI/*Nae*I *CYC1*$_{term}$ with His(8)tag and KDEL fragment (forJOE0055/revJOE0054, 340 bp) in *Eco*RI/*Nae*I site of JOE0014	This study

JOE0022	*Eco*RI/*Nae*I *CYC1*$_{term}$ with His(8)tag and KDEL fragment (forJOE0055/revJOE0054, 340 bp) in *Eco*RI/*Nae*I site of JOE0015	This study
JOE0023	*Eco*RI/*Nae*I *CYC1*$_{term}$ with His(8)tag and KDEL fragment (forJOE0055/revJOE0054, 340 bp) in *Eco*RI/*Nae*I site of JOE0016	This study
JOE0024	*Eco*RI/*Nae*I *CYC1*$_{term}$ with His(8)tag and SKL fragment (forJOE0056/revJOE0054, 337 bp) in *Eco*RI/*Nae*I site of JOE0014	This study
JOE0025	*Eco*RI/*Nae*I *CYC1*$_{term}$ with His(8)tag and SKL fragment (forJOE0056/revJOE0054, 337 bp) in *Eco*RI/*Nae*I site of JOE0015	This study
JOE0026	*Eco*RI/*Nae*I *CYC1*$_{term}$ with His(8)tag and SKL fragment (forJOE0056/revJOE0054, 337 bp) in *Eco*RI/*Nae*I site of JOE0016	This study
JOE0027	*Xho*I/*Sac*II *SAG1* fragment (forJOE0058/ revJOE0059, 843 bp) in *Xho*I/*Sac*II site of JOE0020	This study
JOE0028	*Pau*I/*Sph*I *PGK1*$_{prom}$ with *SUC2*$_{secr}$ (forJOE0042/ revJOE0068, 841 bp) and *Sph*I/*Eco*RI yEGFP3 fragment (forJOE0028/revJOE0062, 717 bp) in JOE0023	This study
JOE0029	*Pau*I/*Sph*I *PGK1*$_{prom}$ with *SUC2*$_{secr}$ (forJOE0042/ revJOE0068, 841 bp) and *Sph*I/*Eco*RI yEGFP3 fragment (forJOE0028/revJOE0062, 717 bp) in JOE0027	This study
JOE0030	*Pau*I/*Sph*I *PGK1*$_{prom}$ (forJOE0042/revJOE0066, 784 bp) and *Sph*I/*Eco*RI yEGFP fragment (forJOE0028/revJOE0062, 717 bp) in JOE0023	This study
JOE0031	*Pau*I/*Sph*I *PGK1*$_{prom}$ (forJOE0042/revJOE0066, 784 bp) and *Sph*I/*Eco*RI yEGFP fragment (forJOE0028/revJOE0062, 717 bp) in JOE0026	This study
JOE0032	*Pau*I/*Sph*I *ADH1*$_{prom}$ fragment (forJOE0045/ revJOE0064, 489 bp) in *Pau*I/*Sph*I site of JOE0030	This study
JOE0033	*Pau*I/*Sph*I *ADH1*$_{prom}$ fragment (forJOE0045/ revJOE0064, 489 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0034	*Pau*I/*Sph*I *ACT1*$_{prom}$ fragment (forJOE0048/ revJOE0063, 578 bp) in *Pau*I/*Sph*I site of JOE0030	This study

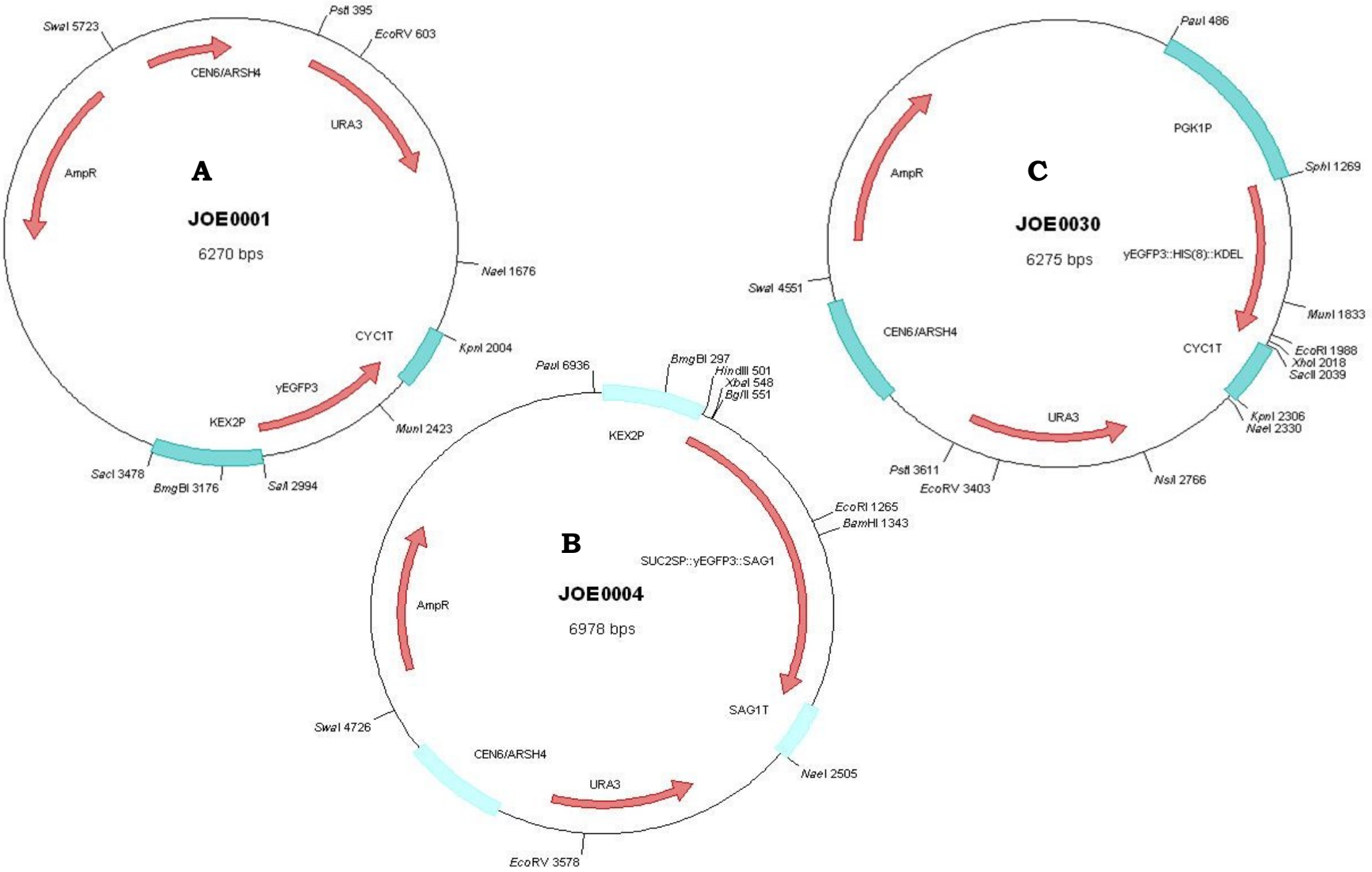
JOE0035	*Pau*I/*Sph*I *ACT1*$_{prom}$ fragment (forJOE0048/ revJOE0063, 578 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0036	*Pau*I/*Sph*I *KEX2*$_{prom}$ fragment (forJOE0052/ revJOE0067, 590 bp) in *Pau*I/*Sph*I site of JOE0030	This study
JOE0037	*Pau*I/*Sph*I *KEX2*$_{prom}$ fragment (forJOE0052/ revJOE0067, 590 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0038	*Sph*I/*Eco*RI mCherry fragment (forJOE0060/ revJOE0037, 711 bp) in *Sph*I/*Eco*RI site of JOE0030	This study
JOE0039	*Sph*I/*Eco*RI mCherry fragment (forJOE0060/ revJOE0037, 711 bp) in *Sph*I/*Eco*RI site of JOE0031	This study
JOE0040	*Pau*I/*Sph*I *TDH3*$_{prom}$ with *SUC2*$_{secr}$ fragment (forJOE0069/ revJOE0070, 763 bp) in *Pau*I/*Sph*I site of JOE0028	This study
JOE0041	*Pau*I/*Sph*I *TDH3*$_{prom}$ with *SUC2*$_{secr}$ fragment (forJOE0069/revJOE0070, 763 bp) in *Pau*I/*Sph*I site of JOE0029	This study
JOE0042	*Pau*I/*Sph*I *PGK1*$_{prom}$ with virus secretion signal fragment (forJOE0042/revJOE0071, 883 bp) in *Pau*I/*Sph*I site of JOE0028	This study
JOE0043	*Pau*I/*Sph*I *PGK1*$_{prom}$ with virus secretion signal fragment (forJOE0042/revJOE0071, 883 bp) in *Pau*I/*Sph*I site of JOE0029	This study
JOE0044	*Sac*I/*Sph*I *KEX2*$_{prom}$ (forJOE0085/revJOE0086, 590 bp) and *Sph*I/*Xho*I CI fragment (forJOE0072/ revJOE0073, 725 bp) in *Sac*I/*Xho*I site of pUG24	This study
JOE0045	*Sph*I/*Xho*I CI with RecA mutation fragment (forJOE0072/revJOE0079//forJOE0078/ revJOE0073, 725 bp) in *Sph*I/*Xho*I site of JOE0044	This study
JOE0046	*Sph*I/*Xho*I CI with RecA mutation and internal NLS fragment (forJOE0072/revJOE0080// forJOE0078/revJOE0073) in *Sph*I/*Xho*I site of JOE0044	This study
JOE0047	*Sph*I/*Xho*I CI with RecA mutation and N-terminal NLS fragment (forJOE0074/ revJOE0079//forJOE0078/ revJOE0073) in *Sph*I/*Xho*I site of JOE0044	This study

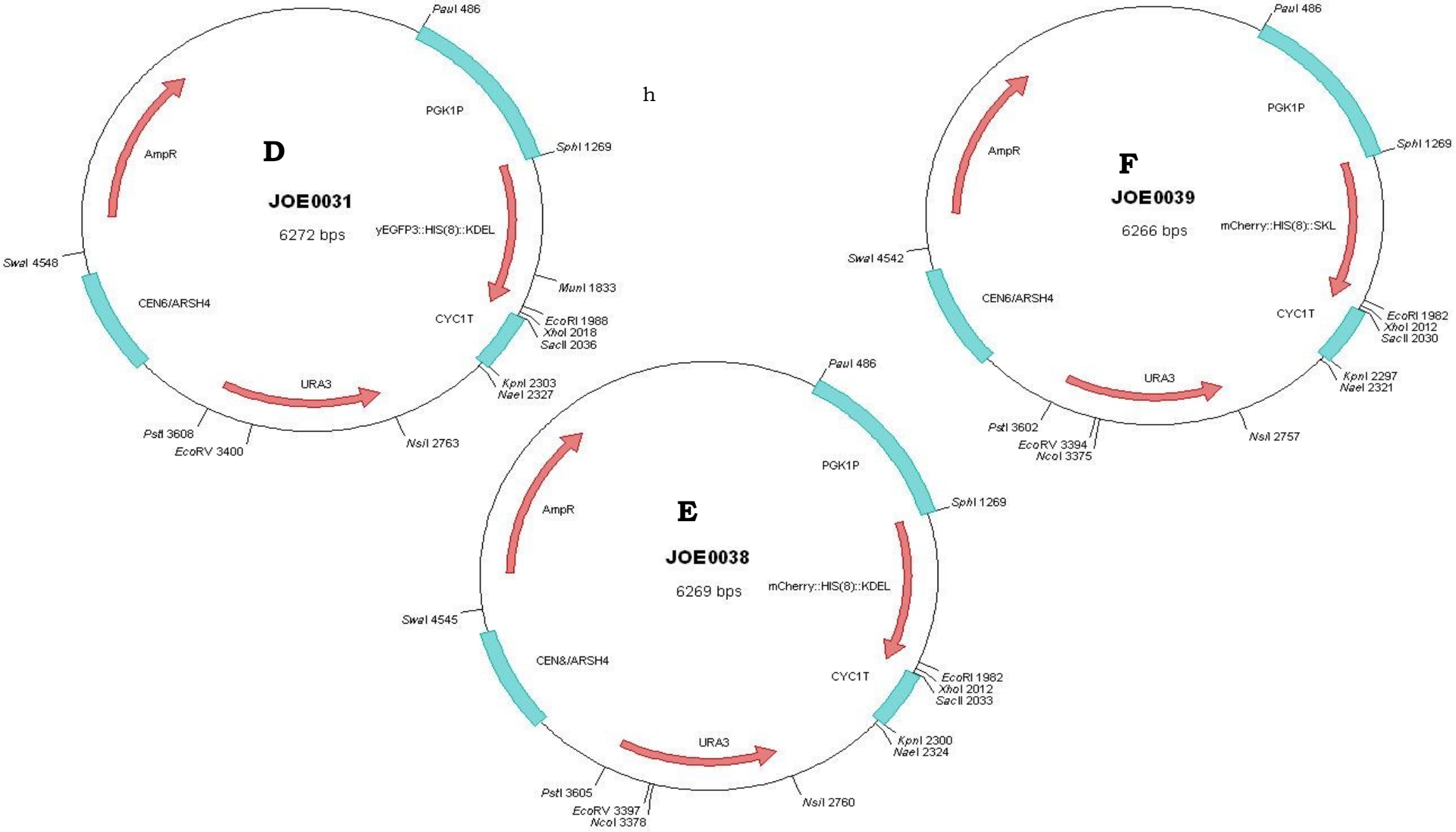
JOE0048	*Sph*I/*Xho*I CI with RecA mutation and C-terminal NLS fragment (forJOE0072/ revJOE0079//forJOE0078/ revJOE0075) in *Sph*I/*Xho*I site of JOE0044	This study
JOE0049	*Sph*I/*Xho*I CI with RecA mutation, internal NLS and N-terminal His(8)tag fragment (forJOE0076/revJOE0080//forJOE0078/ revJOE0073, 749 bp) in *Sph*I/*Xho*I site of JOE0044	This study
JOE0050	*Sph*I/*Xho*I CI with RecA mutation and C-terminal His(8)tag fragment (forJOE0072/ revJOE0079//forJOE0078/ revJOE0077, 749 bp) in *Sph*I/*Xho*I site of JOE0044	This study
JOE0051	*Sac*I/*Sph*I HDA1$_{prom}$ fragment (forJOE0081/ revJOE0082, 437 bp) in *Sac*I/*Sph*I site of JOE0049	This study
JOE0052	*Sac*I/*Sph*I HDA1$_{prom}$ fragment (forJOE0081/ revJOE0082, 437 bp) in *Sac*I/*Sph*I site of JOE0050	This study
JOE0053	*Sac*I/*Sph*I TRL1$_{prom}$ fragment (forJOE0083/ revJOE0084, 296 bp) in *Sac*I/*Sph*I site of JOE0049	This study
JOE0054	*Sac*I/*Sph*I TRL1$_{prom}$ fragment (forJOE0083/ revJOE0084, 296 bp) in *Sac*I/*Sph*I site of JOE0050	This study
JOE0056	*Sac*I/*Sph*I HDA1$_{prom}$ fragment (forJOE0081/ revJOE0082, 437 bp) in *Sac*I/*Sph*I site of JOE0046	This study
JOE0057	*Sac*I/*Sph*I TRL1$_{prom}$ fragment (forJOE0083/ revJOE0084, 296 bp) in *Sac*I/*Sph*I site of JOE0046	This study
JOE0058	*Pau*I/*Sph*I *PGK1*$_{prom}$ with Pos1 fragment (revJOE0072/ revJOE0088//forJOE0087/ revJOE0004, 868 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0059	*Pau*I/*Sph*I *PGK1*$_{prom}$ with Pos2 fragment (revJOE0009/ revJOE0090//forJOE0089/ revJOE0004, 868 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0060	*Pau*I/*Sph*I *PGK1*$_{prom}$ with n Pos1+ fragment (revJOE0009/ revJOE0093//forJOE0094/ revJOE0004, 829 bp) in *Pau*I/*Sph*I site of JOE0031	This study

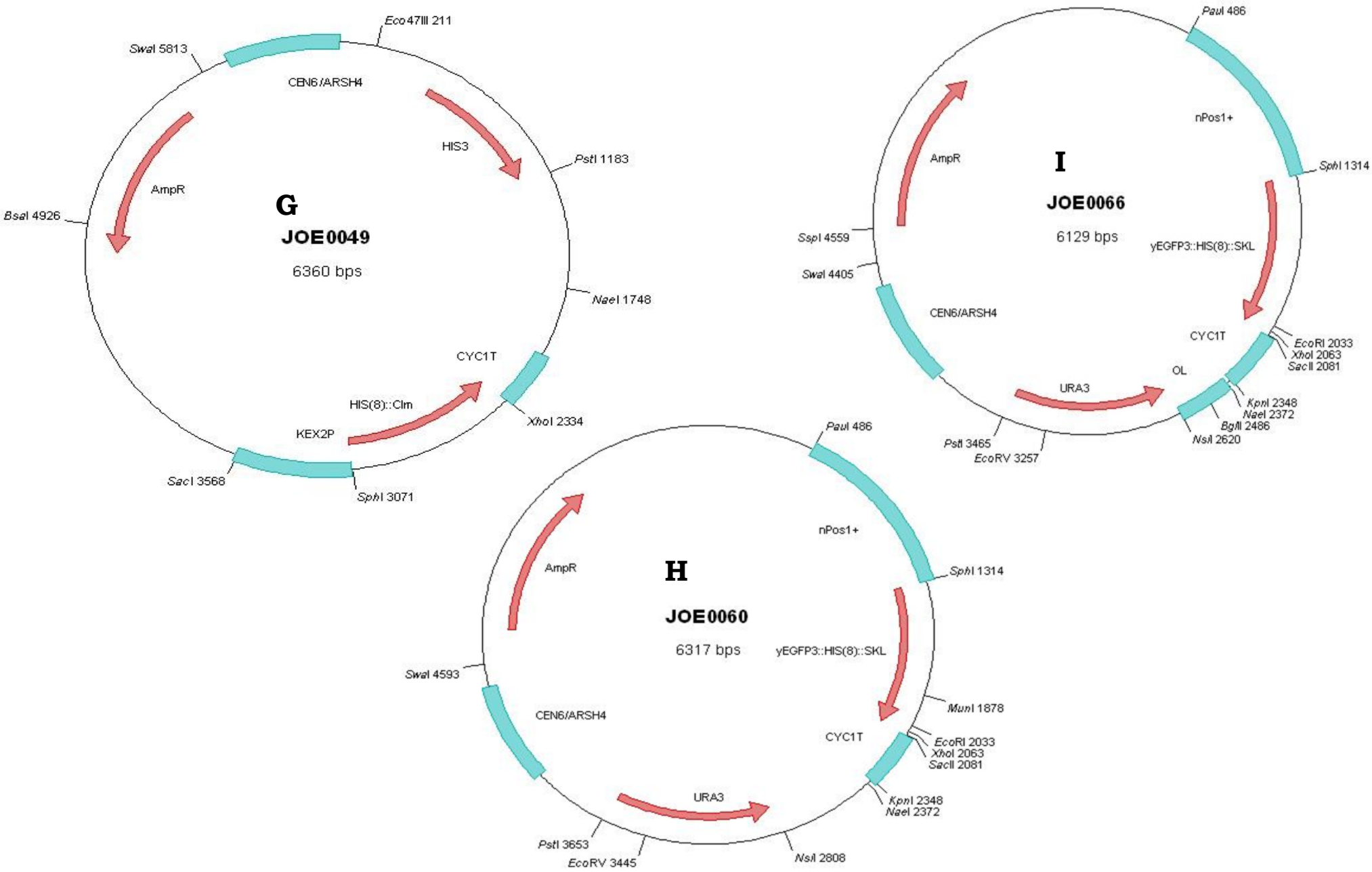
JOE0061	*Pau*I/*Sph*I $PGK1_{prom}$ with n Pos1- fragment (revJOE0009/ revJOE0093//forJOE0095/ revJOE0004, 784 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0062	*Pau*I/*Sph*I $PGK1_{prom}$ with n Pos2+ fragment (revJOE0009/ revJOE0096//forJOE0097/ revJOE0004, 829 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0063	*Pau*I/*Sph*I $PGK1_{prom}$ with n Pos2- fragment (revJOE0009/ revJOE0096//forJOE0098/ revJOE0004, 784 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0064	*Pau*I/*Sph*I $PGK1_{prom}$ with n Pos3+ fragment (revJOE0009/ revJOE0099//forJOE0100/ revJOE0004, 829 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0065	*Pau*I/*Sph*I $PGK1_{prom}$ with n Pos3- fragment (revJOE0009/ revJOE0099//forJOE0101/ revJOE0004, 784 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0066	*Nae*I/*Nsi*I O_L fragment (forJOE0091/ revJOE0092, 246 bp) in *Nae*I/*Nsi*I site of JOE0060	This study
JOE0067	*Nae*I/*Nsi*I O_L fragment (forJOE0091/ revJOE0092, 246 bp) in *Nae*I/*Nsi*I site of JOE0062	This study
JOE0068	*Nae*I/*Nsi*I O_L fragment (forJOE0091/ revJOE0092, 246 bp) in *Nae*I/*Nsi*I site of JOE0063	This study
JOE0069	*Nae*I/*Nsi*I O_L fragment (forJOE0091/ revJOE0092, 246 bp) in *Nae*I/*Nsi*I site of JOE0064	This study
JOE0070	*Nae*I/*Nsi*I O_L fragment (forJOE0091/ revJOE0092, 246 bp) in *Nae*I/*Nsi*I site of JOE0065	This study
JOE0071	*Pau*I/*Sph*I $PGK1_{prom}$ short fragment (forJOE0102/ revJOE0066, 186 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0072	*Sac*I/*Sph*I $TRL1_{prom}$ long fragment (forJOE0103/ revJOE0084, 706 bp) in *Sac*I/*Sph*I site of JOE0046	This study

JOE0073	*Pau*I/*Sph*I *PGK1*$_{prom}$ with n Pos1+ and n Pos2+ fragment (revJOE0009/ revJOE0093//forJOE0094/ revJOE0004 from JOE0062, 874 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0074	*Pau*I/*Sph*I *PGK1*$_{prom}$ with n Pos1+ and n Pos2- fragment (revJOE0009/ revJOE0093//forJOE0094/ revJOE0004 from JOE0063, 829 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0075	*Pau*I/*Sph*I *PGK1*$_{prom}$ with n Pos1+ with n Pos 3+ fragment (revJOE0009/ revJOE0093//forJOE0094/ revJOE0004 from JOE0064, 874 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0076	*Pau*I/*Sph*I *PGK1*$_{prom}$ with n Pos1+ and n Pos3- fragment (revJOE0009/ revJOE0093//forJOE0094/ revJOE0004 from JOE0065, 829 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0077	*Pau*I/*Sph*I *PGK1*$_{prom}$ with TAT fragment (revJOE0009/ revJOE0112//forJOE0113/ revJOE0004, 809 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0078	*Pau*I/*Sph*I *PGK1*$_{prom}$ with n Pos3- and TAT fragment (revJOE0009/ revJOE0112//forJOE0113/ revJOE0004 from JOE0065, 809 bp) in *Pau*I/*Sph*I site of JOE0031	This study
JOE0079	BsaI/*Sph*I fragment (2136 bp) of JOE0075 in BsaI/*Sph*I site of JOE0044	This study
JOE0080	*Sac*I/*Sph*I *HDA1*$_{prom}$ fragment (forJOE0081/ revJOE0082, 437 bp) in *Sac*I/*Sph*I site of JOE0044	This study
JOE0081	Substitution of *CEN6*/*ARSH4* sequence with 2µm fragment (forJOE0118/revJOE0117, 1429 bp) in JOE0044	This study
JOE0082	Substitution of *CEN6*/*ARSH4* sequence with 2µm fragment (forJOE0118/revJOE0117, 1429 bp) in JOE0079	This study
JOE0083	Substitution of *CEN6*/*ARSH4* sequence with 2µm fragment (forJOE0118/revJOE0117, 1429 bp) in JOE0080	This study
JOE0084	Substitution of *CEN6*/*ARSH4* sequence with 2µm fragment (forJOE0118/revJOE0117, 1429 bp) in JOE0046	This study

JOE0085	Substitution of *CEN6/ARSH4* sequence with 2µm fragment (forJOE0118/revJOE0117, 1429 bp) in JOE0056	This study
JOE0086	*Pau*I/*Sph*I $PGK1_{prom}$ with n Pos 2+ with n Pos 3- fragment (revJOE0009/ revJOE0099// forJOE0101/ revJOE0004 from JOE0064, 829 bp) in *Pau*I/*Sph*I site of JOE0062	This study
JOE0087	*Pau*I/*Sph*I $PGK1_{prom}$ with n Pos 2+ with n Pos 3+ fragment (revJOE0009/ revJOE0099// forJOE0100/ revJOE0004 from JOE0064, 874 bp) in *Pau*I/*Sph*I site of JOE0062	This study
JOE0088	*Pau*I/*Sph*I $PGK1_{prom}$ with n Pos 2- with n Pos 3- fragment (revJOE0009/ revJOE0099// forJOE0101/ revJOE0004 from JOE0064, 784 bp) in *Pau*I/*Sph*I site of JOE0063	This study
JOE0089	BsaI/*Sph*I fragment of JOE0031 (2100 bp) in BsaI/*Sph*I site of JOE0046	This study
JOE0090	Substitution of *CEN6/ARSH4* sequence with 2µm fragment (forJOE0118/revJOE0117, 1429 bp) in JOE0089	This study
JOE0091	BsaI/*Sph*I fragment of JOE0031(2100 bp) in BsaI/*Sph*I site of JOE0044	This study
JOE0092	Substitution of *CEN6/ARSH4* sequence with 2µm fragment (forJOE0118/revJOE0117, 1429 bp) in JOE0091	This study
JOE0093	Substitution of *CEN6/ARSH4* sequence with 2µm fragment (forJOE0118/revJOE0117, 1429 bp) in JOE0050	This study
JOE0094	Substitution of *CEN6/ARSH4* sequence with 2µm fragment (forJOE0118/revJOE0117, 1429 bp) in JOE0049	This study
JOE0095	Substitution of *CEN6/ARSH4* sequence with 2µm fragment (forJOE0118/revJOE0117, 1429 bp) in JOE0052	This study
JOE0096	Substitution of *CEN6/ARSH4* sequence with 2µm fragment (forJOE0118/revJOE0117, 1429 bp) in JOE0051	This study


A
JOE0001
6270 bps
SwaI 5723
PstI 395
EcoRV 603
CEN6/ARSH4
URA3
AmpR
NaeI 1676
CYC1T
KpnI 2004
yEGFP3
KEX2P
MunI 2423
SacI 3478
BmgBI 3176
SalI 2994
B
JOE0004
6978 bps
SUC2SP::yEGFP3::SAG1
PauI 6936
BmgBI 297
HindIII 501
XbaI 548
BglII 551
KEX2P
EcoRI 1265
BamHI 1343
AmpR
SAG1T
SwaI 4726
CEN6/ARSH4
NaeI 2505
URA3
EcoRV 3578
C
JOE0030
6275 bps
yEGFP3::HIS(8)::KDEL
PauI 486
PGK1P
SphI 1269
AmpR
SwaI 4551
MunI 1833
CEN6/ARSH4
CYC1T
EcoRI 1988
XhoI 2018
SacII 2039
KpnI 2306
NaeI 2330
URA3
PstI 3611
EcoRV 3403
NsiI 2766


D
JOE0031
6272 bps
PGK1P
yEGFP3::HIS(8)::KDEL
CYC1T
URA3
CEN6/ARSH4
AmpR
PauI 486
SphI 1269
MunI 1833
EcoRI 1988
XhoI 2018
SacII 2036
KpnI 2303
NaeI 2327
NsiI 2763
EcoRV 3400
PstI 3608
SwaI 4548
E
JOE0038
6269 bps
PGK1P
mCherry::HIS(8)::KDEL
CYC1T
URA3
CEN&/ARSH4
AmpR
PauI 486
SphI 1269
EcoRI 1982
XhoI 2012
SacII 2033
KpnI 2300
NaeI 2324
NsiI 2760
PstI 3605
EcoRV 3397
NcoI 3378
SwaI 4545
F
JOE0039
6266 bps
PGK1P
mCherry::HIS(8)::SKL
CYC1T
URA3
CEN6/ARSH4
AmpR
PauI 486
SphI 1269
EcoRI 1982
XhoI 2012
SacII 2030
KpnI 2297
NaeI 2321
NsiI 2757
PstI 3602
EcoRV 3394
NcoI 3375
SwaI 4542


G
JOE0049
6360 bps
Eco47III 211
SwaI 5813
CEN6/ARSH4
HIS3
PstI 1183
AmpR
BsaI 4926
NaeI 1748
CYC1T
HIS(8)::Clm
KEX2P
XhoI 2334
SacI 3568
SphI 3071
I
JOE0066
6129 bps
PauI 486
nPos1+
AmpR
SphI 1314
SspI 4559
yEGFP3::HIS(8)::SKL
SwaI 4405
CEN6/ARSH4
CYC1T
EcoRI 2033
XhoI 2063
SacII 2081
OL
URA3
KpnI 2348
NaeI 2372
BglII 2486
NsiI 2620
PstI 3465
EcoRV 3257
H
JOE0060
6317 bps
PauI 486
nPos1+
AmpR
SphI 1314
yEGFP3::HIS(8)::SKL
SwaI 4593
MunI 1878
CEN6/ARSH4
CYC1T
EcoRI 2033
XhoI 2063
SacII 2081
URA3
KpnI 2348
NaeI 2372
PstI 3653
EcoRV 3445
NsiI 2808

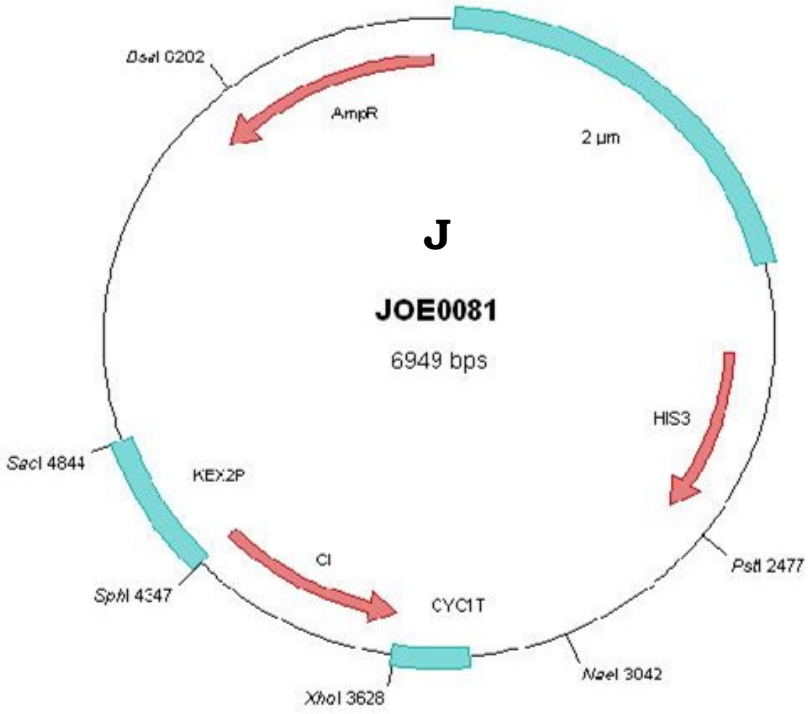


Figure 2: **Plasmid maps and structures**. Schematic illustration of features (red arrows = ORF, blue marks = specific DNA sequence with T=terminator, P=promoter, and replication *ori*), size in bp and unique restriction sites of the plasmids JOE0001 (A), JOE0004 (B), JOE0030 (C), JOE0031 (D), JOE0038 (E), JOE0039 (F), JOE0049 (G), JOE0060 (H), JOE0066 (I) and JOE0081 (J).

2.4 CULTURE MEDIA AND CULTIVATION

2.4.1 MEDIA

All culture media were autoclaved at 121°C for 20 minutes at 2 bar. Non-autoclavable stock solutions were sterile filtered using filters with pore size 0.2 µm (Schleicher and Schuell, Dassel, Germany). For solid media 1.5% (w/v) agar was added.

LB medium (Sambrook and Russell 2001)

tryptone	10 g/L
NaCL	10 g/L
Yeast extract	5 g/L

YPD medium (Sherman 2002)

tryptone	20 g/L
Yeast extract	10 g/L
Glucose	2% (w/v)

Occasionally 50 mg/L-uracil and 50 mg/L-adenine sulfate salt was added.

SD medium (Wickerham 1946)

D+ Biotin	2 ng/L	Folic acid	2 ng/L
D-Ca pantothenat	400 ng/L	Inositol	2 mg/L
Nicotinic acid	400 ng/L	Na_2MoO_4 x 2 H_2O	20 ng/L
p-Aminobenzoic acid	200 ng/L	$ZnSO_4$ x 7 H_2O	40 ng/L
Phyridoxine-HCl	400 ng/L	KH_2PO_4	1 g /L
Riboflavine	200 ng/L	$MgSO_4$ x 7 H_2O	500 mg /L
Thiamine-HCl	400 ng/L	NaCl	100 mg/L
H_3BO_3	50 ng /L	$CaCl_2$ x 2 H_2O	100 mg/L
$CuSO_4$ x 5 H_2O	4 ng/L	$(NH_4)_2SO_4$	5 g/L
KJ	10 ng/L	$MnSO_4$ x H_2O	40 ng/L
$FeCl_3$ x 6 H_2O	20 ng/L	Glucose	2% (w/v)

Corresponding to the needs of the auxotrophic strains, the following supplements were added:

Adenine sulphate	20 mg/L
Uracil	20 mg/L
L-histidine-HCl	20 mg/L
L-tryptophan	20 mg/L
L-methionine	20 mg/L
L-lysine-HCl	30 mg/L

2.4.2 CULTIVATION

The cultivation of *Escherichia coli* was carried out aerobically at 37°C with LB medium and liquid cultures were grown with constant shaking at 170 – 180 rpm (Minitron, Infors HT, Bottmingen, Switzerland) until the required optical density was reached.

Saccharomyces cerevisiae strains were grown aerobically at 28-30°C with SD, YPD or YPDAU-medium, whereby liquid cultures were processed in the same way as described for *Escherichia coli* respectively.

Cell number and growth was controlled by measurement of the optical density (OD) at a wavelength of 600 nm with the SmartSpec PLUS photometer (BIO-RAD, Munich, Germany). The OD_{600} of 0.1 was approximated to be $3*10^6$ (*S. cerevisiae*) or $1*10^8$ cells/ml (*E.coli*) based on www.bionumbers.hms.harvard.edu/default.aspx.

The conservation and storage of strains at -80°C was achieved by re-suspending a pellet of a 2 ml overnight culture in 900 µl of fresh YPDAU for yeast and LB for *E. coli* strains. In each case, 600 ml of 86% glycerol were added to each tube.

2.5 DNA METHODS

2.5.1 ISOLATION OF E. COLI PLASMID DNA

The isolation of plasmid DNA was achieved by following the widely used protocol of Sambrook and Russell (2001) harvesting a 1.5 ml overnight culture from *E. coli* and re-suspending it in 100 µl TE buffer (10 mM Tris-base, pH 8.0 adjusted with HCl, 1 mM EDTA). Through the addition of 200 µl 1% (w/v) SDS /200 mM NaOH and 1.5 min incubation at room temperature (RT) cell lysis was accomplished, which was followed by the addition of 150 µl of 3 M potassium acetate (pH 5.2). Subsequently, cell debris, chromosomal DNA and proteins were removed by centrifugation (16,100 x g, 12 min at RT) and the

same volume of isopropanol was added to the achieved volume of aqueous supernatant with the plasmid DNA for precipitation. The DNA then was pelleted by centrifugation and afterwards salts were removed with 70% (v/v) ethanol. The plasmid DNA was dissolved in TE buffer (pH 8) and stored at -20°C until further processing.

2.5.2 Isolation of Saccharomyces cerevisiae chromosomal DNA

Chromosomal DNA of *Saccharomyces cerevisiae* was isolated following the protocol by Amberg *et al.* (2006) starting with a 6 ml YPDAU overnight culture, which total cell amount was collected by centrifugation into a 1.5 ml microfuge tube. The pellet was re-suspended in 200 µl breaking buffer (2% (v/v) Triton X-100, 1% (w/v) SDS, 0.1 M NaCl, 10 mM Tris-base, pH 8 adjusted with HCl, 1 mM EDTA) and 200 µl phenol/chloroform/isoamyl alcohol (25:24:1) and subsequently 0.3 g glass beads (0.1 mm diameter, acid washed, neutralized and autoclaved) were added. After vortexting the mixture for 6 min and addition of 200 µl TE buffer (pH 8), the phases were separated by centrifugation for 10 min at RT. The supernatant was transferred to a new tube and the same amount of ice cold isopropanol was added. Through incubation for 20 min at -20°C and centrifugation at 4°C, a DNA pellet was produced, which then was re- suspended in 200 µl TE buffer (pH 8) with 10 µg/ml RNase A and incubated for 15 min at 37°C. For another precipation step of DNA 22.5 µl 3 M potassium acetate (pH 5.2) and 230 µl isopropanol were added and the mixture was incubated at -20°C for 20 min. After centrifugation at 4°C salts were removed from the pellet with the help of 70% (v/v) ethanol. The chromosomal DNA was dissolved in TE buffer (pH 8) at 4°C overnight and then stored at -20°C until further processing.

2.5.3 Estimation of DNA concentration and purity

The concentration and purity of isolated DNA was determined by measuring its absorption at 260 nm and 280 nm with a SmartSpec PLUS photometer (BIO-RAD, Munich, Germany). The quotient of the wavelengths served as an indicator for the purity and the DNA concentration was calculated using the

empirical formula for double stranded DNA following for both Sambrook and Russell (2001):

$$\text{DNA in } \mu g/ml = \_ E_{260} \times 50\mu g/ml \times \text{dilution factor}$$

2.5.4 Cleavage of DNA with restriction endonucleases

The enzymes for restriction digestion of PCR fragments, chromosomal or plasmid DNA were purchased from Thermo Fisher Scientific (St. Leon-Rot, Germany) as well as New England Biolabs (NEB, Frankfurt, Germany) and the reactions were performed following the manufacturer's instructions for 3 to 16 hours with 3 - 6 µg DNA.

2.5.5 Agarose gel electrophoresis for DNA separation

Separation of DNA molecules by size was achieved through 0.8–1.2% (w/v) agarose gels in TAE buffer (40 mM Tris-base, pH 8.0 adjusted with acetate, 2 mM EDTA) by applying 3–5 V/cm. Prior to electrophoresis, the samples were mixed with 1/10 loading dye (20% (v/v) Ficoll 400, 0.1% (w/v) bromophenol blue, 0.1% (w/v) xylene cyanol, 1 mg/ml RNase A) and incubated for 15 min at 37°C. As size markers, the DNA ladders 2 Log, 1kpb and 100 bp from NEB (Frankfurt, Germany) were routinely applied. After separation, an ethidium bromide solution (0.5 µg/ml) was used for staining the nucleic acids inside the gel and the bands were detected under UV-light at 312 nm with the DC120 Zoom Digital Camera (KODAK, USA).

RNase A solution

10 mM Tris-base

15 mM NaCl

10% (w/v) RNase A from NEB (Frankfurt, Germany)

pH 7.5 adjusted with HCl

2.5.6 POLYMERASE CHAIN REACTION METHODS

For amplification of specific DNA fragments, different variants of the polymerase chain reaction (PCR) were used within the T3 Thermocycler from Biometra (Goettingen Germany). Generally, all reactions were carried out with the OneTaq 2X Master Mix (20 mM Tris-base, 22 mM KCl, 22 mM NH_4Cl, 1.8 mM $MgCl_2$, 5% (v/v) Glycerol, 0.05% (v/v) Tween 20, 0.06% (v/v) IGEPAL CA-630, 0.2 mM dNTPs, 25 units/ml OneTaq DNA Polymerase, pH 8.9 adjusted with HCl) from NEB (Frankfurt, Germany). Additionally it contained 100-300 ng of template DNA and 50 pmole of each primer (forward and reverse), following the recommended reaction setup for 25 or 50 µl volumes, as well as of the PCR program.

Therefore, routinely a 2 min denaturation step at 94°C was followed by 33 cycles, whereupon each had a 30 s denaturation step at 94°C, a 30 s annealing step at primer depending temperature, and a 1 min/kpb extension step at 68°C. After a final elongation of 10 min at 68°C, the reaction was cooled down to 4°C and then stored at -20°C. The virtual testing, primer design, as well as the specific annealing temperature calculation, was performed with the computer program Clone Manager 9 Professional.

All PCR products were verified by agarose gel electrophoresis and specific differences of the deployed methods are described in the following.

2.5.6.1 Generation of fragments for cloning

DNA fragments generated by PCR for cloning into plasmids were normally purified from agarose gels (see 2.5.7) after digestion with restriction endonucleases, or in the case of low product concentration directly cloned into the pJET1.2/blunt vector with the help of the clone JET PCR cloning set (Thermo Fisher Scientific, St. Leon-Rot, Germany), following the manufacturer's instructions.

2.5.6.2 Overlap Extension PCR

The method of overlap extension PCR by Ho *et al.* (1989) allows the insertion of specific DNA modifications into a defined DNA fragment by a two fragment interim stage, as well as the ligation of two DNA fragments without a ligase. For each fragment, a specific primer pair is designed, whereby the corresponding forward primers carry the possible desired modification and one of them always a 25 bp sequence homologous to the other fragment respectively primer. After generating, the two fragments are diluted 100 fold and mixed in a new reaction only provided with the two reverse primers, whereupon in the first 5 cycles the homologous region serves as primer. In the following 35 cycles amplification of the fusion fragment was achieved by the two reverse primers and so, the annealing temperature was set up from 50°C to the primer specific one.

2.5.6.3 Omega PCR

Substitutions of sequences in plasmids was accomplished with the help of the Omega PCR method, following the two-step approach protocol of Chen *et al.* (2013), whereupon the created fragments had a 46 bp homologous region on each site.

2.5.6.4 Preparation of fragments for homologous recombination

For the homologous recombination of constructs into specific locations of the *Saccharomyces cerevisiae* genome, DNA fragments were modified by PCR subsequent to the protocol of Oldenburg *et al.* (1997). For further specificity of the recombination an additional PCR was carried out, increasing so the homologous region from 52 to 96 bp on each site of the fragment.

2.5.6.5 Control of insertion

The PCR was also used to control the homologous recombination events. Hence, two specific primers were designed, whereupon the forward primer binds inside the inserted construct and the reverse primer outside of it in the

specific genome region, where the constructed fragment should have been inserted. Therefore, a PCR fragment with specific size was detected only in an agarose gel, if the construct was correctly inserted into the genome.

2.5.7 DNA GEL EXTRACTION

With the help of the GeneJET Gel Extraction and DNA Cleanup Micro Kit from Thermo Fisher Scientific (St. Leon-Rot, Germany), DNA fragments digested with restriction endonucleases as well as PCR fragments were extracted from agarose gels following the manufacturer's instructions. The gels previously were stained with GelRed nucleic Acid Stain (Biotium, Hayward, USA).

2.5.8 LIGATION

For the ligation of DNA fragments, T4 ligases from Thermo Fisher Scientific (St. Leon-Rot, Germany) as well as NEB (Frankfurt, Germany) were used and the reactions were performed in 20 µl volume according to the corresponding manufacturer's directions for two hours at RT or overnight at 16°C. Prior to ligation, digested plasmid DNA was treated with the Antarctic phosphatase (NEB, Frankfurt, Germany) to minimize self-ligation, following the manufactures protocol.

2.5.9 SEQUENCING

The partial sequencing of plasmid DNA at the company GATC (Cologne, Germany) required the isolation of plasmid DNA with the help of the high pure plasmid isolation kit from Roche (Mannheim, Germany), according to manufacturer's instructions. After concentration estimation via absorbance measurement, the DNA was diluted to the recommended amount of GATC and sent there for sequencing.

2.6 TRANSFORMATION METHODS

2.6.1 ESCHERICHIA COLI

The method, described in Promega Protocols and Applications Guide (3rd edition) pp. 45-46, was applied to generate competent *E. coli* DH5aF', which were stored at -80°C afterwards. For transformation, the thawed cells were mixed with respective plasmid DNA and incubated for 30 min on ice. Following a heat shock at 42°C for 1.5 min, the cells were mixed after 3 min incubation on ice with 1 ml LB medium and incubated for 60 min at 37°C with constant shaking. Finally, the concentrated cells - in around 100 µl volume - were plated on LB agar containing appropriate antibiotics.

2.6.2 SACCHAROMYCES CEREVISIAE

Transformation of *Saccharomyces cerevisiae* was started with a cell pellet from 0.5 ml of an overnight YPDAU culture to which 10 µl of carrier DNA (denatured salmon sperm DNA, 10 mg/ml) and 1 µg transforming DNA in 10 µl was added and mixed. After incubation at RT for 2 min, the suspension was mixed with 0.5 ml PLATE solution (40% (v/v) PEG 4000, 0.1M LiAc, 10 mM Tris-base, pH 7.5 adjusted with HCl, 1 mM EDTA) and kept for 15 min at RT. Following a heat shock for 15 min at 42°C, the cells were pelleted and 60 µl H_2O was added to spread them onto selective SD plates.

2.7 PLASMID STABILITY TEST AND CONTROL

For the control of plasmid stability, a repeated dilution of appropriate *Saccharomyces* strain cultures in non-selective SD supplemented with His, Leu, Trp, Ade, and Lys or non-selective YPDAU medium was performed by transferring 1 ml of the respective culture to 50 ml of fresh media after 24 hours. Before that, a start culture was incubated in selective SD medium supplemented with Leu, Trp, Ade, and Lys as well as the likewise repeatedly

diluted control of each strain. At timely specific control points, the OD_{600} as well as the fluorescence (see 2.9.2 or 2.9.3) of each culture was measured, whereupon the first dilution of the start culture in a non-selective medium was seen as startpoint respectively t=0 for the experiment. Additionally, at every control point an aliquot of cells from each culture was diluted appropriately for spreading approximately 100-200 cells on YPDAU agar plates. After incubation, the colonies were replica plated on SD without uracil and/or histidine as well as afterwards again on YPDAU to compare the ratio of plasmid bearing and plasmid-free cells after incubation. The colony number on the last YPDAU plate defines thereby the maximal amount of plasmid bearing cells respectively 100%.

2.8 PROTEIN METHODS

2.8.1 ISOLATION OF SACCHAROMYCES CEREVISIAE PROTEINS

Overnight stationary cultures in selective SD medium of 5 ml volume were concentrated as well as adjusted via OD_{600} measurement into single pellets with similar cell numbers ($1*10^7$ cells) and each re-suspended in 300 µl H_2O. Subsequently 300 µl of 0.4 M NaOH was added and the mixture incubated at RT for 5 min. After a centrifugation step, the cell pellet was re-suspended in 300 µl extractions buffer (62 mM Tris–base, 5% (v/v) β–Mercaptoethanol, 10% (v/v) glycerol (87%), 2% (w/v) SDS, pH 6.8 adjusted with HCl) mixed with the recommended amount of “Inhibitor cocktail His tag” by Roth (Karlsruhe, Germany) and immediately boiled for 10 min at 95°C. Following a centrifugation, the supernatant was transferred into a new cup and stored at -20°C until further usage.

2.8.2 POLYACRYLAMIDE GEL ELECTROPHORESES

The separation of proteins by size was achieved by a discontinuous sodium dodecyl sulfate polyacrylamide gel electrophoreses (SDS-PAGE). Therefore, the Mini-PROTEAN II dual slab cell from Bio-Rad (Munich, Germany) was applied.

After preparation of the resolving gel (375 mM Tris–base, 12% (v/v) acrylamide/bisacrylamide (30:0.8), 0.025% (v/v) APS, 0.1% (v/v) TEMED, 0.1% (w/v) SDS, pH 8.8 adjusted with HCl) and of the stacking gel (125 mM Tris–base, 5% (v/v) acrylamide/bisacrylamide (30:0.8), 0.025% (w/v) ammonium persulfate, 0.1% (v/v) TEMED, 0.1% (w/v) SDS, pH 6.8 adjusted with HCl), the samples were loaded. They had previously been mixed and heated at 95°C in 1/4 of loading dye (60% (v/v) glycerol (87%), 9.2% (w/v) SDS, 0.25 M Tris–base, 0.012% (w/v) bromphenolblue, 5% (v/v) β–mercaptoethanol, pH 6.8 adjusted with HCl). Additionally the chamber was filled with 1 x TGS–running buffer (25 mM Tris–base, 192 mM glycine, 0.1% (w/v) SDS, pH 8.4 adjusted with HCl) and the size markers (Roti mark 10-150 PLUS protein marker or Roti mark standard protein marker) from Roth (Karlsruhe, Germany) were added. The system was run with 180V until the desired separation was achieved and the gels were subsequently prepared for western blot analysis or Coomassie staining.

2.8.3 COOMASSIE STAINING

The proteins in the gels were stained with conventional Coomassie Brilliant Blue R 250 (45% (v/v) ethanol, 10% (v/v) acetic acid, 0.25% (w/v) Coomassie Brilliant Blue R-250) for 3 h at RT and afterwards destained in fast (30% (v/v) ethanol, 10% (v/v) acetic acid) or slow (5% (v/v) ethanol, 7% (v/v) acetic acid) destaining solution, until the background was colorless or overnight, respectively. In general, gels were stored in storage solution (3% (v/v) glycerol, 18% (v/v) ethanol) at 4°C.

2.8.4 WESTERN BLOT ANALYSIS

Before proteins could be detected on a membrane (Hybond-ECL GE Healthcare (0,45 µm 20x3) Amersham, Munich, Germany) with the help of specific antibodies (Penta HIS HRP conjugate Quiagen, Hilden, Germany), the proteins were transferred onto it via the Mini Trans-Blot module combined with the standard Trans-Blot cell from Bio Rad (Munich, Germany). Therefore, transfer buffer (20% (v/v) methanol, 25 mM Tris–base, 192 mM glycine, 0.075% (w/v) SDS) was used and the protein transfer was performed with 80 V for 15 min.

Afterwards the membrane was processed following the QIAexpress detection and assay handbook 3th edition 2002 protocol No.7. Now the chemiluminescence reaction was activated by dispersing super signal west pico chemilumiscents substrate on the membrane, according to the protocol of Thermo Fisher Scientific (Ulm, Germany). The signal then was detected with the help of the imaging system Chemi Doc and Quantity One Quantitation software from Bio Rad (Munich, Germany).

2.9 FLUORESCENCE METHODS

2.9.1 FLUORESCENCE MICROSCOPY

The two microscopes Axiostars plus FL (filterset 10, excitation 450-490/ beam spliter 510, emission 515-565) and Axiovert 25 (filterset 25, excitation 400+495+570, beam splitter 410+505+585, emission 460+530+625) all from Carl Zeiss (Jena, Germany), in combination with an HBO 50 ac from Osram (Munich, Germany) were used for observing fluorescence of yeast cultures. The cells were not previously treated and the documentation was done with a digital camera (FE-180, Olympus, Duesseldorf, Germany).

2.9.2 FLUORESCENCE PHOTOMETRY

The measurement of the relative fluorescence was accomplished with the help of the Hitachi fluorescence spectrophotometer F2500 with a detection volume of 2 mL and its connected computer program FL solution (Hitachi, Chiyoda, Japan). Before the measurement, the TBS washed cells were resuspended in 2 ml H_2O and the OD_{600} was measured, whereupon only the calculated OD_{600} values between 0.8 and 4 further have been used. The obtained fluorescence values were relativized with the respective OD_{600}.

2.9.3 FLUORESCENCE DETECTOR

With the combination of a Varian Pro Star 363 Fluorescence Detector (Agilent Technologies, Boeblingen, Germany) and a Pharmacia Biotech pump p-1 (GE Healthcare, Munich, Germany) a fluorescence detector was built with a detection volume of 13 µl. Before the measurement was started, the cells - including a non-fluorescent control - were washed with TBS, re-suspended in H_2O and subsequently the OD_{600} was assimilated. For the assay, a cell suspension volume of at least 20 ml was used and kept on ice. The value of the fluorescence was obtained from Varian Pro Star 363 fluorescence detector and relativized with the respective OD_{600}.

3 RESULTS

The development of an online reporter system to follow plasmid stability in *Saccharomyces cerevisiae* required the examination of the issues reporter genes, promoter strengths and semi artificial promoter control. The results of these experiments will be presented in the following paragraphs, but for completeness, results about the intricate way to an in regard to fluorescence and localization working and expressed reporter gene will be shown first.

3.1 PRE-EXPERIMENTS FOR REPORTER GENE EXPRESSION

As a basis for the experiments, the pUG plasmid series (pUG23, pUG34, pUG35, pUG36, see table 2, Hegemann *et al.* 1999) of the working group around Hegemann (Duesseldorf, Heinrich-Heine-University) was kindly provided. These plasmids contain yeast enhanced GFP (yEGFP3) under the control of the by methionine repressed yeast $MET25_{prom}$ and as auxothrophic markers the functional copies of the yeast *URA3* and *HIS3* genes as well as a copy of the centromeric sequence of the yeast chomosome 6 and a yeast chromosomal replication origin (CEN6/ARSH4). The fluorescent protein was established by Cormack and co-workers (1996) and includes optimized yeast codon usage as well as the amino acid exchange of Ser^{65} to Gly and Ser^{72} to Ala for higher fluorescence and faster chromophore building. For a better understanding in each paragraph the important features of all constructed plasmids are summarized in tables (here tables 4 and 5).

The goal of the first experiments was the alteration of the promoter of pUG35 from a depressed ($MET25_{prom}$) to a constitutive stage. However, an exchange based just on restriction sites *Sac*I and *Sal*I with parts of the *KEX2* promoter (-13 to -490 bp, calculated upstream from the start codon ATG of the gene) let to a seemingly non-functional plasmid JOE0001 (see figure 2 A for structure) in regard of its ability to enable cells to produce fluorescent proteins. In addition, transformation and integration with this plasmid as well as only transformation experiments with pUG35 showed concerns about the suitability

of the *S. cerevisiae* strain DBY947. In a set of PCR experiments with the primers forJOE0003 – revJOE0009 (they generate specific fragments of 721, 680, and 778 bp) it could be observed, that the marker structure of DBY947 was not suitable for further experiments, thus leading to the usage of *S. cerevisiae* strain SY992 from Euroscarf (European *Saccharomyces Cerevisiae* archive, Frankfurt, Germany). With this strain, a stable control of transformation and integration (YJOE0001) was then possible with PCR fragments.

Table 4: Summary of the features of the plasmid constructs JOE0001-JOE0027 and pUG35 as described in the paragraph 3.1

Plasmid	Promoter	N-terminal extension	Reporter protein	C-terminal extension	Intended subcellular localization
pUG35	MET25	-	yEGFP3	-	Cytoplasm
JOE0001	KEX2	-	yEGFP3	-	Cytoplasm
JOE0004	KEX2	$Suc2_{secr}$::SerArg::ΔMet_1	yEGFP3	Sag1p[1] (319 aa)	Cell surface
JOE0005	ACT1	$Suc2_{secr}$::SerArg::ΔMet_1	yEGFP3	Sag1p[1] (319 aa)	Cell surface
JOE0006	ADH1	$Suc2_{secr}$::SerArg::ΔMet_1	yEGFP3	Sag1p[1] (319 aa)	Cell surface
JOE0007	PGK1	$Suc2_{secr}$::SerArg::ΔMet_1	yEGFP3	Sag1p[1] (319 aa)	Cell surface
JOE0008	ACT1	$Suc2_{secr}$::SerArg::ΔMet_1	tdTomato	Sag1p[1] (319 aa)	Cell surface
JOE0009	ADH1	$Suc2_{secr}$::SerArg::ΔMet_1	tdTomato	Sag1p[1] (319 aa)	Cell surface
JOE0010	PGK1	$Suc2_{secr}$::SerArg::ΔMet_1	tdTomato	Sag1p[1] (319 aa)	Cell surface
JOE0011	ACT1	$Suc2_{secr}$::SerArg::ΔMet_1	mCherry	Sag1p[1] (319 aa)	Cell surface
JOE0012	ADH1	$Suc2_{secr}$::SerArg::ΔMet_1	mCherry	Sag1p[1] (319 aa)	Cell surface
JOE0013	PGK1	$Suc2_{secr}$::SerArg::ΔMet_1	mCherry	Sag1p[1] (319 aa)	Cell surface
JOE0014	ADH1	MetProSerArg::ΔMet_1	yEGFP3	Sag1p[1] (319 aa)	Secretion
JOE0015	ACT1	MetProSerArg::ΔMet_1	yEGFP3	Sag1p[1] (319 aa)	Secretion
JOE0016	PGK1	MetProSerArg::ΔMet_1	yEGFP3	Sag1p[1] (319 aa)	Secretion
JOE0017	KEX2	MetProSerArg::ΔMet_1	yEGFP3	Sag1p[1] (319 aa)	Secretion
JOE0018	ACT1	$Suc2_{secr}$::SerArg::ΔMet_1	yEGFP3	His(8)tag::KDEL	ER
JOE0019	ADH1	$Suc2_{secr}$::SerArg::ΔMet_1	yEGFP3	His(8)tag::KDEL	ER
JOE0020	PGK1	$Suc2_{secr}$::SerArg::ΔMet_1	yEGFP3	His(8)tag::KDEL	ER
JOE0021	ADH1	MetProSerArg::ΔMet_1	yEGFP3	His(8)tag::KDEL	Cytoplasm
JOE0022	ACT1	MetProSerArg::ΔMet_1	yEGFP3	His(8)tag::KDEL	Cytoplasm
JOE0023	PGK1	MetProSerArg::ΔMet_1	yEGFP3	His(8)tag::KDEL	Cytoplasm
JOE0024	ADH1	MetProSerArg::ΔMet_1	yEGFP3	His(8)tag::SKL	Peroxisome
JOE0025	ACT1	MetProSerArg::ΔMet_1	yEGFP3	His(8)tag::SKL	Peroxisome
JOE0026	PGK1	MetProSerArg::ΔMet_1	yEGFP3	His(8)tag::SKL	Peroxisome
JOE0027	PGK1	$Suc2_{secr}$::SerArg::ΔMet_1	yEGFP3	His(8)tag::Sag1p[2]	Cell surface

Legend:
1 C-terminal domain of Sag1 (Ser^{331}-Phe^{650})
2 C-terminal domain of Sag1 (Val^{374}-Phe^{650})

Based on the results of the fluorescence tests with induced yEGFP3 of pUG35 in SY992, where cells have either been disintegrated with glass beads or not, it was assumed, that a yEGFP3 outside of the cell could have a higher fluorescence by avoiding a potential quenching effect of cellular compounds

which, in turn, might reduce signal strength. This experiments lead to a strategy where the reporter protein was located on the cell surface. For this reason - following the intermediate constructs JOE0002 and JOE0003 - the plasmid JOE0004 (see figure 2 B) was constructed with the promoter and wild type sequence ATG surrounding (Kozak sequence) of *KEX2* (+3 to -490 bp), and a fusion of yEGFP3 with the N-terminal *Suc2* secretion signal (*SUC2*$_{secr}$) 19 aa, which is short and commonly used in experiments. The sequence for the C-terminal domain (Ser^{331}-Phe^{650}) of the GPI anchored protein Sag1, including its terminator sequence, was also added. Since no fluorescence was observed, *KEX2*$_{prom}$ was considered too weak and was therefore exchanged for the stronger *ACT1*$_{prom}$ (+3 to -515 bp, JOE0005), *ADH1*$_{prom}$ (+3 to -426 bp JOE0006) or *PGK1*$_{prom}$ (+3 to -772 bp, JOE0007) with the same negative result, indicating that likely not the expression itself of the chimeric gene is responsible for the lack of signal detection. After having tested non secretion signal bearing constructs (JOE0014-17) and the reporters mCherry (JOE0011-13) as well as tdTomato (JOE0008-10), whereupon the latter being particularly well known for its high signal intensity, the problem was predicted to be either located at the terminator region or at the C-terminal fusion of Sag1 C-terminal GPI anchoring peptide sequences to yEGFP. Out of these reasons, both sequences were exchanged with the terminator region of the gene *CYC1* (*CYC1*$_{term}$).

In order to obtain a more efficient detection of the reporter proteins a His(8)tag sequence was attached to their C-terminal, which is known for not significantly disturbing the fluorescence activity (Pitson *et al.* 1999). With the aim of retaining a localization inside of different cell compartments, the sequences for the C-terminal ER retention (KDEL, Stornaiuolo *et al.* 2003, JOE0018-23) or peroxisome signal (SKL, Subramani *et al.* 2001, JOE0024-26) were additionally attached to the 3`sequences of the reporter genes. Concentrating first on the reporter yEGFP3, neither of these plasmids nor the plasmid JOE0027 - containing a shorter Sag1 GPI anchor domain (Val^{374}-Phe^{650}) – enabled the cells to display the desired fluorescence. Negative western plot analyses indicated a problem connected to the promoter respectively Kozak sequence, based on the cloning strategy (see Appendix 7.6.1 for western blot results).

Table 5: Summary of the features of the plasmid constructs JOE0028-JOE0039 as described in the paragraph 3.1

Plasmid	Promoter	N-terminal extension	Reporter protein	C-terminal extension	Intended subcellular localization
JOE0028	PGK1	$Suc2_{secr}$::MetGln::ΔMet_1	yEGFP3	His(8)tag::KDEL	ER
JOE0029	PGK1	$Suc2_{secr}$::MetGln::ΔMet_1	yEGFP3	His(8)tag::Sag1[1]	Cell surface
JOE0030	PGK1	MetGln::ΔMet_1	yEGFP3	His(8)tag::KDEL	Cytoplasm
JOE0031	PGK1	MetGln::ΔMet_1	yEGFP3	His(8)tag::SKL	Peroxisome
JOE0032	ADH1	MetGln::ΔMet_1	yEGFP3	His(8)tag::KDEL	Cytoplasm
JOE0033	ADH1	MetGln::ΔMet_1	yEGFP3	His(8)tag::SKL	Peroxisome
JOE0034	ACT1	MetGln::ΔMet_1	yEGFP3	His(8)tag::KDEL	Cytoplasm
JOE0035	ACT1	MetGln::ΔMet_1	yEGFP3	His(8)tag::SKL	Peroxisome
JOE0036	KEX2	MetGln::ΔMet_1	yEGFP3	His(8)tag::KDEL	Cytoplasm
JOE0037	KEX2	MetGln::ΔMet_1	yEGFP3	His(8)tag::SKL	Peroxisome
JOE0038	PGK1	MetGln::ΔMet_1	mCherry	His(8)tag::KDEL	Cytoplasm
JOE0039	PGK1	MetGln::ΔMet_1	mCherry	His(8)tag::SKL	Peroxisome

Legend:
1 C-terminal domain of Sag1 (Val^{374}-Phe^{650})

Interestingly, it was possible to exchange the *Xba*I to a *Sph*I site through a simultaneous ligation of in total three fragments, leading to the plasmids JOE0028 (yEGFP3 ER localization through KDEL) and JOE0029 (yEGFP3 cell wall localization through Sag1 GPI anchor domain) with $SUC2_{secr}$ as well as $PGK1_{prom,}$ which however did not enable the cells to produce any fluorescent protein. In contrary, the plasmids JOE0030 (yEGFP3 cytoplasm localization, see figure 2 C) and JOE0031 (yEGFP3 peroxisome localization through SKL, see figure 2 D) possessing the $PGK1_{prom}$ enabled the cells to create fluorescent proteins in the right localization. These two plasmids built the basis for the test of the reporter mCherry (JOE0038 and JOE0039, see figure 2 E and F) on the one hand, and on the other hand for the comparative analysis of the $ACT1_{prom}$ (+4 to -515 bp, JOE0034 and JOE0035), $ADH1_{prom}$ (+3 to -411 bp, JOE0032 and JOE0033) and $KEX2_{prom}$ (JOE0036 and JOE0037). These results will be discussed in the next paragraph.

3.2 Promoter strength

As mentioned before, it is important to test the actual promoter strength of each construct, because the environmental respectively experiment conditions as well as the combination with the reporter can influence the outcome. The

stability of yEGFP3 does not allow a timely precise subdivision in the range of minutes like other reporters do, but of several hours, which is sufficient to get a general idea though. Additionally, the usage of the single copy CEN6/ARSH4 plasmids simulates the later integration into the genome in regard to gene number and circumvents incomparability through plasmid number differences.

Table 6: Summary of the features of the plasmid constructs as described in the paragraph 3.2

Plasmid	Promoter	N-terminal extension	Reporter protein	C-terminal extension	Intended subcellular localization
JOE0030	PGK1	MetGln::ΔMet1	yEGFP3	His(8)tag::KDEL	Cytoplasm
JOE0032	ADH1	MetGln::ΔMet1	yEGFP3	His(8)tag::KDEL	Cytoplasm
JOE0034	ACT1	MetGln::ΔMet1	yEGFP3	His(8)tag::KDEL	Cytoplasm
JOE0036	KEX2	MetGln::ΔMet1	yEGFP3	His(8)tag::KDEL	Cytoplasm

Hence, for each - with a focus on the constructs leaving the yEGFP in the yeast cytoplasm - JOE0030, JOE0032, JOE0034 and JOE0036 (see table 6), three parallel tests were conducted with selective SD medium containing only 1% (w/v) of glucose, to simulate over time glucose free medium or more precisely the small local nutrition variation of big scale experiments. Due to this background, a promoter with constant high production despite nutrition differences in the media and with a small variance is preferable.

After 4, 8 and 24 hours samples were taken and the relative fluorescence was measured showing as a first result that the $KEX2_{prom}$ is not able to produce a sufficient amount of reporter to be detected by the fluorescence photometer and has therefore been left out. The remaining results for $ACT1_{prom}$, $ADH1_{prom}$ and $PGK1_{prom}$ are shown in figure 3.

It can be observed, that under the control of $ACT1_{prom}$ a rather constant relative fluorescence of around 6-7 was generated in the first hours (4 and 8 h) and even a relative constant signal in glucose free conditions (24 h). However, the construct with $ADH1_{prom}$ appears different since the cells apparently produced in relation a strong relative fluorescence signal at the beginning (9-10, 4h) which was then decreasing (6-7, 8h) through the apparent dependency of the promoter to glucose and even not detectable at the end (24h). Such a glucose dependency can also be observed for $PGK1_{prom}$, but not in a similar amount and amplitude. The highest relative fluorescence (17-18, 4h) can be observed at the

first measurement point from where it was only slightly diminished with declining glucose concentration (15-16, 8h) and showed no interference by the growth state (16-17, 24h), too. Overall, the observations are comparable to the results of the work from Monfort *et al.* (1999) and Partow *et al.* (2010). Because of all these depicted reasons, the *PGK1*$_{prom}$ was chosen for further experiments concerning the subcellular localization.

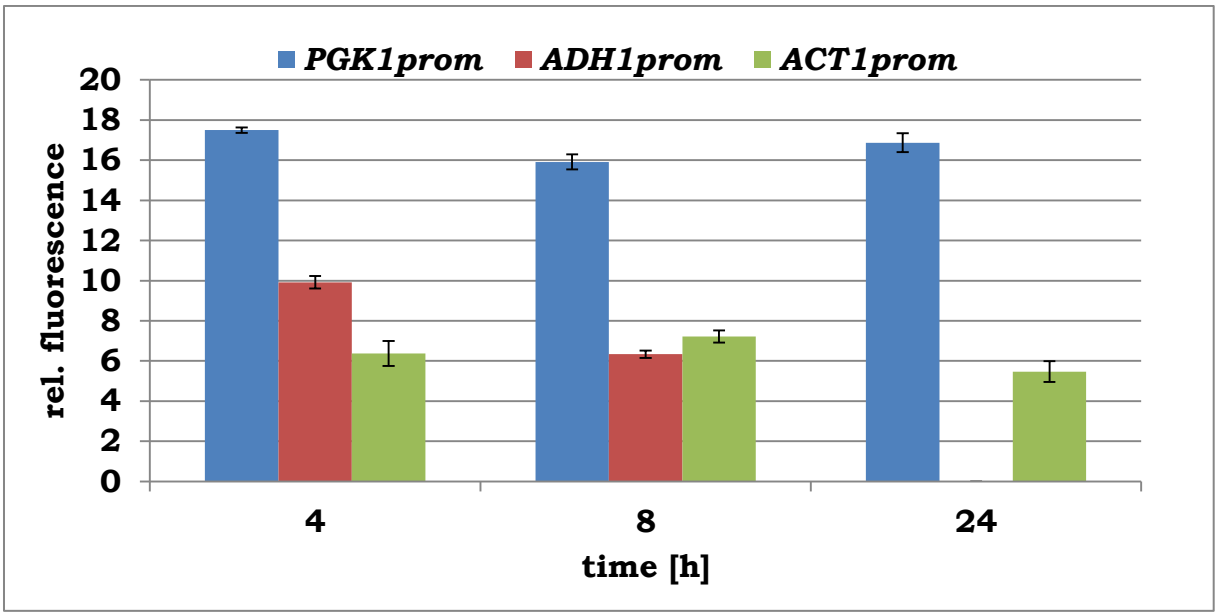


Figure 3: Comparison of yeast protomer activity at different growth phases. At the control points 4, 8 and 24 hours the relative fluorescence (OD_{600} adjusted to 2, see 2.4.2 and 2.9.2 for details) of in selective SD medium (1% glucose) grown SY922 strains containing a centromeric plasmid with yEGFP3 under the control of *ACT1*$_{prom}$ (green), *ADH1*$_{prom}$ (red) or *PGK1*$_{prom}$ (blue) was measured three times respectively (details in table A, Appendix 7.1).

3.3 LOCALIZATION

In order to extend the scope of possible localization of the reporter protein from yet cytoplasm and peroxisome, another effort was undertaken to direct the reporter protein onto the cell wall or ER. Also one of the strongest promoters (*TDH3*$_{prom}$, -1 to -699 bp) in the plasmids JOE0040 (ER) and JOE0041 (cell wall) combined with a 5´ *SUC2*$_{secr}$ fusion (see table 7) to the reporter gene led to no detectable fluorescence in the cells, inducing a literature based fault analysis with a positive result.

Huang and Shusta (2005) describe that the attempts to secrete high concentration of GFP or GFP fusion proteins fail due to the low or even

inexistent expression. This problem also occurs in other cell systems and is therefore not connected to yeast itself, but apparently to the combination of secretion signal respectively leader peptide and target protein. The former results indicate that the fusion of the reporter gene to *SUC2*$_{secr}$ also not enabled cells to secrete fluorescent proteins like yEGFP3. As a potential alternative the secretion signal of the killer toxin virus K28 is mentioned as well as an artificial sequence, but both exceed or conterminous to the possible size for primer manipulation and seemingly work optimal only at low temperatures (Huang and Shusta 2005).

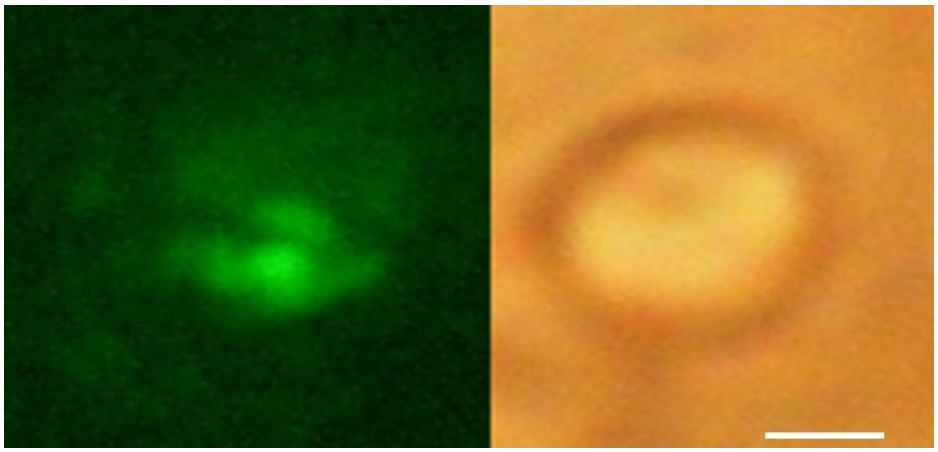


Figure 4: **Subcellular localization of yEGFP3 of SY922 with plasmid JOE0043** (*PGK1prom*; K28 secretion signal; yEGFP3; C-terminal domain of Sag1). Overnight culture conducted in selective SD-medium. Left side fluorescence and right side bright field picture performed with the Axiovert 25. The white bar is equivalent to 5 µm in a 1000 fold magnification.

Nevertheless, the 30 aa K28 secretion signal – sequence based on Schmitt and Tipper (1995) - was tested and inserted into the plasmids JOE0042 (ER) and JOE0043 (cell wall) with yEGFP3 under the control of *PGK1*$_{prom}$ and the result can be seen exemplarily for JOE0043 in figure 4.

Table 7: Summary of the features of the plasmid constructs as described in the paragraph 3.3

Plasmid	Promoter	N-terminal extension	Reporter protein	C-terminal extension	Intended subcellular localization
JOE0031	PGK1	MetGln::ΔMet$_1$	yEGFP3	His(8)tag::SKL	Peroxisome
JOE0030	PGK1	MetGln::ΔMet1	yEGFP3	His(8)tag::KDEL	Cytoplasm
JOE0040	TDH3	Suc2secr::MetGln::ΔMet1	yEGFP3	His(8)tag::KDEL	ER
JOE0041	TDH3	Suc2secr::MetGln::ΔMet1	yEGFP3	His(8)tag::Sag1[1]	Cell surface
JOE0042	PGK1	MetGln::K28::ΔMet1	yEGFP3	His(8)tag::KDEL	ER
JOE0043	PGK1	MetGln::K28::ΔMet1	yEGFP3	His(8)tag::Sag1[1]	Cell surface

Legend:
1 C-terminal domain of Sag1 (Val^{374}-Phe^{650})

While on the one hand a fluorescence signal was visible, on the other hand the growth temperature of in this case 28°C - which was higher than recommended but often used in large scale experiments - seemed to lead to a distribution of the reporter protein presumably mostly into the ER. However, a comparison to localization pictures from “www.yeastgfp.yeastgenome.org” was not clearly.

For this reason, only the cell locations cytoplasm and peroxisome were used for additional tests and improvements. A microscopically picture of both variants can be seen in figure 5. In the case of the cytoplasmic localization, the whole cell was illuminated as expected, whereas in the case of the peroxisome localization few sharp spots were visible even more than expected through the used medium.

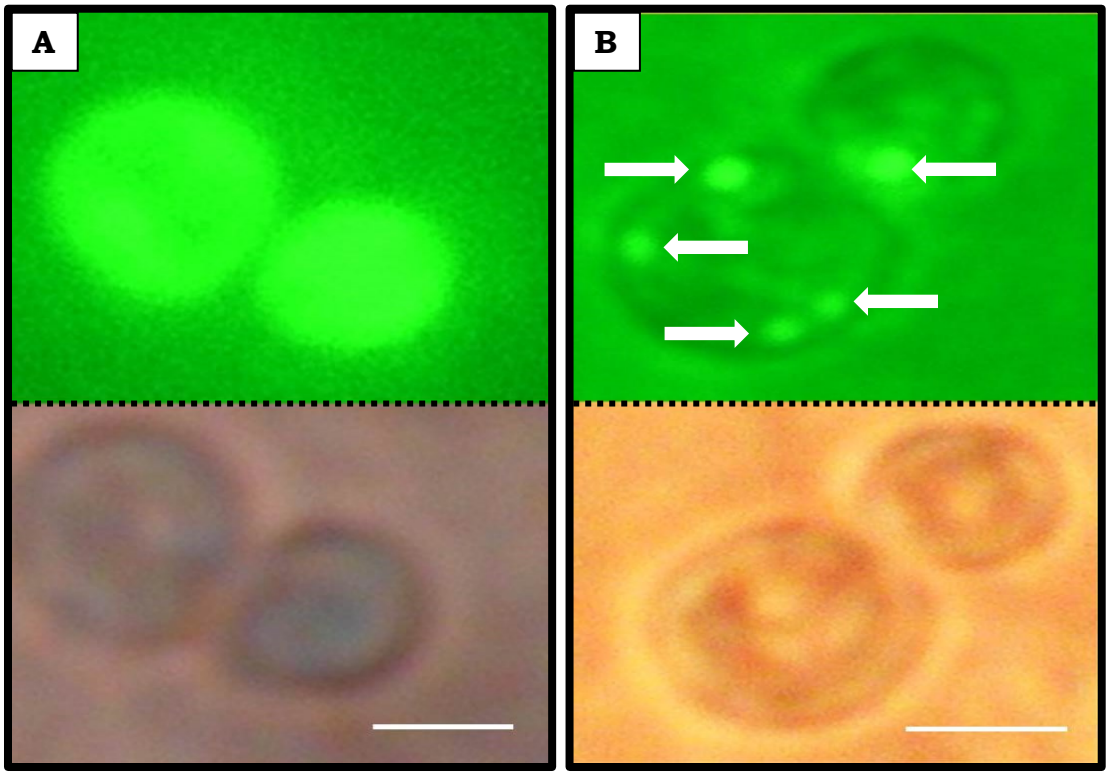


Figure 5: **Subcellular localization of yEGFP3 of SY922 with plasmid JOE0030** (*PGK1*$_{prom}$; yEGFP3; KDEL ER retention signal, left side, **A**) **and JOE0031** (*PGK1*$_{prom}$; yEGFP3; SKL peroxisome location signal, right side, **B**). Overnight cultures conducted in selective SD-medium. At the top the fluorescence and at the bottom the bright field picture performed with the Axiovert 25 is visible. The white bar is equivalent to 5 µm in a 1000 fold magnification and the white arrows indicate the peroxisome location.

For both varieties a similar relative fluorescence of around 18 can be observed for cells in shaken flasks, but apparently not in non-shaken flasks (for both see figure 6). There, a comparatively lower relative fluorescence could be measured as well as an apparent difference in the values for the peroxisome (around 16)

and cytoplasm (around 15) location. With a *p*-value of 11.3, this difference was not significantly sufficient, but sufficient enough to determine the peroxisome as the location of choice for the highest and most constant relative fluorescence signal.

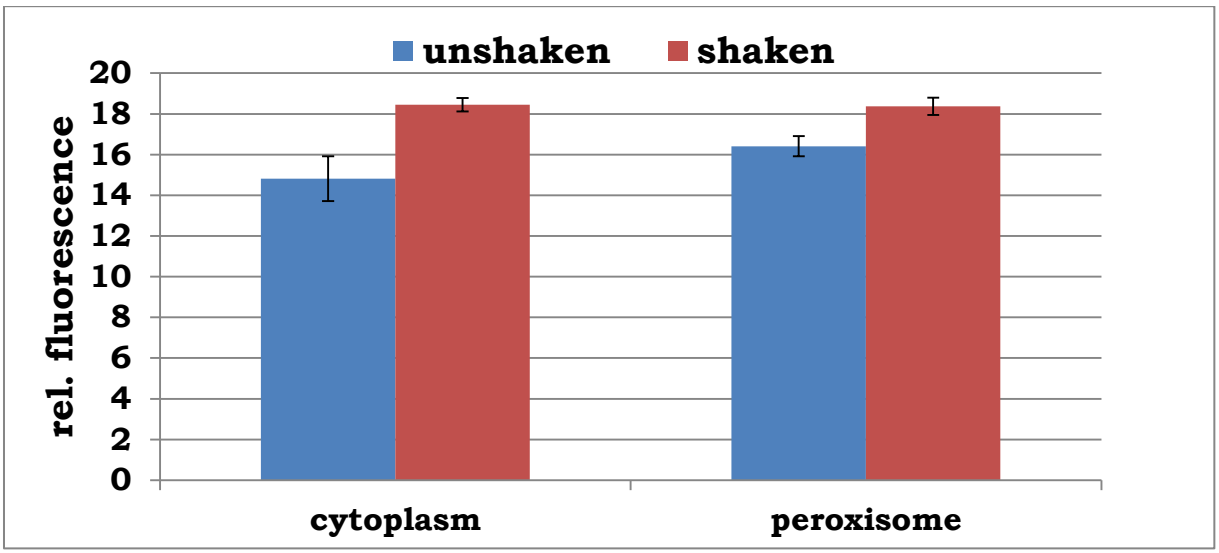


Figure 6: **Comparision of the relative fluorescence of yEGFP3 in the cell location cytoplasm and peroxisome under shaken or unshaken conditions.** Overnight cultures in selective SD medium of SY922 containing the plasmid JOE0030 ($PGK1_{prom}$; yEGFP3; KDEL ER retention signal, left side) and JOE0031 ($PGK1_{prom}$; yEGFP3; SKL peroxisome location signal, right side) were analyzed in respect to their relative fluorescence (see 2.9.2) under shaken (red) or unshaken (blue) conditions respectively three times (details in table B, Appendix 7.1).

3.4 Reporter selection

After determining the localization, the different reporter proteins will be compared in the following. With the self-defined requirement of an integration cassette as small as possible respectively expression burden through the reporter, the tdTomato (476 aa) - with around doubled size compared to mCherry (256 aa) and yEGFP3 (238 aa) - was left out, especially since its codon usage was not yet optimized for yeast like the other two reporters.

Table 8: Summary of the features of the plasmid constructs as described in the paragraph 3.4

Plasmid	Promoter	N-terminal extension	Reporter protein	C-terminal extension	Intended subcellular localization
JOE0038	PGK1	MetGln::ΔMet$_1$	mCherry	His(8)tag::KDEL	Cytoplasm
JOE0039	PGK1	MetGln::ΔMet$_1$	mCherry	His(8)tag::SKL	Peroxisome

The codon usage can have a significant influence on the expression and thus on the concentration respectivly intensity of the particular reporter protein or signal. Both, mCherry (excitation 587 nm/ emission 610 nm) and yEGFP3 (excitation 488 nm/ emission 512 nm) have this optimized codon usage, but differ among others in their excitation and emission wavelengths.

These differences seemed to have a significant influence on the detection limit of the reporter signals. In figure 7 a microscopically picture of SY922 transformed with the plasmids JOE0038 (cytoplasm) and JOE0039 (peroxisome) is depicted (see table 8) and in comparison to figure 5 it showed nearly a similar picture. Only in the peroxisome version simply one sharp spot could be observed.

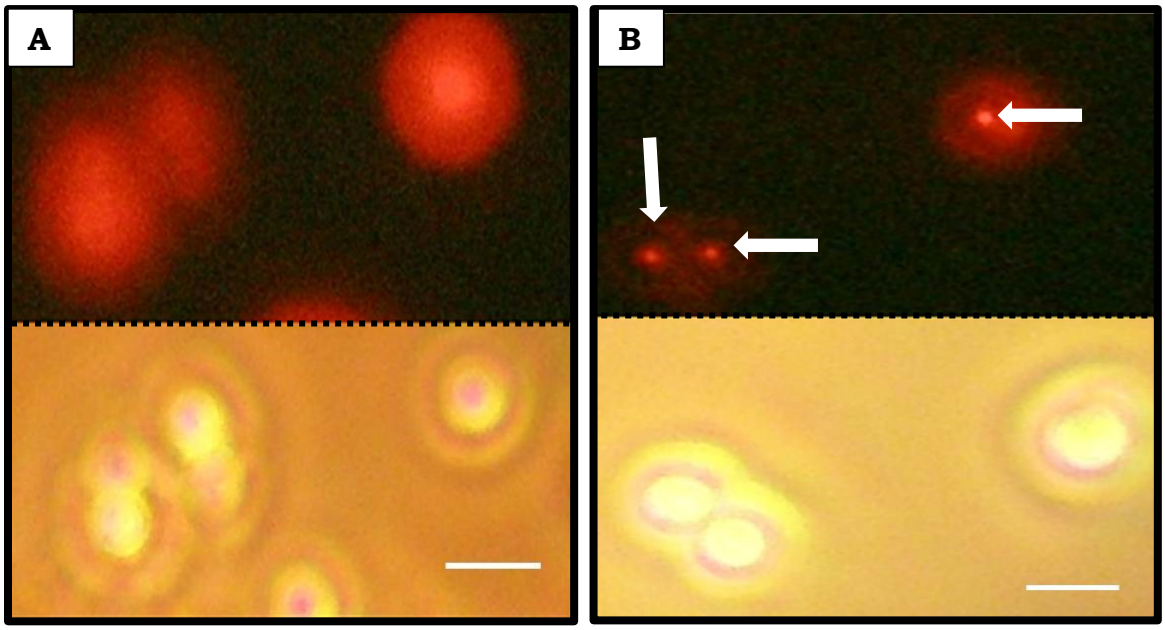


Figure 7: **Subcellular localization of yEGFP3 of SY922 with plasmid JOE0038** (*PGK1*$_{prom}$; yEGFP3; KDEL ER retention signal, left side, **A**) **and JOE0039** (*PGK1*$_{prom}$; yEGFP3; SKL peroxisome location signal, right side, **B**). Overnight cultures performed in selective SD-medium. At the top the fluorescence and at the bottom the bright filed picture conducted with the Axiostar plus FL is visible. The white bar is equivalent to 10 µm in a 400-fold magnification and the white arrows indicate the peroxisome location.

Despite their optical identity, it was surprisingly not possible to detect a fluorescence signal of the cells with JOE0038 and JOE0039 with the help of the fluorescence photometer. This and all the previous results led to the decision to use the combination of yEGFP3 located in the peroxisome under the

expression control of *PGK1*$_{prom}$ as a basis for the development of an phage based detection system.

3.5 DEVELOPMENT OF AN PHAGE BASED DETECTION SYSTEM

3.5.1 CONSIDERATIONS ABOUT THE PHAGE SYSTEM

The hitherto use of the phage system as in Webster and Brammer (1995) is grounded on a plasmid-based constitutive expressed 434 phage CI repressor controlling a genomically incorporated *LacZ* gene under the control of *PGK1*$_{prom}$ with a single 434 operator site integrated at different locations. By that, the activity was measured with a β-galactosidase assay and with integration of the operator site next to the TATA box or next to the upstream activating site of the promoter a relative repression in presence of the repressor of around 65 respectively 31% was achieved.

A first improvement could be made by basing the repressor system on the lambda phage, where the repressor binds more directly to the DNA and to the specific 17 bp operator sequences instead of the 14 bp in 434 phages (Aggarwal *et al.* 1988). Additionally, as a tetramer the repressor develops an even higher affinity and binding to DNA, thus making the insertion of two adjacent operator sites necessary for reaching wild type repression levels.

Wedler and Wambutt followed, as already mentioned, such a strategy in 1995, by integrating the repressor sequence into the yeast genome with constituive expression and the α-amylase being under the control of *TDH3*$_{prom}$ with operator sites on a multi copy plasmid. This made the repression calculation complex and the artificial TATA box was interfering with the promoter structure when inserted in different location than the original TATA box region.

In both approaches the promoter structure - respectively distances between its functional elements - were disregarded as well as the loop forming which improves the repression in the wild type model. Additionally, neither possible improvement of operator sequence and repressor structure nor fine-tuning of the repressors expression was carried out. Such enhancements should improve the repression tremendously and are discussed in the following.

3.5.1.1Loop formation

The forming of DNA loops is widespread in different eukaryotic cell processes (Ptashne 2011), and can approximately be separated into the two forms short and energetic as well as long and entropic. In short loops, about 150 bp or shorter, DNA elasticity arising to sequence or bound protein properties is the crucial factor for forming, which can be observed for example in the lac repressor binding. Additionally, the free energy oscillates depending on the distance and location of the operator sites within the helical DNA periodicity of around 10.9 bp, whereas for long loops free energy increases almost linearly with the distance, but both in small amplitudes. For such long loops as in the case of the lambda phage octomerization, the limiting step is the time and probability of distant repressors associating with each other, which is most of the time between 10–20 ms, if one repressor tetramer is already bound (Saiz and Vilar 2006).

Interestingly, the lambda repressors even form a DNA loop if the operator sites O_L and O_R are close together in a proportion of 5 DNA turns (around 10.5 bp per turn), but only in very high concentrations of the repressor if the proportion is between 4.6 or 5.5 turns (Griffith *et al.* 1986). This effect becomes unnoticeable beyond a total of 20 turns and even the location on the same DNA molecule is then not necessary for octamer cooperativity (Dodd *et al.* 2004). Still uncertain is thereby if a higher concentration of repressors seems to decrease the chance for looping (Saiz and Vilar 2006), or increases it through DNA bending by the unspecific binding of repressors (Liebesny *et al.* 2010), even if this bending is not extensive (Ptashne 2005). Clear is that through loop forming the affinity and specificity of repressors is increased as well as the transcriptional noise is dimmed (Saiz and Vilar 2006). However, it has to be kept in mind that these results like most of the others are based on *E. coli* and can only give hinds for the possibility or reaction of the system respectively components in yeasts.

Therefore, an additional, correctly localized O_L site in the construct will probably also in yeast lead to loop forming and improvement of the repression system especially in a euchromatin region.

3.5.1.2Repressor

The 3D structure of the repressor CI is known under different binding and cooperativity conditions to lead to a wide knowledge about structure and function of the repressor (Jordan and Pabo 1988). Similar to many DNA binding proteins the repressor consist of two domains, whereupon the C-terminal domain disposes the cooperativity by binding the other repressor molecule at the same domain, the oligomerisation domain, whereas the N-terminal binds to the DNA with a HTH motif (Dutreix *et al.* 1984). The repressors thereby recognize the operator sequence in the major groove of the DNA (Ptashne 2005).

The CI repressor has an autocatalytic peptide cleaving activity with the cleavage site being between Ala^{111} and Gly^{112} (Stayrook *et al.* 2008). The catalytic amino acids are thereby located in a small peptide chain, which connects both mentioned domains. There similar to serine proteases the RecA-mediated cleavage is performed, indicating a necessary disruption of the quaternary structure for exposing (Lewis 2011). By exchanging Ala^{111} through Thr, and Gly^{112} through Arg or Glu a repressor can be generated, which is nearly non-degradable by RecA mediation (Gimbel and Sauer 1985). Next to such manipulations, it is also possible to increase the operator binding ability of the CI repressor by amino acid substitutions. Thereby an up to 600 fold higher affinity resulting from changed dissociation and association rate could be reached, adversely connected to a higher unspecific binding (Nelson and Sauer 1985).

Consequently, only the mutation of the cleavage site seems to be useful for the improvement of the repression system. Rad51 is the yeast homologous of RecA and has a certain amount of similarity to it (Ogawa *et al.* 1993), however it is not clear yet if it functions in the same way.

3.5.1.3Operator

The DNA sequences of the operator sites O_L and O_R were completely obtained first in 1977 (Humayun *et al.* 1977), but of course already sequenced partially earlier. The lambda repressor has the highest affinity for the wild type O_R1 site, while the 17 bp are sufficient for binding (Benson *et al.* 1991). However, it is assumed that for example an AT rich spacer of the same length between the operators improves the binding (Maniatis *et al.* 1975) or base changes of the flanking sequences can increase (Benson *et al.* 1987), but also decrease the binding (Brenowitz *et al.* 1989). It has thus to be considered that the wild type distance between the three operators is important for functional cooperativity between the repressor dimers (Hochschild and Ptashne 1986), which is essential for an effective repression (Lewis 2011).

For these reasons, the operator sequences were in some cases particularly improved in the spacer regions in this work.

3.5.1.4Expression

Around 250 CI monomers can be found in a lysogenic lambda phage containing *E. coli* cell, 86% of them being bound unspecifically to DNA and acting as kind of activation buffer respectively reserve and 10 dimers being free (Bakk and Metzler 2004). The switch is normally activated by a threshold of around 60 CI molecules per cell (Little and Michalowski 2010). However, this depends on the loop forming activities, because even without loop forming a stable lysogenic state is possible, which is then more sensitive to CI depletion or mutations (Morelli *et al.* 2009).

It is not clear how many molecules are sufficient for a repression in yeast cells, keeping in mind that the nucleus is bigger than an *E. coli* cell and takes up only around 6-11% of the cell volume (Uchida *et al.* 2011). It has to be noted that the transport into the nucleus is bidirectional and mostly controlled by the nuclear pore complexes as well as directed by various signal sequences. One is for example the nuclear localization signal (NLS) which is essential for the import, consisting mostly of a continuous sequence of arginines or lysines on the cargo surface. Interestingly, small molecules under 60 kDa can also diffuse

through aqueous channels, but are seemingly transported actively as soon as they possess a NLS (Shulga *et al.* 1996). The ability of a NLS to act depends on its localization in the peptide chain and quantity as well as on the cargo protein properties (Nelson and Silver 1989). The existing pores lead to two assumptions: Firstly, that the pH is corresponding to the cytosol and is around 7, but it seems to be slightly higher (Seksek and Bolard 1996), and secondly, that around 30 kDa small NLS free molecules like the here used repressor and fluorescents proteins are equally distributed between nucleus and cytoplasm.

That is why the selected promoter for the CI repressor depends not only on the correct repressor concentration, but also on the chosen strategy. On the one hand, the small CI repressor can diffuse into the nucleus, but must therefore be expressed in a high concentration to enable repression, maybe influencing the fitness of the cell negatively. On the other hand a NLS sequence could increase the concentration of the repressor in the nucleus and a weaker promoter could be sufficient, but it is not clear what kind of influence such an additionally load of positively charged amino acids will have on the DNA binding and/or cooperativity of the repressor.

Yet it is sure that the plasmid carrying the CI repressor expression block should carry histidine as a selection marker, which reduces growth rate and copy number least (Karim *et al.* 2012). All the other options were tested in the course of this work, beginning with the promoter improvements concerning operator sequence insertions.

3.5.2 PROMOTER IMPROVEMENT

However, before going into detail with the results, it is important to get a general idea about the *PGK1* promoter structure and the strategy of the operator insertions.

The *PGK1* promoter (see figure 8 for schematic overview) belongs to the TATA box containing group of promoters and in this case even two functional boxes at -152 bp and -113 bp (calculated upstream from start codon) can be

observed. Although the TATA box at -152 bp seems to be preferred. Additionally, the UAS plays an important role for the high activity (Webster and Brammer 1995), in particular the region between -540 bp and -390 bp bound by Abf1 (-522/-495), Rap1 (-472/-459) and Gcr1 (-458/-427). Several other transcription factors are predicted to bind upstream of this region (www.rulai.cshl.edu/cgi-bin/SCPD/getgene2?PGK1), but do not seem to be very necessary. Interesting is the dynamic interaction between the factors, for example, the Abf1-binding site is only essential for correct initiation in absence of the Rap1 binding site, and otherwise its deletion has only a small negative effect on the promoter activity (Webster and Brammer 1995).

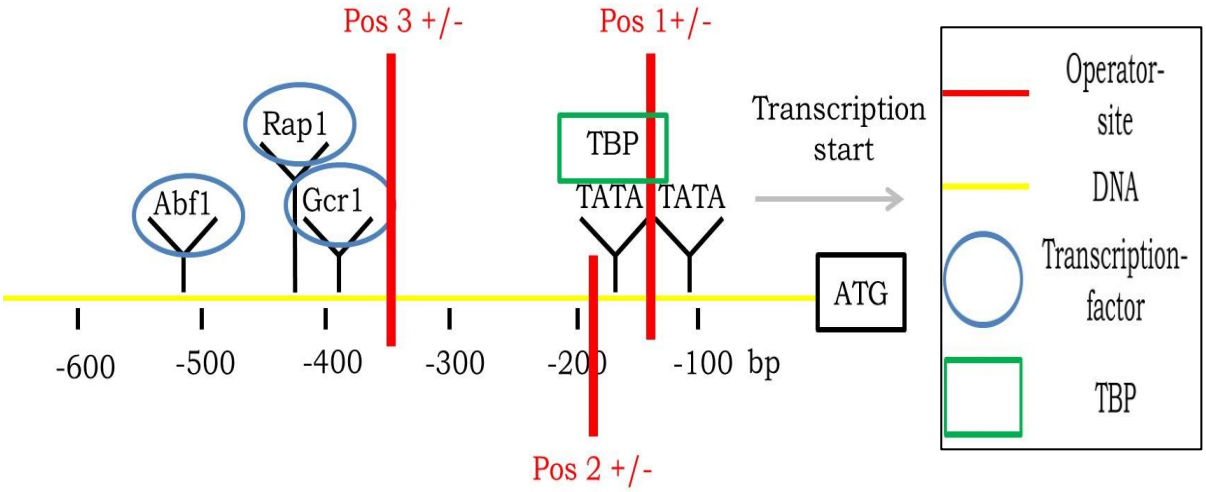


Figure 8: **Schematic illustration of the *PGK1* promoter**. The upstream activating sequences (UAS) are characterized through the binding of the transcription factors (blue circle) Abf1, Rap1 and GcR1. The core promoter region contains two possible TATA boxes where the TATA box binding proteins (TBP, green rectangle) can bind. Additionally indicated are the integration positions of the lambda phage operator sequences (red, Pos 1, 2 and 3), with "**+**" indicating an insertion of the operator site and "**-**" indicating a simultaneous deletion of upstream promoter sequences.

Having this and the studies of Wedler and Wambutt (1995) as well as of Webster and Brammer (1995) in mind, the considerations for the insertions of the lambda phage operator O_L are based on two assumptions: On the one hand, an insertion respectively binding of the repressor next to the 5´ (pos 2, -163/-208 bp) or 3´ (pos1, -130/-175 bp) end of the crucial TATA box should disturb the association of the TATA box binding proteins (TBP) to the TATA box. On the other hand, an insertion of the operator next to the 3´end of the Gcr1-binding site (pos3, -370/-415 bp) should lead to a reduced activity of the promoter, too, if the repressor is bound. All these positions are marked in red

in figure 8 and occur in two variants, with "**+**" indicating an insertion of the operator sequence and "**-**" indicating a simultaneous deletion of upstream promoter sequences to preserve the wild type promoter distances.

Table 9: Features of the plasmid construct used as a basis for the promoter improvements

Plasmid	Promoter	N-terminal extension	Reporter protein	C-terminal extension	Intended subcellular localization
JOE0031	PGK1	MetGln::ΔMet$_1$	yEGFP3	His(8)tag::SKL	Peroxisome

All promoters established by overlap extension PCR were tested for their ability to produce cells with fluorescence signals as plasmid based constructs (established over JOE0031 see table 9) before they were integrated into the *YPRCΔ15* region of SY992 (see appendix 7.5 for PCR control experiment results). The basis for this integration through homologous recombination was a PCR step generating fragments with a size between 3282 bp and 3516 bp depending on the original plasmid and containing next to the homologous sequences the marker *URA3* as well as the reporter gen construct.

In a first attempt, the complete reverse 84 bp O_R lambda operator with O_R1, O_R2 and O_R3 (see figure 9 A) was inserted at positions 1 (JOE0058) and 2 (JOE0059), generating each time a non-functional promoter. For this reason only the sequence of O_R1 and O_R2 (see figure 9 B) - which is in total 45 bp long - was inserted at the positions 1 (see exemplary figure 3 H), 2 or 3 (nPos) or in two of them. In some cases a 246 bp, O_L1 to 3 containing (see partially in figure 9 C) *Nae*I/*Nsi*I lambda phage fragment (ol) was added in the in the reporter gene construct (exemplarily see figure 3 I) to enable potential loop forming. In addition, a variation (TAT) on the method of Wedler and Wambutt (1995) was established by designing an artificial TATA box between two artificial operator sites from Maniatis *et al.* (1975) and exchanging the wild type TATA box with it (see figure 9 D).

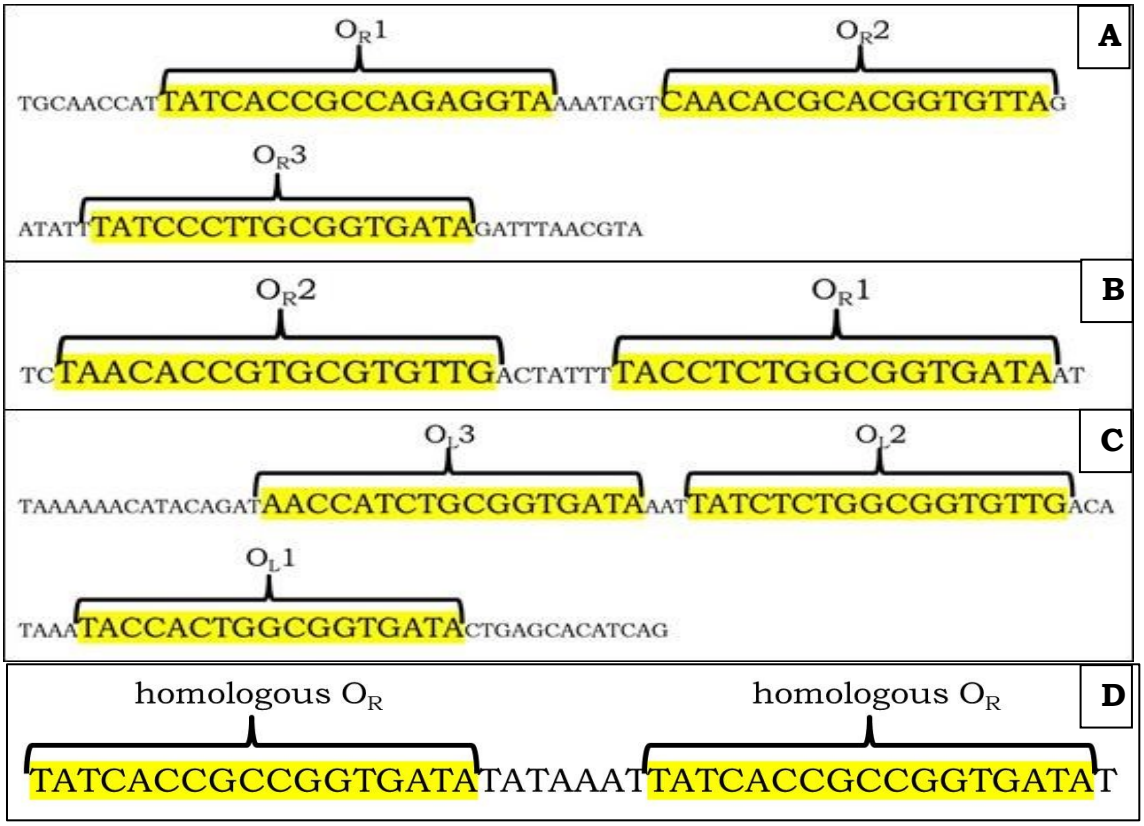


Figure 9: **Illustration of the used operator sites based on lambda phage sequences. A**: The reverse complete 84 bp O_R lambda operator with O_R1, O_R2 and O_R3. **B**: The sequence of O_R1 and O_R2, in total 45 bp long. **C**: Part of the 246 bp, O_L1 to 3 containing *Nae*I/*Nsi*I lambda phage fragment. **D**: TATA box between two homologous operator sites established by Maniatis *et al.* (1975), with the idea based on Wedler and Wambutt (1995).

In table 10, all constructs which have been used more intensively in this study are listed with the name of their origin plasmid, operator integration position, and *Saccharomyces cerevisiae* nomenclature as well as their number respectively possession of the applied operators O_L and O_R.

However, next to these constructs some more variants were tried to be created and tested, with either reduced or no detectable fluorescence (JOE0078, JOE0088) or unaccomplished construction process (addition of O_L *Nae*I/*Nsi*I fragment into JOE0073-76 and PCR of $PGK1_{prom}$ with nPos 2+3).

Table 10: List of the promoter variants (see figure 8) and their features as well as the plasmid basis for integration PCR.

Plasmid basis for integration PCR	Feature		Operator position	Strain name
	O_R	O_L	nPos	
JOE0061	X		-1	YJOE0003
JOE0060	X		1	YJOE0002
JOE0066	X	X	1ol	YJOE0014
JOE0062	X		2	YJOE0004
JOE0067	X	X	2ol	YJOE0015
JOE0063	X		-2	YJOE0005
JOE0068	X	X	-2ol	YJOE0016
JOE0065	X		-3	YJOE0007
JOE0070	X	X	-3ol	YJOE0018
JOE0064	X		3	YJOE0006
JOE0069	X	X	3ol	YJOE0017
JOE0076	XX		1-3	YJOE0011
JOE0075	XX		1+3	YJOE0010
JOE0074	XX		1-2	YJOE0009
JOE0073	XX		1+2	YJOE0008
JOE0077	X		TAT	YJOE0012

Legend: „ol“ addition of O_L fragment into construct after *URA3* marker
„+/positive number“ indicating an insertion of the operator site into the promoter
„-“ indicating a simultaneous deletion of upstream promoter sequences while an insertion of the operator site into the promoter
„TAT“ TATA box between two homologous operator sites established by Maniatis *et al.* (1975), with the idea based on Wedler and Wambutt (1995).
„X” number of features present

In advance it has to be mentioned that for a better understanding only the position terminology (nPos) will be used from now on for the different SY992 strains with integrated reporter gene cassettes listed in table 10. Moreover, cells with the plasmid constructs of -1 and 1+2 although showing no fluorescence, were used as a manner of negative control, but not for further experiments.

In a first test, the fluorescence of cells with integrated and plasmid based reporter constructs were compared. Results are presented exemplarily for nPos 1-2, 1+3 and 1-3 in figure 10.

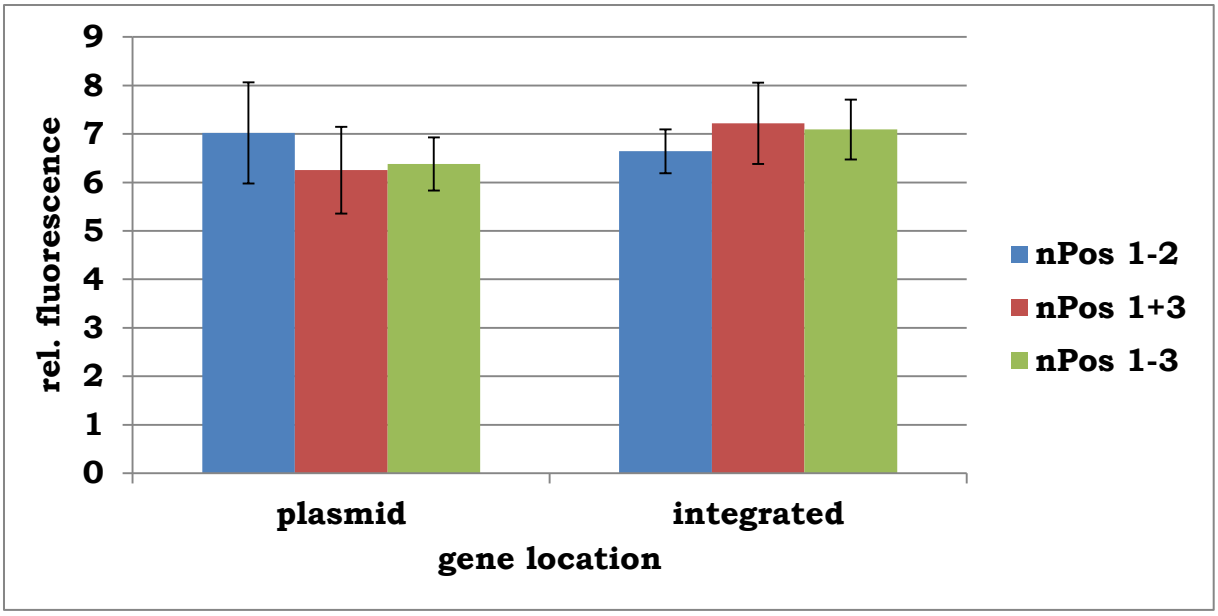


Figure 10: **Comparison of the relative fluorescence of yEGFP3 derived from a plasmid-based or chromosomally integrated expression block**. Overnight cultures in SD medium of SY922 containing the plasmid JOE0074 (blue), JOE0075 (red) and JOE0076 (green) as well as the strains nPos 1-2 (blue), nPos 1+3 (red) and nPos 1-3 (green) were analyzed and compared in respect to their relative fluorescence (see 2.9.2). Details of the measurement results are in table C, Appendix 7.1 and concerning the operator positioning see table 10 as well as figure 8.

As ilustrated in figure 10, no significant difference could be observed between at *YPRCΔ15* genome integrated and plasmid based expression, with calculated *p*-values between 21% and 61% (see Appendix 7.1 table D for details). Important to mention is that a fluorescence test with some rare clones of unknown integration (like YJOE0014 Clone 1, see appendix 7.5) in most cases showed lower signal strength.

In a threefold repeated experiment in SD medium, the fluorescence of cells with respectively each construct from table 10 was measured and illustrated in comparison to each other as well as to cells with the wild type *PGK1* promoter construct (figure 11). It could be noticed that only nPos 1, 1ol, -3 and -3ol reached a *p*-value (see Appendix 7.1 table D for details) high enough to call them comparable and equally strong as the wild type *PGK1* promoter (green). Furthermore, especially constructs with a double insertion (red) of the operator sites (1-3, 1+3, 1-2) as well as the TAT variation (yellow) seemed to have only around half of the wild type promoter activity. Surprisingly constructs with the additional insertion of the O_L fragment (light blue) appeared overall weaker

than the constructs without it (dark blue). Nevertheless, all constructs reached a sufficient high activity to detect the reporter protein easily.

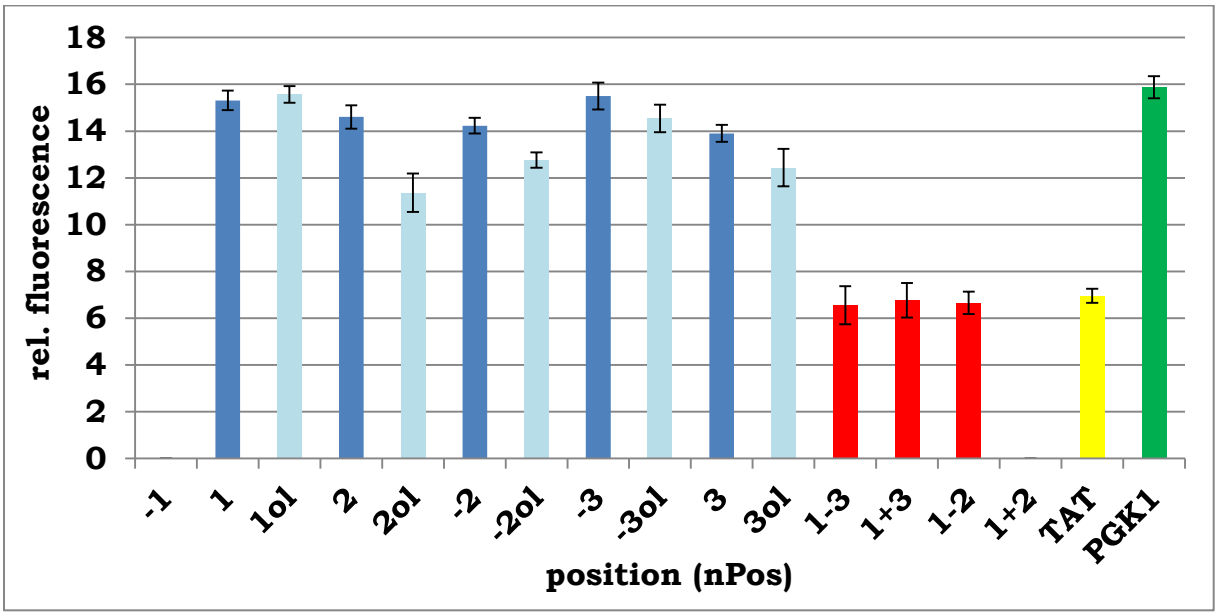


Figure 11: **Relative fluorescence values of cells expressing yEGFP3 under control of *PGK1* promoter variants**. In a threefold repeated experiment, the relative fluorescence (see 2.9.2) of the different yeast strains grown in selective SD medium was measured (details in table D, Appendix 7.1). The value of the wild type *PGK1* promoter (green) is provided as comparison guideline. The constructs with a double insertion of the operator sites (red) the TAT variation (yellow) and constructs with the additional insertion of the O_L fragment (light blue) as well as single constructs without it (dark blue) are marked correspondingly. For the operator positioning and detailed information of nPos -1, 1, 1ol, 2, 2ol, -2, -2ol, -3, -3ol, 3, 3ol, 1-3, 1+3, 1-2, 1+2 and TAT see table 10 as well as figure 8.

The possible influence of the inserted construct respectively reporter gene expression was substantiated by two separate tests of the mean growth rate per hour through OD_{600} measurement in YPDAU medium over time with the results of the different strains listed in table 11 (details in table Q, Appendix 7.2).

The result of the unmodified SY992 served as a benchmark for the following experiments by reaching growth rates between 0.3 and 0.32 h^{-1}. However, only the growth rates of the no yEGFP3 expressing yeast variants like nPos -1 (0.31 h^{-1}) and 1+2 (0.3 h^{-1}) achieved nearly similar levels, indicating maybe a balancing influence of the negative effect by the extra DNA burden and the positive effect by the *URA3* marker. The constantly high expression of yEGFP3 on the other hand in most cases (nPos 1, 1ol, -2, -2ol, -3) seemed to lead to a reduction of the growth rate to around 0.27 h^{-1} or for the double insertions

with lower expression (nPos 1-3, 1+3,1-2) to around 0.28 h^{-1}. However, the growth behavior of the constructs nPos 3 and 3ol seemed to form an exception, since despite a high expression a relatively high growth rate between 0.28 and 0.3 h^{-1} was observed. On the other site were the constructs nPos 2 and 2ol with a comparatively extremly low growth rate between 0.24 and 0.26 h^{-1}, but it has to be kept in mind that these growth experiments were only carried out twice. They are therefore subject to statistic variability especially regarding results of high variance like for example nPos -2, -2ol,-3ol and 1+3.

Table 11: List of the mean growth rate per hour in YPDAU medium from two independent experiments and mean relative fluorescence (see figure 11) of different yeast strains with integrated promoter variants (see table 10 as well as figure 8 for details about strains). The positive control is the strain SY992 without any modification (see appendix 7.2 table Q for details).

nPos	mean growth rate [h^{-1}]		mean rel. fluorescence
-1	0.309	0.267	0.000
1	0.271	0.272	15.310
1ol	0.243	0.279	15.567
2	0.248	0.248	14.607
2ol	0.262	0.269	11.364
-2	0.276	0.225	14.234
-2ol	0.078	0.276	12.767
-3	0.277	0.271	15.500
-3ol	0.191	0.288	14.545
3	0.301	0.284	13.904
3ol	0.237	0.283	12.440
1-3	0.133	0.279	6.554
1+3	0.284	0.286	6.767
1-2	0.295	0.073	6.654
1+2	0.300	0.303	0.000
TAT	0.281	0.298	6.960
SY992	0.296	0.316	0.000

Nonetheless, the results about the promoter improvement can be summarized in different terms. First, it can be recognized that an alteration of the *PGK1* promoter structure in the chosen locations is mostly possible without any severe changes of its strength, at least insofar as the single insertions and changes are seemingly around the size of 45 bp as well as not destroying important promoter sequences. Moreover, a separated insertion of two operator sequences in the case of the additional O_L site (nPos 1ol, 2ol, -2ol, -3ol, 3ol) is leading to a slight reduction and in the case of the insertion of O_R in two

positions (nPos 1-3, 1+3, 1-2, 1+2) to an almost half-reduced activity of the promoter. Additionally, it has to be noted that such an insertion as well as expression of a reporter gene reduces the fitness respectively growth rate of the host organism in most cases only slightly but considerably.

3.5.3 REPRESSOR IMPROVEMENT

After having built a basis for the online reporter system to detect plasmid instability through the insertion of the reporter gene into the genome under the control of a modified *PGK1* promoter, it is now time to present the results of the second part of the system, the repressor system and its enhancements.

The repressor plasmids was built based on the pUG24 (Hegeman *et al.* 1999), which possesses - among other features - a *HIS3* marker as well as *CEN6/ARSH4* sequences. This plasmid was then modified by a ligation of in total three fragments to form the new plasmid JOE0044, containing now the wild type CI repressor (see figure 12 A) not specific optimized for yeast codon usage under the control of the *KEX2* promoter instead of the yEGFP3 expression cassette, but still the CYC_{term}. Following the idea of Gimbel and Sauer (1985) a RecA recognition site mutation was generated in the sequence of the CI repressor by exchanging the sequence for Ala^{111} through Thr and Gly^{112} through Arg (JOE0045) via overlap extension PCR (see 2.5.6.2 and figure 12 B). Additional CI repressor variants with internal (JOE0046, see figure 12 C for details) C-terminal (JOE0048) or N-terminal (JOE0047) NLS - which possess thus additional positively charged aa - were also produced, replacing the wild type CI in JOE0044 (see table 12 for all plasmid details).

On the primarily followed assumption that a NLS marked repressor is transported into the nucleus with a higher affinity, the weak constitutive promoters of the genes *KEX2* (+4 to -490 bp, JOE0044), *HDA1* (+3 to -423 bp, JOE0056), *TRL1* (+4 to -287 bp, JOE0057) were chosen and assumed as sufficient with respect to their strength (Nacken *et al.* 1996, Tochigi *et al.* 2010). Given that hereby a proof of expression must be coupled to an effective detection method, for each promoter a CI variant with C-terminal ($KEX2_{prom}$:

JOE0050, *HDA1*$_{prom}$: JOE0052, *TRL1*$_{prom}$: JOE0054) or N-terminal (*KEX2*$_{prom}$: JOE0049, see exemplary figure 3 G, *HDA1*$_{prom}$: JOE0051, *TRL1*$_{prom}$: JOE0053) His(8)tag was created. Subsequently, the constructs were tested by western blot analysis for their respective expression level ability (see appendix 7.6.2). The negative result of *TLR1*$_{prom}$ was initially assumed to be based on a too short promoter sequence, which was applied again in an elongated form (+4 to -697 bp) in the plasmid JOE0072.

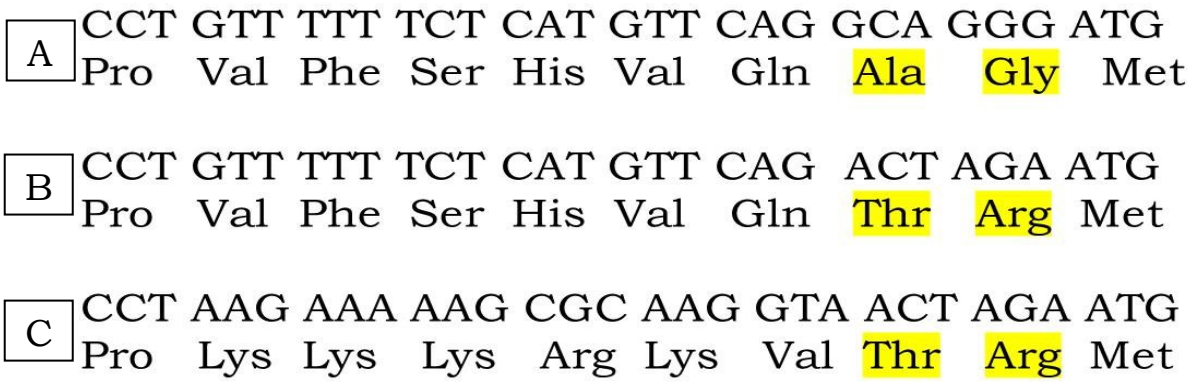


Figure 12: **Illustration of the amino acids 104-113 of the CI repressor and their respective DNA sequence**. The top line of each row represents the DNA sequence and the bottom line the amino acid sequence. A: wild type sequence of CI with a RecA recognition site (yellow marked). B: mutation of the RecA recognition site following the idea of Gimbel and Sauer (1985). C: the 10 aa long yeast codon optimized SV40 virus nuclear localization sequence (www.parts.igem.org/Part:BBa_J63008) integrated in such a way into the sequence, that the RecA recognition site is mutated following the idea of Gimbel and Sauer (1985).

Another hypothesis is that wild type repressor proteins are small enough to diffuse into the nucleus, but need therefore a higher total concentration to reach their suffficient concentration for repression in the nucleus. For this reason the *KEX2* promoter of JOE0044 was exchanged by the strong *PGK1* promoter (JOE0091) and its variation nPos 1+3 (JOE0079) - which next to its diminished expression also could show some fed back abilities preserving cell fitness - as well as for further experiments with *HDA1*$_{prom}$ (JOE0083).

Additionally, some of the plasmids containing the repressor expression block were changed by the method of Omega PCR from a *CEN6/ARSH4* to a 2µm-based plasmid (JOE0044 to JOE0081, see exemplary figure 3 J, JOE0079 to JOE0082, JOE0080 to JOE0083, JOE0046 to JOE0084, JOE0089 to JOE0090

JOE0091 to JOE0092, JOE0050 to JOE0093, JOE0049 to JOE0094, JOE0052 to JOE0095 and JOE0051 to JOE0096).

Table 12: Summary of the plasmids carrying the repressor gene CI with indication of their different changes. After transformation, the protein products of the colored plasmids (yellow) were involved in western blot analyses.

Plasmid	Promoter	CI repressor			Segregation ori
		N-terminal extension	Internal change	C-terminal extension	
pUG24	-	-	-	-	*CEN6/ARSH4*
JOE0044	KEX2	MetGln::ΔMet$_1$	-	-	*CEN6/ARSH4*
JOE0045	KEX2	MetGln::ΔMet$_1$	A111T, G112R[3]	-	*CEN6/ARSH4*
JOE0046	KEX2	MetGln::ΔMet$_1$	NLS[1]	-	*CEN6/ARSH4*
JOE0047	KEX2	MetGln::NLS::ΔMet$_1$	A111T, G112R	-	*CEN6/ARSH4*
JOE0048	KEX2	MetGln::ΔMet$_1$	A111T, G112R[3]	NLS	*CEN6/ARSH4*
JOE0049	KEX2	MetGln::His(8)tag::ΔMet$_1$	NLS[1]	-	*CEN6/ARSH4*
JOE0050	KEX2	MetGln::ΔMet$_1$	NLS[1]	His(8)tag	*CEN6/ARSH4*
JOE0051	HDA1	MetGln::His(8)tag::ΔMet$_1$	NLS[1]	-	*CEN6/ARSH4*
JOE0052	HDA1	MetGln::ΔMet$_1$	NLS[1]	His(8)tag	*CEN6/ARSH4*
JOE0053	TRL1	MetGln::His(8)tag::ΔMet$_1$	NLS[1]	-	*CEN6/ARSH4*
JOE0054	TRL1	MetGln::ΔMet$_1$	NLS[1]	His(8)tag	*CEN6/ARSH4*
JOE0056	HDA1	MetGln::ΔMet$_1$	NLS[1]	-	*CEN6/ARSH4*
JOE0057	TRL1	MetGln::ΔMet$_1$	NLS[1]	-	*CEN6/ARSH4*
JOE0072	TRL1long	MetGln::ΔMet$_1$	NLS[1]	-	*CEN6/ARSH4*
JOE0079	nPos 1+3[2]	MetGln::ΔMet$_1$	-	-	*CEN6/ARSH4*
JOE0080	HDA1	MetGln::ΔMet$_1$	-	-	*CEN6/ARSH4*
JOE0081	KEX2	MetGln::ΔMet$_1$	-	-	2µm-based
JOE0082	nPos 1+3[2]	MetGln::ΔMet$_1$	-	-	2µm-based
JOE0083	HDA1	MetGln::ΔMet$_1$	-	-	2µm-based
JOE0084	KEX2	MetGln::ΔMet$_1$	NLS[1]	-	2µm-based
JOE0089	PGK1	MetGln::ΔMet$_1$	NLS[1]	-	*CEN6/ARSH4*
JOE0090	PGK1	MetGln::ΔMet$_1$	NLS[1]	-	2µm-based
JOE0091	PGK1	MetGln::ΔMet$_1$	-	-	*CEN6/ARSH4*
JOE0092	PGK1	MetGln::ΔMet$_1$	-	-	2µm-based
JOE0093	KEX2	MetGln::ΔMet1	NLS1	His(8)tag	2µm-based
JOE0094	KEX2	MetGln::His(8)tag::ΔMet$_1$	NLS[1]	-	2µm-based
JOE0095	HDA1	MetGln::ΔMet$_1$	NLS[1]	His(8)tag	2µm-based
JOE0096	HDA1	MetGln::His(8)tag::ΔMet$_1$	NLS[1]	-	2µm-based

Legend

1 see figure 12 for details

2 *PGK1* promoter with operator sites at position 1 and 3 (see figure 8 and 11)

3 mutation of the RecA recognition site following the idea of Gimbel and Sauer (1985)

At this point, it has to be mentioned that the 2µm sequence was based on the 2µm ori from yEpGAP-mCherry, which carries four additional bp compared to the wild type 2µm sequence destroying the wild type XbaI site. Table 12 summarizes all plasmid variants used in this study in matters of promoter, mutation and plasmid type, but not functionality, which is only observable in the combination of the plasmids with the yEGFP3 signal expression block under control of the operator improved PGK1 promoter as described in the next paragraphs.

3.6 COMBINATION OF REPRESSOR AND REPORTER GENE

All experiments were conducted three times in test tubes with SD medium as well as overnight cultures and are illustrated similar to figure 11. This figure contains the relative fluorescence test of all used promoter variants without any additional plasmid and can be found for direct comparison as a black box diagrams in every following figure.

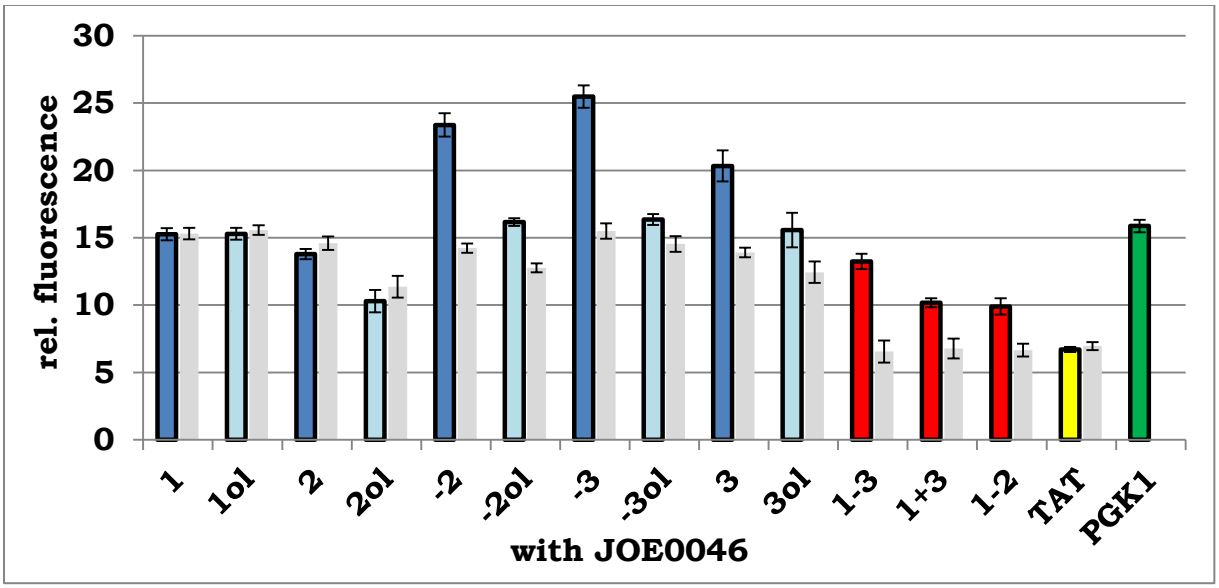


Figure 13: **Relative fluorescence values of cells expressing yEGFP3 under control of *PGK1* promoter variants in combination with JOE0046** (*KEX2*$_{prom}$; CI with internal NLS; *CEN6/ARSH4*). In a threefold repeated experiment, the relative fluorescence (see 2.9.2) of the different combinations grown in selective SD medium was measured (details in table E, Appendix 7.1). The value of the wild type *PGK1* promoter (green) as well as the promoter variants without repressor (grey boxes) are provided as comparison guideline. The constructs with a double insertion of the operator sites (red) the TAT variation (yellow) and constructs with the additional insertion of the O_L fragment (light blue) as well as single constructs without it (dark blue) are marked correspondingly. For the operator positioning and detailed information of nPos 1, 1ol, 2, 2ol, -2, -2ol, -3, -3ol, 3, 3ol, 1-3, 1+3, 1-2 and TAT see table 10 as well as figure 8.

In the first presented experiment the plasmid JOE0046 (*KEX2*$_{prom}$; CI with internal NLS; *CEN6/ARSH4*) is combined with the different promoter variants (figure 13) in cells. Noteworthy is thereby that the variants with C-terminal (JOE0048) or N-terminal (JOE0047) NLS showed either no effect in a combination with the promoter variants nPos 1, 1ol, -3, -3ol (JOE0048) or were not possible to be transformed into *Saccharomyces cerevisiae* (JOE0047).

It can be observed that the repressor CI under these conditions had no significant effect on the relative fluorescence respectively expression of yEGFP3 in combination with nPos 1, 1ol, 2, 2ol and TAT. However, interestingly the repressor had the contrary effect as expected and seemingly increased the expression in combination with nPos -2, -2ol, -3, -3ol, 3, 3ol, 1-3, 1+3 and 1-2 up to 60% in direct comparison of strains without repressor. This effect seemed to be appearing especially in construct variants with no additional O_L site.

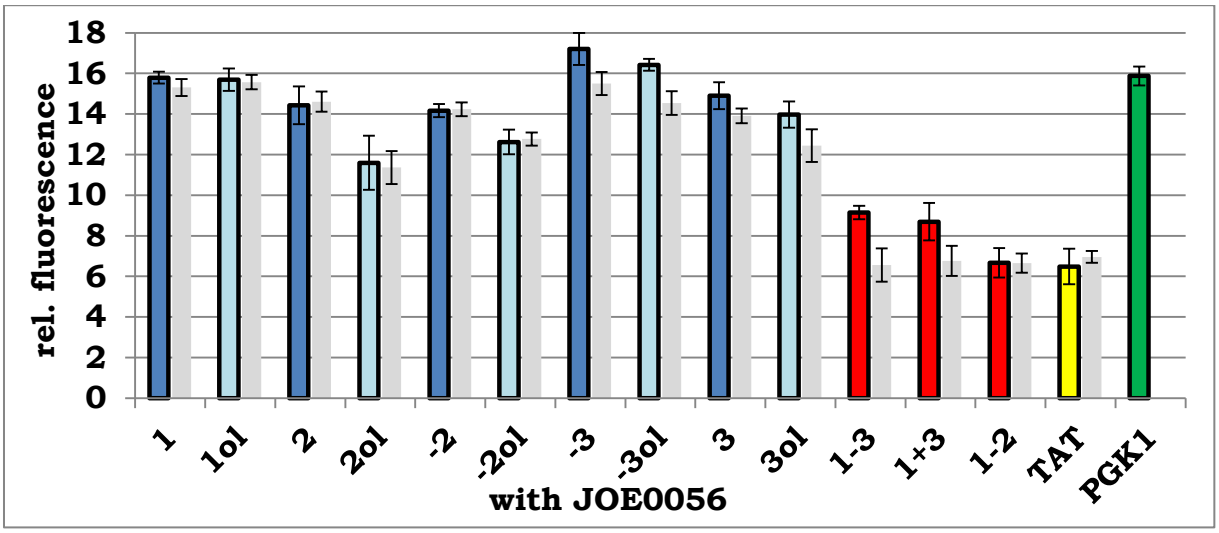


Figure 14: **Relative fluorescence values of cells expressing yEGFP3 under control of *PGK1* promoter variants in combination with JOE0056** (*HDA1*$_{prom}$; CI with internal NLS; *CEN6/ARSH4*). In a threefold repeated experiment, the relative fluorescence (see 2.9.2) of the different combinations grown in selective SD medium was measured (details in table F, Appendix 7.1). The value of the wild type *PGK1* promoter (green) as well as the promoter variants without repressor (grey boxes) are provided as comparison guideline. The constructs with a double insertion of the operator sites (red) the TAT variation (yellow) and constructs with the additional insertion of the O_L fragment (light blue) as well as single constructs without it (dark blue) are marked correspondingly. For the operator positioning and detailed information of nPos 1, 1ol, 2, 2ol, -2, -2ol, -3, -3ol, 3, 3ol, 1-3, 1+3, 1-2 and TAT see table 10 as well as figure 8.

In combination with the plasmid JOE0056 (*HDA1*$_{prom}$; CI with internal NLS; *CEN6/ARSH4*) somehow different results (figure 14) could be detected, indicating an influence of the expression intensity of the repressor. The

outcomes represented in figure 14 showed that in this situation no significant influence on the expression of nearly all tested promoter variants was observed, except for all variants with nPos -3 participation, where again an increase could be noticed. Additionally worth mentioning is, that with *p*-values of around 6 to 10% (see Appendix 7.1 table F for details) also nPos 3 was marginally significant influenced for an increased expression.

A complete divergent result was reached by the combination of JOE0079 (nPos 1+3; CI; *CEN6/ARSH4*) in cells with the different promoter variants (figure 15). With this plasmid in most of the combinations, a percental reduction of the relative fluorescence could be observed (nPos 1, 1ol, -3, -3ol, 1-3, 1+3, 1-2 and TAT), whereupon it was even reaching the lower detection limit for nPos 1, 1ol, 1-3, 1+3, 1-2 and TAT as well as a reduction to 67% (nPos -3) and 77% (nPos -3ol). No influence seemed to be observed for the variants -2, -2ol, 3 as well as 3ol and even an exceptional increase again for the position 2 and 2ol.

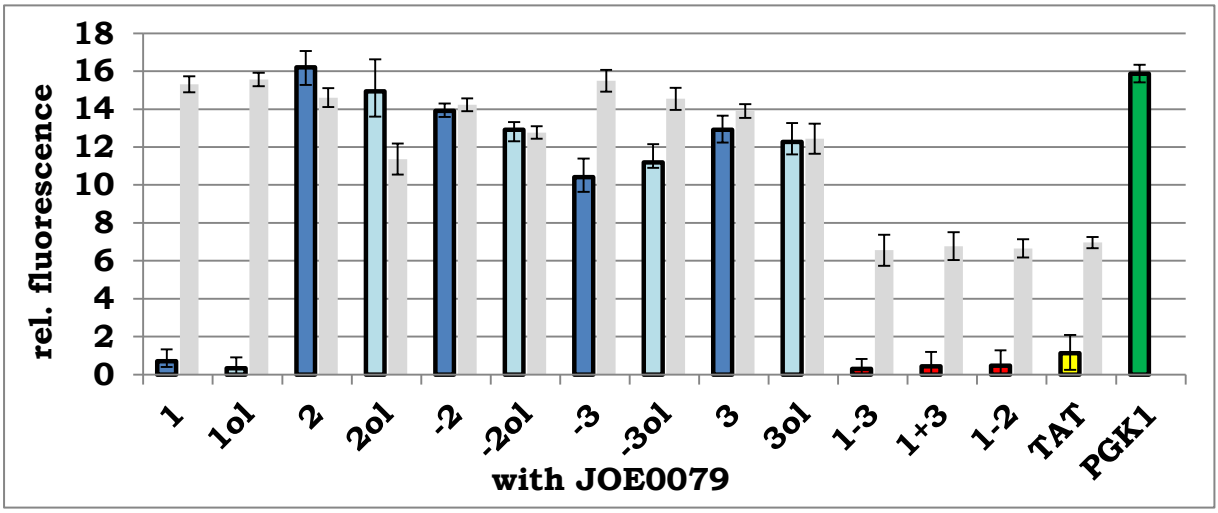


Figure 15: **Relative fluorescence values of cells expressing yEGFP3 under control of *PGK1* promoter variants in combination with JOE0079** (nPos 1+3; CI; *CEN6/ARSH4*). In a threefold repeated experiment, the relative fluorescence (see 2.9.2) of the different combinations grown in selective SD medium was measured (details in table G, Appendix 7.1). The value of the wild type *PGK1* promoter (green) as well as the promoter variants without repressor (grey boxes) are provided as comparison guideline. The constructs with a double insertion of the operator sites (red) the TAT variation (yellow) and constructs with the additional insertion of the O_L fragment (light blue) as well as single constructs without it (dark blue) are marked correspondingly. For the operator positioning and detailed information of nPos 1, 1ol, 2, 2ol, -2, -2ol, -3, -3ol, 3, 3ol, 1-3, 1+3, 1-2 and TAT see table 10 as well as figure 8.

In experiments with JOE0044 (*KEX2*$_{prom}$; CI; *CEN6/ARSH4*, see figure 16) as well as JOE0080 (*HDA1*$_{prom}$; CI; *CEN6/ARSH4*, see figure 17) no effect of the repressor was visible. This indicated a possibly too low repressor concentration.

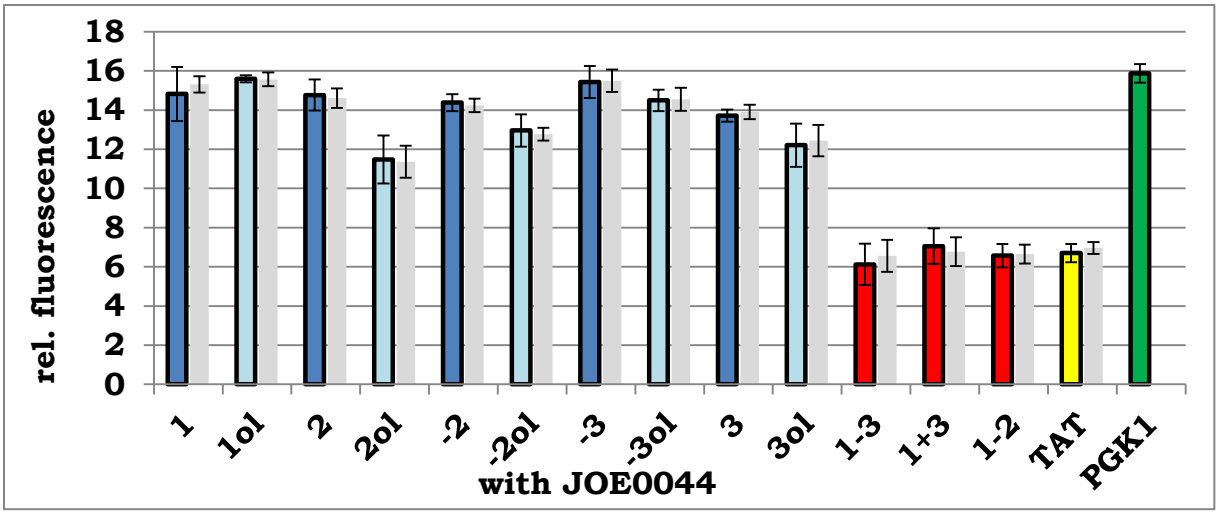


Figure 16: **Relative fluorescence values of cells expressing yEGFP3 under control of *PGK1* promoter variants in combination with JOE0044** (*KEX2*$_{prom}$; CI; CEN6/ARSH4). In a threefold repeated experiment, the relative fluorescence (see 2.9.2) of the different combinations grown in selective SD medium was measured (details in table H, Appendix 7.1). The value of the wild type *PGK1* promoter (green) as well as the promoter variants without repressor (grey boxes) are provided as comparison guideline. The constructs with a double insertion of the operator sites (red) the TAT variation (yellow) and constructs with the additional insertion of the O_L fragment (light blue) as well as single constructs without it (dark blue) are marked correspondingly. For the operator positioning and detailed information of nPos 1, 1ol, 2, 2ol, -2, -2ol, -3, -3ol, 3, 3ol, 1-3, 1+3, 1-2 and TAT see table 10 as well as figure 8.

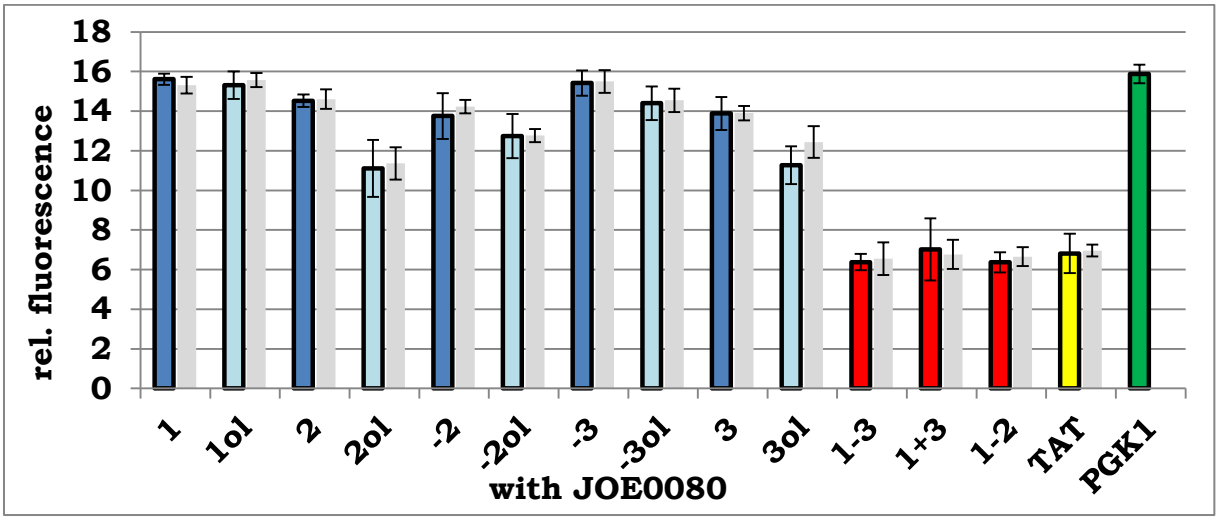


Figure 17: **Relative fluorescence values of cells expressing yEGFP3 under control of *PGK1* promoter variants in combination with JOE0080** (HDA1prom; CI with internal NLS; CEN6/ARSH4). In a threefold repeated experiment, the relative fluorescence (see 2.9.2) of the different combinations grown in selective SD medium was measured (details in table I, Appendix 7.1). The value of the wild type *PGK1* promoter (green) as well as the promoter variants without repressor (grey boxes) are provided as comparison guideline. The constructs with a double insertion of the operator sites (red) the TAT variation (yellow) and constructs with the additional insertion of the O_L fragment (light blue) as well as single constructs without it (dark blue) are marked correspondingly. For the operator positioning and detailed information of nPos 1, 1ol, 2, 2ol, -2, -2ol, -3, -3ol, 3, 3ol, 1-3, 1+3, 1-2 and TAT see table 10 as well as figure 8.

The plasmid JOE0057 (*TRL1*$_{prom}$; CI with internal NLS; *CEN6/ARSH4*, see figure 18) and JOE0072 (*TRL1*$_{prom}$ long; CI with internal NLS; *CEN6/ARSH4*, see figure 19) in combination with different respectively selected promoter variants showed similar results, whereupon assumingly in both cases the repressor concentration is too weak for a reaction.

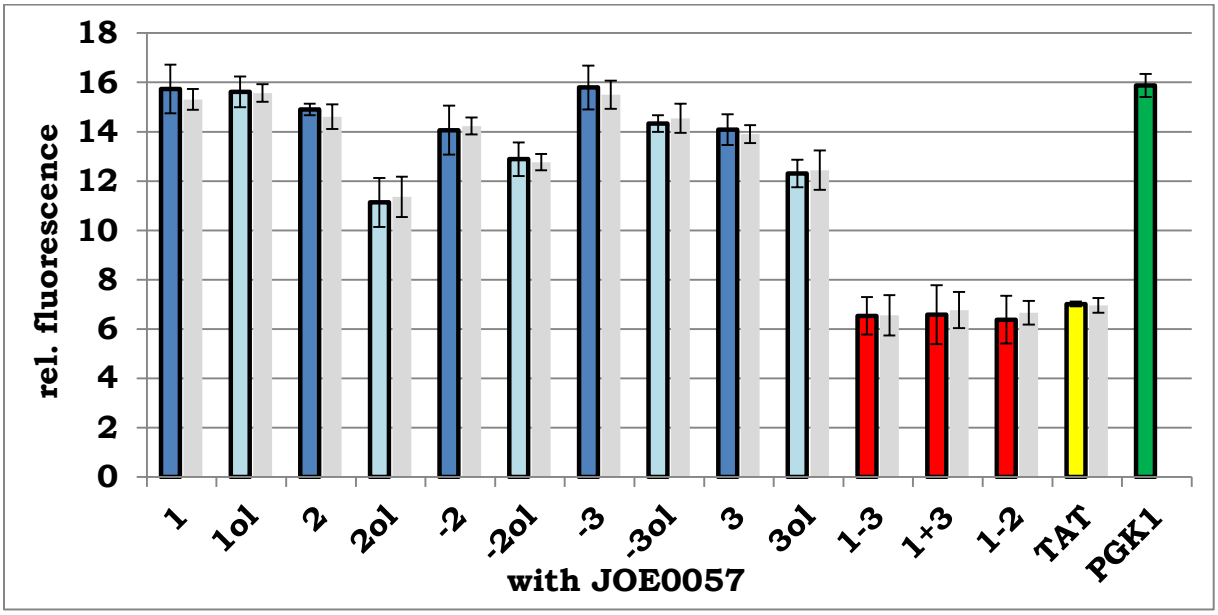


Figure 18: **Relative fluorescence values of cells expressing yEGFP3 under control of *PGK1* promoter variants in combination** with JOE0057 (*TRL1*$_{prom}$; CI with internal NLS; *CEN6/ARSH4*). In a threefold repeated experiment, the relative fluorescence (see 2.9.2) of the different combinations grown in selective SD medium was measured (details in table O, Appendix 7.1). The value of the wild type *PGK1* promoter (green) as well as the promoter variants without repressor (grey boxes) are provided as comparison guideline. The constructs with a double insertion of the operator sites (red) the TAT variation (yellow) and constructs with the additional insertion of the O_L fragment (light blue) as well as single constructs without it (dark blue) are marked correspondingly. For the operator positioning and detailed information of nPos 1, 1ol, 2, 2ol, -2, -2ol, -3, -3ol, 3, 3ol, 1-3, 1+3, 1-2 and TAT see table 10 as well as figure 8.

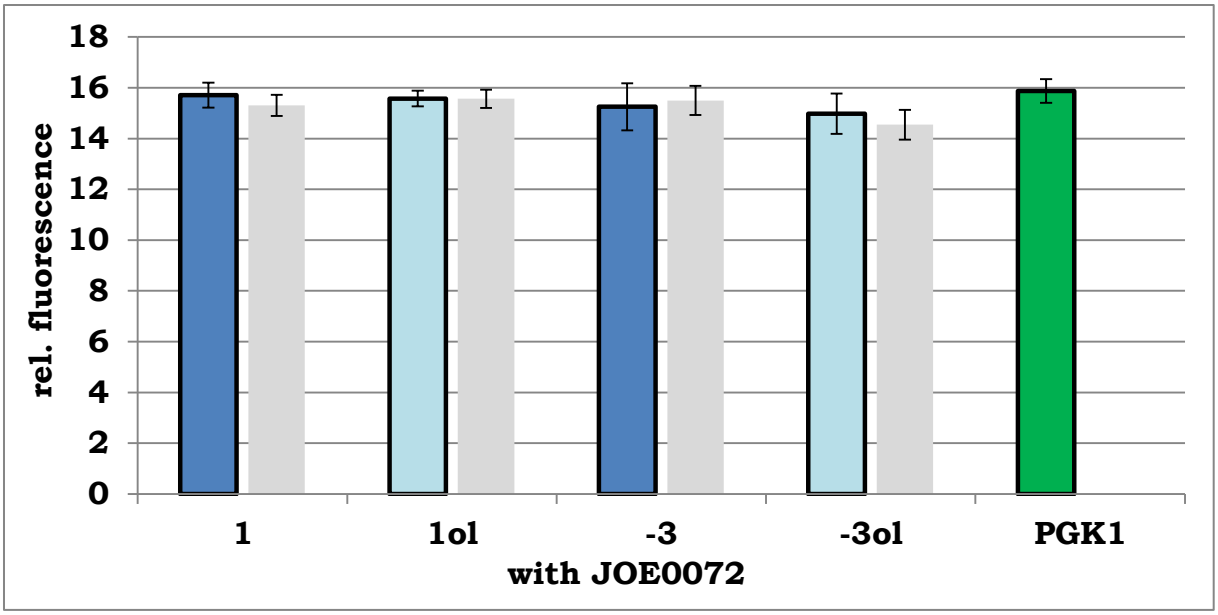


Figure 19: **Relative fluorescence values of cells expressing yEGFP3 under control of *PGK1* promoter variants in combination with JOE0072** (*TRL1*$_{prom}$long; CI with internal NLS; *CEN6/ARSH4*). In a threefold repeated experiment, the relative fluorescence (see 2.9.2) of the different combinations grown in selective SD medium was measured (details in table P, Appendix 7.1). The value of the wild type *PGK1* promoter (green) as well as the promoter variants without repressor (grey boxes) are provided as comparison guideline. The constructs with a double insertion of the operator sites (red) the TAT variation (yellow) and constructs with the additional insertion of the O_L fragment (light blue) as well as single constructs without it (dark blue) are marked correspondingly. For the operator positioning and detailed information of nPos 1, 1ol, -3 and -3ol see table 10 as well as figure 8.

Summarizing the interim results of *CEN6/ARSH4* based plasmids with either CI or CI with internal NLS under the control of different promoters in combination with the different genome inserted *PGK1* promoter variants, led to several preliminary assumptions: Some promoter variants seemed to be influenced more by the repressor than others and a presumed high CI concentration was able to decrease the expression of nPos 1, 1ol, 1+3, 1-3, 1+2, (see figure 15) whereas CI with internal NLS seemed to increase the expression of nPos -2, -2ol, -3, -3ol, 3, 3ol, 1-3, 1+3 and 1-2 (see figure 13 and 14). For this reason, the test of 2µm-based plasmids with either CI or CI with internal NLS will only be based on some selected construct variants, which already showed a significant change respectively interference.

The experiment presented first is therefore based on the plasmid JOE0092 (*PGK1*$_{prom}$; CI; 2µm-based) and it showed (see figure 20) that with seemingly high concentration of CI nearly all tested promoter variants except for 2ol show

a decrease in expression. The highest decrease could be observed for nPos 1 and 1ol - there even with no variance in the measurement - making both seemingly an optimal basis for the detection system.

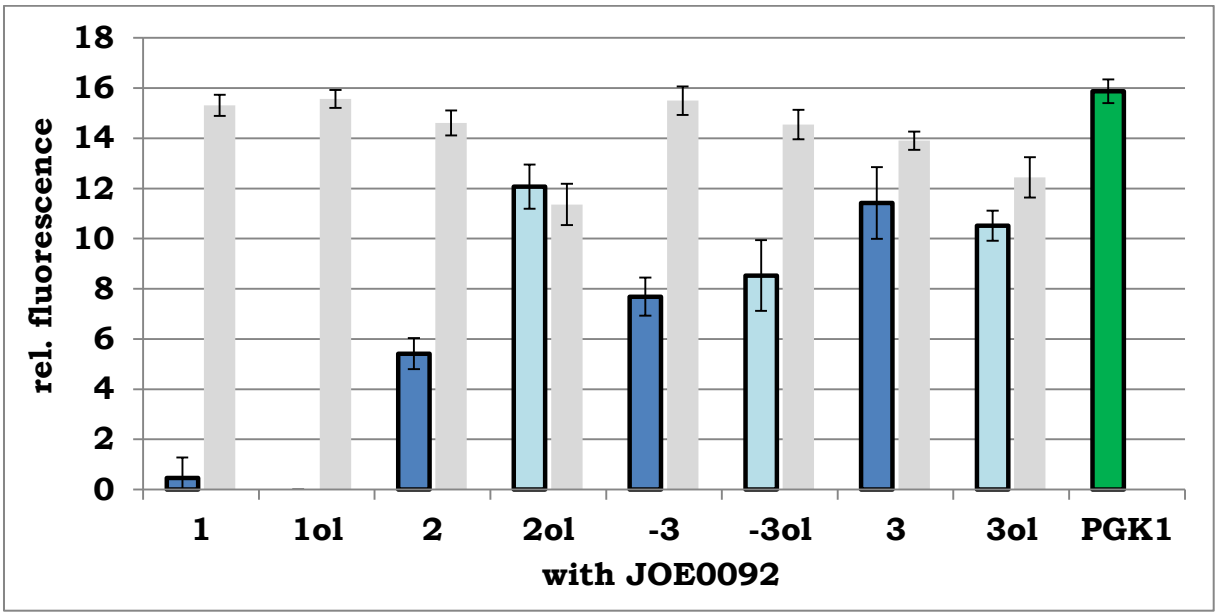


Figure 20: **Relative fluorescence values of cells expressing yEGFP3 under control of *PGK1* promoter variants in combination** with JOE0092 (*PGK1*$_{prom}$; CI; 2µm-based). In a threefold repeated experiment, the relative fluorescence (see 2.9.2) of the different combinations grown in selective SD medium was measured (details in table J, Appendix 7.1). The value of the wild type *PGK1* promoter (green) as well as the promoter variants without repressor (grey boxes) are provided as comparison guideline. The constructs with a double insertion of the operator sites (red) the TAT variation (yellow) and constructs with the additional insertion of the O_L fragment (light blue) as well as single constructs without it (dark blue) are marked correspondingly. For the operator positioning and detailed information of nPos 1, 1ol, 2, 2ol, -3, -3ol, 3 and 3ol see table 10 as well as figure 8.

Again the dependency on the concentration seemed to be visible in the experiments with JOE0081 (*KEX2*$_{prom}$; CI; 2µm-based, figure 21) and JOE0083 (*HDA1*$_{prom}$; CI; 2µm-based, figure 22).

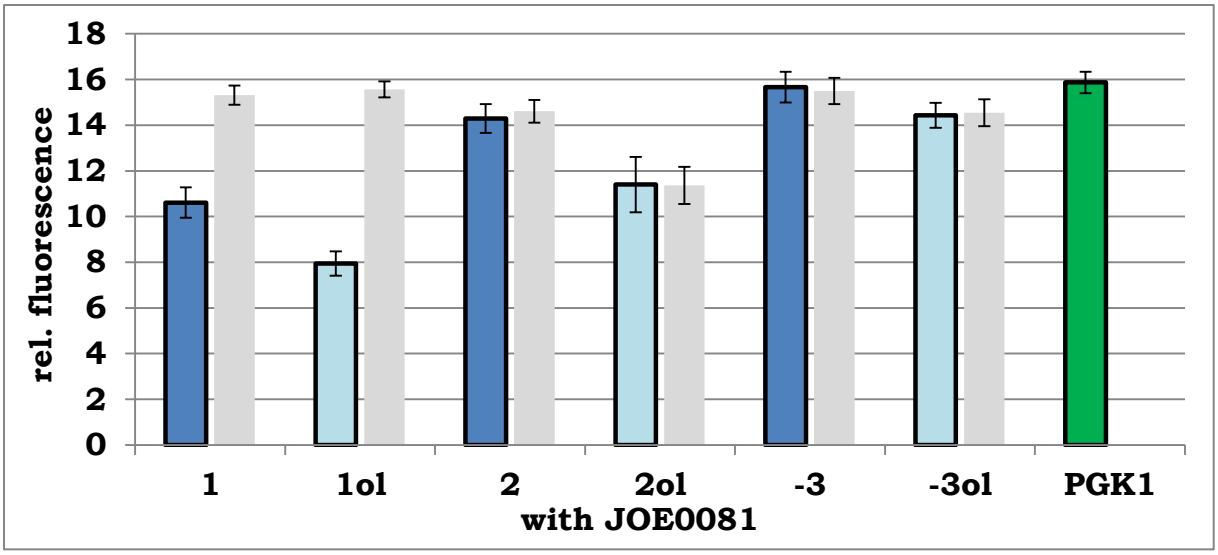


Figure 21: **Relative fluorescence values of cells expressing yEGFP3 under control of *PGK1* promoter variants in combination with JOE0081** (*KEX2*$_{prom}$; CI; 2µm-based). In a threefold repeated experiment, the relative fluorescence (see 2.9.2) of the different combinations grown in selective SD medium was measured (details in table K, Appendix 7.1). The value of the wild type *PGK1* promoter (green) as well as the promoter variants without repressor (grey boxes) are provided as comparison guideline. The constructs with a double insertion of the operator sites (red) the TAT variation (yellow) and constructs with the additional insertion of the O_L fragment (light blue) as well as single constructs without it (dark blue) are marked correspondingly. For the operator positioning and detailed information of nPos 1, 1ol, 2, 2ol, -3 and -3ol see table 10 as well as figure 8.

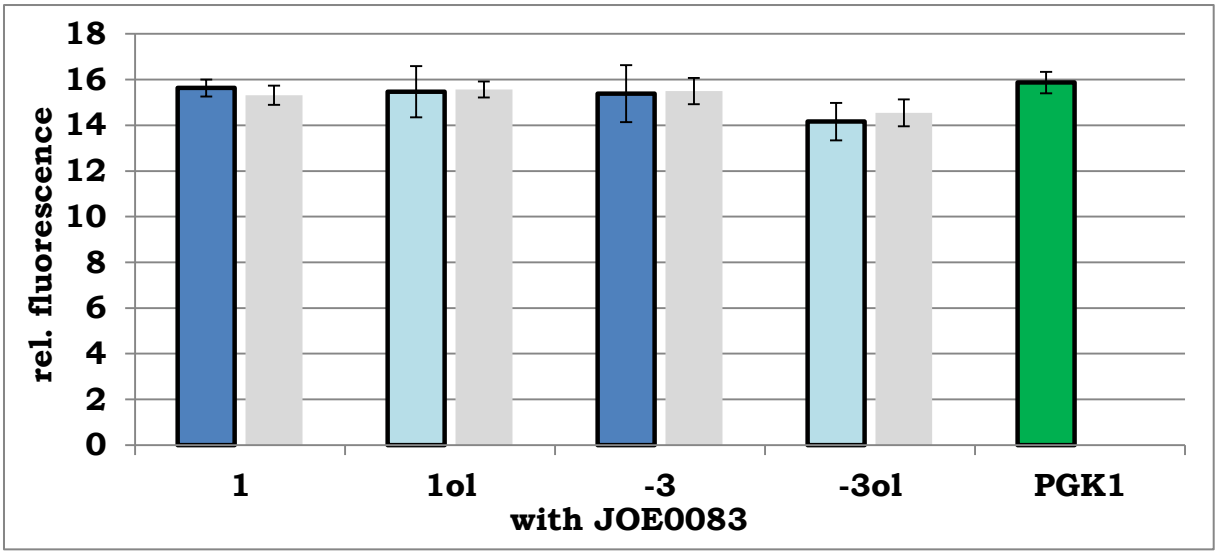


Figure 22: **Relative fluorescence values of cells expressing yEGFP3 under control of *PGK1* promoter variants in combination with JOE0083** (*HDA1*$_{prom}$; CI; 2µm-based). In a threefold repeated experiment, the relative fluorescence (see 2.9.2) of the different yeast strains grown in selective SD medium was measured (details in table M, Appendix 7.1). The value of the wild type *PGK1* promoter (green) as well as the promoter variants without repressor (grey boxes) are provided as comparison guideline. The constructs with a double insertion of the operator sites (red) the TAT variation (yellow) and constructs with the additional insertion of the O_L fragment (light blue) as well as single constructs without it (dark blue) are marked correspondingly. For the operator positioning and detailed information of nPos 1, 1ol, -3 and -3ol see table 10 as well as figure 8.

Only with the plasmid JOE0081 in combination with nPos 1 or 1ol a reduction of the expression intensity of up to 50% could be recognized. In all other combinations, no influence was observed, making nPos 1 and 1ol look more accessible for the repressor, indicating an influence of the promoter respectively chroamatin structure.

In the last mentioned combinations the influence of JOE0084 (*KEX2*$_{prom}$; CI with internal NLS; 2µm-based, figure 23) and JOE0090 (*PGK1*$_{prom}$; CI with internal NLS; 2µm-based, figure 24) was tested. In both cases a similar result could be observed, where even in high concentrations of CI with internal NLS no influence was visible for the nPos 1, 1ol, 2 and 2ol but for the positions nPos -2, -2ol, -3, -3ol, 3 and 3ol. There, an increase of expression intensity respectively relative fluorescence between 10 and 40% could be detected. The prominent result was thereby achieved by the position -3 in combination with the assumed highest concentration of CI with internal NLS by JOE0090.

Before all these results will be summarized, it is interesting to notice how the combination of repressor integration and plasmid based repressor expression influenced the mean growth rate of the explored strains.

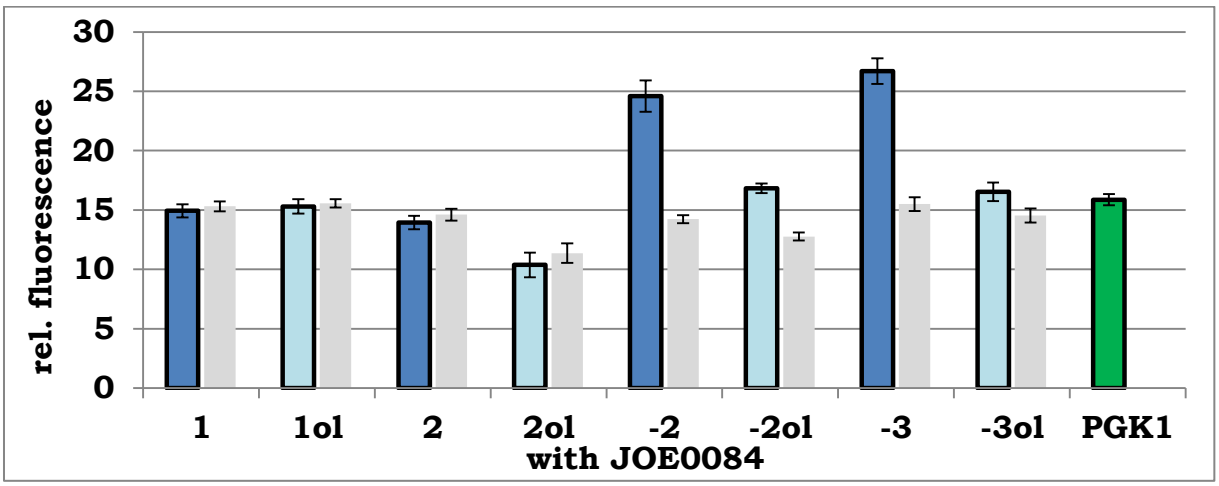


Figure 23: **Relative fluorescence values of cells expressing yEGFP3 under control of *PGK1* promoter variants in combination with JOE0084** (*KEX2*$_{prom}$; CI with internal NLS; 2µm-based). In a threefold repeated experiment, the relative fluorescence (see 2.9.2) of the different combinations grown in selective SD medium was measured (details in table L, Appendix 7.1). The value of the wild type *PGK1* promoter (green) as well as the promoter variants without repressor (grey boxes) are provided as comparison guideline. The constructs with a double insertion of the operator sites (red) the TAT variation (yellow) and constructs with the additional insertion of the O_L fragment (light blue) as well as single constructs without it (dark blue) are marked correspondingly. For the operator positioning and detailed information of nPos 1, 1ol, 2, 2ol, -2, -2ol, -3 and -3ol see table 10 as well as figure 8.

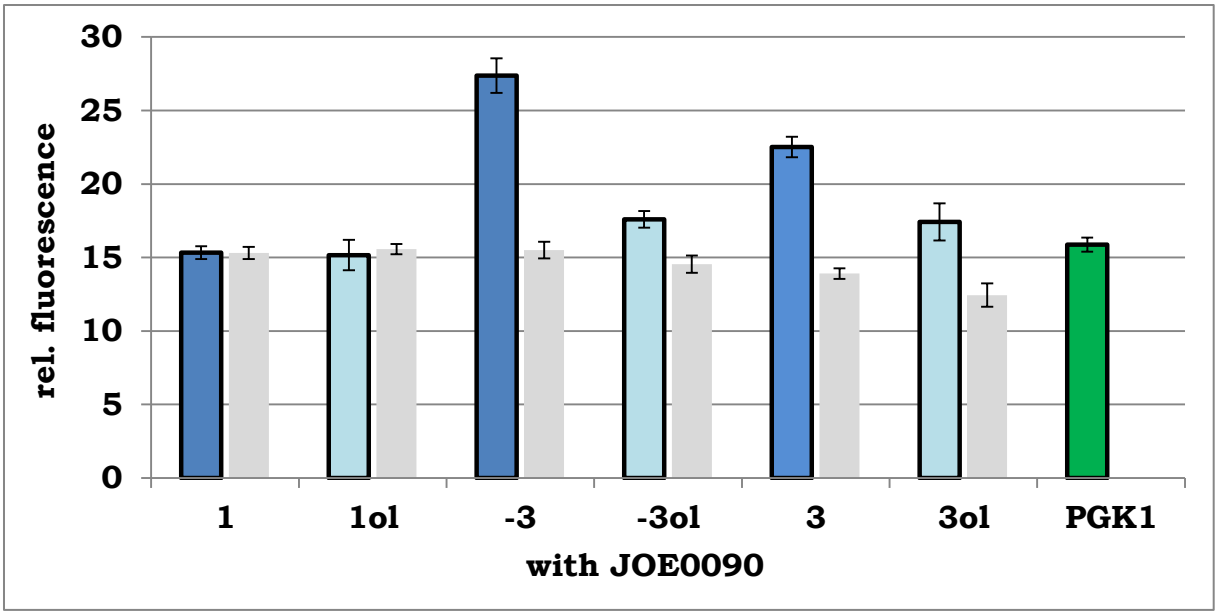


Figure 24: **Relative fluorescence values of cells expressing yEGFP3 under control of *PGK1* promoter variants in combination with JOE0090** (*PGK1*$_{prom}$; CI with internal NLS; 2µm-based). In a threefold repeated experiment, the relative fluorescence (see 2.9.2) of the different combinations grown in selective SD medium was measured (details in table N, Appendix 7.1). The value of the wild type *PGK1* promoter (green) as well as the promoter variants without repressor (grey boxes) are provided as comparison guideline. The constructs with a double insertion of the operator sites (red) the TAT variation (yellow) and constructs with the additional insertion of the O_L fragment (light blue) as well as single constructs without it (dark blue) are marked correspondingly. For the operator positioning and detailed information of nPos 1, 1ol, -3, -3ol, 3 and 3ol see table 10 as well as figure 8.

A twice performed growth experiment of selected strains (nPos 1 and -3) in combination with certain plasmids (JOE0046, JOE0084, JOE0079 and JOE0092) conducted in SD medium is consolidated in table 13 (details in table R, Appendix 7.2).

Table 13: List of the mean growth rate per hour in SD medium and mean relative fluorescence of different promoter variants combined with certain plasmids.The strain SY992 without plasmid and with JOE0031 was used as reference (details in table R, Appendix 7.2).

combination	mean growth rate [h^{-1}]		mean rel. flurorescence
SY992	0.254	0.247	0
JOE0031	0.230	0.235	15.87
1+JOE0046	0.213	0.214	15.25
1+JOE0084	0.258	0.221	14.94
1+JOE0079	0.232	0.241	0.62
1+JOE0092	0.144	0.216	0.47
-3+JOE0046	0.204	0.204	25.49
-3+JOE0084	0.217	0.173	26.7
-3+JOE0079	0.186	0.214	10.42
-3+JOE0092	0.182	0.192	7.69

It could be observed that on the one hand the growth of all tested strains was diminished in comparison to a growth experiment in YPDAU (see table 7). On the other hand, now several factors like additional DNA as well as foreign protein expression influenced the fitness of the cell leading to a complex background for some results. The benchmarks in this experiment were the results of SY922 (0.25 h^{-1}) and SY922 with the plasmid JOE0031 (0.23 h^{-1}), where no tested combination reached the values of SY992 alone. Combinations with relatively high respectively unchanged fluorescence (1+JOE0046, 1+JOE0084 and -3+JOE0079) showed a nearly similar mean growth rate between 0.21 and 0.22 h^{-1}, which seemed to be reduced to around 0.20 h^{-1} in combinations with higher expression rates like in -3+JOE0046 and -3+JOE0084. An influence of the copy number respectively the difference between 2µm-based multicopy and CEN6/ARSH4 single copy plasmids was not as clear as it could be expected. The same applied to the influence of highly reduced expression by the repressor (1+JOE0079 and 1+JOE0092), assumingly both being based on the missing statistical repetition.

Overall, the results can be grouped into two combinations with two specific outcomes: The first one is the combination with wild type CI repressor, which led in most promoter variants (1, 1ol, 2, -3, -3ol, 3, 3ol, 1-3, 1+3, 1-2, TAT, see figure 15, 20 and 21) except -2, -2ol and 2ol to a significant repression. The second one is the combination with CI repressors containing an internal NLS site, which either have no effect (1, 1ol, 2, 2ol, TAT, see figure 18, 19, 23 and 24) on the expression or even increased (-2, -2ol, -3, -3ol, 3, 3ol, 1-3, 1+3, 1-2 see figure 18, 19, 23 and 24) it. Both combinations thereby were dependent on the correct expression intensity of the repressor, whereupon the activity seems to be increased with growing concentration. CI with internal NLS seemingly showed an effect on lower concentrations as CI without internal NLS, yet this could also depend on the promoter variants, which seem to differ in their accessibility for the repressors. Interestingly to mention is also that in most constructs with a double insertion of the operator sites (nPos 1-3, 1+3, 1-2) an increase as well as a decrease is inducible depending on the used repressor.

The most important results for the development of an online reporter system to detect plasmid stability are certainly the combinations with a high respectively

total repression of the reporter protein expression. Therefore, the combinations nPos 1 or 1ol with the plasmids JOE0079 or JOE0092 will be explored further regarding their applicability in a detection system in the next paragraphs.

3.7 Detection of plasmid loss with the reporter system

For detection of plasmid loss with the help of the established reporter and repressor combinations, it was necessary to conduct on the one hand a plasmid stability test respectively control as well as on the other hand a measurement of the relative fluorescence at the same control point. Based on the requirements for these experiments, the direct timely correlation was not possible because of the lengthy plasmid stability test (see 2.7). Through this fact, the presented results are only connected relatively by the respective control point as well as blurred through the additional non-selective media incubation step in the plasmid stability test, which increases the number of plasmid free cells. Keeping this in mind, the results of the initial experiments with the combination of nPos 1 (figure 25) and JOE0079 (nPos 1+3; CI; *CEN6/ARSH4, HIS*) growing in SD-medium without uracil (SD-URA) or selective SD-medium without uracil and histidin (SD-URA-HIS) will be presented next.

Looking particularly at the percentage of plasmid containing cells, it is interesting to recognize that seemingly even in a selective medium – here SD-URA-HIS - not all of the cells carried a plasmid. Indeed the level of plasmid containing cells changed between 84% and 89% in the control points, while the influence of the lengthy method is not clear and the achieved results are not only therefore not certain. In contrast to this, a continuous plasmid loss in the media without uracil seemed to be possible, which manifest assumingly in the slowly but steady decreasing percentage of plasmid containing cells from around 81% to 72% over time. The measurements for the relative fluorescence, as well as the sampling for the plasmid loss assay, were thereby conducted in the intervals of 8, 24 and 30 hours. In contrast to the previously presented relative fluorescence data of the combination of cells containing nPos 1ol with JOE0079 a value of around 10.7 in SD-URA-HIS could be measured in the first control point (6 h). However, after 24 hours no fluorescence was observed, yet

rising again after that (30 h) to 10.6. Regarding the SD-URA medium, a consistent value between 10.4 to 11.6 could be detected.

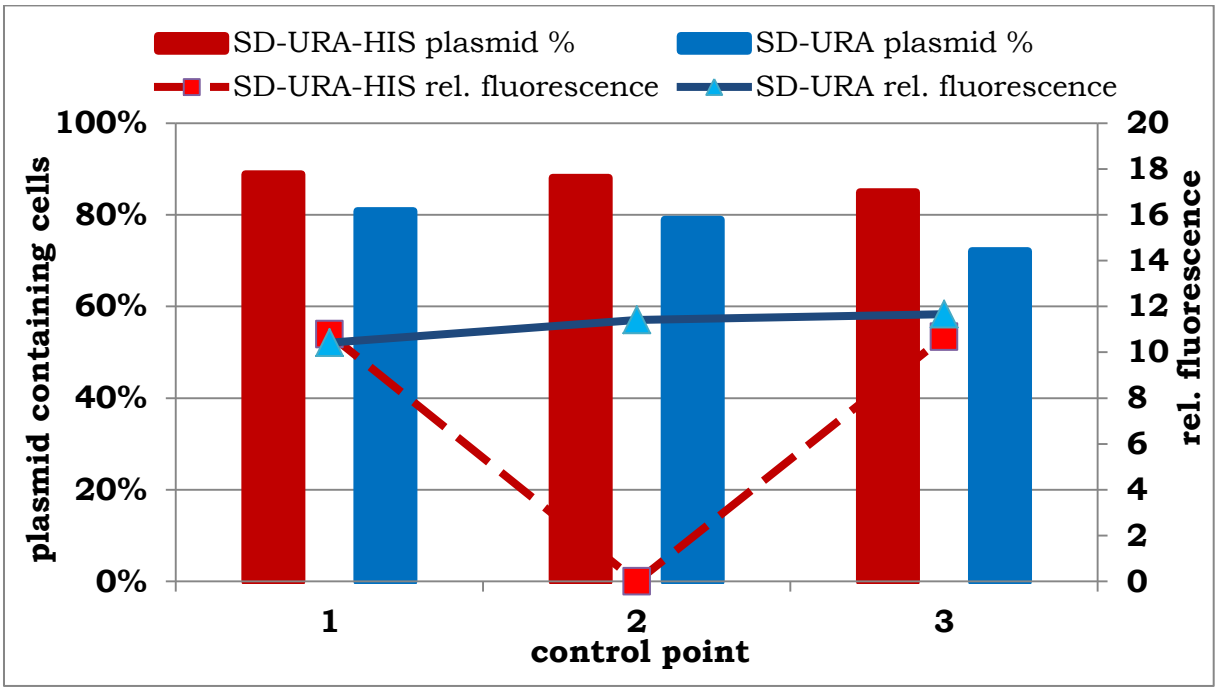


Figure 25: **Illustration of the plasmid loss and corresponding relative fluorescence of cells with plasmid JOE0079** (nPos 1+3; CI; *CEN6/ARSH4*) in **combination with nPos 1**. The filled bars represent the percentage of plasmid containing cells (left axis) of the mentioned combination in selective SD-URA-HIS (red) or unselective SD-URA medium (blue) at the relative control points and the dotted lines stand for the relative fluorescence (right axis, see 2.9.2 for method) of the same culture at the control points 1 (6 h), 2 (24 h) and 3 (30 h); details in Appendix 7.3, table S. The colony number on a control YPDAU plate with around 50-200 cfu defines the maximal amount of plasmid bearing cells respectively 100% (see 2.7 for method).

The combination of nPos 1ol with JOE0079 (nPos 1+3; CI; *CEN6/ARSH4*) produced nearly identical results as it can be seen in figure 26.

The percentage of plasmid containing cells oscillated between 84% and 88% in selective medium (SD-URA-HIS) and declined over time from 82% to 70% in a non-selective medium (SD-URA). Similar to cells containing the combination with 1ol and JOE0079, relative fluorescence of around 11.4 (SD-URA-HIS) and 12.9 (SD-URA) could be measured at the first control point (6 h), which declined for both to either 0 (SD-URA-HIS) or 9.8 (SD-URA) at 24 h and raised again to 11.8 (SD-URA) and 13.5 (SD-URA) at the last control point.

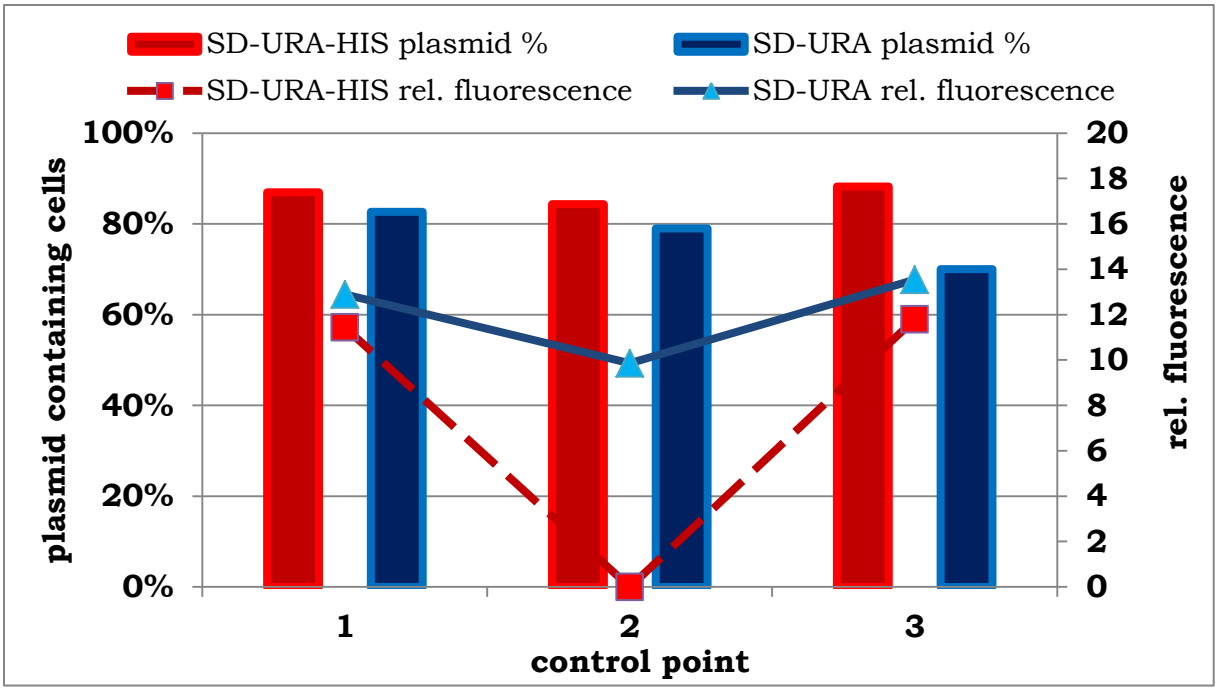


Figure 26: **Illustration of the plasmid loss and corresponding relative fluorescence of cells with plasmid JOE0079** (nPos 1+3; CI; *CEN6/ARSH4*) **in combination with nPos 1ol**. The filled bars represent the percentage of plasmid containing cells (left axis) of the mentioned combination in SD-URA-HIS (red) or SD-URA medium (blue) at the relative control points and the dotted lines stand for the relative fluorescence (right axis, see 2.9.2 for method) of the same culture at the control points 1 (6 h), 2 (24 h) and 3 (30 h); details in Appendix 7.3, table T. The colony number on a control YPDAU plate with around 50-200 cfu defines the maximal amount of plasmid bearing cells respectively 100% (see 2.7 for method).

Thus, no strong difference between nPos 1 and 1ol could be observed and the effect of the additional operator (O_L) site respectively loop forming in matters of system stability was not obvious. Nevertheless, perhaps it was visible in the comparatively lower relative fluorescence at control point 2 (24 h) in a SD-URA medium. The detected fluorescence by itself especially in the first control points in SD-URA-HIS as well as the continuous relative high signal in SD-URA medium was surprising. This could be triggered either by the promoter feed-back mechanisms in JOE0079, which seemingly decreases the promoter activity or by low repressor concentration, both arguments for experiments with the wild type $PGK1_{prom}$ as well as no further tests of a feed-back mechanism.

Additionally, the relatively high stability of *CEN6/ARSH4* plasmids decreased the amplitude of possible signal outcome over time and demanded an extreme length of the experiments about plasmid loss. Moreover, such single copy plasmids are in the industry mostly substituted by chromosomal integration.

In contrast to that, the multi copy 2µm-based plasmids are widely used in the industry despite their instability and are so the main aspirant for a detection system for plasmid stability. This was also the reason for concentrating on experiments with 2µm-based plasmids in the following.

In a nearly similar composition to the already mentioned experiment, but with a slight adjournment of the control points, the combinations nPos 1 (figure 27) as well as nPos 1ol (figure 28) were this time each tested with JOE0092 (*PGK1*$_{prom}$; CI; 2µm-based) again growing in a non-selective (SD-URA) or selective medium (SD-URA-HIS).

Once more, the difference between nPos 1 and 1ol was not evident and not all cells carried a plasmid although growing in selective media. In both experiments the percentage of plasmid bearing cells in SD-URA-HIS varied between 91% and 83% with an outlier of 70% over the measurement points and time. The loss rate was increased especially in comparison to *CEN6/ARSH4* based plasmids, which is visible through a constant and high decline from around 85% / 72% to 47% / 37% in either nPos 1/1ol.

The results of the relative fluorescence measurement are interesting to notice for cells with both combinations. In the control experiments with selective medium, no fluorescence was observable over time (6 h, 30 h and 54 h, see figure 27 and 28), indicating an assumingly high plasmid retention in selective conditions. Whereas in the experiments with non-selective medium, a relative fluorescence value of around 7.1 (nPos 1) respectively 10.1 (nPos 1ol) was measurable at the last control point (54 h). It can be assumed that this fluorescence is connected to the plasmid loss in both combinations and that therefore a plasmid loss of around 60% could be sensed through the developed detection system.

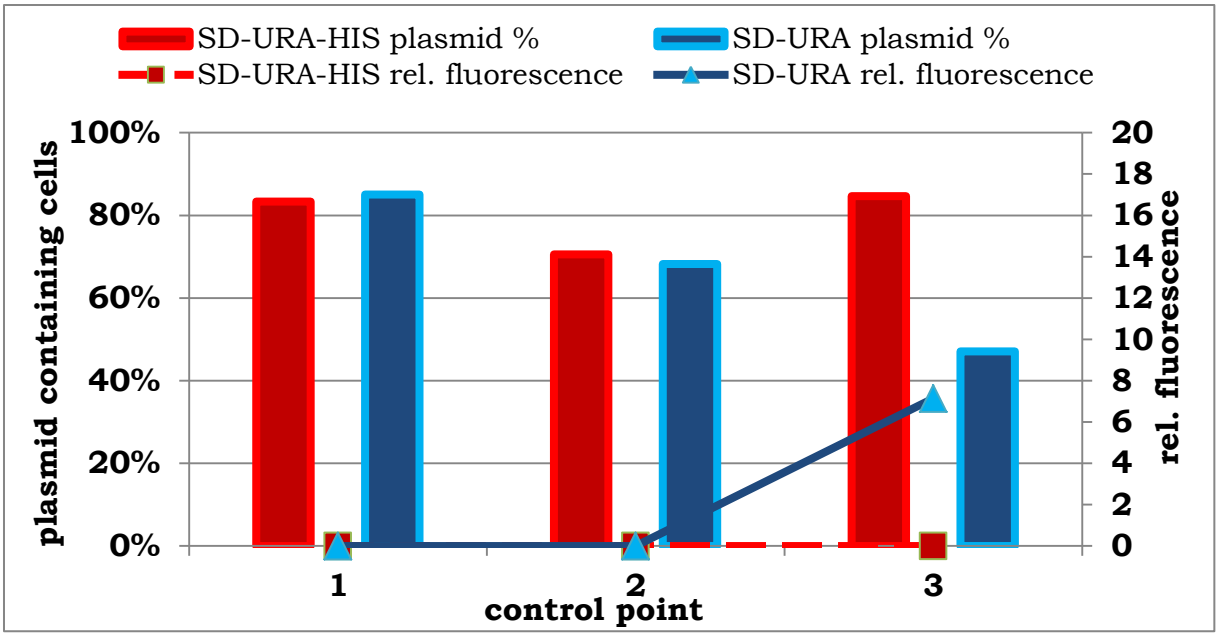


Figure 27: **Illustration of the plasmid loss and corresponding relative fluorescence of cells with cells with plasmid JOE0092** (*PGK1*$_{prom}$; CI; 2µm-based) **in combination with nPos 1**. The filled bars represent the percentage of plasmid containing cells (left axis) of the mentioned combination in SD-URA-HIS (red) or SD-URA medium (blue) at the relative control points and the dotted bars stand for the relative fluorescence (right axis, see 2.9.2 for method) of the same culture at the control points 1 (6 h), 2 (30 h) and 3 (54 h); details in Appendix 7.3, table V. The colony number on a control YPDAU plate with around 50-200 cfu defines the maximal amount of plasmid bearing cells respectively 100% (see 2.7 for method).

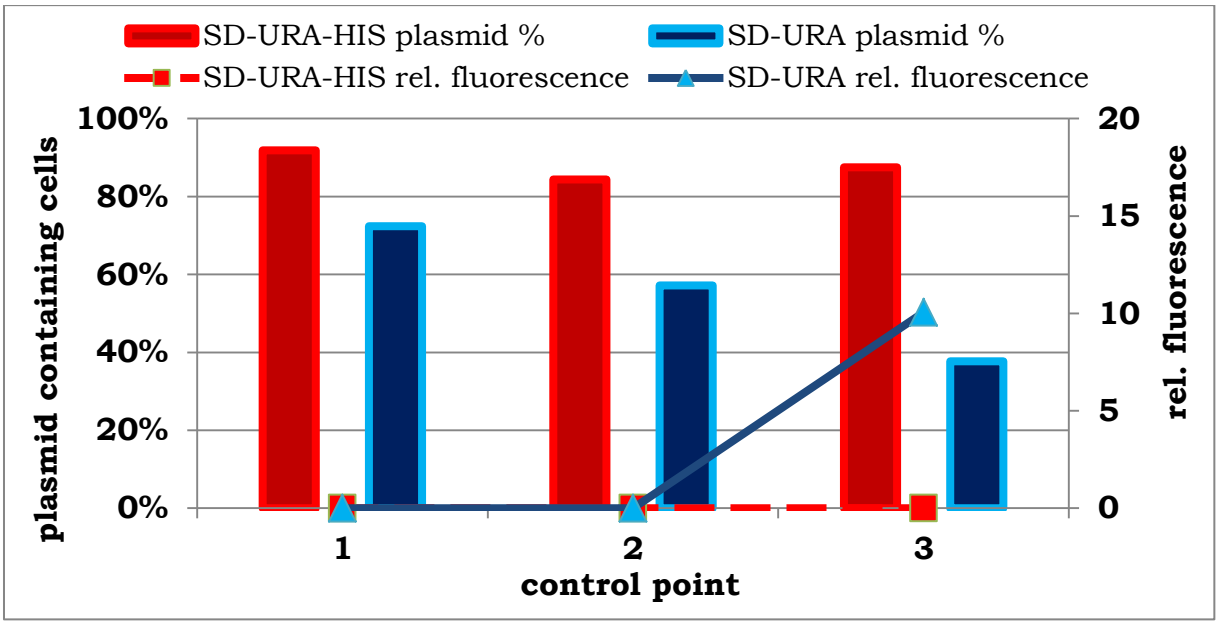


Figure 28: **Illustration of the plasmid loss and corresponding relative fluorescence of cells with plasmid JOE0092** (*PGK1*$_{prom}$; CI; 2µm-based) **in combination with nPos 1ol**. The filled bars represent the percentage of plasmid containing cells (left axis) of the mentioned combination in SD-URA-HIS (red) or SD-URA medium (blue) at the relative control points and the dotted lines stand for the relative fluorescence (right axis, see 2.9.2 for method) of the same culture at the control points 1 (6 h), 2 (30 h) and 3 (54 h); details in Appendix 7.3, table W. The colony number on a control YPDAU plate with around 50-200 cfu defines the maximal amount of plasmid bearing cells respectively 100% (see 2.7 for method).

Due to a seemingly higher relative fluorescence value in the last test as well as in particular the missing variance in the measurement presented in figure 20, the combination of 1ol with JOE0092 will be used exclusively in the following experiments. These experiments aimed to improve on the one hand the detection limit and on the other hand the plasmid loss rate for results that are more significant.

In a first step, the influence of the full medium YPDAU was tested, which is closer to the situation in large-scale experiments and also connected to an assumingly higher plasmid loss than SD-URA medium. It is of course not comparable in all intricacies to experiments in SD medium. Therefore, a test setting was established in order to compare all three media (YPDAU, SD-URA and SD-URA-HIS) with the measurement points for the relative fluorescence being at 6, 24 and 36 hours again.

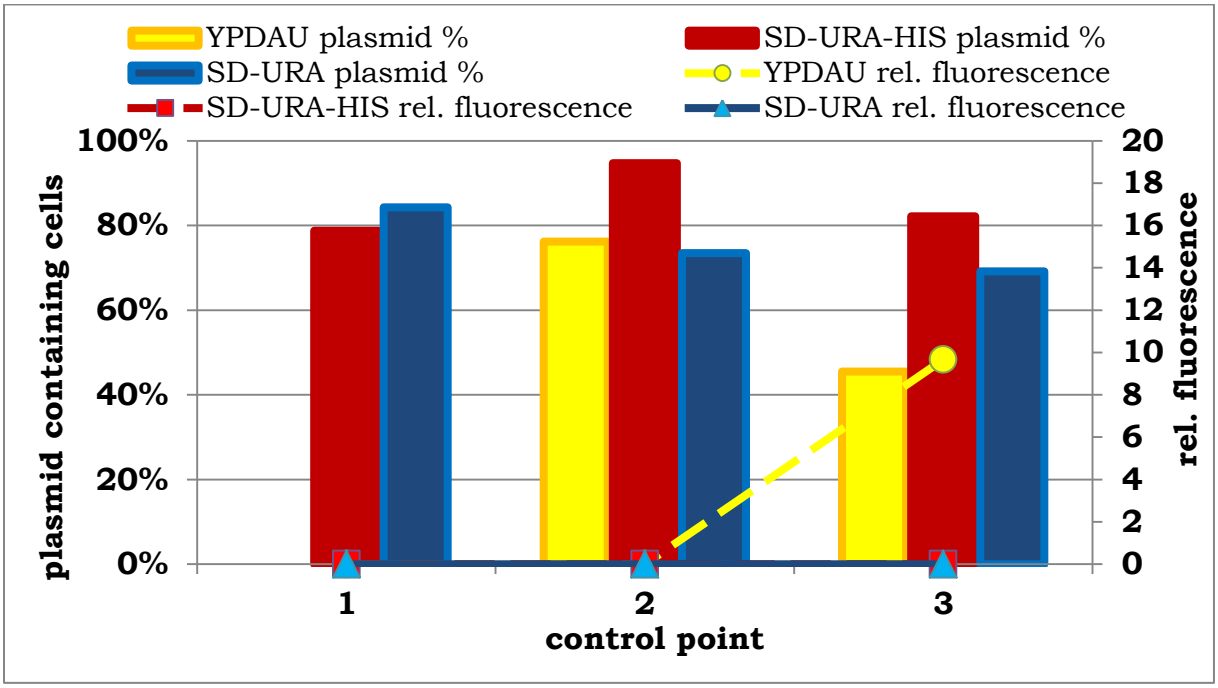


Figure 29: **Illustration of the plasmid loss and corresponding relative fluorescence of cells with plasmid JOE0092** (*PGK1*$_{prom}$; CI; 2µm-based) **in combination with nPos 1ol**. The filled bars represent the percentage of plasmid containing cells (left axis) of the mentioned combination in SD-URA-HIS (red) or SD-URA (blue) or YPDAU medium (yellow) at the relative control points and the dotted lines stand for the relative fluorescence (right axis, see 2.9.2 for method) of the same culture at the control points 1 (6 h), 2 (24 h) and 3 (36 h); details in Appendix 7.3, table U. The colony number on a control YPDAU plate with around 50-200 cfu defines the maximal amount of plasmid bearing cells respectively 100% (see 2.7 for method).

In figure 29, the results are illustrated and the experiment demonstrated that the combination of nPos 1ol with JOE0092 (*PGK1*$_{prom}$; CI; 2µm-based) indeed showed in YPDAU a higher plasmid loss than observed at the same control point for SD-URA medium. While the percentage of plasmid containing cells was decreased only from 84% to 63% in SD-URA and stayed nearly constant in SD-URA-HIS it was decreased in YPDAU from 76% to 45%. The measurement of the relative fluorescence however showed a similar outcome as in the previous test. Only in the late phase of the culture growth respectively experiment, a value of 9.6 was measurable, again connected to a plasmid loss of over 50%. This indicated that the detection limit for the fluorescence measurement seemed to be not sufficient enough to detect early plasmid loss events. However, this would be needed to counteract these events appropriately.

An increase of the detection limit was assumed to be achieved by decreasing the measurement volume of the fluorescence test from 2 ml for the fluorescence photometer to 13 µl using the "Varian Pro Star 363 Fluorescence Detector". However, before further experiments were conducted, the system had been tested for its ability to detect fluorescence with a high efficiency. Therefore, unmodified SY992 cells were mixed in different proportions of the same OD_{600} with SY992 cells containing the plasmid JOE0031 and hence emitting a fluorescence signal with both being overnight cultures.

Table 14: Results of the fluorescence detector test (see 2.9.3). Dilution 1:10 stands for 1 cell SY992+JOE0031 among 10 cells SY922 without plasmid and so on. Adjustment was performed by OD_{600} measurement (Details in Appendix 7.3, table X).

	mean rel. fluorescence	dilution
SD	0	
SY992	8.00	
Mix 1	83.69	1:10
Mix 2	66.02	1:50
Mix 3	27.78	1:100
Mix 4	22.35	1:250
Mix 5	14.31	1:500
Mix 6	14.95	1:750
Mix 7	9.66	1:1000
Mix 8	8.17	1:1500
SY992	8.82	
SD	0	

The results of the test are summarized in table 14 (Details in Appendix 7.3, table X): It can be seen that assumingly at least around one cell under 800-1000 cells must fully express yEGFP3 to be detected by the established system. Additionally, it has been observed that the unmodified SY922 cells compared to the used medium already had a fluorescent background, which served as a threshold for the experiment. The experimental installation with the help of a "p-1" pump and tubes seemingly led despite a long washing step on the one hand to a mixture of the samples as well as a concealment of possible correlation. On the other hand, it led to an increased fluorescence of unmodified SY922 at the end of the experiment, changing the threshold. Despite its drawbacks, the method was used for a possible improved detection of fluorescence in two experiments (see figure 30 and 31) about plasmid loss with the combination nPos 1ol and JOE0092 (*PGK1*$_{prom}$; CI; 2µm-based).

Three consistently-refreshed cultures in SD-URA, SD-URA–HIS and YPDAU were observed over time at four control points (6 h, 24 h, 30 h and 48 h), whereupon on each control point a sufficient amount of cells was withdrawn to create at the end an around 30 ml cell suspension of OD_{600}=0.3 for the measurement of the relative fluorescence. Additionally, the mentioned plasmid stability test was conducted with similar results to the previous experiments.

The percentage of plasmid containing cells varied in SD-URA-HIS medium in both experiments around 80% and 92 % and in the non-selective medium of YPDAU and SD-URA a decrease was observable at all control points. Thus the percentage in SD-URA drops from 77/89% to 62/55% as well as in YPDAU from 86/88% to 54/50%.

Of particular interest is the development of the fluorescence measurements over time. It could be seen that in the selective control medium the mean relative fluorescence in both experiments was almost constantly between 9 and 10 at all control points thus building the threshold for the recognition of increased fluorescence. Some smaller outliers could be detected at the first measurement point of the first test and at the third control point of the second experiment.

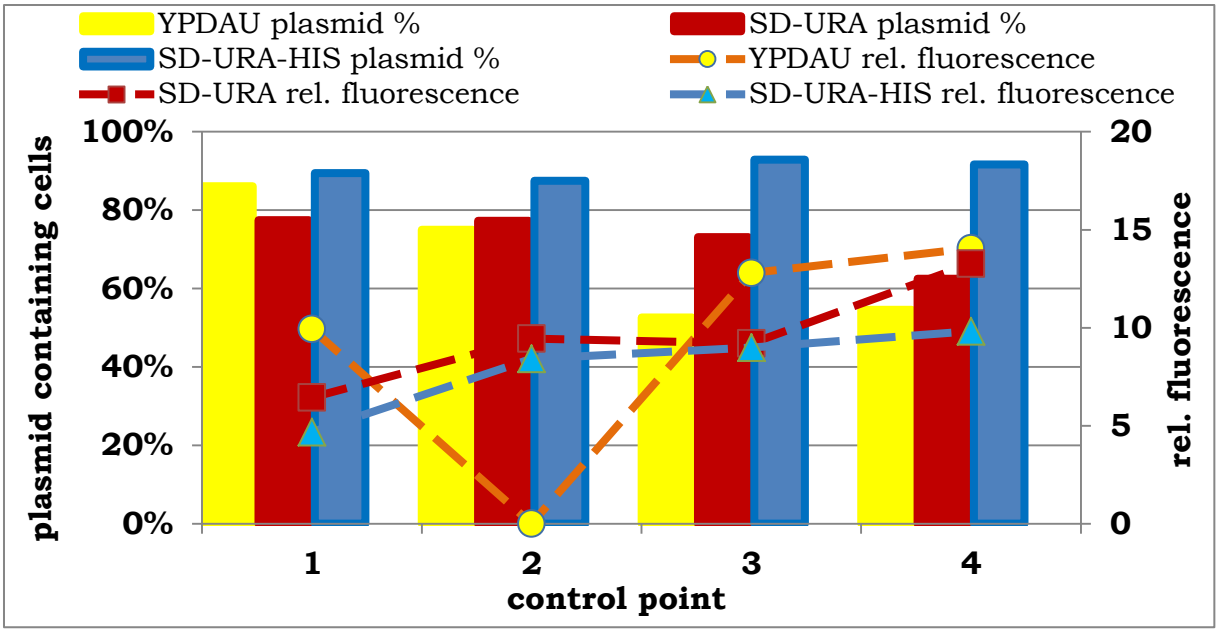


Figure 30: **Illustration of the plasmid loss and corresponding relative fluorescence of cells with plasmid JOE0092** (*PGK1*$_{prom}$; CI; 2µm-based) in **combination with nPos 1ol**. The bars represent the percentage of plasmid containing cells (left axis) of the mentioned combination in SD-URA-HIS (red) or SD-URA (blue) or YPDAU medium (yellow) at the relative control points. The dotted lines stand for the relative fluorescence (right axis, see 2.9.3 for method) of the same culture at the control points 1 (6 h), 2 (24 h), 3 (30 h) and 4 (48 h); details in Appendix 7.3, table Y. The colony number on a control YPDAU plate with around 50-200 cfu defines the maximal amount of plasmid bearing cells respectively 100% (see 2.7 for method).

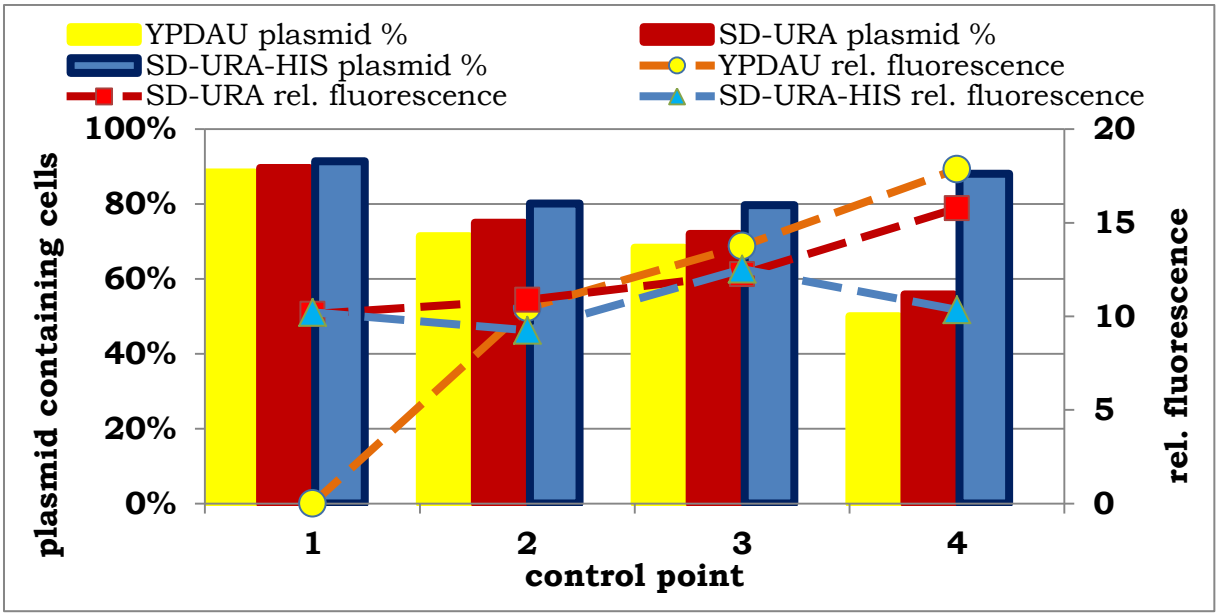


Figure 31: **Illustration of the plasmid loss and corresponding relative fluorescence of cells with plasmid JOE0092** (*PGK1*$_{prom}$; CI; 2µm-based) **in combination with nPos 1ol.** The bars represent the percentage of plasmid containing cells (left axis) of the mentioned combination in SD-URA-HIS (red) or SD-URA (blue) or YPDAU medium (yellow) at the relative control points. The dotted lines stand for the relative fluorescence (right axis, see 2.9.3 for method) of the same culture at the control points 1 (6 h), 2 (24 h), 3 (30 h) and 4 (48 h); details in Appendix 7.3, table Z. The colony number on a control YPDAU plate with around 50-200 cfu defines the maximal amount of plasmid bearing cells respectively 100% (see 2.7 for method).

In the first experiment (figure 30) the threshold was crossed with a mean relative fluorescence of 12.8 in the third control point with the culture in YPDAU, having at this moment only 50% of its plasmid left. With further test time, the fluorescence of the culture in the YPDAU medium has increased to 14.0 and the culture in SD-URA medium also passed the threshold with a relative fluorescence of 13.3, but this time seemingly with already around more than 30% of plasmid losses.

The assumption was confirmed by the second experiment (figure 31), where in measurement point 3 the culture in YPDAU medium significantly crossed the even in comparison to in this situation higher threshold, with a value of around 13.8 and a plasmid loss of around 30%. In the last control point, both cultures in the non-selective media showed a higher relative fluorescence than the control culture, while the plasmid loss was then already at over 50%.

For this reason, it could be assumed and concluded that with the used method and the combination of nPos 1ol and plasmid JOE0092 it might be possible to detect plasmid stability online, as soon as the percentage of plasmid carrying cells is fewer than 70%.

In the last experiment the plasmid loss, respectively the increased expression of yEGFP3 was not only followed by fluorescence measurement but also by microscopic observations at specific control points. The results of these observations are shown in figure 32, where in the top row the situation at 24 h and in the bottom row the situation at 48 h are presented, either with full fluorescence view or with a mixture of bright field and fluorescence view.

In the picture at 24 hours a single cell achieved the full fluorescence intensity, whereas the other cells remarkably also show a faint fluorescent expression, which was not visible in the mixed view. This could be a hint that these cells started either to lose their plasmids, which comes together with an increase in yEGFP3 expression or that this kind of background fluorescence is the normal expression noise. In fresh cells of the combination nPos 1ol with JOE0092, also cells like the mentioned ones could be found as well as cells without any visible fluorescence. Additionally, it could be observed that the originating daughter cell expresses the same fluorescent intensity as its mother cell.

At 48 hours, the picture changed in the way that now more and more cells expressed full fluorescence intensity, which goes hand in hand with the measured increased relative fluorescence at this time point. Nevertheless, still living respectively dividing cells could be observed, which showed no fluorescence signal at all and thus bear assumingly the full amount of plasmid necessary to interrupt the YEGFP3 expression.

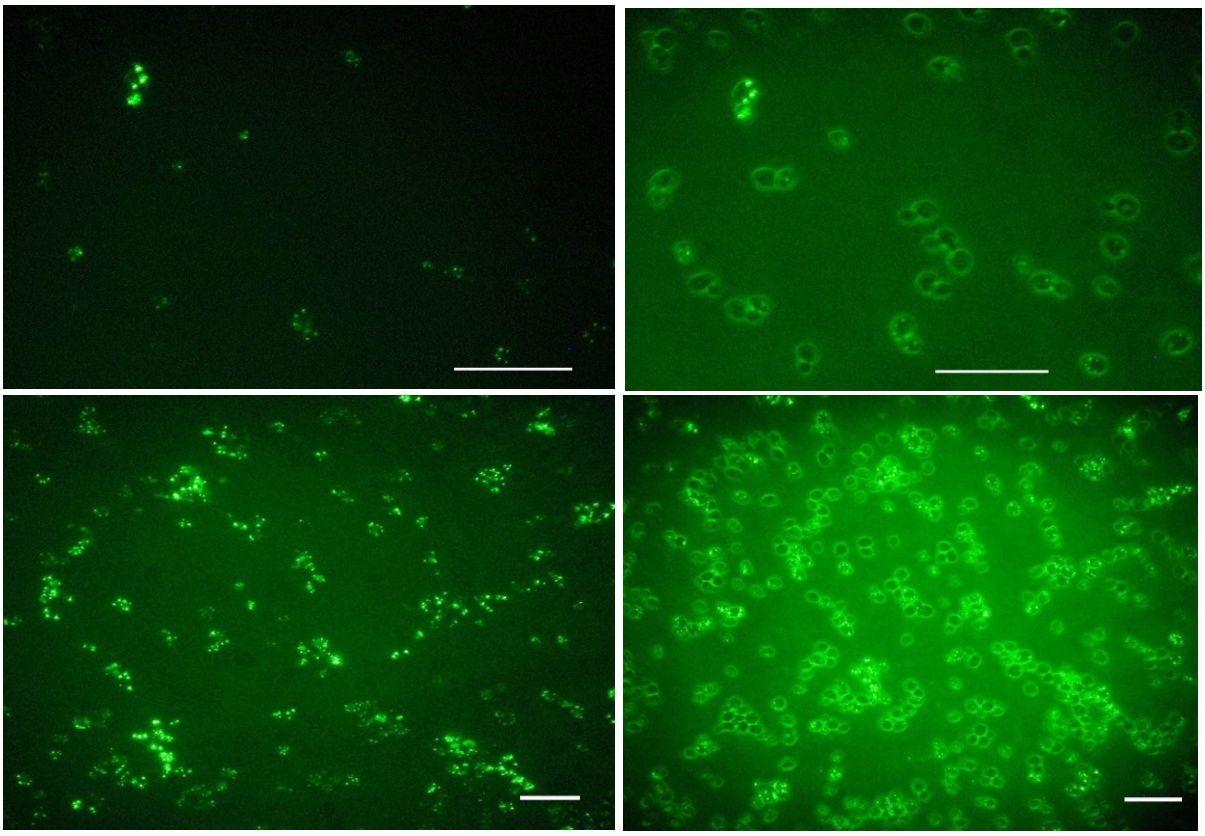


Figure 32: **Observation of the plasmid loss and corresponding fluorescence increase of strain nPos 1ol in combination with plasmid JOE0092** (*PGK1*$_{prom}$; CI; 2µm-based) **over time in YPDAU** (top site: 24 h; bottom site: 48h). On the left side fluorescence and on the right side a mixture between bright field and fluorescence view is visible. Conducted with the axiovert 25, white bar is equivalent to 20 µm with a 100-fold respectively 40-fold magnification.

3.8 SUMMARY OF THE RESULTS

The following overall summary of the achieved results covers briefly the several topics studied in this work. Afterwards, all results will be discussed more intensively in the next paragraph.

The basis for all obtained results was built by the development of a stable and highly expressed reporter gene, which was found out to be yEGFP3 under the

control of the *PGK1* promoter. This reporter additionally seems to have its highest and seemingly most environmental independent signal if it is localized in the peroxisome.

Following these results, the promoter was improved by insertion respectively insertion with connected deletion of different variants of lambda phage based operator sites, which were inserted in three different positions within the promoter and as one fragment outside of the expression cassette. There it could be recognized that a singular insertion of up to 48 bp led in most cases to a highly active *PGK1* promoter, whereas a double insertion of operator fragments led to a decrease in the promoter activity of around 50%.

Interestingly, the insertion of the expression cassettes into a specific genome location did not interfere with the activity of the promoter, making further experiments possible. With this, two variants of the lambda phage repressor – one containing a internal NLS as well as a RecA recognition site mutation and the other the wild type protein sequence – were tested under different expression intensities in combination with the genome integrated expression cassettes variants.

In these experiments, the combination with a highly expressed wild type CI repressor led in most promoter variants to a decreased expression of yEGFP3 whereas in case of the combination with CI repressors containing a internal NLS site, either no effect on the expression or even an increase was observed. The activity of both repressor variants seemed to increase with increasing concentration and CI with internal NLS even seemingly showed an effect on lower concentrations as CI without internal NLS. However, this could also depend on the promoter variants, which seemed to differ in their accessibly for the repressors. Noteworthy is also that in constructs with a specific double insertion of the operator sites an increase as well as a decrease was inducible depending on the applied repressor.

The most important results are of course the combinations with a high, respectively total repression of the reporter protein expression, which were able to detect a plasmid loss of 50% or 30% depending on the detection method.

4 DISCUSSION

On the way to develop an online reporter system to follow plasmid stability in *Saccharomyces cerevisiae* different results were obtained, which cover a wide and differentiated field of biological themes. For this reason, the results achieved in this study will be discussed in the following paragraphs separately and alongside of the result layout, beginning with the used reporter gene respectively protein.

4.1 SELECTION OF THE REPORTER GENE FOR THE STUDY

In this study, mainly the yEGFP3 was used as a reporter, because it showed the optimal requirements for the detection system with regard to sensitivity and size as well as other requested conditions. However, it has to be kept in mind that a variety of factors can influence fluorescence. Among the topics addressed by Tsien (1998) the color tagged will be discussed in detail:

- Autofluorescence of cells or culture media at preferred wavelengths
- Availability of chaperones
- Competition with noise and background signals
- Dimerization
- Efficiency of posttranslational fluorophore formation
- Efficiency of translation including Kozak sequence and codon usage
- Extinction coefficient and fluorescence quantum yield
- Hindrance to folding because of unfortunate fusions to host proteins
- Intrinsic rate of cyclization/oxidation
- Location of GFP
- Molecular properties of mature GFP
- Number of copies of gene, duration of expression
- Quality of excitation and emission filters and dichroic mirrors
- Sensitivity, noise, and dark current of photo detector
- Solubility versus formation of inclusion bodies
- Strength of transcriptional promoters and enhancers
- Susceptibility to photoisomerization/bleaching
- Time, temperature, availability of O_2
- Wavelengths of excitation and emission

4.1.1 DESCRIPTION OF THE USED PROMOTER

The promoter used in this study was the *PGK1* promoter, whose product can reach in wild type cells up to 2-3% of the cell protein (van der Aar *et al.* 1989). Nevertheless, it is often observed, that with heterologous proteins only values between 1-2%, or even only between 0.25-0.5% are reached (Webster and Brammer 1995). An explanation for this could be provided by the experiments of Romanos *et al.* (1992), where it was noticed, that an additional so-called downstream activation sequence (DAS) localized to the first 79 codons influences the expression of the *PGK1* promoter. Consequently, the expression intensity is seemingly depending on the similarity of the heterologous DNA sequence to this DAS.

Another factor influencing the intensity of the *PGK1* promoter is the limiting amount of translation factors. Especially Gcr1p is formed at extremely low levels (Görgens *et al.* 2000). As long as only very few extra copies of the promoter are added additionally - like the genome integration of the reporter expression cassette in this study - no influence of the low transcription factor concentrations can be observed. This seemingly changes if the promoter is based on a plasmid particular a 2µm-based one. In this case, Görgens *et al.* (2000) observed that on the one hand the expression intensity and on the other hand, the mean growth rate is decreasing, which could be confirmed in this study especially through experiments with the repressor CI and the followed mean growth rate decline (see table 13).

The attempt to decrease the transcriptional burden of the cell and therefore possibly increase the mean cell growth through the control of the constant expression of the repressor gene CI by a feedback mechanism was seemingly not successful in this way. Assumingly the low concentration of repressor either through the too weak or the in the feedback mechanism blocked promoter could explain why an oscillation of fluorescence was observed (see figure 25 and 26) instead of a continuous blocking of expression. In this experimental setting mostly in the late phase of growth, a highly repression of

promoter activity was observed because of the accumulated concentration of repressor proteins until then.

Interesting to mention is that the sequence of the promoter *PGK1* used in this study is different in at least two positions to the sequence obtained from the reference strain S288C. In the position -4 bp a C is missing and in position -93 bp a G is exchanged by a T (see Appendix 7.4). These changes might have been introduced by the amplification or cloning process as well as less likely by the used resource for yeast sequences, the wild type strain from Meyen ex E.C. Hansen, which could differ in these positions from the sequence reference strain S288C. Thereby it is not clear what kind of influence these changes have, for example, a start codon adjournment would be possible, thus adding supplementary Met and Gly in front of the yEGFP amino acid sequence, which would have no significant influence on the fluorescence (Tsien 1998). Anyhow, following the explanations about the promoter of the reporter gene, the reporter protein itself is now in the focus of discussion.

4.1.2 Decision about the fluorescence color

After having neglected tdTomato due to its size and its non-optimal codon usage, which could have a 2- to 3-fold effect on the compared protein amount (Romanos *et al.* 1992), the reporter gene mCherry was also rejected. Despite the same condition of expression in comparison to yEGFP3, it was not possible to detect the reporter gene besides by fluorescence microscopy.

Such a result can have different reasons ranking from the expression intensity over the fluorescence abilities to the detection possibilities. For example, the DNA sequence of yEGFP3 could reassemble more closely the earlier mentioned downstream activation sequence and improve thereby the activity of the $PGK1_{prom}$. Furthermore, Shaner *et al.* (2004) stated that there has not yet been developed respectively discovered a red fluorescence, which achieves the same level or intensity in all of the important factors for fluorescence like quantum yields and photostability. Another important factor could be the detection of

fluorescence in high cell density - which was applied here - keeping in mind that the green fluorescence is most easy to detect due to its high energy difference (personal communication with Prof. Fleischmann, University of Applied Sciences Bremen, Germany). This also explains why most often online detections systems for large-scale experiments and continuously monitoring are established with the green fluorescence. A first system was developed by Randers–Eichhorn *et al.* (1997) and then improved by Jones *et al.* (2003) as well as DeLisa *et al.* (1999), enabling the detection system to distinguish specific green fluorescence even up to *E. coli* cell densities of around 1-2 x 10^{11} cells per mL respectively 150 g/L cell weight.

4.1.3 LOCALIZATION OF THE REPORTER GENE AND PROTEIN

In the experiments regarding the localization of the reporter gene inside of the genome, epigenetic effects similar to those described by Flagfeldt *et al.* (2009) could be observed, but were not monitored further through the absent significance for it in this work. It would be interesting for future experiments to determine if there are also spots in the genome, where a higher expression can be detected in comparison to the plasmid respectively *YPRCΔ15* region based expression.

More remarkable were the experiments dealing with the localization of the reporter protein inside the different cell compartments. These experiments were accomplished for the peroxisome and the cytoplasm, but not efficiently for the cell wall as well as the ER.

In most cases, the negative attempts for example to express a heterologous protein are not published (Maya *et al.* 2008). Therefore, it is sometimes difficult to find helpful answers or suggestions about own specific problems - in this situation connected to the secretion of GFP in yeast - in publications. This is especially the case if only positive experiments about *e.g.* the localization of GFP tagged membrane proteins (Newstead *et al.* 2007) or bio monitoring by GFP with the help of yeast cell surface localization (Shibasaki *et al.* 2003) can be found in literature. Even publications with nearly the exact strategy like the

anchoring of GFP to the cell wall by the C-terminal end of a-agglutinin (Ueda and Tanaka 2000) can be discovered.

A combination of works from Hashimoto *et al.* (1998), Li *et al.* (2002) as well as Huang and Shusta (2005) made way for the assumption and confirmation, that the short and relatively often used secretion signal of *SUC2* is not compatible with the expression and effective secretion of yEGFP3.

Nonetheless, a successful localization of the yEGFP3 on the cell wall might have been not as effective as expected. For example Shibasaki *et al.* (2001) discovered a high pH dependency of the around 100-300 GFP molecules attached to the cell wall. However, Tsien (1998) stated that 300-3000 GFP molecules are sufficient for optical detection in a concentrated space, making it more unlikely that the expected fluorescence increase at the cell wall could compensate the missing amount of molecules and pH oscillation.

Other cell compartments like the mitochondria or the vacuole were excluded at the beginning. The acid environment (Amstrong 2010), its volume of around 20-25% of the intercellular volume (Orij *et al.* 2009) and its degrading function (Premsler *et al.* 2009) make the vacuole not to a preferred location for fluorescent proteins, whereby they emerge there, when secretion failed (Prein *et al.* 2000).

In contrary to that, the small volume of all mitochondria of around 1-2% of the total cell volume (Uchida *et al.* 2011) and their optimal pH of about 8 resulting from the electron transport chain activity respectively proton gradient (Orij *et al.* 2009) have to be seen positively. However, despite this and its positive folding environment (Tsien 1998), the mitochondria seem to be connected to a negative effect on fluorescent proteins with respect to photostability (Bogdanov *et al.* 2009).

A different picture can be stated for the localization especially in the peroxisomes. As mentioned at the beginning, localization into the peroxisome can not only increases the folding, but also supports the concentration of the fluorescence proteins onto one spot. Whereas Tsien (1998) suggested, that the

influence by concentration of the molecules on actual non-imaging detection methods seems to be neglectable, it seems like it matters when different environmental conditions are tested respectively compared, even when the results showed a certain variation. In the conducted experiments, it occurred that the physiological conditions in the peroxisomes were additionally helpful in not shaken cultures, where no high substrate and oxygen exchange has been provided. Yet it is unclear what factors are responsible for this observation, because they can be based on higher chemical concentrations or folding supporting proteins as well as optimal physically conditions.

Nevertheless, the signal of the yEGFP3 under the control of the *PGK1* promoter was sufficiently detected with the help of the reporter gene localization into the genome region *YPRCΔ15* as well as into the peroxisome of the reporter protein, thus building a basis for the detection system and the control through the CI repressor, being discussed in the next paragraphs.

4.2 HOMOLOGIES TO THE Λ OPERATOR SEQUENCES IN THE YEAST GENOME

As already mentioned, the wild type promoter has only influence on the different promoter variants, if it is expressed in high concentrations. This depends on the one hand on the diffusion mechanism used for the nucleus transport and on the other hand on the concentration needed in the nucleus to enable repression. A factor not discussed yet is the unspecific respectively specific binding of the repressor at operator like sequences in the genome of *Saccharomyces cerevisiae*. Therefore, after having searched for the six 17 bp operator sites only, a pattern search (www.yeastgenome.org/cgi-bin/PATMATCH/nph-patmatch) for three configurations was conducted, which are listed in table 15.

Table 15: List of the pattern used

Operator site /basis	Sequence/Pattern	Hits in genome
O_L1	TACCACTGGCGGTGATA	0
O_L2	TATCTCTGGCGGTGTTG	0
O_L3	AACCATCTGCGGTGATA	0
O_L pattern	**WAYCWYYKGCGGTGWTR**	1
O_R1	TACCTCTGGCGGTGATA	0
O_R2	TAACACCGTGCGTGTTG	0
O_R3	TATCACCGCAAGGGATA	0
O_R pattern	**TAHCWCYGBVVGKGWTR**	3
O_L/O_R pattern	**WAHCWYYKPBVVGKGWTR**	45

ACTG	Exact match
R	Any purine base (AG)
Y	Any pyrimidine base (CT)
S	G or C
W	A or T
M	A or C
K	G or T
V	A or C or G
H	A or C or T
D	A or G or T
B	C or G or T
N or X	Any base

For the operator sites O_R1, O_R2, O_R3, O_L1, O_L2 and O_L3 no similar sequences were found in the genome of the reference strain S288C, however for an O_L based pattern (see table 16), O_R based pattern (see table 17) and the most unspecific O_L and O_R pattern (see table 18), it was. Therefore in the following tables the affected gene respectively localization in the genome of each hit is shown, with "-" marking a location between two genes and "$_{prom}$" the location in an assumed promoter region.

Table 16: Pattern search in the genome of *S. cerevisiae* S288C with WAYCWYYKGCGGTGWTR.

Chromosome	Hit No.	Matching Pattern	Start	End	Gene/ localization
CHR14	1	TACCATCTGCGGTGATA	689223	689239	*ABZ1*

Table 17: Pattern search in the genome of *S. cerevisiae* S288C with TAHCWCYGBVVGKGWTR.

Chromosome	Hit No.	Matching Pattern	Start	End	Gene/ localization
CHR1	1	TACCACTGTAAGTGATG	204122	204138	*FLO1*
CHR5	1	TAACACTTACGGTGGTA	566580	566596	*YER187W*
CHR7	1	TATCTCCGCAAGTGATA	557870	557886	*YGR035W-A$_{prom}$*

Table 18: Pattern search in *S. cerevisiae* S288C with WAHCWYYKPBVVGKGWTR.

Chromosome	Hit No.	Matching Pattern	Start	End	Gene/ localization
CHR1	1	TACCACTGTAAGTGATG	204122	204138	*FLO1*
CHR2	2	CATCACCCAAGGTGGTT	336116	336132	*RPS11B - YRO2*
		TACCATCTTCAGTGATG	531876	531892	*MAK5 – ADH5*
CHR4	3	TATCTTTTTCCGTGATA	194639	194655	*LOB17 – CCT4*
		AACCATTGTGGGTGATA	417772	417788	*GPM2 – CDC7*
		CAACACCCAAAATGGTT	1501635	1501651	*CAB1 – STL1*
CHR5	1	TAACACTTACGGTGGTA	566580	566596	*YER187W*
CHR7	2	CATCACCGGAAATGTTA	505769	505785	*PEX31 – PRP18*
		TATCTCCGCAAGTGATA	557870	557886	*YGR035W-A$_{prom}$*
CHR8	3	TATCATCTCCCGGGATA	83385	83401	*YHL012W-YHL009W-A*
		CAACACTTCCAGTGGTT	176153	176169	*YHR033W*
		TATCCCCGACAATGGTA	423619	423635	*SOL3$_{prom}$*
CHR9	2	TACCATTTCAAGGGTTA	77812	77828	*RPL40A - TID3*
		TAACCCTTAAGATGATT	338074	338090	*DOT5 – FAA3*
CHR10	4	AACCATTGTAAGGGATG	26766	26782	*HXT8$_{prom}$*
		TATCCCTTCAAAAGATT	83067	83083	*GON7$_{prom}$*
		TATCACTCCAAATGGTA	425390	425406	*CYR1*
		TATCACGCAAAATGGTT	598809	598825	*BUD4*
CHR11	4	TACCTTCTGGGGTGATG	141925	141941	*MRPL38 - PIR3*
		CATCACCCCAGAAGGTA	145231	145247	*PIR3*
		TAACATTTCCAGTGATA	165925	165941	*YKLWΔ3 - MCR1*
		TATCACCCCAAAAGATA	371709	371725	*TUL1$_{prom}$*
CHR12	11	CATCACGTAAAATGTTT	111573	111589	*SPA2 - YLL017W*
		AAACACCGTAGGTGATA	256757	256773	*MNL2*
		TATCACTGCCGGTGTTT	292741	292757	*GAL2 – SMC4*
		TAACCCCTGAAATGATA	333773	333789	*KIN2*
		TACCACCTTCGGTGTTG	410935	410951	*PDC5*
		TACCATTTTGAGTGATG	422885	422901	*NHA1 - RRN5*
		AAACTCCTCAAGTGTTA	863766	863782	*ROM2*
		TATCACTCAAGAAGTTA	1003561	1003577	*IMD3*
		TATCATTTTGGGTGTTG	1039449	1039465	*LEU3 - FMP27*
		TAACACCTCCAAAGTTA	1041450	1041466	*LEU3 - FMP27*
		TATCACTGCCAAAGATT	1047140	1047156	*FMP27*

		TAACTTCTTCAGGGATG	55430	55446	*YML108W – URA5*
CHR13	4	TAACTTCTTGAGTGATA	162745	162761	*CMP2 – SPC2*
		AAACTCTTGGGGTGATA	834046	834062	*GPI12 - YMRWΔ21*
		AATCACTGTCCGTGTTA	866698	866714	*LIP1 – UBP15*
		TAACACGGGCAATGTTT	130320	130336	*TOF1 – BSC4*
CHR14	4	CAACCCTGGAGGAGTTT	510580	510596	$NOP2_{prom}$
		TATCTCTTGGAGGGTTA	515550	515566	*NOP2 – OCA2*
		TACCATCTGCGGTGATA	689223	689239	*ABZ1*
CHR15	2	AACCATTTCCAGTGTTG	161494	161510	*MHF1 – PHM7*
		TATCTTTTTAGGTGATA	186827	186843	$MDM20_{prom}$
CHR16	2	CAACACTGAAGAAGTTA	858431	858447	*KRE6*
		TAACACTGAAAAAGGTA	887246	887262	*YPR172W-DPB2*

By using all O_L sites and O_R sites as a source for a pattern search, around 45 hits can be found in the genome of strain S288C, where the four previous hits are included. Out of this sequence similarities, six are of particular interest, since they are assumingly located in promoter regions of genes (*MDM20, NOP2, TUL1, GON7, HXT8, SOL3*), whereas the rest should have no influence on the cell at all. Based on the henceforth more and more increasing unspecifity of the sequences, it can be expected that the repressor binds only weakly and only at a high concentration.

It can be assumed that only the sequence similarities found with the O_L or the O_R pattern might be important. However, three of the in total only four hits are directly inside of a gene sequence, whereupon assumingly no significante influence of the repressor binding can be presumed as tested in Webster and Brammer (1995). Contrary to this is the last hit, which is assumingly in the promoter region of the putative open reading frame (ORF) *YGR035W-A* with unknown function. It remains unclear whether transcription of this gene is affected through the dimer repressor binding. Yet no abnormal phenotype could be observed in any experiments so far. Nevertheless, the possible phenotypes are listed in table 19.

Tabel 19: Possible phenotypes by genes, if influenced by repressor CI

Gene	Genomic coordinates	Function	Observed phenotypes of null mutant
MDM20	YOL076W	non-catalytic subunit of NatB N-terminal acetyltransferase	Viable; metal and chemical resistance decreased; vaculor and mitochondrial distribution abnormal; absent protein modifications;
NOP2	YNL061W	RNA m(5)C methyltransferase	Inviable; competitive fitness decreased; cell size decreased
TUL1	YKL034W	Subunit of the DSC ubiquitin ligase complex	Viable; metal and stress resistance decreased; haploinsufficient; competitive fitness decreased
GON7	YJL184W	Component of the EKC/KEOPS protein complex; t6A tRNA modification; telomere maintenance and transcription	Viable; inviable; metal and chemical resistance decreased; competitive fitness decreased; respiratory growth absent; utilization of nitrogen source absent;
HXT8	YJl214W	Protein of unknown function with similarity to hexose transporters	Viable; decreased competitive fitness; auxotrophy haploinsufficient;
SOL3	YHR163W	6-phosphogluconolactonase	Viable; decreased competitive fitness; abnormal vacuolar morphology;
	YGR035W-A	Putative protein of unknown function	Uncharacterized

4.3 Modifications on the repressor CI

This problem could be circumvented respectively diminished for instance by controlling the concentration of the repressor in the nucleus by an active nuclear transport, having the additional advantage that the transcriptional burden of the cell decreases. In such a situation, a weaker promoter could be sufficient and an easier fine-tuning of the repressor concentration could decrease unspecific control by it.

For the usage of the active nucleus transport system, a mostly positively charged NLS is essential, whose functionality only requires accessibility on the protein surface. In this study, the NLS was attached to the N-terminal, C-terminal or inter of the repressor CI, but only the internal version showed the desired effect, at the end even in low concentrations. In comparison to that, the repressor with the NLS at the C-terminus did not show any activity at all, indicating that the extra positive charges prevent the important dimerization respectively tetramerization of the repressor molecules. Otherwise could be the C-terminal domain able to bind the negative charged DNA thus eliminating in a non-specific way CI molecules to a degree that for "specific" repression they would no longer be available.

A different effect has been observed for the repressor version with the NLS at the N-terminal end, which is important for the specific DNA binding. There, no transformation into the different yeast strains was possible. This might indicate that the additional positive charge modified the DNA binding domain in such a way, that the unspecific binding was increased and the growth of the cell was inhibited. Such a growth decreasing repressor variation has also been observed by Webster and Brammer (1995), who tested a fusion protein between different phage repressors.

It could not be conclusively clarified, whether the repressor with internal NLS is really using the NLS to reach the nucleus or if the induced structural change is responsible for the effect at low concentrations with the repressor reaching the nucleus via diffusion. The wild type repressor shows no effect at the same expression intensities – neither does the internal NLS version at lower expression intensities. Important to mention is however, that all statements about the expression intensity of CI repressors are only proven indirectly, because of the incompatibility of the detection by western blot method with the ability to function as a repressor. It is also not clear to what extent the on the cloning strategy based additional Gln influences the repressor and its structure. In order to get a conception and a comparison a 3D model was developed with the help of the website “www.toolkit.tuebingen.mpg.de/hhpred” for the wild type CI and the CI with internal NLS (see figure 33 and 34 for two different perspectives and illustration foci).

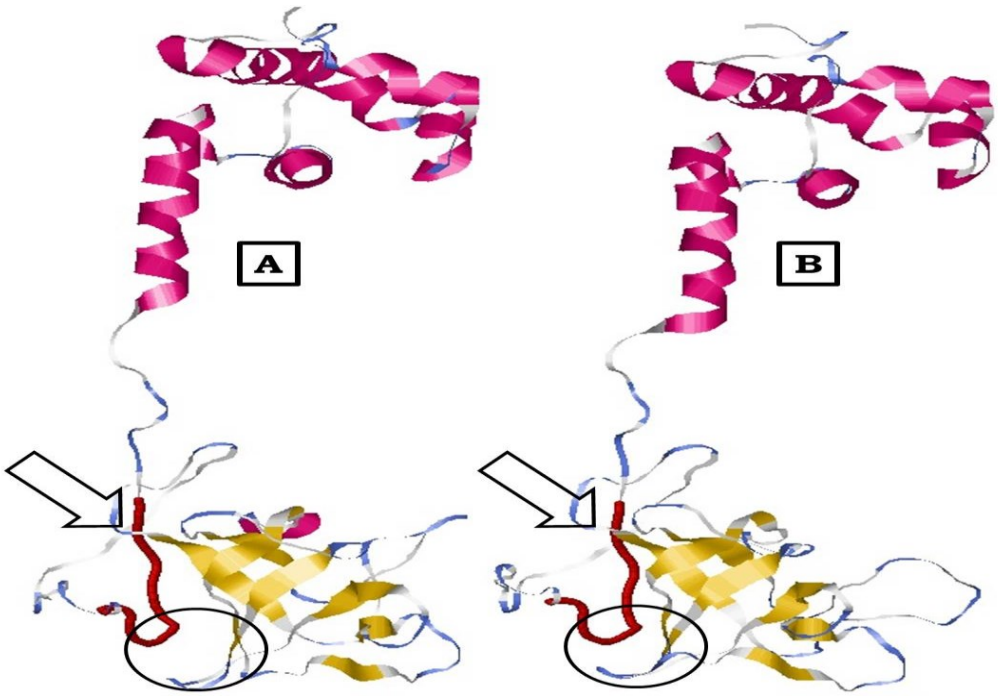


Figure 33: **Schematical illustration of the CI repressor with internal NLS (A) and the wild type CI represssor (B) 3D structure.** The α helices (purple) are dominant in the upper N-terminal DNA binding domain, whereas the β sheets (yellow) are more dominant in the lower C-terminal cooperativity domain. The black arrow indicates the beginning of the backbone from the modified amino acids 104-113 (red, see figure 12 for sequence) and the black circle one of the prominent areas of alteration between the wild type and the CI repressor with internal NLS.

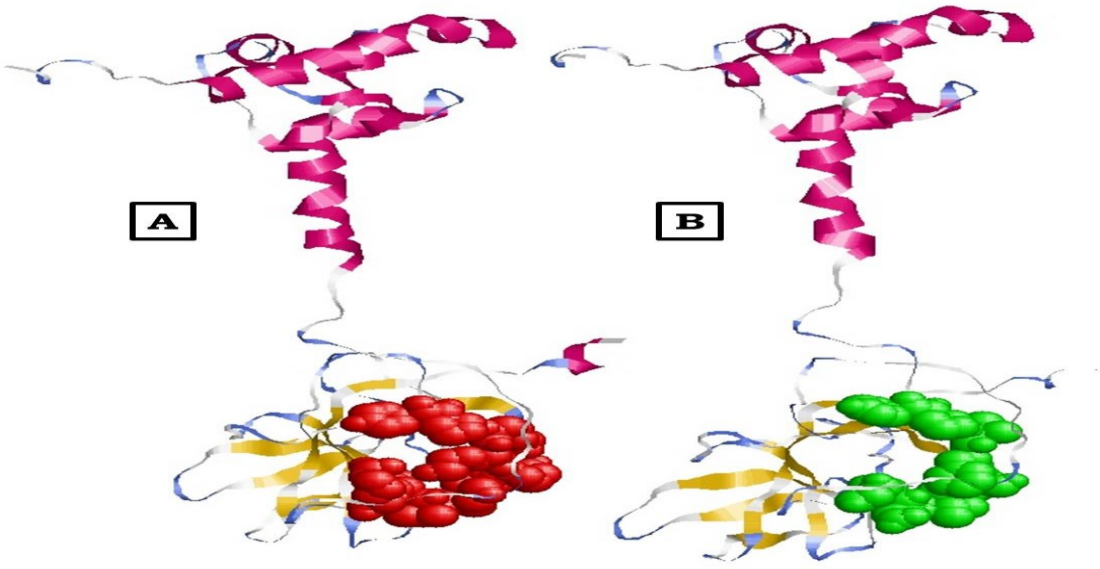


Figure 34: **Schematical illustration of the CI repressor with internal NLS (A) and the wild type CI represssor (B) 3D structure.** The α helices (purple) are dominant in the upper N-terminal DNA binding domain, whereas the β sheets (yellow) are more dominant in the lower C-terminal cooperativity domain. The modified amino acids 104-113 (see figure 12 for sequence) are illustrated in a spacefill modification with NLS (red) and wild type (green) sequence marked differently.

From both perspectives (figures 33 and 34) it can be interpreted that the DNA binding domain is not influenced by the changes at amino acids 104-113 and that even the cooperativity domain seems to be structurally identical. Comparing the backbone structure of the altered amino acid sequences 104-113 (see figure 33) a nearly similar form can be observed. However, through the amino acid exchange the whole cooperativity domain seems to be shifted, which could be explained by the more voluminous amino acids incooperated (see figure 34). Whereupon it has to be kept in mind that the 3D structure of the CI repressor with internal NLS is due to the computer programme and algorithm based on an alignment with the wild type CI repressor, which could hide or underestimate more severe structural changes in reality. Nevertheless, to what extent the secondary structure influences CI in its repression ability will be discussed in the following.

4.4 OPERATOR SITE INSERTION INTO THE *PGK1* PROMOTER

After discussing the experiments about the reporter and the repressor as well as its operator sequences, it is now time to debate about the different positions of the operator site insertion, first of all without the additional O_L site, starting with position 1.

4.4.1 POSITION 1

The assumption for the position 1 insertion (-130 bp, see figure 35) was, that an operator bound by a repressor at this location next to the seemingly preferred TATA box at -153 bp would interfere with important transcription factors. Next to the TATA box binding proteins (TBP), this could affect other important transcription factors of the of the (pre-) initiation complex like the protein complex of transcription factor D II (TFDII).

The area around the position 1 seems to be very sensitive to sequence and distance changes of the promoter structure. This can be seen on the one hand at the results of the insertion of 85 bp, where a non-functional promoter was

produced or a transcription start shift brought out a non-functional yEGFP3. On the other hand, the insertion of 45 bp at the same position leads to a functional yEGFP3 and promoter, being even similarly strong in comparison to the wild type promoter. This indicates that there seems to be a high amount of flexibility in terms of insertions in the area of the TATA box region especially unless a certain amount of additional bp length is crossed. Such mutations of the promoter sequence cannot only be connected to a decrease of activity, for example in the work of Nevoigt *et al.* (2006) an increase up to 120% has been observed for mutant *TEF1* promoters.

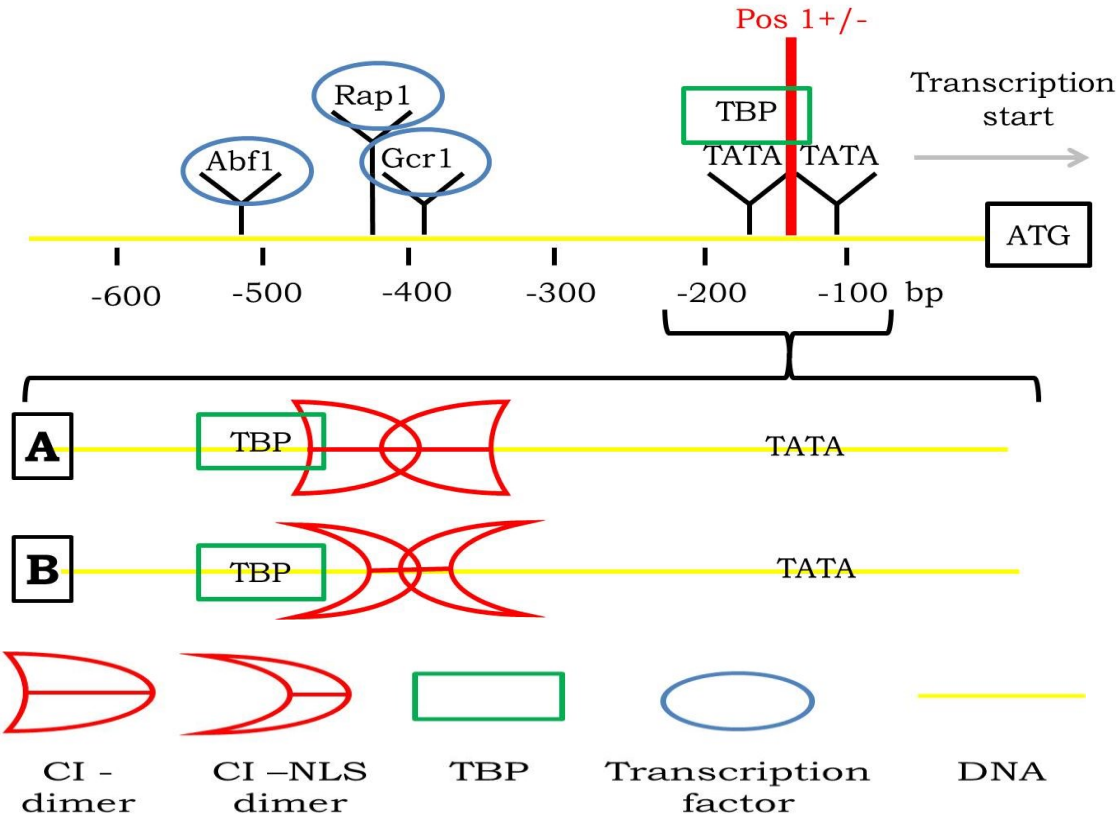


Figure 35: **Schematical illustration of the *PGK1* promoter**. The upstream activating sequences (UAS) are characterized through the binding of the transcription factors Abf1, Rap1 and GcR1 (blue circle). The core promoter region contains two possible TATA boxes where the TATA box binding proteins (TBP, green rectangle) can bind. Additionally indicated is the insertion at position 1 (red line) of the lambda phage operator sites, with "**+**" indicating an insertion of the operator sequence and "**-**" indicating a simultaneous deletion of upstream promoter sequences. The assumed binding of the CI dimer (A) and the CI with internal NLS (B) is shown, too.

However, an insertion of 45 bp at position 1 with an additional upstream deletion of the same base pairs length leads to a non-functional promoter. This occurs most probably due to the almost complete deletion of the TATA box

around -153 bp, indicating that the assumed TATA box at -113 bp is not functional or, at least, able to compensate the loss of the upstream TATA box. This makes the sequence around -153 bp essential for a functional promoter and gives another argument for the localization of position 1 next to this TATA box. Further deletion of sequences between the essential TATA box and the transcription start have not been conducted but could lead next to promoter inactivation also to a reduction of the initiation efficiency from mRNAs of up to 50% (Romanos *et al.* 1992).

In the following, the effect of the two repressor variants on the position 1 will be discussed, starting with the wild type CI. This repressor - if present in a sufficient high concentration - is capable of decreasing the expression of the yEGFP3 100% or respectively to detection limit of the used methods by binding at position 1. In the needed high concentrations, the repressor is bound assumingly as a tetramer enabling the high affinity and binding strength of the repressor CI. This repressor complex is therefore capable to block the TATA box at -153 bp and decreases the binding chance of TBP or TFDII tremendously. With this result, the repression at this position can be assumed stronger than observed before in other experiments with the lambda phage repressor like in the works of Webster and Brammer (1995) as well as Wedler and Wambutt (1995).

A different effect is noticeable in the combination of position 1 with the CI repressor containing an internal NLS site. There, even under high concentrations no influence on the expression is visible. The previously discussed results indicate that this effect is not based on the position, but on the repressor itself. The structural change through the additional mostly positively charged amino acids can only be assumed, but it either changes the overall form of the protein complex, allowing binding of other proteins, or it circumvents the building of the tetramer respectively even the dimer, thus decreasing the binding strength to the DNA as a complex. In any case, it changes the overall isoelectric point from 4.87 to 6.74 and even more so for its central connecting peptide. Most probable is thereby the last situation. Because of the close orientation of the NLS to the C-terminal, respectively

dimerization domain of the CI repressor, an influence on the dimerization or tetramerization is imaginable, but seemingly not as severe as in the case of the CI repressor with NLS on the C-terminal region. As already mentioned, the binding strength of single CI repressors is not sufficient enough, leading most likely to a kick off of it by the TBP respectively TFIID complexes.

4.4.2POSITION 2

Another picture is noticeable for the position 2 (-163 bp, see figure 36), which is in front of the important -153 bp TATA box and shows in the work of Webster and Brammer (1995) - where it is 5 bp in front of the TATA box - the highest repression of around 65%, even with only one operator site. Similar to position 1 an effect on the binding affinity respectively possibility of TBP or TFDII is suspected, as soon as a repressor binds the lambda phage operator at this position. Analogous as well is the influence of the insertion of additional DNA sequences. The initially 85 bp lead again to a non-functional promoter, whereas the insertion respectively replacement of 45 bp had almost no effect on the promoter strength.

The influence on position 2 by the wild type CI, which showed the high repression at position 1 before, will be described first. Nevertheless, despite the promising pre-experiments of Webster and Brammer (1995) with the 434 phage in a similar position, it was not possible to improve the repression by the usage of an additional operator site and the lambda phage repressor with its higher affinity compared to the 434 repressor. In very high repressor concentration, only the approximately same decrease of the promoter activity has been observed at nPos 2 with around 57% (see figure 20). It can only be speculated in what expanse the in comparison additional distance of 5 bp is thereby the crucial factor.

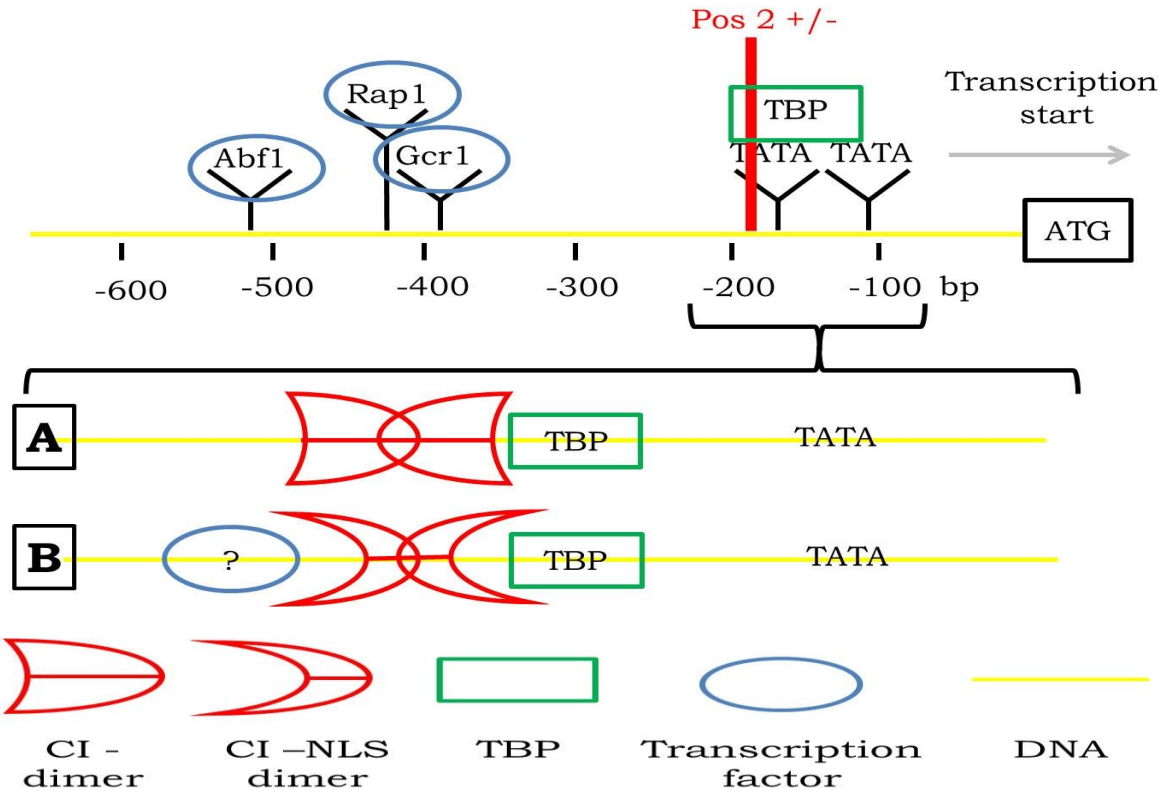


Figure 36: **Schematical illustration of the *PGK1* promoter**. The upstream activating sequences (UAS) are characterized through the binding of the transcription factors Abf1, Rap1 and GcR1 (blue circle). The core promoter region contains two possible TATA boxes where the TATA box binding proteins (TBP, green rectangle) can bind. Additionally indicated are the integration position 2 (red line) of the lambda phage operator sites, with "**+**" indicating an insertion of the operator sequence and "**-**" indicating a simultaneous deletion of upstream promoter sequences. The assumed binding of the CI dimer (A) and the CI with internal NLS (B) is shown, too.

However, no other position 2 variation was tested with such high repressor concentration, based on the results achieved by them in combination with the *CEN6/ARSH4* plasmids, where no influence on the expression has been observed. Nevertheless, it can be assumed that nPos -2 will act similar to nPos 2 in high concentration of wild type CI repressor since the repressor tetramer binds near the TATA box and disturbs the binding of the TPB respectively TFDII, which is again the explanation for the observed decreased expression.

Of particular interest is now the combination of the two position 2 variants with the repressor containing the internal NLS site. Even in high concentration, no effect can be seen in nPos 2. Again, it looks like that the internal NLS site changes the repressor in such a way, that the previously observed repression is not detectable. Probably is that this occurs through a modified binding affinity to each other followed by a changed forming ability of dimers or tetramers. It

can be clearly documented that there is at least a binding to the DNA when nPos 2 is compared directly to nPos -2 in combination with the CI repressor with internal NLS. In the combination with nPos -2, an increase of up to 70% depending on the repressor concentration in the expression rate of yEGFP3 can be detected. The assumed explanation for this could be either that the repressor is able to recruit TF faster or that the repressor retains the necessary factors, which both leads to an increased expression. Hence, the two variants just differ in the presence or absence of additional 45 bp of DNA and not in the distance to the TATA box. It is clear that this increase cannot be connected to for example an increased binding or retaining of the TBP respectively TFDII as well as an factor connected to DNA structural changes. It is likely that by the additional positive charge either the 3D structure is changed in this way that TFs retain at their position or it is based just on the positive charge itself. This theory will be solidified in the results for the position 3.

4.4.3 POSITION 3

The insertion of operator sequences into position 3 (-370 bp, see figure 37) is based on another assumption than the one which served to place O_R in position 1 or 2. In position 3, the binding of the repressor on the operator site should interfere with the binding of the different additional TFs in the UAS of *PGK1prom*. An only around 200 bp long promoter (JOE0071) showed no fluorescence underlining the importance of the UAS confirming previous deletions analyses at Ogden *et al.* (1986). This was confirmed also by Webster and Brammer (1995), which furthermore added an operator site at nearly the same position reaching a reduction of expression with a wild type promoter of around 31%. Again, the insertion of the 45 bp operator sequences with the two lambda phage sites O_L1 and O_L2 diminished the promoter activity just slightly, but the results of the combination with either the wild type repressor or the CI repressor with internal NLS are quite different. In a high concentration of wild type lambda repressor the position nPos 3 (20%) as well as nPos -3 (50%) show a reduced activity of the promoter. Thereby it can be observed that the operator sites seem to follow the statute that the closer it is to the binding site of Gcr1

the stronger the reduction of the expression gets. This indicates that the binding of the wild type lambda phage repressor assumingly disturbs the binding of the transcription factors.

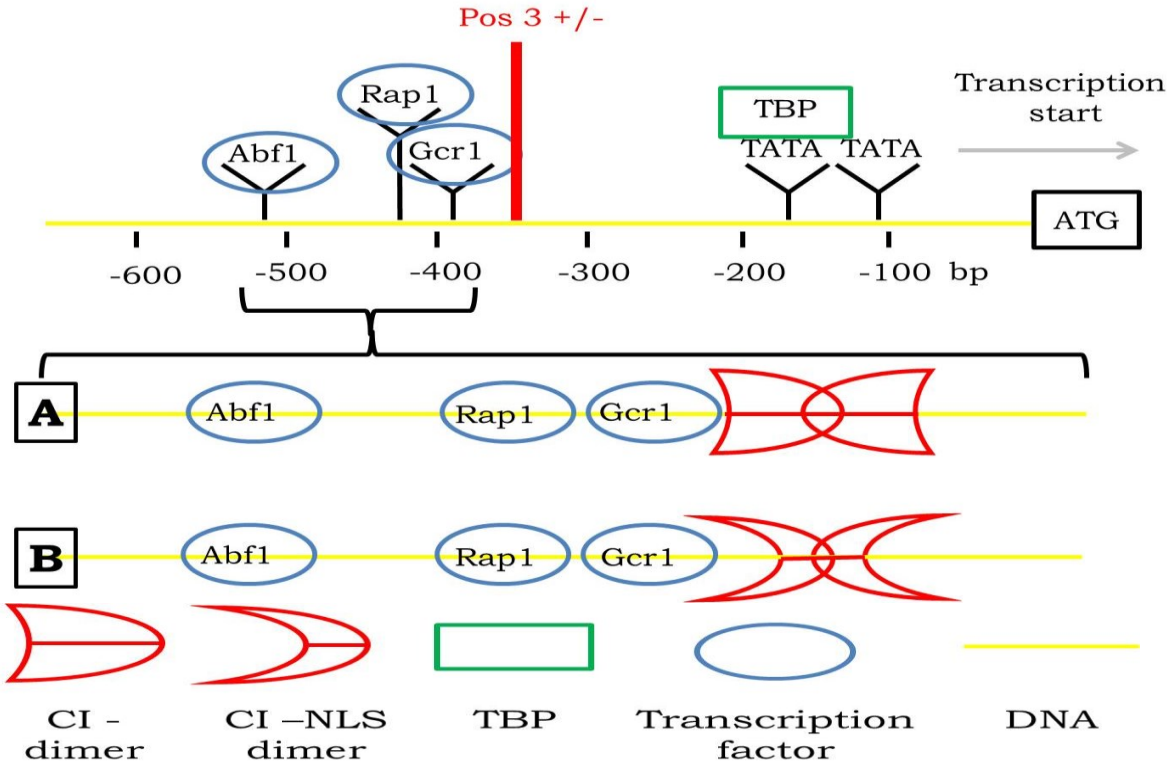


Figure 37: **Schematical illustration of the *PGK1* promoter**. The upstream activating sequences (UAS) are characterized through the binding of the transcription factors Abf1, Rap1 and GcR1 (blue circle). The core promoter region contains two possible TATA boxes where the TATA box binding proteins (TBP, green rectangle) can bind. Additionally indicated are the integration position 3 (red line) of the lambda phage operator sites, with "**+**" indicating an insertion of the operator sequence and "**-**" indicating a simultaneous deletion of upstream promoter sequences. The assumed binding of the CI dimer (A) and the CI with internal NLS (B) is shown, too.

In contrast to this, the repressor with internal NLS increases the expression in both position variants nPos -3 and nPos 3. Similar to the detected repression there it also can be observed, that the closer the operator sequence is to the TF binding site the more the expression increases. This leads to an upsurge of the expression at nPos 3 of around 57% and at nPos -3 at around 81%. Thus the theory is strengthened that the additional internal NLS alterations of the repressor leads to such a structural change or charge change or both that the transcription factors – in this case probably the low concentrated Gcr1p - can assumingly either bind more effectively or are retained.

An insertion of the operator sites between the binding sites of Rap1 and Abf1 was disregarded based on the initial assumption that this would have been accompanied by a destruction of interaction. However, the results of Webster and Brammer (1995) showed in this area a high repression and it would be interesting to find out in future experiments, if the repressor with internal NLS would increase the expression there, too.

4.4.4 Effect of the combination of positions

Interesting to figure out now is certainly the combined insertion of the positions, which differ in their reaction to the repressors into one promoter variants. Although the combinations nPos 2-3, 2+3 and -2-3 have also been constructed, only the combinations nPos 1-3, 1+3 and 1-2 were tested more intensively because of the assumed higher repression of combinations with position 1, which equates more to the overall goal in this work.

However, the first drawback, which can be recognized for the combination of the positions, is that seemingly an insertion of two operator sites of 45 bp each containing two O_L sites decreases the activity of the promoter to around 50%. This is somehow remarkable because the single insertions lead to an almost not recognizable reduction of promoter activity. As mentioned regarding the transcription factor Abf1, it is known that some TF can compensate the loose respectively weaker binding of other TF. But in this case, it can be assumed that through the insertion of 90 bp this is seemingly not possible anymore.

Next to this difference, the positions nPos 1-3, 1+3 and 1-2 react similar to their single variants. In combination with the assumingly high concentration of wild type CI repressor, all of them show a reduced expression of yEGFP3 and in combination with high concentrations of the CI repressor with internal NLS an increase of promoter activity can be observed. These results support the positioned thesis that the wild type repressor is able to block the binding of either TFIID or TBP if it can form dimers or tetramers. Notable is thereby, that the result of Webster and Brammer (1995) with only one operator site suggests that dimer formation is sufficient for repression.

Moreover, the thesis is sustained that the repressor with internal NLS is assumingly able to retain transcription factors and thus increase the expression. For both theories, the distance between the operator site and its assumed place of activity is of course crucial for the efficiency and effectivity of the repressor variation.

Noteworthy is that for the TAT variation, which was based on the method of Wedler and Wambutt (1995) as well as on the two artificial operator sites from Maniatis *et al.* (1975), only a promoter activity of around 50% could be detected, too. This observation is contradictory to the results of Wedler and Wambutt (1995), where promoter activities between 80% and 100% were measured. However, in their study, the improved promoter is located on a 2µm-based plasmid and the wild type O_R sites are used, both factors being a possible reason for the promoter activity discrepancy not taking into account possible copy number variations. Nevertheless, a similar reduction of expression (84%) was detected if the TAT variation is combined with a high expression of wild type repressor. Similar to the previously described constructs it can be assumed that a tetramer or dimer binding of the repressor blocks the binding of TBP respectively of the TFDII complex thus decreasing the activity of the promoter.

4.4.5 THE POSSIBILITY OF LOOP FORMING

In the wild type repression mechanism of the lambda phage loop forming plays an important role in regard to control and stability. In several expression cassettes with promoter variants, an additional O_L site is added to enable loop forming (JOE0066/1ol, JOE0067/2ol, JOE0068/-2ol, JOE0070/-3ol, Jo0069/3ol). With distances of 1391 bp to 1641 bp between the operator sits O_L and O_R - depending on the promoter variants – the possible loop is long and entropic as well as therefore independent of the repressor concentration as described in Griffith *et al.* (1986) and Dodd *et al.* (2004).

Remarkable is that the pure addition of the O_L site in nearly all the promoter variants led to a slight decrease of the activity. This could be explained either

by the additional cloning step, increasing the probability for mutations or by the construction scheme itself. The O_L sites were inserted 3' to the CYC_{term} and therefore did not have a direct influence on the expression cassette of yEGFP3. However, through this insertion the *URA3* marker terminator has been shortened by 100 bp, which could cause a drop in efficiency of transcriptional termination of the marker gene and thus maybe to a decrease expression of yEGFP3 via transcription blocking effects. This random inhibition could therefore explain the variation of decrease observed but not its constancy in most of the measurements.

Nevertheless, all requirements for loop forming are met, indicating the possibility of it. It can thus be assumed that the repressor with internal NLS, which is showing most probably an inability to form dimers or tetramers, is also not able to form an octamer as a basis for the loop. For these reasons, hints for loop forming have to be searched for in the results with the wild type CI repressor. An indication for loop forming can thus be seen either in the comparatively lower relative fluorescence at control point 2 of test with nPos 1ol and 1 combined with JOE0079 (see figure 25 and 26) or in the high repression without any variance of nPos 1ol in combination with JOE0092 (see figure 20). However, further experiments have to be conducted to control the influence of the additional O_L site as well as the proof of loop forming. Despite all that, Morelli *et al.* (2009) describes, that a repression is also possible without the loop, which is then not as stable as well as frailer to repressor concentrations differences.

4.5 LIMITS AND POSSIBILITIES OF THE DETECTION SYSTEM

The developed detection system for plasmid stability in *Saccharomyces cerevisiae* is relatively complex and therefore contains a multiplicity of regulation screws and factors to consider. Many factors are still remaining uncertain. For example is it not clear what kind of influences the insertion of operator sites have on the chromatin structure of a DNA region and so on the accessibility of transcription factors or CI proteins, explaining maybe an effect

of repressors at insertion position 1 in assumingly low repressor concentrations (see figure 21).

However, an important point for the effectivity of the system is its detection limit, which could already be increased by the decrease of the control volume from 2 ml to 13 µl. Through this step, only 30% instead of 50% of the total cells are required in the culture to have lost their plasmid for detection. The usage of flow cytometer, where up to 10 000 cells per second are controlled, or the more and more important microfluidic 'lab on a chip' technology, where individual cells are measured (Longo and Hasty 2006), could further increase the detection limit. However, the impact of the reaction time, which is comprehended as how fast plasmid loss is leading to a signal output, seems to be the most important factor for the detection system.

This reaction time is depending first of all on the forming respectively oxidation step necessary for the chromophore building of the yEGFP3 neglecting thereby the time for its transcription and translation. Depending on the oxygen concentration respectively other conditions, this step occurs for wild type GFP in around 2 (Tsien 1998) to 4 hours (Hackett *et al.* 2006). However, through genetic improvements like the exchange of Ser^{65} to Thr a chromophore forming which is four times (Tsien 1998) faster or even claimed as immediately (De Wulf *et al.* 2000) can be observed. Even a more than 7.5 time faster chromophore building is possible (Welsh and Kay 1997), through the exchange of Ser^{65} to Gly and Ser^{72} to Ala, which was used in this study. At the moment, this is the fastest and cannot be improved further on the level of the fluorescence protein. Nevertheless, it is of course also possible to follow the argumentation of Li *et al.* (2000), claiming that in long lasting large scale experiments and processes with yeast the delay of approximately 30 minutes in this case can be neglected.

Another factor considered to affect the reaction time is the repressor and reporter protein stability, which is assumed to have the most important and influential significance in this detection system. As a guideline for protein stability, *e.g.* the so-called N-end rule can be used; yet keeping in mind that a statement on the protein stability plainly based on its amino acid sequence can only be seen as estimation.

The basic assumption of the N-end rule, which is biquitous in the eukaryotic kindom, is that the half-life of a cytoplasmic protein in a cell is connected to its N-terminal residue as well as to an internal lysine residue serving as a substrate for the ubiquitin system (Romanos *et al.* 1992). The residues can thereby grouped based on their instability into primary, secondary and tertiary amino acids, whereupon a change from tertiary to secondary (N-terminal amidohydrolase) as well as then to primary (Arg-tRNA-protein transferase) is possible in yeast, but of course in slightly different ways in other organsims as well (Varshavsky 1997). Additionally, the Met aminopeptidase can remove the initial Met at the N- terminal, depending mostly on the size respectively gyration radius of the subsequent amino acid (Romanos *et al.* 1992) and its impact on the stability, increasing amino acids seemingly being preferred. However, also other degradation signals for this pathway have been recognized (Varshavsky 1997) and the rule itself seems to differ between different proteins (Hackett *et al.* 2006), making the prediction even more complex.

Both, repressor and reporter are apparently stable, because they normally carry an amino acid combination of Met followed by Gln due to the used strategy with the restriction enzyme *Sph*I. Thereby a removal of the Met based on the N-end rule is unlikely (Romanos *et al.* 1992), giving both proteins a half-life of around 30 hours (Varshavsky 1997). Even the assumed start codon adjournment by the point mutation of the *PGK1* promoter would lead to a stable protein with the combination of Met followed by Gly, where the Met is assumingly removed (Romanos *et al.* 1992) and the Gly - connected also to a protein half-life of 30 h (Varshavsky 1997) - is the starting amino acid of the proteins. Additionally, it has to be mentioned, that both proteins carry naturallywise a proximal internal Lys, which is important as a substrate for the ubiquitin system.

Whereas the high stability of the reporter protein does not seem to be problematic for an efficient and functional detection system, the results in this work suggest that it seems to be for the effectivity of the repressor proteins though. At least it explains the unexpected long reaction time of the detection system, because even if the cells loosed their plasmid, a high concentration of

stable repressor molecules remains, leading to an elongated reaction time of the system. A solution would be to decrease the stability of the repressor though without decreasing the sufficient high concentration of the repressor at any time point to ensure a functional repression system. Such a balanced concentration is influenced by the used plasmid basis, whose impact and differences will be discussed in the following.

4.6 PLASMID DIFFERENCES

In the present study, not only a *CEN6/ARSH4* plasmid has been used, but also a 2µm-based plasmid, which was developed from the *CEN6/ARSH4* plasmid via Omega PCR (see table 12 and 2.5.6.3). This apparently small difference in the segregation origin has an enormous effect on the plasmid stability observed in the conducted experiments.

The *CEN6/ARSH4* plasmid was tested in two experiments for plasmid stability and seemed to be relatively stable under the here applied conditions of a culture in non-selective medium, containing after 30 h still around 70% (see figure 25 and 26) cells with a plasmid. It has to be kept in mind that such results are always depending on the used conditions. For example, the results of Chen *et al.* (1997) indicate that already after 20 hours of cultivation only 40% of the cells in rich medium and 55% in non-selective medium possess a centromere plasmid, while Baruffini *et al.* (2009) shows that after 72 hours still around 60% of the cells carry a centromere plasmid in non-selective conditions. However, it has to be kept in mind that limited comparison likewise for the following reference results is possible as all these plasmids differ in multiple sequences.

In different selective media of course higher plasmid stability can be observed at any time point for centromere plasmids from 80% in Hegemann *et al.* (1999) to around 85% in Baruffini *et al.* (2009) or between 84% and 89% in this study (see figure 24 and 25) or 90% in Chen *et al.* (1997) as well as Falcon and Aris (2003).

Thereby it remains unclear – interestingly it is also not discussed - why not all cells carry a plasmid especially in selective media and thus the assumed

explanations for it range from active or passive cross feeding of plasmid free cells to impreciseness of the detection method of plasmid loss.

The same phenomenon has been observed for 2µm-based plasmids with the percentage of plasmid carrying cells in selective media varying between 79% and 94% in this study (see figure 27-31) or around 90% in Falcon and Aris (2003) as well as around 80% in Baruffini *et al.* (2009).

Interesting is, that 2µm-based plasmids show a higher plasmid loss rate than centromere plasmids, despite of the high stability of wild type 2µm plasmids. In this study after 48 hours of growth in non-selective and rich medium around 40% to 50% of the cells had lost their plasmid. Such a high instability has also been observed in the work of e.g. Baruffini *et al.* (2009), where after 72 hours only around 30% of the cells carried a plasmid. The mechanism for the - in comparison to centromere plasmids - higher plasmid instability of 2µm-based plasmids as well as of the plasmid stability by itself remains unknown as both plasmid types normally should be segregated actively.

Another not yet fully understood phenomenon is the plasmid number of 2µm-based plasmids, because especially used for the multicopy appearance there seems to be a discrepancy in the assumed copy number as well as the thereby expected increase of protein expression.

An explanation for this has been presented by the extensive work of Karim *et al.* (2013), which suggested that two points are most important for the actual plasmid number. The selection marker is one of them, whereby the assumed order from high number to low number is *HIS3~TRP1* > *URA3* > *LEU2* > *KanMX* and the other is the used promoter for the heterologous protein production, whereby an increased promoter strength decreases the plasmid number. However, there seem to be more important factors, because for example Shibasaki *et al.* (2001) observed only a threefold increase in expression intensity (*TRP1* marker, $GAPDH_{prom}$), despite an around 20 times increase of copy number, based on values from Shibasaki *et al.* (2001) in comparison to Karim *et al.* (2013). Additionally, Baruffini *et al.* (2009) detected an approximately 11-14 fold increase of heterologous gene expression ($MET17_{prom}$) in an 30 to 35 numbered plasmid reaching high copy numbers despite the use

of *KanMX*. An even more extreme value can be observed in the publication of Nacken *et al.* (1996), where a switch to a 2µm-based plasmid (*URA3* marker) increases the productivity of the weak $KEX2_{prom}$ up to 140-fold, which has not been noticed in a similar extent in the present work (see Appendix 7.6.2).

These results show that on the one hand once more it seems to be obvious that the expression intensity - as soon as another limit, like for example transcription factors is reached – is not solely depending on a high number of expression cassette. On the other hand, an explanation for the copy number differences in artificial 2µm plasmids has not been developed yet, when compared to the wild type situation of the 2µm.

As mentioned at the beginning, regarding 2µm-based plasmids it is assumed that they undergo homologous recombination with the resident 2µm to remain in the cell, whereupon it seems to be more of an equilibrium, because it is possible to isolate unchanged 2µm-based plasmids again from the cell. To what extent the previously described factors, selection pressure or the quantity of transformed plasmids, influences this mechanism is unclear and not understood yet - despite the wide usage of 2µm-based plasmids. Maybe the artificial 2µm plasmid outgrows the wild type plasmids even without the in the introduction mentioned high selective pressure.

Nonetheless, in experiments is differentiated only between plasmid bearing cells and cells without plasmids and no steady decline has been observed respectively investigated. Unfortunately, in Karim *et al.* (2013) the influence of the plasmid stability was also not tested, but with the here described reporter system it would be easily possible to improve the knowledge about the mechanism of plasmid stability respectively 2µm-based plasmid copy number.

5 CONCLUSION AND OUTLOOK

After having discussed the obtained results in this study, it is now time to summarize which goals and objectives have been reached and what kind of application can be expected with it as well as some to name improvements and suggestions for future experiments at the end.

5.1 GOALS REACHED

First and foremost it was necessary to build a basis for the online reporter system to detect plasmid instability in *Saccharomyces cerevisiae* through the development of a stable expression of fluorescent proteins with an optimized location for signal strength. This was achieved and approved by the constant high expression of yEGFP3 located mostly unaffected through environmental changes in the peroxisome.

In the following experiments, it was possible to control the promoter by a lambda phage based repressor system. By this, the highest repression of 85% reached in Wedler and Wambutt (1995) as well as 61% in Webster and Brammer (1995) was seemingly increased through the developed optimizations in regard to operator and repressor to almost 100% respectively the detection limit of the used method.

Finally, with the complete detection system plasmid stability was investigated under selective, non-selective and rich media conditions, whereby a correlation between plasmid loss and increasing fluorescence signal has been observed. Despite the assumption, a signal increase would be easier and more confident to detect than the signal decrease of the fluorescence in the detection system of Hegemann *et al.* (1999), this was not possible at the present time. However, where this plasmid based system is subjected to wild type oscillations in the plasmid circumstances and therefore at its limits, the here developed detection system is still possible to be improved further.

Next to the repression of the fluorescence, an increase of the expression intensity could be observed in specific combinations of operator integration sites and modified repressor molecules. Such an expression increasing system like the repression method could have different application possibilities, which will be described in the following.

5.2 APPLICATIONS FOR THE SYSTEM

Besides the conceived implementation of detecting plasmid stability of *Saccharomyces cerevisiae* online, the observed effects and mechanisms can be used for further applications.

One additional use could be seen in the production of toxic products respectively substrates, which should only be produced in the late phase of growth with high cell densities. Nowadays this is often mastered either by inducible promoters using an additional substrate for activation like mostly expensive antibiotics or by promoters, which are only active if a substrate like substrate is missing in the medium and thus controlled by the initial substrate concentration respectively addition. With the knowledge about plasmid instability, the developed system could be used in a way that the stable version of the repressor is for example based on a relatively unstable *ARS* plasmid. This will be lost over time, but can help to provide an initial stable repressor concentration, which is slowly degrading and thus activating the promoter in a late phase of culture growth. Such a method would have several advantages on the one hand would it be medium independent and on the other hand no additional substrate has to be added, but of course it requires an intensive study about the given plasmid stabilities.

The relatively new field of synthetic biology is another area, where the results of this study could be deployed usefully. Synthetic biology tries to implement engineering principles into the design of new systems or developments (Blount *et al.* 2012) and covers thereby artificial genetic circuits, metabolic and protein engineering as well as the now increasingly cheaper de novo DNA synthesis (Sayut *et al.* 2007). In the world of engineering, many processes follow a logic

gate function, which expresses itself in regulatory networks with genetic inverters and transistors or oscillatory dynamics as well as toggle switches (Drubin *et al.* 2008). The here developed system represents such a switch and additionally the bistable combinations of positions could open a new possibility of control. It would be interesting to see the effect respectively the behavior of this position combination in the presence of both repressor types.

However, similar as Hong and Nielsen (2012) requested, more work has to be done to characterize different promoters like the here developed, especially with industrially relevant fermentation conditions. The here obtained results reinforce the thesis of Boyle and Silver (2009), that "rationally designed biological devices rarely function entirely as predicted" and "that the success of synthetic biology endeavors depends heavily on understanding how biological systems are regulated". Out of these reasons, such kind of developed systems have to be customized constantly and some improvements for the system will be described next.

5.3 IMPROVEMENT OF THE DETECTION SYSTEM

Like described for multi gene pathways by Flagfeldt *et al.* (2009) is it important for the developed detection system that the expression levels respectively protein stabilities of the single elements are balanced to achieve an optimal repression and detection on the same time. However, it has to be kept in mind that the intrinsic and extrinsic noises have strong effects on genetical circuits (Sprinzak and Elowitz 2005), thus limiting the optimization steps. Furthermore, it is clear that a decision in favor of a faster detection, by for example a decrease of the repressor promoter strength, also the stability of the repression decreases and the other way around.

Nevertheless, the obtained results show that assumingly either the repressor protein stability is too high or the repression is too tight, for a fast detection, which is required. Therefore, it could be better to use - in contrast to the initial assumptions - promoter variants without additional O_L sites, since Morelli *et al.* (2009) describes that even the wild type repression system of phage lambda

works under high repressor concentration without loop forming. It has still to be considered that the achieved results in this study showed a higher variation for the repression and Morelli *et al.* (2009) mentions that the repression is more affected by repressor concentration fluctuation.

More important seems to be the reduction of the basal expression to improve the gene expression control like Maya *et al.* (2008) describes it and thus improve the detection system. A possible compromise and the seemingly best improvement for the developed detection system of plasmid instability in *Saccharomyces cerevisiae* would be to decrease the protein stability of the repressor. As a basis for this, the work of Hackett *et al.* (2006) could be used, where through exploitation of the N-end rule degradation pathway and an ubiquitin fusion strategy different protein stabilities could be reached.

Overall, it can be summarized that for the application of this simplified repressor system from the phage lambda it is not possible yet to circumvent the decision between a fast responsive unstable and a stable system with reduced basal expression. Only with an additional complexity respectively repression level this could be possible in future.

6 REFERENCES

Achstetter, T., Nguyen-Juilleret, M., Findeli, A., Merkamm, M., Lemoine, Y. (1992): A new signal peptide useful for secretion of heterologous proteins from yeast and its application for synthesis of hirudin; in: Gene, No.1 (Vol 110), pp. 25-31.

Aggarwal, A. K., Rodgers, D. W., Drottar, M., Ptashne, M., Harrison, P. C. (1988): Recognition of a DNA operator by the repressor of phage 434: a view at high resolution; in: Science, No. 4880 (Vol. 242), pp. 899-907.

Ahn, Y. T., Wu, X. L., Biswal, S., Velmurugan, S., Volkert, F. C., Jayaram, M. (1997): The 2micron-plasmid-encoded Rep1 and Rep2 proteins interact with each other and colocalize to the *Saccharomyces cerevisiae* nucleus; in: J Bacteriol, No. 23 (Vol. 179), pp. 7497-506.

Ahner, A., Brodsky, J. L. (2004): Checkpoints in ER-associated degradation: excuse me, which way to the proteasome? in: Trends in cell biology, No. 9 (Vol. 14), pp. 474-8.

Amberg, D. C., Burke, D. J., Strathern, J. N. (2006): Preparation of genomic DNA from yeast using glass beads, Cold Spring Harb Protoc, doi:10.1101/pdb.prot4151.

Anderson, L.M., Yang, H. (2008): DNA looping can enhance lysogenic CI transcription in phage lambda; in: PNAS, No. 15 (Vol. 105), pp. 5827-32.

Armstrong, J. (2010): Yeast vacuoles: more than a model lysosome; in: Trends Cell Biol, No. 10 (Vol. 20), pp. 580-5.

Avlund, M., Dodd, I. B., Sneppen, K., Krishna, S. (2009): Minimal gene regulatory circuits that can count like bacteriophage lambda; in: Journal of molecular biology, No. 4 (Vol. 394), pp. 681-93.

Bahl, M. I., Sorensen, S. J., Hestbjerg Hansen, L. (2004): Quantification of plasmid loss in *Escherichia coli* cells by use of flow cytometry; in: FEMS Microbiol Lett, No. 1 (Vol. 232), pp. 45-9.

Baird, G. S., Zacharias, D. A., Tsien, R. Y. (2000): Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral; in: Proc Natl Acad Sci, No. 22 (Vol. 97), pp. 11984-9.

Bakk, A., Metzler, R. (2004): Nonspecific binding of the O_R repressors CI and Cro of bacteriophage lambda; in: Journal of theoretical biology, No. 4 (Vol. 231), pp. 525-33.

Bakk, A., Metzler, R., Sneppen, K. (2004): Sensitivity of O_R in phage lambda; in: Biophysical journal, No. 1 (Vol. 86), pp. 58-66.

Bandhu, A., Ganguly, T., Jana, B., Mondal, R., Sau, S. (2010): Regions and residues of an asymmetric operator DNA interacting with the monomeric repressor of temperate mycobacteriophage L1; in: Biochemistry, No. 19 (Vol. 49), pp. 4235-43.

Bar-Even, A., Paulsson, J., Maheshri, N., Carmi, M., O'Shea, E., Pilpel, Y., Barkai, N. (2006): Noise in protein expression scales with wild type protein abundance; in: Nature genetics, No. 6 (Vol. 38), pp. 636-43.

Baruffini, E., Serafini, F., Lodi, T. (2009): Construction and characterization of centromeric, episomal and GFP-containing vectors for *Saccharomyces cerevisiae* prototrophic strains; in: J Biotechnol, No. 4 (Vol. 143), pp. 247-54.

Basehoar, A. D., Zanton, S. J., Pugh, B. F. (2004): Identification and distinct regulation of yeast TATA box-containing genes; in: Cell, No. 5 (Vol. 116), pp. 699-709.

Bayer, T. S., Smolke, C. D. (2005): Programmable ligand-controlled riboregulators of eukaryotic gene expression; in: Nature biotechnology, No. 3 (Vol. 23), pp. 337-43.

Benson, N., Adams, C., Youderian, P. (1992): Mutant lambda repressors with increased operator affinities reveal new, specific protein-DNA contacts; in: Genetics, No. 1 (Vol. 130), pp. 17-26.

Benson, N., Sugiono, P., Youderian, P. (1988): DNA sequence determinants of lambda repressor binding in vivo; in: Genetics, No. 1 (Vol. 118), pp. 21-9.

Bhaumik, S., Gambhir, S. S. (2002): Optical imaging of *Renilla* luciferase reporter gene expression in living mice; in: Proc Natl Acad Sci, No. 1 (Vol.99), pp. 377-82.

Billinton, N., Barker, M. G., Michel, C. E., Knight, A. W., Heyer, W. D., Goddard, N. J., Fielden, P. R., Walmsley, R. M. (1998): Development of a green fluorescent protein reporter for a yeast genotoxicity biosensor; in: Biosens Bioelectron, No. 7-8 (Vol. 13), pp. 831-8.

Blake, W. J., Kaern, M., Cantor, C. R., Collins, J. J. (2003): Noise in eukaryotic gene expression; in: Nature, No. 6932 (Vol. 422), pp. 633-7.

Blazeck, J., Garg, R., Reed, B., Alper, H. S. (2012): Controlling promoter strength and regulation in *Saccharomyces cerevisiae* using synthetic hybrid promoters; in: Biotechnology and Bioengineering, No. 11 (Vol. 109), pp. 2884-95.

Blount, B. A., Weenink, T., Ellis, T. (2012): Construction of synthetic regulatory networks in yeast; in: FEBS Letters, (Vol. 586), pp. 2112-21.

Bogdanov, A. M., Bogdanova, E. A., Chudakov, D. M., Gorodnicheva, T. V., Lukyanov, S., Lukyanov, K. A. (2009): Cell culture medium affects GFP photostability: a solution; in: Nat Methods No. 6 (Vol. 12), pp. 859-60.

Boyle, P. M., Silver, P. A. (2009): Harnessing nature's toolbox: regulatory elements for synthetic biology; in: J R Soc Interface, No. 6 (Vol. 4), pp. S535-46.

Brenowitz, M., Senear, D. F., Ackers, G. K. (1989): Flanking DNA-sequences contribute to the specific binding of CI-repressor and O_R1; in: Nucleic acids research, No. 10 (Vol. 17), pp. 3747-55.

Brew, C. T., Huffaker, T. C. (2002): The yeast ubiquitin protease, Ubp3p, promotes protein stability; in: Genetics, No. 3 (Vol. 162), pp. 1079-89.

Brocard, C., Kragler, F., Simon, M. M., Schuster, T., Hartig, A. (1994): The tetratricopeptide repeat-domain of the PAS10 protein of *Saccharomyces cerevisiae* is essential for binding the peroxisomal targeting signal-SKL; in: Biochemical and biophysical research communications, No. 3 (Vol. 204), pp. 1016-22.

Buhler, M., Gasser, S. M. (2009): Silent chromatin at the middle and ends: lessons from yeasts; in: EMBO J, No. 15 (Vol. 28), pp. 2149-2161.

Bushman, F. D. (1993): The bacteriophage 434 right operator. Roles of O_R1, O_R2 and O_R3; in: Journal of molecular biology, No. 1 (Vol. 230), pp. 28-40.

Calam, E., Porte, S., Fernandez, M. R., Farres, J., Pares, X., Biosca, J. A. (2013): Biocatalytic production of alpha-hydroxy ketones and vicinal diols by yeast and human aldo-keto reductases; in: Chemico-biological interactions, No. 1-3 (Vol. 202), pp. 195-203.

Cappellaro, C., Baldermann, C., Rachel, R., Tanner, W. (1994): Mating type-specific cell-cell recognition of *Saccharomyces cerevisiae*: cell wall attachment and active sites of a- and alpha-agglutinin; in: The EMBO journal, No. 20 (Vol. 13), pp. 4737-44.

Castillon, G. A., Watanabe, R., Taylor, M., Schwabe, T. M. E., Riezman, H. (2009): Concentration of GPI-anchored proteins upon ER exit in yeast; in: Traffic, No. 2 (Vol. 10), pp. 186-200.

Chen, L., Wang, F., Wang, X., Liu, Y. (2013): Robust one-tube Omega-PCR strategy accelerates precise sequence modification of plasmids for functional genomics; in: Plant and cell physiology, No. 4 (Vol. 54), pp. 634-42.

Cheng, C., Huang, Y. L., Yang, S. T. (1997): A novel feeding strategy for enhanced plasmid stability and protein production in recombinant yeast fedbatch fermentation; in: Biotechnology and bioengineering, No. 1 (Vol. 56), pp. 23-31.

Corbacho, I., Teixido, F., Velazquez, R., Hernandez, L. M., Olivero, I. (2011): Standard YPD, even supplemented with extra nutrients, does not always compensate growth defects of *Saccharomyces cerevisiae* auxotrophic strains; in: Antonie van Leeuwenhoek, No. 3 (Vol. 99), pp. 591-600.

Cormack, B. P., Bertram, G., Egerton, M., Gow, N. A., Falkow, S., Brown, A. J. (1997): Yeast-enhanced green fluorescent protein (yEGFP) a reporter of gene expression in *Candida albicans*; in: Microbiology, No.2 (Vol. 143), pp. 303-11.

Crabtree, G. R., Graef, I. A. (2008): Bursting into the nucleus; in: Sci Signal, No. 51 (Vol. 1), pp. 54-61.

Cubitt, A. B., Heim, R., Adams, S. R., Boyd, A. E., Gross, L. A., Tsien, R. Y. (1995): Understanding, improving and using green fluorescent proteins; in: Trends Biochem Sci, No. 11 (Vol. 20), pp. 448-55.

Cui, H., Ghosh, S. K., Jayaram, M. (2009): The selfish yeast plasmid uses the nuclear motor Kip1p but not Cin8p for its localization and equal segregation; in: J Cell Biol, No. 2 (Vol. 185), pp. 251-64.

Dahm, P., Jennewein, S. (2010): Introduction of the early pathway to taxol biosynthesis in yeast by means of biosynthetic gene cluster construction using SOE-PCR and homologous recombination; in: Methods in molecular biology, (Vol. 643), pp. 145-63.

De Wulf, P., Brambilla, L., Vanoni, M., Porro, D., Alberghina, L. (2000): Real-time flow cytometric quantification of GFP expression and Gfp-fluorescence generation in *Saccharomyces cerevisiae*; in: J Microbiol Methods, No. 1 (Vol. 42), pp. 57-64.

DeLisa, M. P., Li, J., Rao, G., Weigand, W. A., Bentley, W. E. (1999): Monitoring GFP-operon fusion protein expression during high cell density cultivation of *Escherichia coli* using an on-line optical sensor; in: Biotechnol Bioeng, No. 1 (Vol. 65), pp. 54-64.

Demple, B., Harrison, L., Wilson, D. M., Bennett, R. A., Takagi, T., Ascione, A. G. (1997): Regulation of eukaryotic abasic endonucleases and their role in genetic stability; in: Environmental health perspectives, No. 4 (Vol. 105), pp. 931-4.

Dodd, I. B., Shearwin, K. E., Egan, J. B. (2005): Revisited gene regulation in bacteriophage lambda; in: Current opinion in genetics and development, No. 2 (Vol. 15), pp. 145-52.

Dodd, I. B., Shearwin, K.E., Perkins, A.J., Burr, T., Hochschild, A., Egan, J.B. (2004): Cooperativity in long-range gene regulation by the λ CI repressor; in: Genes & Development, (Vol. 18), pp. 344-54.

Doherty, G. P., Bailey, K., Lewis, P. J. (2010): Stage-specific fluorescence intensity of GFP and mCherry during sporulation in *Bacillus Subtilis*; in: BMC research notes, (Vol. 3), p. 303.

Drubin, D. A., Way, J. C., Silver, P. A. (2007): Designing biological systems; in: Genes Dev, No. 3 (Vol. 21), pp. 242-54.

Dujon, B. (1996): The yeast genome project: what did we learn?; in: Trends in genetics, No. 7 (Vol. 12), pp. 263-70.

Dutreix, M., Bailone, A., Devoret, R. (1985): Efficiency of induction of prophage lambda mutants as a function of recA alleles; in: Journal of bacteriology, No. 3 (Vol. 161), pp. 1080-5.

Emanuelsson, O., Brunak, S., von Heijne, G., Nielsen, H. (2007): Locating proteins in the cell using TargetP, SignalP and related tools; in: Nature protocols, No. 4 (Vol. 2), pp. 953-71.

Falcon, A. A., Aris, J. P. (2003): Plasmid accumulation reduces life span in *Saccharomyces cerevisiae*; in: J Biol Chem, No. 43 (Vol. 278), pp. 41607-17.

Falcon, A. A., Rios, N., Aris, J. P. (2005): 2-micron circle plasmids do not reduce yeast life span; in: FEMS Microbiol Lett, No. 2 (Vol. 250), pp. 245-51.

Flagfeldt, D. B., Siewers, V., Huang, L., Nielsen, J. (2009): Characterization of chromosomal integration sites for heterologous gene expression in *Saccharomyces cerevisiae*; in: Yeast, No. 10 (Vol. 26), pp. 545-51.

Flashman, S. M. (1978): Mutational analysis of the operators of bacteriophage lambda; in: Molecular and general genetics, No. 1 (Vol. 166), pp. 61-73.

Frank, A. C., Wolfe, K. H. (2009): Evolutionary capture of viral and plasmid DNA by yeast nuclear chromosomes; in: Eukaryot Cell, No. 10 (Vol. 8), pp. 1521-31.

Fujimura, H., Sakuma, Y., Weimer, T. (1996): A quantitative method for the determination of plasmid copy number in recombinant yeast; in: Biologicals, No. 24 (Vol. 1), pp. 67-70.

Futcher, A. B., Cox, B. S. (1984): Copy number and the stability of 2-micron circle-based artificial plasmids of *Saccharomyces cerevisiae*; in: J Bacteriol, No. 1 (Vol. 157), pp. 283-90.

Gari, E., Piedrafita, L., Aldea, M., Herrero, E. (1997): A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in *Saccharomyces cerevisiae*; in: Yeast, No. 9 (Vol. 13), pp. 837-48.

Gedeon, T., Mischaikow, K., Patterson, K., Traldi, E. (2008): Binding cooperativity in phage lambda is not sufficient to produce an effective switch; in: Biophysical journal, No. 9 (Vol. 94), pp. 3384-92.

Ghosh, S. K., Hajra, S., Paek, A., Jayaram, M. (2006): Mechanisms for chromosome and plasmid segregation; in: Annu Rev Biochem, (Vol. 75), pp. 211-41.

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., Arkin, A. P., Astromoff, A., El-Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K., Flaherty, P., Foury, F., Garfinkel, D. J., Gerstein, M., Gotte, D., Guldener, U., Hegemann, J. H., Hempel, S., Herman, (2002): Functional profiling of the *Saccharomyces cerevisiae* genome; in: Nature, No. 6896 (Vol. 418), pp. 387-91.

Gimble, F. S., Sauer, R. T. (1985): Mutations in bacteriophage lambda repressor that prevent RecA-mediated cleavage; in: Journal of bacteriology, No. 1 (Vol. 162), pp. 147-54.

Goll, J., Uetz, P. (2006): The elusive yeast interactome; in: Genome biology, No. 6 (Vol. 7), p. 223.

Görgens, J. F., van Zyl, W. H., Knoetze, J. H., Hahn-Hagerdal, B. (2001): The metabolic burden of the *PGK1* and *ADH2* promoter systems for heterologous xylanase production by *Saccharomyces cerevisiae* in defined medium; in: Biotechnology and bioengineering, No. 3 (Vol. 73), pp. 238-45.

Gottesfeld, J. M., Neely, L., Trauger, J. W., Baird, E. E., Dervan, P. B. (1997): Regulation of gene expression by small molecules; in: Nature, No. 6629 (Vol. 387), pp. 202-5.

Granek, J. A., Magwene, P. M. (2010): Environmental and genetic determinants of colony morphology in yeast; in: PLoS Genet, No. 1 (Vol. 6), doi:10.1371/journal.pgen.1000823.

Greer, L. F. 3th, Szalay, A. A. (2002): Imaging of light emission from the expression of luciferases in living cells and organisms: a review; in: Luminescence, No. 1 (Vol. 17), pp. 43-74.

Griffith, J., Hochschild, A., Ptashne, M. (1986): DNA loops induced by cooperative binding of lambda repressor; in: Nature, No. 6081 (Vol. 322), pp. 750-2.

Grilly, C., Stricker, J., Pang, W. L., Bennett, M. R., Hasty, J. (2007): A synthetic gene network for tuning protein degradation in *Saccharomyces cerevisiae*; in: Molecular systems biology, (Vol. 3), pp. 127.

Gunge, N. (1983): Yeast DNA plasmids; in: Ann Rev Microbiol, (Vol 37), pp. 253-76.

Gupta, J. C., Mukherjee, K. J. (2001): Stable maintenance of plasmid in continuous culture of yeast under non-selective conditions; in: J Biosci Bioeng, No. 4 (Vol. 92), pp. 317-23.

Gupta, J. C., Mukherjee, K. J. (2002): Stability studies of recombinant *Saccharomyces cerevisiae* in the presence of varying selection pressure; in: Biotechnol Bioeng, No. 5 (Vol. 78), pp. 475-88.

Hackett, E. A., Esch, R. K, Maleri, S., Errede, B. (2006): A family of destabilized cyan fluorescent proteins as transcriptional reporters in *S. cerevisiae*; in: *Yeast* (Vol. 23), pp. 333-49.

Harrison, E., Koufopanou, V., Burt, A., MacLean, R. C. (2012): The cost of copy number in a selfish genetic element: the 2µm plasmid of *Saccharomyces cerevisiae*; in: Journal of evolutionary biology, No. 11 (Vol. 25), pp. 2348-56.

Hashimoto, Y., Koyabu, N., Imoto, T. (1998): Effects of signal sequences on the secretion of hen lysozyme by yeast: construction of four secretion cassette vectors; in: Protein engineering, No. 2 (Vol. 11), pp. 75-7.

Hasty, J., Pradines, J., Dolnik, M., Collins, J. J. (2000): Noise-based switches and amplifiers for gene expression; in: Proc Natl Acad Sci, No. 5 (Vol. 97), pp. 2075-80.

Hegemann, J. H., Klein, S., Heck, S., Guldener, U., Niedenthal, R. K., Fleig, U. (1999): A fast method to diagnose chromosome and plasmid loss in *Saccharomyces cerevisiae* strains; in: Yeast, No. 10B (Vol. 15), pp. 1009-19.

Hensing, M. C., Rouwenhorst, R. J., Heijnen, J. J., van Dijken, J. P., Pronk, J. T. (1995): Physiological and technological aspects of large-scale heterologous-protein production with yeasts; in: Antonie van Leeuwenhoek, No. 3 (Vol. 67), pp. 261-79.

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., Pease L. R. (1989): Site-directed mutagenesis by overlap extension using the polymerase chain reaction; in: Gene, (Vol. 77), pp. 51-59.

Hochschild, A., Lewis, M. (2009): The bacteriophage lambda CI protein finds an asymmetric solution; in: Current opinion in structural biology, No. 1 (Vol. 19), pp. 79-86.

Hochschild, A., Ptashne, M. (1986): Cooperative binding of lambda repressors to sites separated by integral turns of the DNA helix; in: Cell, No. 5 (Vol. 44), pp. 681-7.

Hong, K. K., Nielsen, J. (2012): Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries; in: Cell. Mol. Life Sci. (Vol. 69), pp. 2671-90.

Huang, D., Shusta, E. V. (2005): Secretion and surface display of green fluorescent protein using the yeast *Saccharomyces cerevisiae*; in: Biotechnology progress, No. 2 (Vol. 21), pp. 349-57.

Huh, W., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S., O'Shea, E. K. (2003): Global analysis of protein localization in budding yeast; in: Nature, No. 6959 (Vol. 425), pp. 686-91.

Humayun, Z., Jeffrey, A., Ptashne, M. (1977): Completed DNA sequences and organization of repressor-binding sites in the operators of phage lambda; in: Journal of molecular biology, No. 2 (Vol. 112), pp. 265-77.

Ignea, C., Cvetkovic, I., Loupassaki, S., Kefalas, P., Johnson, C. B., Kampranis, S. C., Makris, A. M. (2011): Improving yeast strains using recyclable integration cassettes, for the production of plant terpenoids; in: Microbial cell factories, No. 4 (Vol. 10), pp. 1-18.

Impoolsup A.; Caunt P.; Greenfield P.F. (1989): Stabilisation of a recombinant yeast plasmid in nonselective medium by cyclic growth rate changes; No. 9 (Vol. 2), pp. 605-608.

Jayaram, M., Mehta, S., Uzri, D., Velmurugan, P. (2004): Segregation of the yeast plasmid: similarities and contrasts with bacterial plasmid partitioning; in: Plasmid, No. 3 (Vol. 51), pp. 162-78.

Jones, J. J., Bridges, A. M., Fosberry, A. P., Gardner, S., Lowers, R. R., Newby, R. R., James, P. J., Hall, R. M., Jenkins, O. (2004): Potential of real-time measurement of GFP-fusion proteins; in: J Biotechnol, No. 1-2 (Vol. 109), pp. 201-11.

Jordan, S. R., Pabo, C. O. (1988): Structure of the lambda complex at 2.5 A resolution: details of the repressor-operator interactions; in: Science, No. 4880 Vol. (242), pp. 893-9.

Kachroo, A. H., Jayaram, M., Rowley, P. A. (2009): Metabolic engineering without plasmids; in: Nat Biotechnol, No. 8 (Vol. 27), pp. 729-31.

Kaern, M., Blake, W. J., Collins, J. J. (2003): The engineering of gene regulatory networks; in: Annual review of biomedical engineering, (Vol. 5), pp. 179-206.

Karim, A. S., Curran, K. A., Alper, H. S. (2013): Characterization of plasmid burden and copy number in *Saccharomyces cerevisiae* for optimization of metabolic engineering applications; in: FEMS yeast research, No. 1 (Vol. 13), pp. 107-16.

Kass, E. M., Jasin, M. (2010): Collaboration and competition between DNA double-strand break repair pathways; in: FEBS Lett, No. 17 (Vol. 584), pp. 3703-8.

Keppler-Ross, S., Noffz, C., Dean, N. (2008): A new purple fluorescent color marker for genetic studies in *Saccharomyces cerevisiae* and *Candida albicans*; in: Genetics, No. 1 (Vol. 179), pp. 705-10.

Kilonzo, P. M., Margaritis, A., Bergougnou, M. A. (2009): Plasmid stability and kinetics of continuous production of glucoamylase by recombinant *Saccharomyces cerevisiae* in an airlift bioreactor; in: Journal of industrial microbiology and biotechnology, No. 9 (Vol. 36), pp. 1157-69.

Kim, J. H., Johannes, L., Goud, B., Antony, C., Lingwood, C. A., Daneman, R., Grinstein, S. (1998): Noninvasive measurement of the pH of the endoplasmic reticulum at rest and during calcium release; in: Proc Natl Acad Sci, No. 95 (Vol. 6), pp. 2997-3002.

Kozak, M. (1999): Initiation of translation in prokaryotes and eukaryotes; in: Gene, No. 2 (Vol. 234), pp. 187-208.

Kozak, M. (2007): Lessons (not) learned from mistakes about translation; in: Gene, No. 1-2 (Vol. 403), pp. 194-203.

Kroll, J., Steinle, A., Reichelt, R., Ewering, C., Steinbuchel, A. (2009): Establishment of a novel anabolism-based addiction system with an artificially introduced mevalonate pathway: complete stabilization of plasmids as universal application in white biotechnology; in: Metabolic engineering, No. 3 (Vol. 11), pp. 168-77.

Lange, N., Steinbuchel, A. (2011): beta-Carotene production by *Saccharomyces cerevisiae* with regard to plasmid stability and culture media; in: Applied microbiology and biotechnology, No. 6 (Vol. 91), pp. 1611-22.

Leite, F. C. B., Dos Anjos, R. S. G., Basilio, A. C. M., Leal, G. F. C., Simoes, D. A., de Morais, M. A. J. (2013): Construction of integrative plasmids suitable for genetic modification of industrial strains of *Saccharomyces cerevisiae*; in: Plasmid, No. 1 (Vol. 69), pp. 114-7.

Leskinen, P., Virta, M., Karp, M. (2003): One-step measurement of firefly luciferase activity in yeast; in: Yeast, No. 13 (Vol. 20), pp. 1109-13.

Lewis, M. (2005): The lac repressor; in: C R Biol, No. 328 (Vol. 6), pp. 521-48.

Lewis, M. (2011): A tale of two repressors; in: Journal of molecular biology, No. 1 (Vol. 409), pp. 14-27.

Li, J., Shu, W., VanDusen, W. J., Schultz, L. D., George, H. A., Herber,W. K., Chae, H. J., Rao, G., Bentley, W. E. (2000): Green fluorescent protein in *Saccharomyces cerevisiae*: Real-time studies of the GAL1 promoter; in: Biotechnology and bioengineering, No. 2 (Vol. 70), pp. 187–196.

Li, J., Xu, H., Bentley, W. E., Rao, G. (2002): Impediments to secretion of green fluorescent protein and its fusion from *Saccharomyces cerevisiae*; in: Biotechnology progress, No. 4 (Vol. 18), pp. 831-8.

Liebesny, P., Goyal, S., Dunlap, D., Family, F., Finzi, L. (2010): Determination of the number of proteins bound non-specifically to DNA; in: J Phys condens matter, No. 41 (Vol. 22), pp. 1-12.

Little, J. W., Michalowski, C. B. (2010): Stability and Instability in the lysogenic state of phage lambda; in: Journal of Bacteriology, No. 22 (Vol. 192), pp. 6064-76.

Liu, Y., Ma, C., Jayaram, M. (2013): Co-segregation of yeast plasmid sisters under monopolin-directed mitosis suggests association of plasmid sisters with sister chromatids; in: Nucleic acids research, No. 7 (Vol. 41), pp. 4144-58.

Loew, R., Heinz, N., Hampf, M., Bujard, H., Gossen, M. (2010): Improved Tet-responsive promoters with minimized background expression; in: BMC Biotechnol, No. 81 (Vol. 10), pp. 13-15.

Longo, D., Hasty, J. (2006): Dynamics of single-cell gene expression; in: Molecular systems biology, (Vol. 2), p. 64.

Lu, C., Jeffries, T. (2007): Shuffling of promoters for multiple genes to optimize xylose fermentation in an engineered *Saccharomyces cerevisiae* strain; in: Appl Environ Microbiol, No. 19 (Vol. 73), pp. 6072-7.

Lu-Chau, T. A., Guillan, A., Nunez, M. J., Roca, E., Lema, J. M. (2004): Anaerobic and aerobic continuous cultures of *Saccharomyces cerevisiae*: comparison of plasmid stability and *EXG1* gene expression; in: Bioprocess and biosystems engineering, No. 3 (Vol. 26), pp. 159-63.

Maniatis, T., Ptashne, M., Backman, K., Kield, D., Flashman, S., Jeffrey, A., Maurer, R. (1975): Recognition sequences of repressor and polymerase in the operators of bacteriophage lambda; in: Cell, No. 2 (Vol. 5), pp. 109-13.

Mattanovich, D., Gasser, B., Hohenblum, H., Sauer, M. (2004): Stress in recombinant protein producing yeasts; in: J Biotechnol, No. 1-3 (Vol. 113), pp. 121-35.

Maya, D., Quintero, M. J., de la Cruz Munoz-Centeno, M., Chavez, S. (2008): Systems for applied gene control in *Saccharomyces cerevisiae*; in: Biotechnol Lett, No. 6 (Vol. 30), pp. 979-87.

Meunier, J. R., Choder, M. (1999): *Saccharomyces cerevisiae* colony growth and ageing: biphasic growth accompanied by changes in gene expression; in: Yeast, No. 12 (Vol. 15), pp. 1159-69.

Mnaimneh, S., Davierwala, A. P., Haynes, J., Moffat, J., Peng, W., Zhang, W., Yang, X., Pootoolal, J., Chua, G., Lopez, A., Trochesset, M., Morse, D., Krogan, N. J., Hiley, S. L., Li, Z., Morris, Q., Grigull, J., Mitsakakis, N., Roberts, C. J., Greenblatt, J. F., Boone, C., Kaiser, C. A., Andrews, B. J., Hughes, T. R. (2004): Exploration of essential gene functions via titratable promoter alleles; in: Cell, No. 1 (Vol. 118), pp. 31-44.

Monfort, A., Finger, S., Sanz, P., Prieto, J. A. (1999): Evaluation of different promoters for the efficient production of heterologous proteins in baker's yeast; in: Biotechnology letters, No. 3 (Vol. 21), pp. 225-229.

Morelli, M. J., Ten Wolde, P. R., Allen, R. J. (2009): DNA looping provides stability and robustness to the bacteriophage lambda switch; in: Proc Natl Acad Sci, No. 20 (Vol. 106), pp. 8101-6.

Mumberg, D., Muller, R., Funk, M. (1995): Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds; in: Gene, No. 1 (Vol. 156), pp. 119-22.

Nacken, V., Achstetter, T., Degryse, E. (1996): Probing the limits of expression levels by varying promoter strength and plasmid copy number in *Saccharomyces cerevisiae*; in: Gene, No. 1-2 (Vol. 175), pp. 253-60.

Nelson H. C. M., Sauer, R. T. (1985): Lambda repressor mutations that increase the affinity and specificity of operator binding ; in: Cell, (Vol. 42), pp. 549-58.

Nelson, M., Silver, P. (1989): Context affects nuclear protein localization in *Saccharomyces cerevisiae*; in: Molecular and cellular biology, No. 2 (Vol. 9), pp. 384-9.

Nevoigt, E., Kohnke, J., Fischer, C. R., Alper, H., Stahl, U., Stephanopoulos, G. (2006): Engineering of promoter replacement cassettes for fine-tuning of gene expression in *Saccharomyces cerevisiae*; in: Appl Environ Microbiol, No. 8 (Vol. 72), pp. 5266-73.

Newstead, S., Kim, H., von Heijne, G., Iwata, S., Drew, D. (2007): High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in *Saccharomyces cerevisiae*; in: Proc Natl Acad Sci, No. 35 (Vol. 104), pp. 13936-41.

Ogawa, T., Yu, X., Shinohara, A., Egelman, E. H. (1993): Similarity of the yeast RAD51 filament to the bacterial RecA filament, in: Science, No. 5103 (Vol. 259), pp. 1896-9.

Ogden, J. E., Stanway, C., Kim, S., Mellor, J., Kingsman, A.J., Kingsman, S. M. (1986): Efficient expression of the *Saccharomyces cerevisiae PGK* gene depends on an upstream activation sequence but does not require TATA sequences. Mol Cell Biol. No.6 (Vol. 12), pp. 4335-43.

Oldenburg, K. R., Vo, K. T., Michaelis, S., Paddon, C. (1997): Recombination-mediated PCR-directed plasmid construction in vivo in yeast; in: Nucleic acids research, No. 2 (Vol. 25), pp. 451-2.

Oliveira, P. H., Prather, K. J., Prazeres, D. M. F., Monteiro, G. A. (2009): Structural instability of plasmid biopharmaceuticals: challenges and implications; in: Trends Biotechnol, No. 9 (Vol. 27), pp. 503-11.

Orij, R., Postmus, J., Ter Beek, A., Brul, S., Smits, G. J. (2009): In vivo measurement of cytosolic and mitochondrial pH using a pH-sensitive GFP derivative in *Saccharomyces cerevisiae* reveals a relation between intracellular pH and growth; in: Microbiology No. 155 (Vol. 1), pp. 268-78.

Ormö, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y., Remington, S. J. (1996): Crystal structure of the *Aequorea victoria* green fluorescent protein; in: Science, No. 5280 (Vol. 273), pp. 1392-5.

Pagny, S., Lerouge, P., Faye, L., Gomord, V. (1999): Signals and mechanisms for protein retention in the endoplasmic reticulum; in: Journal of experimental botany, No. 331 (Vol. 50), pp. 157-164.

Pal, A., Chattopadhyaya, R. (2009): RecA-mediated cleavage of lambda CI repressor accepts repressor dimers: probable role of prolyl cis-trans isomerization and catalytic involvement of H^{163}, K^{177}, and K^{232} of RecA; in: Journal of biomolecular structure and dynamics, No. 2 (Vol. 27), pp. 221-33.

Papanikou, E., Glick, B. S. (2009): The yeast Golgi apparatus: insights and mysteries; in: FEBS Lett, No. 23 (Vol. 583), pp. 3746-51.

Partow, S., Siewers, V., Bjørn, S., Nielsen, J., Maury, J. (2012): Characterization of different promoters for designing a new expression vector in *Saccharomyces cerevisiae*; in: Yeast, (Vol. 27.), pp. 955-64.

Peña-Castillo, L., Hughes, T. R. (2007): Why are there still over 1000 uncharacterized yeast genes?; in: Genetics, No. 176 (Vol.1), pp. 7-14.

Perktold, A., Zechmann, B., Daum, G., Zellnig, G. (2007): Organelle association visualized by three-dimensional ultrastructural imaging of the yeast cell; in: FEMS Yeast Res, No. 7 (Vol. 4), pp. 629-38.

Pitson, D. W., Patterson, G. H., Knobel, S. M. (1999): Quantitative Imaging of the Green Fluorescent Protein (GFP); in: Methods in Cell Biology, edited by Sullivan K. F., Kay S. K., Vol 58, p46.

Pittet, M., Conzelmann, A. (2007): Biosynthesis and function of GPI proteins in the yeast *Saccharomyces cerevisiae*; in: Biochim Biophys Acta, No. 3 (Vol. 1771), pp. 405-20.

Prein, B., Natter, K., Kohlwein, S. D. (2000): A novel strategy for constructing N-terminal chromosomal fusions to green fluorescent protein in the yeast *Saccharomyces cerevisiae*; in: FEBS Lett, No. 1 (Vol. 485), pp. 29-34.

Premsler, T., Zahedi, R. P., Lewandrowski, U., Sickmann, A. (2009): Recent advances in yeast organelle and membrane proteomics; in: Proteomics, No. 20 (Vol. 9), pp. 4731-43.

Pronk, J. T. (2002): Auxotrophic yeast strains in fundamental and applied research; in: Appl Environ Microbiol, No. 5 (Vol. 68), pp. 2095-100.

Ptashne, M. (2005): Regulation of transcription: from lambda to eukaryotes; in: Trends Biochem Sci, No. 6 (Vol. 30), pp. 275-9.

Ptashne, M. (2011): Principles of a switch; in: Nature chemical biology, No. 8 (Vol. 7), pp. 484-7.

Randers-Eichhorn, L., Albano, C. R., Sipior, J., Bentley, W. E., Rao, G. (1997): On-line green fluorescent protein sensor with LED excitation; in: Biotechnol Bioeng, No. 6 (Vol. 55), pp. 921-6.

Rando, O. J., Chang, H. Y. (2009): Genome-wide views of chromatin structure; in: Annu Rev Biochem, (Vol. 78), pp. 245-71.

Remington, S. J. (2006): Fluorescent proteins: maturation, photochemistry and photophysics; in: Current opinion in structural biology, No. 6 (Vol. 16), pp. 714-21.

Robertson, J. B., Stowers, C. C., Boczko, E., Johnson, C. H. (2008): Real-time luminescence monitoring of cell-cycle and respiratory oscillations in yeast; in: Proc Natl Acad Sci, No. 46 (Vol. 105), pp. 17988-93.

Romanos, M. A., Scorer, C. A., Clare, J. J. (1992): Foreign gene expression in yeast: a review; in: Yeast, No. 6 (Vol. 8), pp. 423-88.

Russo, P., Sherman, F. (1989): Transcription terminates near the poly(A) site in the *CYC*1 gene of the yeast *Saccharomyces cerevisiae*; in: Proc Natl Acad Sci, No. 21 (Vol. 86), pp. 8348-52.

Saiz L., Vilar, J. M. G. (2006): DNA looping: the consequences and its control; in: Current Opinion in Structural Biology, (Vol. 16), pp. 344-350.

Sambrook, J., Russell D.W. (2001): Molecular Cloning: A Laboratory Manual, 3rd ed., Cold Spring Harbor Laboratory Press.

Sayut, D. J., Kambam, P. K. R., Sun, L. (2007): Engineering and applications of genetic circuits; in: Molecular bioSystems, No. 12 (Vol. 3), pp. 835-40.

Schmitt, M. J., Tipper, D. J. (1995): Sequence of the M28 dsRNA: preprotoxin is processed to an alpha/beta heterodimeric protein toxin; in: Virology, No. 2 (Vol. 213), pp. 341-51.

Scott-Drew, S., Murray, J. A. (1998): Localisation and interaction of the protein components of the yeast 2 μm circle plasmid partitioning system suggest a mechanism for plasmid inheritance; in: J Cell Sci, No. 13 (Vol. 111), pp. 1779-89.

Seksek, O., Bolard, J. (1996): Nuclear pH gradient in mammalian cells revealed by laser microspectrofluorimetry; in: J Cell Sci, No. 109 (Vol. 1), pp. 257-62.

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N. G., Palmer, A. E., Tsien, R. Y. (2004): Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein; in: Nature biotechnology, No. 12 (Vol. 22), pp. 1567-72.

Sherman, F. (2002): Getting started with yeast; in: Methods in enzymology, (Vol. 350), pp. 3-41.

Shibasaki, S., Tanaka, A., Ueda, M. (2003): Development of combinatorial bioengineering using yeast cell surface display--order-made design of cell and protein for bio-monitoring; in: Biosens Bioelectron, No. 2 (Vol. 19), pp. 123-30.

Shibasaki, S., Ueda, M., Iizuka, T., Hirayama, M., Ikeda, Y., Kamasawa, N., Osumi, M., Tanaka, A. (2001): Quantitative evaluation of the enhanced green fluorescent protein displayed on the cell surface of *Saccharomyces cerevisiae* by fluorometric and confocal laser scanning microscopic analyses; in: Appl Microbiol Biotechnol, No. 4 (Vol. 55), pp. 471-5.

Shulga, N., Roberts, P., Gu, Z., Spitz, L., Tabb, M. M., Nomura, M., Goldfarb, D. S. (1996): In vivo nuclear transport kinetics in *Saccharomyces cerevisiae*: a role for heat shock protein 70 during targeting and translocation; in: The Journal of cell biology, No. 2 (Vol. 135), pp. 329-39.

Sigler, K., Knotkova, A., Kotyk, A. (1981): Factors governing substrate-induced generation and extrusion of protons in the yeast *Saccharomyces cerevisiae*; in: Biochim Biophys Acta, No. 3 (Vol. 643), pp. 572-82.

Simon, J. R., Moore, P. D. (1987): Homologous recombination between single-stranded DNA and chromosomal genes in *Saccharomyces cerevisiae*; in: Mol Cell Biol, No. 7 (Vol. 7), pp. 2329-34.

Smits, G. J., Schenkman, L. R., Brul, S., Pringle, J. R., Klis, F. M. (2006): Role of cell cycle-regulated expression in the localized incorporation of cell wall proteins in yeast; in: Mol Biol Cell, No. 7 (Vol. 17), pp. 3267-80.

Snaith, H. A., Anders, A., Samejima, I., Sawin, K. E. (2010): New and old reagents for fluorescent protein tagging of microtubules in fission yeast; experimental and critical evaluation; in: Methods Cell Biol, (Vol. 97), pp. 147-72.

Snaith, H. A., Samejima, I., Sawin, K. E. (2005): Multistep and multimode cortical anchoring of tea1p at cell tips in fission yeast; in: EMBO J, No. 21 (Vol. 24), pp. 3690-9.

Speare, J. O., Offerdahl, D. K., Hasenkrug, A., Carmody, A. B., Baron, G. S. (2010): GPI anchoring facilitates propagation and spread of misfolded Sup35 aggregates in mammalian cells; in: The EMBO journal, No. 4 (Vol. 29), pp. 782-94.

Springer, M., Weissman, J. S., Kirschner, M. W. (2010): A general lack of compensation for gene dosage in yeast; in: Molecular systems biology, (Vol. 6), p. 368.

Sprinzak, D., Elowitz, M. B. (2005): Reconstruction of genetic circuits; in: Nature, No. 7067 (Vol. 438), pp. 443-8.

Stayrook, S., Jaru-Ampornpan, P., Ni, J., Hochschild, A., Lewis, M. (2008): Crystal structure of the lambda repressor and a model for pairwise cooperative operator binding; in: Nature, No. 7190 (Vol. 452), pp. 1022-5.

Stornaiuolo M, Lotti LV, Borgese N, Torrisi MR, Mottola G, Martire G, Bonatti S. (2003): KDEL and KKXX retrieval signals appended to the same reporter protein determine different trafficking between endoplasmic reticulum, intermediate compartment, and golgi complex; in: Mol. Biol. Cell, No. 14, pp. 889–902.

Subramani, S., Koller, A., Snyder, W. B. (2000): Import of peroxisomal matrix and membrane proteins; in: Annu. Rev. Biochem. No. 69, pp. 399–418.

Suk, K., Choi, J., Suzuki, Y., Ozturk, S. B., Mellor, J. C., Wong, K. H., MacKay, J. L., Gregory, R. I., Roth, F. P. (2011): Reconstitution of human RNA interference in budding yeast; in: Nucleic Acids Res, No. 39 (Vol. 7).

Tochigi, Y., Sato, N., Sahara, T., Wu, C., Saito, S., Irie, T., Fujibuchi, W., Goda, T., Yamaji, R., Ogawa, M., Ohmiya, Y., Ohgiya, S. (2010): Sensitive and convenient yeast reporter assay for high-throughput analysis by using a secretory luciferase from *Cypridina noctiluca*; in: Anal Chem, No.82 (Vol. 13), pp. 5768-76.

Tomlin, G. C., Wixon, J. L., Bolotin-Fukuhara, M., Oliver, S. G. (2001): A new family of yeast vectors and S288C-derived strains for the systematic analysis of gene function; in: Yeast, No. 18 (Vol. 6), pp. 563-75.

Tsien, R. Y. (1998): The green fluorescent protein; in: Annu Rev Biochem, (Vol. 67), pp. 509-44.

Uchida, M., Sun, Y., McDermott, G., Knoechel, C., Le Gros, M. A. , Parkinson, D., Drubin, D. G., Larabell, C. A. (2011): Quantitative analysis of yeast internal architecture using soft X-ray tomography; in: Yeast, Nr 28 (Vol. 3), pp. 227–236.

Ueda, M., Tanaka, A. (2000): Genetic immobilization of proteins on the yeast cell surface; in: Biotechnol Adv, No. 2 (Vol. 18), pp. 121-40.

Uffenbeck, S. R., Krebs, J. E. (2006): The role of chromatin structure in regulating stress-induced transcription in *Saccharomyces cerevisiae*; in: Biochem Cell Biol, No. 4 (Vol. 84), pp. 477-89.

Valli, M., Sauer, M., Branduardi, P., Borth, N., Porro, D., Mattanovich, D. (2005): Intracellular pH distribution in *Saccharomyces cerevisiae* cell populations, analyzed by flow cytometry; in: Appl Environ Microbiol, No. 71 (Vol. 3), pp. 1515-21.

Valli, M., Sauer, M., Branduardi, P., Borth, N., Porro, D., Mattanovich, D. (2006): Improvement of lactic acid production in *Saccharomyces cerevisiae* by cell sorting for high intracellular pH; in: Appl Environ Microbiol, No. 72 (Vol. 8), pp. 5492-9.

van der Aar, P. C., Lopes, T. S., Klootwijk, J., Groeneveld, P., Van Verseveld, H. W., Stouthamer, A. H. (1990): Consequences of phosphoglycerate kinase overproduction for the growth and physiology of *Saccharomyces cerevisiae*; in: Applied microbiology and biotechnology, No. 5 (Vol. 32), pp. 577-587.

van der Aar, P. C., van den Heuvel, J. J., Roling, W. F., Raue, H. A., Stouthamer, A. H., van Verseveld, H. W. (1992): Effects of phosphoglycerate kinase overproduction in *Saccharomyces cerevisiae* on the physiology and plasmid stability; in: Yeast, No. 1 (Vol. 8), pp. 47-55.

van der Klei, I. J., Veenhuis, M. (1997): Yeast peroxisomes: function and biogenesis of a versatile cell organelle; in: Trends in microbiology, No. 12 (Vol. 5), pp. 502-9.

van Roermund, C. W. T., de Jong, M., Ijlst, L., van Marle, J., Dansen, T. B., Wanders, R. J. A., Waterham, H. R. (2004): The peroxisomal lumen in *Saccharomyces cerevisiae* is alkaline; in: Journal of cell science, No. 18 (Vol. 117), pp. 4231-7.

Varshavsky, A. (1997): The N-end rule pathway of protein degradation; in: Genes to Cells, (Vol. 2), pp. 13-28.

Velmurugan, S., Mehta, S., Uzri, D., Jayaram, M. (2003): Stable propagation of 'selfish' genetic elements; in: J Biosci, No. 5 (Vol. 28), pp. 623-36.

Verma, H. K., Singh, J. (2012): New multi-purpose high copy number vector with greater mitotic stability for diverse applications in fission yeast Schizosaccharomyces pombe; in: Plasmid, No. 3 (Vol. 68), pp. 186-94.

Volkert, F. C., Wilson, D. W., Broach, J. R. (1989): Deoxyribonucleic acid plasmids in yeasts; in: Microbiol Rev, No. 3 (Vol. 53), pp. 299-317.

von Heijne, G. (1990): The signal peptide; in: The Journal of membrane biology, No. 3 (Vol. 115), pp. 195-201.

Waters, L. S., Storz, G. (2009): Regulatory RNAs in bacteria; in: Cell, No. 4 (Vol. 136), pp. 615-28.

Webster, C. I., Brammar, W. J. (1995): The bacteriophage 434 operator/repressor system in yeast; in: Microbiology, No. 9 (Vol. 141), pp. 2191-200.

Wedler, H., Wambutt, R. (1995): A temperature-sensitive lambda Cl repressor functions on a modified operator in yeast cells by masking the TATA element; in: Mot. Gen. Genet., (Vol. 248), pp. 499-505.

Welsh, S., Kay, S. A. (1997): Reporter gene expression for monitoring gene transfer; in: Curr Opin Biotechnol, No. 5 (Vol. 8), pp. 617-22.

Westfall, P. J., Pitera, D. J., Lenihan, J. R., Eng, D., Woolard, F. X., Regentin, R., Horning, T., Tsuruta, H., Melis, D. J., Owens, A., Fickes, S., Diola, D., Benjamin, K. R., Keasling, J. D., Leavell, M. D., McPhee, D. J., Renninger, N. S., Newman, J. D., Paddon, C. J. (2012): Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin; in: Proc Natl Acad Sci, No. 3 (Vol. 109), pp. 111-8.

Wheatley, R. W., Lo, S., Jancewicz, L. J., Dugdale, M. L., Huber, R. E. (2013): Structural explanation for allolactose (lac operon inducer) synthesis by lacZ β-galactosidase and the evolutionary relationship between allolactose synthesis and the lac repressor; in: J Biol Chem No. 288 (Vol. 18).

WHO Memorandum (1985): Hepatitis B vaccines prepared from yeast by recombinant DNA techniques: memorandum from a WHO meeting; in: Bulletin of the World Health Organization, No. 1 (Vol. 63), pp. 57-61.

Wickerham, L. J. (1946): A Critical Evaluation of the Nitrogen Assimilation Tests Commonly Used in the Classification of Yeasts; in: J Bacteriol, No. 3 (Vol. 52), pp. 293-301.

Wickner, R. B. (1996): Prions and RNA viruses of *Saccharomyces cerevisiae*; in: Annu Rev Genet, (Vol. 30), pp. 109-39.

Wilson, C. J., Zhan, H., Swint-Kruse, L., Matthews, K. S. (2007): The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding; in: Cellular and molecular life sciences, No. 1 (Vol. 64), pp. 3-16.

Yang, F., Moss, L. G., Phillips, G. N. J. (1996): The molecular structure of green fluorescent protein; in: Nat Biotechnol, No. 10 (Vol. 14), pp. 1246-51.

Ying, B., Ito, Y., Shimizu, Y., Yomo, T. (2010): Refined method for the genomic integration of complex synthetic circuits; in: Journal of bioscience and bioengineering, No. 5 (Vol. 110), pp. 529-36.

Zhang, Z., Moo-Young, M., Chisti, Y. (1996): Plasmid stability in recombinant *Saccharomyces cerevisiae*; in: Biotechnology advances, No. 4 (Vol. 14), pp. 401-35.

Zhang, Z., Scharer, J. M., Moo-Young, M. (1997): Plasmid instability kinetics in continuous culture of a recombinant *Saccharomyces cerevisiae* in airlift bioreactor; in: J Biotechnol, No. 1 (Vol. 55), pp. 31-41.

Zurla, C., Manzo, C., Dunlap, D., Lewis, D. E. A., Adhya, S., Finzi, L. (2009): Direct demonstration and quantification of long-range DNA looping by the lambda bacteriophage repressor; in: Nucleic acids research, No. 9 (Vol. 37), pp. 2789-95.

Websites: http://rulai.cshl.edu/cgi-bin/SCPD/getgene2?PGK1 (1.10.2013)

http://parts.igem.org/Part:BBa_J63008 (1.10.2013)

http://bionumbers.hms.harvard.edu/default.aspx (20.8.2013)

http://depatisnet.dpma.de (10.9.2013)

http://google.de/?tbm=pts (10.9.2013)

http://yeastgenome.org/cgi-bin/PATMATCH/nph-patmatch (15.10.2013)

http://toolkit.tuebingen.mpg.de/hhpred (10.1.2014)

http://yeastgfp.yeastgenome.org (20.8.2013)

7 APPENDIX

7.1 FLUORESCENCE DATA

Table A: Data for the promoter test

	4 h / rel. fluorescence			mean	standard deviation	8 h/ rel. fluorescence			mean	standard deviation	24 h/ rel. fluorescence			mean	standard deviation
PGK1 prom	17.47	17.64	17.38	17.50	0.13	15.70	16.34	15.70	15.91	0.37	17.30	16.36	16.94	16.87	0.47
ADH1 prom	9.74	10.28	9.74	9.92	0.32	6.20	6.27	6.54	6.34	0.18	0.00	0.00	0.00	0.00	0.00
ACT1 prom	7.42	6.34	6.37	6.71	0.62	6.90	7.28	7.49	7.22	0.30	5.90	5.63	4.89	5.47	0.52

Table B: Data for the difference between cytoplasm and peroxisome

Cultivation	cytoplasm/ rel. fluorescence			mean	standard deviation	peroxisome/ rel. fluorescence			mean	standard deviation	*p*-value
unshaken	15.93	14.78	13.73	14.81	1.10	16.79	16.59	15.85	16.41	0.49	11.2774%
shaken	18.12	18.46	18.78	18.45	0.33	17.98	18.82	18.32	18.37	0.42	80.5597%

Table C: Data for the difference between integrated and plasmid based expression cassettes

nPos	plasmid/ rel. fluorescence			mean	standard deviation	integrated/ rel. fluorescence			mean	standard deviation	*p*-value
1-2	6.07	8.14	6.86	7.02	1.04	7.01	6.14	6.79	6.64	0.45	60.8224%
1+3	7.24	5.49	6.02	6.25	0.90	8.07	7.20	6.39	7.22	0.84	24.4665%
1-3	6.37	6.93	5.83	6.38	0.55	6.38	7.42	7.48	7.09	0.62	20.9107%

Table D: Data for the relative fluorescence of the different nPos

nPos	rel. fluorescence			mean	standard deviation	*p*-value
-1	0.00	0.00	0.00	0.00	0.00	0.04%
1	15.04	15.90	14.99	15.31	0.42	27.48%
1ol	15.08	15.70	15.92	15.57	0.36	50.55%
2	13.93	14.77	15.12	14.61	0.50	5.95%
2ol	12.40	11.29	10.40	11.36	0.82	0.54%
-2	14.10	13.90	14.70	14.23	0.34	1.94%
-2ol	12.40	13.20	12.70	12.77	0.33	0.24%
-3	14.70	16.00	15.80	15.50	0.57	51.64%
-3ol	13.73	14.80	15.10	14.54	0.59	6.97%
3	14.34	13.92	13.45	13.90	0.36	1.09%
3ol	11.44	12.48	13.40	12.44	0.80	1.13%
1-3	7.70	5.84	6.12	6.55	0.82	0.06%
1+3	7.70	6.70	5.90	6.77	0.74	0.03%
1-2	6.10	7.27	6.59	6.65	0.48	0.00%
1+2	0.00	0.00	0.00	0.00	0.00	0.04%
TAT	7.16	6.54	7.18	6.96	0.30	0.01%
PGK1	16.05	16.34	15.23	15.87	0.47	100.00%
SY992	0.00	0.00	0.00	0.00	0.00	0.00

Table E: Data for the relative fluorescence of the different nPos in combination with JOE0046

nPos	**rel. fluorescence**			**mean**	**standard deviation**	**with JOE0046**			**mean**	**standard deviation**	***p*-value**
1	15.04	15.90	14.99	15.31	0.42	15.76	15.10	14.90	15.25	0.45	89.2533%
1ol	15.08	15.70	15.92	15.57	0.36	15.77	15.23	14.89	15.30	0.44	49.5725%
2	13.93	14.77	15.12	14.61	0.50	13.96	14.05	13.34	13.78	0.39	13.2989%
2ol	12.40	11.29	10.40	11.36	0.82	9.38	10.97	10.54	10.30	0.82	22.9743%
-2	14.10	13.90	14.70	14.23	0.34	24.29	23.26	22.56	23.37	0.87	0.0619%
-2ol	12.40	13.20	12.70	12.77	0.33	15.98	16.49	16.03	16.17	0.28	0.0527%
-3	14.70	16.00	15.80	15.50	0.57	25.73	26.17	24.56	25.49	0.83	0.0110%
-3ol	13.73	14.80	15.10	14.54	0.59	16.30	16.79	15.99	16.36	0.40	2.9083%
3	14.34	13.92	13.45	13.90	0.36	21.42	20.45	19.13	20.33	1.15	0.4988%
3ol	11.44	12.48	13.40	12.44	0.80	15.72	14.23	16.78	15.58	1.28	3.1092%
1-3	7.70	5.84	6.12	6.55	0.82	13.90	12.83	13.02	13.25	0.57	0.1647%
1+3	7.70	6.70	5.90	6.77	0.74	9.92	10.56	10.07	10.19	0.33	1.3809%
1-2	6.10	7.27	6.59	6.65	0.48	10.55	9.79	9.35	9.90	0.61	0.2644%
TAT	7.16	6.54	7.18	6.96	0.30	6.54	6.91	6.65	6.70	0.19	34.8977%
PGK1	16.05	16.34	15.23	15.87	0.47						

Table F: Data for the relative fluorescence of the different nPos in combination with JOE0056

nPos	**rel. fluorescence**			**mean**	**standard deviation**	**with JOE0056**			**mean**	**standard deviation**	***p*-value**
1	15.04	15.90	14.99	15.31	0.42	15.83	16.07	15.48	15.79	0.29	24.6500%
1ol	15.08	15.70	15.92	15.57	0.36	15.08	15.83	16.16	15.69	0.55	77.6814%
2	13.93	14.77	15.12	14.61	0.50	13.43	14.62	15.26	14.44	0.93	80.5160%
2ol	12.40	11.29	10.40	11.36	0.82	13.08	11.23	10.48	11.60	1.34	82.2331%
-2	14.10	13.90	14.70	14.23	0.34	14.47	13.84	14.20	14.17	0.32	84.3692%
-2ol	12.40	13.20	12.70	12.77	0.33	11.92	13.03	12.90	12.62	0.61	74.4316%
-3	14.70	16.00	15.80	15.50	0.57	17.89	17.38	16.34	17.20	0.79	4.9858%
-3ol	13.73	14.80	15.10	14.54	0.59	16.46	16.69	16.12	16.42	0.29	3.1849%
3	14.34	13.92	13.45	13.90	0.36	14.16	15.12	15.44	14.91	0.67	10.5897%
3ol	11.44	12.48	13.40	12.44	0.80	14.44	14.24	13.23	13.97	0.65	9.7508%
1-3	7.70	5.84	6.12	6.55	0.82	9.41	9.24	8.77	9.14	0.33	3.6028%
1+3	7.70	6.70	5.90	6.77	0.74	7.93	8.43	9.73	8.70	0.93	6.1159%
1-2	6.10	7.27	6.59	6.65	0.48	5.86	6.90	7.25	6.67	0.73	97.7155%
TAT	7.16	6.54	7.18	6.96	0.30	7.39	5.65	6.42	6.49	0.87	45.6088%
PGK1	16.05	16.34	15.23	15.87	0.47						

Table G: Data for the relative fluorescence of the different nPos in combination with JOE0079

nPos	rel. fluorescence			mean	standard deviation	with JOE0079			mean	standard deviation	*p*-value
1	15.04	15.90	14.99	15.31	0.42	1.20	0.00	0.90	0.70	0.62	0.0009%
1ol	15.08	15.70	15.92	15.57	0.36	0.00	0.00	1.00	0.33	0.58	0.0007%
2	13.93	14.77	15.12	14.61	0.50	16.32	17.01	15.30	16.21	0.86	6.4574%
2ol	12.40	11.29	10.40	11.36	0.82	13.00	15.70	16.10	14.93	1.69	4.5745%
-2	14.10	13.90	14.70	14.23	0.34	13.61	13.77	14.35	13.91	0.39	38.6015%
-2ol	12.40	13.20	12.70	12.77	0.33	12.44	13.17	13.11	12.91	0.41	69.5205%
-3	14.70	16.00	15.80	15.50	0.57	11.45	10.27	9.54	10.42	0.96	0.2542%
-3ol	13.73	14.80	15.10	14.54	0.59	11.50	10.11	11.96	11.19	0.96	1.0238%
3	14.34	13.92	13.45	13.90	0.36	13.76	12.55	12.42	12.91	0.74	13.2176%
3ol	11.44	12.48	13.40	12.44	0.80	13.40	11.90	11.50	12.27	1.00	84.0890%
1-3	7.70	5.84	6.12	6.55	0.82	0.00	0.90	0.00	0.30	0.52	0.2401%
1+3	7.70	6.70	5.90	6.77	0.74	0.00	1.30	0.00	0.43	0.75	0.0850%
1-2	6.10	7.27	6.59	6.65	0.48	0.00	1.40	0.00	0.47	0.81	0.0684%
TAT	7.16	6.54	7.18	6.96	0.30	1.52	0.00	1.83	1.12	0.98	0.4495%
PGK1	16.05	16.34	15.23	15.87	0.47						

Table H: Data for the relative fluorescence of the different nPos in combination with JOE0044

nPos	rel. fluorescence			mean	standard deviation	with JOE0044			mean	standard deviation	*p*-value
1	15.04	15.90	14.99	15.31	0.42	13.40	16.15	14.91	14.82	1.38	60.9498%
1ol	15.08	15.70	15.92	15.57	0.36	15.41	15.60	15.78	15.59	0.19	92.8466%
2	13.93	14.77	15.12	14.61	0.50	13.90	15.46	14.94	14.76	0.79	80.1469%
2ol	12.40	11.29	10.40	11.36	0.82	12.81	11.18	10.42	11.47	1.22	91.2187%
-2	14.10	13.90	14.70	14.23	0.34	14.52	13.89	14.72	14.38	0.43	70.1958%
-2ol	12.40	13.20	12.70	12.77	0.33	12.15	13.80	12.93	12.96	0.83	74.2847%
-3	14.70	16.00	15.80	15.50	0.57	14.51	15.69	16.08	15.43	0.82	91.5932%
-3ol	13.73	14.80	15.10	14.54	0.59	13.94	14.53	15.03	14.50	0.55	93.0684%
3	14.34	13.92	13.45	13.90	0.36	14.01	13.75	13.40	13.72	0.31	58.9528%
3ol	11.44	12.48	13.40	12.44	0.80	11.14	12.13	13.35	12.21	1.11	79.8581%
1-3	7.70	5.84	6.12	6.55	0.82	7.24	5.15	5.98	6.12	1.05	63.4592%
1+3	7.70	6.70	5.90	6.77	0.74	8.08	6.76	6.33	7.06	0.91	71.4918%
1-2	6.10	7.27	6.59	6.65	0.48	5.91	7.09	6.70	6.57	0.60	86.8006%
TAT	7.16	6.54	7.18	6.96	0.30	6.38	7.23	6.49	6.70	0.46	48.8858%
PGK1	16.05	16.34	15.23	15.87	0.47						

Table I: Data for the relative fluorescence of the different nPos in combination with JOE0080

nPos	**rel. fluorescence**			**mean**	**standard deviation**	**with JOE0080**			**mean**	**standard deviation**	***p*-value**
1	15.04	15.90	14.99	15.31	0.42	15.35	15.59	15.91	15.61	0.28	43.1535%
1ol	15.08	15.70	15.92	15.57	0.36	14.53	15.87	15.53	15.31	0.69	62.0921%
2	13.93	14.77	15.12	14.61	0.50	14.30	14.88	14.39	14.52	0.31	84.8062%
2ol	12.40	11.29	10.40	11.36	0.82	12.72	10.67	9.94	11.11	1.44	81.6902%
-2	14.10	13.90	14.70	14.23	0.34	13.19	12.99	15.09	13.76	1.16	55.7567%
-2ol	12.40	13.20	12.70	12.77	0.33	11.62	13.85	12.74	12.74	1.12	96.8747%
-3	14.70	16.00	15.80	15.50	0.57	14.69	15.69	15.88	15.42	0.64	88.9539%
-3ol	13.73	14.80	15.10	14.54	0.59	13.42	14.84	14.94	14.40	0.85	83.2762%
3	14.34	13.92	13.45	13.90	0.36	13.78	14.76	13.11	13.89	0.83	97.6323%
3ol	11.44	12.48	13.40	12.44	0.80	11.03	12.33	10.47	11.28	0.95	21.4887%
1-3	7.70	5.84	6.12	6.55	0.82	6.64	6.59	5.91	6.38	0.41	80.1780%
1+3	7.70	6.70	5.90	6.77	0.74	8.65	6.91	5.51	7.02	1.57	82.1184%
1-2	6.10	7.27	6.59	6.65	0.48	5.88	6.89	6.34	6.37	0.51	55.9383%
TAT	7.16	6.54	7.18	6.96	0.30	6.93	6.72	6.80	6.82	1.00	57.0676%
PGK1	16.05	16.34	15.23	15.87	0.47						

Table J: Data for the relative fluorescence of the different nPos in combination with JOE0092

nPos	rel. fluorescence			mean	standard deviation	with JOE0092			mean	standard deviation	*p*-value
1	15.04	15.90	14.99	15.31	0.42	1.40	0.00	0.00	0.47	0.81	0.0046%
1ol	15.08	15.70	15.92	15.57	0.36	0.00	0.00	0.00	0.00	0.00	0.0261%
2	13.93	14.77	15.12	14.61	0.50	6.10	4.87	5.28	5.42	0.62	0.0054%
2ol	12.40	11.29	10.40	11.36	0.82	12.37	12.76	11.08	12.07	0.88	41.1939%
-3	14.70	16.00	15.80	15.50	0.57	7.34	7.16	8.56	7.69	0.76	0.0205%
-3ol	13.73	14.80	15.10	14.54	0.59	9.82	8.73	7.02	8.53	1.41	0.7348%
3	14.34	13.92	13.45	13.90	0.36	10.10	11.24	12.94	11.43	1.43	8.4048%
3ol	11.44	12.48	13.40	12.44	0.80	9.97	10.41	11.16	10.51	0.60	5.5246%
PGK1	16.05	16.34	15.23	15.87	0.47						

Table K: Data for the relative fluorescence of the different nPos in combination with JOE0081

nPos	rel. fluorescence			mean	standard deviation	with JOE0081			mean	standard deviation	*p*-value
1	15.04	15.90	14.99	15.31	0.42	10.55	11.30	9.99	10.61	0.66	0.0837%
1ol	15.08	15.70	15.92	15.57	0.36	7.33	8.24	8.27	7.95	0.53	0.0058%
2	13.93	14.77	15.12	14.61	0.50	13.62	14.39	14.86	14.29	0.63	56.4411%
2ol	12.40	11.29	10.40	11.36	0.82	12.80	10.77	10.63	11.40	1.21	97.0858%
-3	14.70	16.00	15.80	15.50	0.57	15.92	14.90	16.17	15.66	0.67	78.7083%
-3ol	13.73	14.80	15.10	14.54	0.59	13.80	14.78	14.72	14.43	0.55	83.9692%
PGK1	16.05	16.34	15.23	15.87	0.47						

Table L: Data for the relative fluorescence of the different nPos in combination with JOE0084

nPos	**rel. fluorescence**			**mean**	**standard deviation**	**with JOE0084**			**mean**	**standard deviation**	***p*-value**
1	15.04	15.90	14.99	15.31	0.42	14.71	14.52	15.57	14.94	0.56	44.0688%
1ol	15.08	15.70	15.92	15.57	0.36	15.85	14.64	15.43	15.31	0.62	58.4736%
2	13.93	14.77	15.12	14.61	0.50	14.18	13.29	14.37	13.95	0.57	24.4955%
2ol	12.40	11.29	10.40	11.36	0.82	11.13	10.79	9.20	10.37	1.03	29.8831%
-2	14.10	13.90	14.70	14.23	0.34	23.23	24.67	25.87	24.59	1.32	0.2852%
-2ol	12.40	13.20	12.70	12.77	0.33	16.91	17.19	16.38	16.83	0.41	0.0261%
-3	14.70	16.00	15.80	15.50	0.57	25.59	27.74	26.78	26.70	1.08	0.0294%
-3ol	13.73	14.80	15.10	14.54	0.59	17.44	16.07	16.09	16.53	0.78	3.2134%
PGK1	16.05	16.34	15.23	15.87	0.47						

Table M: Data for the relative fluorescence of the different nPos in combination with JOE0083

nPos	**rel. fluorescence**			**mean**	**standard deviation**	**with JOE0083**			**mean**	**standard deviation**	***p*-value**
1	15.04	15.90	14.99	15.31	0.42	15.36	16.05	15.49	15.63	0.37	43.2042%
1ol	15.08	15.70	15.92	15.57	0.36	14.24	16.45	15.72	15.47	1.12	89.9730%
-3	14.70	16.00	15.80	15.50	0.57	13.96	16.27	15.93	15.38	1.25	89.7667%
-3ol	13.73	14.80	15.10	14.54	0.59	13.22	14.72	14.54	14.16	0.82	57.3176%
PGK1	16.05	16.34	15.23	15.87	0.47						

Table N: Data for the relative fluorescence of the different nPos in combination with JOE0090

nPos	rel. fluorescence			mean	standard deviation	with JOE0090			mean	standard deviation	*p*-value
1	15.04	15.90	14.99	15.31	0.42	15.62	15.54	14.82	15.33	0.44	96.5618%
1ol	15.08	15.70	15.92	15.57	0.36	16.34	14.73	14.42	15.16	1.03	57.9681%
-3	14.70	16.00	15.80	15.50	0.57	28.65	27.13	26.33	27.37	1.18	0.0402%
-3ol	13.73	14.80	15.10	14.54	0.59	18.04	16.95	17.80	17.60	0.57	0.5246%
3	14.34	13.92	13.45	13.90	0.36	22.07	22.14	23.31	22.51	0.70	0.0174%
3ol	11.44	12.48	13.40	12.44	0.80	17.44	16.15	18.67	17.42	1.26	0.6686%
PGK1	16.05	16.34	15.23	15.87	0.47						

Table P: Data for the relative fluorescence of the different nPos in combination with JOE0072

nPos	rel. fluorescence			mean	standard deviation	with JOE0072			mean	standard deviation	*p*-value
1	15.04	15.90	14.99	15.31	0.42	15.64	16.24	15.25	15.71	0.50	38.5058%
1ol	15.08	15.70	15.92	15.57	0.36	15.24	15.83	15.66	15.58	0.31	97.7461%
-3	14.70	16.00	15.80	15.50	0.57	14.18	15.70	15.87	15.25	0.93	73.1566%
-3ol	13.73	14.80	15.10	14.54	0.59	14.15	15.73	15.08	14.98	0.79	51.7090%
PGK1	16.05	16.34	15.23	15.87	0.47						

Table O: Data for the relative fluorescence of the different nPos in combination with JOE0057

nPos	**rel. fluorescence**			**mean**	**standard deviation**	**with JOE0057**			**mean**	**standard deviation**	***p*-value**
1	15.04	15.90	14.99	15.31	0.42	14.75	16.71	15.74	15.73	0.98	55.3681%
1ol	15.08	15.70	15.92	15.57	0.36	15.01	15.57	16.25	15.61	0.62	92.2871%
2	13.93	14.77	15.12	14.61	0.50	14.79	14.74	15.17	14.90	0.24	49.8880%
2ol	12.40	11.29	10.40	11.36	0.82	12.24	10.32	10.85	11.14	0.99	79.4748%
-2	14.10	13.90	14.70	14.23	0.34	14.23	12.99	14.96	14.06	1.00	80.2828%
-2ol	12.40	13.20	12.70	12.77	0.33	12.11	13.20	13.34	12.88	0.68	81.1174%
-3	14.70	16.00	15.80	15.50	0.57	14.78	16.22	16.39	15.79	0.89	67.6025%
-3ol	13.73	14.80	15.10	14.54	0.59	13.96	14.40	14.63	14.33	0.34	67.5181%
3	14.34	13.92	13.45	13.90	0.36	14.23	14.61	13.40	14.08	0.62	71.0509%
3ol	11.44	12.48	13.40	12.44	0.80	11.66	12.62	12.62	12.30	0.56	84.5552%
1-3	7.70	5.84	6.12	6.55	0.82	7.38	5.89	6.34	6.53	0.76	97.9563%
1+3	7.70	6.70	5.90	6.77	0.74	7.43	7.10	5.22	6.58	1.19	84.2770%
1-2	6.10	7.27	6.59	6.65	0.48	5.28	7.09	6.77	6.38	0.97	69.9058%
TAT	7.16	6.54	7.18	6.96	0.30	7.12	6.93	6.99	7.01	0.10	82.8072%
PGK1	16.05	16.34	15.23	15.87	0.47						

7.2 GROWTH RATE

Table Q: Data for the mean growth rate of different nPos in YPDAU medium

nPos	60 min (OD_{600})		120 min (OD_{600})		240 min (OD_{600})		300 min (OD_{600})		360 min (OD_{600})		mean growth rate		mean rel. fluorescence
-1	0.056	0.031	0.067	0.032	0.127	0.062	0.190	0.089	0.320	0.144	0.309	0.267	0.000
1	0.051	0.022	0.062	0.024	0.120	0.045	0.091	0.063	0.185	0.090	0.271	0.272	15.310
1ol	0.123	0.044	0.088	0.054	0.294	0.101	0.362	0.121	0.284	0.205	0.243	0.279	15.567
2	0.107	0.100	0.141	0.138	0.252	0.203	0.254	0.253	0.359	0.397	0.248	0.248	14.607
2ol	1.052	1.015	1.257	1.190	2.869	2.820	2.911	2.900	2.832	2.800	0.262	0.269	11.364
-2	0.099	0.093	0.122	0.045	0.276	0.190	0.270	0.260	0.322	0.230	0.276	0.225	14.234
-2ol	0.080	0.017	0.092	0.020	0.111	0.038	0.091	0.051	0.127	0.071	0.078	0.276	12.767
-3	0.145	0.132	0.170	0.142	0.327	0.286	0.446	0.402	0.581	0.528	0.277	0.271	15.500
-3ol	0.076	0.029	0.088	0.025	0.132	0.060	0.157	0.110	0.213	0.122	0.191	0.288	14.545
3	0.130	0.126	0.142	0.149	0.316	0.287	0.422	0.384	0.618	0.566	0.301	0.284	13.904
3ol	1.333	1.040	1.624	1.370	3.178	3.045	4.171	3.531	2.641	2.624	0.237	0.283	12.440
1-3	0.101	0.058	0.128	0.072	0.129	0.132	0.196	0.180	0.216	0.180	0.133	0.279	6.554
1+3	0.081	0.125	0.102	0.155	0.198	0.316	0.258	0.388	0.304	0.473	0.284	0.286	6.767
1-2	0.041	0.114	0.051	0.094	0.093	0.147	0.151	0.132	0.172	0.186	0.295	0.073	6.654
1+2	0.104	0.078	0.115	0.093	0.254	0.189	0.316	0.140	0.523	0.371	0.300	0.303	0.000
TAT	0.152	0.122	0.205	0.149	0.365	0.294	0.434	0.393	0.639	0.572	0.281	0.298	6.960
SY992	0.139	0.885	0.190	1.090	0.340	2.340	0.459	3.100	0.592	4.200	0.296	0.316	0.000

Table R: Data for the mean growth rate of different nPos in combination with several plasmids in SD medium

combination	60 min (OD_{600})		120 min (OD_{600})		240 min (OD_{600})		300 min (OD_{600})		360 min (OD_{600})		mean growth rate	mean rel. fluorescence	
SY992	0.114	0.053	0.123	0.042	0.251	0.111	0.289	0.137	0.434	0.191	0.254	0.247	0
PGK1	0.969	0.952	1.251	1.174	1.819	1.809	2.848	2.867	2.788	2.798	0.230	0.235	15.87
1+JOE0046	0.097	0.066	0.098	0.072	0.200	0.125	0.219	0.153	0.255	0.196	0.213	0.214	15.25
1+JOE0084	0.028	0.014	0.031	0.012	0.059	0.029	0.017	0.029	0.107	0.046	0.258	0.221	14.94
1+JOE0079	0.149	0.123	0.093	0.139	0.284	0.249	0.382	0.326	0.512	0.419	0.232	0.241	0.62
1+JOE0082	0.058	0.067	0.070	0.076	0.093	0.146	0.077	0.093	0.161	0.310	0.144	0.216	0.47
-3+JOE0046	0.169	0.126	0.170	0.102	0.300	0.213	0.386	0.289	0.494	0.395	0.204	0.204	25.49
-3+JOE0084	1.420	1.194	1.561	1.199	3.159	2.254	4.087	2.216	2.588	2.575	0.217	0.173	26.7
-3+JOE0079	0.105	0.037	0.117	0.055	0.196	0.054	0.186	0.129	0.296	0.103	0.186	0.214	10.42
-3+JOE0082	0.105	0.046	0.112	0.054	0.154	0.082	0.25	0.085	0.287	0.145	0.182	0.192	7.69

7.3 PLASMID STABILITY

Table S: Data for the relative fluorescence and plasmid loss for plasmid JOE0079 in combination with nPos 1.

Control point	medium	rel. fluorescence	colonies SD-URA	per plate YPDAU	plasmid %
1	SD-URA	10.41	74	385	80.78%
	SD-URA-HIS	10.78	35	313	88.82%
2	SD-URA	11.41	71	337	78.93%
	SD-URA-HIS	0.00	53	444	88.06%
3	SD-URA	11.66	12	43	72.09%
	SD-URA-HIS	10.67	8	53	84.91%

Table T: Data for the relative fluorescence and plasmid loss for plasmid JOE0079 in combination with nPos 1ol.

Control point	medium	rel. fluorescence	colonies SD-URA	per plate YPDAU	plasmid %
1	SD-URA	12.89	31	179	82.68%
	SD-URA-HIS	11.44	39	299	86.96%
2	SD-URA	9.85	70	333	78.98%
	SD-URA-HIS	0.00	21	134	84.33%
3	SD-URA	13.55	9	30	70.00%
	SD-URA-HIS	11.79	18	153	88.24%

Table U: Data for the relative fluorescence and plasmid loss for plasmid JOE0092 in combination with nPos 1ol.

Control point	medium	rel. fluorescence	colonies SD-URA	per plate YPDAU	plasmid %
1	YPDAU	0.00	0	0	0.00%
	SD-URA	0.00	16	102	84.31%
	SD-URA-HIS	0.00	20	94	78.72%
2	YPDAU	0.00	35	147	76.19%
	SD-URA	0.00	13	49	73.47%
	SD-URA-HIS	0.00	12	227	94.71%
3	YPDAU	9.66	6	11	45.45%
	SD-URA	0.00	41	133	69.17%
	SD-URA-HIS	0.00	5	28	82.14%

Table V: Data for the relative fluorescence and plasmid loss for plasmid JOE0092 in combination with nPos 1.

Control point	medium	rel. fluorescence	colonies per plate SD-URA	YPDAU	plasmid %
1	SD-URA	0.00	30	201	85.07%
	SD-URA-HIS	0.00	10	60	83.33%
2	SD-URA	0.00	27	85	68.24%
	SD-URA-HIS	0.00	23	78	70.51%
3	SD-URA	7.14	9	17	47.06%
	SD-URA-HIS	0.00	2	13	84.62%

Table W: Data for the relative fluorescence and plasmid loss for plasmid JOE0092 in combination with nPos 1ol.

Control point	medium	rel. fluorescence	colonies per plate SD-URA	YPDAU	plasmid %
1	SD-URA	0.00	32	116	72.41%
	SD-URA-HIS	0.00	12	148	91.89%
2	SD-URA	0.00	12	28	57.14%
	SD-URA-HIS	0.00	11	64	82.81%
3	SD-URA	10.06	33	53	37.74%
	SD-URA-HIS	0.00	3	24	87.50%

Table X: Data for the sensitivity test of the fluorescence detector

	readings		OD_{600}	rel. fluorescence	mix
SD	0	0	0.0	0	
SY992	29	31	0.300	8.00	
Mix 1	355	348	0.250	83.69	1:10
Mix 2	300	314	0.280	66.02	1:50
Mix 3	120	130	0.310	27.78	1:100
Mix 4	86	95	0.300	22.35	1:250
Mix 5	66	80	0.270	14.31	1:500
Mix 6	68	71	0.340	14.95	1:750
Mix 7	41	43	0.310	9.66	1:1000
Mix 8	36	40	0.290	8.17	1:1500
SY992	38	44	0.310	8.82	
SD	6	8	0.0	0	

Table Y: Data for the relative fluorescence and plasmid loss for plasmid JOE0092 in combination with nPos 1ol.

Control point	medium	readings		OD_{600}	rel. fluorescence	colonies SD-URA	per plate YPDAU	plasmid %
1	YPDAU	49	44	0.312	9.94	14	101	86.14%
	SD-URA	27	25	0.269	6.44	21	93	77.42%
	SD-URA-HIS	19	24	0.306	4.68	15	142	89.44%
2	YPDAU	0	0	0.0	0.00	12	48	75.00%
	SD-URA	42	43	0.300	9.44	5	22	77.27%
	SD-URA-HIS	38	40	0.309	8.41	1	8	87.50%
3	YPDAU	55	59	0.297	12.79	9	19	52.63%
	SD-URA	43	45	0.319	9.20	64	238	73.11%
	SD-URA-HIS	40	38	0.289	9.00	9	127	92.91%
4	YPDAU	64	68	0.313	14.06	15	33	54.55%
	SD-URA	61	58	0.299	13.27	27	72	62.50%
	SD-URA-HIS	43	46	0.302	9.82	4	48	91.67%

Table Z: Data for the relative fluorescence and plasmid loss for plasmid JOE0092 in combination with nPos 1ol.

Control point	medium	readings		OD_{600}	rel. fluorescence	colonies SD-URA	per plate YPDAU	plasmid %
1	YPDAU	0	0	0.0	0.00	19	164	88.41%
	SD-URA	38	44	0.270	10.12	3	29	89.66%
	SD-URA-HIS	46	48	0.307	10.21	8	93	91.40%
2	YPDAU	46	48	0.300	10.44	14	49	71.43%
	SD-URA	53	55	0.331	10.88	2	8	75.00%
	SD-URA-HIS	42	44	0.310	9.25	24	121	80.17%
3	YPDAU	69	63	0.320	13.75	13	41	68.29%
	SD-URA	61	65	0.343	12.24	14	50	72.00%
	SD-URA-HIS	59	53	0.298	12.53	25	123	79.67%
4	YPDAU	81	85	0.310	17.85	8	16	50.00%
	SD-URA	69	73	0.300	15.78	34	77	55.84%
	SD-URA-HIS	46	48	0.303	10.34	5	42	88.10%

7.4 DNA SEQUENCES

Sequences received by GATC and remodeled with Clone Manager 9 Professional. Specific sites are indicated as follows:

X Missmatch

X O_R site or NLS or HIS(8)tag or aa mutation site

X Start codon

JOE0031 (promoter region)

```
Clone 3                3
actc-agacgcacagatattataacatctgcataataggcatttgcaagaattactcgcngagtaaggaa
JOE0031              507
actcaagacgcacagatattataacatctgcataataggcatttgcaagaattactcg-tgagtaaggaa

Clone 3               72
agagtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttc
JOE0031              576
agagtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttc

Clone 3              142
accctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcat
JOE0031              646
accctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcat

Clone 3              212
aaagcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaact
JOE0031              716
aaagcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaact

Clone 3              282
gaaaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaa
JOE0031              786
gaaaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaa

Clone 3              352
ggtcctagcgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttag
JOE0031              856
ggtcctagcgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttag

Clone 3              422
taccacatgctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctc
JOE0031              926
taccacatgctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctc

Clone 3              492
tttcaaacagaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaa
JOE0031              996
tttcaaacagaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaa

Clone 3              562
gggggtggtttagtttagtagaacctcgtgaaacttacatttacatatatataaacttgcataaattggt
JOE0031            1066
gggggtggtttagtttagtagaacctcgtgaaacttacatttacatatatataaacttgcataaattggt
```

```
Clone 3              632
caatgcaagaaatacatatttggtcttttctaattcttagtttttcaagttcttagatgctttctttttc
JOE0031             1136
caatgcaagaaatacatatttggtcttttctaattcgtagtttttcaagttcttagatgctttctttttc

Clone 3              702
tctttttacagatcatcaaggaagtaattatctacttttacaacanatataaaacaatg-ggcatgca
JOE0031             1206
tctttttacagatcatcaaggaagtaattatctacttttacaacaaatataaaacaatgcggcatgca

Clone 3              771
atctanaggtgaagaattattcactggtgttgtcccaattttggttgaattagatggtgatgttaatggt
JOE0031             1276
atctaaaggtgaagaattattcactggtgttgtcccaattttggttgaattagatggtgatgttaatggt

Clone 3              841
cacaaattttctgtctccggtgaaggtgaaggtgatgctacttacggtanattgaccttaaaatttattt
JOE0031             1346
cacaaattttctgtctccggtgaaggtgaaggtgatgctacttacggtaaattgaccttaaaatttattt

Clone 3              911
gtactactggtaaattgccagttccatggccaaccttagtcactactttccggttatgntgttcaatgnt
JOE0031             1416
gtactactggtaaattgccagttccatggccaaccttagtcactacttt-cggttatggtgttcaatgtt

Clone 3              981
ttgnnnagatacccnagatcatatgaaaacaacatggacttttttcaag
JOE0031             1485
ttg-ctagatacccc-agatcatatg-aaacaacat-gac-tttttcaag
```

JOE0031 (promoter region, point mutation)

```
JOE0031              510
-ca---agacgcacagatattataacatctgcataataggcatttgcaagaattactcgt-gagtaagga
Clone 1                1
-cactcagacgcacagatattataacatctgcataataggcatttgcaagaattactcgn-gagtaagga
Clone 3                1
tcactcagacgcacagatattataacatctgcataataggcatttgcaagaattactcgcngagtaagga


JOE0031              575
aagagtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggctt
Clone 1               69
aagagtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggctt
Clone 3               71
aagagtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggctt

JOE0031              645
caccctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctca
Clone 1              139
caccctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctca
Clone 3              141
caccctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctca

JOE0031              715
taaagcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaac
Clone 1              209
taaagcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaac
Clone 3              211
taaagcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaac
```

```
JOE0031             785
tgaaaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaaca
Clone 1             279
tgaaaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaaca
Clone 3             281
tgaaaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaaca

JOE0031             855
aggtcctagcgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggttta
Clone 1             349
aggtcctagcgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggttta
Clone 3             351
aggtcctagcgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggttta

JOE0031             925
gtaccacatgctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctct
Clone 1             419
gtaccacatgctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctct
Clone 3             421
gtaccacatgctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctct

JOE0031             995
ctttcaaacagaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaacca
Clone 1             489
ctttcaaacagaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaacca
Clone 3             491
ctttcaaacagaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaacca

JOE0031            1065
agggggtggtttagtttagtagaacctcgtgaaacttacatttacatatatataaacttgcataaattgg
Clone 1             559
agggggtggtttagtttagtagaacctcgtgaaacttacatttacatatatataaacttgcataaattgg
Clone 3             561
agggggtggtttagtttagtagaacctcgtgaaacttacatttacatatatataaacttgcataaattgg

JOE0031            1135
tcaatgcaagaaatacatatttggtcttttctaattcgtagttttcaagttcttagatgctttcttttt
Clone 1             629
tcaatgcaagaaatacatatttggtcttttctaattcttagttttcaagttcttagatgctttcttttt
Clone 3             631
tcaatgcaagaaatacatatttggtcttttctaattcttagttttcaagttcttagatgctttcttttt

JOE0031            1205
ctctttttacagatcatcaaggaagtaattatctacttttacaacaaatataaaacaatgcggcatgc
Clone 1             699
ctctttttacagatcatcaaggaagtaattatctacttttacaacaaatataanacaatg-ggcatgc
Clone 3             701
ctctttttacagatcatcaaggaagtaattatctacttttacaacanatataaaacaatg-ggcatgc

JOE0031            1275
aatctaaaggtgaagaattattcactggtgttgtcccaattttggttgaattagatggtgatgttaatgg
Clone 1             768
aatctaaaggtgaagaattattcactggtgttgtcccaattttggttgaattagatggtgatgttaatgg
Clone 3             770
aatctanaggtgaagaattattcactggtgttgtcccaattttggttgaattagatggtgatgttaatgg

JOE0031            1345
tcacaaattttctgtctccggtgaaggtgaaggtgatgctacttacggtaaattgaccttaaaatttatt
Clone 1             838
tcacaaattttctgtctccggtgaaggtgaaggtgatgctacttacggtaaattgaccttaaaatttatt
Clone 3             840
tcacaaattttctgtctccggtgaaggtgaaggtgatgctacttacggtanattgaccttaaaatttatt
```

```
JOE0031            1415
tgtactactggtaaattgccagttccatggccaaccttagtcactactttc-ggttatggtgttcaatgt
Clone 1             908
tgtactactggtaaattgccagttccatggccaaccttagtcactactttc-ggttatggtgttcaatgt
Clone 3             910
tgtactactggtaaattgccagttccatggccaaccttagtcactactttccggttatgntgttcaatgn

JOE0031            1484
tttgct-agataccc-agatcatatgaaa-caacatgacttttcaagtctgccatgccagaaggttatg
Clone 1             977
tttgcg-agataccc-agatcatatgaaa-caacatgacttttcaagtctgccatgccagaaggttatg
Clone 3             980
tttgnnnagatacccnagatcatatgaaaacaacat----------------------------------

JOE0031            1551
ttcaagaaagaactatttttttcaaagatgacggtaactac-aagaccagagctgaagtcaagtttgaag
Clone 1            1044
ttcaagaaagaactattnnnttcaaannntgacggtaactaccaagaccagagctgaagtcaagtttgaag
Clone 3            1016
----------ggacttttt----------------------c-aagnnc----------------------
```

JOE0031 (His(8)tag region)

```
Clone 3               6
gaaaaaaaagaaaaatttgaaatannncncnnnnannnnnatactaacataactataaaaaaataaatag
JOE0031            2215
gaaaaaaaagaaaaatttgaaata-taaataacgttcttaatactaacataactataaaaaaataaatag

Clone 3              76
ggacctagacttcaggttgtctaactccttccttttcnnttagagnggatgtgggggggagggcgtgaatg
JOE0031            2146
ggacctagacttcaggttgtctaactccttccttttcggttagagcggatgtgggggggagggcgtgaatg

Clone 3             146
taagcgtgacataactaattacatgactcgaccagccgcggttatagtttactctcgagatgatgatgat
JOE0031            2076
taagcgtgacataactaattacatgactcgaccagccgcggttatagtttactctcgagatgatgatgat

Clone 3             216
gatgatgatggaattctttgtacaattcatccataccatgggtaataccagcagcagtaacaaattc
JOE0031  HIS       2006
gatgatgatggaattctttgtacaattcatccataccatgggtaataccagcagcagtaacaaattc

Clone 3             286
taacaagaccatgtggtctctcttttcgtttggatctttggataaggcagattgagtggataagtaatgg
JOE0031            1936
taacaagaccatgtggtctctcttttcgtttggatctttggataaggcagattgagtggataagtaatgg

Clone 3             356
ttgtctggtaacaagactggaccatcaccaattggagtattttgttgataatggtcagctaattgaacag
JOE0031            1866
ttgtctggtaacaagactggaccatcaccaattggagtattttgttgataatggtcagctaattgaacag

Clone 3             426
aaccatcttcaatgttgtgtctaattttgaagttaactttgataccattcttttgtttgtcagccatgat
JOE0031            1796
aaccatcttcaatgttgtgtctaattttgaagttaactttgataccattcttttgtttgtcagccatgat
```

```
Clone 3              496
gtaaacattgtgagagttatagttgtattccaatttgtgacctaaaatgttaccatcttctttaaaatca
JOE0031             1726
gtaaacattgtgagagttatagttgtattccaatttgtgacctaaaatgttaccatcttctttaaaatca

Clone 3              566
ataccttttaattcgattctattaactaaggtatcaccttcaaacttgacttcagctctggncttgtagt
JOE0031             1656
ataccttttaattcgattctattaactaaggtatcaccttcaaacttgacttcagctctggtcttgtagt

Clone 3              636
taccgtcatctttgaaaaaaatagttctttcttgaacataaccttctggcatggcagacttgaaaaagtc
JOE0031             1586
taccgtcatctttgaaaaaaatagttctttcttgaacataaccttctggcatggcagacttgaaaaagtc

Clone 3              706 atgttgtttcatatgatctgggtatct
JOE0031             1516 atgttgtttcatatgatctgggtatct
```

JOE0032 (promoter region)

```
Clone 2                1
tttctctctcccccgttgttgtctcnccatatccgcaatgacanannnntgatggaagacactaaaggaa
JOE0032             6434
tttctctctcccccgttgttgtctcaccatatccgcaatgacaaaaaaatgatggaagacactaaaggaa

Clone 2               71
aaaattaacgacaaagacagcaccaacagatgtcgttgttccagagctgatgaggggtatctcgaagcac
JOE0032             6504
aaaattaacgacaaagacagcaccaacagatgtcgttgttccagagctgatgaggggtatctcgaagcac

Clone 2              141
acgaaacttttccttccttcattcacgcacactactctctaatgagcaacggtatacggccttccttcc
JOE0032             6574
acgaaacttttccttccttcattcacgcacactactctctaatgagcaacggtatacggccttccttcc

Clone 2              211
agttacttgaatttgaaataaaaaaaagtttgctgtcttgctatcaagtataaatagacctgcaattatt
JOE0032             6644
agttacttgaatttgaaataaaaaaaagtttgctgtcttgctatcaagtataaatagacctgcaattatt

Clone 2              281
aatcttttgtttcctcgtcattgttctcgttccctttcttccttgtttctttttctgcacaatatttcaa
JOE0032             6714
aatcttttgtttcctcgtcattgttctcgttccctttcttccttgtttctttttctgcacaatatttcaa

Clone 2              351
gctataccaagcatacaatcaactatctcatatacaatgcttttgcaagctttccttttccnnttggctg
JOE0032             6784
gctataccaagcatacaatcaactatctcatatacaatgcttttgcaagctttccttttccttttggctg

Clone 2              421 gttttgcagccaaaatatctgca
JOE0032             6854 gttttgcagccaaaatatctgca
```

JOE0035 (promoter region)

```
Clone 7                9
ttcttccttatcggatcctcaaaaccctaaaaacatatgcctcaccctaacatattttccaattaaccc
JOE0035             6418
ttcttccttatcggatcctcaaaaccctaaaaacatatgcctcaccctaacatattttccaattaaccc
```

```
Clone 7              79
tcaatatttctctgtcacccggcctctattttccattttcttctttacccgccacgcgtttttttctttc
JOE0035            6488
tcaatatttctctgtcacccggcctctattttccattttcttctttacccgccacgcgtttttttctttc

Clone 7             149
aaatttttttcttccttcttctttttcttccacgtcctcttgcataaataaataaaccgttttgaaacca
JOE0035            6558
aaatttttttcttccttcttctttttcttccacgtcctcttgcataaataaataaaccgttttgaaacca

Clone 7             219
aactcgcctctctctctcctttttgaaatatttttgggtttgtttgatcctttccttcccaatctctctt
JOE0035            6628
aactcgcctctctctctcctttttgaaatatttttgggtttgtttgatcctttccttcccaatctctctt

Clone 7             289
gtttaatatatattcatttatatcacgctctctttttatcttcctttttttcctctctcttgtattcttc
JOE0035            6698
gtttaatatatattcatttatatcacgctctctttttatcttcctttttttcctctctcttgtattcttc
Clone 7             359
cttccccttttctactcaaaccaagaagaaaaagaaaaggtcaatctttgttaaagaataggatcttctac
JOE0035            6768
cttccccttttctactcaaaccaagaagaaaaagaaaaggtcaatctttgttaaagaataggatcttctac

Clone 7             429
tacatcagcttttanatttttcncgcttactgctttttttcttcccaagatcgaaaatttantgaattaac
JOE0035            6838
tacatcagcttttagatttttcacgcttactgctttttttcttcccaagatcgaaaatttactgaattaac

Clone 7             499 aatgggcatgcaa
JOE0035            6908 aatgctttttgcaa
```

JOE0044 (promoter region)

```
Clone 5               6
ct-aaggg-ac-naagctggagctctttcagcagctctgatgtagatacacgtatctcgatatgttttat
JOE0044            1243
ctaaagggaacaaaagctggagctctttcagcagctctgatgtagatacacgtatctcgacatgttttat

Clone 5              73
ttttactatacatacataaaagaaataaaaaatgataacgtgtatattattattcatataatcaatgagg
JOE0044            1173
ttttactatacatacataaaagaaataaaaaatgataacgtgtatattattattcatataatcaatgagg

Clone 5             143
gtcattttctgaaacgcaaaaaacggtaaatggaaaaaaaataaagatagaaaaagaaaacaaacaaagg
JOE0044            1103
gtcattttctgaaacgcaaaaaacggtaaatggaaaaaaaataaagatagaaaaagaaaacaaacaaagg

Clone 5             213
aaaggttagcatattaaataactgagctgatacttcaacagcatcgctgaagagaacagtattgaaaccg
JOE0044            1033
aaaggttagcatattaaataactgagctgatacttcaacagcatcgctgaagagaacagtattgaaaccg

Clone 5             283
aaacattttctaaaggcaaacaaggtactccatatttgctggacgtgttctttctctcgtttcatatgca
JOE0044             963
aaacattttctaaaggcaaacaaggtactccatatttgctggacgtgttctttctctcgtttcatatgca
```

```
Clone 5              353
taattctgtcataagcctgttctttttcctggcttaaacatcccgttttgtaaaagagaaatctattcca
JOE0044              893
taattctgtcataagcctgttctttttcctggcttaaacatcccgttttgtaaaagagaaatctattcca

Clone 5              423
catatttcattcattcggctaccatactaaggataaactaatcccgttgtttttggcctcatcacataa
JOE0044              823
catatttcattcattcggctaccatactaaggataaactaatcccgttgtttttggcctcgtcacataa

Clone 5              493
ttataaactactaacccattatcagatgggcatgcaaagcacaaaaaagaaaccattaacacaagagcag
JOE0044              753
ttataaactactaacccattatcagatgggcatgcaaagcacaaaaaagaaaccattaacacaagagcag

Clone 5              563
cttgaggacgcacgtcgccttaaagcaatttatgaaaaaaagaaaaatgaacttggcttatcccaggaat
JOE0044              683
cttgaggacgcacgtcgccttaaagcaatttatgaaaaaaagaaaaatgaacttggcttatcccaggaat

Clone 5              633
ctgtcgcagacaagatggggatggggcagtcaggcgttggtgctttatttaatggcatcaatgcattaaa
JOE0044              613
ctgtcgcagacaagatggggatggggcagtcaggcgttggtgctttatttaatggcatcaatgcattaaa

Clone 5              703
tgcttataacgccgcattgcttacaaaaattctcaaagttagcgttgaagaatttagcccttcaatcgcc
JOE0044              543
tgcttataacgccgcattgcttgcaaaaattctcaaagttagcgttgaagaatttagcccttcaatcgcc

Clone 5              773
agagaaatctacgagatgtatgaagcggttagtatgcagccgtcacttagaagtgagtatgagtaccctg
JOE0044              473
agagaaatctacgagatgtatgaagcggttagtatgcagccgtcacttagaagtgagtatgagtaccctg

Clone 5              843
ttttttctcatgttcaggcagggatgttctcacctaagcttagaacctttaccaaaggtgatgcggagag
JOE0044              403
ttttttctcatgttcaggcagggatgttctcacctgagcttagaacctttaccaaaggtgatgcggagag

Clone 5              913
atgggtaagcacaaccaannanagccagtgattctgcattctggcttgaggttgaaggtaattccatgacc
JOE0044              333
atgggtaagcacaaccaaaaaagccagtgattctgcattctggcttgaggttgaaggtaattccatgacc

Clone 5              983
gcaccaacaggctccaagccaagctttcctgacggaatgttaattctcgttgaccctgagcaggctgttg
JOE0044              263
gcaccaacaggctccaagccaagctttcctgacggaatgttaattctcgttgaccctgagcaggctgttg

Clone 5             1053
agccaggtgatttctgcatagccagacttgggggtgannannttaccttcaannaactgatcagggatac
JOE0044              193
agccaggtgatttctgcatagccagacttgggggtgatgagtttaccttcaagaaactgatcagggatag

Clone 5             1123 cggtcaggt
JOE0044              123 cggtcaggt
```

JOE0046 (promoter region)

```
Clone 1                6
ct-aaggg-acnaaagctggagctctttcagcagctctgatgtagatacacgtatctcgatatgttttat
JOE0046            3866
ctaaagggaacaaaagctggagctctttcagcagctctgatgtagatacacgtatctcgacatgttttat

Clone 1               74
ttttactatacatacataaaagaaataaaaaatgataacgtgtatattattattcatataatcaatgagg
JOE0046            3796
ttttactatacatacataaaagaaataaaaaatgataacgtgtatattattattcatataatcaatgagg

Clone 1              144
gtcattttctgaaacgcaaaaaacggtaaatggaaaaaaaataaagatagaaaaagaaaacaaacaaagg
JOE0046            3726
gtcattttctgaaacgcaaaaaacggtaaatggaaaaaaaataaagatagaaaaagaaaacaaacaaagg

Clone 1              214
aaaggttagcatattaaataactgagctgatacttcaacagcatcgctgaagagaacagtattgaaaccg
JOE0046            3656
aaaggttagcatattaaataactgagctgatacttcaacagcatcgctgaagagaacagtattgaaaccg

Clone 1              284
aaacattttctaaaggcaaacaaggtactccatatttgctggacgtgttctttctctcgtttcatatgca
JOE0046            3586
aaacattttctaaaggcaaacaaggtactccatatttgctggacgtgttctttctctcgtttcatatgca

Clone 1              354
taattctgtcataagcctgttctttttcctggcttaaacatcccgttttgtaaaagagaaatctattcca
JOE0046            3516
taattctgtcataagcctgttctttttcctggcttaaacatcccgttttgtaaaagagaaatctattcca

Clone 1              424
catatttcattcattcggctaccatactaaggataaactaatcccgttgtttttggcctcatcacataa
JOE0046            3446
catatttcattcattcggctaccatactaaggataaactaatcccgttgtttttggcctcgtcacataa

Clone 1              494
ttataaactactaacccattatcagatgggcatgcaaagcacaaaaaagaaaccattaacacaagagcag
JOE0046            3376
ttataaactactaacccattatcagatgggcatgcaaagcacaaaaaagaaaccattaacacaagagcag

Clone 1              564
cttgaggacgcacgtcgccttaaagcaatttatgaaaaaaagaaaaatgaacttggcttatcccaggaat
JOE0046            3306
cttgaggacgcacgtcgccttaaagcaatttatgaaaaaaagaaaaatgaacttggcttatcccaggaat

Clone 1              634
ctgtcgcagacaagatggggatggggcagtcaggcgttggtgctttatttaatggcatcaatgcattaaa
JOE0046            3236
ctgtcgcagacaagatggggatggggcagtcaggcgttggtgctttatttaatggcatcaatgcattaaa

Clone 1              704
tgcttataacgccgcattgcttacaaaaattctcaaagttagcgttgaagaatttagcccttcaatcgcc
JOE0046            3166
tgcttataacgccgcattgcttgcaaaaattctcaaagttagcgttgaagaatttagcccttcaatcgcc

Clone 1              774
agagaaatctacgagatgtatgaagcggttagtatgcagccgtcacttagaagtgagtatgagtaccctа
JOE0046            3096                                                NLS
agagaaatctacgagatgtatgaagcggttagtatgcagccgtcacttagaagtgagtatgagtaccctа
```

```
Clone 1              844
agaaaaagcgcaaggtaactagaatgttctcacctgagcttagaacctttaccaaaggtgatgcggagag
JOE0046             3026
agaaaaagcgcaaggtaactagaatgttctcacctgagcttagaacctttaccaaaggtgatgcggagag

Clone 1              914
atgggtaagcacaaccaaaaaagccagtgattctgcattctggcttgaggttgaaggtaattccatgacc
JOE0046             2956
atgggtaagcacaaccaaaaaagccagtgattctgcattctggcttgaggttgaaggtaattccatgacc

Clone 1              984
gcaccaacaggctccaagccaagctttcctgacggaatgttaattctcgttgaccctgancaggctgttg
JOE0046             2886
gcaccaacaggctccaagccaagctttcctgacggaatgttaattctcgttgaccctgagcaggctgttg

Clone 1             1054
agccaggtgatttctgcatagccagacttgggggtgatgag-ttaccttcaa-aaactgatcagggnnag
JOE0046             2816
agccaggtgatttctgcatagccagacttgggggtgatgagtttaccttcaagaaactgatcagggatag

Clone 1             1122 cggtcaggtgttt
JOE0046             2746 cggtcaggtgttt
```

JOE0049 (promoter region)

```
Clone 5                2
aaggga--nnaagctggagctctttcagcagctctgatgtagatacacgtatctcgatatgttttatttt
JOE0049             3589
aagggaacaaaagctggagctctttcagcagctctgatgtagatacacgtatctcgacatgttttatttt

Clone 5               70
tactatacatacataaaagaaataaaaaatgataacgtgtatattattattcatataatcaatgagggtc
JOE0049             3519
tactatacatacataaaagaaataaaaaatgataacgtgtatattattattcatataatcaatgagggtc

Clone 5              140
attttctgaaacgcaaaaaacggtaaatggaaaaaaaataaagatagaaaaagaaaacaaacaaaggaaa
JOE0049             3449
attttctgaaacgcaaaaaacggtaaatggaaaaaaaataaagatagaaaaagaaaacaaacaaaggaaa

Clone 5              210
ggttagcatattaaataactgagctgatacttcaacagcatcgctgaagagaacagtattgaaaccgaaa
JOE0049             3379
ggttagcatattaaataactgagctgatacttcaacagcatcgctgaagagaacagtattgaaaccgaaa

Clone 5              280
cattttctaaaggcaaacaaggtactccatatttgctggacgtgttctttctctcgtttcatatgcataa
JOE0049             3309
cattttctaaaggcaaacaaggtactccatatttgctggacgtgttctttctctcgtttcatatgcataa

Clone 5              350
ttctgtcataagcctgttctttttcctggcttaaacatcccgttttgtaaaagagaaatctattccacat
JOE0049             3239
ttctgtcataagcctgttctttttcctggcttaaacatcccgttttgtaaaagagaaatctattccacat

Clone 5              420
atttcattcattcggctaccatactaaggataaactaatcccgttgtttttggcctcatcacataatta
JOE0049             3169
atttcattcattcggctaccatactaaggataaactaatcccgttgtttttggcctcgtcacataatta
```

```
Clone 5              490
taaactactaacccattatcagatgggcatgcaacatcatcatcatcatcatagcacaaaaaagaaacca
JOE0049             3099                                HIS
taaactactaacccattatcag-t--gcatgcaacatcatcatcatcatcatagcacaaaaaagaaacca

Clone 5              560
ttaacacaagagcagcttgaggacgcacgtcgccttaaagcaatttatgaaaaaaagaaaaatgaacttg
JOE0049             3032
ttaacacaagagcagcttgaggacgcacgtcgccttaaagcaatttatgaaaaaaagaaaaatgaacttg

Clone 5              630
gcttatcccaggaatctgtcgcagacaagatggggatggggcagtcaggcgttggtgctttatttaatgg
JOE0049             2962
gcttatcccaggaatctgtcgcagacaagatggggatggggcagtcaggcgttggtgctttatttaatgg

Clone 5              700
catcaatgcattanatgcttataacgccgcattgcttacaaaaattctcanagttagcgttgaagaattt
JOE0049             2892
catcaatgcattaaatgcttataacgccgcattgcttgcaaaaattctcaaagttagcgttgaagaattt

Clone 5              770
agcccttcaatcgccagagaaatctacgagatgtatgaagcggttagtatgcagccgtcacttanaagtg
JOE0049             2822
agcccttcaatcgccagagaaatctacgagatgtatgaagcggttagtatgcagccgtcacttagaagtg

Clone 5              840
agtatgagtacctaagaaaaagcgcaaggtaactagaatgttctcacctgancttanaacctttaccaa
JOE0049             2752                NLS
agtatgagtacctaagaaaaagcgcaaggtaactagaatgttctcacctgagcttagaacctttaccaa

Clone 5              910
aggtgannncgnaaanatgggtaagcacaaccaaaaaancccnnntgattctncattctggcttgnggttg
JOE0049             2682
aggtga-tgcggagagatgggtaagcacaaccaaaaaagcc-agtgattctgcattctggcttgaggttg

Clone 5              980 aaagtaatt
JOE0049             2614 aaggtaatt
```

JOE0056 (promoter region)

```
Clone 5                1
aagctggagctccagtaccagcgtctgcatagagccggaaatccttttcttttgcagcctccttctcaata
JOE0056             3509
aagctggagctccagtaccagcgtctgcatagagccggaaatccttttcttttgcagcctccttctcaata

Clone 5               71
tcttctagcacccatgttgcatccctgtagaaattcatgctttttcgcagctaacttctcagtgaaacac
JOE0056             3439
tcttctagcacccatgttgcatccctgtagaaattcatgctttttcgtagctaacttctcagtgaaacac

Clone 5              141
ttggtgaagaaccagctcttctcttaacctcatcgaatattttaggccagtttttttttcctcacttgc
JOE0056             3369
ttggtgaagaaccagctcttctcttaacctcatcgaatattttaggccagtttttttttcctcacttgc

Clone 5              211
cattttcccgcaccgtttggaaatctcgaggaaaatgcttctttccaaaagtttaatcaggtcactaatt
JOE0056             3299
cattttcccgcaccgtttggaaatctcgaggaaaatgcttctttccaaaagtttaatcaggtcactaatt
```

```
Clone 5              281
aggtaaatagagctgggaggaagcatacacagtgtttcgtctgtggaaggaacttttatttttcttaaat
JOE0056             3229
aggtaaatagagctgggaggaagcatatacagtgtttcgtctgtggaaggaacttttatttttcttaaat

Clone 5              351
gttttaatttttagtgacctaattcacataacaaaatattgagaaagggaaagttgagcactgtaatacg
JOE0056             3159
gttttaatttttagtgacctaattcacataacaaaatattgagaaagggaaagttgagcactgtaatacg

Clone 5              421
ccgaacagattaagcatgcaacatcatcatcatcatcatagcacaaaaaagaaaccattaacacaagagc
JOE0056             3089              HIS
ccgaacagattaagcatgcaacatcatcatcatcatcatagcacaaaaaagaaaccattaacacaagagc

Clone 5              491
agcttgaggacgcacgtcgccttaaagcaatttatgaaaaaaagaaaaatgaacttggcttatcccagga
JOE0056             3019
agcttgaggacgcacgtcgccttaaagcaatttatgaaaaaaagaaaaatgaacttggcttatcccagga

Clone 5              561
atctgtcgcagacaagatggggatggggcagtcaggcgttggtgctttatttaatggcatcaatgcatta
JOE0056             2949
atctgtcgcagacaagatggggatggggcagtcaggcgttggtgctttatttaatggcatcaatgcatta

Clone 5              631
aatgcttataacgccgcattgcttacaaaaattctcaaagttagcgttgaagaatttagcccttcaatcg
JOE0056             2879
aatgcttataacgccgcattgcttgcaaaaattctcaaagttagcgttgaagaatttagcccttcaatcg

Clone 5              701
ccagagaaatctacgagatgtatgaagcggttagtatgnanccgtcacttagaagtgagtatgagtaccc
JOE0056             2809
ccagagaaatctacgagatgtatgaagcggttagtatgcagccgtcacttagaagtgagtatgagtaccc

Clone 5              771 taannanaaagcgcaaggtaactannaatgttctcacc
JOE0056             2739 taa-gaaaaagcgcaaggtaacta-gaatgttctcacc
```

JOE0059 (promoter region)

```
Clone 1              514
acgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgag
JOE0059                7
acgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgag

Clone 1              584
gaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcat
JOE0059               77
gaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcat

Clone 1              654
actattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacg
JOE0059              147
actattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacg

Clone 1              724
tggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaac
JOE0059              217
tggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaac
```

```
Clone 1              794
ccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaaggtcctag
JOE0059              287
ccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaaggtcctag

Clone 1              864
cgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttagtaccacat
JOE0059              357
cgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttagtaccacat

Clone 1              934
gctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaac
JOE0059              427
gctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaac

Clone 1             1004
agaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaaggggtgg
JOE0059              497
agaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaaggggtgg

Clone 1             1074
tttagtttagtagaacctcgtgaaacttacatttgcaaccattatcaccgccagaggtaaaatagtcaa
JOE0059              567           OR
tttagtttagtagaacctcgtgaaacttacatttgcaaccattatcaccgccagaggtaaaatagtcaa

Clone 1             1144
cacgcacggtgttagatatttatcccttgcggtgatagatttaacgtacatatatataaacttgcataaa
JOE0059              637
cacgcacggtgttagatatttatcccttgcggtgatagatttaacgtacatatatataaacttgcataaa

Clone 1             1214
ttggtcaatgcaagaaatacatatttggtcttttctaattcgtagttttcaagttcttagatgctttct
JOE0059              707
ttggtcaatgcaagaaatacatatttggtcttttctaattcttagttttcaagttcttagatgctttct

Clone 1             1284
ttttctctttttacagatcatcaaggaagtaattatctacttttacaacaaatataaaacaatgcggc
JOE0059              777
ttttctctttttacagatcatcannggaagtaantatctacttttacaacaaatataaaacaatg-ggc

Clone 1             1354
atgcaatctaaaggtgaagaattatt-cactggtgttgtcccaattttg
JOE0059              846
atgcnatctnnaggtgaanaattattncnctggtgttgtcccaattttg
```

JOE0060 (promoter region)

```
Clone 4                1
ttataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgaggaactatcgca
JOE0060              525
ttataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgaggaactatcgca

Clone 4               71
tacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcatactattatcag
JOE0060              595
tacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcatactattatcag

Clone 4              141
ggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacgtggcctcttat
JOE0060              665
ggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacgtggcctcttat
```

```
Clone 4             211
cgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaacccagacacgct
JOE0060             735
cgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaacccagacacgct

Clone 4             281
cgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaaggtcctagcgacggctcac
JOE0060             805
cgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaaggtcctagcgacggctcac

Clone 4             351
aggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttagtaccacatgctatgatgcc
JOE0060             875
aggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttagtaccacatgctatgatgcc

Clone 4             421
cactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaacagaattgtccg
JOE0060             945
cactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaacagaattgtccg

Clone 4             491
aatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaangggggtggtttagtttagt
JOE0060            1015
aatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaaggggggtggtttagtttagt

Clone 4             561
agaacctcgtgaaacttacatttacatatatataaacttgcataanttggtcaatgcaagaaatactctaa
JOE0060            1085                                                 OR
agaacctcgtgaaacttacatttacatatatataaacttgcataaattggtcaatgcaagaaatactct-

Clone 4             631 aacaccgtgcgtgttgactattttacctctggcngngataata
JOE0060            1154 aacaccgtgcgtgttgactattttacctctggcggtgataata
```

JOE0061 (promoter region)

```
Clone 1               5
cacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgaggaa
JOE0061             517
cacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgaggaa

Clone 1              75
ctatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcatact
JOE0061             587
ctatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcatact

Clone 1             145
attatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacgtgg
JOE0061             657
attatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacgtgg

Clone 1             215
cctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaaccca
JOE0061             727
cctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaaccca

Clone 1             285
gacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaaggtcctagcga
JOE0061             797
gacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaaggtcctagcga
```

```
Clone 1              355
cggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttagtaccacatgct
JOE0061              867
cggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttagtaccacatgct

Clone 1              425
atgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaacaga
JOE0061              937
atgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaacaga

Clone 1              495
attgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaaggggtggttt
JOE0061             1007
attgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaaggggtggttt

Clone 1              565
agtttagtagaacctcgtgaaacttacattctaacaccgtgcgtgttgactattttacctctggcggtga
JOE0061             1077      OR
agtttagtagaacctcgtgaaacttacattctaacaccgtgcgtgttgactattttacctctggcggtga

Clone 1              635
taatatatttggtcttttcgaattcttagtttttcaagttcttagatgctttctttttctctttttttaca
JOE0061             1147
taatatatttggtcttttctaattcgtagtttttcaagttcttagatgctttctttttctctttttttaca

Clone 1              705
gatcatcaaggaagtaattatctacttttttacaacaaatataaaacaatg-ggcatgcaatctaaaggtg
JOE0061             1217
gatcatcaaggaagtaattatctacttttttacaacaaatataaaacaatgcggcatgcaatctaaaggtg

Clone 1              774
aagaattattcactggtgttgtcccaattttggttgaattagatggtgatgttaatggtcacaaattttc
JOE0061             1287
aagaattattcactggtgttgtcccaattttggttgaattagatggtgatgttaatggtcacaaattttc

Clone 1              844
tgtctccggtgaaggtgaaggtgatgctacttacggtaaaattgaccttaaaatttattt
JOE0061             1357
tgtctccggtgaaggtgaaggtgatgctacttacggt-aaattgaccttaaaatttattt
```

JOE0062 (promoter region)

```
Clone 1                2
actcagaacgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaa
JOE0062              507
actcaagacgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaa

Clone 1               72
gagtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttca
JOE0062              577
gagtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttca

Clone 1              142
ccctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcata
JOE0062              647
ccctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcata

Clone 1              212
aagcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactg
JOE0062              717
aagcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactg
```

```
Clone 1              282
aaaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaag
JOE0062              787
aaaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaag

Clone 1              352
gtcctagcgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttagt
JOE0062              857
gtcctagcgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttagt

Clone 1              422
accacatgctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctct
JOE0062              927
accacatgctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctct

Clone 1              492
ttcaaacagaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaag
JOE0062              997
ttcaaacagaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaag

Clone 1              562
ggggtggtttagtttagtagaacctcgtgaaacttacatttctaacaccgtgcgtgttgactattttac
JOE0062            1067                   OR
ggggtggtttagtttagtagaacctcgtgaaacttacatttctaacaccgtgcgtgttgactattttac

Clone 1              632
ctctggcggtgataatacatatatataaacttgcataaattggtcaatgcaagaaatacatatttggtct
JOE0062            1137
ctctggcggtgataatacatatatataaacttgcataaattggtcaatgcaagaaatacatatttggtct

Clone 1              702
tttctaattcttagtttttcaagttcttagatgctttctttttctctttttttacagatcatcaaggaagt
JOE0062            1207
tttctaattcgtagtttttcaagttcttagatgctttctttttctctttttttacagatcatcaaggaagt

Clone 1              772 aattatctacnnnttacaacaaatataaaa
JOE0062            1277 aattatctactttttacaacaaatataaaa
```

JOE0063 (promoter region)

```
Clone 1                1
acgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgag
JOE0063              514
acgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgag

Clone 1               71
gaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcat
JOE0063              584
gaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcat

Clone 1              141
actattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacg
JOE0063              654
actattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacg

Clone 1              211
tggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaac
JOE0063              724
tggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaac
```

```
Clone 1              281
ccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaaggtcctag
JOE0063              794
ccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaaggtcctag

Clone 1              351
cgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttagtaccacat
JOE0063              864
cgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttagtaccacat

Clone 1              421
gctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaac
JOE0063              934
gctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaac

Clone 1              491
agaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaactctaacaccgt
JOE0063             1004                                              OR
agaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaactctaacaccgt

Clone 1              561
gcgtgttgactattttacctctggcggtgataatacatatatataaacttgcataaattggtcaatgcaa
JOE0063             1074
gcgtgttgactattttacctctggcggtgataatacatatatataaacttgcataaattggtcaatgcaa

Clone 1              631
gaaatacatatttggtcttttctaattcttagttttcaagttcttagatgctttctttttctctttttt
JOE0063             1144
gaaatacatatttggtcttttctaattcgtagttttcaagttcttagatgctttctttttctctttttt

Clone 1              701
acanatcatcaaggaannaantatctacttttacaacnantataaaacaatg-ggcatgcaatctaaag
JOE0063             1214
acagatcatcaaggaagtaattatctacttttacaacaaatataaaacaatgcggcatgcaatctaaag
```

JOE0064 (promoter region)

```
Clone 1                1
cacagatattataacntctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgaggaa
JOE0064              517
cacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgaggaa

Clone 1               71
ctatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcatact
JOE0064              587
ctatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcatact

Clone 1              141
attatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacgtgg
JOE0064              657
attatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacgtgg

Clone 1              211
cctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaaccca
JOE0064              727
cctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaaccca

Clone 1              281
gacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaaggtcctagcga
JOE0064              797
gacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaaggtcctagcga
```

```
Clone 1              351
cggctcacaggttttgtaacaagcaatcgaaggttctaacaccgtgcgtgttgactattttacctctggc
JOE0064              867           OR
cggctcacaggttttgtaacaagcaatcgaaggttctaacaccgtgcgtgttgactattttacctctggc

Clone 1              421
ggtgataattctggaatggcgggaaagggtttagtaccacatgctatgatgcccactgtgatctccagag
JOE0064              937
ggtgataattctggaatggcgggaaagggtttagtaccacatgctatgatgcccactgtgatctccagag

Clone 1              491
caaagttcgttcgatcgtactgttactctctctctttcaaacagaattgtccgaatcgtgtgacaacaac
JOE0064             1007
caaagttcgttcgatcgtactgttactctctctctttcaaacagaattgtccgaatcgtgtgacaacaac

Clone 1              561
agcctgttctcacacactcttttcttctaaccaaggggtggtttagttnagnagaacctcgtgaaactt
JOE0064             1077
agcctgttctcacacactcttttcttctaaccaaggggtggtttagtttagtagaacctcgtgaaactt

Clone 1              631
acatttacatatatataaacttgcataaattggtnnntgcaagaaatacatatttggnctttctaattc
JOE0064             1147
acatttacatatatataaacttgcataaattggtcaatgcaagaaatacatatttggtctttctaattc
```

JOE0065 (promoter region)

```
Clone 7                5
acgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgag
JOE0065              514
acgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgag

Clone 7               75
gaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcat
JOE0065              584
gaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcat

Clone 7              145
actattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacg
JOE0065              654
actattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacg

Clone 7              215
tggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaac
JOE0065              724
tggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaac

Clone 7              285
ccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacatctaacacc
JOE0065              794                                              OR
ccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacatctaacacc

Clone 7              355
gtgcgtgttgactattttacctctggcggtgataattctggaatggcgggaaagggtttagtaccacatg
JOE0065              864
gtgcgtgttgactattttacctctggcggtgataattctggaatggcgggaaagggtttagtaccacatg

Clone 7              425
ctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaaca
JOE0065              934
ctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaaca
```

```
Clone 7             495
gaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaaggggtggt
JOE0065            1004
gaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaaggggtggt

Clone 7             565
ttagtttagtagaacctcgtgaaacttacatttacatatatataaacttgcataaattggtcaatgcaag
JOE0065            1074
ttagtttagtagaacctcgtgaaacttacatttacatatatataaacttgcataaattggtcaatgcaag

Clone 7             635
aaatacatatttggtcttttctaattcttagtttttcaagttcttagatgctttctttttctctttttta
JOE0065            1144
aaatacatatttggtcttttctaattcgtagtttttcaagttcttagatgctttctttttctctttttta

Clone 7             705
cagatcatcaaggaagtaattatctacttttttacaacaaatataaaacaatg-ggcatgcaatctaaagg
JOE0065            1214
cagatcatcaaggaagtaattatctacttttttacaacaaatataaaacaatgcggcatgcaatctaaagg

Clone 7             774
tgaagaattattcactggtgttgtcccaattttggttgaattagatggtgatgttaatggtcac
JOE0065            1284
tgaagaattattcactggtgttgtcccaattttggttgaattagatggtgatgttaatggtcac
```

JOE0073 (promoter region)

```
Clone 4              9
gacgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtga
JOE0073             27
gacgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtga

Clone 4             79
ggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctca
JOE0073             97
ggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctca

Clone 4            149
tactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcac
JOE0073            167
tactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcac

Clone 4            219
gtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaa
JOE0073            237
gtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaa

Clone 4            289
cccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaaggtccta
JOE0073            307
cccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaaggtccta

Clone 4            359
gcgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttagtaccaca
JOE0073            377
gcgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttagtaccaca

Clone 4            429
tgctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaa
JOE0073            447
tgctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaa
```

```
Clone 4              499
cagaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcatctaaccaaggggggtg
JOE0073              517
cagaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaaggggggtg

Clone 4              569
gtttagtttagtagaacctcgtgaaacttacatttctaacgccgtgcgtgttgactattttacctctgg
JOE0073              587            OR
gtttagtttagtagaacctcgtgaaacttacatttctaacaccgtgcgtgttgactattttacctctgg

Clone 4              639
cggtgataatacatatatataaacttgcataaattggtcaatgcaagaaatactctaacaccgtgcgtgt
JOE0073              657                                  OR
cggtgataatacatatatataaacttgcataaattggtcaatgcaagaaatactctaacaccgtgcgtgt

Clone 4              709
tgactattttacctctggcggtgataatacatatataaacttgcataaattggtcaatgcaagaaata
JOE0073              727
tgactattttacctctggcggtgataat----atat----------------ttggtc---t-------t-

Clone 4              779
ctctaacaccgtgcgtgttgactattttacctctggcggtgataatatatttggtcttttctaattctta
JOE0073              767                          OR
-tt---c----t----------aatt----c---------g----ta----g-----------------

Clone 4              849
nttttcaagttcttagatgctttcttttttctctttttttacagatcatcaaggaagtaattatctacttt
JOE0073              780
-tttttcaagttcttagatgctttcttttttctctttttttacagatcatcaaggaagtaattatctacttt

Clone 4              919 ttacaac-aanataa
JOE0073              849 ttacaacaaatataa
```

JOE0074 (promoter region)

```
Clone 1                1
ttc-actc-nnacgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaag
JOE0074               17
ttcaactcaagacgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaag

Clone 1               69
gaaagagtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggc
JOE0074               87
gaaagagtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggc

Clone 1              139
ttcaccctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccct
JOE0074              157
ttcaccctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccct

Clone 1              209
cataaagcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaa
JOE0074              227
cataaagcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaa

Clone 1              279
actgaaaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaa
JOE0074              297
actgaaaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaa
```

```
Clone 1             349
caaggtcctagcgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtt
JOE0074             367
caaggtcctagcgacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtt

Clone 1             419
tagtaccacatgctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctct
JOE0074             437
tagtaccacatgctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctct

Clone 1             489
ctctttcaaacagaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaac
JOE0074             507
ctctttcaaacagaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaac

Clone 1             559
tctaacaccgtgcgtgttgactattttacctctggcggtgataatacatatatataaacttgcataaatt
JOE0074             577                    OR
tctaacaccgtgcgtgttgactattttacctctggcggtgataatacatatatataaacttgcataaatt

Clone 1             629
ggtcaatgcaagaaatactctaacaccgtgcgtgttgactattttacctctggcggtgataatatatttg
JOE0074             647                                          OR
ggtcaatgcaagaaatactctaacaccgtgcgtgttgactattttacctctggcggtgataatatatttg

Clone 1             699
gtcttttctaattcttagtttttcaagttcttagatgctttctttttctctttttttacagatcatcaagg
JOE0074             717
gtcttttctaattcgtagtttttcaagttcttagatgctttctttttctctttttttacagatcatcaagg

Clone 1             769
aagtaattatctacttttttacaacaaatataaaacaatg-ggcatgcaatctaaaggtgaagaattattc
JOE0074             787
aagtaattatctacttttttacaacaaatataaaacaatgcggcatgcaatctaaaggtgaagaattattc

Clone 1             838
actggngttgtcccaattttggttgaattagatggtgatgttaatggncacaaattttctgtctccggtg
JOE0074             857
actggtgttgtcccaattttggttgaattagatggtgatgttaatggtcacaaattttctgtctccggtg

Clone 1             908
aagnngaagnngatgctacttacngtaaattgaccttaaaatttanttgtactactg
JOE0074             927
aaggtgaaggtgatgctacttacggtaaattgaccttaaaatttatttgtactactg
```

JOE0075 (promoter region)

```
Clone 8               1
actc-agacgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaa
JOE0075             507
actcaagacgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaa

Clone 8              70
gagtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttca
JOE0075             577
gagtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttca

Clone 8             140
ccctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcata
JOE0075             647
ccctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcata
```

```
Clone 8             210
aagcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactg
JOE0075             717
aagcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactg

Clone 8             280
aaaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaag
JOE0075             787
aaaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaag

Clone 8             350
gtcctagcgacggctcacaggttttgtaacaagcaatcgaaggttctaacaccgtgcgtgttgactattt
JOE0075             857                      OR
gtcctagcgacggctcacaggttttgtaacaagcaatcgaaggttctaacaccgtgcgtgttgactattt

Clone 8             420
tacctctggcggtgataattctggaatggcgggaaagggtttagtaccacatgctatgatgcccactgtg
JOE0075             927
tacctctggcggtgataattctggaatggcgggaaagggtttagtaccacatgctatgatgcccactgtg

Clone 8             490
atctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaacagaattgtccgaatcgtg
JOE0075             997
atctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaacagaattgtccgaatcgtg

Clone 8             560
tgacaacaacagcctgttctcacacactcttttcttctaaccaaggggtggtttagtttagtagaacct
JOE0075            1067
tgacaacaacagcctgttctcacacactcttttcttctaaccaaggggtggtttagtttagtagaacct

Clone 8             630
cgtgaaacttacatttacatatatataaacttgcataaattggtcaatgcaagaaatactctaacaccgt
JOE0075            1137                                          OR
cgtgaaacttacatttacatatatataaacttgcataaattggtcaatgcaagaaatactctaacaccgt

Clone 8             700
gcgtgttgactattttacctctggcggtgataatatatttggtcttttctaattcttagttttcaagtt
JOE0075            1207
gcgtgttgactattttacctctggcggtgataatatatttggtcttttctaattcgtagttttcaagtt

Clone 8             770
cttagatgctttcttttttctctttttttacagatcatcaaggaagtaattatctacttttttacaacanata
JOE0075            1277
cttagatgctttcttttttctctttttttacagatcatcaaggaagtaattatctacttttttacaacaaata

Clone 8             840
taaaacaatg-ggcatgcaatctaaaggtgaagaattattcactggtgttgtcccaattttggttgaatt
JOE0075            1347
taaaacaatgcggcatgcaatctaaaggtgaagaattattcactggtgttgtcccaattttggttgaatt

Clone 8             909
agatggtgatgttaatggtcacaaattttctgtctccggtgaaggtgaaggtgatgctacttacggtaaa
JOE0075            1417
agatggtgatgttaatggtcacaaattttctgtctccggtgaaggtgaaggtgatgctacttacggtaaa

Clone 8             979
ttgaccttaaaatttatttgnactactggtaaattgccagttccatggccaaccttagtcactactttcc
JOE0075            1487
ttgaccttaaaatttatttgtactactggtaaattgccagttccatggccaaccttagtcactactttc-c

Clone 8            1049 ggttatgg
JOE0075            1556 ggttatgg
```

JOE0076 (promoter region)

```
Clone 1               1
acgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgag
JOE0076             514
acgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgag

Clone 1              71
gaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcgccctcat
JOE0076             584
gaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcat

Clone 1             141
actattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacg
JOE0076             654
actattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacg

Clone 1             211
tggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaac
JOE0076             724
tggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaac

Clone 1             281
ccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacatctaacacc
JOE0076             794                                                 OR
ccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacatctaacacc

Clone 1             351
gtgcgtgttgactattttacctctggcggtgataattctggaatggcgggaaagggtttagtaccacatg
JOE0076             864
gtgcgtgttgactattttacctctggcggtgataattctggaatggcgggaaagggtttagtaccacatg

Clone 1             421
ctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaaca
JOE0076             934
ctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaaca

Clone 1             491
gaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaaggggtggt
JOE0076            1004
gaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaaggggtggt

Clone 1             561
ttagtttagtagaacctcgtgaaacttacatttacatatatataaacttgcataaattggtcaatgcaag
JOE0076            1074
ttagtttagtagaacctcgtgaaacttacatttacatatatataaacttgcataaattggtcaatgcaag

Clone 1             631
aaatactctaacaccgtgcgtgttgactattttacctctggcggtgataatatatttggtcttttctaat
JOE0076            1144                              OR
aaatactctaacaccgtgcgtgttgactattttacctctggcggtgataatatatttggtcttttctaat

Clone 1             701
tcttagtttttcaagttcttagatgctttcttttttctcttttttacagatcatcaaggaagtaattatct
JOE0076            1214
tcgtagtttttcaagttcttagatgctttcttttttctcttttttacagatcatcaaggaagtaattatct

Clone 1             771
acttttttacaacaaatataaaacaatg-ggcatgcaatctaaaggtgaagaattattcactggngttgtc
JOE0076            1284
acttttttacaacaaatataaaacaatgcggcatgcaatctaaaggtgaagaattattcactggtgttgtc
```

```
Clone 1              840
ccaattttggttgaattanatggtgatgttaatggtcacaaattttctgtctccggtgaangntgaangt
JOE0076             1354
ccaattttggttgaattagatggtgatgttaatggtcacaaattttctgtctccggtgaa-ggtgaaggt

Clone 1              910 gatgctacttacggtaaattgaccttaaaatttatttgtactactg
JOE0076             1423 gatgctacttacggtaaattgaccttaaaatttatttgtactactg
```

JOE0081 (promoter region)

```
Clone 2              129
catttatcagggttattg-ctcatgagcggatacatatttgaatgtactcttgttacccatcattgaatt
JOE0081                1
catttatcagggttattgtctcatgagcggatacatatttgaatgtactcttgttacccatcattgaatt

Clone 2              198
ttgaacatccgaacctgggagttttccctgaaacagatagtatatttggacctgtataataatatatagt
JOE0081               71
ttgaacatccgaacctgggagttttccctgaaacagatagtatatttgaacctgtataataatatatagt

Clone 2              268
ctagcgctttacggaagacaatgtgtgtatttcggttcctggagaaactgttgcatctattgcataggta
JOE0081              141
ctagcgctttacggaagacaatgtatgtatttcggttcctggagaaactattgcatctattgcataggta

Clone 2              338
atcttgcacgtcgcatccccggttcattttctgcgtttccatcttgcacttcaatagcatatctttgtta
JOE0081              211
atcttgcacgtcgcatccccggttcattttctgcgtttccatcttgcacttcaatagcatatctttgtta

Clone 2              408
acgaagcatctgtgcttcattttgtagaacgaaaatgcaacgcgagagcgctaatttttcaaacaaagaa
JOE0081              281
acgaagcatctgtgcttcattttgtagaacaaaaatgcaacgcgagagcgctaatttttcaaacaaagaa

Clone 2              478
tctgagctgcatttttacagaacagaaatgcaacgcgaaagcgctattttaccaacgaagaatctgtgct
JOE0081              351
tctgagctgcatttttacagaacagaaatgcaacgcgaaagcgctattttaccaacgaagaatctgtgct

Clone 2              548
tcatttttgtaaaacaaaaatgcaacgcgagagcgctaatttttcaaacaaagaatctgagctgcatttt
JOE0081              421
tcatttttgtaaaacaaaaatgcaacgcgagagcgctaatttttcaaacaaagaatctgagctgcatttt

Clone 2              618
tacagaacagaaatgcaacgcgagagcgctattttaccaacaaagaatctatacttctttttgttctac
JOE0081              491
tacagaacagaaatgcaacgcgagagcgctattttaccaacaaagaatctatacttctttttgttctac

Clone 2              688
aaaaatgcatcccgagagcgctatttttctaacaaagcatcttagattactttttttctcctttgtgcgc
JOE0081              561
aaaaatgcatcccgagagcgctatttttctaacaaagcatcttagattactttttttctcctttgtgcgc

Clone 2              758
tctataatgcagtctcttgataacttttgcactgtaggtccgttaaggttagaagaaggctactttggt
JOE0081              631
tctataatgcagtctcttgataacttttgcactgtaggtccgttaaggttagaagaaggctactttggt
```

```
Clone 2              828
gtctatttttctcttccataaaaaaagcctgactccacttcccgcgtttactgattactagcgaagctgcg
JOE0081              701
gtctatttttctcttccataaaaaaagcctgactccacttcccgcgtttactgattactagcgaagctgcg

Clone 2              898
ggtgcatttttttcaagataaaggcatccccgattatattctataccgatgtggattgcgcatactttgtg
JOE0081              771
ggtgcatttttttcaagataaaggcatccccgattatattctataccgatgtggattgcgcatactttgtg

Clone 2              968
aacagaaagtgatagcgttgatgattcttcattggtcagaaaattatggacggtttcttctattttgtct
JOE0081              841
aacagaaagtgatagcgttgatgattcttcattggtcagaaaattatgaacggtttcttctattttgtct

Clone 2             1038
ctatatactacgtataggaaatgtttacattttnngtattgttttcgattcactctatgaannagttctta
JOE0081              911
ctatatactacgtataggaaatgtttacatttt-cgtattgttttcgattcactctatgaatagttctta

Clone 2             1108 ctacaatttttt
JOE0081              980 ctacaatttttt
```

JOE0086 (promoter region)

```
Clone 1                1
ctc-agacgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaag
JOE0086               22
ctcaagacgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaag


Clone 1               70
agtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcac
JOE0086               92
agtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcac

Clone 1              140
cctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataa
JOE0086              162
cctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataa

Clone 1              210
agcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactga
JOE0086              232
agcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactga

Clone 1              280
aaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacatct
JOE0086              302                                                 OR
aaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacatct

Clone 1              350
aacaccgtgcgtgttgactattttacctctggcggtgataattctggaatggcgggaaagggtttagtac
JOE0086              372
aacaccgtgcgtgttgactattttacctctggcggtgataattctggaatggcgggaaagggtttagtac

Clone 1              420
cacatgctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctttt
JOE0086              442
cacatgctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctttt
```

```
Clone 1              490
caaacagaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaaggg
JOE0086              512
caaacagaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaaggg

Clone 1              560
ggtggtttagtttagtagaacctcgtgaaacttacatttctaacaccgtgcgtgttgactattttacct
JOE0086              582                 OR
ggtggtttagtttagtagaacctcgtgaaacttacatttctaacaccgtgcgtgttgactattttacct

Clone 1              630
ctggcggtgataatacatatatataaacttgcataaattggtcaatgcaagaaatacatatttggtcttt
JOE0086              652
ctggcggtgataatacatatatataaacttgcataaattggtcaatgcaagaaatacatatttggtcttt

Clone 1              700
tctaattcttagtttttcaagttcttagatgctttctttttctctttttttacagatcatcaaggaagtaa
JOE0086              722
tctaattcgtagtttttcaagttcttagatgctttctttttctctttttttacagatcatcaaggaagtaa

Clone 1              770
ttatctacttttttacaacaaatataaaacaatg-ggcatgcaatctaaaggtgaagaattattcactggt
JOE0086              792
ttatctacttttttacaacaaatataaaacaatgcggcatgcaatctaaaggtgaagaattattcactggt

Clone 1              839
gttgtcccaattttggttgaattagatggtgatgttaatggtcacaaattttctgtctccggtgaaggtg
JOE0086              862
gttgtcccaattttggttgaattagatggtgatgttaatggtcacaaattttctgtctccggtgaaggtg

Clone 1              909
aaggtgatgctacttacgnnnaattgaccttaaaatttatttgtactactggtaaattgccagttccatg
JOE0086              932
aaggtgatgctacttacggtaaattgaccttaaaatttatttgtactactggtaaattgccagttccatg

Clone 1              979 gccaaccttagtcactactttcggttatgg
JOE0086             1002 gccaaccttagtcactactttcggttatgg
```

JOE0087 (promoter region)

```
Clone 1                1
ctc-agacgcacagatattataacatctgcataataggcatttgcaagaattactcgt-nnnaaggaaag
JOE0087               22
ctcaagacgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaag

Clone 1               69
agtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcac
JOE0087               92
agtgaggaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcac

Clone 1              139
cctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataa
JOE0087              162
cctcatactattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataa

Clone 1              209
agcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactga
JOE0087              232
agcacgtggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactga
```

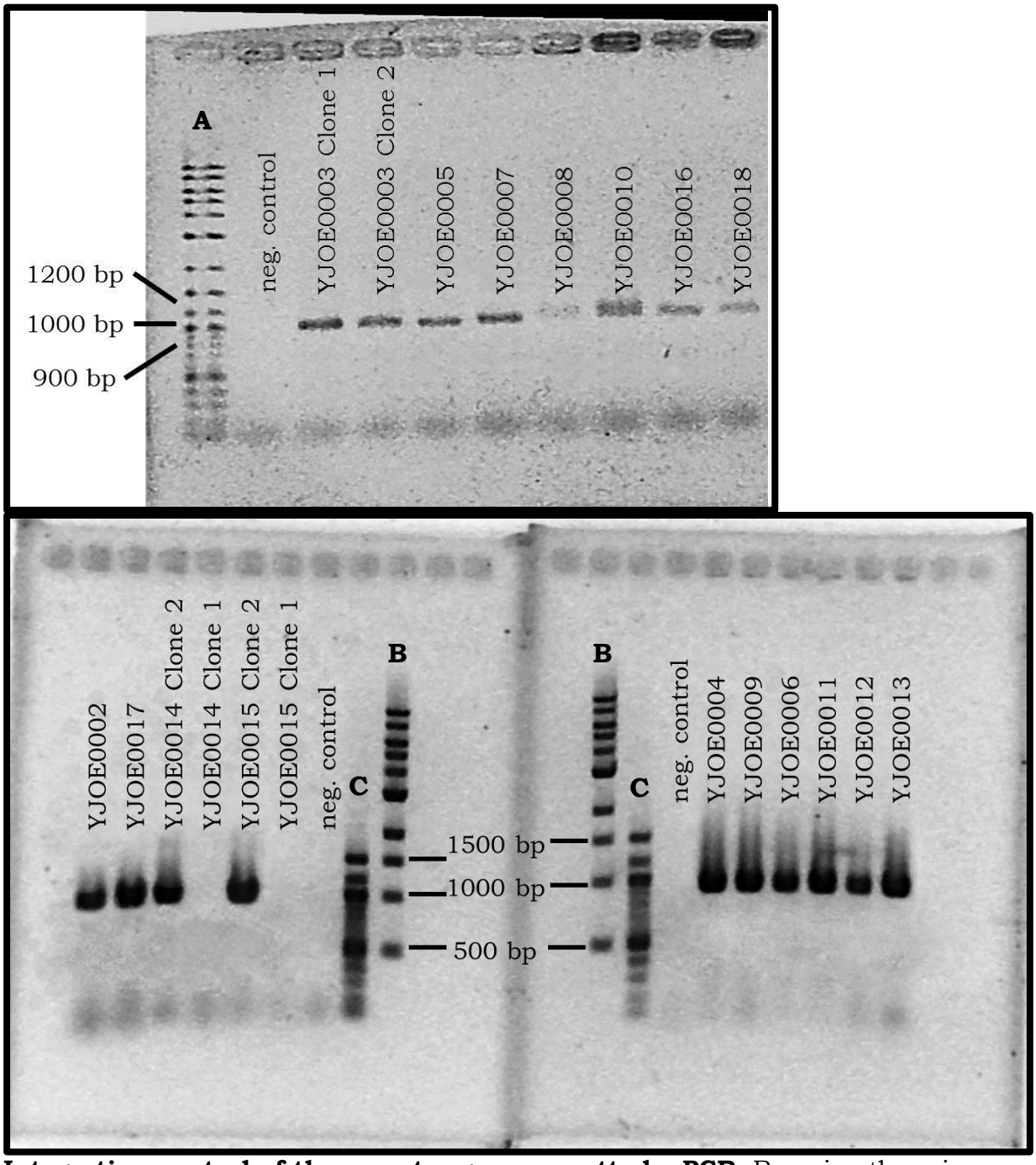
```
Clone 1             279
aaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaagg
JOE0087             302
aaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaagg

Clone 1             349
tcctagcgacggctcacaggttttgtaacaagcaatcgaaggttctaacaccgtgcgtgttgactatttt
JOE0087             372                    OR
tcctagcgacggctcacaggttttgtaacaagcaatcgaaggttctaacaccgtgcgtgttgactatttt

Clone 1             419
acctctggcggtgataattctggaatggcgggaaagggtttagtaccacatgctatgatgcccactgtga
JOE0087             442
acctctggcggtgataattctggaatggcgggaaagggtttagtaccacatgctatgatgcccactgtga

Clone 1             489
tctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaacagaattgtccgaatcgtgt
JOE0087             512
tctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaacagaattgtccgaatcgtgt

Clone 1             559
gacaacaacagcctgttctcacacactcttttcttctaaccaaggggtggtttagtttagtagaacctc
JOE0087             582
gacaacaacagcctgttctcacacactcttttcttctaaccaaggggtggtttagtttagtagaacctc

Clone 1             629
gtgaaacttacatttctaacaccgtgcgtgttgactattttacctctggcggtgataatacatatatat
JOE0087        OR   652
gtgaaacttacatttctaacaccgtgcgtgttgactattttacctctggcggtgataatacatatatat

Clone 1             699
aaacttgcataaattggtcaatgcaagananatacatatttggtcttttctaattcttagttttcaagttc
JOE0087             722
aaacttgcataaattggtcaatgcaagaaatacatatttggtcttttctaattcgtagttttcaagttc

Clone 1             769
ttagatgctttcttttttctctttttttacagatcatcaaggaagtaattatctacttttttacaacaaatat
JOE0087             792
ttagatgctttcttttttctctttttttacagatcatcaaggaagtaattatctacttttttacaacaaatat

Clone 1             839
aaaacaatg-ggcatgcaatctaaaggtgaagaattattcactggtgttgtcccaattttggttgaatta
JOE0087             862
aaaacaatgcggcatgcaatctaaaggtgaagaattattcactggtgttgtcccaattttggttgaatta

Clone 1             908
gatggtgatgttaatggtcacaaattttctgtctccggtgaaggtgaaggtgatgctacttacggtaaat
JOE0087             932
gatggtgatgttaatggtcacaaattttctgtctccggtgaaggtgaaggtgatgctacttacggtaaat

Clone 1             978
tgaccttaaaatttatttggactactggnaaattgccagttccatggccaacctta
JOE0087            1002
tgaccttaaaatttatttgtactactggtaaattgccagttccatggccaacctta
```

JOE0088 (promoter region)

```
Clone 5               5
acgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgag
JOE0088              28
acgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgag
```

```
Clone 5              75
gaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcat
JOE0088              98
gaactatcgcatacctgcatttaaagatgccgatttgggcgcgaatcctttattttggcttcaccctcat

Clone 5             145
actattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacg
JOE0088             168
actattatcagggccagaaaaaggaagtgtttccctccttcttgaattgatgttaccctcataaagcacg

Clone 5             215
tggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaac
JOE0088             238
tggcctcttatcgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaac

Clone 5             285
ccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacatctaacacc
JOE0088             308                                                 OR
ccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacatctaacacc

Clone 5             355
gtgcgtgttgactattttacctctggcggtgataattctggaatggcgggaaagggtttagtaccacatg
JOE0088             378
gtgcgtgttgactattttacctctggcggtgataattctggaatggcgggaaagggtttagtaccacatg

Clone 5             425
ctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaaca
JOE0088             448
ctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctctctttcaaaca

Clone 5             495
gaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaactctaacaccgtg
JOE0088             518                                       OR
gaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaactctaacaccgtg

Clone 5             565
cgtgttgactattttacctctggcggtgataatacatatatataaacttgcataaattggtcaatgcaag
JOE0088             588
cgtgttgactattttacctctggcggtgataatacatatatataaacttgcataaattggtcaatgcaag

Clone 5             635
aaatacatatttggtcttttctaattcttagttttcaagttcttagatgctttcttttttctctttttta
JOE0088             658
aaatacatatttggtcttttctaattcgtagttttcaagttcttagatgctttcttttttctctttttta

Clone 5             705
cagatcatcaaggaagtaattatctacttttttacaacaaatataaaacaatg-ggcatgcaatctaaagg
JOE0088             728
cagatcatcaaggaagtaattatctacttttttacaacaaatataaaacaatgcggcatgcaatctaaagg

Clone 5             774
tgaagaattattcactggtgttgtcccaattttggttgaattagatggtgatgttaatggncacaaattt
JOE0088             798
tgaagaattattcactggtgttgtcccaattttggttgaattagatggtgatgttaatggtcacaaattt

Clone 5             844
tctgtctccggtgaaggtgaaggngatgctacttacggtaaattgaccttaaaatttatttgtactactg
JOE0088             868
tctgtctccggtgaaggtgaaggtgatgctacttacggtaaattgaccttaaaatttatttgtactactg

Clone 5             914 gtaaattgccagttccatggcca
JOE0088             938 gtaaattgccagttccatggcca
```

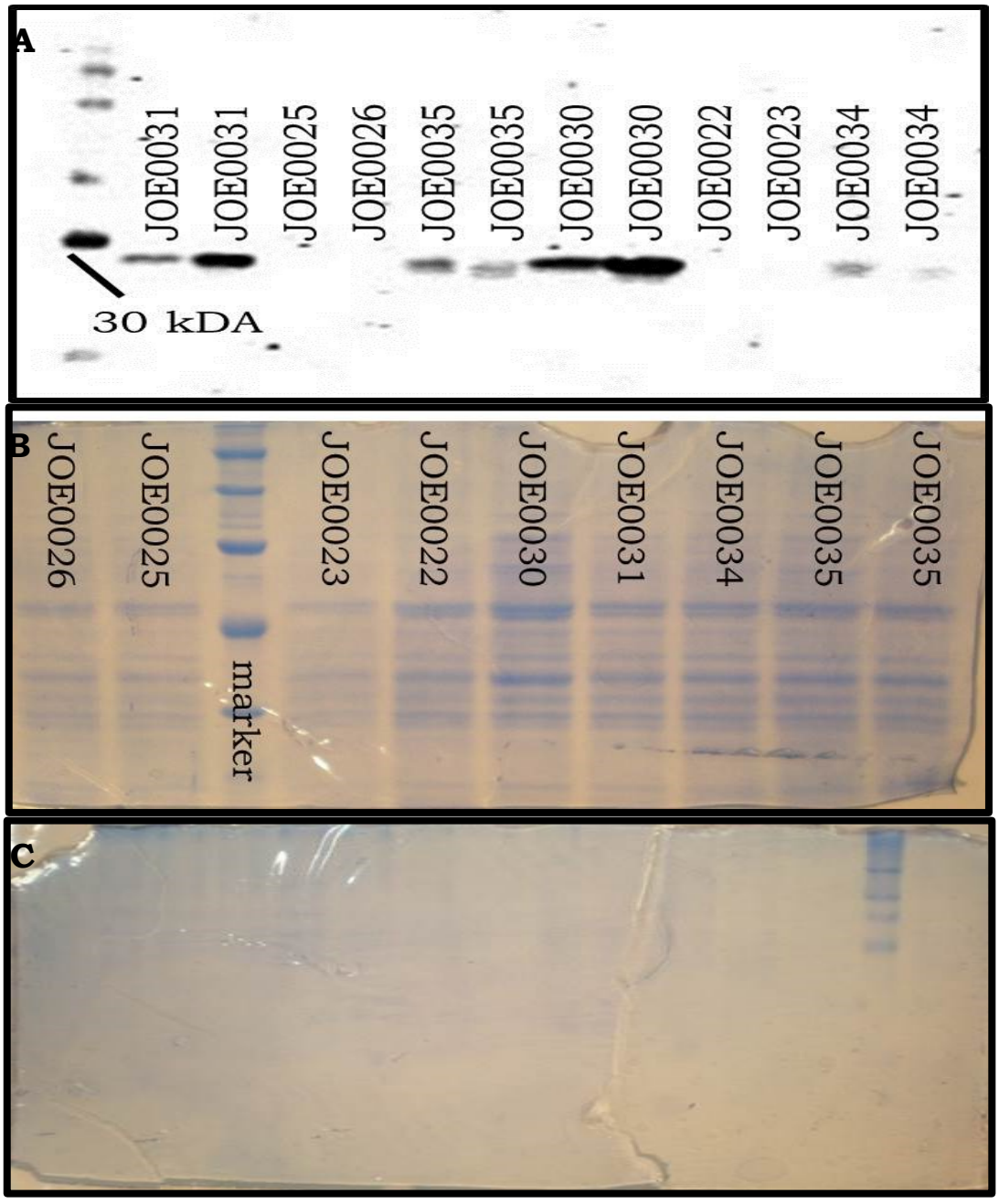
7.5 INTEGRATION CONTROL



Integration control of the reporter gene cassette by PCR. By using the primer pair forJOE0008 and revJOE0109 (see table 1) a specific DNA fragment could only be produced in combination with the respective chromosomal DNA isolation (see 2.5.2 and 2.5.6 for method details), if the expression cassette was integrated correctly into the genome. 15µl of each PCR were loaded on agarose gels with the size of the possible fragment depending on the promoter variation used in the reporter gene cassette beeing between 1022 bp (YJOE0003, YJOE0005, YJOE0007, YJOE0016, YJOE0018), 1047 bp (YJOE0012 YJOE0013),1067 bp (YJOE0002, YJOE0004, YJOE0006, YJOE0009, YJOE0017, YJOE0014, YJOE0015, YJOE0011) and 1112 bp (YJOE0008, YJOE0010), see table 2 for details. The negative control was based on 15 µl of a PCR with the chromosomal DNA isolation of SY992 and as size marker 8µl of the DNA ladders 2 Log (**A**), 1kpb (**B**) or 100 bp (**C**) from NEB (Frankfurt, Germany) were used and loaded.

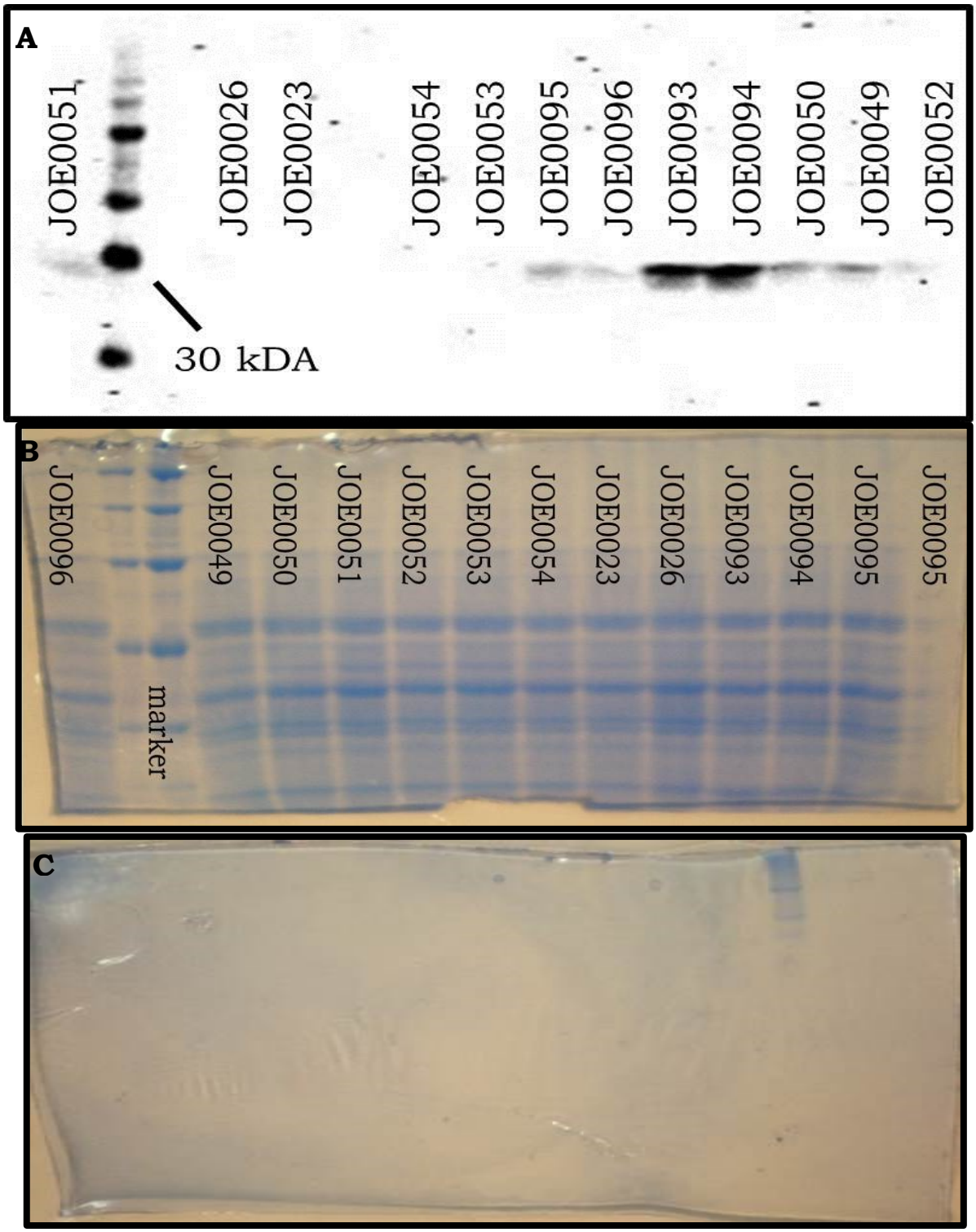
7.6 WESTERN BLOT ANALYSIS

7.6.1 REPORTER PROTEIN



Western blot analysis of different reporter protein variants (around 29 kDa each). Results of the western blot (**A**) with 5 µl and 10 µl of a total protein isolation (see 2.8 for methods) loaded for the yeasts containing plasmids (see table 3, 4 and 5 for details) JOE0030, JOE0031, JOE0034 as well as JOE0035. 15 µl were loaded for JOE0022, JOE0023, JOE0025 and JOE0026 after analysis of the control gel (**B**) for the quality of the used protein isolation, where always 7 µl was loaded. **C** represents the control gel of the protein transfer to the membrane. Additionally the markers Roti mark 10-150 PLUS protein marker (5 µl) or Roti mark standard protein marker (7µl) from Roth (Karlsruhe, Germany) were used.

7.6.2REPRESSOR PROTEIN



Western blot analysis of different reporter protein (around 29 kDa each) **and repressor variants** (around 28 kDa each). Results (**A**) of the western blot with 10 µl of a total protein isolation loaded (see 2.8 for methods) for the yeasts containing plasmids JOE0023, JOE0026, JOE0049, JOE0050, JOE0051, JOE0052, JOE0053, JOE0054 JOE0093, JOE0094, JOE0095 as well as JOE0096 (see table 3, 4 and 12 for details). **B** represents the control gel for the quality of the used protein isolation, where always 7 µl was loaded and **C** the control gel of the protein transfer to the membrane. Additionally the markers Roti mark 10-150 PLUS protein marker (5 µl) or Roti mark standard protein marker (7µl) from Roth (Karlsruhe, Germany) were used.

7.7 ACKNOWLEDGEMENTS

An dieser Stelle möchte ich mich bei den zahlreichen Menschen bedanken, die diese Arbeit mit ermöglicht und begleitet haben.

Besonders möchte ich mich dabei bei Herrn Prof. Dr. Tilman Achstetter bedanken, welcher mir über die ganze Zeit hinweg mit seiner Erfahrung und Wissen zur Seite stand und mich auch bei Rückschlägen neu motiviert hat.

Außerdem möchte ich mich bei Frau Prof. Dr. Barbara Reinhold-Hurek bedanken, dass Sie dieses Projekt und die Kooperation der Hochschule und Universität Bremen ermöglicht hat.

Dank auch an Herrn Prof. Dr. Johannes Hegemann und Frau Dr. Sabine Keppler-Ross für die Bereitstellung ihrer Plasmide.

Mein Dank gilt ebenso den vielen guten Seelen des UB Gebäudes, insbesondere Irina, Anne, Angela und Tanja, sowie meinen Leidensgenossen Simone und Dominik, welche mir schöne, lustige und langanhaltene Erinnerungen an die Hochschulzeit beschert haben, genauso wie die zahlreichen Studenten mit denen ich zusammen arbeiten und staunen durfte.

Vielen Dank auch an Lenchen für Rat, Tat und Spielplatzpausen.

Ein großer Dank gilt auch meiner ganzen Familie, welche mir jederzeit ein großer Rückhalt war und ist.

Und ganz besonders möchte ich mich bei Louisa bedanken, ohne dich und deine bedingungslose Unterstützung hätte ich es nicht geschafft. Du hast in der Zeit vom Labor bis zum Im-Stuhl-Schreiben für meine glücklichsten Momente gesorgt und wirst es auch in Zukunft. Danke.

7.8 DECLARATION

I hereby assure that I wrote the dissertation with the title “Development of an online reporter system to follow plasmid stability in *Saccharomyces cerevisiae*” independently and without any prohibited material. All used resources are stated with the respective reference.

Hiermit versichere ich, dass ich die vorliegende Dissertation mit dem Thema „Development of an online reporter system to follow plasmid stability in *Saccharomyces cerevisiae*“ selbstständig und ohne unerlaubte fremde Hilfe verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Johannes Öhlein