Régulation of transcription dynamics in single cells

Aymoz Delphine

Year : 2017

Regulation of transcription dynamics in single cells

Delphine Aymoz

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Département de Microbiologie Fondamentale

Regulation of transcription dynamics in single cells

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Delphine AYMOZ

Master de l'Université Joseph Fourier de Grenoble

Jury

Prof. Nicolas Perrin, Président
Dr. Serge Pelet, Directeur de thèse
Prof. Sophie Martin, Co-directrice
Prof. Nouria Hernandez, expert
Dr. Gaël Yvert, expert

Lausanne 2017

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Ecole Doctorale
Doctorat ès sciences de la vie

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président·e	Monsieur	Prof. Nicolas **Perrin**
Directeur·trice de thèse	Monsieur	Prof. Serge **Pelet**
Co-directeur·trice	Madame	Prof. Sophie **Martin**
Expert·e·s	Madame	Prof. Nouria **Hernandez**
	Monsieur	Prof. Gaël **Yvert**

le Conseil de Faculté autorise l'impression de la thèse de

Madame Delphine Aymoz

Master de l'Université Joseph Fourier, Grenoble, France

intitulée

Regulation of transcription dynamics in single cells

Lausanne, le 3 novembre 2017

pour le Doyen
de la Faculté de biologie et de médecine

Prof. Nicolas Perrin

Abstract

In mammalian cells, the MAPK pathways regulate vital processes such as differentiation, growth or death in response to a wide array of stimuli, like hormones or stresses. To achieve this, they interpret extracellular cues and reprogram the cell transiently or on long-term, through the transcriptional regulation of target genes. Understand the regulation of the MAPK signaling and resulting transcription is of high importance because of the involvement of mis-regulation of these pathways in many diseases (cancer, diabetes, cardiovascular diseases). As these signaling cascades are conserved among various organisms, we study them in the budding yeast *Saccharomyces cerevisiae*. This widely used model organism possesses its own MAPK network, and is easy to genetically modify.

The transcription is a crucial process in cellular life, as it drives the production of proteins involved in all possible aspects of life. As such, it is highly regulated through complex combination of factors. Moreover, the transcription is a dynamic process that can occur either continuously or in a bursty manner. These bursts can be variable in their frequencies and duration, leading to various mRNA productions within cells from a clonal population. Hence, the transcription regulation needs to be studied at the single cell level and in real-time.

To this purpose, we developed a new kind of gene expression reporters able to quantify in real-time and at the single cell level the expression arising from a promoter, that we named dPSTR. Using this system, we measured the extremely fast and transient gene expression resulting from hyper osmotic stress, and observed a difference on the sub-minute timescale in gene induction. Then, we used the dPSTR to quantify gene expression triggered by the mating signaling. We found that despite a signaling activity occurring minutes after addition of exogenous pheromone, resulting gene expression occurs on different timescales. Finally, by means of genetic analysis and modification of endogenous promoters, we defined some rules governing the expression kinetics of mating-induced genes.

Résumé

Dans les cellules mammifères, les cascades MAPK régulent des processus vitaux tels que la différentiation, la croissance ou la mort en réponse à de nombreux stimuli, tels que les hormones ou les stress. Pour parvenir à cela, les cascades interprètent des signaux extracellulaires et reprogramment la cellule de manière transitoire ou sur le long terme, en régulant la transcription de gènes cibles. Comprendre la régulation de la transduction du signal par ces cascades MAPK et de la transcription qui en découle est d'une importance capitale du fait de l'implication de la dérégulation de ces cascades dans de nombreuses maladies (cancer, diabète, maladies cardiovasculaires). Comme ces cascades de signalisation sont conservées dans divers organismes, nous les étudions dans la levure à bourgeon *Saccharomyces cerevisiae*. Cet organisme modèle, très souvent utilisé, possède son propre réseau de cascades MAPK, et il peut être facilement génétiquement modifié.

La transcription est un processus crucial dans la vie d'une cellule, puisqu'il contrôle la production de protéines impliquées dans tous les aspects possibles de la vie. De ce fait, elle est extrêmement régulée par de complexes combinaisons de facteurs. De plus, la transcription est un processus dynamique qui peut se produire de manière continue ou de manière explosive, que l'on appelle burst. Ces bursts varient selon leurs fréquences et leurs durées, ce qui amène à des rendements d'ARNm différents entre les cellules issues d'une population clonale. En conséquence, la transcription doit être étudiée au niveau des cellules uniques et en direct.

Dans ce but, nous avons développé une nouvelle classe de rapporteurs de l'expression génique capables de quantifier en temps réel et dans chaque cellule individuellement l'expression issue d'un promoteur, que nous avons appelé dPSTR. En utilisant ce système, nous avons mesuré des évènements d'expression génique très rapides et transitoires en réponse à un stress hyper-osmotique. Nous avons pu observer des différences temporelles de quelques dizaines de secondes dans l'induction de gènes. Ensuite, nous avons utilisé le dPSTR pour quantifier l'expression de gènes résultant de la transduction du signal de reproduction. Nous avons trouvé que malgré le fait que l'activité de transduction se produise dans les minutes suivant l'addition de phéromone exogène, l'expression génique qui en découle se produit à différentes échelles de temps. Pour finir, grâce à des analyses génétiques et des modifications de promoteurs endogènes, nous avons défini certaines règles gouvernant le moment de l'induction des gènes induit par la phéromone.

Résumé pour le grand public

Les cellules qui composent tout organisme vivant sont capables d'analyser leurs environnements proches, et d'en détecter les variations. Ces variations, ou signaux, sont décelées et interprétées par des cascades de signalisations. Elles sont formées de plusieurs protéines qui relaient à la chaine une information spécifique, jusqu'à ce qu'elle atteigne la protéine capable d'établir une réponse appropriée. Les cascades de signalisations sont capables de reconnaître différents signaux, comme les hormones ou certains stress, et les interpréter pour y répondre correctement. Par exemple, lorsqu'une cellule détecte une hormone de croissance, elle va croître et se diviser. Pour cela, elle va devoir produire de nouvelles protéines qui sont encodées dans son ADN sous la forme de gènes. Ce processus d'expression des gènes s'appelle la transcription. C'est un processus extrêmement régulé pour éviter de déclencher une réponse physiologique si elle n'est pas parfaitement adaptée à la situation.

Au cours de ma thèse, j'ai étudié la transcription établie en réponse à des cascades de signalisation chez la levure de boulanger. Cette levure est un micro-organisme formé d'une seule cellule, qui est très facile à étudier en laboratoire. Les cascades de signalisations qu'elle possède sont apparentées à celles qui régulent nos propres cellules, qui sont elles bien plus délicates à étudier en laboratoire. Comme pour les humains, chacune des cellules d'une population de levure est unique et va répondre à sa manière à un stimulus. De ce fait, nous essayons de mesurer dans chaque cellule la transcription qui résulte d'une stimulation externe transmise par une cascade de signalisation.

Le but de ma thèse a tout d'abord été de développer un nouvel outil permettant cela. Cet outil, que nous avons appelé dPSTR, est en fait une protéine fluorescente rouge que l'on peut visualiser dans les cellules à l'aide d'un microscope spécial. Elle nous permet de savoir dans quel état se trouve un gène: exprimé ou non exprimé. Pour cela, le dPSTR se déplace entre deux compartiments de la cellule: lorsque le signal rouge s'accumule dans le noyau, le gène est exprimé, alors qu'à l'inverse, lorsque le gène n'est pas exprimé, le signal rouge se situe dans toute la cellule.

Ce nouveau système apporte un point de vue différent sur de nombreuses questions biologiques car il permet de mesurer la transcription dans chaque cellule en temps réel, ce qui n'était pas possible avec les techniques disponibles jusqu'ici. Par exemple, nous avons pu voir à quel point la levure était capable de s'adapter rapidement à un stress de type osmotique grâce à l'expression de gènes dans les minutes suivant le stress. De plus, nous avons pu observer la grande variabilité dans l'expression de ces gènes entre cellules voisines d'une population.

A terme, nous espérons que notre système dPSTR sera utilisé pour de nombreuses études et dans différents types de cellules.

Sommaire

List of Figures

Remerciements

Merci aux membres du **jury** d'avoir accepté de faire partie de mon comité et d'avoir le courage de lire ce manuscrit.

Je voudrais remercier énormément **Serge Pelet**, mon superviseur, pour m'avoir donné cette superbe opportunité de venir travailler dans un environnement aussi incroyable que Lausanne. Merci d'avoir cru en moi, de m'avoir fait grandir intellectuellement, et de nous avoir poussé à parler anglais. Tu m'as donné le parfait encadrement pour que je puisse aboutir à ce travail, exigeant mais flexible, disponible, patient, de bon conseils, et à l'écoute des suggestions (et qu'est-ce que ça fait du bien d'entendre qu'on a de bonnes idées!). Je suis très heureuse d'avoir fait partie de ton équipe depuis le quasi-début, et j'espère vraiment qu'elle va continuer à grandir.
Je voudrais aussi remercier **Sophie Martin** pour tous ses conseils lors des lab meetings.
Je vais ensuite bien sûr remercier le **groupe Pelet** en général, parce que travailler avec des amis c'est quand même beaucoup plus sympas que travailler avec des inconnus, surtout en science où la communication est essentielle. **Eric**, toi et moi on s'aime d'amour vache, sans toi je ne serais pas allée au bout, et je ne serais probablement même pas venue au début. Toutes ces discussions qu'on a pu avoir, au bureau, en lab meeting, c'est ça qui nous a fait avancer tous les deux, et je te remercie pour ça. Je dois remercier aussi toutes mes petites mains, **Benoit, Barbara, Amy, Kévin, Réjane, Joan, Clémence, Christine, Marta** sans qui je ne serais jamais, mais alors jamais, allée aussi loin. Je remercie aussi les stagiaires de master **Jean-Jerrold et Mathieu,** pour leur très bon travail et aussi pour m'avoir supportée. Merci à tous ces jeunes d'avoir été aussi sympas au quotidien. Spéciale dédicace au **#CouchPeople** #LovUall #Hashtags #Licornasses
I want to thank the **Martin's lab**, previous and current members, for all the stimulating conversations, meetings, and advices, and for providing a friendly environement. Particulièrement, je voudrais remercier **Omaya**, parce que sans lui, je n'aurais jamais su que "il y a ce mec sympas qui monte son groupe à Lausanne, viens.", et je ne serais jamais venue… Then I want to thank **Felipe**, who was my first neighbor in the office, because he also gave me a lot of advices. **Vincent**, merci d'être aussi cool et de savoir autant de choses sur le troubleshooting et les PCR. **Aleks**, **BB**, it was really amazing working with you, you are such an inspiration! **Bita**, thank you for everything sweetie I am really sorry about how things happened.
I would like to thank the entire **DMF** of course, former and present members, this department is amazing, so many great people around, it is such a great workplace by the lake (and such a great couch).
Je vais maintenant passer à des remerciements (encore) plus personnels.
Pour la famille, merci d'abord à **la Flo**, pour tous ces moments sociaux du lundi, et toutes les bonnes bouffes. Merci ensuite à toute la famille proche, **mes parents**, qui m'ont soutenue dès le début, ma soeur **Emi**, mon frère&co (**TofAlexEleStebanOre**), merci d'avoir cru en moi et désolée de ne pas avoir été aussi présente que je l'aurais voulu.
Yannick, on sait, tous les deux, que tu es pour beaucoup dans l'aboutissement de cette thèse. Merci de m'avoir poussée à partir, à rester, et à m'accrocher.
Ma **Sousou**, merci pour ton soutien et ton amitié depuis toutes ces années! Pour **Tam**: tu vois j'aurais bien finit par publier deux - trois écrits, même si c'est pas du même genre qu'on imaginait au lycée..
Lapinou, merci pour tous ces lapins et ces discussions enflammées..

Je dois ensuite dire un millier d'énormes mercis à **Marta-Giulia-Mama**. Merci d'aimer le bench et pas le reste, cette thèse aurait été a moitié vide sans toi. Merci de m'avoir maternée, nourrie, corrigée, consolée, conseillée. Tu es plus qu'une amie pour moi, tu as littéralement changé ma vie et je ne te remercierais jamais assez d'être la personne que tu es. **#Bazmeg**

Jordan, merci infiniment pour ton soutien ces derniers mois, d'abord dans l'écriture de la thèse, et puis dans la vie. Merci de m'avoir montré que tout était possible, y compris d'écrire tout ça en un mois. Merci de m'avoir conseillée, réconfortée, supportée, rachetée, et donné enfin ce qu'on attendait.

General Introduction

Cells are able to decide of their fate according to their surrounding habitat. They can modify their metabolism, transcriptional program and physiology through sensing and interpretation of various signals emanating from the environment, such as hormones, nutrient concentrations, or biochemical properties of the medium.

Signaling is a crucial process to maintain homeostasis in cells and organisms. During embryonic development, cells differentiate upon sensing of morphogens gradients, which trigger the activation of specific transcriptional programs (Christian, 2012; Gurdon and Bourillot, 2001; Rogers and Schier, 2011). In humans, communication between cells from distinct organs can occur by the means of hormones. These signaling molecules travel through the circulatory system and drive various responses by activating signaling cascades upon receptor binding. For instance, production of insulin by pancreatic cells will regulate the concentration of sugar in the blood, by stimulating glucose uptake by muscle and adipose tissues. Another example of signaling molecules is growth factors, regulating cellular proliferation and differentiation.

In the signaling network of eukaryotic cells, the Mitogen-Activated Protein Kinase (MAPK) signaling pathways play an important role by transducing a wide array of stimuli, including mitogens, growth factors and cytokines (Chen et al., 2001; Dhillon et al., 2007; Kyriakis and Avruch, 2001; Marshall, 1994). They are also able to trigger responses to a variety of stresses that threaten cell survival, like UV, heat stress or osmotic shock. These pathways are not only able to sense and interpret a signal, but also to drive the adapted physiological response. They regulate very important processes in mammalian cells such as cell division, cell differentiation, or apoptosis (Dhillon et al., 2007; Sun et al., 2015; Zhang and Liu, 2002). The sensing of an input signal by a receptor leads to a cascade of events resulting in the activation of the MAPK pathway. A MAPK pathway is composed of kinases that are activated in series by phosphorylation: activated MAPK kinase kinase (MAP3K) phosphorylates a MAPK kinase (MAP2K), which in turn activates the MAPK. This last kinase controls the adaptation to the stimulus, through a transcriptional response and/or a post-translational regulation by phosphorylation. Even though MAPK pathways are often presented as linear routes, they usually are embedded in complex networks.

As these cascades are widely solicited in organisms, any perturbation could result in disastrous effects. Indeed, it was shown that the human ERK pathway, which normally regulates cell proliferation in response to growth factors and mitogens, is deregulated in a third of cancers (Dhillon et al., 2007; Sun et al., 2015). Constitutive activation of the pathway through mutations of upstream components leads to hyper-activity of the MAPKs ERK1/2, which drives tumor growth through an overexpression of genes whose products trigger cell cycle entry, and repression of genes inhibiting proliferation (Dhillon et al., 2007). Mis-regulation of MAPK signaling has moreover been associated with numerous other diseases touching various organs. For instance, the ERK pathway may be involved in diabetes and obesity by impairing insulin production (Chao et al., 2006; Lawrence et al., 2008). Also, p38, JNK and ERK MAPKs signaling have been shown to play a role in cardiac and vascular diseases (Muslin, 2008). Hence, the understanding of these pathways, their regulation and their output is crucial for scientists to fathom a wide array of diseases, and hope to develop efficient treatments.

MAPK signaling impacts over different timescales

MAPK signaling results in adaptation and modification of the cell physiology over different timescales (Murphy and Blenis, 2006; Murphy et al., 2002). The activated MAPK can

phosphorylate specific targets and modify their activity state, by having either a positive of a negative impact. This response occurs on a short timescale following signaling, and promotes a fast response to a specific stimulation (Chen and Thorner, 2007). Moreover, this effect can be transient, as the signaling is often counteracted by negative feedback mechanisms occurring through dephosphorylation of the substrates and the kinases (Jacoby et al., 1997; Mattison et al., 1999; Zhan et al., 1997). However, MAPK signaling can often profoundly modify the cell fate in the long term (Santos et al., 2007). In order to drive this change, the MAPK controls a transcriptional response that leads to a more stable response. Indeed, by newly producing specific proteins, the cells can modify their fate on a longer timescale by inducing more important changes on their physiology. For instance, in the case of differentiation, the cells rewire their transcriptional program to commit to a specific cell type and perform specific functions (Murphy et al., 2004). As such, the transcriptional output driven by a MAPK pathway is a critical aspect of the signaling response. This transcriptional process is highly regulated over different layers that connect the input stimulus to the output response. Indeed, wrongful induction of genes in inappropriate conditions can lead to deleterious fates, as previously illustrated with the examples of diseases.

The MAPK pathways are present in all eukaryotic organisms from yeast to mammals, where they control crucial cell-fate decisions (Widmann et al., 1999). In our group, we use the budding yeast *Saccharomyces cerevisiae* as a model organism for their study.

The budding yeast as a model organism

Saccharomyces cerevisiae is a unicellular eukaryote that for decades has been widely used as a model organism. The budding yeast has a short generation time, allowing one to obtain large amount of cells in a short time. It is the first eukaryotic organism that has been fully sequenced, and it has become a reference for genomic studies (Engel et al., 2014; Goffeau et al., 1996). It can be genetically engineered thanks to straightforward methods of cloning and transformation. Because budding yeast has a very efficient homologous recombination system, it is easy to mutate, insert or delete any gene of interest by insertion of heterologous DNA (Winzeler et al., 1999). Furthermore, this organism can grow in two forms, haploid or diploid, and it is possible to obtain combinations of mutants by crossing different haploids single mutants, dissecting the spores resulting from the sporulation of the diploid and grow them separately. Moreover, numerous molecular techniques are available in yeast, like co-immuno-precipitation or fluorescent tagging, which allow the identification of protein interactions and localizations.

Saccharomyces cerevisiae is a good model organism to study the regulation of biological processes that are conserved between eukaryotes. For instance, regulations of cell cycle progression or transcription have been widely studied in yeast, allowing to set a basis for studies in higher eukaryotes. Moreover, yeast has been used to understand several human diseases. Indeed, some human genes linked to diseases possess orthologs in yeast (roughly 400, (Smith and Snyder, 2006)). For instance, yeast allowed the study of mitochondrial myopathies by means of deletions, as yeast can survive without functional mitochondria. As another example, a study identified a mechanism in yeast, which led to the design of a drug to treat the human disease Friedreich's ataxia (Fleming et al., 2005; Muhlenhoff et al., 2002). Additionally, the yeast system was used to observe and understand the beginning of prion related diseases, because it also possesses proteins that can form prions. These widespread neurodegenerative diseases are difficult to study in human as they cause protein aggregates in brain cells. Yeast prion proteins are not homologous to the ones causing diseases in humans, but the introduction of the human proteins in yeasts allowed very interesting studies on prion formation (Lindquist et al., 2001; Smith and Snyder, 2006). For instance, essential work on Huntington's or

Alzheimer's diseases was performed in yeast (Colby et al., 2004; Komano et al., 1998; Zhang et al., 1997).

The budding yeast also possesses a well-characterized MAPK network, and its study has set grounds for understanding MAPK networks of higher eukaryotes (Chen and Thorner, 2007; Gustin et al., 1998). Interestingly, human ERK1 protein sequence alignment to yeast MAPKs shows a relative similarity between them (Atienza et al., 2000). In fact, some studies showed that human MAPKs were functional in the budding yeast and were able to fulfill yeast's MAPK functions (Atienza et al., 2000; Han et al., 1994; Takekawa et al., 1997).

The budding yeast MAPK network

Yeast cells are constantly analyzing the environment surrounding them. They are able to detect changes in nutrient availability, temperature, osmolarity or pressure and trigger an appropriate reaction to counteract these stresses. For this purpose, the budding yeast has evolved a complex MAPK network that controls cell fate in response to various signals (Figure 0.1). The network is composed of 4 pathways that are present in haploid vegetatively growing cells and a fifth one regulating sporulation, which will not be discussed here (Gustin et al., 1998). Three of these pathways promote cell survival upon a variety of stresses, and are functional in both the haploid and the diploid form (Gustin et al., 1998). These pathways promote survival to osmotic stress, cell wall damages, and nutrient starvation (Hohmann, 2002; Cullen and Sprague, 2012). The fourth pathway drives the response to mating pheromones, and is related to cellular communication. It drives mating of two haploid partners in order to form a diploid zygote (Bardwell, 2005).

Depending on the stimulus and the response triggered, the MAPK pathways act on different timescales. For instance, hyperosmotic stress drives a rapid and transient response to allow the cell to adapt to its new environment (Hohmann, 2002; Pelet et al., 2011). Once the cell has accommodated, the signaling is no longer required and the cell can resume its growth. On the opposite, nutrient starvation will drive pseudo-hyphal growth of the yeast cells in around 15-20 hours (Cullen and Sprague, 2012; Kron and Gow, 1995).

MAPK signaling results in modification of target effectors of the response by phosphorylation, and also in modification of the transcriptional program of the cells. These two effects allow the cell to respond to stimulation over different timescales, and are dissociable. Thanks to the advent of microarray analysis, the transcriptional programs driven by the yeast MAPK pathways are pretty well characterized (Breitkreutz et al., 2003; Capaldi et al., 2008; Gasch et al., 2000; Madhani et al., 1999; Roberts et al., 2000; Zeitlinger et al., 2003). Moreover, multiple genetic studies have led to the identification of transcription factors (TFs) targeted by MAPK signaling and controlling transcriptional response (Gustin et al., 1998).

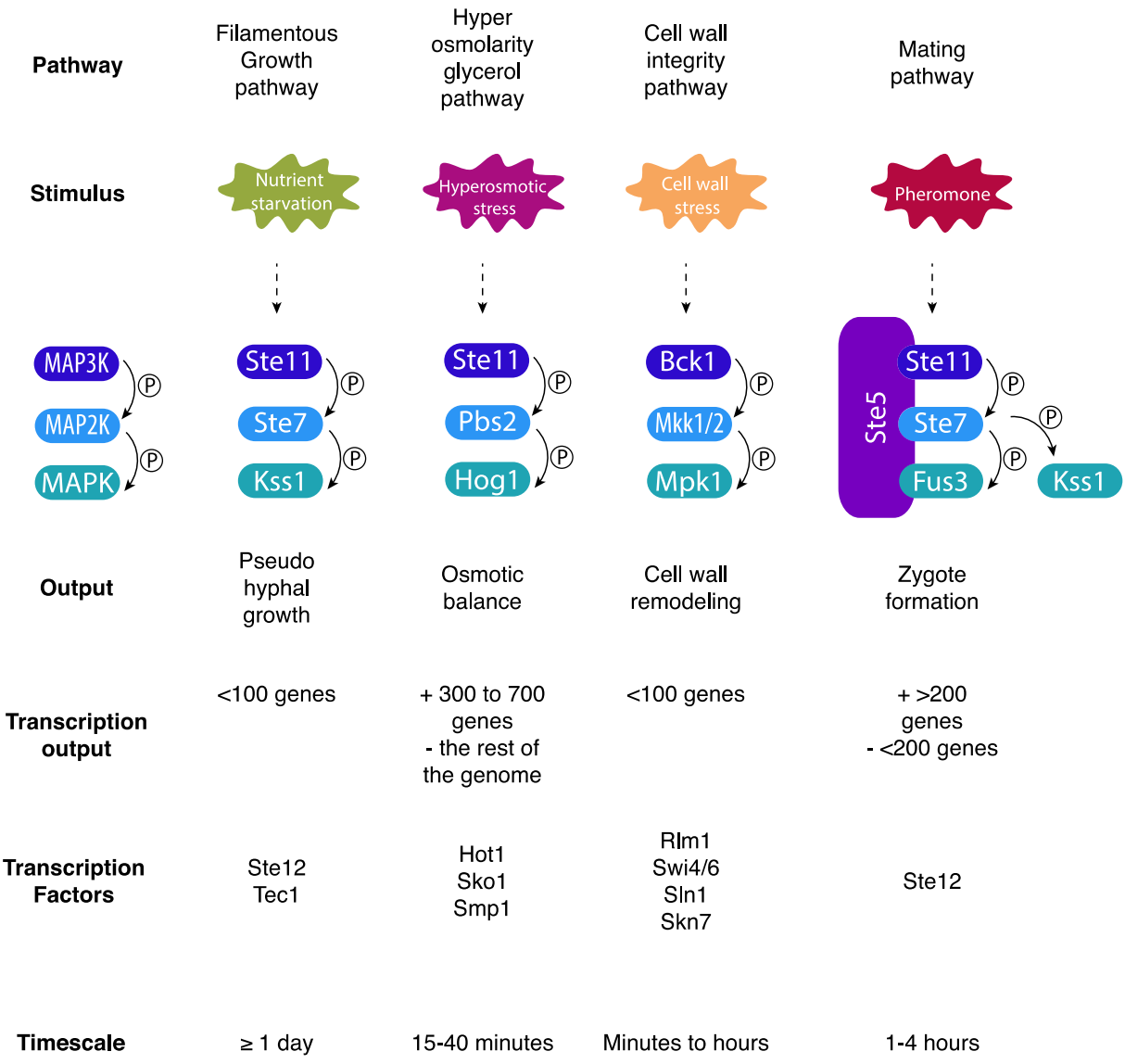


Figure 0.1 The MAPK network of the yeast *Saccharomyces cerevisiae*

The MAPK network of the budding yeast is formed of 5 pathways, including one solicited during sporulation and that is not represented here. Here, only the core MAPK pathway is represented (MAP3K -> MAP2K -> MAPK). The four pathways respond to a different stimulus, and promote various physiological responses (output). The transcriptional output is summarized with number of induced genes (+) and repressed genes (-). The main transcription factors for each pathway are indicated. These pathways act on different timescales to promote a response.

The High-Osmolarity Glycerol (HOG) pathway

Sudden increase in extracellular osmolarity leads to an immediate loss of cellular volume by outflowing water (Hohmann, 2002; Saito and Posas, 2012). The activation of the MAPK Hog1 within minutes following stress eventually allows the cell to restore its osmotic pressure, and to resume growth through the accumulation of a compatible solute (Maeda et al., 1995). To achieve this, two sensing branches promote the activation of the MAPK Hog1 (Figure 0.2). Activated Hog1 promotes glycerol accumulation in the cell to increase its osmolarity, which causes influx of water inside the cell and cell size recovery (Hohmann, 2002; Miermont et al., 2011; O'Rourke et al., 2002; Saito and Posas, 2012).

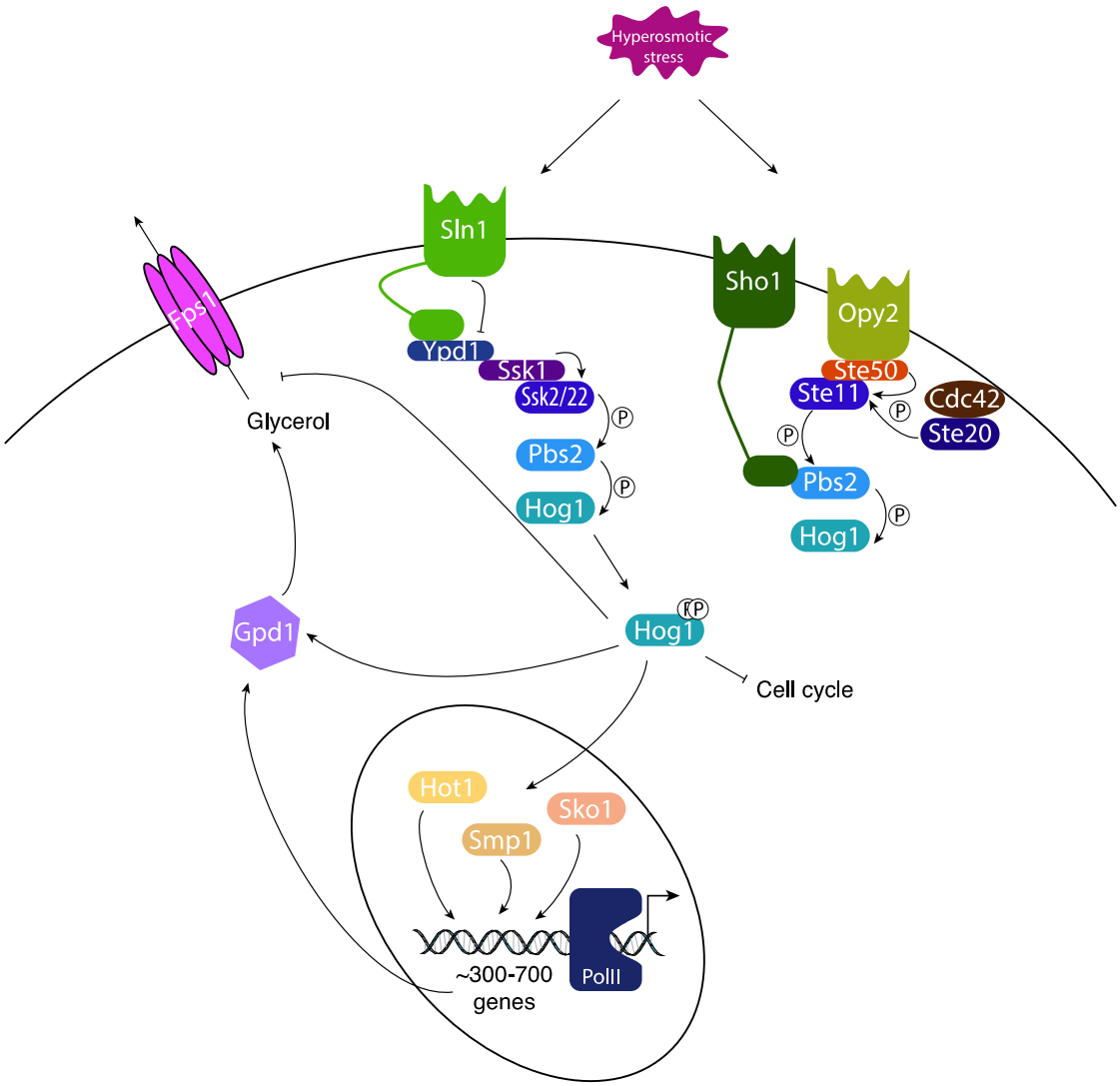


Figure 0.2 The HOG signaling pathway

Hyperosmotic stress drives the activation of two branches: the Sln1 branch (left) and the Sho1 branch (right). They both lead to the activation of the MAP2K Pbs2, which phosphorylates Hog1. The MAPK then controls different processes: transient inhibition of the cell cycle, induction of transcription of hundreds of genes, stimulation of glycerol production, closure of channels exporting glycerol outside of the cell.

Signaling in the HOG pathway

The first branch, the Sln1 branch, starts with the transmembranar osmo-sensor Sln1 (Maeda et al., 1994; Posas and Saito, 1998; Posas et al., 1996). In constant osmotic conditions, Sln1 interacts with and phosphorylates both Ypd1 and Ssk1, and these proteins act as a phosphor-relay. Upon hyper-osmotic stress, Sln1 loses its kinase activity, and Ypd1 and Ssk1 are dephosphorylated. The non-phosphorylated form of Ssk1 can interact with Ssk2, which drives its auto-phosphorylation. Ssk2 then activates Pbs2, which in turn phosphorylates Hog1.

The second branch is called the Sho1 branch. Sho1 binds Pbs2 in close vicinity to the membrane through its SH3 domain (Maeda et al., 1995; Tatebayashi et al., 2007). For a long time, Sho1 was believed to be the osmosensor but recent results are suggesting that it may rather act like an anchor protein, and other proteins have been proposed to be the osmosensors, like Msb2, Opy2 and Hrk1 (de Nadal et al., 2007; Reiser et al., 1999; Tatebayashi et al., 2007). Pbs2 brings back together the MAP3K Ste11 and the MAPK Hog1 at the membrane. There, Ste20 is activated by membrane osmosensors and phosphorylates Ste11, which in turn activates Pbs2,

which phosphorylates Hog1. The Ste50 protein is also involved as it promotes interaction of Ste11 with the membrane protein Opy2 (Wu et al., 2006).

Phosphorylated Hog1 accumulates in the nucleus and drives two kinds of responses leading to glycerol accumulation. Firstly, it quickly phosphorylates targets in order to change their activity and promote a rapid response. Activated Hog1 delays cell cycle progression by phosphorylating key regulators of different checkpoints (Clotet and Posas, 2007; Radmaneshfar et al., 2013; Yaakov et al., 2009). Very rapidly, cytosolic Hog1 promotes closure of the glycerol channel Fsp1 in order to favor osmolyte accumulation (Tamas et al., 1999). Additionally, Hog1 promotes activation of Gpd1, likely via its indirect dephosphorylation, to stimulate glycerol production (Blomberg and Adler, 1992; Lee et al., 2012; Vaga et al., 2014). Secondly, minutes after stress, Hog1 controls a transient transcriptional response that is explained below (Capaldi et al., 2008; Gasch et al., 2000; Nadal-Ribelles et al., 2012; Posas et al., 2000). 10 to 30 minutes after stress, depending on the severity of the stress, Hog1 is dephosphorylated by phosphatases and exported from the nucleus (Hohmann, 2002; Jacoby et al., 1997).

Negative feedback over HOG signaling

To ensure that the HOG-driven response is transient, several negative feedbacks allow the signal to arrest (Hohmann, 2002). First of all, as glycerol accumulates in the cell and causes volume recovery, the intracellular osmolarity increases to match the external osmolarity, leading to deactivation of the osmosensors. Moreover, Hog1 itself drives negative feedback by preventing interaction of Ste50 with Opy2. It also phosphorylates Sho1. Finally, phosphatases such as Ptp2, Ptp3, Ptc1, Ptc2 and Ptc3, some of which are induced at the transcriptional level by the HOG signaling, deactivate Hog1 (Gasch et al., 2000; Jacoby et al., 1997).

Transcriptional output of the HOG pathway

Activated Hog1 directly drives the transient expression of hundreds of genes whereas a general downregulation of transcription occurs (Nadal-Ribelles et al., 2012). Between 300 and 700 genes have been shown to be upregulated in response to hyper-osmotic shock (Capaldi et al., 2008; Gasch et al., 2000; Nadal-Ribelles et al., 2012; Posas et al., 2000). The transcriptional response generated by Hog1 allows a long-term adaptation to the stress. It is dispensable for immediate survival to the stress, but is required if a cell is challenged again by another osmotic stress (Westfall et al., 2008).

Among induced genes, we can cite *GPD1* and *GPD2*, which are glycerol producers, and *STL1*, which is a glycerol importer (Albertyn et al., 1994; Ferreira et al., 2005; Zhao et al., 1994). The presence of *GPD1* is essential for survival in hyper-osmotic conditions, as deletion results in very low glycerol accumulation under these conditions (Albertyn et al., 1994). Moreover, it is constitutively expressed under normal growth conditions, and is rapidly induced upon Hog1 activation by hyper-osmotic shock (Albertyn et al., 1994; Alepuz et al., 2001; Rep et al., 1999a). The expression of the *STL1* gene has been shown to occur within the 5 minutes following salt stress (Mas et al., 2009). It depends on a switch from repressed to open chromatin state, which requires the presence of the MAPK Hog1 to recruit the RSC and SAGA chromatin remodeler complexes (Mas et al., 2009; Zapater et al., 2007). Following stress induction, *STL1* expression is repressed by the replacement of nucleosomes by another chromatin-remodeling complex, INO80 (Klopf et al., 2009).

The role of Hog1 during transcription

The MAPK Hog1 controls the expression output at various levels. Once activated, it relocates into the nucleus and directly phosphorylates target TFs to activate them (de Nadal and Posas, 2010). Some act as transcriptional activators, like Hot1 and Smp1, and another one, Sko1, switches from repressor to activator upon Hog1 phosphorylation (de Nadal et al., 2011;

de Nadal and Posas, 2010). The TF allows Hog1 to bind to promoter sequences of osmo-stress responsive genes, recruit the RNA Pol II, and mediate transcription initiation (Alepuz et al., 2003). Hog1 drives recruitment at promoters of chromatin remodeler complexes, like Rpd3, SAGA and RSC, to remodel the chromatin and allow transcription (De Nadal et al., 2004; Mas et al., 2009; Zapater et al., 2007). The MAPK then travels with the RNA Pol II and the elongation complex along the gene coding sequences to promote chromatin remodeling (Proft et al., 2006). It was shown that the 3'UTR part of genes is involved in Hog1 traveling with RNA Pol II, which led to the hypothesis that osmostress-induced genes may form a loop between their promoter and terminator to recycle RNA pol II (Proft et al., 2006). The mRNA export is also regulated by Hog1 through phosphorylation of nucleoporins (Regot et al., 2013). Additionally, Hog1 has a role on stability of mRNA of osmostress-induced genes (Molin et al., 2009).

The mating-pheromone response pathway

Yeast cells can grow either in haploid of diploid forms. The haploid cells can exist in two mating types, MAT**a** and MATα. They can communicate and acknowledge the presence of cells of opposite mating type thanks to the constant production of mating pheromone (Figure 0.3). These pheromones, called **a**-factor and α-factor, are sensed through receptors, and drive the activation of the pheromone response MAPK pathway (Bardwell, 2005). The **a**-factor is sensed by the receptor Ste3 present in α cells, and the α-factor is sensed by the receptor Ste2 on the surface of MAT**a** cells. In the presence of potential mating partners of opposing mating type, cells respond to the pheromone present in the medium, and interpret the pheromone

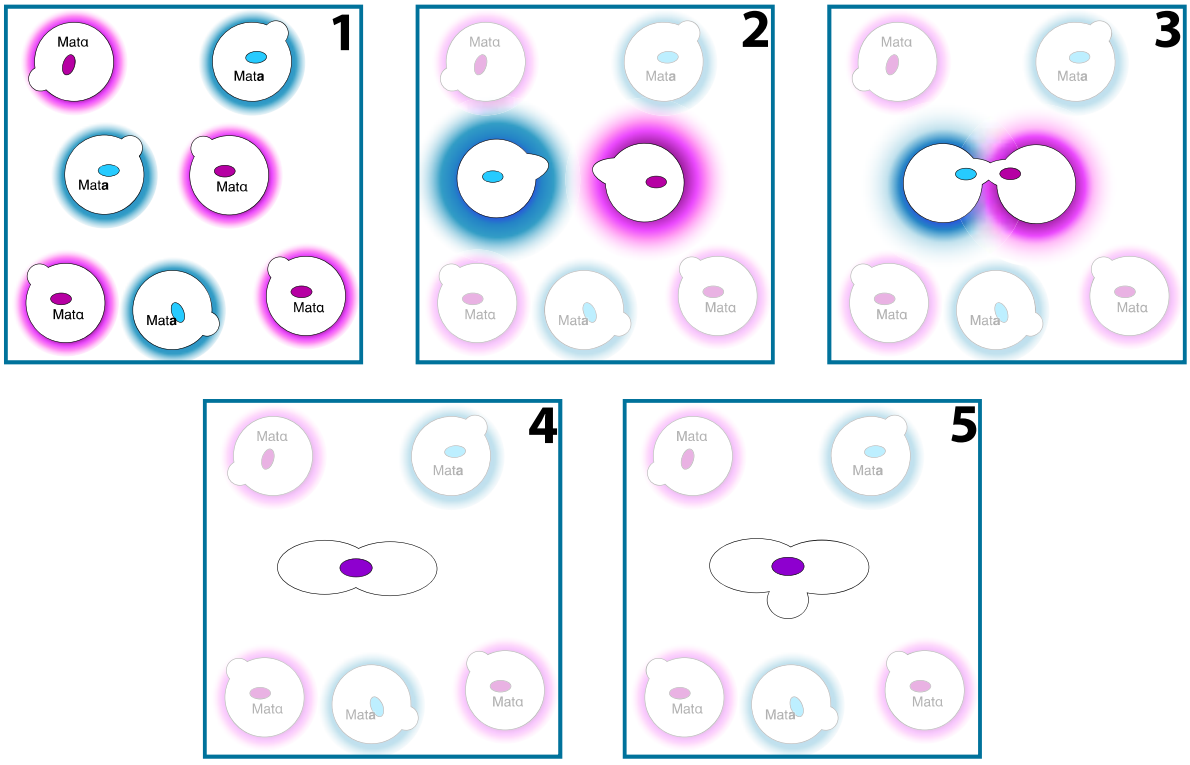


Figure 0.3 Overview of the mating process

In a mating mixture, cell produce mating pheromone at a constant rate (1). When they sense the presence of a mating partner through its pheromone, cells commit to the mating, start to produce more pheromone, and polarize their growth towards the mating partner (2). When mating partner touch, their cell wall and membrane fuse, and cytoplasmic contents are mixed (3). Then, the two nuclei fuse in a process called karyogamy (4). The resulting diploid zygotes then enter in a new round of cell cycle (5).

gradient in order to find the closest source (Bardwell, 2005; Merlini et al., 2013). The cell then initiates polarized growth in the direction of the pheromone gradient, a process also called shmooing, allowing the two partners to come closer to each other (Moore et al., 2008). Then, the cell wall of the two partners is degraded, and membranes fuse (Merlini et al., 2013). After fusion, the two nuclei also fuse in a process called karyogamy, and the resulting diploid cell starts a new round of cell cycle. The process of mating, from signaling to zygote formation, can take several hours to occur (Bardwell, 2005).

Cells growing in a mating mixture have to choose their fate: mate or divide. As mating signaling results in cell cycle arrest in G1, commitment to the mating process can be deleterious in terms of fitness in case of mating failure (Atay and Skotheim, 2017). Indeed, cells spend time in trying to mate instead of proliferating, which is not competitive.

Signaling in the mating pathway

The binding of α-factor to Ste2 triggers a cascade of events leading to the activation of the mating pathway (Figure 0.4A). Ste2 is a G-coupled protein receptor, and interaction with pheromone leads to the detachment of the Gα from the Gβγ proteins. The Gβγ then binds to Ste20 and the Ste5-Ste11 complex. The scaffold Ste5 is recruited to the membrane, which is necessary and sufficient to promote MAPK signaling (Elion, 2001). Ste5 brings the MAP3K Ste11, the MAP2K Ste7 and the MAPK Fus3 closer to each other. Ste20 phosphorylates Ste11, which in turn activates Ste7. The MAP2K Ste7 then phosphorylates Fus3. As both Ste11 and Ste7 participate to the signaling of the filamentous growth (FG) pathway, the FG MAPK Kss1 is also activated. The requirement of Ste5, which is not solicited during FG response, is one way to conserve signaling specificity, as Fus3 is a rather poor substrate for Ste7 in absence of Ste5 (Good et al., 2009; Maleri et al., 2004). The activation of Fus3 by this signaling branch will drive cell cycle arrest in G1 and establishment of a transcriptional program (Atay and Skotheim, 2017). Indeed, Fus3 activates Far1, which in turn inactivates the G1 cyclin Cdk1 complexes, leading to cell cycle arrest in G1 (Bardwell, 2005; Elion et al., 1993; Gartner et al., 1998; Peter et al., 1993; Peter and Herskowitz, 1994; Tedford et al., 1997). Moreover, Fus3 promotes activation of the main TF Ste12 by inhibition of its repressors Dig1/Dig2 (Cook et al., 1996; Tedford et al., 1997). Ste12 then promotes the induction of roughly 200 genes by binding their promoter sequence on precise regulatory elements (Errede and Ammerer, 1989; Roberts et al., 2000; Zeitlinger et al., 2003)(See Chapter 4 for more details). Additionally, the cell cycle arrest mediates the repression of hundreds of genes involved in cell cycle progression (Bardwell, 2005; Roberts et al., 2000).

A second branch is activated by the Gβγ complex, the polarity branch, and drives the polarized growth towards the mating partner (Figure 0.4B) (Atay and Skotheim, 2017). This branch involves the CDK inhibitor Far1, which can associate to Cdc24 in the nucleus (Butty et al., 1998). Upon mating stimulation, Far1-Cdc24 relocates to the cell membrane and binds the Gβγ complex. The GEF Cdc24 then activates the GTPase Cdc42. The activation of Cdc42 promotes the recruitment of polarity patches at the membrane, and these patches recruit formins that tether actin cables to the shmoo tip. These cables serve to deliver vesicles containing proteins for the cell wall synthesis for the polarized growth, and, later, for cell wall degradation preceding the fusion (Atay and Skotheim, 2017; Merlini et al., 2013). In the absence of Far1, cells are still able to promote polarized growth, although the presence of Far1 is required for the shmoo to be correctly oriented in direction of the pheromone gradient (Valtz et al., 1995). Moreover, Cdc42 activates Ste20, leading to the activation of the MAPK pathway via the mechanism described above.

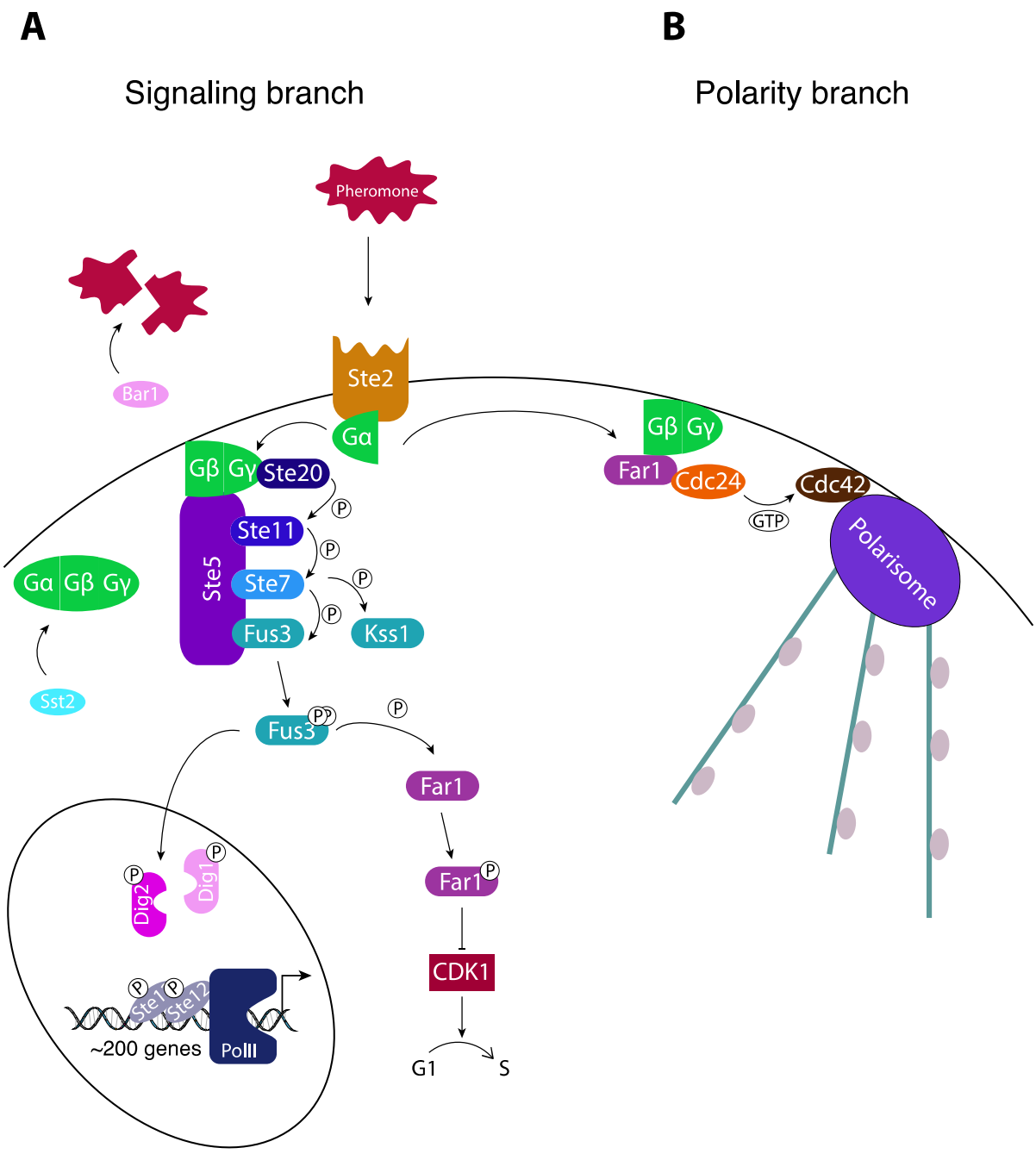


Figure 0.4 The mating signaling pathway

A. Pheromone drives the activation of the mating MAPK pathway, leading to activation of the MAPKs Fus3 and Kss1. Fus3 promotes expression of hundreds of genes through deregulation of the TF Ste12, and the cell cycle arrest in G1 through phosphorylation of Far1.

B. The polarity branch is activated independently, and leads to polarized growth in direction of the pheromone gradient.

Negative feedback for the signaling

To ensure that committing to the mating response is the appropriate choice, the cell has developed several negative feedback mechanisms to attenuate signaling of the pheromone pathway. First of all, the cell secretes a protease, Bar1, able to degrade the α-factor. This shapes the pheromone landscape around the cells (Barkai et al., 1998; Jin et al., 2011). Moreover, pheromone degradation by Bar1 prevents the full activation of the mating response in case of low pheromone concentration, which would correspond to a situation where no partner is available.

Another negative feedback lies in the Sst2 protein, which promotes association of the Gα to the Gβγ proteins (Yu et al., 2008). This leads to the arrest of signaling from the top of the mating cascade.

Also, Ptp2, Ptp3 and Msg5 phosphatases deactivate the MAPK Fus3 by dephosphorylating it (Doi et al., 1994; Zhan et al., 1997). Moreover, Msg5 is also transcriptionally induced by pheromone response (Roberts et al., 2000). This constant deactivation is counterbalanced by activation of Fus3, unless signaling is stopped upstream of the MAPK.

Another negative feedback mechanism is the degradation of the TF Ste12 that occurs about an hour after pheromone signaling (Esch et al., 2006). This leads to arrest in transcription of mating genes in case the mating signaling stopped.

Transcriptional output of the mating pathway

The activation of the mating pathway leads to derepression of the main mating TF Ste12. This transcription activator binds to DNA sequences defined as Pheromone Response Elements (PRE), found in promoter sequences of mating responsive genes (Chou et al., 2006; Dolan et al., 1989; Errede and Ammerer, 1989; Zeitlinger et al., 2003). Ste12 drives the induction of roughly 200 genes encoding various proteins required for correct progression of the mating process (Roberts et al., 2000). For instance, both Far1 and Ste12 are induced by pheromone stimulation. We can also cite Bar1 and Sst2, involved in the negative regulation of the mating pathway. Additionally, Ste12 drives the expression of proteins required for cell aggregation (like Aga1, Fig2), for cell-cell fusion (like Fus1, Fus2, Prm1, Fig1) and for karyogamy (Kar3, Cyk1, Prm3) (Roberts et al., 2000; Zeitlinger et al., 2003). Moreover, dozens of genes are repressed upon mating stimulation, which mostly encode proteins involved in growth (Roberts 2000). This repression has been shown to be dependent on cell cycle arrest driven by Far1 (Bardwell, 2005; Roberts et al., 2000).

Ste12 can associate with other TFs to promote gene expression. It can associate with Mcm1 to induce expression of some mating genes, like *FAR1*, in specific stages of the cell cycle (Errede 1989: Dolan 1989; Oehlen 1996). It can also bind to the TF of the FG response, Tec1 (Chou et al., 2006; Madhani, 1997). Upon mating stimulation, Fus3 phosphorylates Tec1 to target it for degradation, in order to avoid the induction of FG driven genes (Chou et al., 2006). This way, the cell avoids cross-talks between these two pathways.

Fus3 is also recruited to the chromatin of 9 mating target genes, including *AGA1*, *PRM1*, *FUS1* and *KAR4*, with both Kss1 and Ste5 (Pokholok et al., 2006).

In the absence of Ste12, cells are unable to establish the transcriptional response and to mate (Hartwell, 1980; Roberts et al., 2000). Thus, the transcriptional response is essential for mating occurrence.

Transcription is essential

As we illustrated before, transcription is a crucial process required for many aspects of cellular life, whether they are related to growth, like cell cycle or metabolism, or whether they are adaptive and induced by deleterious environmental conditions. It is the first step by which

cells produce proteins, central actors of all living processes. Therefore, transcription is a highly regulated process.

A typical gene is composed of a coding sequence, also called open reading frame (ORF), which is formed by a certain number of codons, or tri-nucleotides, encoding a specific amino acid. The sequence upstream of the ORF is called promoter, and controls to a large extent the expression of the gene thanks to the presence of upstream regulatory elements. The pre-initiation complex assembles on the core promoter sequence, from where it starts transcribing DNA into mRNA. This complex is formed of general transcription factors (TFIIA-B-D-E-F-H) that scan the DNA sequence and recruit the RNA polymerase II at the correct site. Downstream of the ORF, the terminating sequence allows the RNA Pol II complex to dissociate and free the mRNA produced (Hahn and Young, 2011).

Nowadays, it is straightforward to identify the beginning and the end of a coding sequence, and to determine the amino acid sequence of the encoded protein. Furthermore, it is even possible to predict, to a certain extent, the function of the encoded protein. Indeed, computational biology now allows to recognize similarities between reference proteins and unknown proteins, and infer the presence of domains that can define the roles of a protein. Moreover, it is also possible to predict some structural features of the encoded protein, without even knowing its function. However, prediction of the expression level of a gene based on its DNA sequence remains challenging. Hence, a central remaining question in genetics is the understanding of rules underlying the gene expression (Weingarten-Gabbay and Segal, 2014).

Many studies of gene expression provide snapshots of a large population of cells, mixing millions of individual cells together. They lead to the false impression that protein levels are homogeneous between different cells and constant in time. However, the picture changes drastically when gene expression is directly measured at the single cell level. Indeed, mRNA synthesis measurements in single cells revealed that gene expression is a highly stochastic and heterogeneous process (Larson, 2011; Larson et al., 2011; Zenklusen et al., 2008). Transcription occurs in bursts defined by periodic recruitment of the transcription machinery. These bursts may vary in frequency and duration, leading to the production of different mRNA levels. In constitutively expressed genes, initiation of transcription is believed to occur as distinct events uniformly spaced in time that lead to a relatively constant rate of mRNA production. However, transcription can also occur in bursts of multiple initiation events very close in time, taking advantage of the fact that chromatin is already in an open state. These bursts are followed by periods during which the gene is not transcribed (Zenklusen et al., 2008). This mode of expression often occurs at induced genes, such as stress-responsive genes, and relies on the recruitment of chromatin remodelers to maintain the promoter in an open conformation. It was shown that the variability in expression of constitutively expressed genes between cells of a population is very low. This variation depends mostly on the abundance of general transcription factors and ribosomes in the cells, or on the cell cycle stage (Raser and O'Shea, 2004). However, the transcription of the bursty genes is highly heterogeneous between cells because the transcription bursts can vary greatly in their frequencies and durations, which leads to various amounts of transcripts produced between cells. Furthermore, the dynamic processes occurring at the promoter, such as TF recruitment or chromatin remodeling, contribute significantly to the expression noise (Coulon et al., 2013; Pelet et al., 2011).

Hence, the transcription is a dynamic and noisy process. To understand its regulation, it has to be studied in real-time and at the single cell level.

Towards the understanding of grammatical rules

To decipher grammatical rules of transcription, different aspects have to be understood (Weingarten-Gabbay and Segal, 2014). The first one is the transcription factors. In the budding yeast genome, at least 169 genes encode TFs, which makes them one of the most abundant

classes of proteins (Hahn and Young, 2011; Teixeira et al., 2014). These TFs are regulated via different means, and via different inputs, as we saw in the examples of the HOG and the mating pathways. They can interact with promoters by direct binding to specific sequences, termed transcription factor binding sites (TFBS). For some TFs, consensus sequences have been identified and grouped in a database that allows to predict putative TFs interaction for any DNA sequence (Teixeira et al., 2014). For instance, a consensus sequence has been identified for the binding of Ste12, and its presence is often correlated to a mating-induced expression (Chou et al., 2006; Dolan et al., 1989; Errede and Ammerer, 1989; Zeitlinger et al., 2003). For other TFs, like Hot1, no consensus binding site has been identified. However, depending on how close or remote the sequence is to the consensus, the affinity of the TF for the TFBS may be increased or decreased, and it was shown that affinity to the TFBS could be a determinant for the expression (Kheradpour et al., 2013; Su et al., 2010; Weingarten-Gabbay and Segal, 2014). Besides, knowing that some regulatory elements are present in a promoter sequence is not sufficient to predict whether the TF will be active or not. Indeed, in the case of mating-induced genes, the TF Ste12 is repressed under normal growth condition, preventing their expression, even if they possess a consensus TFBS (Cook et al., 1996; Olson et al., 2000; Tedford et al., 1997). Moreover, the orientation of a TFBS can be of high importance for its capacity to promote induction, in cases where the binding site is not a palindromic sequence, like for Ste12. Additionally, the number of TFBSs and the distance between them can be crucial for gene expression. TFs can simply form multimers, but there can be some more complicated cases of cooperative binding or lateral diffusion of TFs over an array of TFBSs (Coleman and Pugh, 1995; Hertel et al., 1997; Weingarten-Gabbay and Segal, 2014). Hence, a relation may exist between number of TFBSs and gene expression. Furthermore, TFs can act in combination with other factors, leading to the integration of multiple signals for the expression. This further complicates the prediction of a promoter output from its sequence. Finally, a major determinant in gene expression is the organization of the chromatin in the promoter region. Indeed, DNA is compacted into chromatin, where it is wrapped around histone proteins in the form of nucleosomes. These nucleosomes can mask important regulatory sequences, and the recruitment of chromatin remodelers to the promoter may be required to observe transcription. The position of these nucleosomes on the genome is only poorly predicted by the DNA sequence (Field et al., 2008; Kaplan et al., 2009; Segal et al., 2006). However, studies reported the positions of nucleosomes over the entire yeast genome. (Brogaard et al., 2012; Jiang and Pugh, 2009). Hence, this may be the more complex obstacle in the final understanding of the rules determining transcription.

Objective of the thesis

To summarize, we have established that transcription is a dynamic, noisy and regulated process. The promoter is one major determinant of gene expression output. In order to understand grammatical rules of transcription, we need to study promoter-driven expression in real-time and at the single cell level.

MAPK-driven transcription is crucial for cell adaptation to environmental cues. Through the examples of gene expression driven by the stress-induced HOG pathway and the pheromone-induced mating pathway, we will try to decipher basic rules for transcription induction in response to MAPK signaling. These examples are very interesting models for different reasons. First, the genes targeted by these pathways are weakly induced or repressed in normal growth conditions, and strongly induced in response to stimulation. Furthermore, these pathways can easily be induced by addition of external molecules. Second, they drive a very dynamic response. Third, they intrinsically drive a very heterogeneous response between single cells, because of stochastic biochemical processes required, like chromatin remodeling, and also because of cell cycle regulation in case of the mating response (Colman-Lerner et al.,

2005; Pelet et al., 2011) As such, they offer a good opportunity to study rapid transcription establishment in response to a stimulus, and observe variability between cells from the same population.

At the time when I started this thesis, methods of gene expression measurements failed to fulfill the requirements we defined before: real-time study of gene expression at the single cell level. Hence, our first goal was to develop a new family of gene expression reporters, able to provide live quantification of protein expression driven by a promoter. This system, which we named dPSTR, is based on the quantification of fluorescent protein relocation, and allows to measure gene expression levels and kinetics in hundreds of single living cells.

Outline of the thesis

In the first chapter of this thesis, I detail some preexisting methods for gene expression measurement, and discuss their advantages and drawbacks. I then expose the design of our new gene expression reporter, the dPSTR. We developed our system in the budding yeast, which provides the advantage to have a short generation time and to be easily genetically modifiable.

As a proof of concept, we applied the dPSTR technology to the study of the transient response to the hyper-osmotic shock driven by the HOG pathway. This study, published as a research paper in Nature Communications, is presented in the second Chapter (Aymoz 2016). We were able to observe gene expression within minutes following stress, and quantify differences of several minutes in the expression dynamics between promoters. We were also able to quantify gene expression noise in real-time in course of the cell response to stress. Thanks to the design of our system, we were able to observe multiple subsequent transcription events driven by successive stresses.

In the Chapter 3, we applied the dPSTR method to the study of the transcription driven by the mating response. We were able to show that pheromone-driven transcription occurs on a broad timescale depending on the promoter measured. We showed that the timeline of expression is conserved from cell to cell, and occurs both in response to exogenous pheromone stimulation and in mating mixtures.

In the following Chapter 4, we identified some regulation mechanisms controlling the mating-induced gene chronology of expression. Through genetic analysis and construction of systematic variants of two prototypical promoters, we were able to identify some basic rules controlling gene expression dynamics.

Finally, in Chapter 5, we discuss the limitations of the dPSTR system, and propose some perspectives on how the dPSTR could be used in the future.

Chapter 1

Gene expression measurements

Preamble

In this thesis, we report the development of a new class of gene expression reporters. The dynamic Protein Synthesis Translocation reporter, or dPSTR, faithfully reports on the levels and kinetics of gene expression. This Chapter will first describe some available methods to study gene expression. As none of these methods allow to faithfully report on gene expression kinetics at the single cell level, we decided to develop a new kind of gene expression reporter. Hence, the second part of this Chapter explains the principle of the dPSTR system, and possible applications coming out from it.

Abstract

The gene expression resulting from MAPK pathway signaling is dynamic, heterogeneous, and highly regulated. A variety of methods are available to study gene expression at various layers: population or single cell level, end-point or continuous measurements, low or high throughput. In this Chapter, we describe some of the current methods and discuss their limitations for the study of the MAPK-driven transcription. Following this, we present the principle and design of the gene expression reporter we developed in this thesis, the dPSTR. The dynamic Protein Synthesis Translocation Reporter quantifies protein expression driven by a promoter of interest in real-time and at the single cell level. Finally, we describe some of the features that the dPSTR allows to measure, like expression levels and kinetics.

Introduction

Transcription is a tightly regulated process that controls all cellular life. It is important for the metabolism of the cell, for cell survival to stress, or cell differentiation. Often, the gene expression is regulated via extracellular inputs that are sensed and integrated by signaling cascades. Downstream of these pathways, transcription factors are activated, and drive the expression of the specific genes. According to the nature of the stimulus that a cell will sense, the response can be triggered over various timescales. For instance, the transcriptional response to morphogens during embryogenesis can occur over days (Ashe and Briscoe, 2006; Christian, 2012). In other cases, when the life of the cell is threatened, transcription can arise within minutes following stress (Mas et al., 2009; Pelet et al., 2013). Moreover, in response to the same input, single cells from a subclonal population can display very different behaviors (Altschuler and Wu, 2010; Elowitz et al., 2002). Indeed, the ability of cells to generate gene expression depends on amounts of proteins performing transcription, like polymerase and transcription factors. These amounts are variable from cell to cell, and lead to heterogeneous transcriptional outputs in response to a stimulus. Moreover, their expression capacity, which is regulated by the amount of ribosomes performing translation, is adding an additional layer of noise between cells. Additionally, stochastic processes can occur during gene expression resulting in intrinsic noise within cells.

Thus, it is now clear that single cell behavior can provide insights on biological questions that population measurements cannot offer (Altschuler and Wu, 2010). Indeed, population measurements hide the behavior of single cells by averaging out subpopulations with different gene expression behaviors, which may even result in a pattern that does not represent either subpopulation. Moreover, it is important to get information on gene expression in real-time, as transcription is a dynamic process that can occur with different kinetics in cells from a same population. Snapshots of gene expression in cells that belong to a population only provide information at that specific time point. It does not allow to follow the behavior of a cell throughout its entire response. As such, some important behaviors can be missed. For instance, asynchronous oscillating gene expression may be mistaken for bimodality with snapshots population measurements, because only a part of the population is expressing at each time point.

Overall, this highlights the need to study gene expression dynamically and at the single cell level.

In a first part, we will describe the current methods available to study gene expression at either population or single cell level, at the level of the mRNA or the proteins, and low or high-throughput.

Then, we will present the design of the fluorescent reporter we developed to study gene expression in real-time and at the single cell level, and how we can use it.

Part I: The current methods for gene expression measurements

Many techniques have been developed throughout the years to study gene expression (Kalisky et al., 2011; Lenstra et al., 2016; Roth, 2002; Shav-Tal et al., 2004). They can provide different layers of information, at different levels of clarification. Some methods provide information at the population level, other at the single cell level. Some allow a temporal monitoring of the gene expression, whereas other do not. Finally, some methods are low-throughput, and provide information on the expression of one or a few number of genes at the same time, whereas some high throughput methods provide information at the genome-wide level.

Gene expression measurements can be achieved at two levels: mRNA or protein. Genes are transcribed in mRNAs which are then excluded from the nucleus and translated into proteins. There can be some regulatory steps between transcription and translation. For instance, the stability of the mRNA can be increased or decreased according to environmental conditions. Or the mRNA can be present in the cell but its translation can be repressed. This allows the cell to rapidly produce certain proteins in response to specific environmental conditions. For instance, Gcn4, which regulates the induction of genes involved in amino acid biosynthesis in amino acid starved cells, is regulated at the translational level by this means (Hinnebusch and Natarajan, 2002). Global-scale studies reported a poor correlation between mRNA levels and protein abundances in yeast (Greenbaum et al., 2003; Gygi et al., 1999). Indeed, mRNA levels do not inform on translational regulation or protein half-life, which directly impact the amount of proteins present in a cell.

mRNA-based measurements

One way to measure gene expression is to quantify the mRNA level resulting from the transcription of a gene of interest. Several methods allow to do this with various degrees of precision. They start by a first step of total RNA purification, consisting of extracting RNA from cells, and isolating it from proteins and DNA. It is very difficult to perform the extraction step efficiently, mostly because of the fragility of the RNA and the presence of RNAses in the cells. After this extraction, several techniques can be applied for quantification.

Northern Blot

The Northern Blot is a widely used technique, which has been developed in 1977. The mRNA can be loaded on a denaturing electrophoresis gel and separated according to its size, and then blotted on a membrane. The use of a complementary probe, either radiolabelled or chemiluminescently labeled, allows the detection of a specific transcript. Usually, a housekeeping gene is also monitored in order to normalize the levels of the gene of interest. This method has been widely used to study transcriptional outputs resulting from MAPK signaling. For instance, to demonstrate the involvement of chromatin remodeling in the expression of hyper-osmostress-induced genes, two studies used Northern Blots (Klopf et al., 2009; Mas et al., 2009). This method provides a measurement of the population behavior at a specific moment in time.

RT-PCR

Alternatively, a reverse transcription polymerase chain reaction (RT-PCR) can be performed to quantify a specific mRNA. Briefly, the mRNA is reverse transcribed into cDNA, which is then amplified by a PCR. In contrast to Northern blot analysis, this method allows the

detection of mRNA expressed at very low copy number because of the amplification step, which is not doable using Northern Blot analysis.

Microarray transcriptome analysis

The total mRNA extracted from a population of cells can be analyzed at a genomic scale by the use of microarray chips. The chips contain thousands of specific probes fixed on a surface, which are able to hybridize with cDNA derived from all the genes of a genome. mRNA is reverse transcribed into cDNA and hybridized on the chips. A ratio between a controlled condition (unstimulated) and a tested condition (stimulated) will report on the relative variation of a gene expression between the two conditions analyzed. For instance, Gasch *et al.* used this method to assess the yeast transcriptome variations in response to various stresses, for various durations (Gasch et al., 2000). This high throughput method provides a big picture of the remodeling of a transcriptome, since these data are often combined to bioinformatics methods that correlate changes in expression to the function encoded by the genes.

This method is also widely used in biomedical research, for diagnosis purposes (Trevino et al., 2007).

Single-cell RNA-seq

RNA-seq consists in sequencing all RNAs of a sample using next-generation sequencing of cDNAs (Mortazavi et al., 2008). To estimate gene expression, reads are quantified and aligned to coding regions of the entire genome. This method can also be applied to single cells via amplification steps, which allow to get insights on the full transcriptome of a single cell (Tang et al., 2009). This is interesting in the cases of embryonic development or tumorigenesis studies, in order to fully understand cell-to-cell heterogeneity. However, this method provides a snapshot measurement of the transcriptome of the cells, as it requires cell destruction to be performed.

mRNA FISH

Detection of transcription sites or sole mRNA molecules is possible by Fluorescence *in situ* hybridization (FISH). Probes specific to either 5'-UTR or 3'-UTR of a gene are coupled to a dye, and hybridized with mRNA in single fixed cells. Visualization is then achieved by fluorescent microscopy. This technique allows the detection of transcription sites, by visualization of nascent RNA via 5'-UTR probes, or the detection of completed transcripts, via 3'-UTR probes (Femino, 1998) (Shav-Tal et al., 2004). By designing probes for different genes coupled to different dye, it is possible to monitor the expression of multiple genes at the same time (Femino, 1998; Raj et al., 2008). However, this technique requires the fixation of cells, and precludes the follow-up of single-cell behavior.

MS2/PP7 system

Another way to detect transcription site is the use of the MS2/PP7 system. This technique is based on fluorescent *in vivo* hybridization. The mRNA of interest is tagged by stem loop sequences, which are recognized by a phage-coat protein fused to a GFP (Bertrand et al., 1998). As the mRNA is synthetized, stem loops appear and are bound by the labeled protein, resulting in the formation of foci, observable by fluorescent microscopy. Two couples of phage coat proteins and loops have been characterized MS2 and PP7 (Bertrand et al., 1998; Larson et al., 2011). Using these two proteins coupled to different fluorophores, Hocine *et al.* were able to quantify the elongation rate of the RNA Polymerase II at the *MDN1* locus, by tagging both 5'- and 3'-UTR (Hocine et al., 2013). Another report used the same method to quantify antisense transcription at the *GAL10* locus (Lenstra et al., 2015). These methods are very powerful to quantify transcription kinetics in live cells. However, for now, only two probes are

available, limiting the number of genes quantifiable simultaneously. Moreover, the imaging of these foci requires extensive imaging of the cells, with the measurement of several stacks in the Z- direction. This can lead to photobleaching of the fluorescent proteins and loss of the signal, but also phototoxicity for the cells. Finally, it is still a challenge to achieve automatized detection properly.

Protein-based measurements

Roughly half of the budding yeast's genome encodes protein (Mackiewicz et al., 2002). Hence, the major function in gene expression results in production of mRNA that is then translated into proteins. The amount of proteins present in cells can be quantified through different methods.

Western Blot

The levels of a protein can be determined by means of Western Blot. Briefly, proteins are extracted from cells, and separated by SDS-PAGE electrophoresis based on their sizes. After transfer on a membrane, the protein of interest is stained via an antibody directly recognizing it, or via a tag fused to the protein. This method allows population readout of protein amounts, providing proper loading controls. It has the advantage of distinguishing between different forms of the same protein, either modified by phosphorylation, like in the case of activated MAPKs, or different isoforms of the same protein, like in the case of the short and long forms of Kar4 (Brewster et al., 1993; Errede, 1993; Gammie et al., 1999).

Mass spectrometry

Mass spectrometry can also be used in gene expression analysis to quantify a fold change in abundance (Pasa-Tolic et al., 2002). In theory, this method allows an identification of all proteins expressed in an organism and has been developed to propose an equivalent to microarray measurements for proteins. However, a study reported the identification of 1484 yeast proteins in ~90 hours, which is a low output number compared to the time required for this experiment (Peng et al., 2003; Washburn et al., 2001). Another method using internal Accurate Mass Tags (AMTs) allows to have a better view of relative protein abundance changes of in different conditions or in course of time (Smith et al., 2002a; Smith et al., 2002b; Smith et al., 2001). More recently, a study reported the generation of a reference map of the full yeast proteome (Picotti et al., 2013). However, these methods are time-consuming, costly to use, and require a population of cell. Mass spectrometry is mostly used to detect protein modification, like phosphorylation, as illustrated by this study on cross-talk between two MAPK pathways in yeast (Vaga et al., 2014).

Protein reporters

Another way to study gene expression is to fuse the promoter sequence of a gene to a DNA sequence encoding a reporter protein. When the promoter of interest is induced, the reporter protein will be detectable. Protein reporter techniques are usually costless and straightforward to use. They allow the measurement of the induction of a promoter placed elsewhere in the genome, but not directly of the gene of interest. These methods report on gene expression levels without perturbing the cellular physiology.

Beta-galactosidase

One widely used reporter gene is the bacterial *lacZ* gene, encoding the beta-galactosidase enzyme (Rose and Botstein, 1983). It degrades ß-galactosides, and its activity can easily be detected by using a derivative called X-gal. This compound can be degraded by the ß-gal enzyme expressed by cells bearing the fusion. Its degradation will lead to the appearance

of a blue dye that can easily be detected with a spectrophotometer. Although this method is very easy and cheap to use, and allows the quantification of relative expression of different promoters, it fails to provide real-time data or single-cell level data. To analyze the constraints of PREs positions on mating-induced transcription, Su and collaborators used this technique (Su et al., 2010).

Fluorescent proteins

Fluorescent proteins (FP) are another type of expression reporter, and their genes can be fused to a promoter or a coding gene of interest. The green fluorescent protein was first observed in the jellyfish *Aequorea victoria* in 1962, but it took another 30 years to tag the first protein with GFP and express it in cells (Chalfie et al., 1994; Tsien, 2009). Since then, different variants of the GFP have been obtained, with different properties: excitation and emission wavelengths, maturation time or oligomerization ability (Olenych et al., 2007; Remington, 2006; Shaner et al., 2005). In 2008, a Nobel Prize in Chemistry was attributed to the discovery and development of the GFP (Tsien, 2009). The FPs exhibit many advantages over the previously described methods for the study of gene expression. First of all, as for the *lacZ* gene, these constructs are easy to design and clone. Second, they are available in different variants that can have separated emission and excitation spectra, providing a way to measure several promoters at the same time, up to 4-5 according to the settings of the machine used. Third, they can be measured at the single cell level without requiring the destruction of the sample for measurements. Hence, they are suitable for living single cell measurements of gene expression (Pelet et al., 2013). Indeed, time-lapse microscopy experiments provide long-term measurements of single cells, providing that the cell is supplemented by media and at the appropriate temperature. For short-term experiments, from minutes to a couple of hours, a simple glass well can be sufficient. For longer experiments, a microfluidic chip can be used to provide a continuous flux of fresh media. Both can be used in a chamber controlling the temperature for optimal growth of the cells, and both allow to follow single cells for the duration of the experiment. The quantification of the fluorescent measurements can be performed by algorithm in an automatized manner, as will be described in a further section. This technique was used to describe the noise in gene expression in bacteria (Elowitz et al., 2002). Time-lapse measurements of fluorescent protein expression have been used in various organisms to study stress response (Loewer et al., 2010; Pelet et al., 2011).

Fluorophores require a long maturation time before being fluorescent and visible, from 15 to 60 minutes (Nagai et al., 2002; Olenych et al., 2007). This may not be a problem for quantification of gene expression in mammalian cells, which can occur on the timescale of hours. However, this precludes quantifications of fast genes expression events occurring on a short timescale (less than an hour). For instance, Northern Blots experiments showed that hyperosmotic stress in yeast yielded a transient mRNA production within minutes following stress. Such kinetics cannot be observed with the use of a fluorescent protein fusion directly quantified (see Chapter 2). Instead, an alternative method can be used.

The protein synthesis can be blocked by addition of a drug, the cycloheximide, at various times following stimulation. This will freeze the cell in the state it was at this precise time point, allowing the experimenter to wait a couple of hour for the pool of produced FP to be mature prior to perform his measurements. Typically, this is done for measurements with a flow cytometer, which provides very fast quantification of several thousands of cells at once. The cytometer will illuminate each cell separately, and monitor its granularity and size, and all relevant fluorescent channels (Shapiro, 2003). A software then allows the selection of single-cells, as opposition to aggregates or cell fragments, for further analysis. As this method requires the blocking of cells, it is not suitable to follow single cells in course of time. Instead, it provides snapshots of a population at various times after stimulation (Pelet et al., 2013). These kinds of

measurements allowed the identification of a bimodal response in gene expression following hyperosmotic stress in budding yeast (Pelet et al., 2011), feature that was hidden when the response was studied at the population level. It shows that only part of the population is inducing gene expression at low stress strength, illustrating another level of cell-to-cell heterogeneity.

Luciferase assay

Luciferases are enzymes able to produce light upon catalyzing the production of a molecule in a high quantum state, which emits a photon to go back to its ground state (Wilson and Hastings, 1998). They do not require maturation, do not bleach and are functionally active just after translation (Frydman et al., 1999). The reaction catalyzed by the luciferase produces less photons than the GFP, which makes it harder to detect (Mazo-Vargas et al., 2014). Mazo-Vargaz *et al.* developed a method measuring gene expression in budding yeast using time-lapse microscopy, but in order to achieve a sub-minute temporal resolution, they had to record 5 Z-stacks with a 10 seconds exposure for each plan, which means 50 seconds for only one field of view, which greatly reduces the number of cells quantified at once (Mazo-Vargas et al., 2014). With this technique, they monitored cell cycle induced transcription for more than 500 minutes. However, the plasmid used in this study was inserted in multiple copies in the genome to promote a quantifiable output. This leads to erroneous quantifications, as the number of inserted copies may differ from cell to cell, amplifying the expression noise.

We presented an overview of some useful techniques for gene expression. Table 1 summarizes their main qualities and drawbacks. However, none of them allows to monitor gene expression in real-time, in living cells, at the single-cell level and with no effect on cell physiology. As such, we decided to develop a new gene expression technique based on fluorescent microscopy that would fulfill these requirements.

Table 1.1 Overview of gene expression measurement techniques

Technique	Single cell or population	Number of genes	Measurement of	Advantages	Downsides
Northern Blot	Population	1	mRNA	No tagging required	Snapshot measurement
RT-PCR	Population	1	mRNA	Quantitative	Sensitive Snapshot measurement
Microarray analysis	Population	Whole-genome	mRNA	Genome-wide	Snapshot measurement Quantification of fold induction
Single-cell RNA-seq	Single cell	Whole genome	mRNA	High number of targets	Snapshot measurement
mRNA FISH	Single-cell	5	mRNA	Single-molecule visualization	Requires cell fixation
MS2/PP7	Single-cell	Up to 2	mRNA	Real-time	Challenging automatized quantification Phototoxicity
Western Blot	Population	1	Protein	Detection of protein isoforms	Snapshot measurement Requires tagging or existence of a specific antibody
Mass spectrometry	Population	>1500 proteins	Protein	High number of targets	Cost Time-consuming
Beta-galactosidase	Population	1	Promoter	Costless	Snapshot measurement
Fluorescent proteins by microscopy	Single-cell	3-4	Promoter	Single-cell visualization	Long maturation time of FP
Fluorescent proteins by FACS	Single-cell	3-4	Promoter	High number of cells	Snapshots measurement
Luciferase	Single-cell	1	Promoter	Single-cell visualization Real-time	Low signal slows down acquisition

Part II: A new gene expression reporter

The dPSTR design

The dynamic Protein Synthesis Translocation Reporter (dPSTR) converts the expression arising from a promoter into the signal of relocation of a fluorescent protein (FP) into the nucleus of yeast cells (Figure 1.1) (Aymoz et al., 2016). It is composed of two protein parts, encoded by two transcriptional units, both placed on the same plasmid, integrated in a single copy in the genome (Wosika et al., 2016). The FP is under the control of a constitutive promoter, and a pool of mature and fluorescent proteins is present in the cell at all times. The FP can diffuse freely in and out of the nucleus, because of its size, leading to a homogenous repartition throughout the cell (Wang and Brattain, 2007). The second transcriptional unit is driven by the promoter of interest, which can be any promoter: stress- or signaling-induced, constitutively expressed or with oscillating patterns. The encoded peptide will drive the accumulation of the FP in the nucleus. To achieve this, it contains two Nuclear Localization Signals (NLS), which are driving the active import of proteins into the nucleus (Kaffman and O'Shea, 1999). The interaction between this induced peptide and the FP is caused by the presence of a synthetic domain (SynZip or SZ) in each protein (Reinke et al., 2010; Thompson et al., 2012). The SynZips are small coiled coils peptides of 50 AA, which are able to interact as zippers in a specific and strong way. As soon as the induced moiety is translated, it attaches the NLSs to the FP, and promotes its accumulation in the nucleus by increasing its nuclear import rate. The measurement of the nuclear accumulation of the FP, that is the difference between nuclear and cytoplasmic fluorescence, provides a readout of the expression of the promoter of interest.

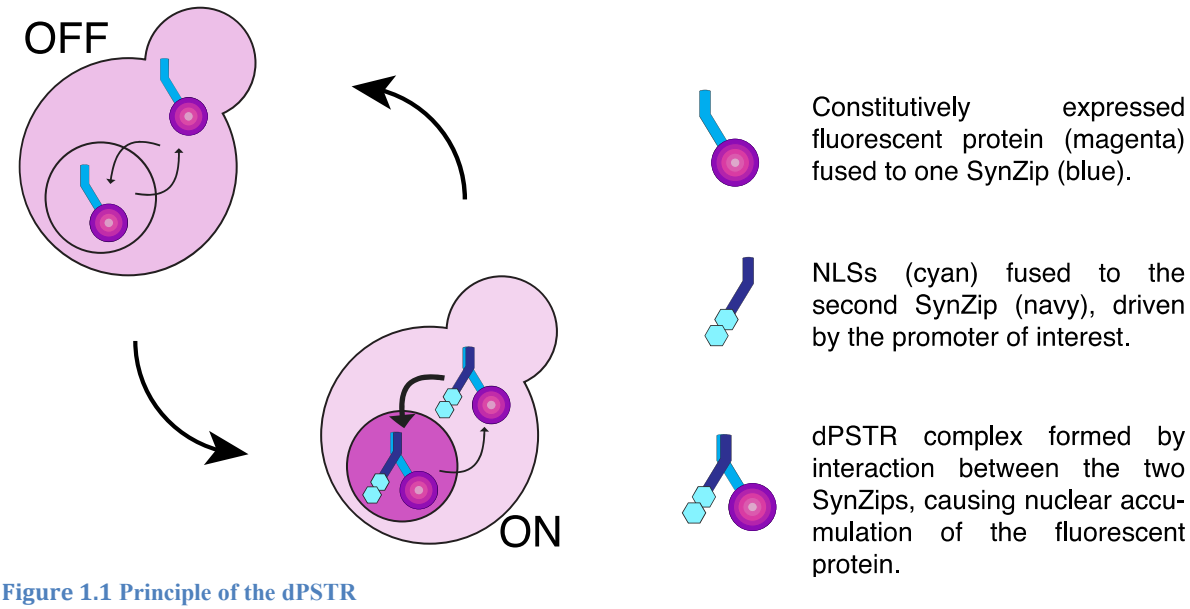


Figure 1.1 Principle of the dPSTR

Schematic representation of the dynamic Protein Synthesis Translocation Reporter (dPSTR) (Aymoz et al., 2016). The dPSTR converts protein expressoin arising from a promoter of interest into a relocation of a constitutive fluorescent signal from the cytoplasm into the nucleus of cells. It is based on two transcriptional units carried on a sole uniquely integrated vector. The fluorescent protein in constitutively expressed and present all over the cell in the dPSTR OFF state. A second peptide driven by the promoter of interest carries two Nuclear Localization Signals (NLS), allowing the nuclear recruitment of any protein. The physical interaction between the fluorescent protein and the induced peptide is taking place through the interaction of small synthetic peptides called SynZips, providing strong and specific interaction. The whole complex is then recruited in the nucleus, changing the fluorescent signal arising from the cytoplasm and the nucleus.

In order to quantify transient gene expression, we designed an unstable version of the dPSTR (dPSTR*), by addition of the destabilization sequence in front of the inducible peptide (UbiY). The ubiquitin amino-terminal part of the peptide is cleaved, and reveals an amino acid that determines the stability of the peptide (Varshavsky, 1996). As a result, the half-life of the protein is decreased to several minutes. With a tyrosine as N-terminal amino acid, we expect the half-life of the peptide to be within the order of 10 minutes (Varshavsky, 1996). However, we quantified a half-life of 2 minutes for our UbiY-2xNLSs-SZ1 peptide, by blocking protein translation with cycloheximide (Aymoz et al., 2016). If the revealed amino acid is a methionine, the peptide is expected to be stable, but in the case, we still observe a degradation of this peptide, suggesting an inherent short half-life for the construct (data not shown). With this unstable construct, there is a balance between production and degradation rate of the induced moiety. As long as the promoter is induced, the UbiY-2xNLSs-SZ1 protein is produced and drives the accumulation of the FP in the nucleus. When promoter expression stops, the degradation rate becomes higher than the production rate, and the FP looses the ability to be actively imported in the nucleus, decreasing its nuclear concentration.

We wanted to be able to measure at least two promoters at the same time using the dPSTR. To do so, we designed a second dPSTR that uses a different SZ pair and a different FP. The $dPSTR^{R}$ possesses a red FP and contains SZ1 and SZ2, whereas the $dPSTR^{Y}$ has a yellow FP variant and uses SZ3 and SZ4 (Aymoz et al., 2016). These two pairs of SZ were found to not cross-interact within pairs (Reinke et al., 2010; Thompson et al., 2012). Hypothetically, it is possible to construct a third compatible dPSTR, as Reinke *et al.* described a network of three orthogonal pairs, but due to compatibility limitations at the FP level, we did not build this construct (Reinke et al., 2010).

Because of their design, it is really easy to modify the dPSTR plasmids in order to measure any promoter of interest. Moreover, any FP variant can be used by replacing the current ones (Shaner et al., 2005).

Quantification during time-lapse experiments

The dPSTR system is used in combination with time-lapse microscopy. We use 96-glass-well plates in order to measure different conditions in one experiment. The microscope is automatized, can visit a number of positions at the chosen frequency, and take images with the chosen illumination settings (wavelength and exposure time) (Edelstein et al., 2010; Pelet et al., 2013). Of course, the more positions are visited, the bigger will be the number of cells monitored. However, there are technical constraints regarding the number of positions that can be visited and the frequency of imaging, as the microscope stage has to move to each position and then take the appropriate images. This process requires a certain time that cannot be bypassed for now. In a typical dPSTR experiment, we image up to 8 wells, 5 positions per well, with the two dPSTRs (300ms exposure each) every 2min30. In Chapter 2, we were able to increase the frequency of imaging up to 30 seconds for 2 wells with 4 positions each.

Automatized quantifications

We developed an automated platform allowing to efficiently segment and track the cells throughout an experiment, YeastQuant (Pelet et al., 2012) (See Supplementary Figure 2 from Chapter 2). Briefly, two bright-field images are used to determine the cell compartment. The nuclear compartment is detected thanks to the presence of a histone tag, and is subtracted from the cell element to obtain the cytoplasm compartment. The YeastQuant output is a big matrix of average fluorescence intensities in any color channel and any compartment chosen before quantification. We then use Matlab (The Mathworks) software to design appropriate script to plot the data. We only select cells tracked throughout an experiment, and apply further filters

on the cell and nuclear size to ensure a good data quality. In a typical dPSTR experiment, we can quantify from 100 to 600 of single cells tracked in course of an experiment lasting 1h30.

Overall, imaging and quantification are both very easy to achieve and automatize.

Effect on the cell physiology

The dPSTR plasmids are integrated in two loci mutated in the lab strains we are using specifically to allow plasmid selection (*URA3* locus for dPSTR^{R} and *LEU2* for dPSTR^{Y}). Hence, the use of dPSTR does not require in any case the disruption of the gene of interest. Indeed, some reporter techniques require either tagging the protein of interest with a fluorescent protein or sometimes complete replacing of the coding sequence by the reporter gene. Both methods can result in partial or total loss of function of the protein of interest, which can sometimes be problematic. For instance, in the context of our study of the HOG-induced transcription, the cell survival would have been impaired if we had to delete the *GPD1* coding gene, which may distort the measurements (Albertyn et al., 1994; Aymoz et al., 2016).

Moreover, as the SynZips are synthetic, we do not expect them to interact with any other protein in the cell. Hence, the presence of the dPSTR should not perturb the physiology of the cell. We did not notice any impact on the growth rate either (data not shown).

Measurement of expression kinetics and levels

The dPSTR allows to quantify gene expression kinetics and levels in single cells. As we show throughout this thesis, we can quantify the expression output for a gene in response to a stimulus. This corresponds to the amount of protein produced during the transcription event. It is quantified as the difference between the basal nuclear enrichment of the dPSTR, before stimulation, and the maximal nuclear enrichment, measured after stimulation. We can use this value as a readout of the expression level of a gene. This value is comparable between promoters measured with the same colored dPSTR only, as its value is defined by the fluorescence intensities measured in the nucleus and in the cytoplasm. Due to nuclear fluorescence fluctuations because of nucleus movement, this value may be negative if a cell does not express the promoter of interest. We can then use this threshold to determine whether a cell is expressing or not, based on this expression output. We decided that if the cell expression output was equal or higher than 20% of the population average expression output, it would be considered as expressing.

Among expressing cells, we can quantify the expression kinetics by measuring what we termed Response Time. It represents the time that each cell needs to overcome 20% of its expression output.

Basal nuclear enrichment of the dPSTR

Some promoters that we measure with the dPSTR have basal levels of induction, independent of any stimulation. With the dPSTR, we can quantify the amount of cells having a basal expression. However, we need to define a threshold of expression. As the FP can diffuse freely through the nuclear membrane, it does not require the presence of the NLS moiety to enter in the nucleus. The presence of the NLS will only increase the nuclear import rate of the FP, resulting into a higher import rate than export rate, causing the accumulation of the FP in the nucleus. We quantified the nuclear fluorescence in the RFP channel for a strain that does not carry any mCherry protein, compared to a strain carrying only the constitutive moiety of the dPSTR^{R} (Figure 1.2). We can see that the strain possessing the mCherry part of the dPSTR has a higher nuclear fluorescence intensity than the background fluorescence, although there is a large discrepancy between cells (Figure 1.2A). This strain without NLS-SZ moiety was used as a threshold to determine expressing cells in basal conditions. Indeed, Figure 5B shows two mating-induced promoters that are also induced during vegetative growth. Here, we can clearly

see that they are expressed in different proportions of cells, and at different levels. The p*FAR1*-dPSTR^R is strongly accumulated in all cells, whereas p*AGA1*-dPSTR^R accumulates in the nucleus of only 12% of the cells and at much lower levels (Figure 1.2B).

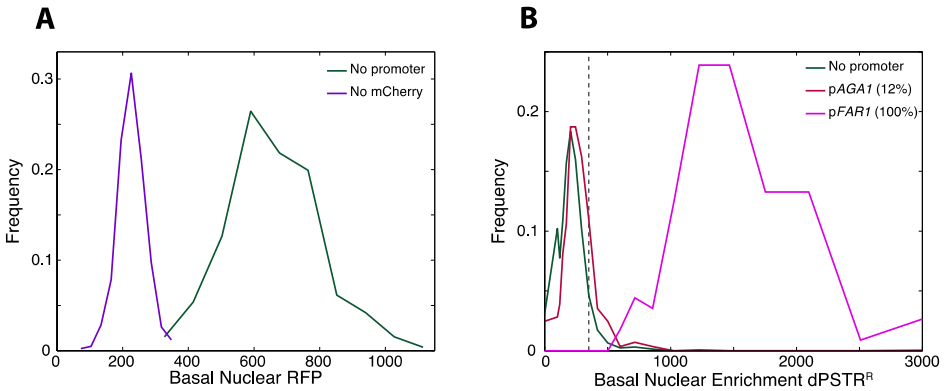


Figure 1.2 Nuclear enrichment of the dPSTR and FP alone

A. Distribution of the basal RFP nuclear fluorescence in a strain with no RFP (No mCherry, purple) or in a strain with only the constitutive part of the dPSTR^R (mCherry-SZ2, "No promoter", green).

B. Distribution of the basal nuclear enrichment (nuclear minus cytoplasmic fluorescence) for the strain with only the constitutive part of the dPSTR^R (green), or for p*AGA1*-dPSTR^R (red) or p*FAR1*-dPSTR^R (pink).

Combination of dPSTR and other sensors

The dPSTR can be used in combination with other fluorescent sensors, like signaling sensors, allowing the correlation of signaling activity with gene expression. In Chapter 3, we combine the dPSTR with the kinase activity sensor SKARS, and are able to measure signaling activity and gene expression at the same time. The two versions of the dPSTR can also be mixed, allowing combinatorial measurement of the expression of two promoters. In this case, we can quantify kinetic differences in gene expression within a single cell.

If the two dPSTRs are used to measure the same promoter, it allows the quantification of expression noise in real-time, as we do in Chapters 2, 3 and 4 (Elowitz et al., 2002).

Moreover, the use of the unstable version of the dPSTR (dPSTR*) allows to quantify transient gene expression events, as we do for gene expression resulting from hyper-osmotic stress in Chapter 2. We show that this brings the opportunity to study memory effect in transcription, which cannot be achieved with FP reporters that have a long half-life (Natarajan et al., 1998). Hence, the dPSTR can be used to monitor oscillating gene expression, as in the case of promoters regulated by the cell cycle. Potentially, the dPSTR* can also be used to quantify gene repression. For instance, a large number of genes are repressed in response to mating signaling (Roberts et al., 2000). The dPSTR* could help to analyze study the kinetics and duration of gene repression.

To summarize, the dPSTR allows real-time quantification of gene expression at the single cell level. It bypasses the maturation time of FP, which previously precludes measurements of fast kinetics of induction. The dPSTR can easily be implemented and combined with other fluorescent sensors. Measurements by time-lapse microscopy in combination with computational analysis allow automatized quantification of hundreds of live cells over a few hours.

Supplementary Table

Supplementary Table 1.1: List of yeast strains used Chapter 1

Strain	Genotype	Plasmid
W303	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*	
ySP580	*HTA2-CFP*	
yDA278	*HTA2-CFP*	
	URA3: pRPL24A mCherry-SZ1	pDA157
	LEU2: pAGA1-dPSTR$^{Y4\text{-}3}$	pSP368
yDA351	*HTA2-CFP*	
	LEU2: pGPD1-dPSTR$^{Y4\text{-}3}$	pDA200
ySP642	*HTA2-CFP*	
	URA3: pAGA1-dPSTR$^{R2\text{-}1}$	pSP365
	LEU2: pAGA1-dPSTR$^{Y4\text{-}3}$	pSP368
yDA146	*HTA2-CFP*	
	URA3: pFAR1-dPSTR$^{R2\text{-}1}$	pDA229
	LEU2: pAGA1-dPSTR$^{Y4\text{-}3}$	pSP368

Chapter 2

dPSTR development

Preamble

The content of this Chapter have been published as a research article in the review *Nature Communications* in April 2016 :

Aymoz, D., Wosika, V., Durandau, E. & Pelet, S. Real-time quantification of protein expression at the single-cell level via dynamic protein synthesis translocation reporters. *Nature communications* **7**, 11304, doi:10.1038/ncomms11304 (2016).

ARTICLE

Received 13 Aug 2015 | Accepted 13 Mar 2016 | Published 21 Apr 2016

DOI: 10.1038/ncomms11304 OPEN

Real-time quantification of protein expression at the single-cell level via dynamic protein synthesis translocation reporters

Delphine Aymoz[1], Victoria Wosika[1], Eric Durandau[1] & Serge Pelet[1]

Contributions :
Serge Pelet, Delphine Aymoz, Victoria Wosika and Eric Durandau designed the experiments. Delphine Aymoz performed the dPSTR experiments. Serge Pelet performed the flow cytometry experiments. V.W. and S.P. performed the PP7 experiments. Serge Pelet and Delphine Aymoz analysed the data and wrote the manuscript.

Abstract

Real-time quantification of protein expression at the single-cell level *via* dynamic protein synthesis translocation reporters

Protein expression is a dynamic process, which can be rapidly induced by extracellular signals. It is widely appreciated that single cells can display large variations in the level of gene induction. However, the variability in the dynamics of this process in individual cells is difficult to quantify using standard fluorescent protein (FP) expression assays, due to the slow maturation of their fluorophore. Here we have developed expression reporters that accurately measure both the levels and dynamics of protein synthesis in live single cells with a temporal resolution under a minute. Our system relies on the quantification of the translocation of a constitutively expressed FP into the nucleus. As a proof of concept, we used these reporters to measure the transient protein synthesis arising from two promoters responding to the yeast hyper osmolarity glycerol mitogen-activated protein kinase pathway (p*STL1* and p*GPD1*). They display distinct expression dynamics giving rise to strikingly different instantaneous expression noise.

Introduction

Protein synthesis is a multi-step process. It is typically initiated by the activation of a transcription factor (TF), which binds the promoter sequence of a gene. This active TF allows the recruitment of the polymerase resulting in the formation of an initiation complex. In parallel, chromatin remodelling enzymes act on the locus to enable the efficient transcription of the gene. The polymerase travels along the locus to produce the mRNA. After transcription, the mRNA is exported out of the nucleus to be translated into the amino acid chain that will form the protein. Many complexes and enzymes implicated in this process have been characterized, allowing a detailed mechanistic understanding of the entire protein expression machinery (de Nadal et al., 2011; Weake and Workman, 2010). Comparatively, little is known about how a given DNA sequence influences the final amount of protein produced, the dynamics at which it is expressed or the cell-to-cell variability in the level of protein synthesized. Since the promoter sequence of a gene controls the first steps in protein synthesis, it plays a key role in controlling the final protein levels (Dadiani et al., 2013; de Nadal et al., 2011; Kadonaga, 2004; Knight et al., 2014; Raveh-Sadka et al., 2012; Weake and Workman, 2010).

To quantify mRNA and protein levels, numerous techniques have been developed. However, most of these measurements are performed at the population level (northern blots, western blots) and/or provide snapshots of the cell content at a given point in time (flow cytometry). Only live-cell microscopy combined with fluorescent protein (FP)-based technologies provides a tool to quantify, at the single-cell level, the temporal evolution of protein expression in a given cell.

Using this technique, Elowitz *et al.* measured for the first time the noise associated with protein production using a set of two FPs controlled by identical promoters (Elowitz et al., 2002). To dissect the origin of the fluctuations in protein levels in the same cell, they defined the intrinsic and extrinsic expression noises. This allowed them to observe that the variability between individual cells (extrinsic) as well as stochastic intracellular processes (intrinsic) contribute to the total expression noise.

Many studies have since used FP variants to quantify expression levels in individual cells, either by fusing the FP to a target protein or by placing the FP under the control of a promoter of interest (Colman-Lerner et al., 2005; Lam et al., 2008; Pelet et al., 2011; Raser and O'Shea, 2004). Unfortunately, the maturation time of FP, which varies from tens of minutes to more than an hour, sets a low limit to the dynamics that can be observed (Miyawaki et al., 2003; Shaner et al., 2005). In many signalling pathways, the appropriate timing of gene expression is tightly controlled since it can influence the output of the system (Lahav et al., 2007; Purvis et al., 2012; Skotheim et al., 2008). Moreover, the stable fold of the FP results in a very long half-life in the cell, which hinders the monitoring of oscillatory or transient protein synthesis.

To circumvent these problems, we set out to design a new assay to monitor protein expression induced by a promoter element based on the relocation of a constitutively expressed FP, thereby avoiding the maturation delay. We report here on the development and validation of the dynamic protein synthesis translocation reporter (dPSTR) that provides real-time measurement of protein expression arising from a promoter element in live single cells. As a proof of concept, we adapted the dPSTR system for the study of the protein production driven by the hyper osmolarity glycerol (HOG) pathway in the model organism *Saccharomyces cerevisiae*. We show that by measuring the relocation of an FP into the nucleus, we are able to accurately and dynamically quantify protein expression in hundreds of single cells, with a time resolution under a minute. We also demonstrate that two dPSTRs can be combined in the same strain allowing the first real-time measurements of intrinsic and extrinsic expression noises.

Finally, we further prove the dynamic nature of this assay by stimulating cells multiple times to quantify successive rounds of protein expression.

Results

Design of the dynamics expression reporter

As the maturation time of an FP hinders the quantification of the dynamics of protein expression, we designed an expression reporter that bypasses this rate-limiting step. Our assay uses the change in subcellular localization of a mature FP as a read-out of protein synthesis driven by a promoter of interest. The dPSTR is a protein heterodimer which is encoded by two transcriptional units present on a unique plasmid integrated in the yeast genomic DNA. The first unit constitutively expresses the FP that can freely diffuse between the nucleus and the cytoplasm. The second unit is under the control of the promoter of interest and encodes two nuclear localization signals (NLS), which promote the active import of proteins into the nucleus (Kaffman and O'Shea, 1999). The interaction between the FP and the NLSs is driven by the presence of synthetic bZip domains (SynZips or SZ) in each unit (Reinke et al., 2010; Thompson et al., 2012). These SynZips form strong and specific heterodimers that induce a relocation of the FP into the nucleus, proportionally to the expression level of the NLSs (Figure 2.1).

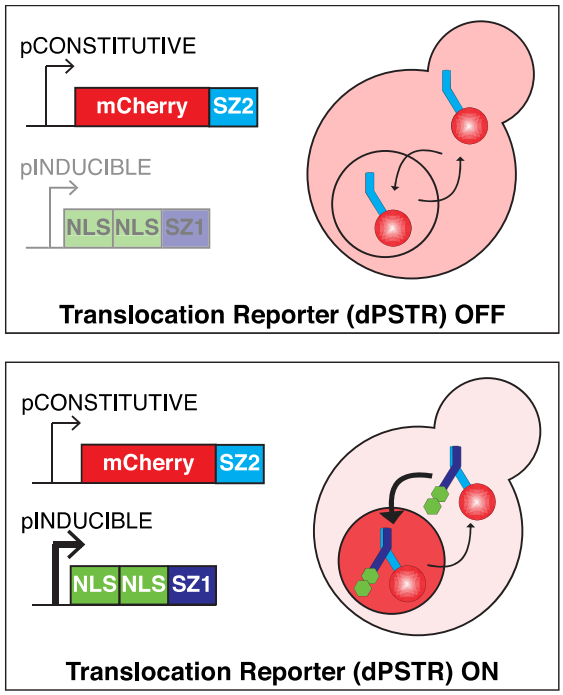


Figure 2.1 Principle of the dPSTR

A FP is fused to a SynZip (SZ2), expressed under the control of a constitutive promoter and can freely diffuse between the cytoplasm and the nucleus (dPSTR OFF, top). The induction of the promoter of interest drives the expression of the second peptide of the reporter, composed of two NLSs fused to a compatible SynZip (SZ1). The strong interaction between the SynZip peptides leads to the enrichment of the FP in the nucleus (dPSTR ON, bottom).

Validation of the method

As a proof of concept, we integrated a dPSTR measuring the expression of the osmostress responsive promoter p*STL1* by the relocation of the red fluorescent protein (RFP) variant mCherry (p*STL1*-dPSTR^{R}, R denotes the fluorescent channel used: RFP), in a strain bearing a histone tagged with cyan fluorescent protein (CFP). Hyper-osmotic shock triggers the transient activation of the HOG pathway, which culminates in the activation of the mitogen-activated protein kinase (MAPK) Hog1. When activated, Hog1 increases the intracellular glycerol production, driving the adaptation of the cells to the high-osmolarity medium. In addition, Hog1 induces the transcription of 300-700 stress response genes (Capaldi et al., 2008; Gasch et al., 2000; Nadal-Ribelles et al., 2012). Among them, *STL1* has been used as a model for stress activated gene expression and is widely studied (De Nadal et al., 2004; Mas et al., 2009; Neuert et al., 2013; Pelet et al., 2011).

To demonstrate the validity of our approach and compare it to traditional expression reporters, we fused a fast maturing Venus FP to the inducible construct controlled by p*STL1* (2xNLS-Venus-SZ) (Nagai et al., 2002). Before induction, the RFP signal is homogeneously distributed between the nucleus and the cytoplasm and no Venus fluorescence can be detected (Figure 2.2). Upon addition of NaCl, the HOG pathway is activated and the inducible part of the dPSTR is expressed. This results in a detectable enrichment of the mCherry in the cell nucleus 10 min after induction, while at the same time point, no fluorescent signal from the

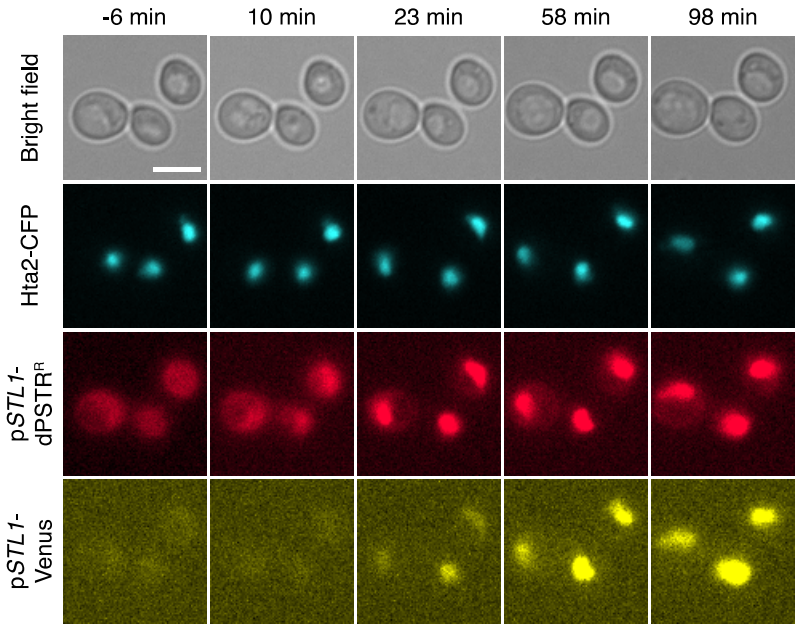


Figure 2.2 Microscopy images of p*STL1*-dPSTR^R compared to classic promoter-FP fusion

Microscopy images of cells with histone Hta2 tagged with CFP and carrying the p*STL1*-dPSTR^R submitted to a 0.2M NaCl stress. The inducible peptide is fused to a Venus FP. Scale bar, 5 µm

Venus can be detected. Note that this nuclear enrichment is dependent on the formation of the SZ heterodimer and on the nature of the promoter element (Supplementary Figure 1A,B).

To quantify the dynamics of protein production, time-lapse movies for three different concentrations of salt (0.1, 0.2 and 0.4 M NaCl) and a control were measured in parallel. Using an automated image analysis pipeline, the nucleus and cytoplasm of the cells were segmented and tracked during the entire experiment (Pelet et al., 2012). At each time point, their average fluorescence intensity was measured (Supplementary Figure 2). Figure 2.3A represents the temporal evolution of the difference between nuclear and cytoplasmic fluorescence in RFP, which is a measure of the level of dPSTR nuclear enrichment. Note that the small increase in nuclear enrichment happening at time zero is an artefact of the shrinking of the cells upon hyper-osmotic stress addition and is not a transcriptional response of the cell (Supplementary Figure 1A,B). The dynamics in dPSTR nuclear enrichment can be compared with those obtained for the expression of the Venus FP (Figure 2.3B). Although the graded protein production due to increasing salt concentration is observed with the two methods, there is a clear kinetic difference between the dPSTR behaviour and the Venus fluorescence signal. The latter appears with a delay, rises more slowly, and reaches its maximum later. We attribute this difference to

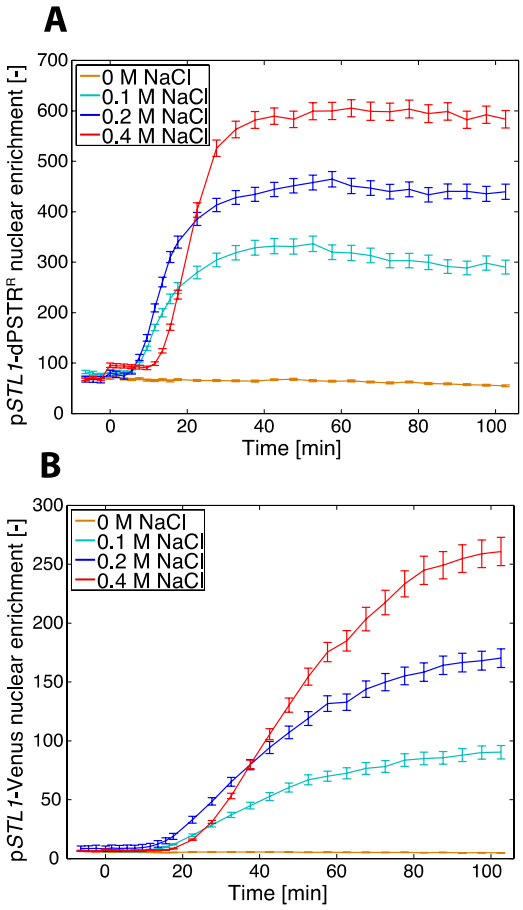


Figure 2.3 Quantification of p*STL1* expression measured with the dPSTR or the promoter-FP fusion

Quantifications of the nuclear enrichment in the dPSTR^R (A) and the Venus (B) channels for cells stressed with 0 (orange, N_C.285), 0.1 (cyan, N_C.266), 0.2 (blue, N_C.294) or 0.4M NaCl (red, N_C.265). Nuclear enrichment is measured as the difference between nuclear and cytoplasmic fluorescence for each single cell. For all similar graphs throughout the paper, the solid lines represent the population average and the error bars are the s.e.m. N_C represents the number of single cells measured.

the maturation step required to form the FP fluorophore (Nagai et al., 2002; Pelet et al., 2012; Remington, 2006).

To quantify this difference more precisely, we measured the time when half of the maximal nuclear enrichment of each FP was obtained for each single-cell trace (Figure 2.4, see Methods). At 0.2 M NaCl, with the dPSTR sensor, the majority of cells need between 10 and 20 min to reach this value, while the halfmaximum of the Venus fluorescence signal is reached later and with a larger spread (between 30 and 60 min). The delay in protein production at 0.4 M NaCl relative to the other concentrations has been attributed to a strong compression of the cell (Miermont et al., 2013). This temporal difference is clearly identified with the dPSTR, while with Venus expression, the distributions at 0.2 and 0.4 M are overlapping, indicating similar kinetics of expression for many cells.

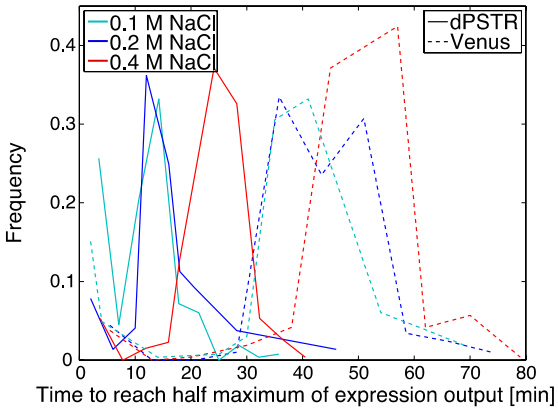


Figure 2.4 Time to reach half of each FP maximal nuclear enrichment

Histograms of the time needed for each single cell to reach half of its expression output for either the dPSTR^R (solid lines) or the Venus (dashed lines). The expression output represents the maximal amplitude of the nuclear enrichment (see Methods).

To verify whether the dynamics measured with the dPSTR reflected the real kinetics of protein production, we used flow cytometry, a method that provides snapshot measurements of the dynamics of protein expression (Pelet et al., 2013). The cells bearing the expression reporter p*STL1*-qVenus were treated with NaCl. At specific time points, translation was blocked by addition of cycloheximide. All the qVenus produced at that point was allowed to mature for 2 h before the measurement was performed. The evolution of protein production quantified by this method aligns well with the live-cell measurements performed with the dPSTR (Supplementary Figure 1C), showing that the dPSTR quantifies the expression dynamics precisely. Indeed, both dPSTR and flow cytometry measurements indicate that proteins start to be synthesized 10 min after the hyper-osmotic stress.

While the dPSTR provides a faster and more accurate determination of the expression kinetics, we wanted to verify whether the level of protein expression measured with the dPSTR was comparable to the one measured with the classical promoter-FP fusion. By setting a threshold based on the non-induced control, we verified that the percentage of expressing cells in the population based on the p*STL1*-dPSTR^R or p*STL1*-Venus signals provided similar proportions (Supplementary Figure 3A,B). Figure 2.5 also demonstrates that there is a high correlation (R^2=0.74 at 0.2 M) between the amounts of Venus measured and the nuclear enrichment of the dPSTR. Note that this correlation is higher at 0.1 M and slightly drops at 0.4 M (Supplementary Figure 3C to E).

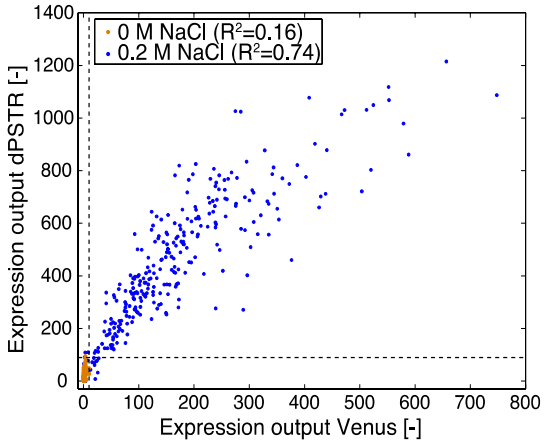


Figure 2.5 Correlation of the expression output measured by dPSTR of protein-FP fusion

Single cell correlation of the expression output measured by either the p*STL1*-dPSTR^R or the p*STL1*-Venus assay, for control cells (orange) or cells induced with 0.2M NaCl (blue). The dashed lines represent the expression thresholds, above which cells are considered as expressing.

This difference could be explained by a saturation effect of the nuclear enrichment of the sensor (Supplementary Note 1 and Supplementary Figure 4). Taken together, these results demonstrate that the relocation of the FP in the dPSTR assay provides a real-time measurement of protein synthesis in live single cells, allowing accurate quantification of both levels and kinetics of protein expression arising from a promoter of interest.

Transient expression

Environmental stresses cause a profound but transient modification in the yeast transcriptional program (Gasch et al., 2000). Northern blot measurements confirmed that the *STL1* mRNAs are produced within 4 min after hyper-osmotic shock but remain in the cell for less than an hour, in agreement with the transient activation of the HOG pathway (Mas et al., 2009; Nadal-Ribelles et al., 2012). To obtain a more precise estimate of the dynamics of mRNA production arising from p*STL1*, we implemented the PP7 system, which allows the identification of the transcription site as a bright fluorescent focus in the nucleus (Larson et al., 2011) (Figure 2.6A). Twenty-four mRNA stem loops, placed under the control of the *STL1* promoter, are recognized by the bacteriophage coat protein PP7 tagged with a double-GFP. Upon induction of the cells with NaCl, bright fluorescent dots appears in a majority of the nuclei. Quantification of the intensity of these nuclear foci provides a dynamic read-out of the mRNA production (Figure 2.6B). Upon 0.2 M NaCl stress, foci can be observed in few cells already 3 min after stress. The intensity and the number of the nuclear foci tend to decrease 10 min after induction. The delay between the stimulus and mRNA production corresponds to the time required for signal transduction, association of Hog1 with TFs, induction of chromatin remodelling and recruitment of the polymerase (Mas et al., 2009; Zapater et al., 2007).

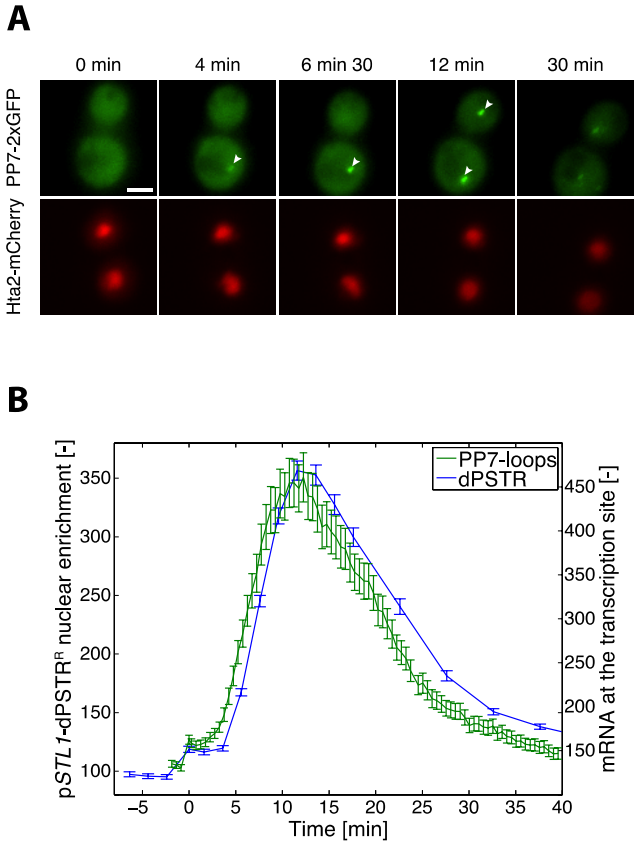


Figure 2.6 Real-time mRNA measurements of p*STL1* transient expression

A. Maximum intensity projections from Z-stacks of cells bearing an Hta2-mCherry tag and transformed with the PP7-2xGFP system with 24 mRNA PP7-stem loops under the control of the *STL1* promoter stimulated with 0.2M NaCl. The presence of transcription site in the nucleus is highlighted by white arrowheads. Scale bar represents 5 µm.

B. Comparison between the mRNA apparition at the transcription site using the PP7-2xGFP (green, N_C.285) and the unstable p*STL1*-dPSTR^R nuclear enrichment (blue, N_C.655) in two different strains, under stimulation by 0.2M NaCl. The fluorescence of the transcription site was quantified by measuring the difference between the 20 brightest pixels in the nucleus and the average nuclear fluorescence.

The PP7 measurements confirm the transient nature of the transcription induced by the activation of the HOG pathway. To obtain a read-out of protein synthesis induction and arrest, we modified the inducible unit by adding a destabilization sequence (UbiY-2xNLS-SZ, Figure

2.7). Upon translation of the peptide, the leading ubiquitin is cleaved off and the exposed amino acid (Y) decreases the half-life of the protein to a few minutes (Varshavsky, 1996). With this construct, protein production is counterbalanced by protein degradation. As long as the rate of protein production is larger than protein degradation, the dPSTR accumulates in the nucleus (Figure 2.6B). Shortly after the mRNA production reaches its maximum, the dPSTR fluorescence which has accumulated in the nucleus starts to return slowly to a uniform localization as can be seen by the decline in nuclear enrichment. Moreover, using the unstable version of the dPSTR prevents the accumulation of the inducible peptide in the cell thereby avoiding any saturation effect (Supplementary Figure 2.5A to C).

Using cycloheximide inhibition, we quantified a half-life for the unstable peptide of 2 min (Supplementary Figure 5D). Therefore, the observed decline in nuclear enrichment, with a half-life close to 10 min, is not only solely limited by the dPSTR degradation rate but also reflects the implication of other biological factors, such as the arrest of transcription and the stability of the mRNA. The comparison between the PP7 and dPSTR measurements shows a short expected temporal delay between mRNA transcription and protein synthesis comprising processes such as mRNA export and translation (Shahrezaei and Swain, 2008). This close consecutive apparition of PP7 and dPSTR signals further confirms that the dynamics of protein production measured with the dPSTR correspond to the genuine kinetics of protein expression.

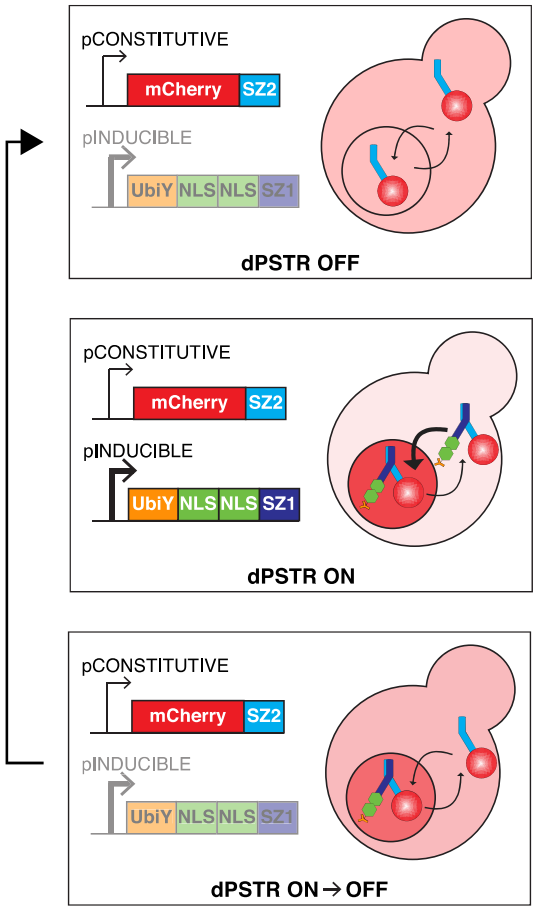


Figure 2.7 Principle of the unstable dPSTR

The dPSTR was modified to measure transient gene expression by addition of an UbiY destabilization sequence at the N-terminus of the induced peptide (2xNLS-SZ). The degradation of the induced construct allows the FP to recover its initial homogenous distribution throughout the cell after stimulation.

Correlation of signalling activity and protein expression

In order to correlate signalling activity and protein expression dynamics, the degradable reporter construct (comprising the UbiY destabilization sequence) was transformed in a strain bearing the MAPK Hog1 tagged with yellow fluorescent protein (YFP) (Figure 2.8). Hog1 nuclear accumulation upon hyper-osmotic stress is linked to its activity and has been extensively used to quantify the dynamics of signal

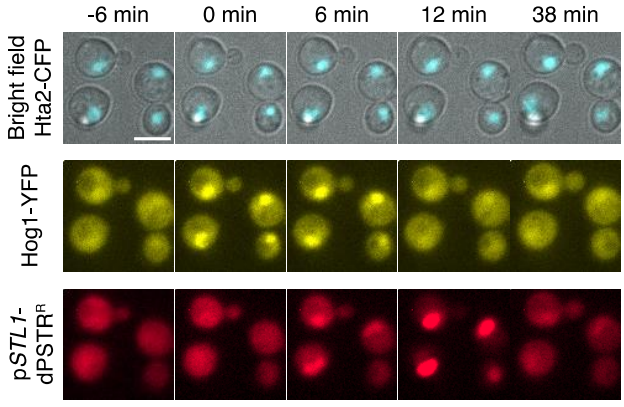


Figure 2.8 Microscopy images of joint signaling and gene expression measurements

Microscopy images of a strain with Hog1 tagged with mCitrine and carrying the unstable p*STL1*-dPSTR^R that was challenged by 0.2M NaCl. The nuclear accumulation of Hog1-YFP precedes protein expression. Scale bar, 5 µm.

transduction in the HOG pathway (Hersen et al., 2008; Mettetal et al., 2008; Pelet et al., 2011; Reiser et al., 1999). A few minutes after Hog1 relocates in the nucleus, the p*STL1*-dPSTR^{R} starts to accumulate in the nucleus.

Figure 2.9A displays the changes in cell area upon increasing osmotic challenges, which trigger an immediate shrinking of the cells. Depending on the severity of the stress, the cells need between 10 and 30 min to recover their original sizes. Figure 2.9B depicts Hog1 relocation, quantified as the ratio of nuclear to cytoplasmic YFP fluorescence as a function of time, which is almost a mirror image of the cellular adaptation process. Indeed, Hog1 enters the nucleus quickly after stress, and it returns to a uniform localization when cells recover their original sizes. The MAPK drives the adaptation process by increasing the production of glycerol, causing a negative feedback on its own activity (Albertyn et al., 1994; Hohmann et al., 2007).

In comparison, the dynamics of protein production measured by the p*STL1*-dPSTR^{R} is delayed because a number of events need to be completed before proteins can be produced (Figure 2.9C). These biological processes include promoter activation, which can require TF recruitment and chromatin remodelling, and mRNA synthesis and translation (de Nadal and Posas, 2010). The maximal protein production, corresponding to the peak in nuclear enrichment, is reached when Hog1 returns to its basal level. It has been shown that active Hog1 is closely associated with all the steps of transcription and even travels along the ORFs with the elongation complex (de Nadal and Posas, 2010). Therefore, transcription is expected to stop as soon as Hog1 activity returns to its basal level.

When measured at the population level, both Hog1 activity and p*STL1* expression increase with the strength of the stress; however, this correlation does not hold true for single-cell measurements (Figure 2.10A). This discrepancy has been attributed to slow chromatin-remodelling steps occurring at the promoter, which decouple the expression from the level of MAPK activity (Pelet et al., 2011).

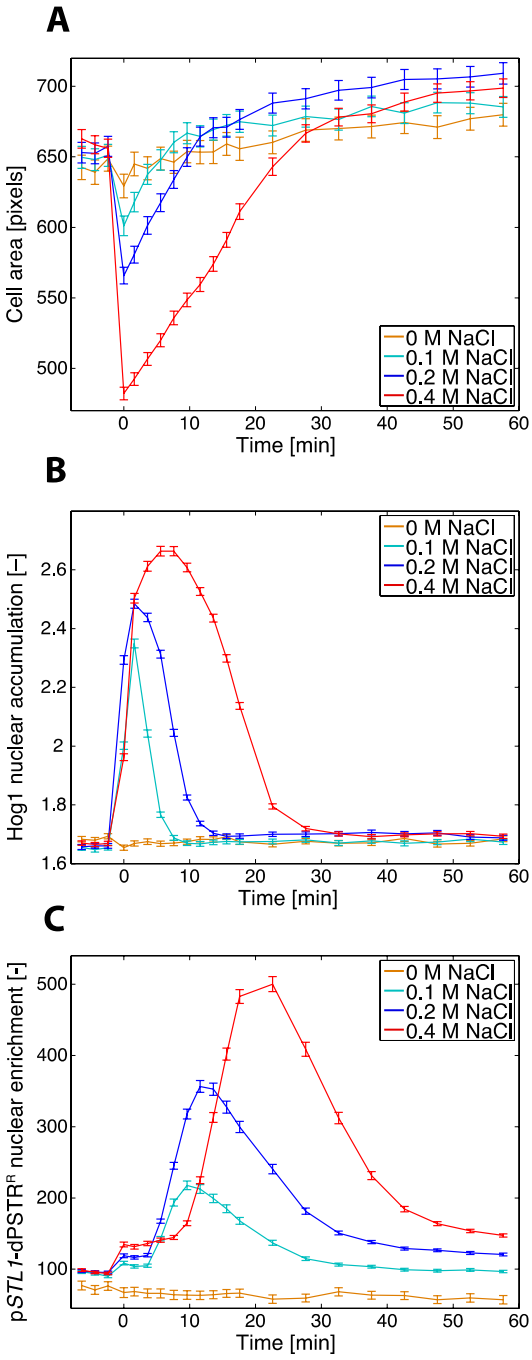


Figure 2.9 Quantifications of cell area, signaling and gene expression

Quantification of the cell area (A) Hog1 nuclear accumulation, measured as the ratio between nuclear and cytoplasmic YFP fluorescence (B) and p*STL1*-dPSTR^{R} nuclear enrichment (C) for cells stimulated with 0 (orange, N_C.467), 0.1 (cyan, N_C.558), 0.2 (blue, N_C.655) and 0.4M NaCl (red, N_C.802).

One prediction from such a model is that cells that become transcriptionally active earlier tend to express more proteins, since they profit from a longer temporal window of gene expression for the same MAPK activity. Indeed, it can be seen in Figure 2.10B that there is an inverse correlation between the time when the protein production is detected and the level of

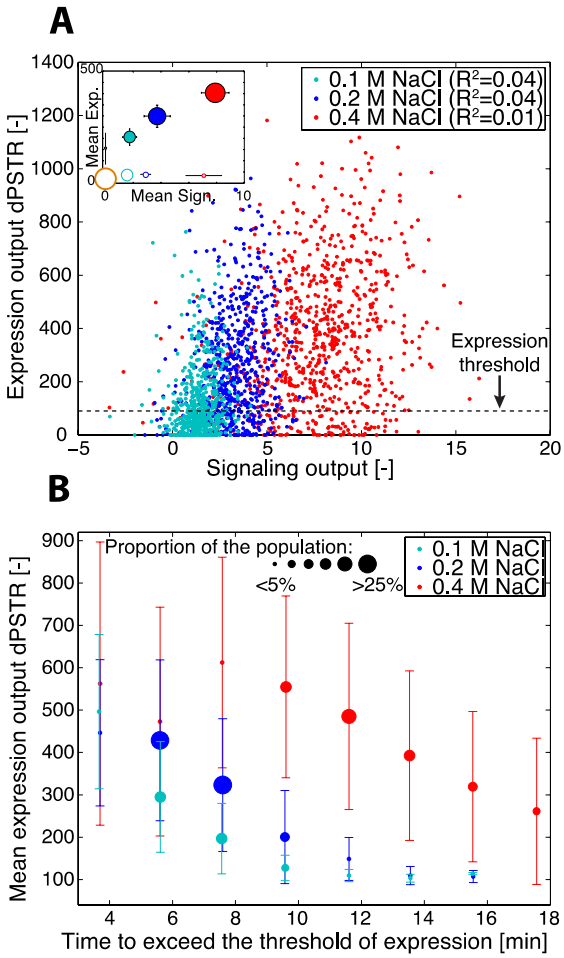


Figure 2.10 Correlation between signaling and gene expression output

A. Scatter plot of the signalling output measured as the integral below the Hog1 nuclear accumulation curve versus the expression output measured as the maximum in p*STL1*-dPSTR^{R} nuclear enrichment. The dashed line represents the expression threshold. The mean signalling output versus the mean expression output for the expressing cells (filled circles) and the non-expressing cells (empty circles) is plotted in the inset. The size of the marker is indicative of the percentage of cell in each category.

B. Correlation between the time needed to overcome the expression threshold and the expression output. The mean expression output and the standard deviation were calculated for groups of cells, which exceed the expression threshold at the same time point. The marker size is indicative of the percentage of cells (from the total population) in each group.

protein produced. The unique ability of the dPSTR to measure the dynamics of protein production in real time allows us to confirm that there is a direct influence of the time when the promoter is activated on the output in protein production.

Similar experiments were performed with another stress-inducible promoter p*GPD1* (Supplementary Figure 6; (Alepuz et al., 2001; Rep et al., 1999b)). In comparison with p*STL1*, which is repressed under normal growth conditions, p*GPD1* has a low basal level of transcription, which is increased under conditions of hyper-osmotic stress in a Hog1-dependent manner. The combined measurements of the Hog1-YFP relocation dynamics and the p*GPD1*-dPSTR^{R} expression provided in general a very similar picture to the p*STL1* measurements. One obvious difference is that a larger fraction of the population expresses the p*GPD1*-dPSTR^{R} at low stress levels (Inset in Supplementary Figure 6E). In addition, we noticed that p*GPD1*-dPSTR^{R} seems to be expressed more rapidly than the p*STL1*-dPSTR^{R} (compare panels D and F from Supplementary Figure 6 to Figure 2.10).

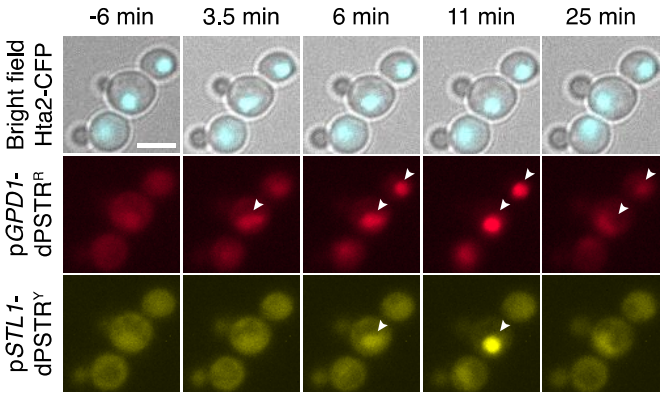


Figure 2.11 Microscopy images of a strain bearing two dPSTRs

Microscopy images of cells carrying p*GPD1*-dPSTR^{R} (RFP channel) and p*STL1*-dPSTR^{Y} (YFP channel) stimulated with 0.2M NaCl. The two reporters are built with two orthogonal pairs of SynZips. The arrowheads indicate the nuclei with accumulated FPs, highlighting the different timings of accumulation of each dPSTR. Scale bar, 5 µm.

To better characterize this difference in expression dynamics between the two stress-inducible promoters, we combined a p*GPD1*-dPSTR and a p*STL1*-dPSTR in the same cells (Figure 2.11). This combination is possible because they possess two sets of orthogonal SynZips (SZ1/SZ2 and SZ3/SZ4 (Reinke et al., 2010)) and drive the relocation of either a red or a yellow FP variant (resp. p*GPD1*-dPSTR^{R} and p*STL1*-dPSTR^{Y}). Following a 0.2 M stimulus, the cells were imaged with 35 s resolution (Figure 2.12A). The average response of the population indicates a 1.5 min delay between p*GPD1* and p*STL1* expression in favour of p*GPD1* (Figure 2.12B). This delay remains constant during the complete period of expression of the two promoters. Moreover, since both reporters are present in the same cell, we can correlate their expression output within individual cells. Approximately a third of the cells expressing p*GPD1* do not express p*STL1*, while only a few cells were identified as p*STL1* positive only (Figure 2.12C).

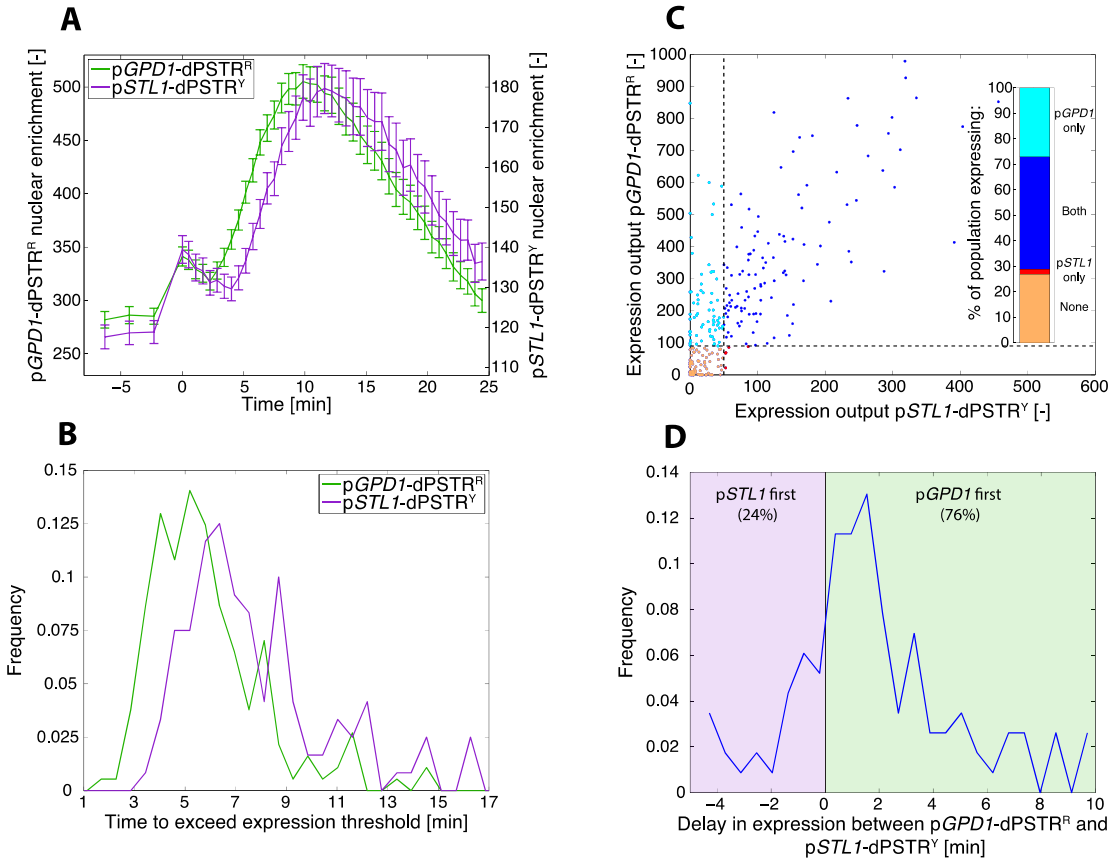


Figure 2.12 Dynamic measurements of two osmostress-induced promoters

A. Quantification of the nuclear enrichment of p*GPD1*-dPSTR^{R} (green, left axis) and p*STL1*-dPSTR^{Y} (purple, right axis) in course of time (N_C.260).

B. Histograms of the time needed to overcome the expression threshold for cells expressing the indicated promoter.

C. Correlation between the expression output of p*GPD1*-dPSTR^{R} and the one of p*STL1*-dPSTR^{Y} in single cells (R^2=.0.48). The dashed lines represent the expression thresholds for each dPSTR. The inset is showing the fraction of the population expressing either p*GPD1* alone (cyan), p*STL1* alone (red), both promoters (blue) or none (orange).

D. The delay between p*GPD1* and p*STL1* expression in cells that express both dPSTRs calculated from the difference in time to overcome the expression threshold for both reporters. Positive times represent cells where p*GPD1* overcomes the expression threshold first (green area, 76% of the cells expressing both promoters), and negative or null times indicates that p*STL1* is expressed before or at the same time as p*GPD1* (purple, 24%).

In the cells expressing both reporters, we could observe that p*GPD1* expression precedes p*STL1* expression in a large majority of cells (Figure 2.12D). We can infer that this temporal delay observed for two promoters within the same cell, which are thereby experiencing the same level of Hog1 activity and are controlled by the same TF Hot1, can be attributed either to different efficiency of the TFs associated with each promoters to recruit the transcriptional machinery or to the chromatin remodelling step (Capaldi et al., 2008). We propose that the important remodelling taking place on the repressed p*STL1* promoter is largely absent from the p*GPD1* (Mas et al., 2009).

Dynamic noise quantification

A prediction from our observations and from previous studies (Neuert et al., 2013; Pelet et al., 2011) is that slow stochastic chromatin remodelling is responsible for a large intrinsic noise in p*STL1* expression. Indeed, the recruitment of the chromatin remodelling machinery is thought to occur stochastically at each locus, creating large variability in expression within the same cell. However, this variability should be mostly absent from the p*GPD1*-dependent expression. To monitor the temporal fluctuations of this noise, we combined two dPSTRs in the same cell controlled either by two *STL1* promoters or two *GPD1* promoters (Figure 2.13 and Supplementary Figure 7). The absolute intensity of nuclear relocation measured in both channels is different, due to disparities in FP brightness, but the dynamics of relocation are similar for both reporters (Supplementary Figure 8). For each time point in the data set, we can correlate the instantaneous amplitude of nuclear relocation in the yellow and in the red channels in every single cell. Three time points in the early, intermediate and highest phase of expression have been selected (arrows in Figure 2.13) and are plotted in Figure 2.14. It is apparent that throughout the time-lapse, the two p*GPD1*-dPSTRs display a very tight correlation, which is largely absent from the double p*STL1*-dPSTRs strain. For example, the upper cell shown in panel D is expressing first the p*STL1*-dPSTR^{R} copy, and only later the p*STL1*-dPSTR^{Y}. Thus, both the dynamics and the levels of expression can significantly vary between the two p*STL1*-dPSTRs within the same cell.

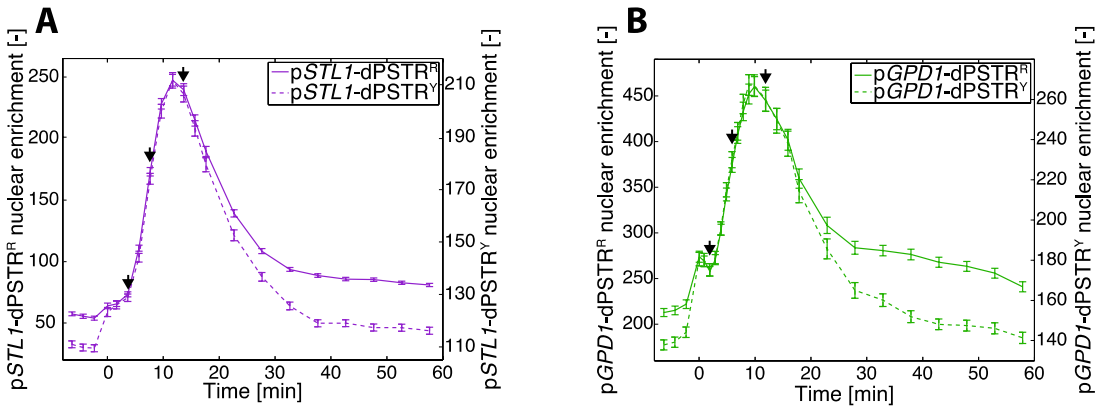


Figure 2.13 Measurements of osmostress induction of two copies of p*STL1* or p*GPD1* in the same cells

Quantification of the nuclear enrichment of dPSTR^{R} (right axis) and dPSTR^{Y} (left axis) for a strain carrying two p*STL1*-dSPTRs (A) or two p*GPD1*-dPSTRs (B) in cells stimulated with 0.2M NaCl (resp. N_C.958 and N_C.368).

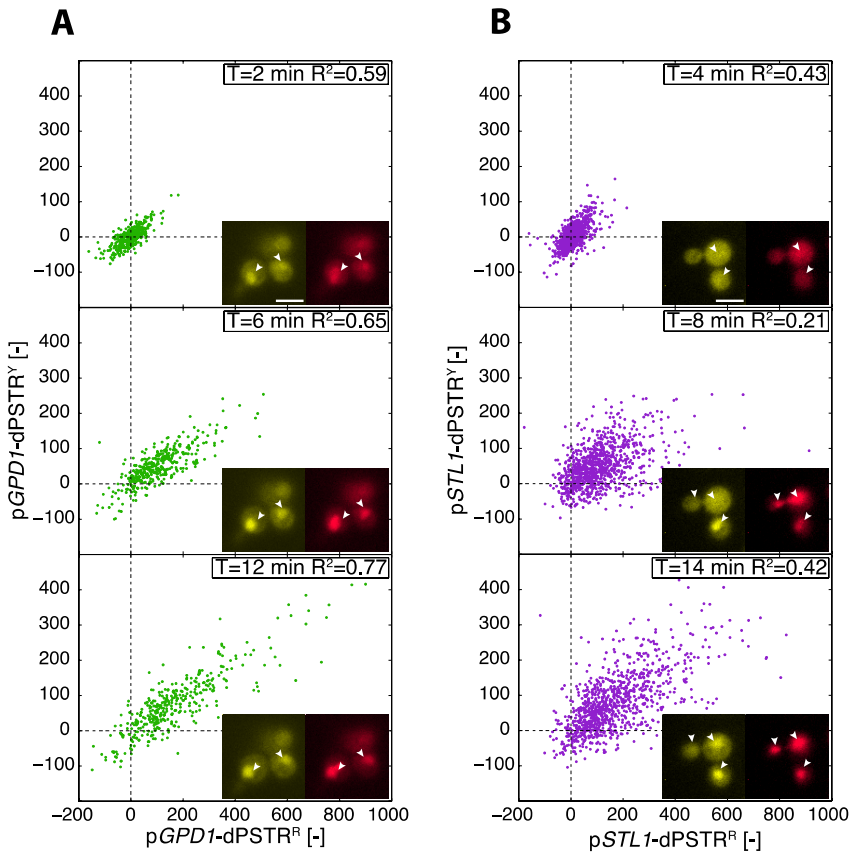


Figure 2.14 Single-cell correlation of expression from two copies of the same promoter measured by the two dPSTRs

Correlation of the instantaneous corrected nuclear enrichment of p*GPD1*-dPSTRs (A) or p*STL1*-dPSTRs (B) in YFP and RFP channels at the indicated times after induction. Pictures indicate representative cells at the same time points. Arrowheads are highlighting nuclei in the focal plane Scale bars, 5µm.

Using these data, we calculated the evolution of the intrinsic expression noise over time (Elowitz et al., 2002; Raser and O'Shea, 2004) (Figure 2.15; see Methods). In the p*GPD1* case, at time zero, the intrinsic noise is relatively low and drops further as the two dPSTRs are expressed synchronously. Interestingly, in the p*STL1* case, protein production can arise stochastically from either locus, resulting in an initial increase in the proportion of intrinsic noise. Later on, as the two loci are expressed, this component of the noise tends to decrease, because transcripts in the same cell share the same translational machinery. The expression capacity of a cell, which is linked to the number of ribosomes, is thought to be a major determinant of extrinsic noise (Colman-Lerner et al., 2005). A similar behaviour can be observed with cells induced with 0.1 and 0.4 M NaCl (Supplementary Figure 9). To conclude, the dPSTR system allows the accurate measurement of the evolution of the expression noise in

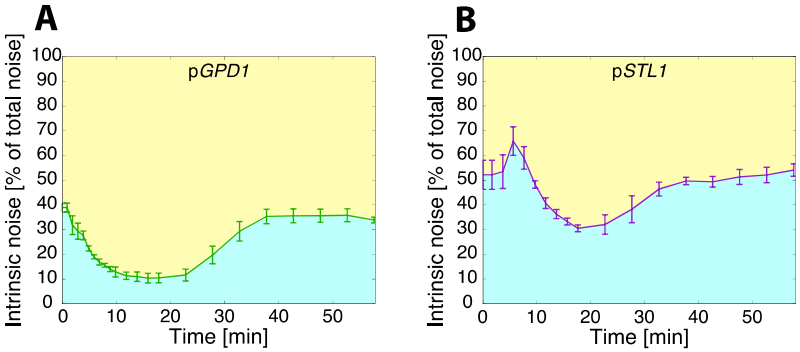


Figure 2.15 Real-time measurements of expression noise

Evolution of the intrinsic expression noise for p*GPD1* (A) and p*STL1* (B). The blue area under the curve represents the proportion of intrinsic noise, and the yellow area above represents the proportion of extrinsic noise.

real time, which cannot be accomplished using other current methods.

Induction of successive rounds of protein expression

The property of the dPSTR to return to its initial cytoplasmic distribution after degradation of the induced moiety should allow the measurement of multiple rounds of protein expressions. To test this, cells bearing a Hog1-YFP combined either with a p*STL1*-dPSTR^R or a p*GPD1*-dPSTR^R were subjected to a first hyper-osmotic stress. Forty-three minutes later, the NaCl concentration was further increased to double the osmolarity in the medium (Figure 2.16A). The two hyper-osmotic events led to two shrinking and recovery phases driven by the activation of Hog1 (Figure 2.16B and Supplementary Figure 10A). Each period of Hog1 activity resulted in p*STL1*-dPSTR^R and p*GPD1*-dPSTR^R relocation (Figure 2.16C to D). At the population level, there is a linear correlation between Hog1 signalling output and the expression output of the two promoters, which is maintained for the first step and the second step of stimulation (Supplementary Figure 10B, C).

Interestingly, the level of Hog1 activity in individual cells is weakly correlated between each stress (Figure 2.16E). Cells that have responded strongly in the first step are also more likely to respond strongly in the second step. As expected from our previous single induction experiments, no single cell correlation between Hog1 activity and subsequent protein expression is observed in either pulse, neither for p*STL1* nor for p*GPD1* (Supplementary Figure 10D to G). Because of the stochastic activation induced by the chromatin-remodelling step, the p*STL1* expression cannot be correlated between each pulse (Figure 2.16F). We can identify cells that responded to the first pulse but not the second one or vice-versa, or cells that responded to both steps or not at all. This large diversity in responses demonstrates that the ORF does not retain a memory of previous transcription events. This is in agreement with previous studies which have demonstrated that histones are reassembled rapidly once the transcription has stopped in order to repress the locus (Klopf et al., 2009). It is however more surprising to see that the p*GPD1* expression in the first and second pulse is not correlated either (Figure 2.16G). A large majority of the cells express in both pulses but do it to a different extent. This suggests that the cellular parameters that allowed a strong correlation of the *GPD1* expression during a single pulse (Figure 2.15) are not maintained from one stimulus to the next to allow a correlation across time (Figure 2.16G).

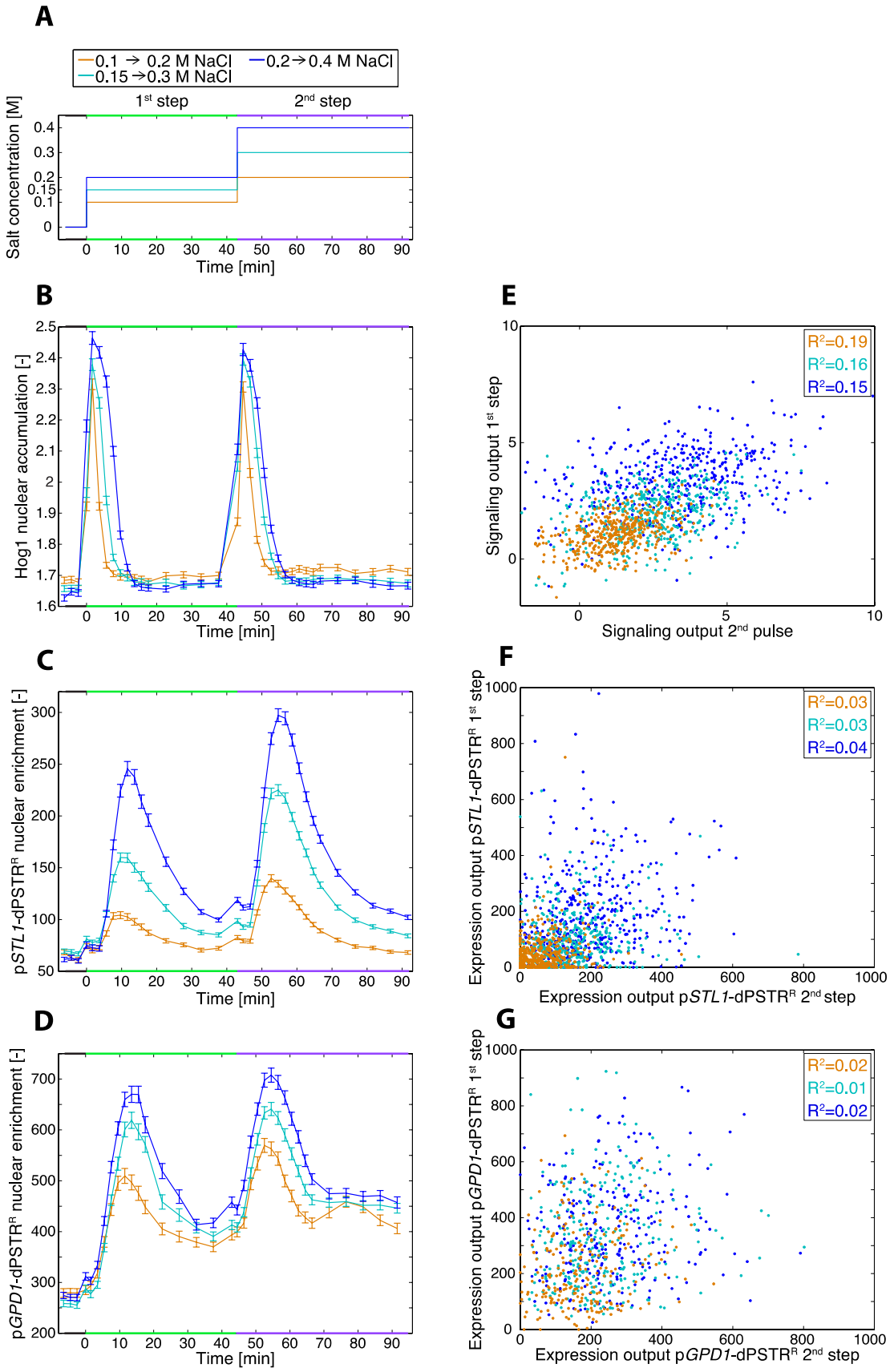
Figure 2.16 Consecutive hyper-osmotic stresses result in uncorrelated transcription events. (on the next page)

A. Evolution of the NaCl concentration over time. Cells were stimulated at time 0 with a given concentration of NaCl. A second hyper-osmotic stress of similar amplitude was performed 43 min later by doubling the concentration of NaCl in the well: 0.1-0.2 (orange), 0.15-0.3 (cyan), 0.2-0.4 (blue) (throughout the entire figure).

B-C. Quantification of the average Hog1 nuclear accumulation (B) and p*STL1*-dPSTR^R nuclear enrichment (C) for cells subjected to the steps in NaCl concentrations depicted in A (0.1-0.2: N_C.429; 0.15-0.3: N_C.450; 0.2-0.4: N_C.449).

D. Quantification of the average p*GPD1*-dPSTR^R nuclear enrichment, from a different strain, subjected to the steps in NaCl concentrations depicted in a (0.1-0.2: N_C.235; 0.15-0.3: N_C.296; 0.2-0.4: N_C.276). These cells also bear Hog1- YFP that showed the same behaviour as in B.

E-G. Correlations of the signalling outputs (E), of the expression outputs of p*STL1*-dPSTR^R (F) or of the expression outputs of p*GPD1*-dPSTR^R (G) for the two stress events.


A
0.1 → 0.2 M NaCl
0.2→0.4 M NaCl
0.15 →0.3 M NaCl
1st step
2nd step
Salt concentration [M]
Time [min]
B
Hog1 nuclear accumulation [-]
Time [min]
C
pSTL1-dPSTR^R nuclear enrichment [-]
Time [min]
D
pGPD1-dPSTR^R nuclear enrichment [-]
Time [min]
E
R²=0.19
R²=0.16
R²=0.15
Signaling output 1st step
Signaling output 2nd pulse
F
R²=0.03
R²=0.03
R²=0.04
Expression output pSTL1-dPSTR^R 1st step
Expression output pSTL1-dPSTR^R 2nd step
G
R²=0.02
R²=0.01
R²=0.02
Expression output pGPD1-dPSTR^R 1st step
Expression output pGPD1-dPSTR^R 2nd step

Discussion

In this paper, we have demonstrated the use of a synthetic translocation reporter to quantify the dynamics of protein synthesis emanating from a promoter of interest in live single cells. The clear advantage of the dPSTR system is that it provides measurements at the minute timescale of protein expression events that FP cannot offer, due to the slow maturation time of the fluorophore. Even the fast folding sfGFP needs on average 6 min to become fluorescent, which precludes the fast measurement of the protein expression dynamics (Houser et al., 2012; Khmelinskii et al., 2012). Moreover, there is no comparable fast maturing FP in other spectral channels, limiting this technique to only one promoter in a given strain. In our assay however, the quantification of the dynamics of protein expression are limited only by the import rate into the nucleus, which happens on the sub-minute timescale (Timney et al., 2006). Since the assay relies on constitutively expressed FPs, the FP can be easily exchanged without affecting the dynamics of expression. Therefore, multiple reporters can be combined using appropriate FP spectral variants and orthogonal SynZip pairs (Figures 2.11 to 2.16). Based on the palette of available SynZips and FPs, we can envision to combine up to three dPSTRs in the same cell (Reinke et al., 2010; Shaner et al., 2005). Note also that while it is feasible to estimate the nuclear enrichment only based on the fluorescent channel of the dPSTR and the whole-cell object, an exact quantification of the nuclear and cytoplasmic intensities will require a nuclear tag occupying one of the few fluorescent channels available for FP measurements.

Luminescence microscopy has also demonstrated the capability of recording the fast dynamics of protein expression (Mazo-Vargas et al., 2014; Suter et al., 2011). However, due to the low photon flux generated by the luciferase, long integration times that can last several minutes are required. To achieve sub-minute temporal resolution, Mazo-Vargas *et al.* recorded a Z-stack with five planes with 10 s exposure for each plane (Mazo-Vargas et al., 2014). In microscopy experiments, a tradeoff has to be reached between the frequency of acquisition, and the number of XY-stage positions visited. The longer the acquisition at each field of view lasts, the fewer XY-positions can be imaged. The parallel imaging of multiple positions can greatly improve the throughput of an experiment by allowing to increase the number of single-cell recordings, thereby improving the statistics of the measurements. To reach a 35 s time resolution, we have imaged three positions in three different wells recording close to 50 images per time point with more than 250 cells monitored for each well. The long exposure time required for luminescence data acquisition would clearly prevent reaching such imaging frequency and high number of cells, thereby lowering the resolution and statistical significance of the acquired data set.

The MS2 or PP7 technologies, which allow the detection of mRNA transcription in individual cells, offer complementary information to our reporter. We have implemented it for the detection of mRNA production arising from the p*STL1* promoter. The dynamics of the transcription site apparition indicate that there is roughly a 2 min delay between the production of the mRNA and appearance of the protein, during which protein synthesis is occurring. While the PP7 signal at the transcription site can provide rich information about the transcription dynamics, it has to be noted that its detection and automated quantification is not straightforward. PP7 measurements require the acquisition of Z-stacks with high magnification, thereby limiting the speed of acquisition and the size of the field of view, and hence, the number of cells that can be imaged. Moreover, the automated detection of the fluorescent signal at the transcription site is more complex than the measurement of the relocation of the dPSTR sensor from the cytoplasm to the nucleus of the cell.

We believe that our assay offers important advantages over current techniques. However, one limitation arises from the level of expression of the constitutive FP, which has to

match to a certain extent the amount of the inducible binding partner (Supplementary Figure 4). If the constitutive FP is expressed at too high levels, it becomes difficult to quantify small changes in protein expression. At the other extreme, low levels of FPs can lead to the saturation of the nuclear signal when all the FPs are bound to a 2xNLS-SZ moiety. In addition to providing important advantages in the quantification of the expression dynamics, the presence of the destabilizing sequence strongly reduces the chances of observing a saturation of the signal in the nucleus, since the 2xNLS-SZ peptide does not accumulate in the cell (Supplementary Figure 5A to C). However, the sensitivity of the measurement is slightly decreased and therefore, when using the same threshold, fewer expressing cells are detected with the unstable version of the dPSTR (Figure 2.12C). In this study, we used a mildly expressed promoter p*RPL24A* to control the FP abundance. We have tested another stronger constitutive promoter and have observed a lower sensitivity of this construct to detect weak expression at 0.1 M NaCl (Supplementary Figure 4E).

The dPSTR technique applied to the *STL1* and *GPD1* promoters allowed to quantify events that could not be observed before with techniques based on FP expression. Our previous analysis of the p*STL1*-induced expression led us to postulate that stress responsive genes were induced with a large temporal variability due to slow chromatin remodelling events taking place at the induced locus (Pelet et al., 2011). The real-time observation of protein synthesis with dPSTR allows now to confirm the large variability in dynamics and levels of p*STL1* expression. In comparison with a *GPD1* promoter, p*STL1* is induced with slower dynamics and larger intrinsic noise. Since there is basal expression of *GPD1* in unstressed cells, we believe that switch from a repressed to an active promoter via the chromatin remodelling step could be largely responsible for the striking kinetic differences measured between these two promoter elements. Interestingly, and somewhat in contradiction with our prediction, the correlation between the Hog1 signalling output as measured by the integral below relocation curve and the p*GPD1*-dPSTR output is not better than with p*STL1*. This implies that it is not only the chromatin-remodelling step that uncouples signalling and protein expression outputs. However, in the case of p*GPD1*, we cannot attribute this to a stochastic activation of the transcribed locus, since two p*GPD1* promoters in the same cells are strongly correlated. Therefore other extrinsic factors such as the amount of polymerases or the number of ribosomes could strongly influence the expression output independently of the Hog1 signal. It remains to be tested what parameters and how much of the kinase activity profile are encoded in the expression output.

More generally, this novel assay will now allow us to investigate the contribution of the various factors active at the promoters that control the kinetics of gene activation. Moreover, thanks to the destabilized nature of the reporter, it will become possible to study the processes implicated in the memory in successive stress events, such as the repositioning of the chromatin on the transcribed locus. In addition, due to the conservation of NLS sequences, this reporter could be easily adapted to quantify the dynamics of protein synthesis in higher eukaryotes.

Methods

Strains and plasmids.

Yeast strains and plasmids are listed in Supplementary Tables 1 and Annexe 1. dPSTR plasmids were constructed by cloning different parts of the reporter into the single integration vectors pSIVU or pSIVL vector backbone (see below). The p*STL1*-dPSTR^{R} monitors expression arising from the promoter p*STL1*, and is based on an RFP variant (mCherry) (R for red), and the SynZips SZ1 and SZ2 (Reinke et al., 2010; Thompson et al., 2012). The p*STL1*-dPSTR^{Y} is based on the YFP variant mCitrine (A206K L221K) and the pair of SynZips SZ3 and SZ4. The FPs were expressed from the constitutive promoters p*RPL24A* or p*RPL15A*, cloned between SacI and XbaI. The FPs mCherry or mCitrine were cloned HindIII-SalI. The SynZips were cloned SalI-NheI. For the inducible part, a second MCS was designed and subcloned between AatII-SphI. The promoter p*STL1* (-800 to -1) is cloned SacI-XbaI. The destabilization sequence UbiY is cloned XbaI-HindIII (Pelet et al., 2011). To generate the p*GPD1*-dPSTRs, the p*STL1* promoter was replaced by p*GPD1* (-1,000 to -1) in all the constructs. The SynZips were cloned between SalI-NheI, and the CYC1 or SIF2 terminators XhoI-KpnI. We deposited at Addgene a set of plasmids for the dPSTR system, along with maps and sequences.

The plasmids were transformed in a yeast strain from a W303 background, bearing a Hta2-CFP nuclear marker (ySP37 (Durandau et al., 2015)). Hog1 was tagged with mCitrine using the plasmid pKT139 (Sheff and Thorn, 2004).

To integrate both transcriptional units of the dPSTR, we developed a set of single integration vectors that entirely replace the selection marker gene by recombination with its promoter and terminator (Wosika et al., 2016). The pSIV carries two MCSs, separated by the transcription unit that will compensate for the auxotrophy of the strain. The first MCS is constructed in the pSIV backbone, whereas the second MCS is cloned in a different intermediate vector, and subcloned into the pSIV between AatII and SphI. The pSIVL integrates into the *leu2* locus and carries a scrambled sequence of the *LEU2* gene, whereas the pSIVU integrates into the ura3 locus and comprises the *Candida albicans URA3* gene, in order to avoid any undesirable recombination. The plasmids were transformed into our reference strain after digestion with PacI, separating the bacterial part from the yeast part of the plasmid.

For each transformation, 8–10 clones were screened based on their fluorescence intensities, and four clones with similar expression levels of the FP were further analysed by a time-lapse experiment upon stimulation with 0.2 M NaCl, to discard clones that would display an aberrant relocation behaviour.

Sample preparation

The cells were grown overnight in synthetic medium to saturation (YNB: CYN3801, CSM: DCS0031, ForMedium). They were diluted to an OD_{600} of 0.05 in the morning and grown for at least 4 h before the start of the experiment. All the time-lapse experiments were performed in well slides, for which selected wells of 96-well-plates (MGB096-1-2LG, Matrical Bioscience) were coated with a filtered solution of Concanavalin A in H2O (0.5 mg/ml, C2010-250MG, Sigma-Aldrich) for 30 min, rinsed with H2O and dried for at least ten hours. Before the experiment, the cells were diluted to an OD_{600} of 0.04, briefly sonicated, and 200 ml of cell suspension were added to a well. Imaging was started 30 min later to let cells settle to the bottom of the well. To stimulate the cells, 100 ml of a 3X stress solution was added in the well (see Supplementary Table 2 for the concentrations).

Microscopy

Images were acquired on a fully automated inverted epi-fluorescence microscope (Ti-Eclipse, Nikon) controlled by micro-manager and placed in an incubation chamber set at 30°C, with a 40X oil objective and appropriate excitation and emission filters(Edelstein et al., 2010). The excitation is provided by a solid-state light source (SpectraX, Lumencor). The images were recorded with an sCMOS camera (Flash4.0, Hamamatsu). A motorized XY-stage allowed recording multiple fields of view at every time point. CFP (50 ms), RFP (300 ms) and YFP (150 ms for Hog1, 300 ms for dPSTR^{Y}) and two bright-field (10 ms) images were recorded at time intervals varying from 35 s to 5 min.

Data analysis

Time-lapse movies were analysed with the YeastQuant platform (Supplementary Figure 2) (Pelet et al., 2012). The nuclei of the cells were segmented by thresholding of the CFP images. The contour of the cell around each nucleus was detected using two bright-field images. The cytoplasm object was obtained by removing the nucleus object expanded by two pixels from the cell object. Dedicated scripts in Matlab (The Mathworks) were written to further analyse the data. Only cells tracked from the beginning to the end of the movie were taken into consideration. In addition, a quality control was applied on each trace and only the traces with low variability in nuclear and cell area, and nuclear CFP fluorescence were kept for further analysis (typically more than 65% of the tracked cells). The curves displayed in the figures represent the mean and s.e.m. of these selected traces for one representative experiment out of three true biological replicates.

For each cell, the difference between its average intensity in the nucleus and in the cytoplasm was calculated at every time point to plot the nuclear enrichment of the dPSTR and of the Venus. The basal level is calculated as the mean of the two time points immediately following the addition of NaCl (T ¼ 0 min and T ¼ 2 min), in order to take in account the sudden increase in fluorescence intensity triggered by the abrupt nuclear enrichment upon shrinking of the cells when NaCl is added. This is an artefact of the measurement and not a transcriptional response of the cell (Supplementary Figure 1). The maximal enrichment was obtained for each single-cell trace smoothed by a moving average of three points. For each single cell, the corrected nuclear enrichment of the dPSTR was calculated as the smoothed trace subtracted by its basal level. The expression output represents the maximal corrected nuclear enrichment of the dPSTR. The expression thresholds were determined based on the expression outputs of the non-induced populations. The time to overcome the expression threshold was defined as the first time point when the corrected nuclear enrichment of the dPSTR is equal to or greater than the expression threshold. Finally, the time to reach half of the expression output was extracted from each non-smoothed trace as the first time point at which the nuclear enrichment is equal or greater than half of the maximal nuclear enrichment. The intrinsic and extrinsic noise were calculated according to the formula from Elowitz *et al.*(Elowitz et al., 2002):

$$\eta^2_{int} = \frac{\langle (r_i - y_i)^2 \rangle}{2\langle r_i \rangle \langle y_i \rangle} \quad \eta^2_{ext} = \frac{\langle r_i y_i \rangle - \langle r_i \rangle \langle y_i \rangle}{\langle r_i \rangle \langle y_i \rangle} \quad \eta^2_{tot} = \frac{\langle r_i^2 + y_i^2 \rangle - 2\langle r_i \rangle \langle y_i \rangle}{2\langle r_i \rangle \langle y_i \rangle}$$

since: $\eta^2_{tot} = \eta^2_{int} + \eta^2_{ext}$ we plot the fraction of the intrinsic noise as: $\frac{\eta^2_{int}}{\eta^2_{tot}}$

r_i and y_i are the normalized nuclear accumulation from the i^{th} cell in the red and yellow channels, respectively. The normalization factors were obtained from the highest and lowest average intensity from the entire data set for one replicate.

Hog1 nuclear accumulation was calculated as the ratio of the nuclear over the cytoplasmic intensities. The signaling output of each single cell was calculated as the area under Hog1 curve during a time interval determined as the period of Hog1 activity for the average population.

mRNA transcription sites measurements

The measurement of mRNA transcription sites was performed using the PP7 technique (Larson et al., 2011). The PP7-2xGFP under the control of the p*MET* promoter was cloned in an integrative vector pRS304 integrated in the *TRP1* locus. The original promoter p*POL1* from vector pDZ306 (Addgene# 35196) was replaced by the promoter p*STL1* and drives the transcription of 24 stem loops, which can be tightly bound by a dimer of PP7. The loops were integrated in the *glt1* locus to generate a very long transcript of 6.4 kb to facilitate the visualization of the PP7 foci. Cells were imaged with the same set-up as described above. A 60 objective and piezo Z-stage (Nano-Z200, Mad City Labs) was used and 3 Z-stacks from -2 to +2 µm were acquired every 30 s. For the identification of the nucleus, an Hta2-mCherry tag was present in the strain and imaged with similar Z-stacks. Two bright-field images were also recorded for segmentation of the cells. The maximum intensity projections of the GFP and RFP images were used in the YeastQuant pipeline. As a measurement of the transcription site intensity, the difference between the average intensity of the 20 brightest pixels in the nucleus and the average intensity of the nucleus is calculated.

Flow Cytometry

The flow cytometry experiments were performed as previously described in (Pelet et al., 2013). Briefly, a cell population was induced by NaCl and samples were taken at different time points and immediately blocked with cycloheximide (0.1 mg/ml). Cells were incubated for at least 2 h to allow the maturation of the FP before being measured by flow cytometery (FACSCalibur, BD). Ten thousand events were measured and a gating was applied on the forward and side scattering to discard clusters of multiple cells.

Supplementary Tables

Supplementary Table 1: List of yeast strains used in this study

Strain	Background	Genotype	Plasmid
ySP2	W303	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*	
ySP37	W303	*HTA2-CFP*	
yDA119	ySP37	*HTA2-CFP* *URA3: pSTL1-Venus-dPSTR^{R2-1}* (*)	 pDA176
ySP9	W303	*LEU2: pSTL1-quadrupleVenus*	pSP34
ySP374	ySP37	*URA3: HTA2-mCherry* *trp1: PP7-GFP* *glt1: his3 pSTL1-24xPP7-SL*	 pSP268 pSP264
yDA134	ySP37	*HTA2-CFP* *URA3: pSTL1-dPSTR^{R2-1}* *HOG1-mCitrine: HIS3*	 pDA183
yDA137	ySP37	*HTA2-CFP* *URA3: pSTL1-dPSTR^{R2-1}* *LEU2: pSTL1-dPSTR^{Y4-3}*	 pDA183 pDA199
yDA139	ySP37	*HTA2-CFP* *URA3: pGPD1-dPSTR^{R2-1}* *LEU2: pGPD1-dPSTR^{Y3-4}*	 pDA193 pDA200
yDA155	ySP37	*HTA2-CFP* *URA3: pGPD1-dPSTR^{R2-1}* *HOG1-mCitrine: HIS3*	 pDA193
yDA156	ySP37	*HTA2-CFP* *URA3: pGPD1-dPSTR^{R2-1}* *LEU2: pSTL1-dPSTR^{Y3-4}*	 pDA193 pDA199
yDA93	ySP37	*HTA2-CFP* *URA3: pSTL1-dPSTR^{R2-1}* *HOG1-mCitrine: HIS3*	 pDA140
yDA85	ySP37	*HTA2-CFP* *URA3: pSTL1-dPSTR^{R2-3}*	 pDA137
yDA112	ySP37	*HTA2-CFP* *URA3: pGAL1-dPSTR^{R2-1}*	 pDA169

* The numbers in the superscript dPSTR^{R2-1} indicate the pair of SynZip used.

Supplementary Table 2: Concentrations of inducers

Final concentration	Concentration of inducer solution	Figure	Final concentration after second stress	Concentration of inducer solution	Figure
0.1 M	0.3 M	1, 3, 6 S2, S3, S4, S6, S8, S9	0.2 M	0.5 M	6, S10
0.15 M	0.45 M	6, S10	0.3 M	0.75 M	6, S10
0.2 M	0.6 M	1, 2, 3, 4, 5, 6, S1, S3, S4,, S6, S7, S8, S9	0.4 M	1 M	6, S10
0.4 M	1.2 M	1, 3,S1, S3, S4, S6, S8, S9			

Supplementary Table 3: List of yeast strains used for fluorescence microscopy calibration

Genotype: W303, *Protein-mCherry:URA3*

Strain	Protein tag	Protein Number
ySP37[1]	-	0
ySP408	Ygr117c	1280
ySP426[1,2]	Ste12	1920
ySP534[1]	Kss1	5480
ySP399	Hog1	6'780
ySP400	Apt1	11'200
ySP405	Tda1	10'200

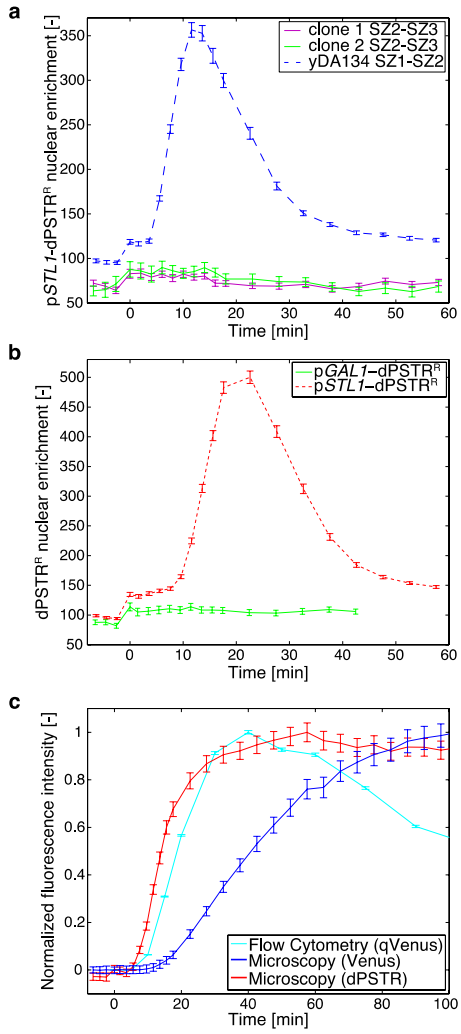
[1]These strains also bear a Hta2-CFP tag

[2] Protein-mCherry:HIS

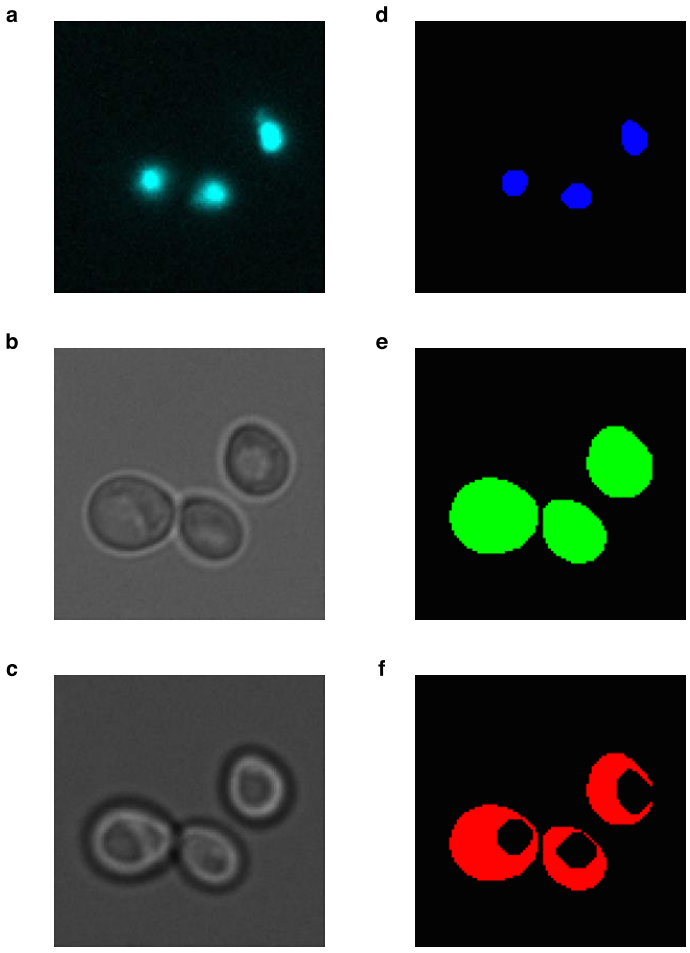
Acknowledgments

We thank the members of the Pelet and Martin lab for helpful discussions, Marta Schmitt, Benoit Robert, Barbara Brandani, Amy Pittet and Kevin Gross for technical assistance. Sophie Martin and Sara Mitri for critically reading the manuscript. Fabian Rudolf at the ETHZ for generating the UbiY destabilization sequence. We thank Amy Keating at MIT for sharing some SynZip plasmids. Work in the Pelet lab is supported by the Swiss National Science Foundation, SystemsX.ch and the University of Lausanne.

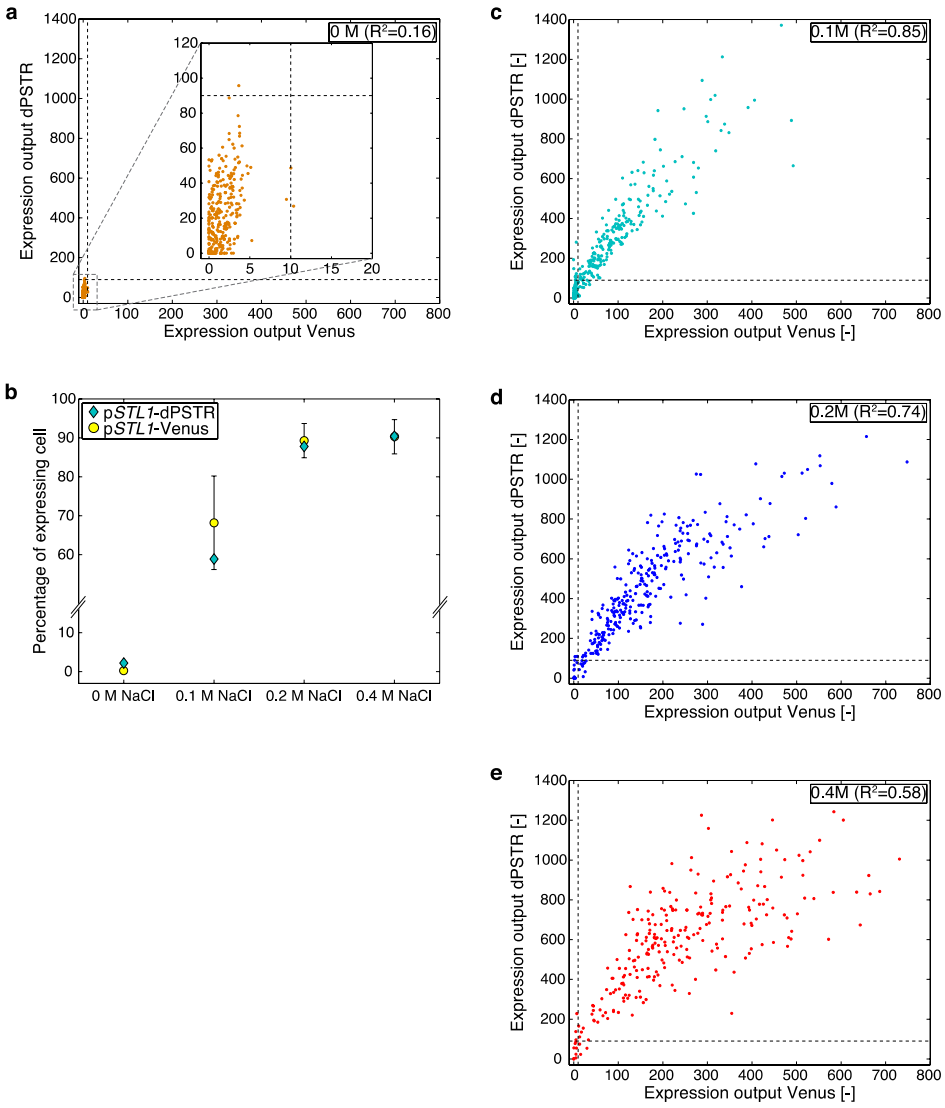
Supplementary materials



Supplementary Figure 1: Control experiments for the dPSTR.
a. Stimulation with 0.2 M NaCl of two different clones carrying a p*STL1*-dPSTR^{R2,3} containing the SynZips SZ2 and SZ3, which are not interacting. The non-functional p*STL1*-dPSTR^{R2,3} does not display any nuclear enrichment. For comparison, the functional p*STL1*-dPSTR^{R2,1} (yDA134) is plotted with a dashed blue line.
b. Stimulation with 0.4 M NaCl of a strain carrying a p*GAL1*-dPSTR^{R}, which is not expressing upon hyperosmotic shock, and does not display any nuclear enrichment. For comparison, the p*STL1*-dPSTR^{R} (yDA134) is plotted with a dashed red line.
c. Comparison of p*STL1* expression induced by 0.2 M NaCl measured by flow cytometry (p*STL1*-qVenus, cyan, N_C~10'000) or microscopy for Venus expression or p*STL1*-dPSTR^{R} relocation (blue and red respectively, data from **Fig. 1**). Note the close overlap between the rise of transcription quantified with flow cytometry or the dPSTR. To allow for a direct comparison of the three methods, the fluorescent values measured are normalized.



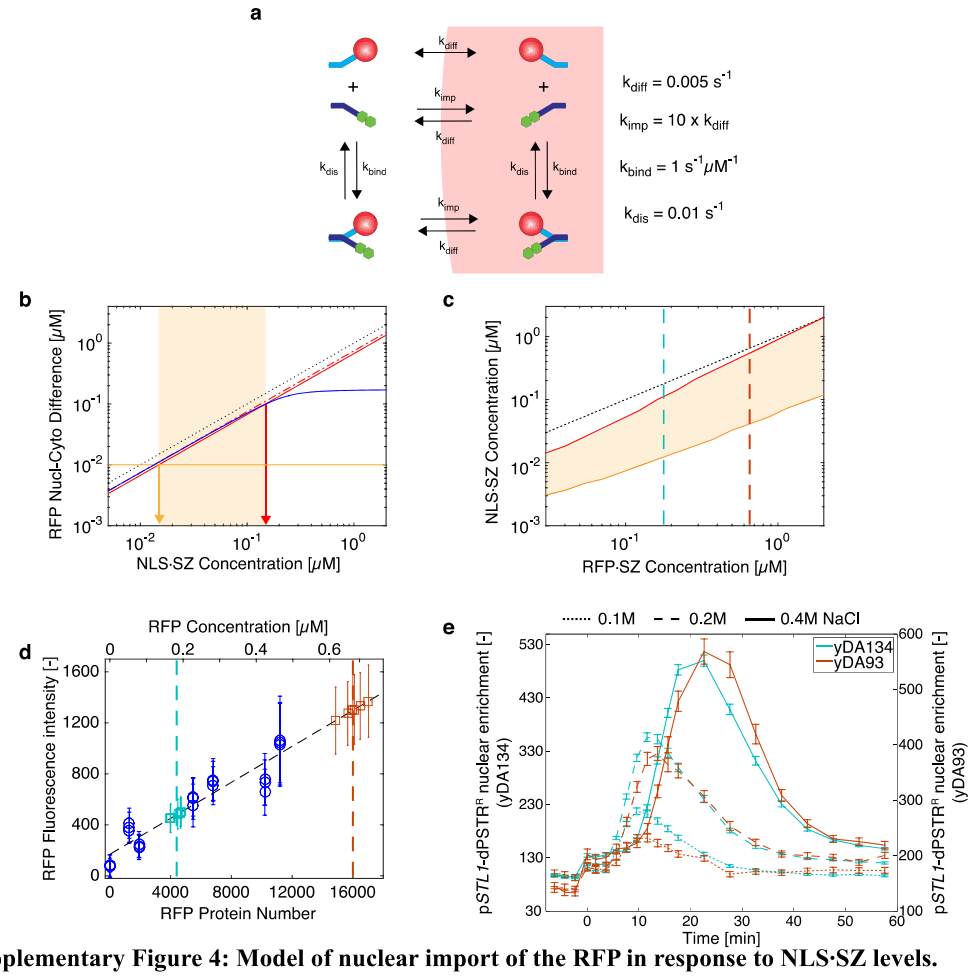
Supplementary Figure 2: Image segmentation process using YeastQuant
a. - b. - c. Microscopy images used for the segmentation: histone tag CFP (**a**), bright field image in the focal plane (**b**), and out of focus bright field image (**c**, z=-2.5µm).
d. - e. - f. Different objects are defined by the segmentation process. First, the Nucleus is characterized using the CFP image (**d**). The two bright field images are used to find the cell contour and define the Cell object (**e**). Then, the Nucleus object, enlarged by 2 pixels, is subtracted to the Cell object, to define the Cytoplasm object (**f**).



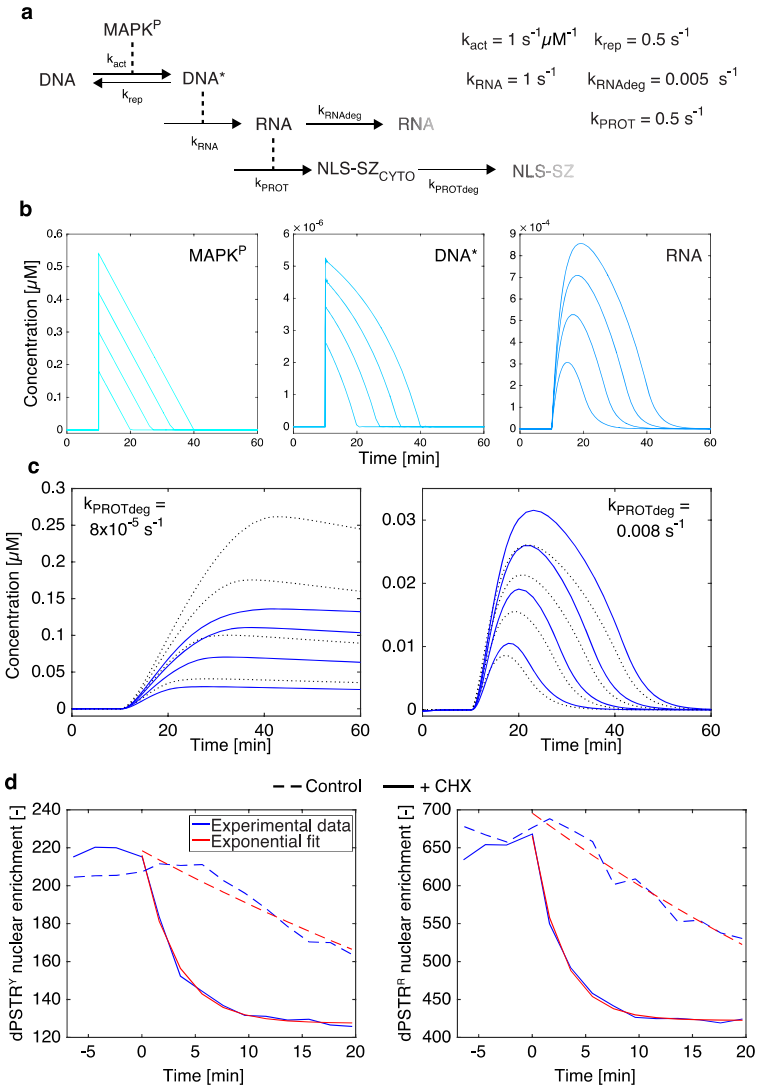
Supplementary Figure 3: Comparison of the level of expression and the dynamics of expression measured by the dPSTR and by the Venus fluorescence apparition.

a. - c. -d. - e. Single cell correlation of the expression output measured by the dPSTR or the Venus, for the non-induced control 0 M NaCl (**d**) and each concentration: 0.1 M NaCl (**c**), 0.2 M NaCl (**d**), 0.4 M NaCl (**e**). The dashed lines represent the thresholds for expression.

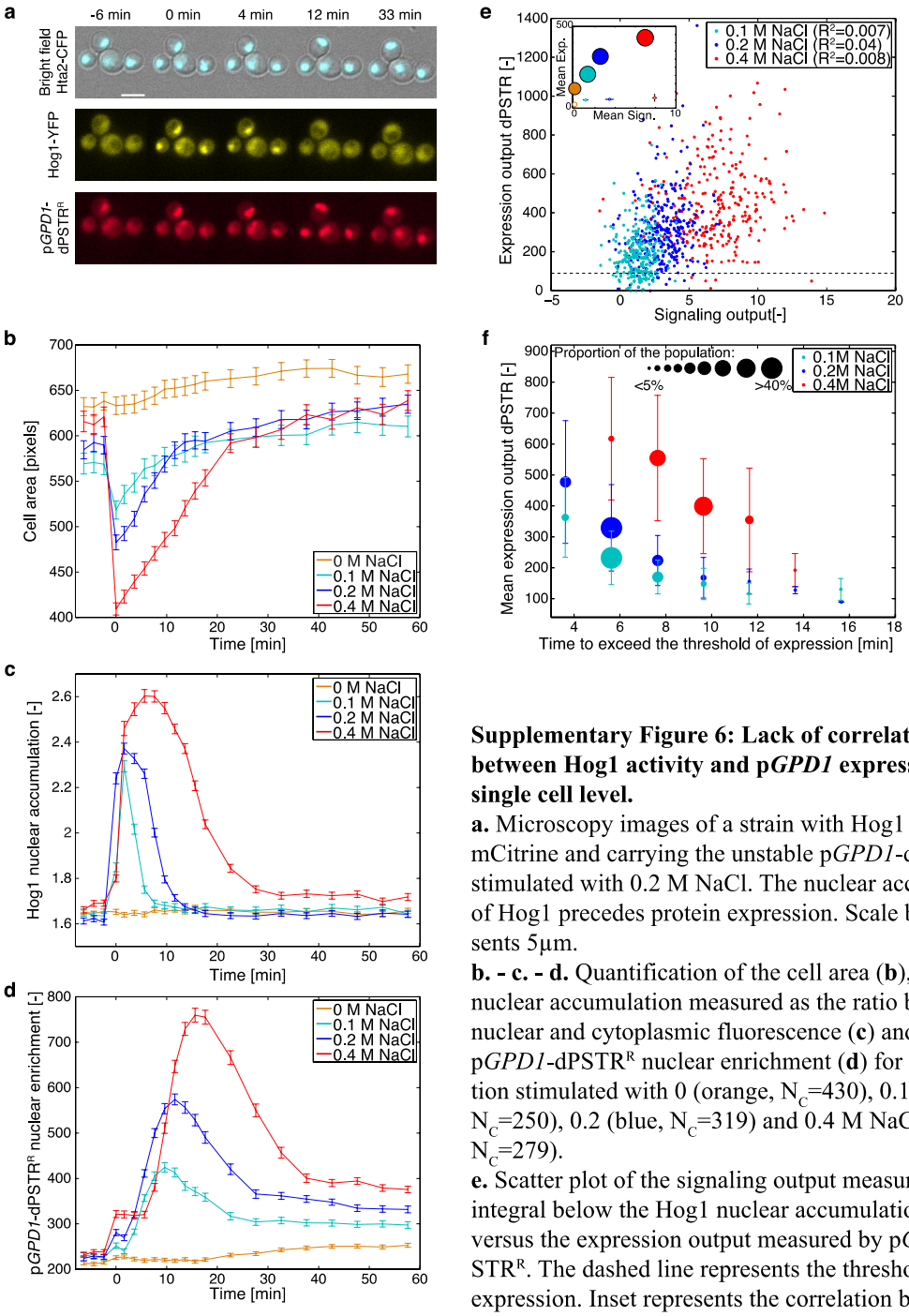
b. Percentage of expressing cells measured by the dPSTR or by the promoter-FP method, for each concentration. Cells having an expression output above the expression threshold plotted in **a**, c, **d** and **e** are defined as expressing cells. Note that the proportions are similar using the two methods. Each dot represents the mean of three experiments and the error bars, the standard deviations.



Supplementary Figure 4: Model of nuclear import of the RFP in response to NLS·SZ levels.
a. Schematic of the reactions implemented in the model and their corresponding rate constants. See Supplementary Note for further details.
b. Difference between nuclear and cytoplasmic RFP concentrations expected for a fixed total RFP·SZ concentration (0.2 µM) and a range of NLS·SZ concentration (blue line). The dotted black line indicates the total NLS·SZ concentration. At low concentrations, there is a linear relationship between NLS·SZ level and nuclear minus cytoplasmic RFP concentration (red dash-dotted line). At higher concentration, the nuclear enrichment of the RFP saturates since an increasingly higher fraction of the RFP is already localized in the nucleus. If we tolerate 10% deviation from linearity (solid red curve), we obtain an upper limit for the concentration of NLS·SZ that can be measured (red arrow). The sensitivity of the detection allows to observe a 10% increase in RFP nuclear enrichment, thereby setting a lower limit to the concentration of NLS·SZ that can be measured (orange arrow). Therefore, the shaded area represents the sensitivity range of the assay for the selected RFP·SZ expression level.
c. The lower and upper limits of sensitivity of the system are calculated for a range of RFP·SZ concentrations. An NLS·SZ sensitivity window is depicted with a shaded area bound by lower and higher limits (orange and red lines, respectively). The dotted black line indicates the case where NLS·SZ = RFP·SZ.
d. Calibration of the protein number and concentration of RFP·SZ in the reporter strains (yDA134, blue squares and yDA93 orange, 6 biological replicates) relative to endogenously tagged proteins with different expression levels (blue circles, three biological replicates, see Supplementary Table 4). The mean and standard deviation of more than 500 single cells is plotted and fitted by linear regression (dashed black line). The expression level of RFP is found to be 4400±300 protein per cell or 0.18µM for yDA134 and 16'000±750 protein per cell or 0.66µM for yDA93 (dashed line in panel c and d).
e. Normalized nuclear enrichment in course of time for yDA134 (blue) and yDA93 (orange) stimulated with 0.1 (resp. N_C=558 and N_C=341), 0.2 (resp. N_C=655 and N_C=297), or 0.4M NaCl (resp. N_C=802 and N_C=268). Note the good overlap between the responses of the two strains at 0.2 and 0.4M NaCl and the lower sensitivity at 0.1 M.



Supplementary Figure 5: Model of NLS·SZ synthesis in presence or absence of degradation
a. Schematic of a simple expression model based on mass-action kinetics for the synthesis of the NLS·SZ. The NLS·SZ is produced in the cytoplasm and enters in equilibrium with the RFP·SZ as described in Supplementary Figure 4. See also Suplementary Note.
b. Dynamics of the MAPK^P which is the input to the model and the resulting gene activation DNA* and RNA production.
c. The dynamics of total NLS·SZ production (dotted line) is compared to the RFP nuclear-cytoplasmic enrichment (solid line) in the case where the NLS·SZ is stable or if it has a half-life of 2 min. From this model, it can be observed that the dynamics of RFP nuclear enrichment can be limited by two reactions, the formation of the complex between NLS·SZ and the RFP·SZ and the import of the NLS-SZ in the nucleus. Both of these reactions happen with fast dynamics thus allowing a monitoring of the protein production in the sub-minute time scale. The deactivation of the system depends on three reactions, the repression of the active gene, the degradation of the mRNA and the degradation of the NLS-SZ. In our experimental data, the return of the dPSTR nuclear enrichment to pre-stimulus level takes place on the order of 10 minutes suggesting that the NLS·SZ degradation is not the limitting factor for this process.
d. Characterization of destabilization sequence half-life. Cells bearing the p*GPD1*-dPSTR^Y and p*GPD1*-dPSTR^R were grown overnight in SD+1M sorbitol to increase the basal level of p*GPD1* expression (note the differences in nuclear enrichment before time 0 compared to Fig. 5b). Cells were treated with cycloheximide (solid line, N_C = 380) or not (dashed line, N_C = 383). A half-life for the UbiY-NLS-SZ is estimated to 2.1±0.5 min (3 experiments, mean and standard deviation of the RFP and YFP dynamics) based on an exponential fit of the data (red line). (One replicate is shown)



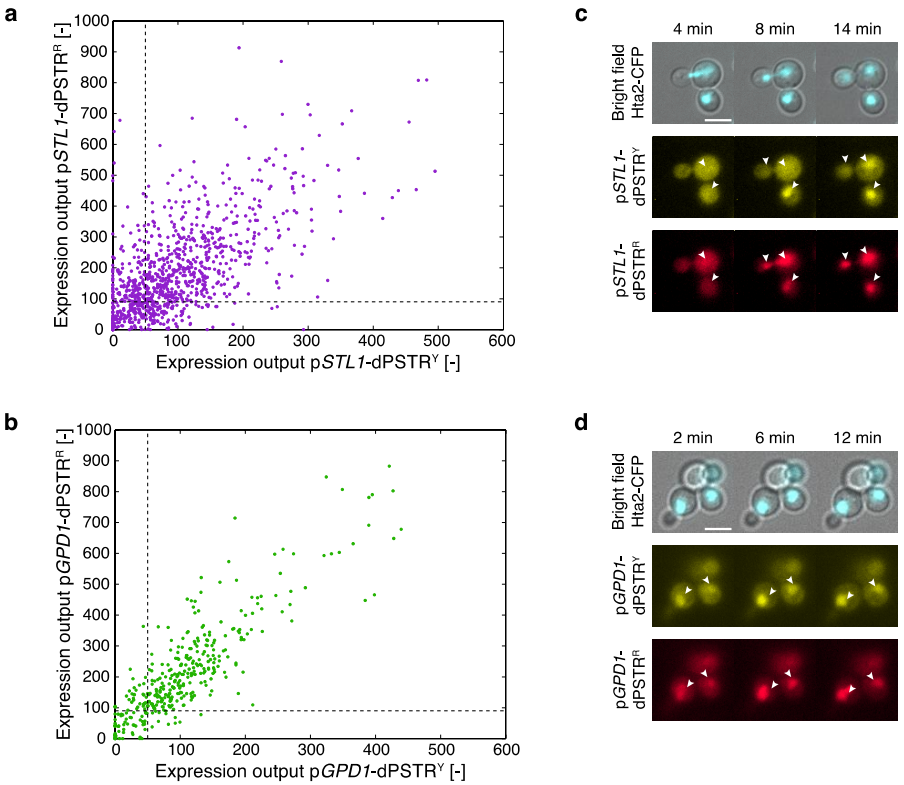
Supplementary Figure 6: Lack of correlation between Hog1 activity and p*GPD1* expression at the single cell level.

a. Microscopy images of a strain with Hog1 tagged with mCitrine and carrying the unstable p*GPD1*-dPSTR^R stimulated with 0.2 M NaCl. The nuclear accumulation of Hog1 precedes protein expression. Scale bar represents 5µm.

b. - c. - d. Quantification of the cell area (**b**), Hog1 nuclear accumulation measured as the ratio between nuclear and cytoplasmic fluorescence (**c**) and p*GPD1*-dPSTR^R nuclear enrichment (**d**) for cell population stimulated with 0 (orange, N_C=430), 0.1 (cyan, N_C=250), 0.2 (blue, N_C=319) and 0.4 M NaCl (red, N_C=279).

e. Scatter plot of the signaling output measured as the integral below the Hog1 nuclear accumulation curve versus the expression output measured by p*GPD1*-dPSTR^R. The dashed line represents the threshold for expression. Inset represents the correlation between the mean signaling output (x axis) and the mean expression output (y axis) for the expressing cells (filled circles, R^2=0.91) and the non expressing cells (empty circles, R^2=0.66). The size of the marker is indicative of the percentage of cells in each category.

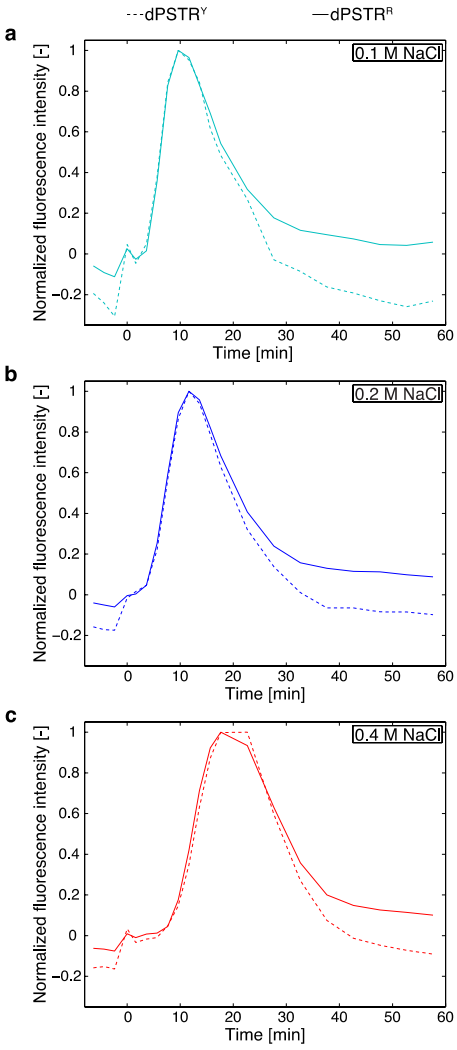
f. Correlation between the time needed to overcome the expression threshold and the expression output. The mean expression output and the standard deviation were calculated for groups of cells which exceed the expression threshold at the same time point. The marker size is indicative of the percentage of cells (from the total population) in each group. Note that the marker size is bigger than in Figure 3, indicating that a higher proportion of the population expresses p*GPD1* compared to p*STL1*.



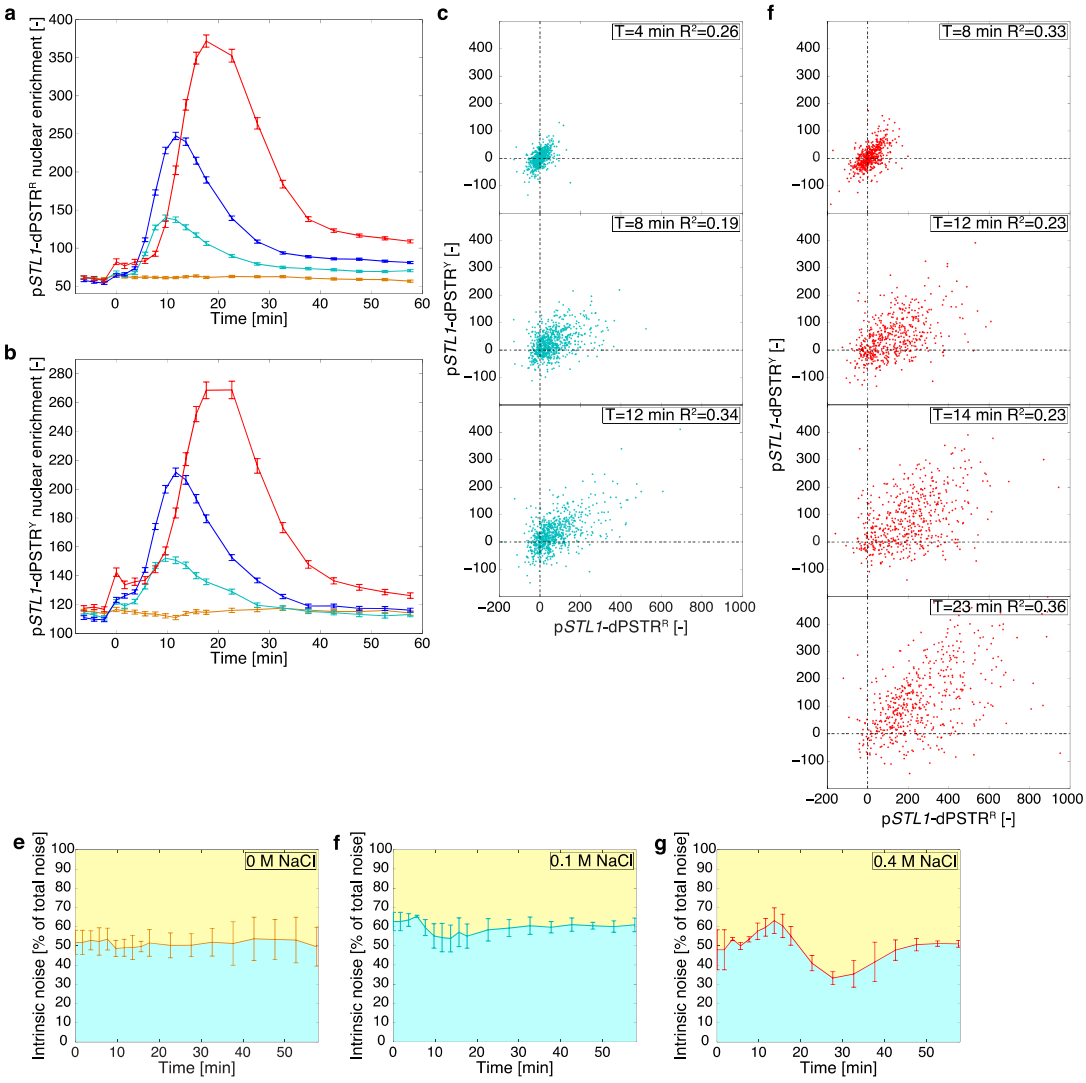
Supplementary Figure 7: Expression output of cells containing double p*STL1* or p*GPD1* dPSTRs.

a. - b. Expression output measured for cells from **Figure 5** expressing two p*STL1*-dPSTRs (**a**) or two p*GPD1*-dPSTRs (**b**).

c. - d. Microscopy images of cells bearing an Hta2-CFP and expressing two p*STL1*-dPSTRs (**c**) or two p*GPD1*-dPSTRs (**d**). The white arrowheads show nuclei in the focal plane. Scale bars represent 5µm.



Supplementary Figure 8: Control experiments for the p*STL1*-dPSTR^{R} and p*STL1*-dPSTR^{Y} sensors. **a. - b. - c.** Normalized traces of p*STL1*-dPSTR^{Y} (dashed line) and p*STL1*-dPSTR^{R} (filled line) for cells stimulated with 0.1 M NaCl (**a**), 0.2 M NaCl (**b**) and 0.4 M NaCl (**c**). The weaker overlap observed at later time-points could be due to a differential bleaching of the two fluorescent proteins.



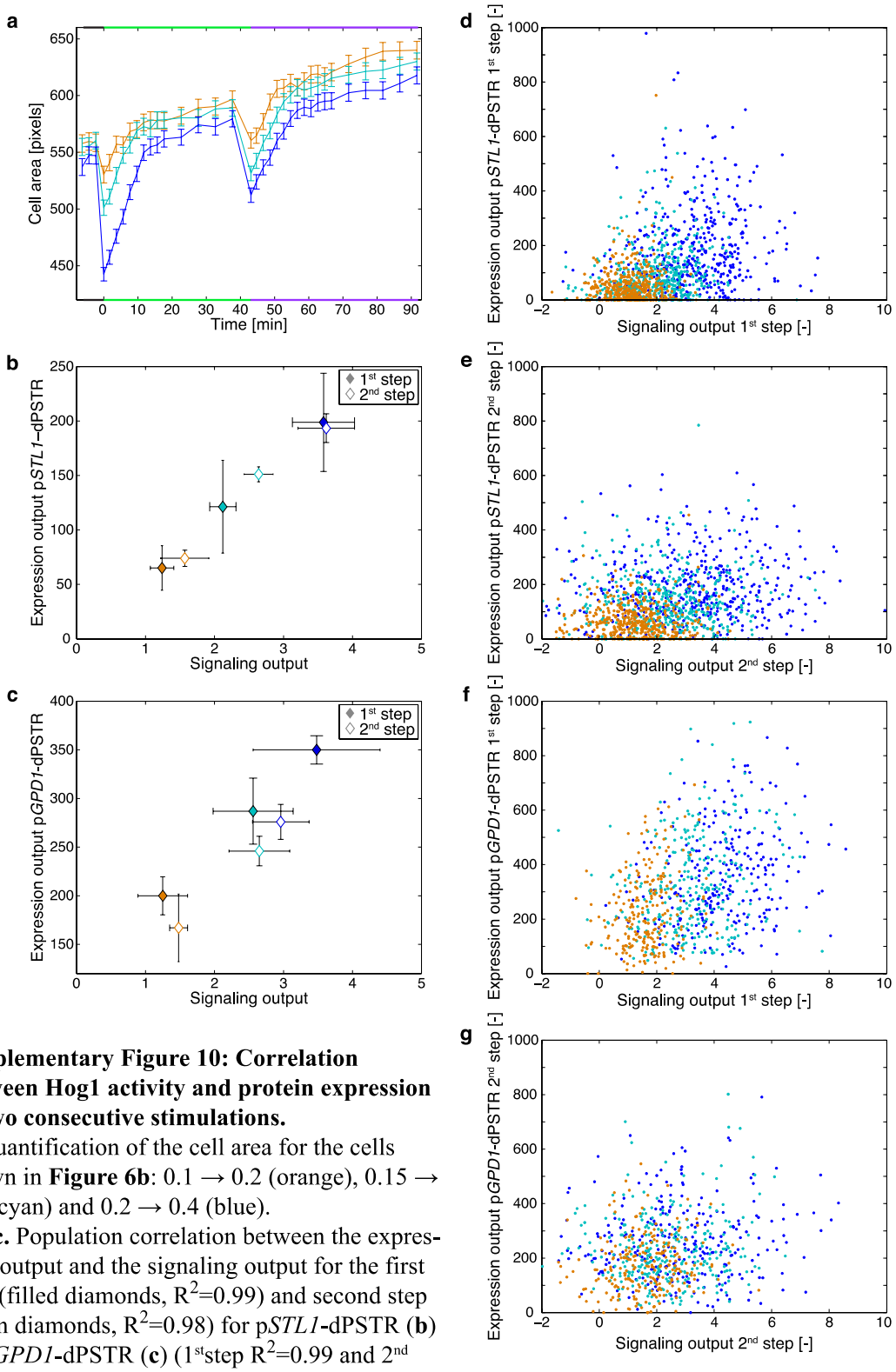
Supplementary Figure 9: Dynamic quantification of the expression noise of p*STL1* with different stimuli.

a. - b. Quantification of p*STL1*-dPSTR^{R} (**a**) and p*STL1*-dPSTR^{Y} nuclear enrichment (**b**) for cell population stimulated with 0 (orange, N_C=670), 0.1 (cyan, N_C=765), 0.2 (blue, N_C=958) and 0.4 M NaCl (red, N_C=530). The 0.2M induction was used in **Figure 5**.

c. Single cell correlation between the corrected nuclear enrichment of each p*STL1*-dPSTR nuclear enrichment at given times, for cells stimulated with 0.1M NaCl: T=4 min, T=8 min, T=12 min.

d. Same as above for cells stimulated with 0.4M NaCl at times: T=8 min, T=12 min, T=14 min, T=23 min.

e. - f. - g. Evolution of the intrinsic noise for unstimulated cells (**e**) or cell stimulated with 0.1 M NaCl (**f**) or 0.4 M NaCl (**k**). The mean of 3 biological replicate with the SD is ploted. Note the transient rise in intrinsic noise observed at the onset of gene expression.



Supplementary Figure 10: Correlation between Hog1 activity and protein expression in two consecutive stimulations.

a. Quantification of the cell area for the cells shown in **Figure 6b**: 0.1 → 0.2 (orange), 0.15 → 0.3 (cyan) and 0.2 → 0.4 (blue).

b. - c. Population correlation between the expression output and the signaling output for the first step (filled diamonds, R^2=0.99) and second step (open diamonds, R^2=0.98) for p*STL1*-dPSTR (**b**) or p*GPD1*-dPSTR (**c**) (1st step R^2=0.99 and 2nd step R^2=0.99) . Each point represents the average of three biological replicates. Error bars are the standard deviations.

d. - g. Single-cell correlation between the expression output and the signaling output for the 1st step (**d** and **f,** data from **Figure 6b** and **c**) and the 2nd step (**e** and **g,** data from **Figure 6d**). For all the correlations, $R^2<10^{-3}$.

Supplementary Note

dPSTR relocation model

We have used a simplified model to estimate the amount of nuclear enrichment of the fluorescent protein as function of the expression of the inducible peptide (Supplementary Figure 4). The model contains three species: the fluorescent protein linked to the SynZip (FP·SZ), the inducible peptide (NLS·SZ) and the complex formed between those two proteins via the SynZip (FP·SZ·NLS). The K_d of the complex formed has been estimated to 10µM (Reinke et al., 2010; Thompson et al., 2012). These three species are partitioned between two compartments: the nucleus and the cytoplasm. We estimate a passive diffusion of the FP·SZ into and out of the nucleus to be 0.005 s^{-1} (Timney et al., 2006). Based on the level of nuclear enrichment of the Venus·SZ (Figure 2.1d), we estimate that the double NLS present on the NLS·SZ peptide allows at least a 10-fold enrichment of the proteins in the nucleus. Using these parameters, we have simulated the steady-state nuclear and cytoplasmic concentrations for each component for a range of total NLS·SZ and FP·SZ. For each condition, the nuclear to cytoplasmic difference is calculated as:

$$\Delta Nucl - Cyto = ([FP \cdot SZ_N] + [FP \cdot SZ \cdot NLS_N]) - ([FP \cdot SZ_C] + [FP \cdot SZ \cdot NLS_C])$$

Figure 2.S4b displays the outcome of the model for a concentration of 0.2 µM of RFP·SZ. For low concentrations of the NLS·SZ, there is a linear correlation between nuclear enrichment of the RFP and the NLS·SZ concentration. However, there is a limit to the detection ability of the microscope and we estimate that only a 10% enrichment can be reliably detected at the single cell level (~10^{-2} µM), thereby setting a lower limit to the NLS·SZ that can be measured with the dPSTR (1.5x10^{-2} µM). If the NLS·SZ concentration reaches too high levels, the linear relationship with RFP nuclear enrichment is lost, as can be seen from the saturation of the blue curve in Supplementary Figure 4b. Allowing for 10% error in linearity between NLS·SZ level and nuclear enrichment measurement results in an upper limit in NLS·SZ that can be reached before saturation (1.6x10^{-1} µM). These upper and lower limits were calculated for a range of RFP·SZ concentrations (Supplementary Figure 4c).

In parallel to this modeling effort, we have quantified the level of RFP·SZ expression in the dPSTR strain. We measured a calibration curve based on the fluorescence intensity of mCherry tagged proteins with known expression levels (Ghaemmaghami et al., 2003; Wu and Pollard, 2005; Zechner et al., 2014). After linear regression between protein numbers and fluorescent intensities, we can get a good estimate of the protein number of the constitutively expressed fluorescent protein in the dPSTR.

At 4 400 proteins per cell or 0.18 µM (for a 40 fl volume (Jorgensen et al., 2002; Tyson et al., 1979)), we can estimate the low and high limits of expression that can be quantified with the dPSTR between 1.2·10^{-2} to 1.2·10^{-1} µM of synthesized NLS·SZ. For comparison, we also quantified the fluorescence of another dPSTR strain with higher expression levels (16'000 prot. per cell, yDA93). The concentration range that can be quantified varies from 3.9·10^{-2} to 6·10^{-1} µM. These two strains have different sensitivity windows. As predicted by the model, we observe with this strain a lower sensitivity at detecting protein expression from the p*STL1* promoter at 0.1 M NaCl (Supplementary Figure 4e).

Model of NLS·SZ synthesis and degradation

We further developed this model to include a minimal transcriptional model in order to simulate the dynamics and level of nuclear import with a stable and unstable dPSTR (Supplementary Figure 5). The reactions and rate constants used in the model are presented in Supplementary Figure 2.5a. Briefly, the input to the model is the level of activated MAPK

($MAPK^P$), which increases after 10 min of simulation and declines in a linear fashion to reach zero after a time that is function of the level of activity upon stimulation. This mimics the different temporal windows of Hog1 activity upon various hyper-osmotic stresses. The active kinase turns the promoter into an active state. From this state, mRNA can be produced, which ultimately leads to the expression of the NSL·SZ. This peptide is produced in the cytoplasm, and it can bind to the FP·SZ and accumulate in the nucleus following the same reactions as described above. Once kinase activity returns to pre-stress levels, the promoting region of the DNA returns to its off-state, the mRNA molecules are degraded and the peptide NLS·SZ will be degraded as well.

To test the influence of the destabilizing sequence on the output of the model, we have selected two variants, the first one where the NLS·SZ is stable, and the second one where the NLS·SZ is actively degraded, which increases the rate of this reaction by 100 fold (Supplementary Figure 5c). To estimate the stability of the UbiY-NLS-SZ in cells, we have performed an experiment where protein transcription is blocked by cycloheximide and the degradation kinetics of the protein can thus be quantified. A strain bearing the p*GPD1*-$dPSTR^R$ and p*GPD1*-$dPSTR^Y$ was grown overnight in SD-full containing 1M Sorbitol, diluted in the morning with the same medium, placed under the microscope and treated with 0.1mg/ml cycloheximide (Supplementary Figure 5d). In this high osmolarity medium, the basal expression level of p*GPD1* is increased due to the higher basal activity of Hog1. We measured a half-life of 2.1 ±0.5 min for the UbiY-NLS-SZ peptide which was used in the model. As expected and in agreement with our experimental measurements of the dPSTR, due to accumulation of all NLS·SZ expressed in the cell, the stable inducible peptide can lead to more saturation than the degraded one (Supplementary Figure 5c). However, the degraded peptide will be expressed at lower levels and thus nuclear enrichment of the fluorescent protein might be more difficult to detect at low concentrations. Therefore, the level of expression of the RFP·SZ might have to be adapted to accommodate the transcriptional output of the inducible promoter. We have achieved this by testing different combinations of ribosomal protein gene promoters, which span a large range of expression levels (Knight et al., 2014) and different terminators (Yamanishi et al., 2013). Alternatively, for promoters with very low expression outputs, we could envision to multiply the SZ motif on the inducible protein, such that each expressed peptide would induce the recruitment of two or four fluorescent proteins.

Chapter 3

Timeline of the mating-induced transcription

Preamble

This chapter is part of a research article under revision at *Molecular Systems Biology* (when this version is printed in November 2017), that is also published on BioRxiv (https://www.biorxiv.org/content/early/2017/10/02/197103) (https://doi.org/10.1101/197103).

Contributions:
Marta Schmitt and Delphine Aymoz built plasmids and strains.
Delphine Aymoz performed the time-lapse experiments, designed the experiments and wrote the manuscript.
Serge Pelet designed the experiments and performed experiments.

Abstract

The budding yeast mating response is often considered as a typical example of cell-fate decision. Upon sensing their presence via pheromone communication, two haploid cells of opposing mating type can decide to mate in order to form a diploid zygote. To do so, they activate signaling through a MAPK pathway, which culminates in the activation of the MAPK, Fus3. The activation of the kinase results in cell cycle arrest, polarized growth towards the pheromone gradient and the partner, and transcription of hundreds of target genes. While all the actors of the mating response are well known, very little is understood on the kinetics at which the signaling and resulting transcription occur. In this Chapter, we dynamically measure the expression of 14 mating-induced genes in single cells. We report that the mating genes induction occurs at various kinetics following pheromone sensing. Furthermore, we found that pheromone concentration is one of the determinants of their time of induction.

Introduction

The mating response of the budding yeast has been widely used as a model for cell-fate decision-making, where haploid cells can decide to continue their vegetative growth, or find a partner and form a diploid cell (Atay and Skotheim, 2017; Bardwell, 2005). This decision-making process requires the activation of the pheromone MAPK pathway, which is activated in presence of the pheromone produced by the opposing mating type. The binding of α-factor to the membranar G-protein coupled receptor Ste2 will drive a series of events leading to the activation of the MAPK pathway, formed by the MAP3K Ste11, the MAP2K Ste7, and the MAPK Fus3, all recruited to the membrane via the scaffold Ste5 (Bardwell, 2005). As Ste11 and Ste7 are both shared with the filamentous growth (FG) pathway, the FG MAPK Kss1 is also activated by pheromone sensing (Gartner et al., 1992; Ma et al., 1995). Fus3 phosphorylates numerous targets. Among them, Far1 is a CDK inhibitor, which drives the cell cycle arrest in G1 and allows correct establishment of the polarization site towards the partner (Atay and Skotheim, 2017; Bardwell, 2005; Elion et al., 1993; Gartner et al., 1998; Peter et al., 1993; Peter and Herskowitz, 1994). The active MAPK will also promote the induction of the mating transcriptional program, via the phosphorylation of the inhibitors of the transcription factor (TF) Ste12 (Cook et al., 1996; Tedford et al., 1997).

Target genes of the mating response

Pheromone-induced genes have been identified through microarray experiments in different conditions of stimulation and genetic backgrounds (Roberts et al., 2000). It has been found that more than 200 genes were upregulated by pheromone stimulation, in a signaling-dependent manner, and in a Ste12-dependent manner. In addition, more than 200 genes were shown to be downregulated by the pheromone stimulation, in a Far1-dependent manner. This downregulation can be attributed the cell cycle arrest that it controls. Among downregulated targets, genes that are involved in cell cycle progression were identified, as well as genes involved in DNA replication, budding or mitosis. This is consistent with the fact that when cells commit to the mating, they need to stop cell cycle progression, and hence will not divide.

Among genes upregulated by the mating response, numerous protein with well-identified functions can be cited. First of all, Ste12 is itself induced by pheromone response. In addition to this, a large number of the proteins involved in the signaling in response to mating are induced (Fus3, Far1, Ste2, the pheromone receptor), as well as negative regulators of the mating pathway (Sst2, which downregulates the signaling, and Bar1, which degrades the pheromone). Proteins involved in the fusion process are induced, like Fus1, Fig1, Prm1, which were all widely used as mating transcription reporters. The Kar proteins are induced, and are implicated in the karyogamy process, the fusion of the nuclei of the two partners.

Ste12 is involved in cell cycle dependent transcription

It was shown that if Ste12 is required for pheromone-induced transcription of mating genes, it is also required for their induction during the cell cycle. Indeed, some mating genes like *FAR1*, *AGA1*, *FUS1* or *STE12* itself, have been shown to be induced in a cell cycle controlled manner (Oehlen et al., 1996). This basal transcription is dependent on Ste12 presence. The genes have a maximal induction in G1, and low mRNA levels in S phase. Interestingly, this pattern corresponds to the cell cycle stages at which cells are respectively pheromone permissive (G1) or restrictive (S) (Durandau et al., 2015).

While the mating-induced genes have been identified, only few details about the kinetics of their induction in response to pheromone signaling are known. In this Chapter, we describe

the induction of 14 mating-induced genes in response to pheromone signaling either in exogenous system or in mating mixture. We will see that the mating-induced transcription response is temporally controlled.

Results

Combinatorial quantification of mating signaling and mating-genes expression

In presence of pheromone of the opposing mating type, yeast cells will activate the two MAPKs Fus3 and Kss1, after signaling through a MAP Kinase pathway. To quantify this signaling activity in real-time and at the single-cell level, we developed in our lab a kinase activity sensor (Colman-Lerner et al., 2005; Durandau et al., 2015). The Synthetic Kinase Activity Relocation Sensor (SKARS) is based on the relocation of a fluorescent protein from the nucleus to the cytoplasm of the cell, upon specific phosphorylation of the sensor by the kinase of interest. The Ste7 docking site promotes specificity towards the mating MAPKs, and allows the activated kinases to phosphorylate the Nuclear Localization Signal (NLS) of the fluorescent sensor (Remenyi et al., 2005). This phosphorylation will inactivate the NLS, leading to a decrease in SKARS active nuclear import. The ratio of cytoplasmic-to-nuclear fluorescence is the readout of the sensor, and corresponds to the kinase activity. Following pheromone stimulation, the SKARS exits the nucleus within minutes, meaning that the kinases Fus3 and Kss1 are activated rapidly upon pheromone sensing. This result was previously shown by Western Blot quantifications or the SKARS itself (Durandau et al., 2015; Nagiec and Dohlman, 2012; Yu et al., 2008).

The pheromone signaling will lead to a rewiring of the transcriptional program, with more than 200 genes induced. Among them, *FIG1* has been identified by microarray experiments as one of the most induced. Using the dPSTR system, we quantified its expression kinetics, in cells that also carry the SKARS. In the two cells shown in Figure 3.1A, the nuclear enrichment of the p*FIG1*-dPSTR^{R} occurs 45 minutes after the nuclear export of the SKARS, meaning that the expression of p*FIG1* is detectable 45 minutes after kinase activation. Using the automated platform YeastQuant, we can quantify the nuclear enrichment of both sensors for the duration of the experiment in hundreds of cells (Figure 3.1B; (Pelet et al., 2012); see Methods; see Sup. Figure 2.1, in Chapter 2).By doing so, we can see that the kinase is activated

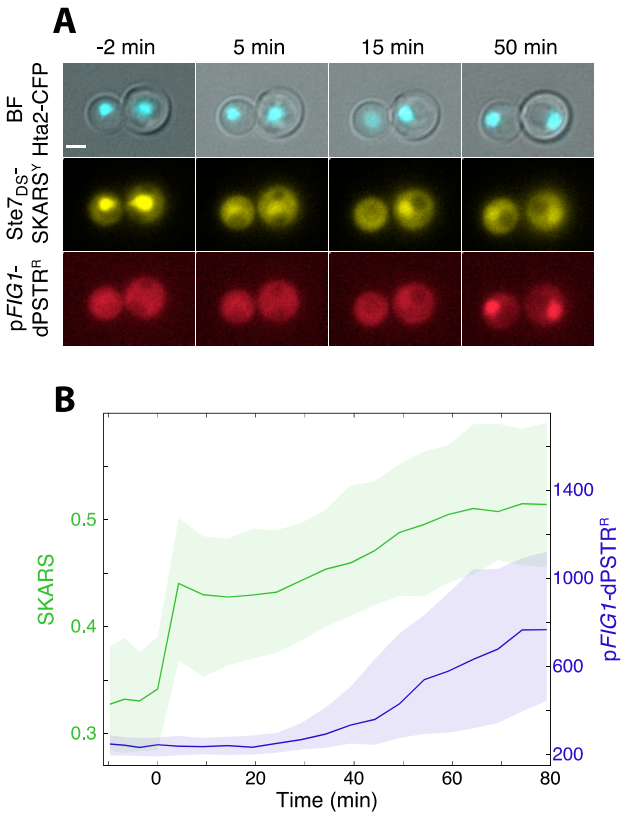


Figure 3.1 Interplay between kinase activity and p*FIG1* induction in the mating pathway

A. Microscopy images of cells stimulated with a saturating pheromone concentration (1µM) at time 0min. The cells bear a histone tagged with CFP, a yellow SKARS reporting on Fus3 and Kss1 activities, and a red dPSTR reporting on p*FIG1* induction.
Scale bar 2.5µm.

B. Quantifications of the SKARS (green, left axis), measured by the ratio of cytoplasmic to nuclear YFP, and of the p*FIG1*, measured by the difference between nuclear and cytoplasmic fluorescence of the dPSTR (blue, right axis).

For all similar graphs, the solid line is the median response and the shaded area represents the 25 to 75 percentiles of the population.

within the 10 minutes following stimulation, and that this activation is increasing throughout the experiment. The induction of the promoter p*FIG1* is occurring with a delay of 15 to 30 minutes compared to the kinase activation. This result is quite surprising considering that in response to hyperosmotic stress, cells are able to initiate gene expression within the 2 minutes following the stress and after MAPK activation (Chapter 2; (Aymoz et al., 2016)). This result prompted us to quantify the expression kinetics of other mating-induced genes, including *AGA1* (Chou et al., 2006; Roy et al., 1991). As we can see in Figure 3.2, this promoter is induced 10 minutes after kinase activation, when looking at the population response. As single cells can display a large variability in their signaling capacity, we quantified p*FIG1* and p*AGA1* induction only in cells that activated the MAPKs within the 10 minutes following stimulation (Supplementary Figure 1) (Bendezu et al., 2015; Durandau et al., 2015). Even by choosing fast responding cells, there is a large variability in the induction kinetics of p*FIG1*, ranging from 15 to 70 minutes, whereas the induction of p*AGA1* is tighter, from 15 to 30 minutes. This suggests that there is no temporal correlation between the signaling activity and the resulting transcriptional response.

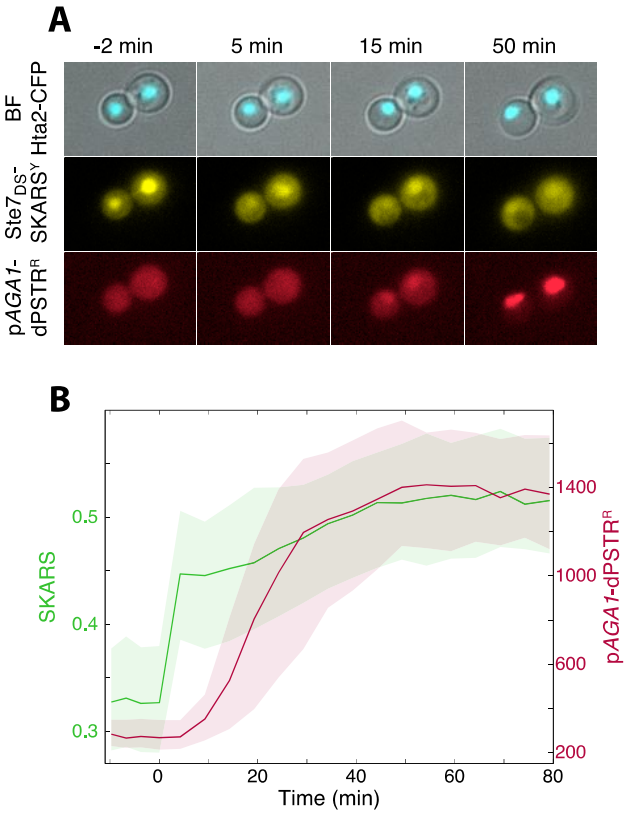


Figure 3.2 Interplay between kinase activity and p*AGA1* induction in the mating pathway

A. Microscopy images of cells stimulated with a saturating pheromone concentration (1µM) at time 0min. The cells bear a histone tagged with CFP, a yellow SKARS reporting on Fus3 and Kss1 activities, and a red dPSTR reporting on p*AGA1* induction.
Scale bar 2.5µm.

B. Quantifications of the SKARS (green, left axis), measured by the ratio of cytoplasmic to nuclear YFP, and of the p*AGA1*, measured by the difference between nuclear and cytoplasmic fluorescence of the dPSTR (red, right axis).

In order to confirm these results, we compared the induction in the two promoters within the same cells using the two colors of the dPSTRs (Figure 3.3). We can see that in the two cells from the picture, the induction of p*AGA1* occurs 25 minutes before the expression of p*FIG1*.

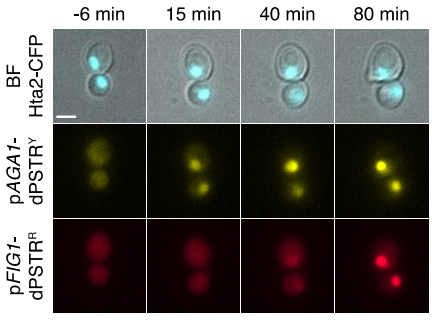


Figure 3.3 Microscopy images of a strain carrying p*FIG1*-dPSTR^{R} and p*AGA1* dPSTR^{Y}

Microscopy images of cells stimulated with a saturating pheromone concentration (1µM) at time 0min. The cells bear a histone tagged with CFP, a red dPSTR reporting on p*FIG1* induction, and a yellow dPSTR reporting on p*AGA1* induction.
Scale bar 2.5µm.

We then quantified for both promoters the response time (Figure 3.4). This metrics represents the time at which all expressing cells are overcoming 20% of their maximal expression (see Methods; Supplementary Figure 2). We found that the histograms of response time are very different between the two promoters. For p*AGA1*, there is a tight distribution centered around 15 minutes, with 83% of the cells inducing the promoter within the first 30 minutes. For p*FIG1*, most of the expressing cells have a response time between 20 and 45 minutes, but this is highly variable from cell to cell. We then quantified the difference between the response times of the two promoters within the same single-cells (Figure 3.4, inset). We can see that 87% of cells expressing both promoters are inducing p*AGA1* before p*FIG1*, with an average delay of 23 minutes. To look at the expression at the single cell level in a more dynamic way, we correlated the nuclear enrichment of both dPSTRs at different times following stimulation, in all single cells (Figure 3.5). Before stimulation, the correlation of the two dPSTRs is random, but 10 minutes after induction, p*AGA1* starts to be induced in some of the cells, displacing the cloud toward the right. 35 minutes after stimulation, a large fraction of the cells are expressing p*AGA1* to higher levels, and some of the cells are starting to induce p*FIG1*, displacing the cloud in the upper direction. 60 minutes after stimulation, most of the cells express both promoters. There is no correlation between the levels of induction of the two promoters. Overall, we showed that although signaling occurs quickly in response to pheromone stimulation, two promoters induced by this signal can be expressed with different dynamics.

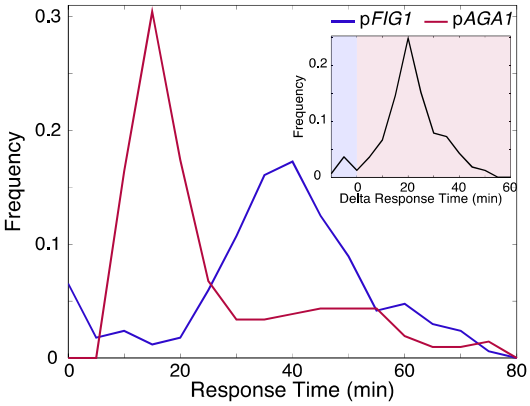


Figure 3.4 Quantification of response times of p*FIG1* and p*AGA1*

The response time has been quantified for p*AGA1* (red) or p*FIG1* (blue) in all cells expressing each promoter. The inset is the difference between the response time of p*AGA1*-dPSTR^{Y} and p*FIG1*-dPSTR^{R}, for all cells expressing both promoters.

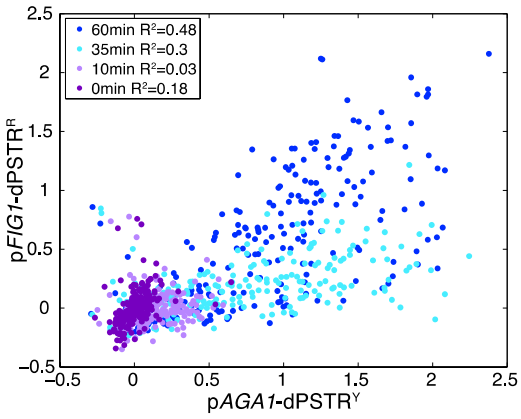


Figure 3.5 Single-cell instantaneous dPSTR correlation

Correlation of normalized dPSTR nuclear enrichments from all single cells at different time points after stimulation.

Characterization of mating-induced promoters expression kinetics

Having established that the two promoters p*FIG1* and p*AGA1* are induced with different kinetics following pheromone stimulation, we tested the induction time of other mating-induced promoters, compared to p*AGA1*. Using the dPSTR^{R}, we quantified the induction of a total set of 14 mating-induced genes, in combination with the p*AGA1*-dPSTR^{Y} (Supplementary Figure 3). These genes were chosen based on the role of the proteins encoded, microarray data or analysis of TF binding sites presence (Bardwell, 2005; Chou et al., 2006; Roberts et al., 2000). For all these promoters, we quantified their expression level that we define as the maximal dPSTR nuclear enrichment measured after stimulation. This value was correlated to the response time measured in all expressing cells (Figure 3.6). We can clearly see that the selected promoters have different behaviors in terms of kinetics and levels of induction. Indeed, some promoters behave like the early-induced p*AGA1*, and are induced rapidly after stimulation and

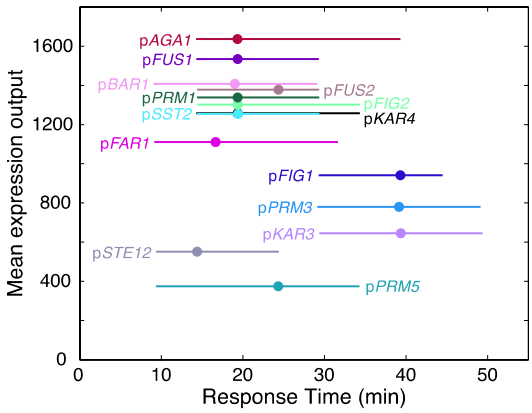


Figure 3.6 Dynamics of induction of mating promoters after pheromone stimulation

Response time versus mean expression output for the 14 mating-dependent promoters. Dots represent the median response times of the cell population and lines represent the 25th and 75th percentiles. All promoters were measured with the dPSTR^{R}. The strains also bear the p*AGA1*-dPSTR^{Y} for direct comparison of the dynamics of promoter induction.

have high expression levels. Among them, we can cite the protease Bar1, the negative regulator of pheromone signaling Sst2, and the main effector of the mating response Far1 (Bardwell, 2005; Roberts et al., 2000). Other promoters are induced later, like p*FIG1*, and two other proteins involved in karyogamy (Kurihara et al., 1996; Shen et al., 2009). Interestingly, the main mating TF Ste12 is also induced early in the response, but is expressed at low levels. For all promoters, we quantified the difference between their response time and the response time of p*AGA1* in the same cells (Figure 3.7). The early promoters have a distribution centered around 0, showing a strong temporal correlation with p*AGA1* (Figure 3.7A). We identified a third category of promoters that are slightly delayed compared to p*AGA1* (Figure 3.7B). Finally, the distribution of late promoters is shifted towards positive values, meaning that they are delayed compared to p*AGA1* (Figure 3.7C).

We can correlate the induction of two promoters in single cells. We can see that there is a strong correlation in the induction of p*FUS1* and p*AGA1* that holds true throughout the experiment (Figure 3.8A). Interestingly, we can see that there is also a good correlation between the two late-induced promoters p*FIG1* and p*KAR3*, which suggests a common regulatory mechanism (Figure 3.8B).

In order to better characterize the induction of these promoters, we quantified the Correlative Promoter Variability (CPV). This metric is based on the formula that quantifies the intrinsic noise in gene expression (Elowitz et al., 2002). It reflects the variation between two promoters measured at the

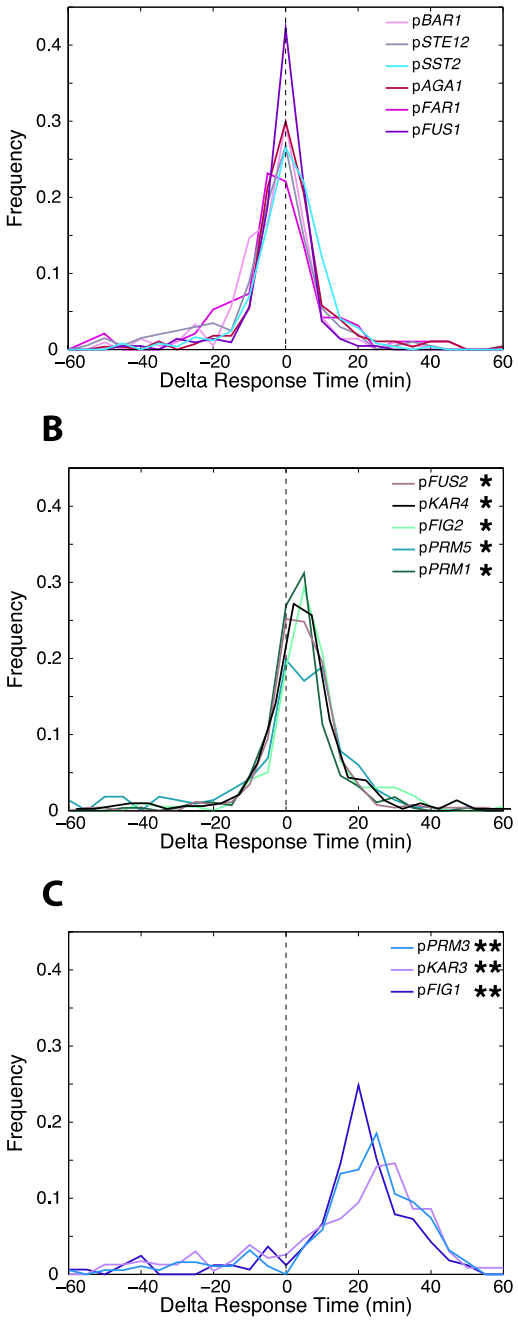


Figure 3.7 Distribution of the response time of all promoters versus p*AGA1*

Distribution of the difference between the response time for p*AGA1*-dPSTR^{Y} and the specified promoter measured with dPSTR^{R} for one representative experiment. A Signtest was performed to assess distribution centered around 0 (*10^{-20} <pVal<10^{-3}; **pVal<10^{-20}).

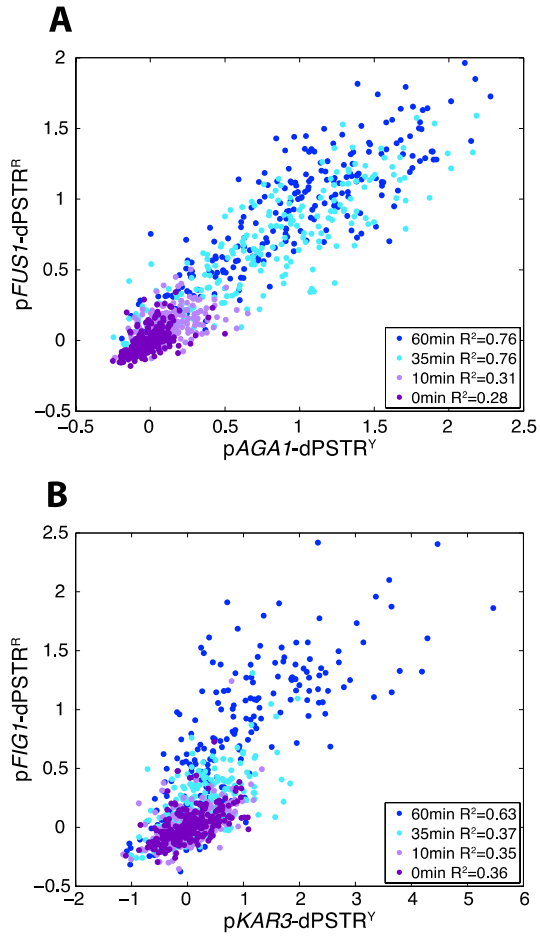


Figure 3.8 Correlation of two pairs of promoters

Correlation of the normalized dPSTR nuclear enrichments from all single cells of a representative experiment at different time points after stimulation in a strain with p*FUS1*-dPSTR^R and p*AGA1*-dPSTR^Y (A) or p*FIG1*-dPSTR^R and p*KAR3*-dPSTR^Y (B).

stimulation by pheromone, the CPV value reflects the difference between the two promoters expression in log phase-growing cells. For the strain quantifying two copies of p*AGA1*, the CPV increases slightly right after stimulation, but decreases 10 minutes after stimulation, when the two copies start to be concomitantly induced, and goes below 20% (Figure 3.9A). Some promoters, like p*FAR1*, p*BAR1* p*SST2* or p*FUS2*, display a high CPV in basal conditions. This is due to the cell cycle dependent expression of these promoters, which is not correlated to p*AGA1*'s (Oehlen et al., 1996). However, p*FUS1* and p*AGA1* have a rather low CPV before and after stimulation, as their cell cycle dependent expression occurs at the same time, suggesting a co-regulation of these two promoters. The intermediate genes behave similarly to the early-induced genes, with a fast decay of the CPV

same time in the same cell. Before

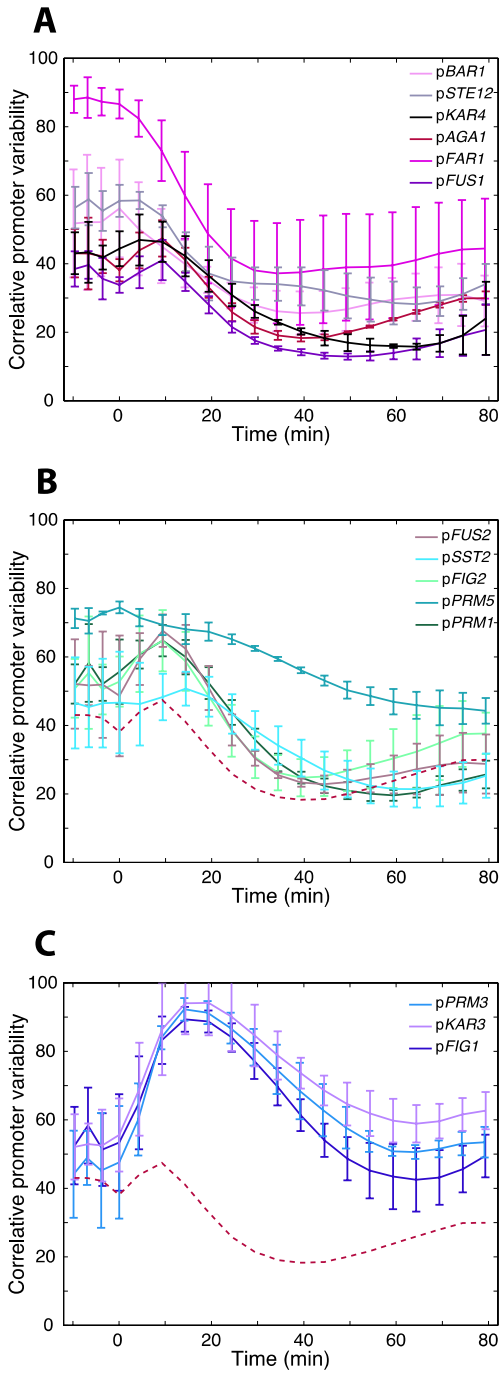


Figure 3.9 Evolution of the Correlative Promoter Variability for various pairs of promoters

The Correlative Promoter Variability (CPV) was measured in strains with different dPSTR^R and the p*AGA1*-dPSTR^Y, either early induced promoters (A), intermediate (B), or late (C). The curves represent the mean of 3 replicates and the error bar the standard deviation between replicates. A low CPV corresponds to a similar expression between two promoters in the same cell (see Methods). The dotted red line in B and C represents the CPV for p*AGA1* vs p*AGA1*.

after pheromone stimulation (Figure 3.9B). On the opposite, when looking at the CPV of p*AGA1* vs. p*FIG1*, there is a strong increase right after stimulation, due do the high variability between the two promoters, as p*AGA1* is induced but not p*FIG1*. As soon as the induction of p*FIG1* starts, around 20 minutes, the CPV decreases up to 40% (Figure 3.9C).

Basal induction of the promoters

Using the dPSTR, we also quantify the basal expression of the promoters. It was shown that some of them are induced during specific phases of the cell cycle (Oehlen et al., 1996). For instance, p*AGA1* is induced in G1 and p*FAR1* in G2/M. In order to determine the proportion of expressing cells in basal conditions, we quantify the nuclear enrichment of a dPSTR^R where the inducible moiety of the dPSTR is absent from the cells. This reflects the inherent capacity of the fluorescent part of the dPSTR to accumulate in the nucleus. Any higher nuclear enrichment will be due to expression of the NLSs-SZ peptide through induction of the promoter of interest. Figure 3.10 shows the basal expression of the 14 promoters of the study. We can see that most of the early genes have a basal induction due to their cell cycle regulation. Two out of three late promoters do not have a basal induction. In other word, there is no correlation between a basal expression and the kinetics of pheromone-induced expression.

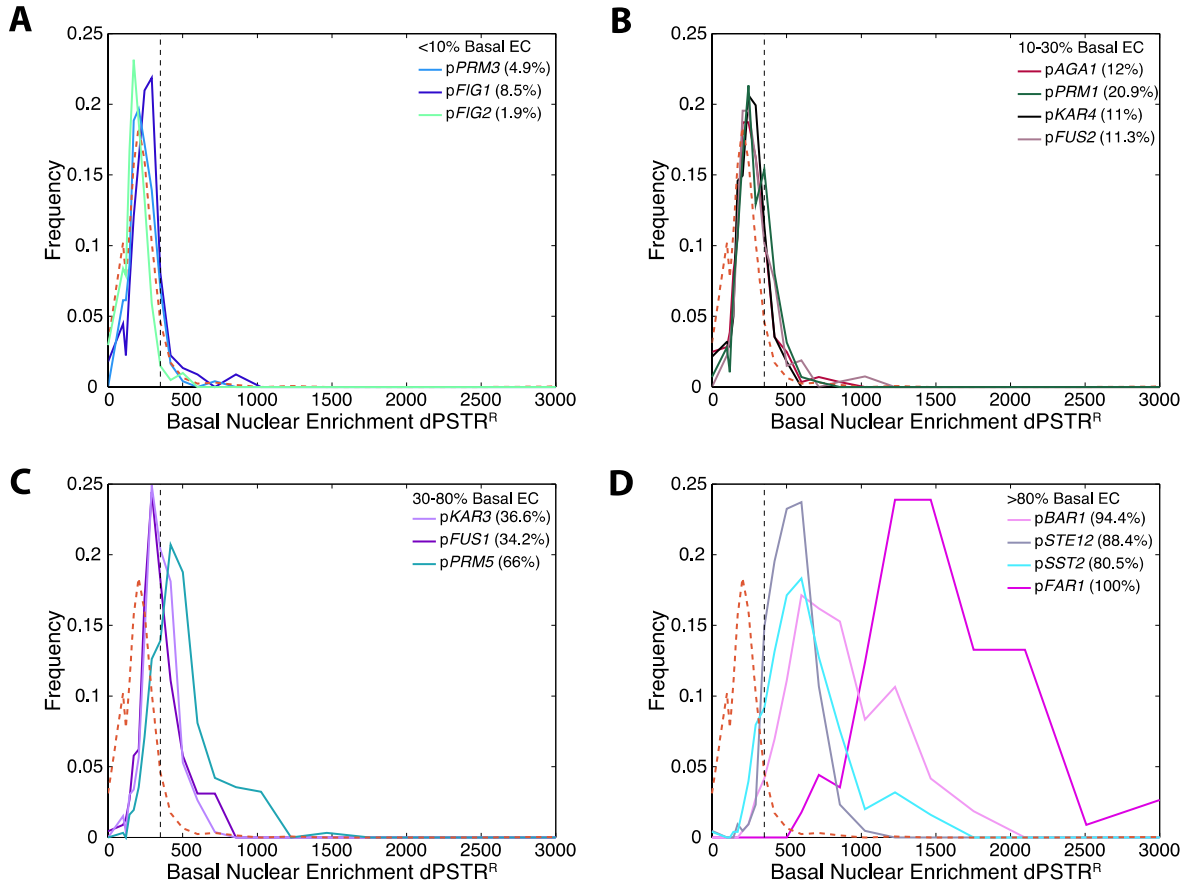


Figure 3.10 Basal level of expression

Histograms of the basal nuclear enrichment of the dPSTR^R are plotted for all cells from the representative experiments chosen for each promoter of the study. The orange dotted line on all plot is the nuclear enrichment observed in a strain carrying only the constitutive part of the dPSTR^R (mCherry-SZ2), without the second half of the construct (MCS2, where the inducible promoter is). From this strain, the threshold for basal enrichment, due to the fluorescent protein itself, was determined (black dotted line, N-C=350, N_C=2888), with 5% of false positive cells. Promoters were classified in 4 categories: no basal level (A), low basal level (B), mid basal level (C) and high basal level (D).

Expression of mating genes in presence of different concentrations of pheromone

In a mating mixture, the concentration of α-factor sensed by an **a** cell is evolving through two mechanisms. First of all, MATα will produce pheromone in an increasing manner, and it will diffuse in all directions. Then, the MAT**a** cell will produce a protease, Bar1, which degrades constantly the surrounding pheromone (Ballensiefen and Schmitt, 1997). These two events will lead to the establishment of a gradient of pheromone, which eventually allows the cell to find its partner (Barkai et al., 1998; Jin et al., 2011). Therefore, we assessed the expression of 10 promoters in response to a range of pheromone concentrations. We found that early-induced genes display a dose-dependent behavior, with increasing expression level as the concentration of pheromone increases (Figure 3.11A). Moreover, the proportion of expressing cells is also increasing with the pheromone concentration (Figure 3.11B). This means that as the concentration increases, more cells will induce the promoters to higher levels. This behavior is also true, although to a lower extent, for intermediate genes (Figure 3.11C-D). Late-induced promoters, however, show a different behavior. They are not induced if the pheromone concentration is lower than 300nM, which is characteristic of a switch-like behavior (Figure 3.11E-F).

We also performed dose-response experiments in a *bar1*Δ background, in which cells are not able to degrade the pheromone in the medium (Supplementary Figure 4) (Ballensiefen and Schmitt, 1997). In this background, p*FIG1* is induced at lower concentrations than in a WT background. The fact that in this condition the pheromone is not degraded will change the amount sensed by the cells. Indeed, in the WT background, the degradation will lead to lower sensed amounts, whereas in a *bar1*Δ the cells sense more pheromone for longer times. Hence, even at low concentration, the signaling is sufficient to activate p*FIG1*. This shows that the induction of this promoter is driven by the sensed pheromone concentration or the duration of the signaling.

Pre-stimulation experiment

As p*FIG1* and p*AGA1* depict different expression patterns in response to pheromone concentration, we monitored their induction in response to increasing amounts of pheromone. To do this, we measured the cell 30 minutes prior to addition of saturating pheromone concentration (1μM) during which we added or not low amounts of pheromone (10nM; Figure 3.12A). The induction of p*AGA1* is low during the first 30 minutes of pre-stimulation, as expected (Figure 3.11). Following addition of saturating concentration of pheromone, the induction of p*AGA1* increases quickly, and reaches higher level in the case of pre-stimulation (Figure 3.12B). The promoter p*FIG1* is not induced within the first 30min, whether there is pre-stimulation or not, which was expected considering the dose-response experiment in Figure 3.11. In pre-stimulated cells, p*FIG1* is induced starting from only 10 minutes after saturating pheromone concentration addition, whereas without pre-stimulation the expression starts at 30 minutes. Overall, cells that were inducing p*AGA1* during the pre-stimulation step are inducing p*FIG1* before cells that did not express p*AGA1* within the first 30 minutes. Cells that did not express p*AGA1* during pre-stimulation express p*FIG1* with the same kinetics as cells that were not pre-stimulated (Figure 3.12D). We showed before that early genes were correlated to p*AGA1* (Figures 3.7 to 3.9). Hence, the induction of early genes during pre-stimulation allows faster induction kinetics of the late p*FIG1*. Among the early genes, there may be one or several genes that encode proteins involved in the induction of p*FIG1*. There could be induction of a transcription factor, which was presumably absent before pheromone addition, and needs to be expressed for p*FIG1* expression to happen. At this point, we can suggest the protein Kar4, a transcription factor involved in the regulation of the late-induced p*KAR3*, as a good candidate regulator (Kurihara et al., 1996).

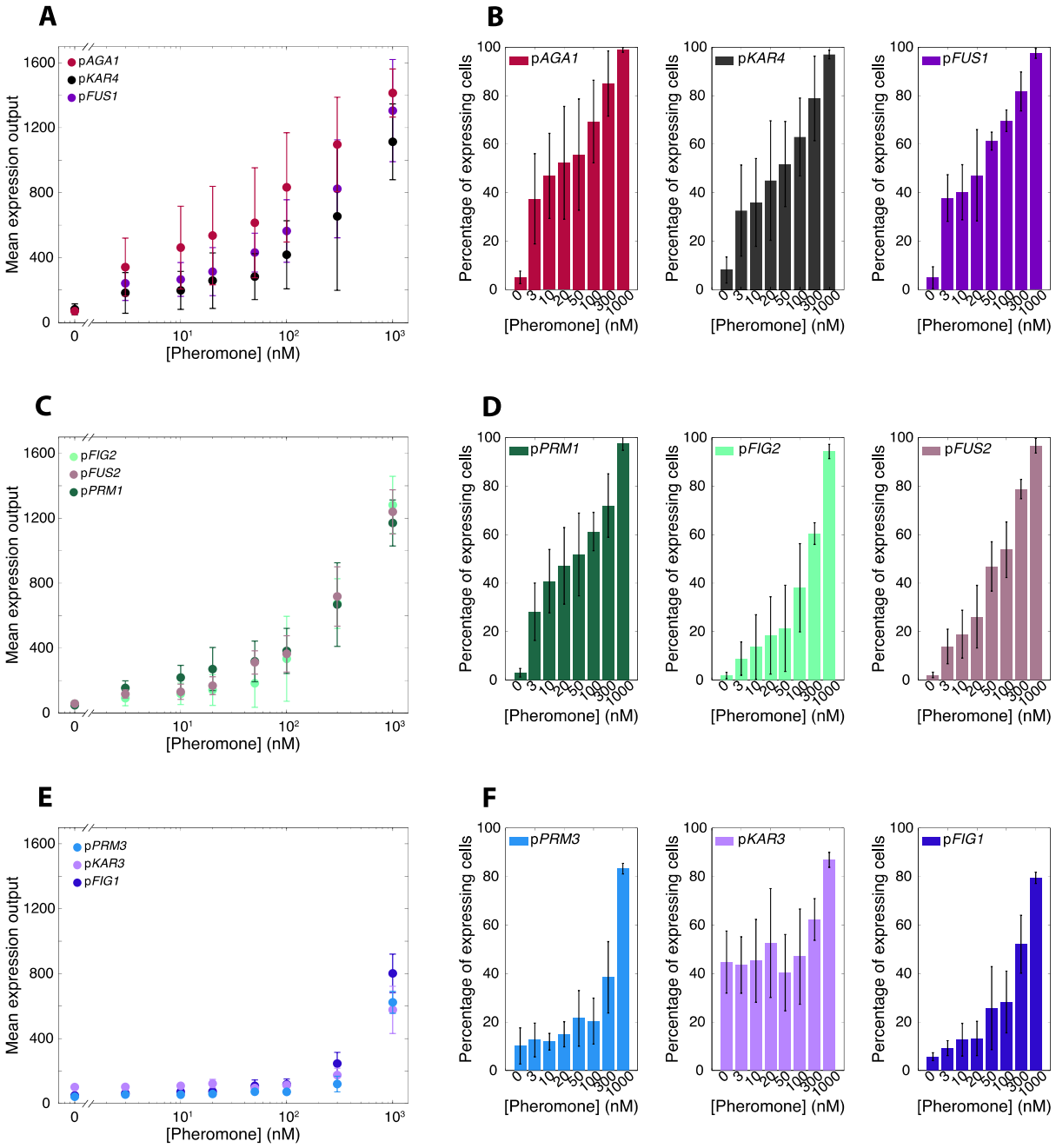


Figure 3.11 Dose-response of various mating-induced promoters

Mean expression output (A, C, E) and percentage of expressing cells (B, D, F) for the indicated promoters in response to different pheromone concentration, for either early-induced promoters (A, B), intermediate promoters (C, D) or late promoters (E, F). Error bars represent the standard deviation of 3 replicates.

Induction during mating

The previous experiments were performed in rather artificial conditions, where cells were settled in wells and exposed to exogenous synthetic pheromone in saturating concentrations at a specific time. However, in a mating mixture, the pheromones are not homogeneously distributed. Indeed, MATα cells produce the α-factor but MAT**a** cells degrade it. Hence, depending on cell repartitions, different patterns of pheromone repartition will arise. Moreover, sensing of the pheromone from the opposite mating type leads to a positive feedback loop that increases the pheromone production. All of these elements will shape the pheromone

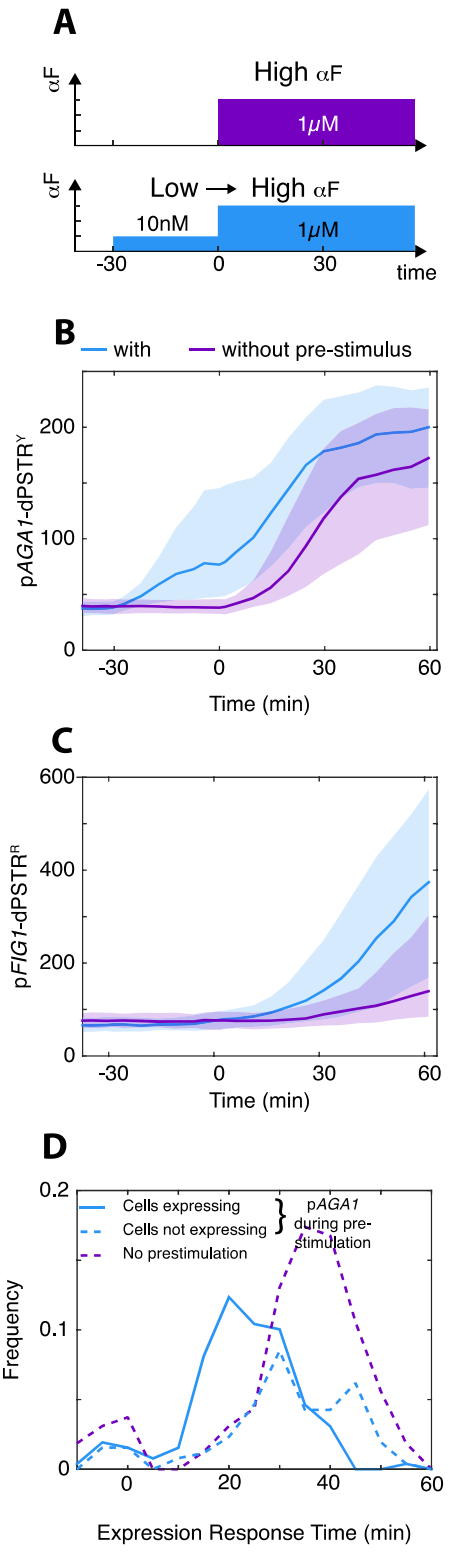


Figure 3.12 Pre-stimulation of strains by pheromone impacts on gene expression

A. Cells were imaged for 30min prior to stimulation by saturating concentration of pheromone (1μM), during which they were pre-exposed (blue) or not (violet) to low pheromone concentration (10nM).
B, C. p*AGA1*-dPSTR^{Y} (A) and p*FIG1*-dPSTR^{R} (D) nuclear enrichment in the conditions described in A.
D. Response time for p*FIG1* in cells that expressed p*AGA1* during pre-stimulation (filled blue line), did not express p*AGA1* during pre-stimulation (dotted blue line) or that were not pre-stimulated (dotted purple line).

distribution in a spatial and temporal manner. Hence, the cells will be exposed to changing amounts of pheromone.

We thus quantified the expression of some of our promoters in mating mixtures. To do this, we mixed the MAT**a** strains carrying the dPSTRs with MATα cells carrying a cytosolic tdi-RFP, allowing to distinguish them. The cytosolic FP allows to quantify the time of fusion, which we define as a sudden increase in iRFP signal in the cell containing the dPSTRs upon mixing of the two cytosols. With the use of the automated platform YeastQuant, we can now quantify events of fusion and monitor gene expression relative to the fusion events.

We noticed that at the beginning of the experiment, about a fourth of the cells are already expressing p*AGA1* at low levels, probably due to a low concentration of α-factor. Figure 3.13A shows images of a pair of cells of opposing mating types, which undergo fusion at time 0 minute. The promoter p*AGA1* is induced 25 minutes before fusion, whereas the promoter p*FIG1* is induced only 10 minutes prior to fusion. This suggests that the chronology in gene expression that we observed in the artificial conditions is conserved during mating. To confirm this, we measured the induction of the two promoters in hundreds of fusing cells (Figure 3.13B). We can clearly see that in fusing cells, p*AGA1* is induced gradually from 80 minutes before fusion, and this until fusion occurs. However, p*FIG1* is only induced within the 30 minutes preceding fusion. We also quantified the response time for the two promoters in this condition (Figure 3.13C). However, the response time is rather difficult to quantify for p*AGA1*, as part of the cells already induce it at the beginning of the experiment. As we quantify a fold increase in the nuclear enrichment, the response time of p*AGA1* may be earlier that what we measure here. Hence, the distribution of p*AGA1* response time is quite broad, but most of the cells induce it more than 35 minutes prior to fusion. On the other hand, p*FIG1* is mostly induced within the 30 minutes preceding fusion. We also noticed that there are

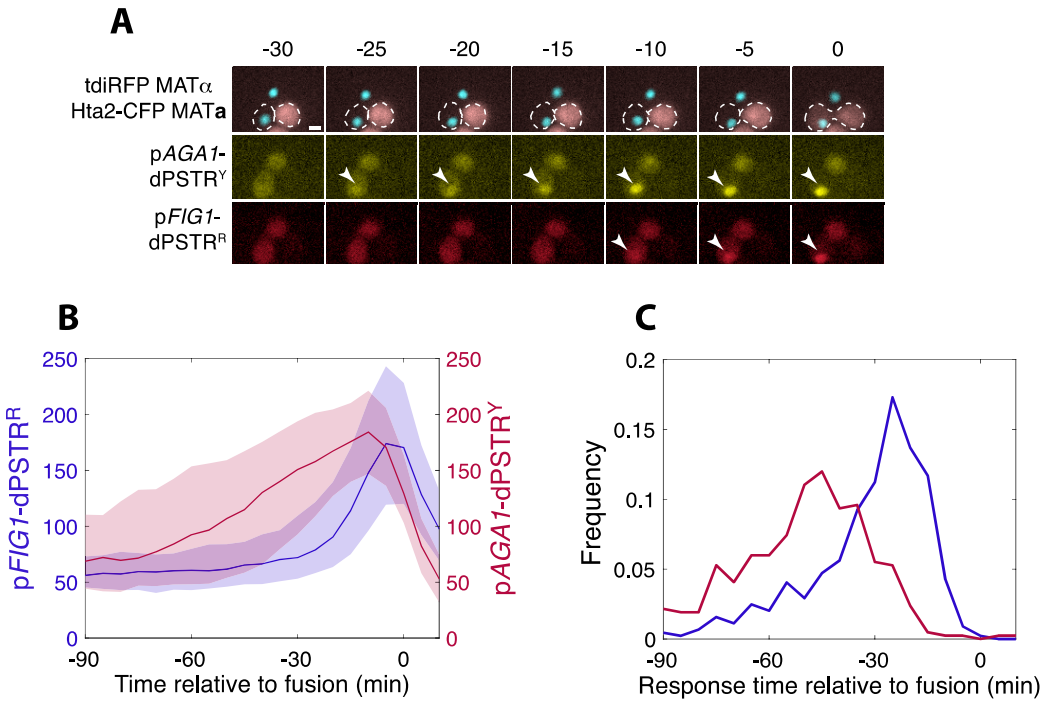


Figure 3.13 Dynamics of p*FIG1* and p*AGA1* during the mating process

A. Microscopy images of a mating mixture containing the MATa strain (Hta2-CFP, p*FIG1*-dPSTR^{R} and p*AGA1*-dPSTR^{Y}) and a MATα (cytoplasmic tdiRFP) at different times after beginning of the imaging (time 0).

B. Quantification of the nuclear enrichment of p*FIG1*-dPSTR^{R} (blue, left axis) and of p*AGA1*-dPSTR^{Y} (red, right axis). Single cell traces were synchronized relative to their fusion time, identified by a sudden increase in tdiRFP signal into the MATα cells.

C. Distribution of the response time of p*AGA1* and p*FIG1* relative to the fusion time.

only few cells expressing p*FIG1* without undergoing fusion, and that these cells were close to fusing pairs. We can assume that around fusing cells, the concentration of pheromone is high enough to promote induction of p*FIG1*.

We then performed mating experiments with strains measuring different promoters in combination with p*AGA1*-dPSTR^{Y}. To summarize our results, we plotted the cumulative probability to observe the expression of promoters in course of the time until fusion (Figure 3.14). The two promoters p*AGA1* and p*FIG1* are showing the most opposite behaviors. For p*AGA1*, the probability to find 50% of the cells that will undergo fusion expressing it occurs at 65 minutes before fusion. For p*FIG1*, the same number is reached only 30 minutes before fusion. Interestingly, some promoters classified as early-induced are expressed with a delay compared to p*AGA1*. The p*FUS1* and p*KAR4* are induced in 50% of the fusing cells 55 minutes before fusion. Moreover, promoters classified as intermediate have different behaviors than in the stimulation experiments. For instance, p*PRM1* behaves more or less like p*KAR4* and p*FUS1*, whereas p*FUS2* and p*FIG2* are closer to late-induced genes compared to the previous experiments.

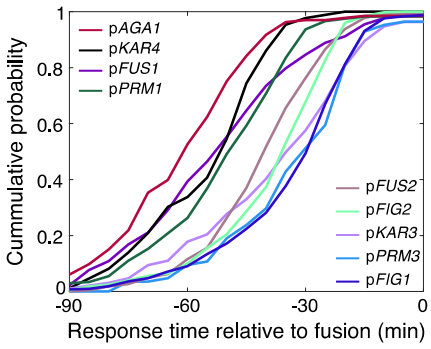


Figure 3.14 Induction of other promoters during the mating process

Cumulative probability of the response time relative to fusion for 9 pheromone-induced promoters measured in mating conditions.

Overall, we confirmed that even in real mating conditions, the mating genes are induced with a chronology relative to the time of fusion.

Link between timing of expression and encoded function

This really interesting behavior that we described must be of importance in the good proceedings of the mating. In order to understand this, we correlated the function encoded by the genes to their expression kinetics (Figure 3.15). We can see that among proteins expressed early in the mating response, we can find some regulator of the signaling. Among them, there is Far1, the regulator of important downstream events like cell cycle arrest and positioning of the shmoo site. This protein is induced only during the G1 phase, and likely has a very high turnover rate, so it needs to be strongly and quickly induced for commitment to the mating process. Indeed, its induction will promoter cell cycle arrest in G1, allowing the cells to induce later genes. If the amount of Far1 is not sufficient to arrest the growth, the cell will have to wait until the next G1 phase, one cycle later, before being able to signal again and induce late genes. Also, the cells will induce Bar1, degrading the pheromone, and Sst2, which is downregulating the signaling. This means that the cell, right after pheromone sensing, will activate negative feedback loops (Atay and Skotheim, 2017; Bardwell, 2005). At low pheromone concentrations, the negative feedback will turn off the signaling, both by decreasing the input signal, and turning off the signaling pathway, and cells will not commit to the mating response. However, if the concentration of pheromone is high enough, the cell will bypass the negative feedbacks

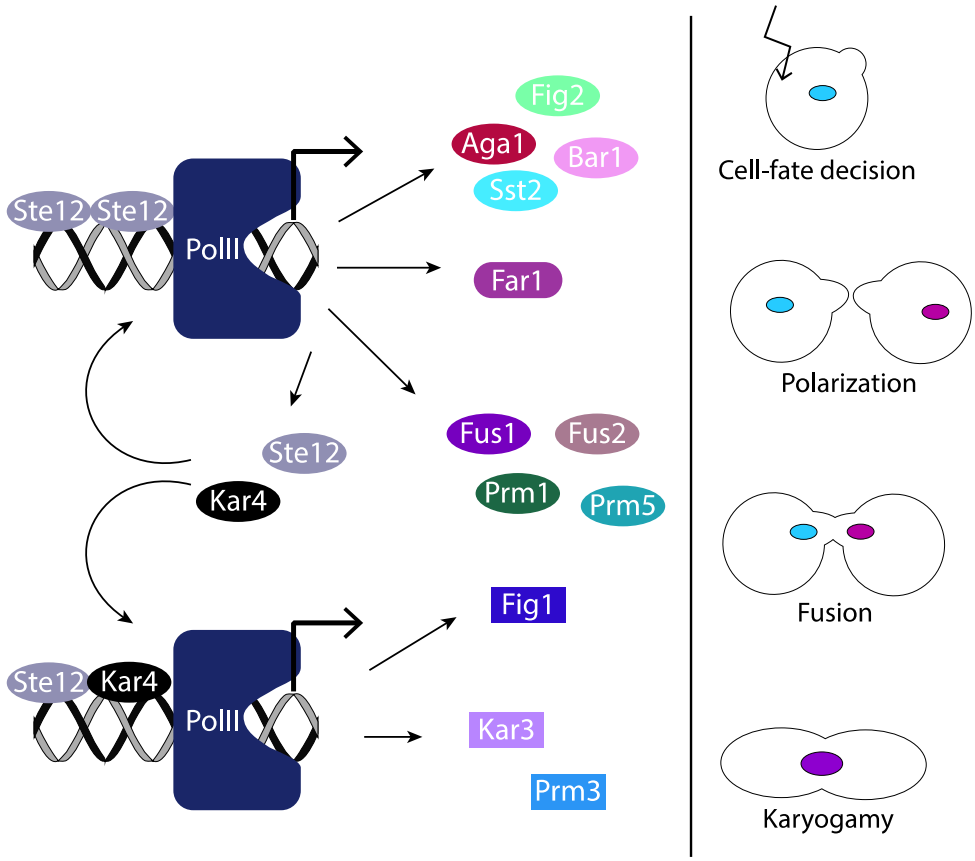


Figure 3.15 Summary of the pheromone-induced transcription

The 14 promoters of the study are expressed at various time following stimulation, which correspond to different steps in the mating process. The mating starts by a decision taken by the cell to proceed to mating. The cell will then start to polarize its growth towards the partner, after having arrested in G1. Later on, the cell wall of the two cells will be degraded, and their membrane will fuse. There will be exchange of cytoplasm, and finally fusion of the nuclei, resulting in a diploid zygote.

and proceed in the mating. Other early induced genes like *AGA1* and *FIG2* encode agglutinin proteins that allow the cell to bind its partner of the opposing mating type (Erdman et al., 1998). We think that this also plays a part in the cell-fate decision-making process. The two proteins Fus1 and Fus2 act in the fusion step, although it was suggested that Fus1 acts before Fus2 (Gammie et al., 1998). In agreement with this, we found that p*FUS1* was one of the early-induced promoters, and p*FUS2* one of the intermediates. Prm1 has a role in plasma membrane fusion, and has similar kinetics to p*FUS2* (Jin et al., 2004). Among early-induced genes, we can also find the two transcription factors Ste12 and Kar4, both known to be required for transcription of at least some mating-induced genes (Lahav et al., 2007; Roberts et al., 2000; Zeitlinger et al., 2003). Then, among the late-induced genes, we can find *FIG1*, implicated in the influx of Ca2+ during mating and required for efficient membranes fusion (Aguilar et al., 2007; Muller et al., 2003). The other two late-induced genes we identified are involved in the last step of zygote formation, the karyogamy (Kurihara et al., 1996; Shen et al., 2009).

To summarize, we think that there is a correlation between the time at which a mating-induced gene will be expressed, and the function performed by the encoded protein, in agreement with the timeline of mating steps.

Perturbation of genes induction

We showed that the chronology of mating-induced genes expression was correlated to the mating steps. We wondered if and how the perturbation of this chronology could alter the mating process. The protein Fus1 is involved in cell wall degradation, probably by allowing correct recruitment of vesicles at the shmoo tip (Gammie et al., 1998; Paterson et al., 2008). It may also be an anchor for the recruitment of Fus2, which is also required for efficient cell wall degradation. We showed previously that p*FUS1* is induced among the early genes. As it has a defined role in cell-cell fusion, we tried to delay the induction of the *FUS1* gene by placing it under the control of the late-induced p*FIG1*. We deleted the endogenous copy of *FUS1*, and placed another copy on a plasmid, under the control of p*FIG1* promoter. We quantified the number of diploid cells formed after 2 hours of mating on a filter by flow cytometry. The MAT**a** cells possess a $dPSTR^{R}$, which makes them RFP positive, and a $dPSTR^{Y}$ conferring them a low YFP signal. The MATα cells have a strong YFP signal. These features allow us to quantify diploid cells, which have a strong YFP signal and RFP signal. However, the results were not convincing. Indeed, the number of diploid cells was really low in general. Moreover, there is no significant difference in the number of diploids measured in the WTxWT cross or the WTx*fus1*Δ cross compared to the WTxdelayed-*FUS1* cross.

However, this experiment was performed only once as part of a master project (Jean-Jerrold Pierre), and needs to be reproduced and improved. First of all, the fluorescent signal of our cells was not optimal to efficiently categorize them with the flow cytometer. Indeed, we are now constructing strains with appropriate fluorescent proteins to match the flow cytometer requirements. Second, we should reproduce this experiment with different mating durations, as the number of WT mating pairs we obtained was fairly low compared to other studies (Salzman et al., 2015).

Discussion

In this Chapter, we described the chronology of the transcriptional response induced by mating stimulation, both during exogenous stimulation and in mating mixture. We found that all the mating-induced genes are not induced simultaneously, but rather that there is a succession of transcriptional events. Indeed, we showed that proteins shaping the different steps of the mating were induced in a chronological manner. Shortly after pheromone sensing and MAPK activation, proteins that are involved in the cell-fate decision process are induced. They will either lead to the downregulation of the mating pathway, or the commitment to the mating response. TFs are also induced at the same time, and are necessary for amplifying the transcriptional response to late induced genes. Next, genes required for the fusion step are expressed. Finally, among the late-induced genes, we found proteins involved in the fusion of the nuclei to form the diploid zygote.

In this study, we characterized the induction kinetics of 14 mating-induced promoters. For now, the dPSTR is a low-throughput technique that allows quantification of one promoter at the time. The process of strain construction requires several steps that are time-consuming: cloning of the promoter of interest in the cloning plasmid, subcloning into the dPSTR plasmid, and transformation in yeast. As such, the dynamic quantification of 14 promoters as a first set represents already an important step, and we hope that these promoters represent well the diversity of mating gene induction. However, we quantified the induction of only 14 promoters out of the more than 200 induced in response to the mating (Roberts et al., 2000). We could select and measure other important proteins involved in different mating steps, and see whether their expression time correlates with their roles, as we found for our set of promoters. Among these other 200 genes, we may also identify more late-induced genes, which are underrepresented in our set.

We think that the correct regulation of this transcription timeline is required to efficiently perform the mating. Indeed, if the expression of proteins involved in the fusion occurs too early, the cell might attempt to fuse before having established contact with the partner, which could result in cell death by lysis. On the other hand, if the induction of proteins involved in the fusion was to be delayed, the two partners would not attempt to fuse at the same time, which could also result in cell death of both partners. Of course, there may also be some intermediate phenotypes like delays in every step of the mating, from the cell cycle arrest to the fusion process.

In order to test whether the chronology of gene expression was important, we tried to perturb it by delaying the induction of the protein Fus1. Deletion of *fus1* results in partial fusion deficiency, but it is thought that the presence of Fus2 may compensate for the loss of Fus1, as only the double mutant completely fails to fuse (Gammie et al., 1998). As p*FUS1* is one of the early-induced promoters, we placed the gene encoding the protein Fus1 under the control of the late-induced p*FIG1*. With our experimental settings, we did not observe any significant phenotype. However, several elements may explain why this experiment did not work. First, it must be noticed that the expression level of these two promoters are different, as p*FIG1* reaches only half of the expression output of p*FUS1* (Figure 3.5). As we think that the expression output we measure with the dPSTR is correlated to the protein levels arising from the corresponding gene, any resulting effect may be due to either a lower amount of protein or a delay in the appearance of the protein, which we could not distinguish. Secondly, this experiment was performed in a *fus1*Δ background, where Fus2 might still be able to compensate for the loss of Fus1 (Gammie et al., 1998). We are currently building the *fus1*Δ*fus2*Δ double mutant in order to delay the induction of either Fus1 or Fus2 protein. With these strains, we hope to obtain different mating efficiency compared to the WT. Indeed, we expect a delay in the cell wall

degradation of our strains, and this should decrease the mating efficiency. We also could try to perform a mating competition assay, where we mix a WT and one of our strains, both MAT**a**, with a MATα strain. We could then quantify the proportion of mated WT versus the proportion of mated Fus-delayed strain. An alternative could be to perform time-lapse microscopy, and observe whether there is a visible effect of the delay of Fus1 or Fus2.

Understanding the regulation

The main perspective from this project is now to understand the temporal regulation of these promoters. One link between the promoters and the temporality lies in the expression behavior regarding different pheromone concentrations. We showed that early-induced genes have a dose-dependent behavior, whereas late-induced genes have a switch-like behavior (Figure 3.15). One link between pheromone concentration and the promoters is the amount of activated MAPKs present in the cells. Indeed, measurements with the SKARS showed that the activity levels of the MAPKs were dependent on the pheromone concentration (Durandau et al., 2015). This could mean that early-induced promoters and late-induced promoters require different amounts of active MAPKs to be induced. It has been shown that MAPKs can in some instances directly participate in the transcription process, by recruitment of factors at the promoters (de Nadal et al., 2011). It is possible that in these examples the late-induced promoters require the presence of active MAPK at their regulatory sequences to be induced, which could only happen when sufficient levels of active MAPKs have been reached.

Another hypothesis is that the expression of late-induced genes requires the presence of a factor that is absent or inactivated prior to mating stimulation. A good target for this is the TF Kar4 that we identified as one of the early-induced genes. This TF has been shown to be involved in the regulation of genes involved in karyogamy, such as *CIK1* and *KAR3* (Kurihara et al., 1996; Lahav et al., 2007).

To summarize, we describe in this part the timeline of gene expression ensuing from the mating activation. The remaining work is to understand what regulates this chronology, and this is the subject of the next Chapter.

Methods

Strains and plasmids.

Yeast strains are listed in Supplementary Table 1 and plasmids in Annex 1.

The dPSTR plasmids were transformed in a yeast strain from a W303 MAT**a** background, bearing a Hta2-CFP marker (ySP580).

Each dPSTR is fully carried by a single integration vector (pSIV (Wosika et al., 2016)) and integrated in the genome. The red (and yellow) variants of the dPSTR (dPSTR^{R} and dPSTR^{Y}, respectively) are integrated in the *URA3* (resp. *LEU2*) locus and based on interaction of the SynZips SZ1-SZ2 (resp. SZ3-SZ4) (Thompson et al., 2012), and the mCherry (resp. MCitrine) fluorescent variant (Aymoz et al., 2016). The relevant promoters of interest (-1000 to -1) were amplified and cloned upstream of the inducible stable part of the dPSTR, in pSP360 for the dPSTR^{R}, and pSP363 for dPSTR^{Y}, and checked by sequencing. The inducible part was then further cloned in the pSIV vector containing the FP part of the dPSTR (pDA157 for the dPSTR^{R}, and pDA223 for dPSTR^{Y}).

To quantify the kinase activity, SKARS plasmids were transformed in a strain carrying p*FIG1*- or p*AGA1*-dPSTR^{R} (Durandau et al., 2015).

For each transformation, 8 clones were screened based on their fluorescence intensities, and four clones with similar fluorescence levels were further analyzed by a time-lapse experiment upon stimulation with 1µM of α-factor, to discard clones that would display an aberrant relocation behavior.

Sample preparation.

The cells were grown overnight in selective synthetic medium to saturation (YNB:CYN3801, CSM:DCS0031, ForMedium). They were diluted to an OD600 of 0.05 in the morning and grow for 4h before the start of the experiment. All the time-lapse experiments were performed in well slides, for which selected wells of 96-well plates (MGB096-1-2LG, Matrical Bioscience) were coated with filtered solution of Concanavalin A in H2O (0.5 mg.mL-1, C2010-250MG, Sigma-Aldrich) for 30min, rinsed with H2O and dried for at least ten hours. Before the experiments, the cells were diluted to an OD600 of 0.04, briefly sonicated, and 200µL of cell suspension were added to a well. Imaging was started 30 min later to let the cells settle to the bottom to the well. To stimulate the cells, 100µL of a 3µM solution of synthetic exogenous α-factor was added in the cell to reach a final 1µM concentration of pheromone.

Microscopy

Images were acquired on a fully automated inverted epi-fluorescence microscope (Ti-Eclipse, Nikon) controlled by micro-manager (Edelstein et al., 2010) and placed in an incubation chamber set at 30°C, with a 40X oil objective and appropriate excitation and emission filters. The excitation is provided by a solid-state light source (SpectraX, Lumencor). The images were recorded with an sCMOS camera (Flash4.0, Hamamatsu). A motorized XY-stage allowed recording multiple fields of view at every time points, typically 5 positions per well, 8 wells per experiment. CFP (50 ms), RFP (300 ms) and YFP (300 ms) and 2 bright-field (10 ms) images were recorded at time intervals of 2 min before induction and 5 min after.

Data analysis.

Time-lapse movies were analyzed with the YeastQuant platform (Pelet et al., 2012). Briefly, the nuclei of the cells were segmented by thresholding of the CFP images. The contour of the cell around each nucleus was detected using two bright-field images. The cytoplasm object was obtained by removing the nucleus object expanded by two pixels from the cell

object. Dedicated scripts in Matlab (The Mathworks) were written to further analyze the data. Only cells tracked from the beginning to the end of the movie were taken into consideration. In addition, a quality control was applied on each trace and only cells with low variability in nuclear and cell area, nuclear CFP fluorescence, and a ratio of RFP to YFP fluorescence lower than a certain threshold, were kept for further analysis. The curves displayed in the figures are from one representative experiment out of at least three true biological replicates.

For each cell, the difference between its average intensity in the nucleus and the cytoplasm was calculated at each time point to plot the nuclear enrichment of $dPSTR^R$ and $dPSTR^Y$.

For further analysis, all retained cell traces were smoothed by a moving average of three points. The basal level is calculated as the mean of the three time-points preceding the stimulation. The corrected nuclear enrichment of the dPSTR was calculated by subtracting the basal level to the smooth trace. The expression output represents the maximal corrected nuclear enrichment of the dPSTR. The population average expression output was calculates on the mean trace of all cells. A threshold to qualify cells as expressing was defined as 20% of the population average expression output. For all expressing cells, dPSTRs traces were normalized between 0 and 1, and the response time was identified as the first time point, after stimulation, to exceed 0.2. For plots of population average correlation, and instant correlations, all cell traces were normalized by the mean trace of all cells.

The Correlative Promoter Variability (CPV) is calculated based on the formula from Elowitz *et al*. (Elowitz et al., 2002) for noises as the ratio of intrinsic and total noise:

$$CPV = \frac{\eta_{int}^2}{\eta_{tot}^2} \qquad \eta_{int}^2 = \frac{\langle (r_i - y_i)^2 \rangle}{2\langle r_i \rangle \langle y_i \rangle} \qquad \eta_{tot}^2 = \frac{\langle r_i^2 + y_i^2 \rangle - 2\langle r_i \rangle \langle y_i \rangle}{2\langle r_i \rangle \langle y_i \rangle}$$

r_i and y_i are the normalized nuclear accumulation from the i^{th} cell at a specific time point in the red and yellow channels, respectively. The normalization factors were obtained from the highest and lowest population averaged intensities from the entire data set for one replicate.

Mating experiments

Mating experiments were performed in agar pads loaded into 96-well plates to monitor multiple strains in parallel. Agarose 2% in Synthetic medium was heated 5 minutes at 95°C. Approximately 150µL of this solution was placed in a small aluminum frame to form a square pad of the proper dimension to fit in a well plate. After cooling at 4°C for 5 minutes the pad was removed from the frame. In parallel, 500 µL of cells at OD_{600} 0.1 were centrifuged. MAT**a** cells were resuspended in 10 µL of synthetic media and this cell suspension was used to resuspend the MAT□ cells. 1 µL of this mixture was deposited on the agar-pad and placed up-side down in a well. Imaging was started roughly 30 minutes later, after selecting appropriate fields of view. CFP (50ms), RFP (100ms), YFP (100ms), tdiRFP (100ms) and 2 bright-field (30ms) images were recorded every 5 minutes for 2 hours. Cells were segmented based on the CFP and bright-field images. After quality control, cells tracked for at least 10 frames were taken into consideration for analysis. Fusion events were defined by a sudden increase of more than 50 in the average nuclear fluorescence in the tdiRFP channel. Cells were considered as non-fusing if their average nuclear fluorescence in the tdiRFP channel did not increase by more than 10 throughout the track.

Supplementary Table

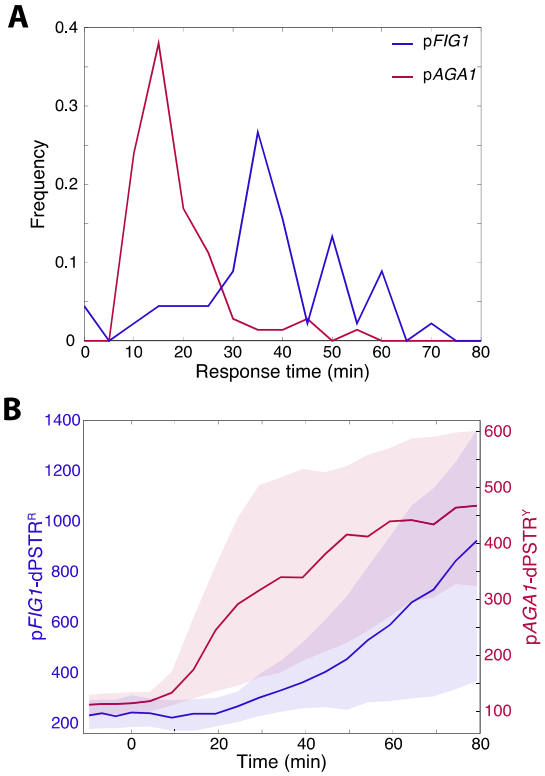
Supplementary Table 1: List of yeast strains used in this study

Strain	Background	Genotype	Plasmid
ySP2	W303	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*	
ySP580	W303	*HTA2-CFP*	
ySP644	ySP580	*HTA2-CFP*	
		LEU2: pFIG1-dPSTR^{Y4-3}	pSP367
ySP641	ySP580	*HTA2-CFP*	
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
ySP671	ySP580	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
yDA278	ySP641	*HTA2-CFP*	
		URA3: pRPL24A mCherry-SZ1	pDA157
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
ySP643	ySP641	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA146	ySP641	*HTA2-CFP*	
		URA3: pFAR1-dPSTR^{R2-1}	pDA229
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA186	ySP641	*HTA2-CFP*	
		URA3: pSTE12-dPSTR^{R2-1}	pDA233
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA189	ySP641	*HTA2-CFP*	
		URA3: pFUS1-dPSTR^{R2-1}	pDA239
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA198	ySP641	*HTA2-CFP*	
		URA3: pKAR4-dPSTR^{R2-1}	pDA228
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA244	ySP641	*HTA2-CFP*	
		URA3: pBAR1-dPSTR^{R2-1}	pDA300
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA245	ySP641	*HTA2-CFP*	
		URA3: pSST2-dPSTR^{R2-1}	pDA301
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA247	ySP641	*HTA2-CFP*	
		URA3: pKAR3-dPSTR^{R2-1}	pDA306
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA249	ySP641	*HTA2-CFP*	
		URA3: pPRM3-dPSTR^{R2-1}	pDA305
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
ySP664	ySP641	*HTA2-CFP*	
		URA3: pPRM5-dPSTR^{R2-1}	pSP384
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
ySP642	ySP641	*HTA2-CFP*	
		URA3: pAGA1-dPSTR^{R2-1}	pSP365
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA296	ySP641	*HTA2-CFP*	
		URA3: pFUS2-dPSTR^{R2-1}	pDA334
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA302	ySP641	*HTA2-CFP*	
		URA3: pFIG2-dPSTR^{R2-1}	pDA340
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA306	ySP641	*HTA2-CFP*	

		URA3: pPRM1-dPSTR^{R2-1}	pDA338
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
ySP646	ySP644	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pFIG1-dPSTR^{Y4-3}	pSP367
yDA203	yDA199	*HTA2-CFP*	
		URA3: pKAR4-dPSTR^{R2-1}	pDA228
		LEU2: pKAR3-dPSTR^{Y4-3}	pDA268
ySP663	ySP644	*HTA2-CFP*	
		URA3: pPRM5-dPSTR^{R2-1}	pSP384
		LEU2: pFIG1-dPSTR^{Y4-3}	pSP367
yDA207	ySP671	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pKAR3-dPSTR^{Y4-3}	pDA268
ySP614	ySP580	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R3-4}	pSP343
		TRP3: Ste7$_{DS}$-SKARS^{Y}	pED73
ySP615	ySP580	*HTA2-CFP*	
		URA3: pAGA1-dPSTR^{R3-4}	pSP350
		TRP3: Ste7$_{DS}$-SKARS^{Y}	pED73
yCS418	ySP580	*HTA2-CFP*	
		Ste12-9myc::NAT	
		Kar4-6HA::HPH	
ySP711	ySP122	MATα	pSP397
		URA3: pGDP-tdiRFP	

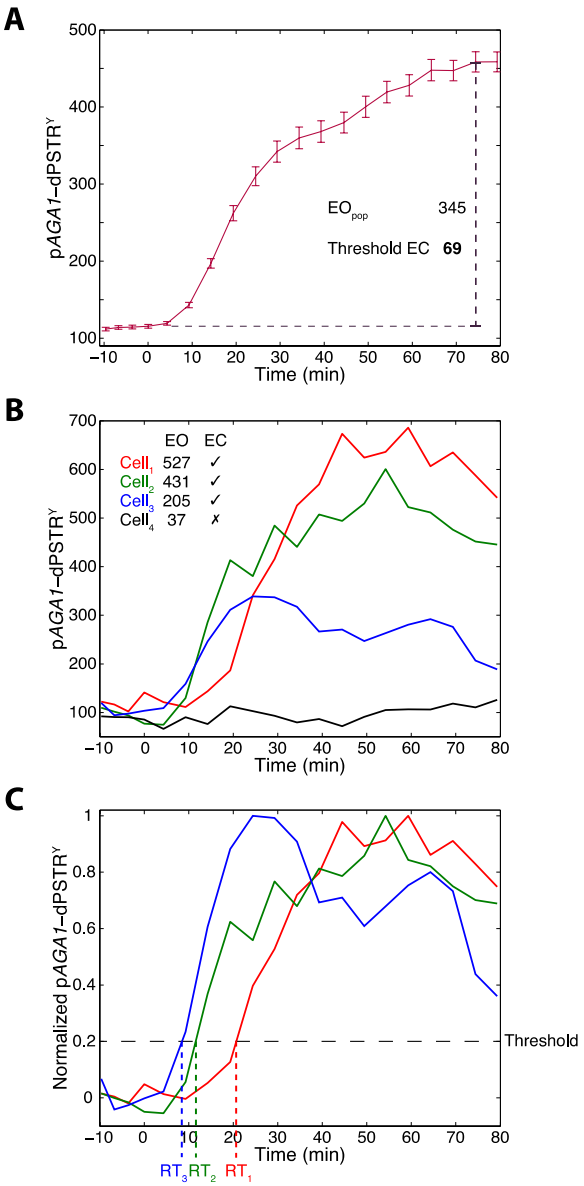
* The numbers in the superscript dPSTR^{R2-1} indicate the pair of SynZip used, and the letter the color of the fluorescent protein (R for mCherry and Y for MCitrine).

Supplementary figures



Supplementary Figure 1

Expression response of signaling competent cells
A. Distribution of the response time for the indicated dPSTR for cells activating the kinase within the 10min following the stimulation. These data are extracted from the same experiments used in panels C and D of Figure 1.
B. Population median (solid line) of the nuclear enrichment of the red p*FIG1*-dPSTR^{R} (left axis, blue) or the p*AGA1*-dPSTR^{Y} (right axis, red) for the strain used in panels E to G of Figure 1. Shaded areas represent the 25-75 percentiles of the single cell responses.



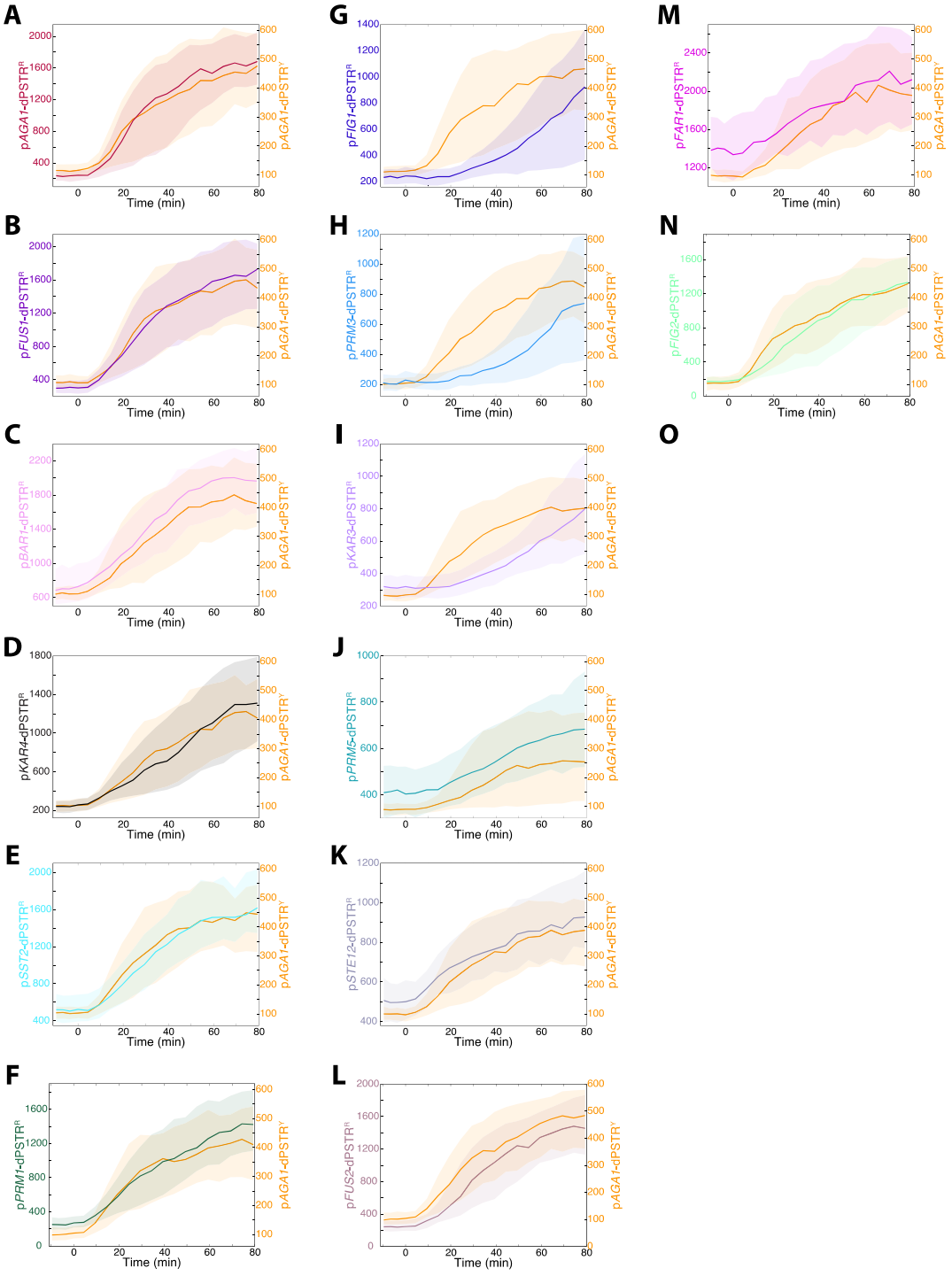
Supplementary Figure 2

Metrics calculation

A. Population average nuclear enrichment of the p*AGA1*-dPSTR^{Y} in course of time. Nuclear enrichment is the difference between nuclear and cytoplasmic fluorescence. All cells of the sample that passed quality control are taken in account here (low variability on nuclear and cell area and in nuclear CFP fluorescence). Error bars represent the SEM. The expression output (EO) of the population (EO_{pop}) is calculated from the average of all smoothed traces as the difference between the maximum of the trace and the mean of the three first time points. The threshold for expressing cells (EC) is arbitrary set at 20% of EO_{pop}.

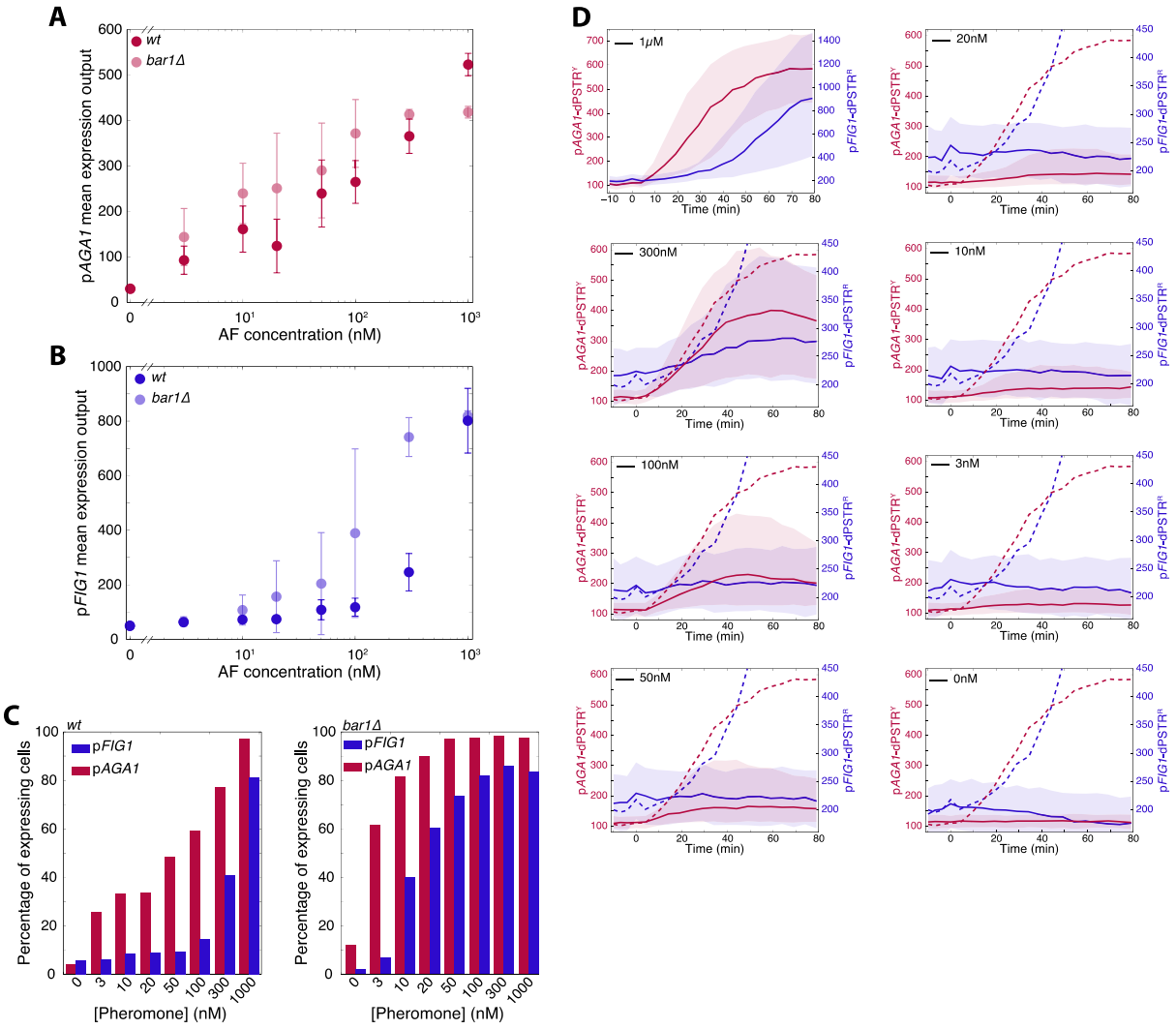
B. Single cell traces of nuclear enrichment of the p*AGA1*-dPSTR^{Y} in course of time. EO of each single cell is calculated as the difference between the maximum of the single cell smoothed trace and the mean of the three first time points. An expressing cell (EC) is a cell that overcomes a threshold defined arbitrarily as 20% of the EO_{pop}.

C. Normalized traces of expressing single cells are used to determine the response time (RT). All expressing cells are normalized between 0 and 1. The response time is determined as the first time point exceeding 0.2, meaning 20% of their own expression.



Supplementary Figure 3

Dynamics and expression level of mating-dependent promoters.
A to N. Population median (solid line) of the nuclear enrichment of the red dPSTR^R (left axis) or the p*AGA1*-dPSTR^Y (right axis, yellow curves) for the 11 promoters of the study. Shaded areas represent the 25-75 percentiles of the single cell responses. Note that the scale of the dPSTR^Y is identical for all graphs, whereas dPSTR^R scales are different. The basal level and induced level can vary according to the measured promoter. For instance, p*FAR1* has a high basal level, due to its cell cycle dependent induction.



Supplementary Figure 4
Dose response of p*AGA1* and p*FIG1* induction.

A and B. Mean expression output for p*AGA1* (A) and p*FIG1* (B) in response to different pheromone concentrations, in a *wt* or *bar1Δ* (shaded) background. Error bars represent the standard deviation of 3 replicates. Note that the induction of p*AGA1* gradually increases with the pheromone concentration, whereas p*FIG1* displays a switch-like response in *wt*. Note that in a *bar1Δ* background, the expression occurs at lower concentrations and with higher level for both promoters. The first dot is the non-induced control.

C. Percentage of cells expressing p*AGA1* (red) or p*FIG1* (blue) in a *WT* (left panel) or *bar1Δ* (right panel) background, according to various pheromone concentrations. Note that only at high concentrations a significant proportion of the population expresses p*FIG1*.

D. Median nuclear enrichment of the p*AGA1*-dPSTR^{Y} (red, left axis) and of the p*FIG1*-dPSTR^{R} (blue, right axis) in course of time, upon the stated stimulation, in the *wt* background. The shaded area represents the 25th - 75th percentile. The reference curves at 1μM are represented in dashed line for comparison. Note that the scale of the p*FIG1*-dPSTR^{R} is different between the 1μM and the other concentrations.

Preamble to Chapter 4

In the last Chapter, we showed that mating signaling resulted in a chronology of gene expression. Indeed, genes are induced at various times after sensing the stimulus, and this timing correlates with the role of the encoded proteins. For instance, early-induced genes comprise genes involved in the cell-fate decision-making, and the commitment to the mating response. Some intermediate genes are encoding proteins required for the fusion with the partner. Finally, some late-induced genes encode proteins required for the final stage of zygote formation, the karyogamy. We showed that this chronology is maintained both in artificial conditions, with stimulation by exogenous pheromone, and in physiological conditions, in a mating mixture. We also found that genes expressed at different timings were differently induced according to the pheromone concentration sensed by the cell, and hence the level of MAP kinase activity. We propose that this chronology in gene expression is of high importance for the cell to preserve a correct mating efficiency.

In this Chapter, we will discuss a model of regulation of expression kinetics of mating-induced genes. To decipher this regulation, we mainly focused on two promoters, the late p*FIG1* and the early p*AGA1*, as models for the two classes of genes. We used two different strategies to understand how the mating-induced transcription was temporally controlled.

For the first one, we deleted several proteins that we thought might be involved in the control of this chronology. Among them, various transcription factors activated by different signals, MAP kinases and signaling proteins, as well as chromatin remodelers were mutated.

For the second strategy, we tested the hypothesis that the kinetics of induction could be controlled by the promoter architecture itself. To address this, we mapped all the binding sites for the main mating transcription factor Ste12. We systematically mutated Ste12 binding sites in the two promoters, changed the distances between them, and built chimeric promoters in order to understand which features were important to predict the induction kinetics. As a proof that we understand the parameters that control the induction kinetics, we aimed at delaying the induction of the early p*AGA1* and accelerating the induction of the late p*FIG1*.

Using these two strategies, we came up with a model for the regulation of the induction kinetics of p*FIG1* and p*AGA1*.

Chapter 4

Deciphering the regulation of mating-induced transcription

Preamble

Parts of this Chapter are included in a research article under revision at *Molecular Systems Biology* (when this version is printed in November 2017), that is also published on BioRxiv (https://www.biorxiv.org/content/early/2017/10/02/197103) (https://doi.org/10.1101/197103).

Contributions:
Jean-Jerrold Pierre built and measured most of the strains presented in part IV.
Marta Schmitt and Delphine Aymoz built plasmids and strains.
Delphine Aymoz performed the time-lapse experiments, designed the experiments and wrote the manuscript.
Serge Pelet designed the experiments.

Abstract

Via MAPK signaling, cells establish a physiological program occurring over several hours leading to the formation of a diploid zygote following fusion with a partner of opposing mating type. In order to organize this event, the yeast modifies its transcriptional program through activation of hundreds of genes, regulated by the transcription factor Ste12. In the previous Chapter, we have established that the mating-induced transcription was occurring over different timescales following signaling. We also showed that pheromone concentration was one of the determinants of these temporal differences in gene expression. Here, we identify additional factors that control these expression kinetics. Using a genetic approach, we found the involvement of the TF Kar4 in the regulation of the late-induced gene *FIG1*. We then tested the importance of the promoter architecture in the expression kinetics, and found that the organization of Ste12 cis-regulatory elements on the promoters was regulating the expression kinetics.

Introduction

We showed previously that the induction of the mating-induced transcriptional program occurs with a specific chronology (Chapter 3). Some genes are induced early in the mating response and encode proteins that will lead to commitment of the cell to the mating process. This will lead to the induction of other genes, later, encoding proteins that will participate to fusion or karyogamy processes. This part aims at understanding the regulation of two prototypical mating-induced genes, *AGA1* and *FIG1*. We showed in the Chapter 3 that they have very different expression behaviors, as *AGA1* is induced within the 15 minutes following pheromone addition, whereas *FIG1* is induced much later. The gene *AGA1* is a cell cycle regulated promoter, induced in G1, which is strongly and quickly expressed upon pheromone stimulation (Oehlen et al., 1996; Roy et al., 1991). It encodes an agglutinin that allows agglutination with cells from the opposing mating type. It is not essential for mating in our static conditions, but is very important for mating in liquid medium or in nature, as its deletion strongly impairs mating in liquid medium but causes only a mild mating defect on solid medium (Lipke et al., 1989; Roy et al., 1991). *FIG1* is not expressed during vegetative growth, and is induced at late times following pheromone stimulation. Cells lacking *FIG1* show a very mild defect in fusion, which can be compensated by a high concentration of Ca^{2+} in the medium (Muller et al., 2003). The protein Fig1 plays a role in the membrane fusion process, by controlling the Ca2+ influx (Erdman et al., 1998; Merlini et al., 2013; Muller et al., 2003).

It was shown that the main TF of the mating, Ste12, binds to conserved regions termed PREs (Kronstad et al., 1987; Van Arsdell et al., 1987). These cis-regulatory elements have to be organized correctly in order to promote efficient pheromone-dependent induction (Su et al., 2010). However, their involvement in the regulation of timing of expression is still poorly understood, and we found that the architecture of mating-induced promoters is a key determinant in their expression kinetics.

As the two promoters p*AGA1* and p*FIG1* display very different kinetics of expression, we chose them as models and decided to study in details how their induction kinetics are controlled. First, we performed a targeted genetic deletions in order to identify the key regulators of gene expression. Then, we studied the architecture of the two promoters, and modified them in order to identify crucial elements involved in their kinetic regulation. Finally, we describe a model of the regulation of their induction kinetics.

Results

Part I: Ste12 is the main transcription factor for the yeast mating response

An experiment searching for MAT**a** cells that would grow in presence of α-factor allow the first observation of Ste12 involvement in the mating response (Hartwell, 1980). Indeed, cells mutated for *ste12* were able to grow in presence of the mating pheromone. Later, it was understood that transcription of mating-specific genes was altered in this mutant, leading to sterility and insensitivity to mating pheromone.

Since this first observation, numerous studies have been performed in order to understand the function, properties and regulation of Ste12, the main mating transcription factor. This first part will describe the knowledge accumulated about Ste12.

Properties of the protein Ste12

The gene *STE12* encodes a protein of 688 amino acids, with a predicted mass of 77.8 KDa (Dolan et al., 1989). Figure 4.1 represents all the identified domains and interactions of Ste12 that will be detailed in this part. The structure of the protein Ste12 is not characterized yet.

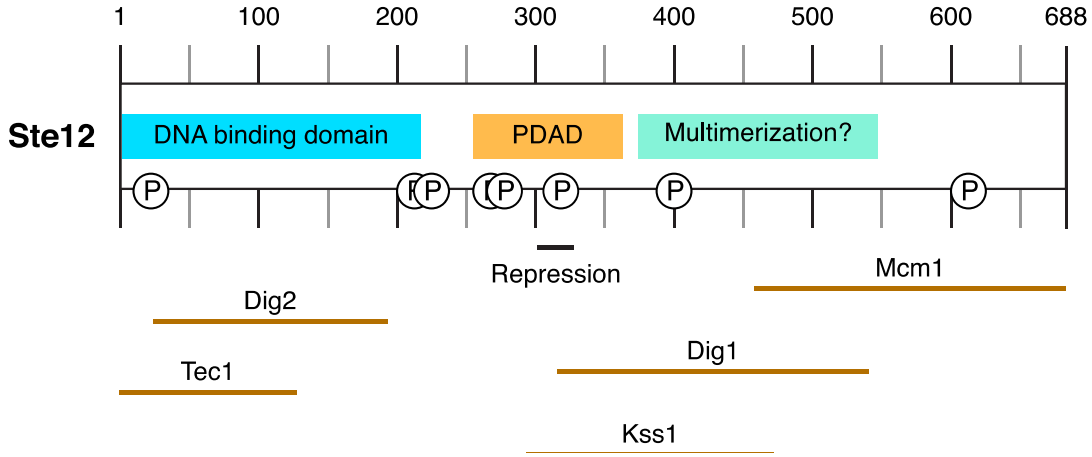


Figure 4.1 Structure of the protein Ste12

Through several studies, Ste12 functional domains have been partly identified and mapped. The domain involved in multimerization is not yet precisely mapped. Brown lines below represent sites of interaction with the specified protein. PDAD: Pheromone-Dependent Activation Domain. Circled Ps represent phosphorylation sites. See main text for details.

Ste12 regulates its own expression, which is increased by stimulation with pheromone (Figure 4.2). Surprisingly, we found that the level of induction of Ste12 in response to pheromone is quite low (See Chapter 3).

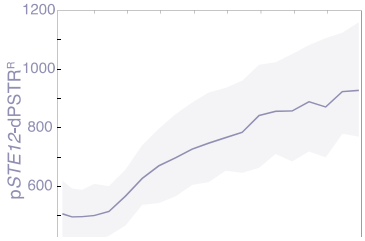


Figure 4.2 Expression of Ste12 measured by the dPSTR

Nuclear enrichment of the p*STE12*-dPSTR^R in course of time, following saturating pheromone stimulation at time 0 min.

The transcription factor Ste12 localizes in the nuclei of the cell (Figure 4.3). This localization does not change in presence of pheromone.

Ste12 is able to activate transcription both in absence and in presence of mating pheromone. Indeed, the fusion of the amino acids 214 to 473 of Ste12 to the Gal4 DNA-binding domain resulted in GAL1-*lacZ* expression even in absence of pheromone (Song et al., 1991). However, the fusion of the full Ste12 peptide is not able to induce transcription unless the cells are stimulated with pheromone, suggesting the existence of a repression mechanism, which will be discussed in details in a later section.

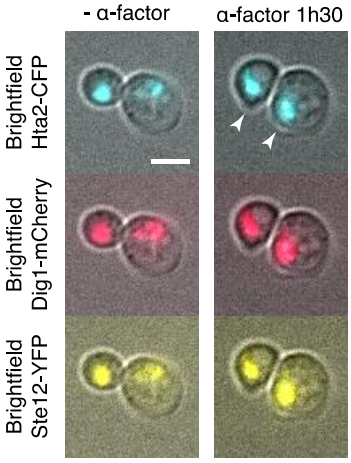


Figure 4.3 Localization of Ste12 and Dig1

Microscopy images of cells with histone, Ste12 and Dig1 tagged, before or after stimulation with saturating pheromone concentration. Arrows show shmoo sites.

DNA binding properties

Ste12 is a protein able to bind DNA at 5'UTR sequences of genes previously identified as mating-induced. Ste12 is able to bind to a specific DNA sequence, which was actually identified prior to the Ste12 functional studies, and named Pheromone Responsive Element (or PRE) (Kronstad et al., 1987; Van Arsdell et al., 1987). The sequence is the following: 5'-ATGAAACA-3'. It was noticed that the presence of these PREs in the promoter of several mating-induced genes was necessary to promote their full induction upon mating stimulation (Dolan et al., 1989; Errede and Ammerer, 1989).

The DNA binding domain of Ste12 was identified through the expression of truncated versions of Ste12 as the 215 N-terminal amino acids of the protein (Yuan and Fields, 1991). The minimal fragment 41-204 has been shown to bind the PRE sequences by DNA footprinting. Interestingly, insertion of linkers in the DNA binding domain, at position 85 or 103, prevented complementation of a *ste12*Δ, suggesting that this domain is very sensitive to mutations (Kirkman-Correia et al., 1993).

Indeed, we performed an alignment of the Ste12 proteins from several *Saccharomyces* species, and we can clearly see that the DNA binding domain is highly conserved (Figure 4.4). More broadly, the sequence KQKVFFWFSVA localized at position 149 in *S. cerevisiae* Ste12 has been shown to be extremely conserved in Ste12 homologous from various fungi (Wong Sak Hoi and Dumas, 2010).

Several studies addressed the genome-wide localization of Ste12 on DNA in basal conditions or following mating stimulation. The first identified 29 genes as directly regulated by Ste12, meaning bound by the protein (Ren et al., 2000). They claim that the binding of Ste12 to genes is very limited in absence of stimulation. The second study distinguishes the two transcriptional programs regulated by Ste12, and shows that different stimuli lead to different binding profile of Ste12 (Zeitlinger et al., 2003). Indeed, butanol activation of the filamentous growth pathway leads to accumulation of Ste12 at filamentous growth genes, in a Tec1 dependent manner. On the opposite, stimulation of the mating pathway leads to increased binding of Ste12 at mating genes targets.

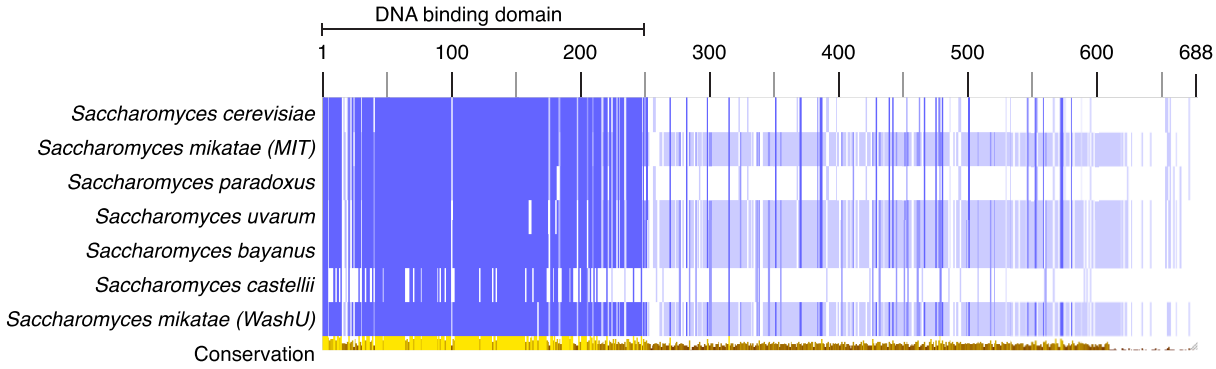


Figure 4.4 Alignment of *Saccharomyces* Ste12

Alignment of Ste12 proteins from the indicated *Saccharomyces* species, performed with Jalview. Darker blue amino acids are the more conserved. The conservation degree is depicted in a histogram in shades of brown to yellow (low to high conservation). Note that the DNA binding domain lies within the first N-terminal 250 amino acids.

Pheromone dependent activation domain

A domain of Ste12 has been identified as the minimal pheromone induction domain (Kirkman-Correia et al., 1993; Pi et al., 1997). This 35AA long part is located from amino acids 301 to 334. Its deletion confers a high basal activity to Ste12, with target genes being transcribed even in absence of pheromone. This suggests that this domain promotes repression of Ste12 in basal condition.

Phosphorylation of Ste12

Ste12 protein is highly phosphorylated in vegetatively growing cells. A study reports that at least 7 sites are phosphorylated in absence of pheromone (Hung et al., 1997). Interestingly, pheromone stimulation increases the phosphorylation level of Ste12, by at least two more sites. This modification requires a functional signaling pathway to occur, and happens within minutes after stimulation. *In vitro*, Ste12 was identified as a substrate for the two MAPKs Fus3 and Kss1 (Elion et al., 1993). It has been suggested that this phosphorylation of Ste12 was leading to its transcriptional activation, as the kinetics of phosphorylation were similar to the mRNA production kinetics measured (Song et al., 1991). However, the opposite, meaning that the phosphorylation could rather decrease Ste12 activity, has also been proposed (Pi et al., 1997).

Another hypothesis could be that the hyper-phosphorylated form of Ste12 could bind to different protein partners, in order to differentially regulate gene expression (Hung et al., 1997). In particular, phosphorylation of Ste12 could promote its interaction with itself and allow oligomerization. Alternatively, it was suggested that Ste12 phosphorylation could modify its DNA-binding properties and allow it to bind to multiple adjacent PREs (Su et al., 2010).

In summary, the role of Ste12 phosphorylation is still not completely understood.

Negative effect of Ste12

The transcriptional activator Ste12 has also been shown to have negative effects on the transcription. Indeed, the binding of Ste12 to two PRE sequences near the TATA box in response to pheromone modifies the expression of the *PRY3* gene (Bickel and Morris, 2006). A long transcript is induced during vegetative growth, and the binding of Ste12 not only represses its transcription, it also promotes the expression of a shorter transcript, through a switch in the TATA box choice for transcription start. The role of this shorter protein has not been identified yet, but cells can mate efficiently in absence of this product. However, lack of

repression of *PRY3* by Ste12 strongly impairs mating efficiency (to 50-60% of the WT) (Bickel and Morris, 2006).

Multimerization

It is widely accepted in the field that Ste12 forms homodimers in order to activate transcription. A study showed that, *in vitro*, the C-terminal part of the protein was able to interact with a full length Ste12 (Olson et al., 2000). The same study suggested that the repressor Dig1 might only be able to interact with dimers of Ste12. However, the experiments performed in this study may need to be reproduced, as the *in vitro* assay was performed with truncated recombinant Ste12 tagged with GST, mixed with yeast extract expressing full-length Ste12. Ste12 detection was performed with an antibody targeting Ste12 (residues 216-688), which may also bind to the recombinant Ste12, as the selected fragments are containing a big part of the targeted Ste12 fragment. It was also reported that Ste12 is able to bind *in vitro* to a single PRE 2010 (Su et al., 2010). In this study, they failed to observe multimerization of recombinant Ste12, and proposed that Ste12 may require additional factors, or post-translational modifications, to bind multiple binding sites *in vivo*.

Regulation of Ste12

Repressors of Ste12

Ste12 is mainly regulated by two proteins called Dig1 and Dig2 (previously named Rst1 and Rst2) (Cook et al., 1996; Tedford et al., 1997). These two proteins share 27% of identity, and are encoded by paralog genes that arose from the whole genome duplication. Nevertheless, the levels at which they are produced differs: Dig2 is present in cells in much lower amount than Dig1 during vegetative growth, and it was reported to be difficult to observe Dig2 in pull-down assays (Olson et al., 2000). According to the YeastGFPdatabase, there are 1460 molecules of Dig1 per cell, against 1310 of Dig2 in the S288C background, which differs from the one of the previous studies (Ghaemmaghami et al., 2003). The proteins both localize in the nucleus during vegetative growth, because they both possess an NLS, where they can interact with and regulate Ste12 (Tedford et al., 1997). Figure 4.3 shows that Dig1 localizes in the nucleus both in presence and in absence of pheromone. Additionally, they possess Ser/Thr sites that can be phosphorylated by MAP Kinases, and that were shown to be phosphorylated in response to pheromone stimulation. Also, they can interact with Ste12 independently of each other, suggesting that they bind to different regions of Ste12. This was indeed confirmed by assessing the binding of the Digs proteins to truncated version of Ste12 (Olson et al., 2000). Interestingly, Dig2 binds a region overlapping with the DNA binding domain of Ste12, potentially preventing it to interact with PREs (Figure 4.1). Dig1 seems to bind the region that was termed pheromone-dependent activation domain, which was described as sufficient to promote pheromone-dependent activation of Ste12 (Pi et al., 1997). The region contains a tyrosine residue that, when mutated to alanine, prevents the binding of Dig1. This region was previously thought to interact with both Dig proteins, which was later questioned by contradictory results (Olson et al., 2000; Pi et al., 1997). Altogether, these interactions suggest two distinct modes of regulation by the Dig proteins (Olson et al., 2000).

Regulation by the MAP Kinases

The two Dig proteins can both bind Fus3, and both bind Ste12. However, the interaction of Ste12 with Fus3 is indirect and requires the presence of one Dig protein (Tedford et al., 1997). In agreement with this, it was found that Fus3 enters in the nucleus upon pheromone stimulation, a process that may depend on the presence of the Digs proteins (Blackwell et al., 2007). As the Digs are substrates of Fus3, it was suggested that upon pheromone stimulation, Fus3 enters the nucleus, phosphorylates Ste12/Dig1 or Ste12/Dig2, which ultimately break the

repressor complex, freeing Ste12 from its repressors, allowing it to activate transcription of its target genes. In this process, it is still unclear which phosphorylation events are relevant (Blackwell et al., 2007; Tedford et al., 1997).

The role of the MAPK Kss1 in Ste12 regulation has also been assessed, as this MAPK is also activated during pheromone response (Gartner et al., 1992; Ma et al., 1995). Kss1 was found to directly bind Ste12 (region 298-473), and this binding is required to repress Ste12 (Bardwell et al., 1998a). This repression by non-activated Kss1 requires the presence of the two Digs protein (Bardwell et al., 1998b). The MAPK interacts with the TF through its activation loop, where the MAPKK phosphorylation sites are. Upon activation by the MAPKK, Kss1 is less able to bind Ste12. Moreover, it was shown that Kss1 can phosphorylate both Dig1 and Ste12 (Bardwell et al., 1998b).

Signaling-dependent degradation of Ste12

After long pheromone exposure, the cells are able to repress the mating response and enter a new round of cell cycle. One of the mechanisms allowing this down regulation is the degradation of Ste12 that has been observed (Esch et al., 2006). 60 to 90 min after stimulation, Ste12 is ubiquitinated and targeted for degradation, reducing its half-life from 230 minutes to 25 minutes. This process depends on the presence of the mating MAPK Fus3 and its main effector Far1. This implies a transient activation of mating-induced genes after pheromone stimulation, probably helping to cancel the fate decision taken by the cell, if they need to go back to vegetative growth.

Other proteins interacting with Ste12

Interaction with Tec1

Ste12 can interact with the transcription factor for the FG response, Tec1 to promote gene expression in response to the FG pathway (Chou et al., 2006). They can be found in a complex with Dig1, but little or no Dig2 was observed there. It seems that Tec1 competes with Dig2 to bind the same N-terminal region of Ste12. It was shown that the region of Tec1 that interacts with Ste12 is the same region required for its transcriptional activity suggesting that it is Ste12 that provides transcriptional activity to Tec1. The role of Tec1 here is simply to target Ste12 at filamentous growth responsive genes, which carry TCS motifs (CATTCY) (Madhani, 1997). Genome-wide studies identify genes regulated by the FG pathway (Breitkreutz et al., 2003; Roberts et al., 2000).

Interaction with Mcm1

Mcm1 is a DNA binding protein that is involved in cell type specific gene expression, in combination with the protein α1, α2 and Ste12 (Primig et al., 1991). Mcm1 is also implicated in transcription of genes during the M/G1 transition (Spellman et al., 1998). Mcm1 was shown to directly interact with Ste12, most likely through its C-terminal region (Errede and Ammerer, 1989; Primig et al., 1991).

Interaction with Kar4

A screen for karyogamy deficient mutants first identified *KAR4* as a gene that would lead to failure of the nuclei congression when mutated (Kurihara et al., 1994). The role of the protein was then assessed in more details, and it was found that it is required for expression of proteins performing the karyogamy process, specifically Kar3 and Cik1 (Kurihara et al., 1996; Lahav et al., 2007).

These three proteins are also expressed during the cell cycle, and Kar4 induction in G1 precedes its targets expression. However, Kar4 is not required for the induction of its target during vegetative growth. They all are induced by pheromone stimulation (see Chapter 3) (Kurihara et al., 1996). Interestingly, there are two start codons that have been identified for the expression of Kar4, resulting in two forms of the protein (Gammie et al., 1999). The long form of Kar4 is expressed in vegetative growth, whereas a shorter form is induced in response to pheromone treatment (Figure 4.16E). The presence of Ste12 is necessary for induction of Kar4 in presence of pheromone only for the induction of the shorter form. This could be due to a similar regulation as the one observed for the *PRY3* gene (Bickel and Morris, 2006). The protein Kar4 is phosphorylated after pheromone stimulation, although the role of this phosphorylation is unknown (Kurihara et al., 1996). Lahav *et al.* identified a minimal region on the promoter of *KAR3* that is bound by Kar4, in a cooperative manner with Ste12 (Lahav et al., 2007). However, to induce transcription, it requires both pheromone addition and Ste12 presence, indicating that either Kar4 requires indirect activation by Ste12, or they work in cooperation to induce transcription, like Tec1 does (Chou et al., 2006). Microarray analysis in a *kar4*Δ mutant identified a set of 29 genes induced by pheromone in a Kar4-dependent manner (Lahav et al., 2007). Among them, we can find *KAR3* and *CIK1*, the previously identified targets. Figure 4.5 summarizes the data that was presented here. During vegetative growth, Ste12 can be bound by Dig1, in combination with Kss1, which prevents it to recruit the polymerase. Hypothetically, this repressing complex could bind to PRE sequence on mating-responsive genes. Ste12 can also be found in complex with Dig2, and this interaction prevents the TF to bind the DNA. Another complex can exist, between Ste12 and Tec1, repressed by Dig1 and Kss1. Ste12 can also interact with Mcm1 and drive cell cycle dependent transcription. Upon mating signaling, the activation of the mating MAPK Fus3 leads to phosphorylation of several targets. First, the phosphorylation of Tec1 will lead to its degradation, which frees some Ste12. Then, Fus3 and/or Kss1 will phosphorylate the two Dig proteins, leading to their detachment from Ste12. As Ste12 is no longer repressed, it can drive the induction of its mating-induced targets. Among them, the short form of the protein Kar4 will be induced, and will promote the expression of genes involved in the karyogamy.

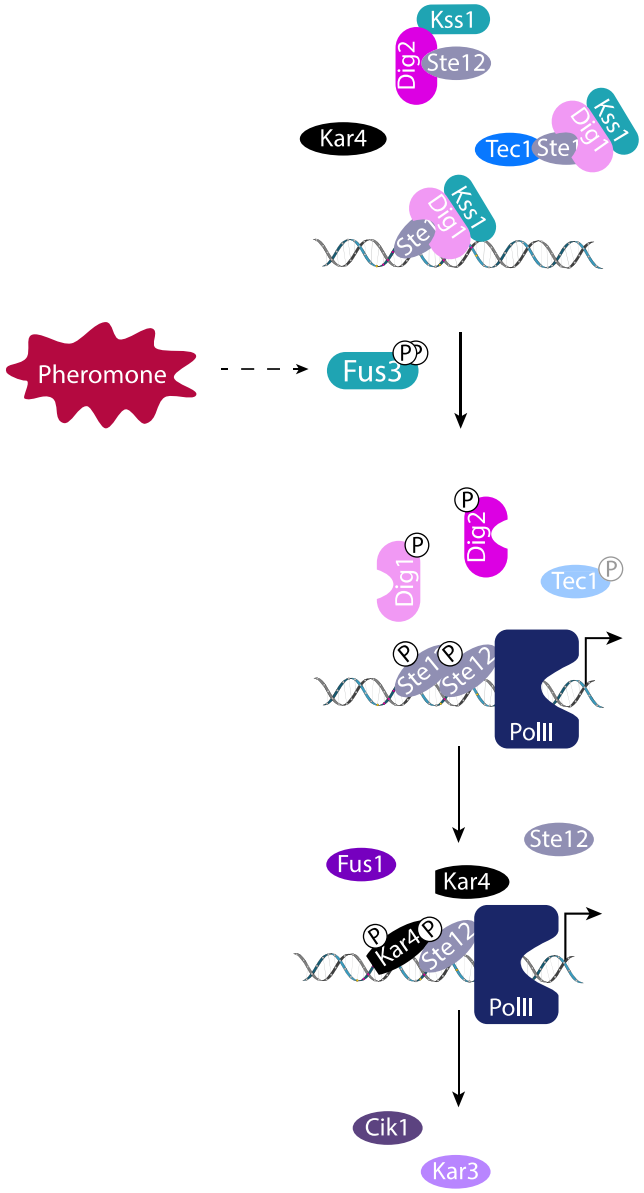


Figure 4.5 Overview of the pheromone-induced events regarding TFs

Part II: Genetic analysis

Deletion of proteins of the mating response

Deletions of key players of the mating signaling

First of all, we verified that the induction of the two promoters was requiring signaling through the mating pathway. We deleted the pheromone receptor encoded by *STE2*, as well as the MAP3K Ste11. These two mutants were totally defective for both p*AGA1* and p*FIG1* inductions (Supplementary Figure 1).

We then deleted the two MAP Kinases activated by the mating signaling, Fus3 and Kss1. We found that in absence of Kss1, the induction of both p*AGA1* and p*FIG1* is the same as in WT conditions (Figure 4.6A-B). Indeed, the function of the mating MAPK Fus3 is conserved and sufficient to promote their inductions. On the opposite, the loss of Fus3 results in different expression patterns. The expression of the early p*AGA1* is delayed by 10 to 30 minutes, although its expression level is comparable to a WT background (Figure 4.6C). The late p*FIG1* is mostly not induced with 27% of expressing cells versus 76% in the WT (Figure 4.6D, Supplementary Figure 2). This suggests that the signaling activity transmitted by Kss1 is not sufficient to promote induction of late genes, although it was shown that in a *fus3*Δ background, Kss1 is expressed at a higher level (Roberts et al., 2000). Indeed, using the SKARS, we showed that the activity of Kss1 reached a lower level in response to pheromone in a *fus3*Δ background (Durandau et al., 2015). As we showed that p*FIG1* induction was dependent on the MAPK activity, it is possible that the low signaling activity in *fus3*Δ is not sufficient to promote induction of lately expressed genes. Moreover, active Fus3 will degrade the FG TG Tec1. In absence of Fus3, Tec1 is active and can target Ste12 to FG responsive genes, decreasing the amount of Ste12 to target the mating genes (Chou et al., 2006; Madhani, 1997; Zeitlinger et al., 2003).

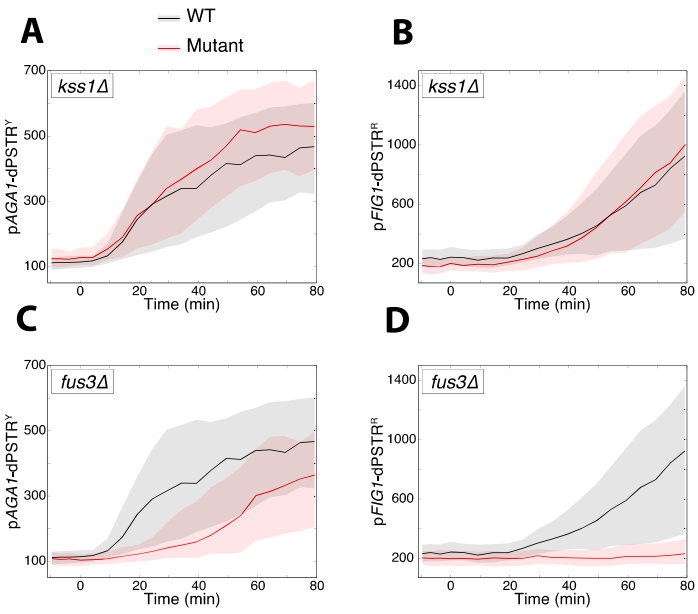


Figure 4.6 Expression of p*FIG1* and p*AGA1* in MAPK mutants

Nuclear enrichment of the p*AGA1*-dPSTR^{Y} (A-C) and p*FIG1*-dPSTR^{R} (B-D) after stimulation by 1µM of pheromone in the *kss1*Δ (A-B) or the *fus3*Δ (C-D) backgrounds. In all similar graphs, curves are the median of either the mutant (red) or the WT strain (black) for one representative experiment, with the shaded area representing the 25-75 percentile.

Deletion of targets of the mating pathway

As the presence of Fus3 was required to correctly induce the two promoters, we deleted some of the targets of the MAPK. Far1 is the main effector of the mating response, and controls the arrest of the cell cycle in G1 and the correct positioning of the polarization site (Barkai et al., 1998; Jin et al., 2011). It is directly phosphorylated by Fus3 in response to pheromone stimulation, which protects Far1 from degradation (Elion et al., 1993; Henchoz et al., 1997; Peter et al., 1993). In absence of *FAR1*, cells are unable to arrest their cell cycle because they cannot inhibit the cyclins (Bardwell, 2005) We found that in a *far1*Δ background, cells are able

to induce p*AGA1* as well as in a WT background (Figure 4.7A). However, the induction of p*FIG1* is strongly impaired (Figure 4.7B). Only 48% of the cells are able to induce it, and at a lower level than in the WT (Supplementary Figure 2). However, it is not possible to conclude whether the lack of induction of this late promoter is due to the absence of Far1 itself, or to the absence of cell cycle arrest. Indeed, the fact that the cells fail to arrest might impair all downstream events.

Two other targets of Fus3 are the repressors Dig1 and Dig2. These two proteins are inhibiting Ste12, probably by different mechanisms (Olson et al., 2000). Deletion of *DIG1* results in a delay in the induction of p*AGA1*, although the expression level is similar (Figure 4.7C). The promoter p*FIG1* is induced in fewer cells but to a similar level to the WT, although a delay can be observed (Figure 4.7D). Dig2 is expressed at lower levels than Dig1 in the cells, and binds to a different region of Ste12 (Olson et al., 2000). Indeed, it is thought to mask the DNA binding domain of Ste12, preventing it to be present at its target genes. Interestingly, Dig2 competes with Tec1 to bind Ste12 at the same region (Chou et al., 2006). The deletion of *DIG2* does not modify the induction kinetics or the levels of neither p*AGA1* nor p*FIG1* (Figure 4.7E-F). These two last results were surprising. Indeed, one could expect that removing an inhibitor of Ste12 should result in higher activity of Ste12, and hence higher and/or faster expression kinetics. As it was suggested that the two Digs proteins have a redundant function, we deleted them both (Cook et al., 1996; Tedford et al., 1997). For both promoters, we observe a high basal level of expression that matches to the maximal level observed after pheromone stimulation in a WT background (Figure 4.7G-H). Mating stimulation of this strain results in induction of p*FIG1* with the same kinetics as in a WT background. We barely observe an induction of p*AGA1*, but in this case, we believe that this is due to a saturation of the dPSTR^{Y} signal. Indeed, we noticed that the yellow version of the dPSTR has a decreased sensitivity compared to the red version, an observation that will be discussed in the Chapter 5 of this thesis. To conclude, the absence of Dig1 seems to delay the two promoters, and the absence of Far1 impairs transcription of the late p*FIG1*.

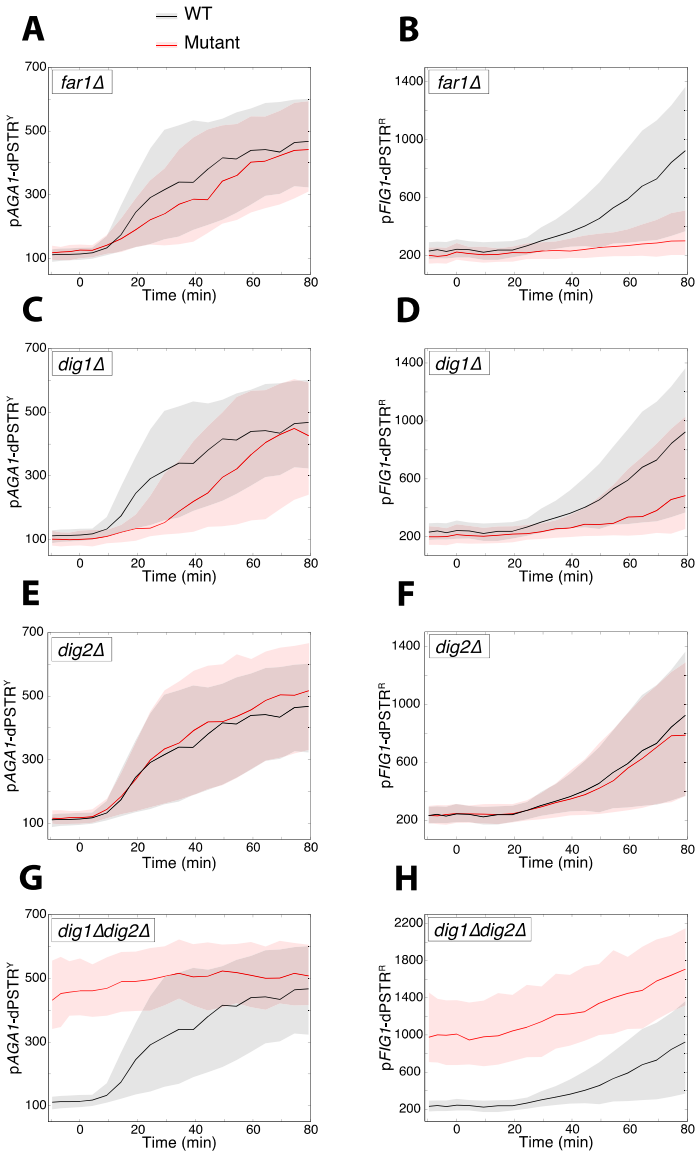


Figure 4.7 Expression of p*FIG1* and p*AGA1* in mutants of MAPK targets

Nuclear enrichment of the p*AGA1*-dPSTR^{Y} (A-C-E-G) and p*FIG1*-dPSTR^{R} (B-D-F-H) after stimulation by 1µM of pheromone in the indicated mutants (red curves) compared to the WT (black curves).

Role of the nucleosome on expression

Deletions of chromatin remodelers

In the Chapter 2, we studied the transcriptional response triggered by hyper-osmotic stress, for which it was previously shown that chromatin remodeling is a crucial step (de Nadal and Posas, 2010). Chromatin remodeling has been shown to be involved in this stress-induced transcription. The SAGA complex has been demonstrated to be required for establishment of the transcriptional response, whereas the INO80 complex is rather involved in the repression of transcription after adaptation occurred (Klopf et al., 2009; Zapater et al., 2007). To our knowledge, the involvement of chromatin remodeling in mating-induced transcription has not been reported. However, we deleted *GCN5*, encoding the histone acetylase of the SAGA complex, and *ARP8*, coding for an actin-related protein part of the INO80 complex. Arp8 does not play any role in the induction of either p*AGA1* or p*FIG1* (Figure 4.8A-B). However, deletion of *GCN5* impaired the induction of both promoters. The induction of p*AGA1* is delayed, and occurs only in 54% of the cells, compared to 93% in the WT (Figure 4.8C, Supplementary Figure 2). The expression of p*FIG1* is completely abrogated in this mutant (Figure 4.8D). The fact that the *gcn5*Δ cells look very sick prompted us to assess the signaling capacity in this mutant. We found that the kinase activity at the population level is reduced compared to a WT strain, which suggests that some cells may still be able to signal and activate transcription (Figure 4.8E). However, since only p*AGA1* is induced, although to a lesser extent than in a WT, we can hypothesize that chromatin remodeling is required for induction of the two promoters, but might be more important at p*FIG1*. Perhaps this could be linked to the fact that *AGA1* is expressed in certain phases of the cell cycle, which could result in a loosen chromatin configuration, compared to *FIG1*, which is not induced under vegetative growth conditions (Oehlen et al., 1996).

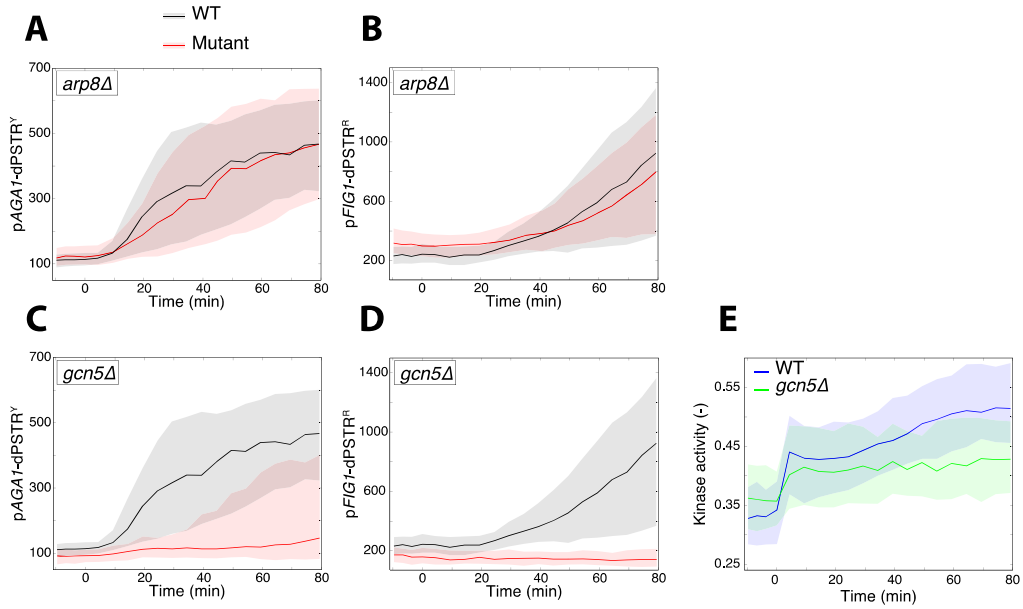


Figure 4.8 Expression of p*FIG1* and p*AGA1* in chromatin remodelers mutants

A-D. Nuclear enrichment of the p*AGA1*-dPSTR^{Y} (A and C) and p*FIG1*-dPSTR^{R} (B and D) after stimulation by 1µM of pheromone in a *arp8*Δ background (top) or *gcn5*Δ background (bottom). Black curves are WT and red curves mutants.

E. Mating MAPKs activity in a WT (blue) or a *gcn5*Δ background (green), measured by the Ste7$_{DS}$-SKARS.

Histone displacement at pFIG1 and pAGA1

This last result convinced us to monitor at various time the histone -1 displacement at p*FIG1* and p*AGA1* various times after pheromone stimulation. We found that the displacement of the histone occurs rapidly at *AGA1* locus in response to pheromone stimulation, in agreement with its fast kinetics of induction (Figure 4.9). However, histone removal at *FIG1* promoting region occurs with a delay compared to *AGA1*, again in agreement with its slow kinetics of induction. Nevertheless, the timing measured in this experiment should be used with caution as the strain used in this assay, which has two tagged transcription factors (yCS418), has a slightly different expression profile for these two promoters (see later section). Some investigations remain to be done to identify precisely which chromatin remodelers are involved in this process.

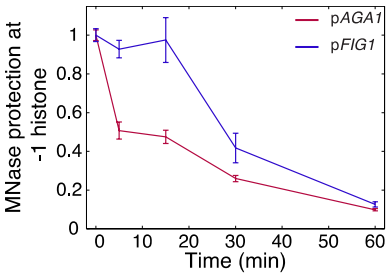


Figure 4.9 Nucleosome -1 displacement at *FIG1* and *AGA1* loci

Normalized nucleosome -1 occupancy at p*AGA1* (red) and p*FIG1* (blue) loci, quantified by Micrococcal Nuclease (MNase) assay.

Deletion of transcription factors

We then deleted several transcription factors activated by the mating but also by other signaling pathways to assess whether or not they would play a role in the induction of p*FIG1* or p*AGA1*. One of our hypotheses was that the induction of the late p*FIG1* could require the presence of an additional transcription factor that could be absent or inactive in the early mating response.

Deletion of the FG transcription factor Tec1

Mating signaling leads to the activation of the mating MAPK Fus3, but also of the MAPK Kss1 from the filamentous growth pathway. The transcription factor Tec1 is activated in combination with Ste12 but is quickly phosphorylated by Fus3, a process that targets it to degradation (Bruckner et al., 2004; Chou et al., 2004). The loss of Tec1 seems to only slightly affect the kinetics of the two promoters, which are slightly delayed (Figure 4.10). This is in agreement with the absence of TCS, the Tec1 binding sites, on both p*AGA1* and p*FIG1* (Chou et al., 2006).

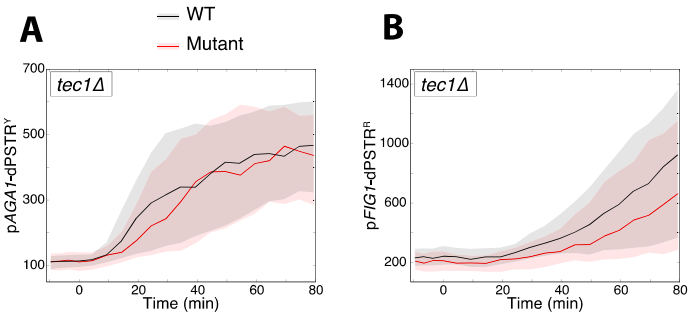


Figure 4.10 Expression of p*FIG1* and p*AGA1* in *tec1*Δ

Nuclear enrichment of the p*AGA1*-dPSTR^{Y} (A) and p*FIG1*-dPSTR^{R} (B) after stimulation by 1µM of pheromone in a *tec1*Δ mutant (red curves) compared to a WT strain (back curves).

Deletion of transcription factors of the Cell Wall Integrity pathway

Cells responding to mating remodel their cell wall in order to polarize their growth towards the source of pheromone. This allows them to meet their partner to fuse and form a zygote. However, this process, called shmooing, triggers the activation of another MAPK pathway: the Cell Wall Integrity pathway (CWI) (Levin, 2005, 2011). The CWI MAPK Mpk1 (or Slt2) is phosphorylated in response to pheromone (Buehrer and Errede, 1997; Zarzov et al., 1996). This is an indirect activation that is due to the polarized growth, which triggers a remodeling of the cell wall. The two events were shown to be correlated, and phosphorylation of the MAPK occurs from 30-60 minutes after pheromone stimulation. One could hypothesize

that the induction of the late p*FIG1* requires additional transcription factor(s) that could be activated by another signaling cascade. This prompted us to delete known transcription factors regulated by the CWI pathway.

Swi6 is a TF involved in the transcription at the G1/S transition, as part of the complexes SBF and MBF. Its deletion causes hypersensitivity to calcofluor white, which is causing damages to the cell wall, and is classically used as stimulus to activate the CWI pathway (Igual et al., 1996). It has been shown to be directly phosphorylated by Mpk1, which could as well regulate Swi6 during the cell cycle.

Skn7 is a transcription factor that has been shown to be involved in survival to various stresses, like osmotic stress, oxidative stress or cell wall integrity stresses (Brown et al., 1993; Krems et al., 1996; Li et al., 1998).

Rlm1 is a transcription factor whose activity depends on phosphorylation by Mpk1 (Dodou and Treisman, 1997). It is required for induction of most of the CWI-induced genes (Garcia et al., 2004).

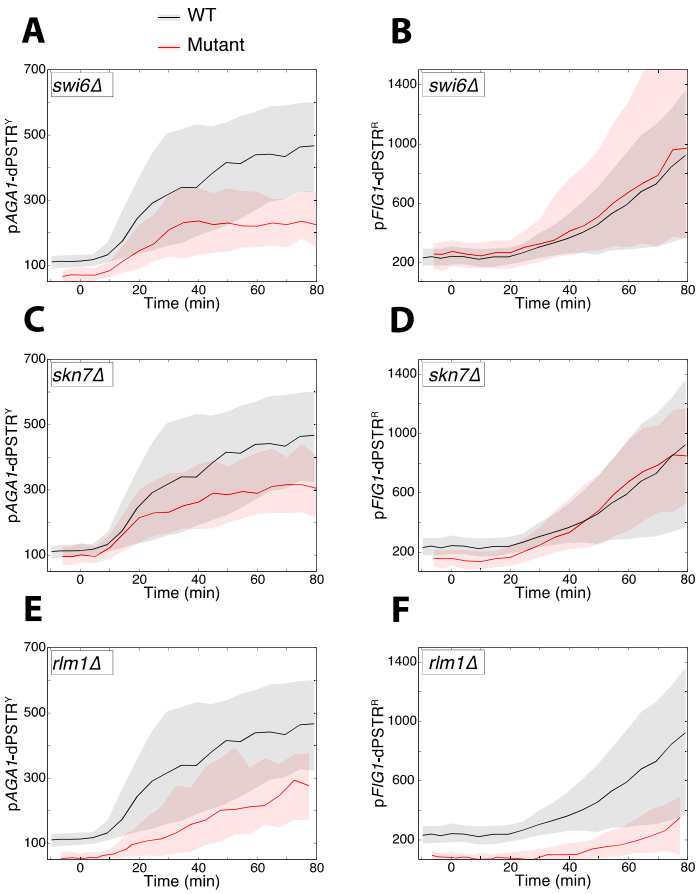


Figure 4.11 Expression of p*FIG1* and p*AGA1* in mutants of transcription factors of the cell wall integrity pathway

Nuclear enrichment of the p*AGA1*-dPSTR^{Y} (A-C-E) and p*FIG1*-dPSTR^{R} (B-D-F) after stimulation by 1µM of pheromone in the indicated mutant (red curves) compared to a WT strain (black curves).

We found that none of these factors seem to play a role in the establishment of p*FIG1* and p*AGA1* induction in response to pheromone (Figure 4.11).

Deletion of mating transcription factors

The most important transcription factor for the mating response is Ste12. It directly regulates the induction of more than 200 genes by binding to their 5'UTR sequences (Roberts et al., 2000; Zeitlinger et al., 2003). Ste12 binds to Pheromone Responsive Elements (PREs) that can be in various amounts and orientations in mating induced genes. This will be extensively discussed in Part III from this Chapter. We confirmed that the induction of the two promoters p*AGA1* and p*FIG1* was dependent on the presence of the TF Ste12. Indeed, no induction of either promoter is observed in *ste12*Δ background (Figure 4.12A-B).

The deletion of *mot3* have been shown to result in the same hypersensitive phenotype as *bar1* deletion, meaning that it is involved in negative regulation of the mating pathway (Grishin et al., 1998). The same study found that Mot3 was downregulating mating-induced genes such as *FUS1*, *AGA1*, and *KAR3*, as they result in more β-galactosidase activity in a *mot3*Δ than in a WT (Grishin et al., 1998). However, we do not report the same effect with the

dPSTR measurements. Indeed, we found that basal levels and pheromone induction of p*AGA1* and p*FIG1* are not affected by the absence of Mot3 (Figure 4.12C-D).

Kar4 is a transcription factor that is required for expression of proteins involved in the latest stage of the mating, the nuclei fusion of the two partners (Kurihara et al., 1996; Lahav et al., 2007). Microarray analysis showed its involvement in the induction of 28 genes, including *KAR3*. Neither *FIG1* nor *AGA1* were classified as Kar4-dependent genes. It was shown that Kar4 can bind cooperatively with Ste12 to a specific region of *KAR3* promoter (Lahav et al., 2007). Kar4 is a very good candidate for the regulation of p*FIG1*, as it was described as a factor involved in transcription of late mating-induced genes, like *KAR3*, although the timing of expression described in the literature looked different from what we measured (Chen and Thorner, 2007). Moreover, *KAR4* is induced during pheromone response, as we show in Figure 4.13, with very similar kinetics to p*AGA1* (Kurihara et al., 1996). The protein Kar4 is induced in two forms of different size depending on whether or not pheromone is present (Gammie et al., 1999). We observed that deletion of *KAR4* did not alter the induction of p*AGA1*, but dramatically reduced p*FIG1* induction (Figure 4.12E-F). Only 36% of the cells are able to induce p*FIG1* against 76% in a WT population (Supplementary Figure 2). In light of these results, we decided to study more broadly the involvement of Kar4 in mating-genes transcription.

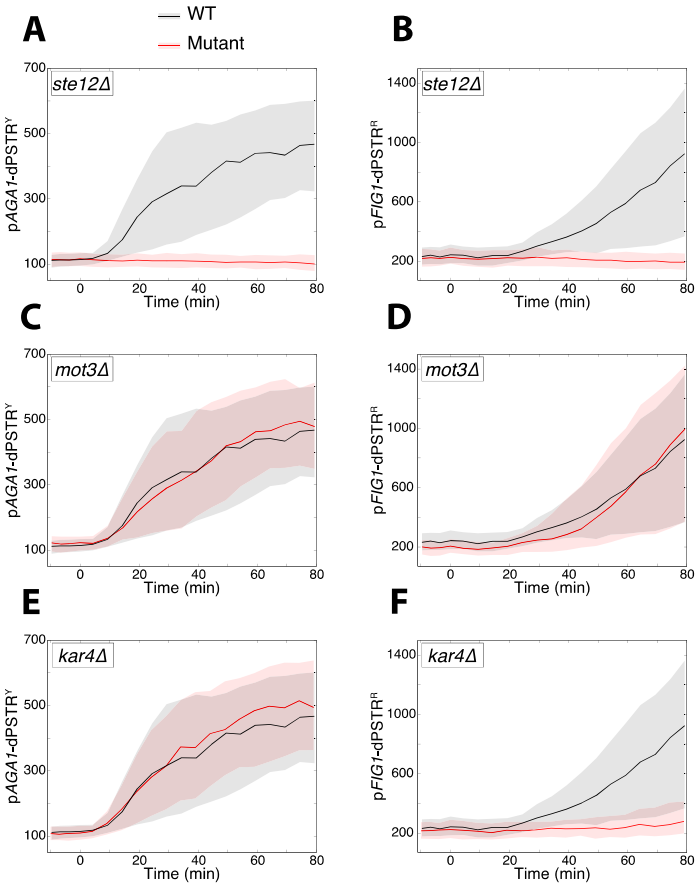


Figure 4.12 Expression of p*FIG1* and p*AGA1* in mutants of transcription factors from the mating pathway

Nuclear enrichment of the p*AGA1*-dPSTR^{Y} (A) and p*FIG1*-dPSTR^{R} (B) after stimulation by 1µM of pheromone in the indicated mutant. Curves are the median of either the mutant (red) or the WT strain (black) for one representative experiment, with the shaded area representing the 25-75 percentile.

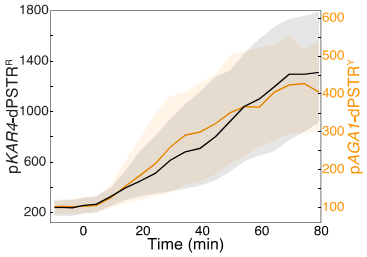


Figure 4.13 *pKAR4* induction measured by the dPSTR

Nuclear enrichment of the p*KAR4*-dPSTR^{R} (black, left axis) compared to the p*AGA1*-dPSTR^{Y} (orange, right axis), in the same cells.

Kar4 role in establishment of mating-induced transcription response

Impact of KAR4 deletion of induction of other mating-induced promoters

As we found *FIG1* to be Kar4-dependent, which was not identified with a different experimentation, we decided to test whether other promoters were Kar4-dependent (Lahav et al., 2007). We assessed the effect of the deletion of *KAR4* on the induction of a subset of promoters quantified with the dPSTR^{R}, in combination with p*AGA1*-dPSTR^{Y} (Figure 4.14). We confirmed that Kar4 was not required for the induction of p*AGA1*, as the distribution of its response time in a WT strain and *kar4*Δ strain is similar (Figure 4.14A-C). Surprisingly, we found that induction of intermediate genes, namely *PRM1*, *FIG2* and to a lesser extent *FUS2*, were also impaired (Figure 4.14E to M). Indeed, we found that the response time of all these genes was delayed compared to a WT condition. Despite this, the expression level is as high as in a WT background, which could explain why these genes were not identified as Kar4-dependent by Lahav *et al.*, as they measured gene expression after 90min of pheromone treatment (Lahav et al., 2007). Finally, we also found that Kar4 was required for the induction of another late gene, *PRM3*, in similar ways to *FIG1* (Figure 4.14N to S).

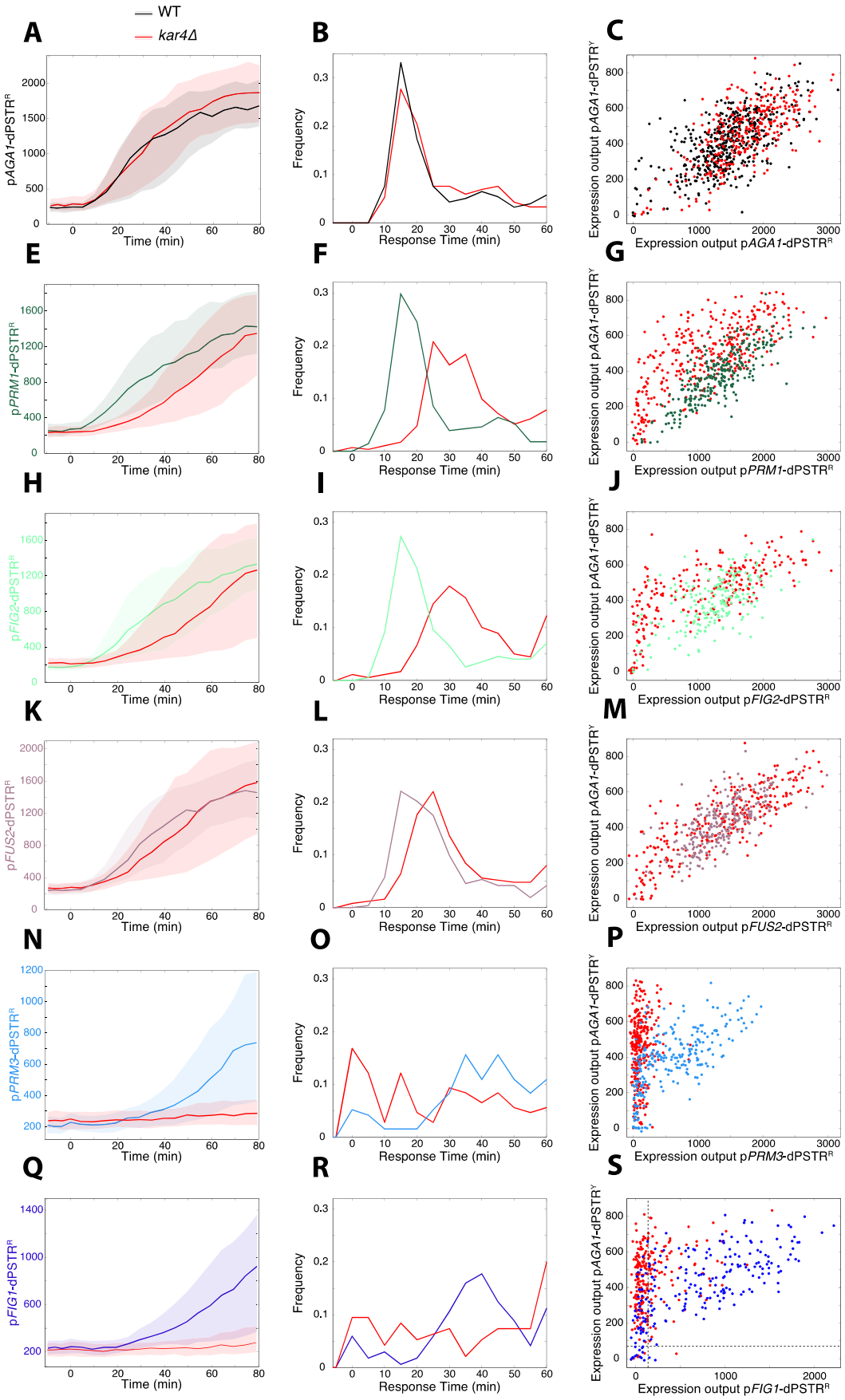
To conclude, we found that Kar4 may be involved in the transcription of most of the mating-induced genes, to different extent. It seems that the presence of Kar4 allows the optimal induction kinetics for intermediate genes, but is not required to reach the optimal expression output. However, we found that the presence of Kar4 is required for expression of late genes.

Figure 4.14 Effect of the loss of Kar4 on the induction of various promoters (on the next page)

A-E-H-K-N-Q. Nuclear enrichment of the dPSTR^{R} measuring the indicated promoter in a WT (same color as the axis) or a *kar4*Δ background (red).

B-F-I-L-O-R. Distribution of the response time of each promoter in a WT (same color as in right panel) or a *kar4*Δ background (red).

C-G-J-M-P-S. Correlation of the expression output of the promoters in a WT (same color as in right panel) or a *kar4*Δ background (red), compared to the p*AGA1*-dPSTR^{Y}.


WT
kar4Δ
A
B
C
pAGA1-dPSTR^R
Time (min)
Frequency
Response Time (min)
Expression output pAGA1-dPSTR^Y
Expression output pAGA1-dPSTR^R
E
F
G
pPRM1-dPSTR^R
Expression output pPRM1-dPSTR^R
H
I
J
pFIG2-dPSTR^R
Expression output pFIG2-dPSTR^R
K
L
M
pFUS2-dPSTR^R
Expression output pFUS2-dPSTR^R
N
O
P
pPRM3-dPSTR^R
Expression output pPRM3-dPSTR^R
Q
R
S
pFIG1-dPSTR^R
Expression output pFIG1-dPSTR^R

Kar4 and Ste12 interactions with DNA

In collaboration with the Posas lab, (Universitat Pompeu Fabra, Barcelona, Spain), we performed biochemistry experiments to learn more about Kar4 involvement in the regulation of p*FIG1*. We wanted to confirm several assumptions. First of all, as it has been suggested that Kar4 was binding DNA in cooperation with Ste12 (Lahav et al., 2007), we wanted to test whether Ste12 and Kar4 could interact together. Secondly, we wanted to know if and when Kar4 was binding to *FIG1* promoting region, as well as to *AGA1*. In order to do Chromatin ImmunoPrecipitation (ChIP) experiments, Kar4 and Ste12 were tagged with 6HA and 9myc tags, respectively. The initial idea was to tag them in separate strains and compare the behavior between these two strains. Unfortunately, tagging of either TF results in modification of pheromone-induced expression of both *AGA1* and *FIG1*, as shown by Northern Blot quantification (Supplementary Figure 3).

In order to be able to address if they could concurrently interact with DNA or not, a strain with the two tags was generated. Figure 4.15 shows the mRNA levels for both *FIG1* and *AGA1* in this strain (yCS418) compared to our reference strain (ySP580). The presence of the tags reduces the differences in induction kinetics, as *FIG1* is induced earlier. Moreover, levels of expression are impaired: *AGA1* levels decrease of about 50%, whereas *FIG1* levels are increasing to more than 130% after 30 minutes of stimulation (Figure 4.15B).

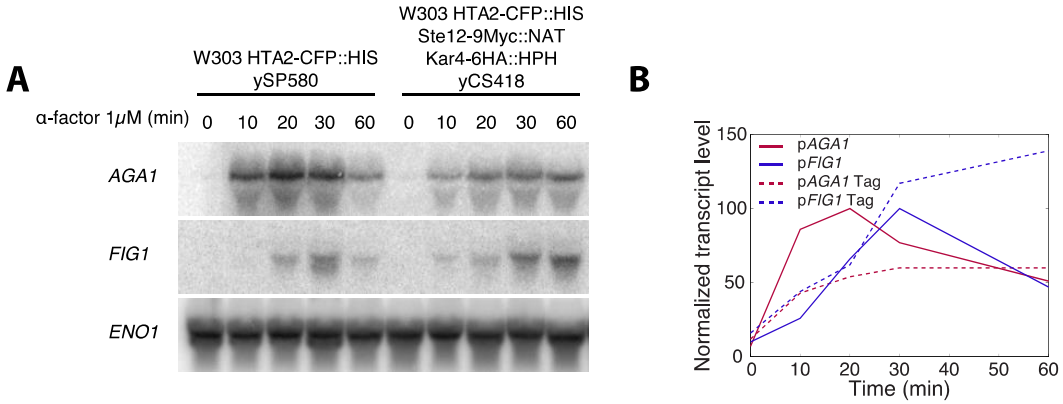


Figure 4.15 Tagging of Ste12 and Kar4 impairs *AGA1* and *FIG1* expression

A. Northern blot measurements of *AGA1* and *FIG1* transcripts in the strain used for dPSTR (ySP580, left) or the strain used for ChIP and CoIP experiments (yCS418, right).

B. Transcript level quantified from the experiment in A, normalized to the maximal level, for *AGA1* (red), *FIG1* (blue), in ySP580 (filled lines) or yCS418 (Tag, dotted lines).

Nevertheless, we quantified in this strain the binding of both Ste12 and Kar4 to *AGA1* and *FIG1* promoters. In basal condition, the two transcription factors are present at both promoters (Figure 4.16A to D). Pheromone stimulation induces further recruitment of both, at the two loci. Ste12 is quickly recruited at both loci, although the fold recruitment is higher for *FIG1* than for *AGA1* (Figure 4.17A). The recruitment of Kar4 is very weak at p*AGA1*, around 3 fold, but occurs immediately after pheromone stimulation (Figure 4.17B). At p*FIG1*, a strong enrichment of Kar4 is observed, peaking 30 minutes after stimulation (Figure 4.17B). This corresponds to the time at which *FIG1* transcripts level reached 85% of its maximal level (Figure 4.15B).

We also tested whether the binding of Kar4 to the DNA would require the presence of Ste12 or not. We found that in a *ste12*Δ background, there is no recruitment of Kar4 at either locus (Figure 4.16B and D). Moreover, we noticed that the shorter form of the protein Kar4,

the one induced by the mating, is absent in the *ste12*Δ, arguing in favor of a regulation of Ste12 for this short form (Figure 4.16E).

As said earlier, we observed that histone -1 was removed with a delay at *FIG1* locus relative to *AGA1* locus (Figure 4.9). In Chapter 2, we measured in real-time the noise in expression of two copies of the same promoter within the same cells, in response to hyper-osmotic stress. We hypothesized that the higher intrinsic noise observed for p*STL1* relative to p*GPD1* was due to the stochasticity of the chromatin remodeling process (Aymoz et al., 2016; Pelet et al., 2011). As we found that chromatin remodeling was also required for induction of p*AGA1* and p*FIG1*, we measured the intrinsic noise of two copies in the same cells (Figure 4.18, see Methods)(Elowitz et al., 2002). What we found is that the intrinsic noise in basal condition is at the same level for both promoters. Upon pheromone stimulation, the

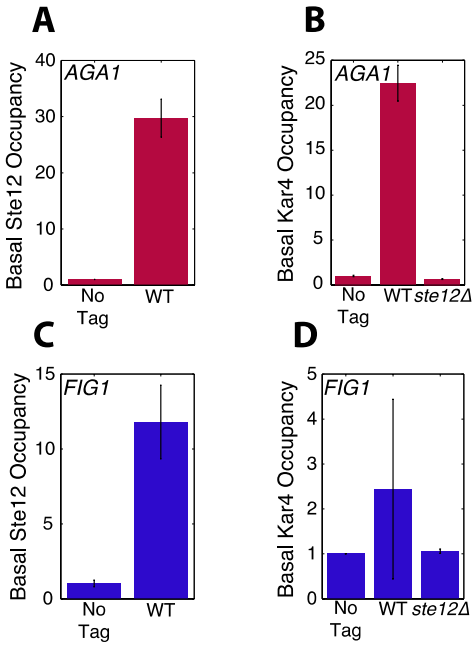
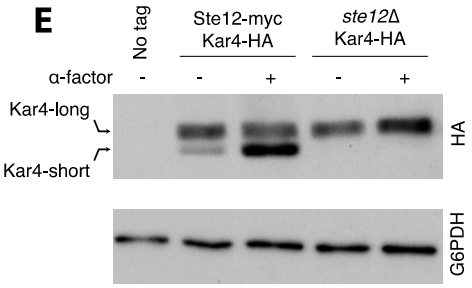


Figure 4.16 Basal recruitment of Ste12 and Kar4 at *FIG1* and *AGA1* promoters

A-D. Basal occupancy of Ste12 (A, C) and Kar4 (B, D) in the promoting region of *AGA1* (red, A, B) or *FIG1* (blue, C, D) assessed by ChIP and real-time PCR. Results are shown normalized to the untagged strain.

E. Western blot quantification of Kar4 levels in a WT background (left) or a *ste12*Δ background (right).



intrinsic noise of p*AGA1* increases during the 10 minutes following stimulation, when induction of this promoter starts (Figure 4.18A, Supplementary Figure 4). After this time point, the intrinsic noise decreases as the two copies of p*AGA1* get concurrently induced. For p*FIG1*, the intrinsic noise increases from 5 to 25 minutes following pheromone stimulation (Figure 4.18B). The maximum level reached is around 50%, which is higher than for p*AGA1*. After that, when the two copies of p*FIG1* are induced, the intrinsic noise decreases to reach levels as low as *AGA1*'s. This could reflect the stochasticity of chromatin remodeling required for the onset of transcription. It occurs later at *FIG1* compared to *AGA1*, and may require more proteins to be efficient.

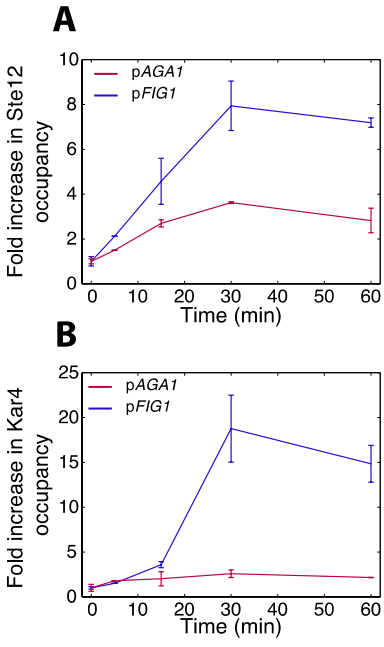


Figure 4.17 Recruitment of Ste12 and Kar4 at promoters in response to pheromone

Fold increase in Ste12 (A) and Kar4 (B) occupancy at p*AGA1* (red) and p*FIG1* (blue) compared to occupancy in the non-treated sample, measured by ChIP assays. All samples were normalized to an internal telomere control.

We also quantified the same metrics in a *kar4*Δ background. There is no big difference between WT and *kar4*Δ backgrounds regarding the intrinsic noise of *AGA1* (Figure 4.12A). The basal intrinsic noise for *FIG1* is lower in the mutant (Figure 4.18B). However, following pheromone stimulation, it increases from 20 to 40 minutes, when it reaches a plateau around 40% in the *kar4*Δ, whereas a minimal value in the WT is observed at the same moment. This could suggest that the absence of Kar4 increases the stochasticity of transcription initiation at p*FIG1*.

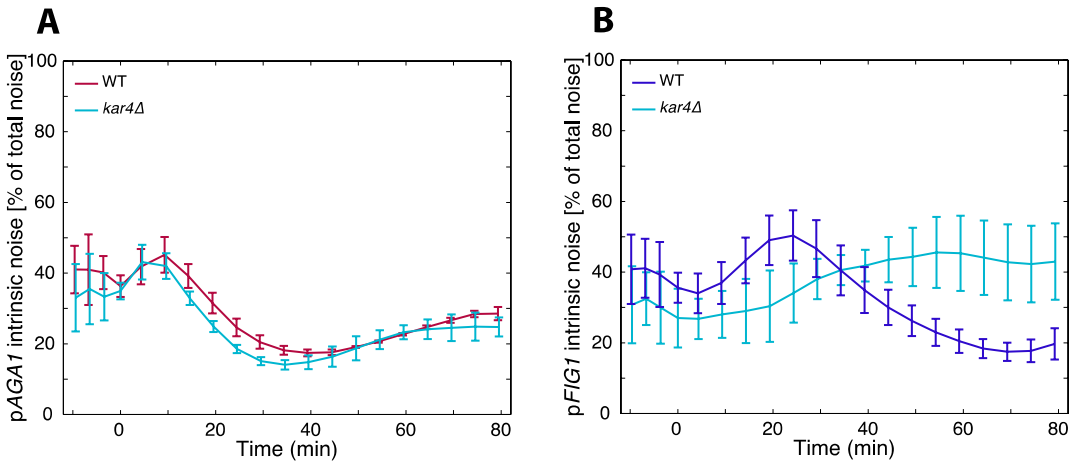


Figure 4.18 Real-time quantification of intrinsic expression noise in *AGA1* and *FIG1* expression

Evolution of intrinsic noise as percentage of total expression noise for p*AGA1* (A) and p*FIG1* (B) in a WT background (resp. red and blue) or *kar4*Δ background (cyan).

In summary, we propose that chromatin remodeling is a crucial process for the induction of *FIG1*. Moreover, it could require the presence of Kar4 to correctly occur and promote *FIG1* induction in the majority of cells.

In this part, we showed a few requirements for the pheromone-dependent induction of p*AGA1* and p*FIG1*. First of all, the signaling pathway needs to be intact, and the MAPK Fus3 activated. Secondly, Ste12 needs to be activated, and present at the two promoters. Thirdly, chromatin remodeling needs to take place to open the chromatin and allow transcription to occur. Finally, the presence of Kar4 is required for correct induction of p*FIG1*.

Part III: Promoter organization and impact on Ste12 transcriptional activation

Constraints on promoter organization and link to expression

Several studies aimed at determining how the architecture of an inducible promoter could control its expression (Hansen and O'Shea, 2015; Su et al., 2010; Weingarten-Gabbay and Segal, 2014). It was shown in these examples that the organization of Transcription Factor Binding Sites (TFBS) was submitted to tight constraints. With the example of the general stress TF Msn2, it was proposed that the levels and kinetics of genes expression was determined by only three factors: the number of TFBS, the nucleosome occupancy over these sites, and their distance to the TATA box (Hansen and O'Shea, 2015). In this part, we address the constraints on Ste12 BS organization at pheromone-inducible promoters in the regulation of the expression kinetics.

Regulatory elements of pheromone-induced genes

Ste12 is able to bind to Pheromone Responsive Elements (PRE) (Dolan et al., 1989; Errede and Ammerer, 1989; Su et al., 2010; Zeitlinger et al., 2003). A consensus sequence has been identified as 5'-ATGAAACA-3' (Kronstad et al., 1987; Van Arsdell et al., 1987). Most of pheromone-induced genes possess these sequences (Chou et al., 2006; Zeitlinger et al., 2003).

Pheromone Response Elements presence is necessary and sufficient to promote gene induction in a pheromone-dependent manner (Hagen et al., 1991). Already in 1991, a few rules for Ste12 DNA binding had been established. Through a β-galactosidase assay, it was found that a single PRE was sufficient to increase the expression of a gene in response to pheromone, but that the expression levels were much higher with multiple repeats of PREs. It was also noticed that multiple PREs with head-to-tail orientations, in the forward direction, were leading to higher expression than the same numbers of PRE in the reverse direction (Hagen et al., 1991). More recently, a thorough study described organizational constraints on PRE for efficient pheromone-induced transcription (Su et al., 2010). The authors assessed the importance of different characteristics of the PREs, such as their sequences, orientation, and relative distances. They assessed the affinity of Ste12 for the PREs *in vitro*, and found that any mutation within the 8 nucleotides of the consensus impairs Ste12 affinity for this sequence. In particular, mutations in the central AAA were highly deleterious for interaction with Ste12.

Another interesting result is that the combination between a consensus PRE and mutated PREs of various decreased affinities for Ste12 allows a pheromone-dependent induction (Su et al., 2010). Nevertheless, the expression level reached will decrease with the affinity of Ste12 for the non-consensus PRE. They also propose that the optimal distance between two PREs for pheromone-dependent induction is 40nt, when these PREs are in the head-to-head or tail-to-tail orientation. However, two PREs in the same orientation are able to promote expression only if they are separated by 3 nucleotides. Interestingly, this configuration would place the 6 core nucleotides of the two PREs on the same side of the DNA helix (Figure 4.19). This configuration could allow two Ste12 molecules binding each PRE to face each other and interact together. However, if the distance between these PREs is increased to 6nt, the two binding sites are on opposite sides of the DNA, which could prevent the two Ste12 to interact together. In conclusion, they proposed a model where Ste12 can form dimers on the DNA only if the PREs are in a certain configuration, enabling the two proteins to face each other, or bend the DNA to interact. This very interesting study could nevertheless be improved. First of all, all experiments were based on β-galactosidase assays, which are sensitive and only informative

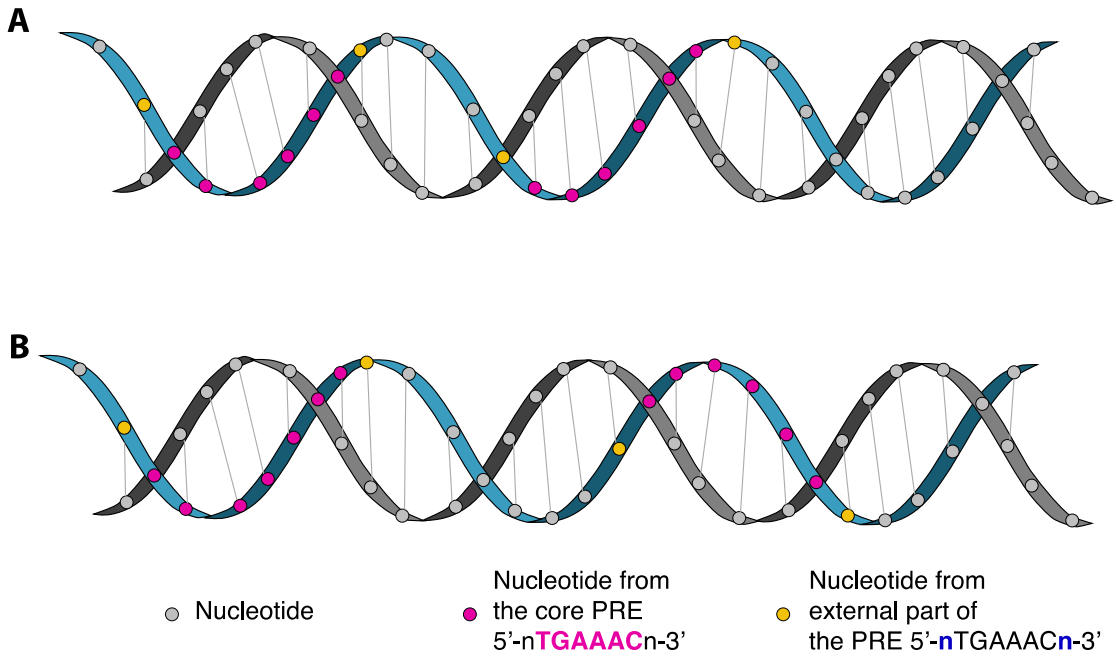


Figure 4.19 Representation of pairs of PREs on the DNA double helix

A. Two PREs in the same direction and separated by 3nt are located on the same side of the DNA helix.

B. If the distance is increased to 6nt, the sites are no longer on the same side of the DNA helix.

of the time point that was selected. Moreover, some of the general rules were actually established in a context where other PREs were present, like in the promoter of *FUS1*.

Yet, this study sets the grounds for our investigations. In this chapter, we systematically mutated the PREs directed by the two promoters p*AGA1* and p*FIG1* and monitored the modification in the expression kinetics and expression levels of these promoters in response to pheromone stimulation. By using the dPSTR, we managed to obtain information on the role of PREs in the kinetics of expression.

Part IV: Promoter architecture as a determinant for expression kinetics

Regulatory elements in AGA1 and FIG1 promoters

We mapped Ste12 binding sites present in the promoting regions of *FIG1* and *AGA1* (Figure 4.20). We defined two types of PREs, according to their sequences and their closeness to the consensus. The sequences of what we define consensus PREs are 5'-nTGAAACn, whereas the non-consensus PREs (PRE^{nc}) can have one additional mutation within the 6 core nucleotides.

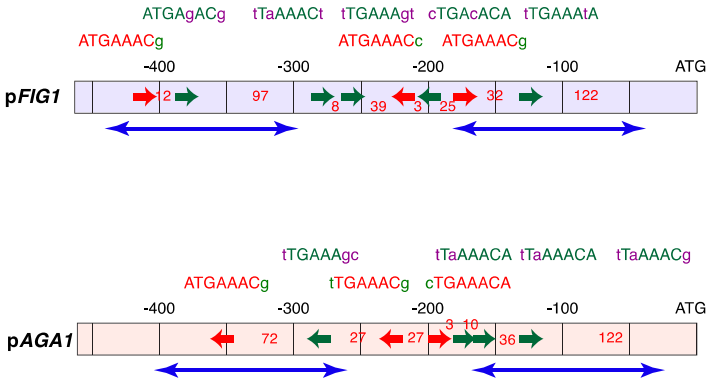


Figure 4.20 Maps of p*AGA1* and p*FIG1*

Positions of the PREs on p*FIG1* (top) of p*AGA1* (bottom). The red arrows represent the location and orientation of consensus Ste12-binding sites (5'-nTGAAACn-3'). The green arrows symbolize the non-consensus binding sites that possess mutations within the 6 core nucleotides of the PREs. The sequences of each binding site is detailed above, in the corresponding color, with capital nucleotides matching the consensus sequences and small nucleotides being mutations from the consensus. The red numbers between sites represent the distance in bp between them or with the ATG. Blue arrows represent nucleosome positions obtained from Brogaard *et al.* (Brogaard et al., 2012).

The two promoters possess 3 consensus PREs, in very similar orientations. The two closest from the ATG are in a tail-to-tail orientation, and this for both promoters. However, the distance between these two PREs is different: 36nt for p*FIG1* against 27nt for p*AGA1*, both different from the optimal distance defined by Su *et al.* (40nt) (Su et al., 2010). However, the furthest PREs from the ATG are in opposite orientations. Also, both p*AGA1* and p*FIG1* possess some PRE^{nc}, 4 for p*AGA1* and 5 for p*FIG1* in various orientations and positions. Nevertheless, there are no obvious features here that would determine whether they would be expressed with different kinetics or not. As we shown in Figure 4.9 that histone remodeling was occurring at different times, we added the nucleosome positions to our maps (blue arrows in Figure 4.20) (Brogaard et al., 2012). This revealed a different topology for the two promoters. Indeed, the presence of histones may mask most of the PREs on both promoters. However, two consensus PREs and a PRE^{nc} are still available at p*AGA1*, whereas only one PRE and three PRE^{nc} are free at p*FIG1*. This could explain the different kinetics of induction of the two promoters. Indeed, we can imagine that Ste12 is already present at p*AGA1* locus, in the form of a dimer, bound to the two consensus sites. Pheromone signaling would result in activation of this dimer, by phosphorylation and detachment of Dig1, leading to fast recruitment of the polymerase complex. At p*FIG1* however, we could imagine that Ste12 binding could only occur at both consensus sites after histone removal, which occurs later at p*FIG1* than at p*AGA1* (Figure 4.9). However, Su *et al.* suggested that a PRE separated by 3nt from a PRE^{nc}, both in the same

orientation, are sufficient for transcription induction (Su et al., 2010). This is the configuration that we observe in p*FIG1*. We can imagine that dimerization can occur on these two sites in presence of an additional factor, which could be Kar4 that we identified as required for p*FIG1* induction. Thus, we can also imagine that histone removal can be a consequence of Ste12 dimer recruitment at PRE- PRE^{nc} sites.

To test all these hypotheses, we systematically deleted all consensus PREs from both promoters, either alone or in combination with each other.

Construction of synthetic promoters

In order to easily test multiple constructs, we designed a new set of dPSTR plasmids in which we added unique restriction sites surrounding the PRE-containing regions of both promoters. Indeed, the classic dPSTR plasmid requires two rounds of cloning to exchange the measured promoter (Figure 4.21; see Methods). With this strategy, we can obtain a dPSTR measuring a promoter with a different PRE-containing region in one cloning experiment. For p*AGA1*, the total construct is the 1kb upstream of the start codon. However, the previous ORF ends 640nt before the ATG. We placed a first ApaI cloning site (GGGCCC) 513-518nt before the start codon, by mutation of 4nt. The second site, ClaI (ATCGAT), was placed 151-144nt before the start codon also by mutations of 5 nucleotides. This strategy allows to maintain the positions of the sites, that would have been displaced by simply add restriction sites. This way, the PRE-containing region is 363nt long, and surrounded by ApaI-ClaI restriction sites. With the same strategy, we obtained a synthetic version of p*FIG1* that possesses a 300nt-long PRE-containing sequence, flanked by ApaI and ClaI sites.

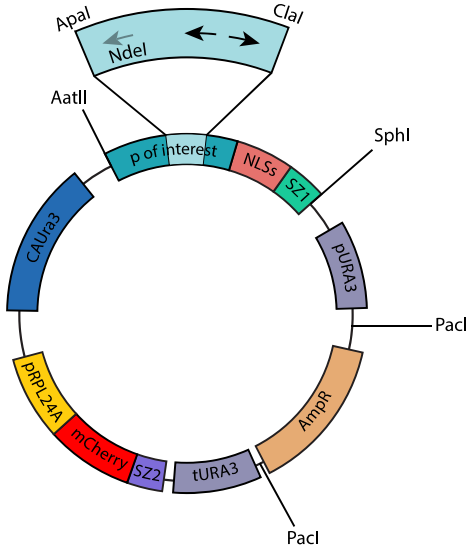


Figure 4.21 Map of the dPSTR plasmid for promoter-variant analysis

The PRE-containing region of each promoter was artificially flanked by two unique cloning sites. Variant promoters were synthesized and clone in these sites.

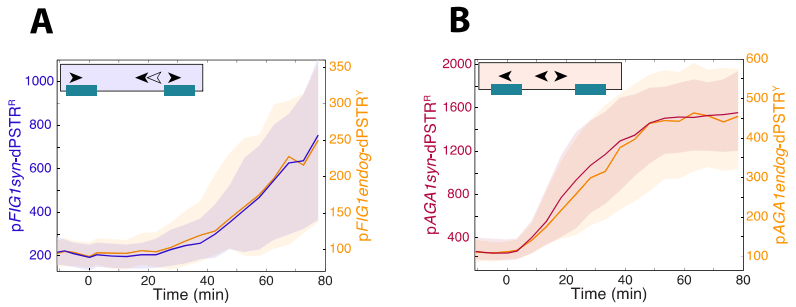


Figure 4.22 Addition of cloning sites does not impair the induction of the promoters

Nuclear enrichment for the dPSTR^R (left axis) measuring endogenous promoters, compared to the dPSTR^Y (orange, right axis) quantifying the synthetic version of the promoter in the same cells, for either p*FIG1* (A, blue) or p*AGA1* (B, red).
Schematic maps of the promoter structure are represented in top left of each plot. Black arrows symbolize PRE, and empty arrows PRE^{nc}. Green lines mark the emplacements of histones -2 and -1, hiding some PREs. These schemes are used throughout the part to symbolize the variants measured. For all similar graph, medians are plotted and shaded areas represent 25th-75th percentile, and 1µM of α-factor was applied at time 0 to stimulate cells.

We verified that these constructs behaved the same as the endogenous promoters we used so far, by measuring them with the dPSTR^R in combination with the endogenous promoters measured by dSPTR^Y (Figure 4.22). We found that the two dPSTRs curves are superimposed, showing that the addition of the restriction sites does not modify the kinetics of induction of either p*AGA1* or p*FIG1*.

Our strategy to delete PREs was actually to mutate them into a restriction site far from the PRE consensus sequence, so that we can easily screen our constructs by digestion. Moreover, this strategy allows, as mention above, to avoid any perturbation of the rest of the DNA. Indeed, a deletion of 8nt will completely change the relative positions of all other putative sites on the DNA helix. As the DNA molecule makes one turn every 10.5nt, a deletion could impair cooperative binding because sites could be positioned on the opposite side of a DNA turn. We chose to mutate consensus PREs into NdeI sites (CATATG) when mutated alone, or into SnaBI sites (TACGTA), when a second PRE was mutated.

Variants of the AGA1 promoter

The loss of the consensus PRE hidden by the -2 nucleosome is not deleterious for p*AGA1* induction (Figure 4.23A). However, loss of either consensus PRE between the two histones causes a delay in the induction, although expression level is the same as for the endogenous p*AGA1* (Figure 4.23B-C). This supports the hypothesis that a dimer of Ste12 could form here. However, transcription can still occur quite fast, and we can imagine that the presence of the free PRE^{nc} plays a role in Ste12 recruitment. Then, we deleted more than one PRE at the time. Combined deletion of the free reverse PRE with the hidden PRE results in a stronger delay than deletion of the reverse PRE alone (Figure 4.23D, compared to 23B). In this construct, one consensus PRE separated by 3nt from a PRE^{nc} in the same reverse orientation is still present and free from the nucleosomes in basal conditions. As shown by Su *et al*. this configuration is still able to allow pheromone-dependent transcription, although we show here that the kinetics of induction is strongly impaired (Su et al., 2010). In fact, the induction of this variant is highly similar to p*FIG1* induction kinetics. However, the expression level is still as high as for the endogenous p*AGA1*. Combined deletion of the two consensus PREs between the two histones strongly impairs the expression of p*AGA1* (Figure 4.16E). Deletion of the hidden consensus PRE combined to deletion of the free forward PRE does not impair much the expression dynamics of p*AGA1* (Figure 4.23F). Finally, we deleted all consensuses PRE from p*AGA1* (Figure 4.23G). This results in a complete loss of pheromone-dependent induction,

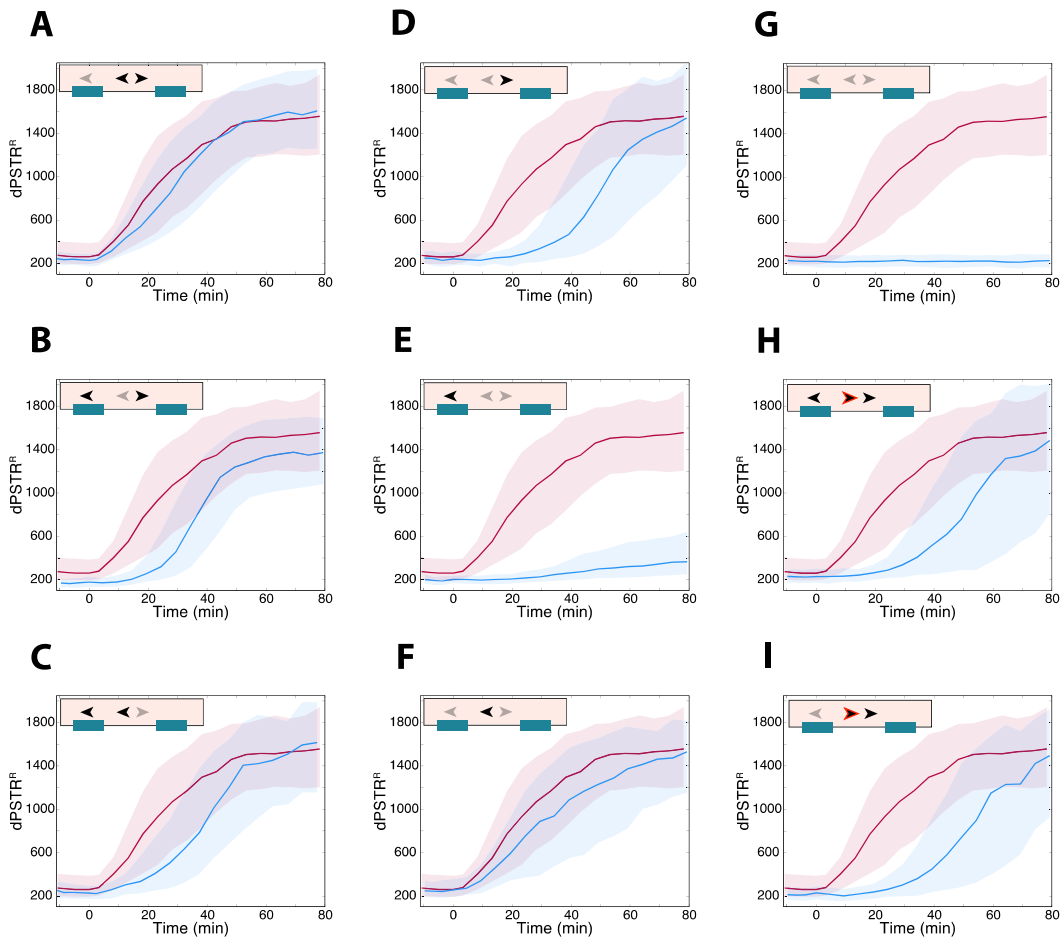


Figure 4.23 dPSTR measurements of the induction of p*AGA1* variants

Nuclear enrichment of the dPSTR^R measuring the indicated variant of p*AGA1* (blue curves), compared to the p*AGA1synth*-dPSTR^R (red) as a reference. Shaded arrows on the schematic promoter maps represent the mutated PRE.
A-C. Single PRE mutations.

D-F. Double PRE mutations.

G. Deletion of all consensus PREs of p*AGA1*.

H-I. Swapping of the free reverse consensus PRE into the forward direction.

suggesting that for p*AGA1*, the presence of at least one consensus Ste12 binding site is required for pheromone-dependent induction.

We also tried to reverse the orientation of the free reverse PRE, so that the two PREs between the nucleosomes would be in the forward orientation (Figure 4.23H-I). This leads to a delay in the induction of p*AGA1*, which has similar kinetics to p*FIG1*. We further deleted the hidden PRE in this construct, to verify that it was not playing a role in the induction here. Once again, these constructs possess two consensus PREs, in the same orientation (forward), separated by 27nt, which is not the optimal distance of 3nt described by Su *et al.* (Su et al., 2010). However, they also possess an additional PRE^{nc} 3nt away from a consensus one, both in the forward orientation. This situation is highly similar to p*FIG1* architecture, although the orientation is the opposite and they show similar induction kinetics.

To summarize, we were able to relatively easily slow down the induction of p*AGA1* to p*FIG1* level.

Variants of the pFIG1 promoter

Deletion of any of the two hidden consensus PREs in the promoter of *FIG1* results in a decrease of the expression level, although the kinetics of expression remain similar (Figure 4.24A and B). However, deletion of the free consensus PRE completely abolishes the induction of p*FIG1*, highlighting the importance to have at least one free consensus PRE (Figure 4.24C). We then deleted the two hidden consensus PREs, and found that this variant is still induced by pheromone, although to a lower level, as expected from Su *et al.* conclusions (Figure 4.24D). These results show that the free consensus PRE of p*FIG1* is of high importance for its expression.

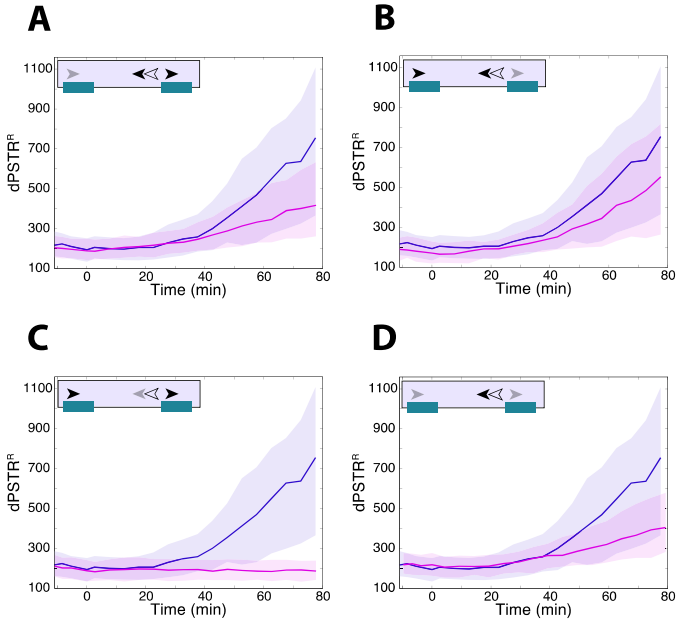


Figure 4.24 dPSTR measurements of the induction of variants of p*FIG1*

Nuclear enrichment of the dPSTR^R measuring the indicated variant of p*FIG1* (pink curves), compared to the p*FIG1synth*-dPSTR^R (blue) as a reference. Shaded arrows on the schematic promoter maps represent the mutated PRE.
A-C. Single PRE mutations.

D. Double PRE mutations.

We then assessed the importance of the free PRE^{nc} that is centered around -200. We found that p*FIG1* cannot be induced in absence of the PRE^{nc}, similarly to deletion of the free consensus PRE (Figure 4.25A and B). This supports the hypothesis that a dimer of Ste12 could form between these two available PREs separated by 3nt at p*FIG1*, leading to its expression. Without one of these sites, no dimer can be recruited, and transcription cannot occur. As Su *et al.* proposed that 3nt was the optimal distance in this configuration, we slid the PRE^{nc} away

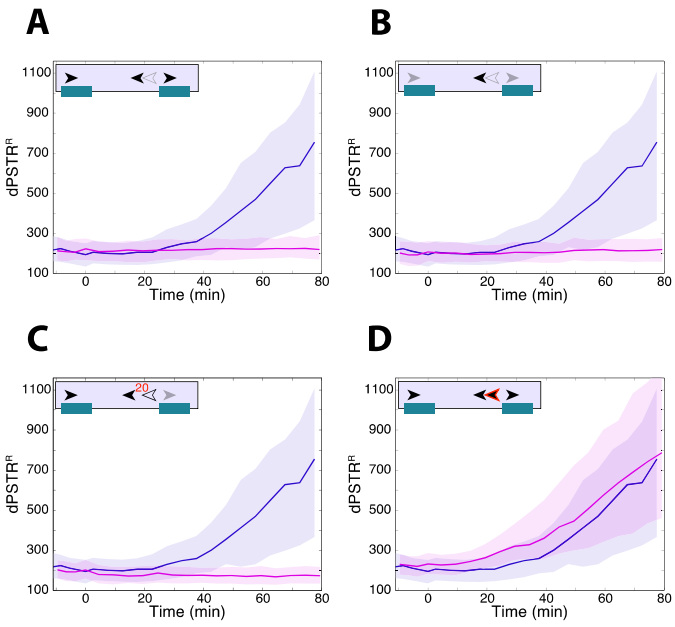


Figure 4.25 dPSTR measurements of the induction of variants of p*FIG1* PRE^{nc}

Nuclear enrichment of the dPSTR^R measuring the indicated variant of p*FIG1* (pink curves), compared to the p*FIG1synth*-dPSTR^R (blue) as a reference. Shaded arrows on the schematic promoter maps represent the mutated PRE.
A, B. Deletion of the free PRE^{nc} alone (A) or in combination with the two hidden PREs (B).

C. Increase of the distance between the free PRE and PRE^{nc} from 3 to 20nt.

D. Mutation of the PRE^{nc} into a consensus PRE.

from the free consensus PRE, to separate them by 20nt (Figure 4.25C). As expected, the expression is lost.

As Ste12 affinity of the PRE^{nc} is decreased compared to a consensus site, we turned the PRE^{nc} into a consensus PRE by mutation of the 5'-cTGAcACA sequence into 5'-cTGAAACA (Figure 4.25D). By doing this, we were hoping to promote recruitment of a Ste12 dimer in basal condition, leading to a faster recruitment of the polymerase in response to pheromone. However, this variant is induced only slightly before p*FIG1*, and to similar levels. This could suggest the requirement of an additional factor for p*FIG1* expression, even if Ste12 is recruited. One could imagine that this factor may be absent until 20-30 minutes following stimulation, preventing a faster induction of p*FIG1*.

Chimeric promoters

Segal and colleagues established a model showing that the organization of nucleosomes genome-wide is partly predictable by the DNA sequence (Field et al., 2008; Kaplan et al., 2009; Segal et al., 2006). Indeed, several DNA features were identified as either favorable or unfavorable to nucleosome binding. For instance, the presence of tracts of adenosine has been shown to prevent nucleosome formation, probably because this changes the biochemical properties of the DNA, which cannot be bent around them (Field et al., 2008). With this in mind, we can imagine that some nucleosomes positions can be encoded in the DNA, especially those important for transcriptional regulation.

As mentioned above, *AGA1* is induced during vegetative growth, whereas *FIG1* is not expressed in absence of pheromone. We looked for poly(dA:dT) tracts in the promoters of *FIG1* and *AGA1* and found an imperfect tract of 38nt long located between the transcription start site and the start codon of *FIG1*. Moreover, the last 18nt preceding the start codon are a perfect tract of adenosines. This specific sequence is expected to cause a 10-fold nucleosome depletion (Field et al., 2008). This led us to think that the position of the -1 nucleosome may be regulated in order to hide the last consensus PRE of *FIG1*, and prevent its inappropriate expression during vegetative growth or too early in the mating process.

We wondered whether by exchange of the last 150nt of each promoter we could somehow modify the stability or position of the -1 nucleosome on both promoters. To do so, we built chimeric promoters. Because of the way we designed the variant plasmids, we could easily build chimeras between p*FIG1* and p*AGA1*, between the PRE-containing sequence and the last 150nt of each promoter (Figure 4.21).We found that these constructs have intermediate expression kinetics (Figure 4.26). The chimera with *AGA1* PRE-containing region is expressed before the other chimera and p*FIG1*, but after *AGA1*. This means that the last 150nt of p*FIG1* are delaying the induction of the promoter, or that the end part of p*AGA1* is responsible for the fast induction kinetics. The other chimera supports the same hypotheses, as the loss of the last 150nt of p*FIG1* (or gain of the last 150nt from p*AGA1*) leads to a faster induction. We could interpret these results in term of nucleosome positioning. If the -1 nucleosome is indeed encoded in the last 150nt, and at p*FIG1* it negatively regulates transcription, this negative effect can be transmitted to a different locus, as we did with the chimeric construct. On the opposite, the -1 nucleosome of p*AGA1* may be in a looser conformation, as this promoter has a basal level of induction. Supporting this, we

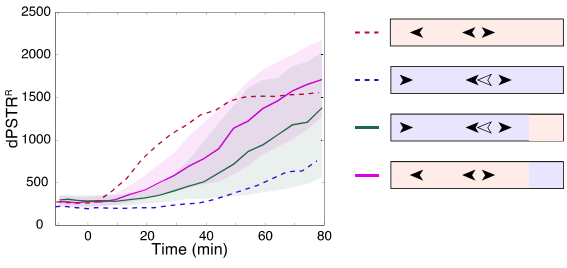


Figure 4.26 dPSTR measurement of the induction of chimeric promoters

Nuclear enrichment of the dPSTR^R measuring p*AGA1* (red dotted line), p*FIG1* (blue dotted line), or the chimeras p*FIG1_AGA1* (green line) or p*AGA1_FIG1* (pink line).

found that the two chimeras are expressed in 22% of the cells prior to stimulation, compared to ~30% for p*AGA1*.

Involvement of Kar4 in the induction of the variants

To get better insights in Kar4 importance and role in mating-induced transcription, we deleted this TF in several variant constructs of interest.

First of all, we found that only one of the two chimeric constructs requires Kar4 for correct induction. Indeed, the chimera composed of the PRE-containing region from p*FIG1* and the last part of p*AGA1* has an impaired expression in *kar4*Δ background (Figure 4.27A). This suggests that Kar4 requirement for p*FIG1* expression occurs within the PRE-containing region, where Ste12 is expected to interact. The last 150nt of p*FIG1* do not confer Kar4-dependence to p*AGA1*, again supporting the idea that Kar4 needs to interact with Ste12 (Figure 4.27B).

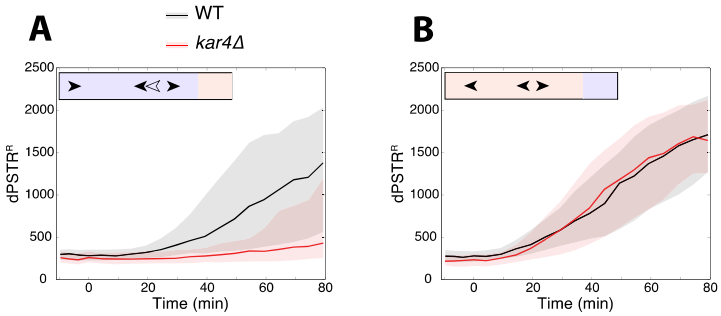


Figure 4.27 Expression of the chimeras in a *kar4Δ* background

Nuclear enrichment of the dPSTR^R measuring the chimeras p*FIG1_AGA1* (A) or p*AGA1_FIG1* (B) in either a WT (black line) or *kar4*Δ (red line) background.

Interestingly, we managed to get some variants of p*AGA1* that were dependent on Kar4. Indeed, variants in which we expect the free reverse consensus PRE to be non-functional (either mutation or directional swap), now show a requirement for Kar4 (Figure 4.28A and B, respectively). This is in agreement with the fact that the expression of these variants is delayed, as we described above. In the absence of the free reverse consensus PRE, only a forward consensus PRE and a forward PRE^{nc} separated by 3nt are available. We can hypothesize that in this configuration, Ste12 efficient binding may require the presence of Kar4.

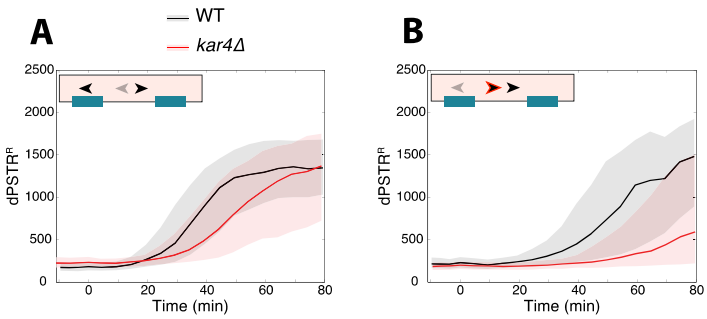


Figure 4.28 Expression of p*AGA1* variants in a *kar4Δ* background

Nuclear enrichment of the dPSTR^R measuring the variants of p*AGA1* deleted for the free reverse PRE (A) or deleted for the hidden reverse PRE in combination with a swap of the free reverse PRE (B) in either a WT (black line) or *kar4*Δ (red line) background.

As p*FIG1* was initially Kar4-dependent, we wondered whether we could obtain some variants that would not require Kar4 presence, as for p*AGA1*. The best candidate for this is the variant in which we mutated PRE^{nc} into a consensus site, as we assumed the binding of Ste12 could occur more easily (Figure 4.25D). Interestingly, this promoter variant, although still late-induced, does not require the presence of Kar4 (Figure 4.29).

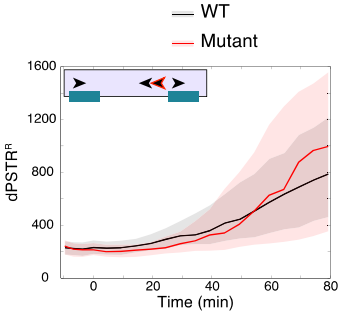


Figure 4.29 Expression of a p*FIG1* variant with 4 PREs in a *kar4Δ* background

Nuclear enrichment of the dPSTR^R the variant of p*FIG1* where the free PRE^{nc} was mutated into a PRE in either a WT (black line) or *kar4*Δ (red line) background.

This supports the hypothesis that Kar4 could provide mechanistic help to Ste12 for its dimerization.

An alternative role for Kar4 can be the recruitment of the chromatin remodeling machinery. However, in the p*FIG1* variant with 2 free consensus PREs, the chromatin remodeling machinery recruitment is still occurring in absence of Kar4, as we observe expression, although the noise in expression is higher, suggesting a stochastic process in the transcription establishment. An alternative suggestion is that Ste12 stable dimer can recruit itself the chromatin remodeling machinery, and that Ste12 dimerization may require Kar4 at weak binding sites. In the cases where Ste12 dimer cannot form appropriately, there could be a failure in histone removal, which would block the transcription, as we observed.

To test this hypothesis, we built a chimeric construct between the PRE-containing region of p*FIG1*, in which the PRE^{nc} was mutated into a consensus site, and the last part of p*AGA1*. We assume that in this construct, the negative effect from the last 150nt of p*FIG1*, probably coming from the histone -1 positioning, is absent. Moreover, we expect an easy Kar4-independent dimerization of Ste12 in this construct. Altogether, this chimera may reach faster expression kinetics, maybe close to p*AGA1*'s. We found that this promoter is expressed slightly after p*AGA1*, and long before p*FIG1*. Moreover, it does not depend on Kar4 to be expressed (Figure 4.30).

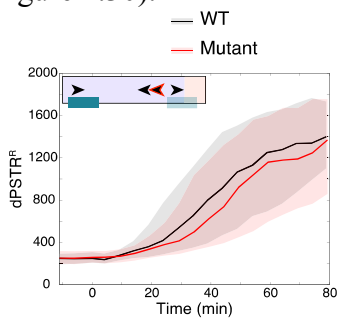


Figure 4.30 Expression of the chimeric promoter with 4 consensus PREs

Nuclear enrichment of the dPSTR^R measuring the chimeric promoter composed of the PRE containing region of p*FIG1* where the free PRE^{nc} was mutated into a consensus PRE, and the last 150nt of p*AGA1* in either a WT (black line) or *kar4*Δ (red line) background.

Overall, we managed to obtain some variants of p*AGA1* that became dependent on the presence of the transcription factor Kar4, although to a lower extent than p*FIG1* (Figure 4.31A). These variants possess only the pair PRE-PRE^{nc} separated by 3nt that is free from nucleosomes. This supports our hypothesis that the Ste12 dimer needs Kar4 support to bind to this configuration of PREs. We also managed to obtain some variants of the promoter p*FIG1* that lost the dependency to Kar4 (Figure 4.31B). These are mutants in which we mutated the PRE^{nc} into a consensus site, and that are expected to have 2 consensus PREs free from any nucleosome.

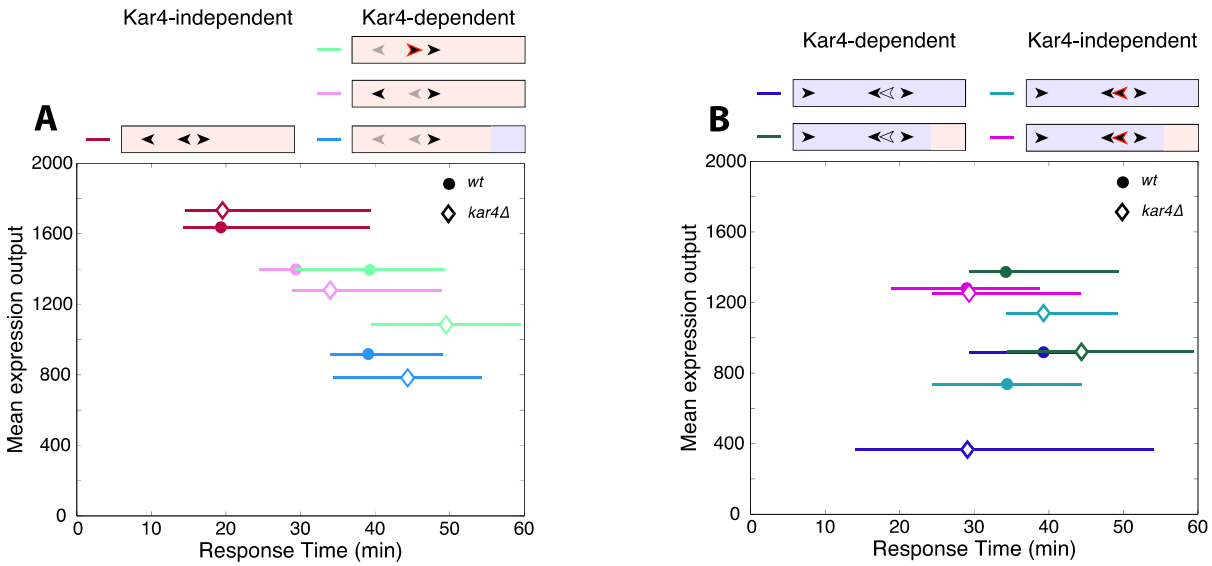


Figure 4.31 Summary of the variants expression in WT and *kar4Δ*

Correlation of the mean expression output versus the response time for the indicated construct, derived from p*AGA1* (A) or p*FIG1* (B), in a WT (dot) or a *kar4*Δ background (diamonds).

To summarize this fourth part, by modifying p*AGA1* and p*FIG1* binding sites for Ste12, we managed to understand some rules governing their transcription. We first learnt that we are able to change their expression kinetics by simply either adding or suppressing PREs. This means that the PRE landscape is what actually controls the expression kinetics. We also gain in understanding the role of Kar4, that we think helps to stabilize dimer formation at non-consensus sites.

Discussion

In this chapter, we tried to understand which features are determining the expression time of p*AGA1* and p*FIG1*. By using the dPSTR system, we showed that these two promoters are induced with different delays following pheromone stimulation. There is an average delay of 23 minutes between the induction of p*AGA1* and the expression of p*FIG1* in single cells (Chapter 3). In order to understand how the mating signaling that occurs within minutes following stimulation could be converted into temporal information, we first mutated a number of factors potentially involved.

We identified Ste2, Ste5, Fus3, Far1, Ste12 and Kar4 as key actors in this process. Indeed, mating-induced transcription requires an intact signaling pathway and its activation through the pheromone receptor Ste2. The TF Ste12 is absolutely required for the expression of both promoters. Ste12 requires MAPK activity to be de-repressed and recruit polymerase. We showed that in absence of the mating MAPK Fus3, Kss1 activity is sufficient to promote the induction of *AGA1*, but not *FIG1*. We showed previously, using the SKARS, that the activity of Kss1 is much lower than the activity of Fus3 in response to pheromone stimulation (Durandau et al., 2015). We can relate this low activation to stimulation with a lower concentration of pheromone. Moreover, we can imagine that the activation of Kss1 maybe transient. If we assume that all early genes behave like p*AGA1*, a low level of MAPK activation will be sufficient to induce proteins involved in downregulation of the mating signaling, such as the protease Bar1, and Sst2, which promotes inhibition of the G proteins coupled to the pheromone receptor. This would lead to faster attenuation of the signal, and deactivation of the MAPK Kss1, eventually leading to repression of Ste12. Moreover, the phosphatases involved in its deactivation should be more available to target Kss1, as Fus3 is absent. Depending on the kinetics of these events, this could happen before p*FIG1* expression has occurred, thereby preventing it. We also found that Far1 presence was required for induction of *FIG1*. This can be explained by the fact that as the cells fail to arrest in G1 and cycle through S phase, the MAPK will be deactivated due to repressing mechanisms in S phase (Strickfaden et al., 2007) (Durandau, personal communication). Hence, levels of active MAPK are too low to allow *FIG1* induction. Finally, we found that the transcription factor Kar4 was also required for induction of *FIG1* but not *AGA1*. This transcription factor is itself induced in response to pheromone, at the same time as p*AGA1*. The protein Kar4 exists in two forms, the long form induced during G1 phase, and the short form induced in response to pheromone signaling (Brizzio et al., 1998; Gammie et al., 1999; Kurihara et al., 1996). The short form is expressed from an alternative start codon 87nt downstream from the regular start codon, resulting in a protein shorter by 29 amino acids. We showed that its induction is dependent on the presence of Ste12 during pheromone response. We also showed that Ste12 and Kar4 are interacting together *in vivo*, and are both present at the promoting regions of *FIG1* and of *AGA1* in presence and in absence of pheromone.

In a second set of experiments, we looked in details at the architecture of the two promoters in order to understand how the temporal information could be encoded in the DNA. We systematically mutated all Ste12 binding sites on both promoters, and further deleted Kar4 in some interesting variants. We placed the Ste12 binding sites in the context of the chromatin. Indeed, the presence of histone can mask some binding sites. We found that the two promoters have different numbers of available binding sites in different conformations. p*AGA1* possesses two consensus PREs in opposite orientations, that can be bound by Ste12 even in basal conditions. p*FIG1* only possesses one consensus PRE very close to a non-consensus PRE, both in the reverse orientation, available in basal conditions. We hypothesize that in this configuration, Ste12 dimerization can only occur in presence of Kar4, as stabilizer of the

complex. Favoring this hypothesis, a p*FIG1* variant where the PRE^{nc} was mutated in a consensus site lost its requirement for Kar4.

We also assessed the kinetics of histone -1 displacement at the two endogenous loci. We found that it occurred 10-25 minutes later at *FIG1* than at *AGA1*. However, the strain from this experiment is the one used for the immunoprecipitation experiments, and the addition of tags to Kar4 and Ste12 impairs the expression of both *FIG1* and *AGA1*. Ideally, we should perform this experiment in the untagged strain, and compare the kinetics with what we measure with the dPSTR.

Taken all together, we can propose a model of p*AGA1* and p*FIG1* regulation. Upon pheromone stimulation, Ste12 is quickly derepressed by Fus3- and Kss1-dependent phosphorylation of Dig1/Dig2 repressors. Ste12 dimers can be quickly recruited, if not already present, at early-induced promoters that possess free pairs of PREs. The chromatin is remodeled in order to displace histone -1, and transcription can start. This entire process should happen within the 10 minutes following pheromone addition. At the same time as p*AGA1*, p*KAR4* is induced, resulting in the accumulation of the short form of Kar4. This helps to promote dimerization of Ste12 at *FIG1* promoter, at the pair of PRE-PRE^{nc} between the histones -2 and -1. The stable dimer of Ste12 drives recruitment of the chromatin remodeling machinery, which displaces histone -1, allowing transcription to occur (Figure 4.32).

This model, drawn from the experiments presented in this chapter, needs more investigations to be tested and completed.

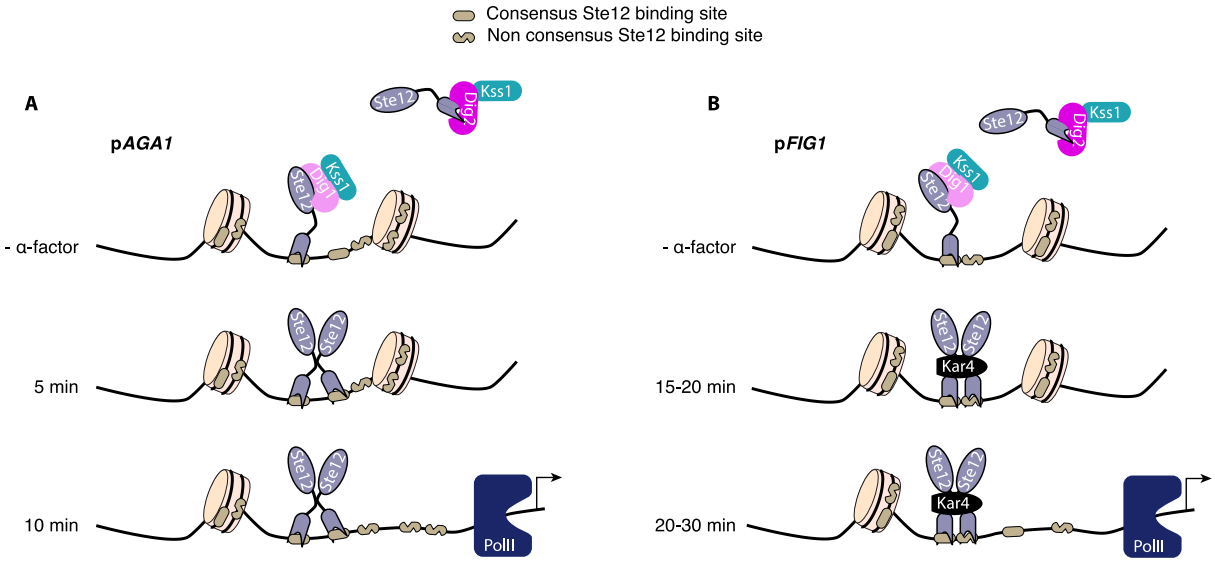


Figure 4.32 Model of the regulation of p*AGA1* and p*FIG1*

Schematic representation of the Ste12 binding sites and nucleosome positioning on p*AGA1* (A) and p*FIG1* (B). Addition of pheromone triggers Ste12 recruitment and RNA Pol II recruitment at different times.

First of all, the fact that tagging Kar4 or Ste12 impairs *AGA1* and *FIG1* expression is problematic to understand what is really happening in terms of dynamics and chronology of the events. Indeed, we could imagine that the presence of the tags can hinder protein-protein interaction, or promote interaction that is not happening in WT conditions. Ideally, antibodies directly targeting Ste12 and Kar4 should be used. Esch *et al.* reported the existence of an antibody recognizing Ste12, but did not use it throughout their study because low amounts were available (Esch et al., 2006). To our knowledge, no antibody targeting Kar4 is available.

Then, the MNase assay should be performed in an untagged strain, for the same reasons as explained above. As all the measurements provided in this study are made with a dPSTR plasmid inserted in the *URA3* locus, a control experiment could verify that the histone

positioning is similar in the *URA3* locus and at the endogenous loci. For this purpose, we built two strains where the endogenous promoter and ORF of the two genes were deleted, and the corresponding dPSTR plasmid inserted at the *URA3* locus. MNase experiments are ongoing to map histones positions on our constructs. Moreover, these strains could be used to assess the histone positions in the chimeric constructs, which would help to provide insights on the ability of the sequences to predict nucleosome positioning.

It would also be interesting to perform more experiments in the *kar4*Δ background. Indeed, we could verify whether the binding of Ste12 is decreased compared to a WT strain in absence of Kar4 to presumably stabilize it. Moreover, we could address whether the histone displacement is occurring or not. These two assays should be performed both at the endogenous p*FIG1* and on a mutant with PRE^{nc} mutated to consensus, both in WT and *kar4*Δ background. We could also monitor Kar4 recruitment at the mutated p*FIG1* in order to see whether it still binds and to what extent. This would allow to improve our model. Indeed, if chromatin remodeling is occurring at the mutated p*FIG1* in absence of Kar4, and Ste12 is more present, as a dimer, it would strongly suggest that Kar4 is required at p*FIG1* to stabilize Ste12 binding to a PRE-PRE^{nc}.

Alternatively, *in vitro* experiments could be performed to assess whether Kar4 helps Ste12 binding to a PRE-PRE^{nc} pair. Such *in vitro* assays have already been performed, in different contexts (Lahav et al., 2007; Su et al., 2010). This could also address the question of whether Kar4 is required for Ste12 to bind DNA or not. Indeed, we found the presence of Kar4 at both loci, even though *AGA1* induction is not impaired by the absence of Kar4.

If we could confirm the interaction of Kar4 and Ste12 at PRE-PRE^{nc} pairs, we could also address the stoichiometry of these interactions. Through *in vitro* binding assays, Chou *et al.* showed that Ste12 could bind with Dig1 and Dig2 with a 1:1.5:1.4 ratio (Chou et al., 2006). It would be interesting to know whether Kar4 can interact with monomers or multimers of Ste12. We could also try to see whether Kar4 can interact with the Digs repressors or not. Indeed, as a long form of Kar4 is produced during vegetative growth (Gammie et al., 1999; Kurihara et al., 1996), it can be present in repressed Ste12 complex in absence of pheromone. More interestingly, the question of the role of having two forms of Kar4 has been asked, and it was found that they both allow to rescue mating in *kar4*Δ cells if they are expressed to "sufficient levels" (Gammie et al., 1999). However, we can imagine that the two forms of Kar4 could have different affinities for Ste12, which would explain the requirement for a certain level of expression. This could be tested by tagging Kar4 in N-ter, either the short or the long form, and perform co-immunoprecipitation with Ste12.

The model that we describe here may be reflecting what happens for p*AGA1* and p*FIG1* regulation. However, we described in the previous Chapter 3 other promoters induced at various time following pheromone stimulation. We also mapped all consensus and non-consensus PREs on these promoters, in addition to the nucleosome positions obtained from genome-wide mapping experiments (Figure 4.20)(Brogaard et al., 2012). We can see that all early-induced or intermediate promoters possess at least 2 consensuses PRE available, in combination or not with PRE^{nc} (Figure 4.33A). One notable exception is p*PRM1*, for which all consensus PREs are predicted to be hidden. However, we can argue that the histone positions that we extract from a previous study might be slightly different from what is happening in our strain (Brogaard et al., 2012). Indeed, the global histone mapping was performed in a strain of the BY4741 background, derived from the S288C strain, whereas we use the W303 background. These two backgrounds are slightly different, with ~15% of differences in their genome, spanning to 799 protein-coding genes (Ralser et al., 2012). Thus, the histone positioning is also likely to differ between the two strains. In fact, we found during MNase experiments a difference in histone landscape in the *AGA1* promoting region (Supplementary Figure 5).

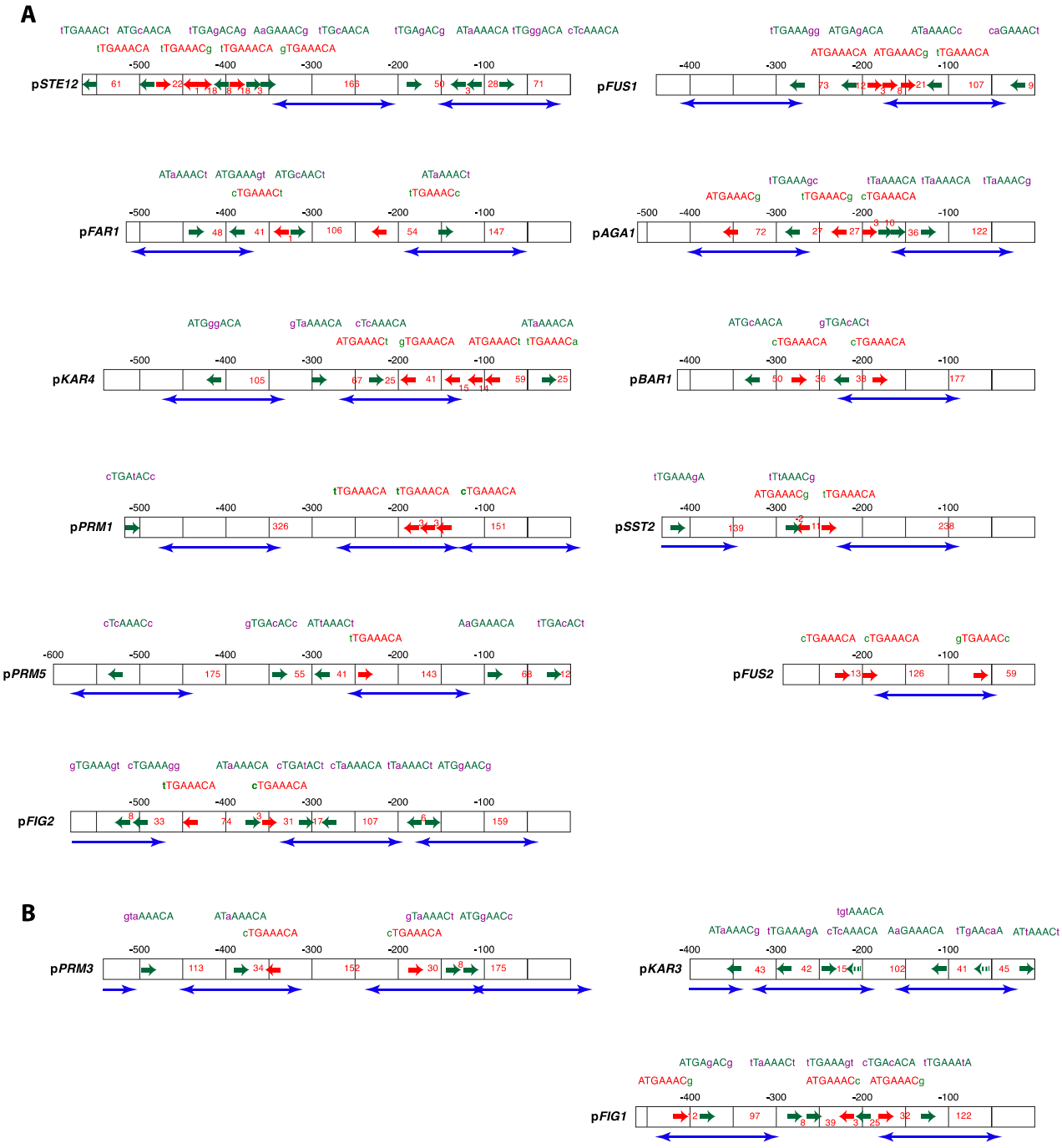


Figure 4.33 Maps of the promoters analyzed in Chapter 3

All Ste12 binding sites were mapped on the 14 promoters of the study, which are defined as nucleotides from the stop codon of the upstream ORF to the ATG of the gene of interest. Only consensus PREs, meaning nTGAAACn, are represented here by red arrows indicating their positions and orientations, with their corresponding sequences above (green lower cases are mutations from the full consensus ATGAAACA). Non-consensus PREs as studied by Su et al. are represented by green arrows and green sequences. The numbers between the PREs represent the distance in nucleotides separating the last nucleotide of the upstream PRE from the first nucleotide of the downstream PRE. The last number is the relative distance between the last PRE and the start codon. Blue arrows represent nucleosome positioning as extracted from Brogaard *et al.* (Brogaard et al., 2012).

A. Maps of the early and intermediate promoters of the study.

B. Maps of the late induced promoters of the study.

The three late-induced promoters have at most one consensus PRE available in basal condition (Figure 4.33 B). Hence, our model could be a general rule for pheromone-induced transcription. However, it would be relevant to enlarge this study to a bigger number of mating-induced genes, in order to validate our model. Indeed, more than 200 genes are induced upon pheromone stimulation, and roughly the same number is repressed (Roberts et al., 2000). The priority would be to increase our pool of late-induced genes. A good target to start with is the *CIK1* gene, which was reported to be Kar4-dependent (Kurihara et al., 1996).

We could also use the unstable version of the dPSTR to quantify gene repression, and assess the required mechanisms for this (Aymoz et al., 2016). Indeed, roughly 200 genes have been found to be downregulated in response to pheromone stimulation (Roberts et al., 2000). They are encoding functions that are not needed if the cell decides to mate, like DNA replication, budding or mitosis. This gene cluster requires the presence of Far1 to be repressed (Roberts et al., 2000). It could be interesting to assess how this inhibition is regulated, and how fast. Indeed, one can imagine that this process may occur immediately, or rather late, when the decision to undergo mating has been made, and the cell cannot go back to vegetative growth.

Interestingly, the unstable dPSTR revealed a previously non-described feature in the induction of mating-regulated genes. We found that the two early promoters p*AGA1* and p*KAR4* were not induced at the same speed throughout the pheromone response (Figure 4.34). Indeed, it seems that the early genes are induced at a high rate during the first 15min following pheromone stimulation, and at a slower rate after that. The plateau from 15 to 30 minutes illustrates an equilibrium between the rate of production of the dPSTR relocating module and its degradation. Following this plateau, the production rate becomes higher than the degradation rate, and the dPSTR can accumulate again in the nucleus, at a time which coincides with induction of late promoters. This very interesting behavior requires further investigations.

We could also try to design variants of other mating-induced genes. Indeed, our model only describes the discrepancy between early and late-induced genes, but we noticed that some genes were induced at intermediate kinetics. Dissection of such promoters (p*FIG2*, p*FUS2*, p*PRM1*) could allow an understanding of their regulation. We could imagine that they have different requirement for Ste12 or Kar4, or may need a different transcription factor to play a role in their regulation.

To conclude, we managed to come up with a model of the regulation of pheromone-induced expression of p*AGA1* and p*FIG1*, using both genetic and promoter mutation approaches. However, we need to confirm this model and verify whether it applies to other promoters as well.

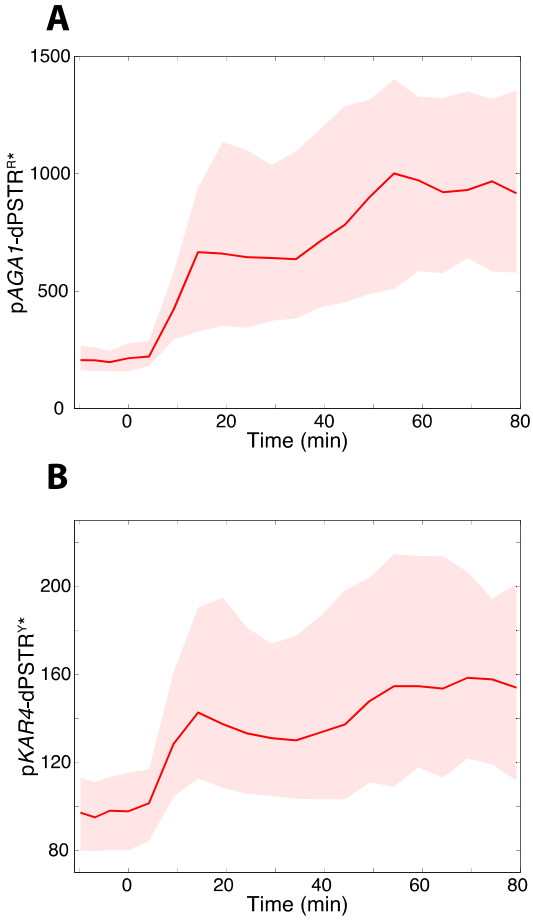


Figure 4.34 Measurements with dPSTR*

Nuclear enrichment of the unstable p*AGA1*-dPSTR*^{R} (A) and of the unstable p*KAR4*-dPSTR*^{Y} (B) in the same strain, upon stimulation with 1µM pheromone at time 0. Medians are plotted and shaded areas represent the 25th-75th percentile.

Methods

Strains and plasmids

Yeast strains and plasmids are respectively listed in Supplementary Table 1 and Annex 2.

The dPSTR plasmids were transformed in a yeast strain from a W303 MAT**a** background, bearing a Hta2-CFP marker ySP580).

Each dPSTR is fully carried by a single integration vector (pSIV (Wosika et al., 2016)) and integrated in the genome. The red (and yellow) variants of the dPSTR (dPSTR^{R} and dPSTR^{Y}, respectively) are integrated in the *URA3* (resp. *LEU2*) locus and based on interaction of the SynZips SZ1-SZ2 (resp. SZ3-SZ4) (Thompson et al., 2012), and the mCherry (resp. MCitrine) fluorescent variant (Aymoz et al., 2016). The relevant promoters of interest (-1000 to -1) were amplified and cloned upstream of the inducible stable part of the dPSTR, in pSP360 for the dPSTR^{R}, and pSP363 for dPSTR^{Y}, and checked by sequencing. The inducible part was then further cloned in the pSIV vector containing the FP part of the dPSTR (pDA157 for the dPSTR^{R}, and pDA223 for dPSTR^{Y}).

For the promoter variants, synthetic versions of p*FIG1* and p*AGA1* were designed, containing unique restriction sites (ApaI and ClaI) surrounding the region containing the PREs. This allowed to obtain dSPTR plasmids of mutants of each promoter in only one cloning (Figure 4.21). Modified fragments of p*FIG1* and p*AGA1* with sequential mutations of PREs into NdeI (CATATG) or SnaBI (TACGTA) restriction sites were designed and synthetized by IDT (gBlocks), and cloned into pDA283 or pDA282 using ApaI-ClaI. All constructs were verified by digestion and sequencing. All strains were sequenced for final confirmation. We verified that the presence of the cloning sites was not altering in any way the induction of the two promoters (data not shown). The variants were transformed in a strain carrying either p*FIG1*-dPSTR^{Y} or p*AGA1*-dPSTR^{Y} (resp. ySP644 and ySP641).

For each transformation, 8 clones were screened based on their fluorescence intensities, and four clones with similar fluorescence levels were further analyzed by a time-lapse experiment upon stimulation with 1µM of α-factor, to discard clones that would display an aberrant relocation behavior.

Sample preparation

The cells were grown overnight in selective synthetic medium to saturation (YNB:CYN3801, CSM:DCS0031, ForMedium). They were diluted to an OD_{600} of 0.05 in the morning and grown for 4h before the start of the experiment. All the time-lapse experiments were performed in well slides, for which selected wells of 96-well plates (MGB096-1-2LG, Matrical Bioscience) were coated with filtered solution of Concanavalin A in H_2O (0.5 mg.mL-1, C2010-250MG, Sigma-Aldrich) for 30min, rinsed with H_2O and dried for at least ten hours. Before the experiments, the cells were diluted to an OD_{600} of 0.04, briefly sonicated, and 200µL of cell suspension were added to a well. Imaging was started 30 min later to let the cells settle to the bottom to the well. To stimulate the cells, 100µL of a 3µM solution of synthetic exogenous α –factor was added in the cell to reach a final 1µM concentration of pheromone.

Microscopy

Images were acquired on a fully automated inverted epi-fluorescence microscope (Ti-Eclipse, Nikon) controlled by micro-manager (Edelstein et al., 2010) and placed in an incubation chamber set at 30°C, with a 40X oil objective and appropriate excitation and emission filters. The excitation is provided by a solid-state light source (SpectraX, Lumencor). The images were recorded with an sCMOS camera (Flash4.0, Hamamatsu). A motorized XY-

stage allowed recording multiple fields of view at every time point, typically 5 positions per well, 8 wells per experiment. CFP (50 ms), RFP (300 ms) and YFP (300 ms) and 2 bright-field (10 ms) images were recorded at time intervals of 2 min before induction and 5 min after.

Data analysis

Time-lapse movies were analyzed with the YeastQuant platform (Pelet et al., 2012). Briefly, the nuclei of the cells were segmented by thresholding of the CFP images. The contour of the cell around each nucleus was detected using two bright-field images. The cytoplasm object was obtained by removing the nucleus object expanded by two pixels from the cell object. Dedicated scripts in Matlab (The Mathworks) were written to further analyze the data. Only cells tracked from the beginning to the end of the movie were taken into consideration. In addition, a quality control was applied on each trace and only cells with low variability in nuclear and cell area, nuclear CFP fluorescence, and a ratio of RFP to YFP fluorescence lower than a certain threshold, were kept for further analysis. The curves displayed in the figures are from one representative experiment out of at least three true biological replicates.

For each cell, the difference between its average intensity in the nucleus and the cytoplasm was calculated at each time point to plot the nuclear enrichment of $dPSTR^R$ and $dPSTR^Y$.

For further analysis, all retained cell traces were smoothed by a moving average of three points. The basal level is calculated as the mean of the three time-points preceding the stimulation. The corrected nuclear enrichment of the dPSTR was calculated by subtracting the basal level to the smooth trace. The expression output represents the maximal corrected nuclear enrichment of the dPSTR. The population average expression output was calculates on the mean trace of all cells. A threshold to qualify cells as expressing was defined as 20% of the population average expression output. For all expressing cells, dPSTRs traces were normalized between 0 and 1, and the response time was identified as the first time point, after stimulation, to exceed 0.2. For plots of population average correlation, and instant correlations, all cell traces were normalized by the mean trace of all cells.

The intrinsic and extrinsic noise were calculated according to the formula from Elowitz *et al.(Elowitz et al., 2002)*:

$$\eta^2_{int} = \frac{\langle (r_i - y_i)^2 \rangle}{2\langle r_i \rangle \langle y_i \rangle} \quad \eta^2_{ext} = \frac{\langle r_i y_i \rangle - \langle r_i \rangle \langle y_i \rangle}{\langle r_i \rangle \langle y_i \rangle} \quad \eta^2_{tot} = \frac{\langle r_i^2 + y_i^2 \rangle - 2\langle r_i \rangle \langle y_i \rangle}{2\langle r_i \rangle \langle y_i \rangle}$$

since: $\eta^2_{tot} = \eta^2_{int} + \eta^2_{ext}$ we plot the fraction of the intrinsic noise as: $\frac{\eta^2_{int}}{\eta^2_{tot}}$

r_i and y_i are the normalized nuclear accumulation from the i^{th} cell in the red and yellow channels, respectively. The normalization factors were obtained from the highest and lowest average intensity from the entire data set for one replicate.

ChIP assays

Yeast cultures were grown to early log phase (OD_{600} 0.4–0.6), then samples (50 ml) were subjected to 1μM α-factor for the indicated times. For crosslinking, yeast cells were treated with 1% formaldehyde for 20 min at room temperature. Glycine was added to a final concentration of 330 mM for 15 min. Cells were collected, washed four times with cold TBS (20 mM Tris-HCl, pH 7.5, 150 mM NaCl), and kept at −20 °C for further processing. Cell pellets were resuspended in 0.3 ml cold lysis buffer (50 mM HEPES-KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, 0.1% sodium deoxycholate, 1% Triton-X 100, 1mM PMSF, 2mM benzamidine, 2 μg/ml leupeptin, 2 μg/ml pepstatin, 2 μg/ml aprotinin). An equal volume of glass beads was added, and cells were disrupted by vortexing (with Vortex Genie) for 13 min on ice. Glass beads were discarded and the crosslinked chromatin was sonicated with water

bath sonicator (Bioruptor) to yield an average DNA fragment size of 350 bp (range, 100–850 bp). Finally, the sample was clarified by centrifugation at 16,000g for 5 min at 4 °C. Supernatants were incubated with 50 µl anti-HA 12CA5 or anti-Myc 9E10 monoclonal antibodies precoupled to pan mouse IgG DynabeadsTM (Invitrogen, 11042). After 120 min at 4 °C on a rotator, beads were washed twice for 4 min in 1 ml lysis buffer, twice in 1 ml lysis buffer with 500 mM NaCl, twice in 1 ml washing buffer (10 mM Tris-HCl pH 8.0, 0.25 M LiCl, 1 mM EDTA, 0.5% N-P40, 0.5% sodium deoxycholate) and once in 1 ml TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA). Immunoprecipitated material was eluted twice from the beads by heating for 10 min at 65 °C in 50 µl elution buffer (25 mM Tris-HCl pH 7.5, 10 mM EDTA, 0.5% SDS). To reverse crosslinking, samples were adjusted to 0.3 ml with elution buffer and incubated overnight at 65 °C. Proteins were digested by adding 0.5mg/ml Proteinase K (Novagen, 71049) for 1.5h at 37°C. DNA was extracted with phenol-chloroform-isoamyl alcohol (25:24:1) and chloroform. It was finally precipitated with 48% (v/v) of isopropanol and 90 mM NaC for 2 h at −20 °C in the presence of 20 µg glycogen, and resuspended in 30 µl of TE buffer. Quantitative PCR analysis of *AGA1* and *FIG1* promoter sequences used the following primers with locations indicated by the distance from the respective ATG initiation codon: *AGA1* promoter (-310/-207); *FIG1* promoter (-400/-197); and TEL (telomeric region on the right arm of chromosome VI). Experiments were done on three independent chromatin preparations and quantitative PCR analysis was done in real time using an Applied Biosystems 7700 sequence detector. Immunoprecipitation efficiency was calculated in triplicate by normalizing the amount of PCR product in the immunoprecipitated sample by that in TEL sequence control. The binding data are presented as fold induction with respect to the non-treated condition.

In vivo coprecipitation assay

Ste12-Myc and/or Kar4-HA-tagged cells in mid-log phase (50 ml) were treated with 1µM α-factor for 30 min or left untreated and then collected by brief centrifugation at 4°C. Pellets were harvested with glass beads in the FastPrep-24 (Qbiogene, 60 s at speed 5) in lysis buffer A (50mM Tris–HCl pH 7.5, 150mM NaCl, 15mM EDTA, 15mM EGTA, 2mM DTT, 0.1% Triton X-100, 1mM PMSF, 1mM benzamidine, 2 µg/ml leupeptin, 2 µg/ml pepstatin, 25mM β-glycerophosphate, 1mM sodium pyrophosphate, 10mM sodium fluoride, 100 µM sodium orthovanadate), and lysates were clarified by centrifugation and quantified by the Bradford assay (Bio-Rad Laboratories). 1.5mg of cleared supernatant was subjected to immunoprecipitation with rabbit polyclonal HA tag antibody (Abcam, ab9110) overnight at 4°C. Immunocomplexes were recovered with Dynabeads^{TM} protein A (Invitrogen, 10002D) and washed with lysis buffer. Finally, they were resolved by SDS–PAGE and blotted with mouse monoclonal anti-HA 12CA5 or anti-Myc 9E10 antibodies. As a control, 50 µg of whole-cell extract were also blotted to check the expression levels of the tagged proteins (total).

Northern blot analysis

Yeast strains were grown to mid-log phase in rich medium and then treated with 1µM α-factor for the length of time indicated. Total RNA and expression of specific genes were probed using radiolabelled PCR fragments containing a fragment of *AGA1* ORF (+145/+936 bp), *FIG1* ORF (+106/+948 bp) and *ENO1* ORF (+1/+1310 bp). Signals were acquired with a Fujifilm BAS-5000 phosphorimager and ImageQuantTL software.

MNase nucleosome mapping

Yeast spheroplast preparation and micrococcal nuclease digestions were performed as described previously with modifications (Nadal-Ribelles et al., 2015; Nadal-Ribelles et al., 2014). Ste12-Myc and Kar4-HA double-tagged strain was grown to early log phase (A660 0.4–0.6) and samples of 500 ml of culture were exposed to 1µM α-factor for the indicated length of

time. The cells were cross-linked with 1% formaldehyde for 15 min at 30°C and the reaction was stopped with 125 mM glycine for 5 min. Cells were washed and resupended in 1M sorbitol TE buffer before cell wall digestion with 100 T zymoliase (USB). Cells were then lysed and immediately digested with 60–240 mU/µl of micrococcal nuclease (Worthington Biochemical Corporation, Lakewood; NJ., USA). DNA was subjected to electrophoresis in a 1.5% (w/v) agarose gel and the band corresponding to the mononucleosome was cut and purified using a QIAquick gel extraction kit (Qiagen). DNA was used in a real-time PCR with specific tiled oligonucleotides covering *AGA1* promoter and partial coding sequence (-928/+470) or *FIG1* promoter and a partial coding sequence (-933/+463) included in the Supplementary Table 2. PCR quantification was referred to an internal loading control (telomeric region in chromosome 6) and nucleosome occupancy was normalized to 1 at the (-1) nucleosome region of the untreated condition.

Supplementary Tables

Supplementary Table 1: List of yeast strains used in Chapter 4

Strain	Background	Genotype	Plasmid
ySP2	W303	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*	
ySP580	W303	*HTA2-CFP*	
ySP644	ySP580	*HTA2-CFP*	
		LEU2: pFIG1-dPSTR^{Y4-3}	pSP367
ySP641	ySP580	*HTA2-CFP*	
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
ySP643	ySP641	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA160	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		KAR4::NAT	
yDA163	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		FAR1::NAT	
yDA218	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		GCN5::NAT	
yDA141	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		KSS1::NAT	
yDA161	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		STE11::NAT	
yDA162	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		STE12::NAT	
yDA164	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		STE2::NAT	
yDA142	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		FUS3::NAT	
yDA208	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		MOT3::NAT	
yDA151	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		DIG1::NAT	
yDA159	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		TEC1::NAT	

yDA219	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		ARP8::NAT	
ySP692	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		BAR1::NAT	
yDA305	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		DIG2::NAT	
		DIG1::KAN	
yDA317	ySP643	*HTA2-CFP*	
		URA3: pFIG1-dPSTR^{R2-1}	pSP366
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		DIG2::NAT	
yDA316	yDA249	*HTA2-CFP*	
		URA3: pPRM3-dPSTR^{R2-1}	pDA305
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		KAR4::NAT	
yDA348	ySP642	*HTA2-CFP*	
		URA3: pAGA1-dPSTR^{R2-1}	pSP365
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		KAR4::NAT	
yDA319	yDA296	*HTA2-CFP*	
		URA3: pFUS2-dPSTR^{R2-1}	pDA334
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		KAR4::NAT	
yDA320	yDA302	*HTA2-CFP*	
		URA3: pFIG2-dPSTR^{R2-1}	pDA340
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		KAR4::NAT	
yDA322	yDA306	*HTA2-CFP*	
		URA3: pPRM1-dPSTR^{R2-1}	pDA338
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		KAR4::NAT	
yDA258	ySP644	*HTA2-CFP*	
		URA3: pFIG1syn_ΔPREIII-dPSTR^{R2-1}	pDA288
		LEU2: pFIG1-dPSTR^{Y4-3}	pSP367
yDA251	ySP644	*HTA2-CFP*	
		URA3: pFIG1syn_4thPRE-dPSTR^{R2-1}	pDA307
		LEU2: pFIG1-dPSTR^{Y4-3}	pSP367
yDA264	ySP644	*HTA2-CFP*	
		URA3: pFIG1syn_PREII_10nt_PREIII-dPSTR^{R2-1}	pDA319
		LEU2: pFIG1-dPSTR^{Y4-3}	
			pSP367
yDA265	ySP644	*HTA2-CFP*	
		URA3: pFIG1syn_PREII_27nt_PREIII-dPSTR^{R2-1}	pDA320
		LEU2: pFIG1-dPSTR^{Y4-3}	
			pSP367
yDA266	ySP644	*HTA2-CFP*	
		URA3: pFIG1syn_PREII_27nt_PREIII_OoN-dPSTR^{R2-1}	pDA321
		LEU2: pFIG1-dPSTR^{Y4-3}	
			pSP367
yDA270	ySP644	*HTA2-CFP*	
		URA3: pFIG1syn_ΔPREIΔPREIII-dPSTR^{R2-1}	pDA324
		LEU2: pFIG1-dPSTR^{Y4-3}	
			pSP367
yDA294	ySP644	*HTA2-CFP*	

		URA3: pFIG1syn_163nt_PREII_ΔPREIII-dPSTR^{R2-1}	pDA329
		LEU2: pFIG1-dPSTR^{Y4-3}	
			pSP367
yDA295	ySP644	*HTA2-CFP*	
		URA3: pFIG1syn_PREII_36nt_PREIII_OoN-dPSTR^{R2-1}	pDA330
		LEU2: pFIG1-dPSTR^{Y4-3}	
			pSP367
yDA313	ySP644	*HTA2-CFP*	
		URA3: pFIG1syn_PREII^{nc}_to_consensus-dPSTR^{R2-1}	pDA344
		LEU2: pFIG1-dPSTR^{Y4-3}	
			pSP367
yDA314	ySP644	*HTA2-CFP*	
		URA3: pFIG1syn_ΔPREII^{nc} -dPSTR^{R2-1}	pDA345
		LEU2: pFIG1-dPSTR^{Y4-3}	pSP367
yDA315	ySP644	*HTA2-CFP*	
		URA3: pFIG1syn_ΔPREIΔPREII^{nc}ΔPREIII -dPSTR^{R2-1}	pDA346
		LEU2: pFIG1-dPSTR^{Y4-3}	pSP367
yDA289	yDA266	*HTA2-CFP*	
		URA3: pFIG1syn_PREII_27nt_PREIII_OoN-dPSTR^{R2-1}	pDA321
		LEU2: pFIG1-dPSTR^{Y4-3}	
		KAR4::NAT	pSP367
yDA298	yDA270	*HTA2-CFP*	
		URA3: pFIG1syn_ΔPREIΔPREIII-dPSTR^{R2-1}	pDA324
		LEU2: pFIG1-dPSTR^{Y4-3}	
		KAR4::NAT	pSP367
yDA326	yDA313	*HTA2-CFP*	
		URA3: pFIG1syn_PREII^{nc}_to_consensus-dPSTR^{R2-1}	pDA344
		LEU2: pFIG1-dPSTR^{Y4-3}	
		KAR4::NAT	pSP367
yDA230	ySP641	*HTA2-CFP*	
		*URA3: pAGA1*FIG1-dPSTR^{R2-1}*	pDA285
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA228	ySP641	*HTA2-CFP*	
		*URA3: pFIG1*AGA1 -dPSTR^{R2-1}*	pDA284
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA303	ySP641	*HTA2-CFP*	
		*URA3: pAGA1*FIG1ΔPREII*AGA1 -dPSTR^{R2-1}*	pDA343
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA330	ySP641	*HTA2-CFP*	
		*URA3: pAGA1*FIG1*AGA1 -dPSTR^{R2-1}*	pDA365
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA352	ySP644	*HTA2-CFP*	
		*URA3: pAGA1*FIG1*AGA1 -dPSTR^{R2-1}*	pDA387
		LEU2: pFIG1-dPSTR^{Y4-3}	pSP367
yDA354	ySP644	*HTA2-CFP*	
		*URA3: pAGA1ΔPREII*FIG1 -dPSTR^{R2-1}*	pDA388
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA355	ySP644	*HTA2-CFP*	
		*URA3: pAGA1ΔPREIΔPREII*FIG1 -dPSTR^{R2-1}*	pDA389
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
yDA235	yDA230	*HTA2-CFP*	
		*URA3: pAGA1*FIG1-dPSTR^{R2-1}*	pDA285
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		KAR4::NAT	
yDA234	yDA228	*HTA2-CFP*	
		*URA3: pFIG1*AGA1 -dPSTR^{R2-1}*	pDA284
		LEU2: pAGA1-dPSTR^{Y4-3}	pSP368
		KAR4::NAT	
yDA329	yDA303	*HTA2-CFP*	

		*URA3: pAGA1*FIG1ΔPREII*AGA1 -dPSTR^{R2-1}* *LEU2: pAGA1-dPSTR^{Y4-3}* *KAR4::NAT*	pDA343 pSP368
yDA337	yDA330	*HTA2-CFP* *URA3: pAGA1*FIG1*AGA1 -dPSTR^{R2-1}* *LEU2: pAGA1-dPSTR^{Y4-3}* *KAR4::NAT*	 pDA365 pSP368
yDA358	yDA352	*HTA2-CFP* *URA3: pAGA1*FIG1*AGA1 -dPSTR^{R2-1}* *LEU2: pFIG1-dPSTR^{Y4-3}* *KAR4::NAT*	 pDA387 pSP367
yDA360	yDA354	*HTA2-CFP* *URA3: pAGA1ΔPREII*FIG1 -dPSTR^{R2-1}* *LEU2: pAGA1-dPSTR^{Y4-3}* *KAR4::NAT*	 pDA388 pSP368
yDA361	yDA355	*HTA2-CFP* *URA3: pAGA1ΔPREIΔPREII*FIG1 -dPSTR^{R2-1}* *LEU2: pAGA1-dPSTR^{Y4-3}* *KAR4::NAT*	 pDA389 pSP368
yCS418	ySP580	*HTA2-CFP* *Ste12-9myc::NAT* *Kar4-6HA::HPH*	
yCS409	ySP580	*HTA2-CFP* *Ste12-9myc::NAT*	
yCS412	ySP580	*HTA2-CFP* *Kar4-6HA::HPH*	
yDA222	ySP580	*HTA2-CFP* *URA3: pAGA1 -dPSTR*$^{R2-1}$* *LEU2: pKAR4-dPSTR*$^{Y4-3}$*	 pSP390 pDA274

* The numbers in the superscript dPSTR^{R2-1} indicate the pair of SynZip used, and the letter the color of the fluorescent protein (R for mCherry and Y for MCitrine).

Supplementary table 2: oligonucleotides for MNase nucleosome mapping.

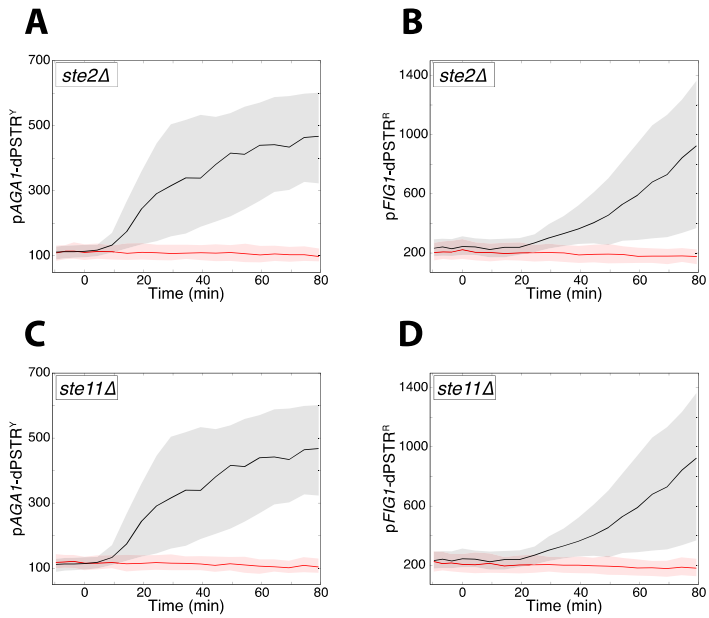
Oligos for *AGA1* gene				
Amplicon	**Forward_primer**	**Position from ATG**	**Reverse_primer**	**Position from ATG**
-928	CACTTCCAAGCGTATCATCAGTT	-977	CAACACAGCATTTGGACCTG	-878
-858	ACGTTTGATGCAGGTCCAA	-907	ATCCAGGAACAGAGCCAAACA	-808
-768	TCTGTTCCTGGATGGGACAA	-820	GCAACTCAAGATCCAATTCACG	-715
-679	TGGATCTTGAGTTGCAAAAGG	-729	GACCAGTCTTTGCGTCAATCA	-628
-570	TGGTCTACCAAAGGAATAAGATCAA	-632	TAAATTGAAGCTTGGTCGTTCC	-507
-488	CGGTATTGGTCGGAACGA	-539	AAAAGCAAAAGAGTAGGCATCAAA	-437
-405	GATGCCTACTCTTTTGCTTTTTCA	-457	GGGTGGACGTACTTGTGCTAA	-353
-342	CGTAGCTGTCTAACAGCACCACT	-395	GAAGATATGTGACAGGTACCCTAAATG	-289

-259	AGGGTACCTGTCACATATCTTCTCA	-310	ATTATGTTACAGCCGCGTTTTG	-207
-182	CAAGTCAAAACGCGGCTGT	-233	TTGCTTGCTGGGAACTGC	-131
-92	GCAGTTCCCAGCAAGCAA	-148	AATGCTCGCCGTTTTTCA	-35
10	AAAACGGCGAGCATTAACAA	-49	CAAGGCAATATTAGTTAATCCCAACA	69
102	TGGGATTAACTAATATTGCCTTGG	47	TAGCGCGGGTGAAACTGTAG	156
195	TCACCCGCGCTAGTCTCC	145	CGGCAGCAGATGAAGTGG	244
268	ACTTGGTGTCCATTGACGGTA	205	TGCCTCATGGGAGCAGAC	330
374	TGCTCCCATGAGGCATGT	316	AATGAGCTGATAGCGGTTGTG	431
470	CGCTATCAGCTCATTATCCGAAGT	417	CGAAAGTGTAGAGGTGACAGGTG	522

Oligos for *FIG1* gene

Amplicon	**Forward_primer**	**position from ATG**	**Reverse_primer**	**position from ATG**
-933	TGATCAACCAAACGCCGATA	-995	CTTGGAAAGTTGGGGCATC	-871
-833	TGCCCCAACTTTCCAAGA	-887	GGGAAGACACTGGGTCATTG	-778
-738	CAATGACCCAGTGTCTTCCCTA	-797	GAACGTTTGCGTCCGTGTC	-679
-629	ATCGACACGGACGCAAAC	-700	TCACCGGCATTCTTGGAA	-558
-517	TTCCAAGAATGCCGGTGA	-575	TCATCCCAAAGAGGAAGCAC	-458
-428	GTGCTTCCTCTTTGGGATGA	-477	TCGTCTCATCAAGTCAAAATTCG	-378
-337	GTGCTTCCTCTTTGGGATGA	-477	TAACTGACACATACATGAAACCATC	-197
-211	TGAAAGTCCTTCTCGCTTTAGG	-264	TTTCTTGGTTCGTTTCATTGC	-157
-123	AAATATGGCTAAGTAGCAATGAAACG	-192	ACCAAAGACAAGCAAGAACCTG	-54
-14	CAGGTTCTTGCTTGTCTTTGGT	-75	TCTGGGCATACGCTTGGTA	48
61	TGGTCGCAATCTCAATGATTT	2	CGGGTTGTAACAGCCGATG	120
168	GGCTGTTACAACCCGTCAAA	106	TCCTCCAAGCCCAGAGTTG	230
277	TGGGCTTGGAGGAAGTCA	218	GCATAGCAAATTGAATTGGAGAA	335
380	TTGCTATGCAAGAAAGAATTTAAGC	327	GACAACGCTTGATTGGGTTTT	432
463	AAACCCAATCAAGCGTTGTCT	413	TGTTAGGATTACAGTTGCCATCAA	513
591	TGTAACTGTCCCTAAGTTACCATTCAA	540	GTCCACATCGCACCAATACC	641
672	TGGGGTATTGGTGCGATG	619	GCTGCCTTCTTGCCCTTC	725
778	GGGCAAGAAGGCAGCAGT	711	TTTTGGGCACATGGAACATT	845
887	AATGTTCCATGTGCCCAAAA	826	TGACATTCTGAAATATCCGCTCA	948

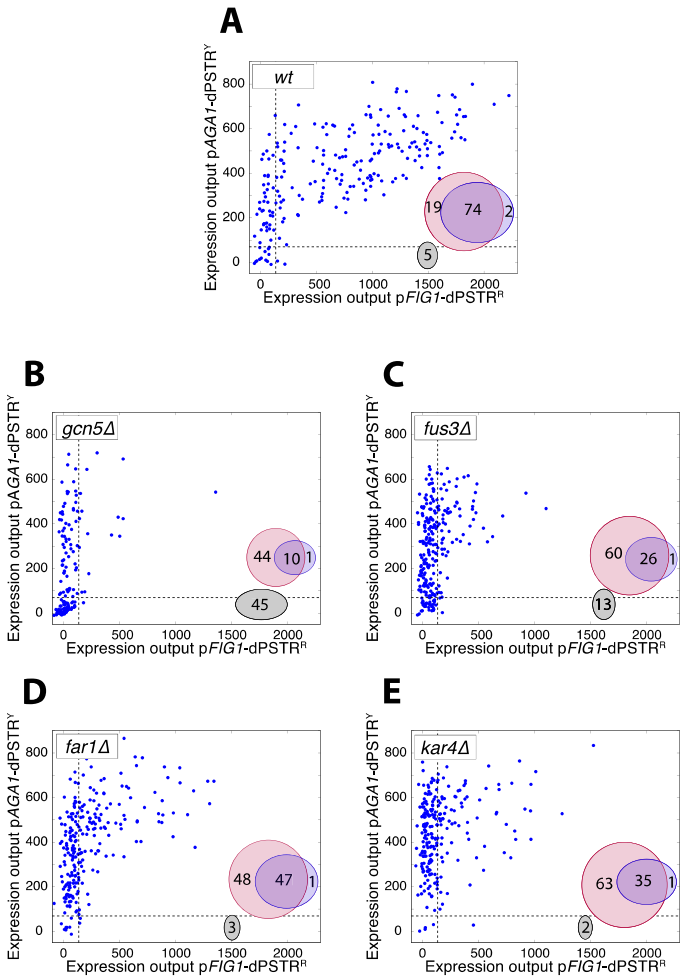
Supplementary Figures



Supplementary Figure 1

Expression of p*FIG1* and p*AGA1* in signaling dead mutants

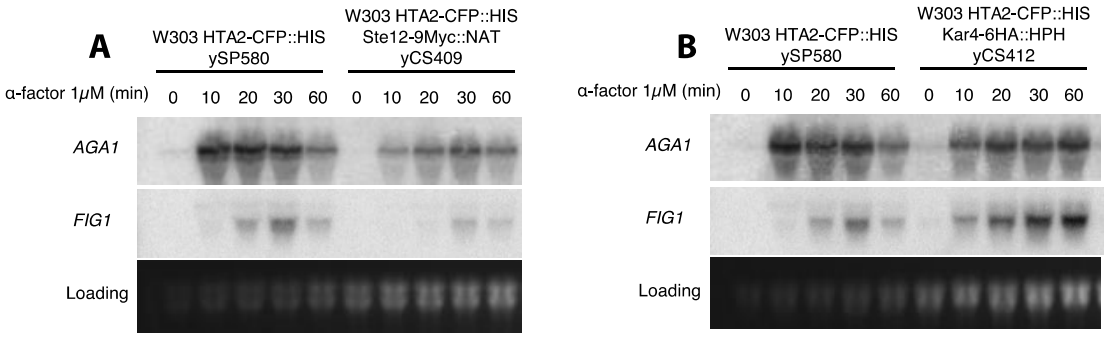
Nuclear enrichment of the p*AGA1*-dPSTR^{Y} (A-B) and p*FIG1*-dPSTR^{R} (C-D) after stimulation by 1μM of pheromone in the indicated mutant. Curves are the median of either the mutant (red) or the WT strain (black) for one representative experiment, with the shaded area representing the 25-75 percentile.



Supplementary Figure 2

Correlation of gene expression output and proportion of expressing cells

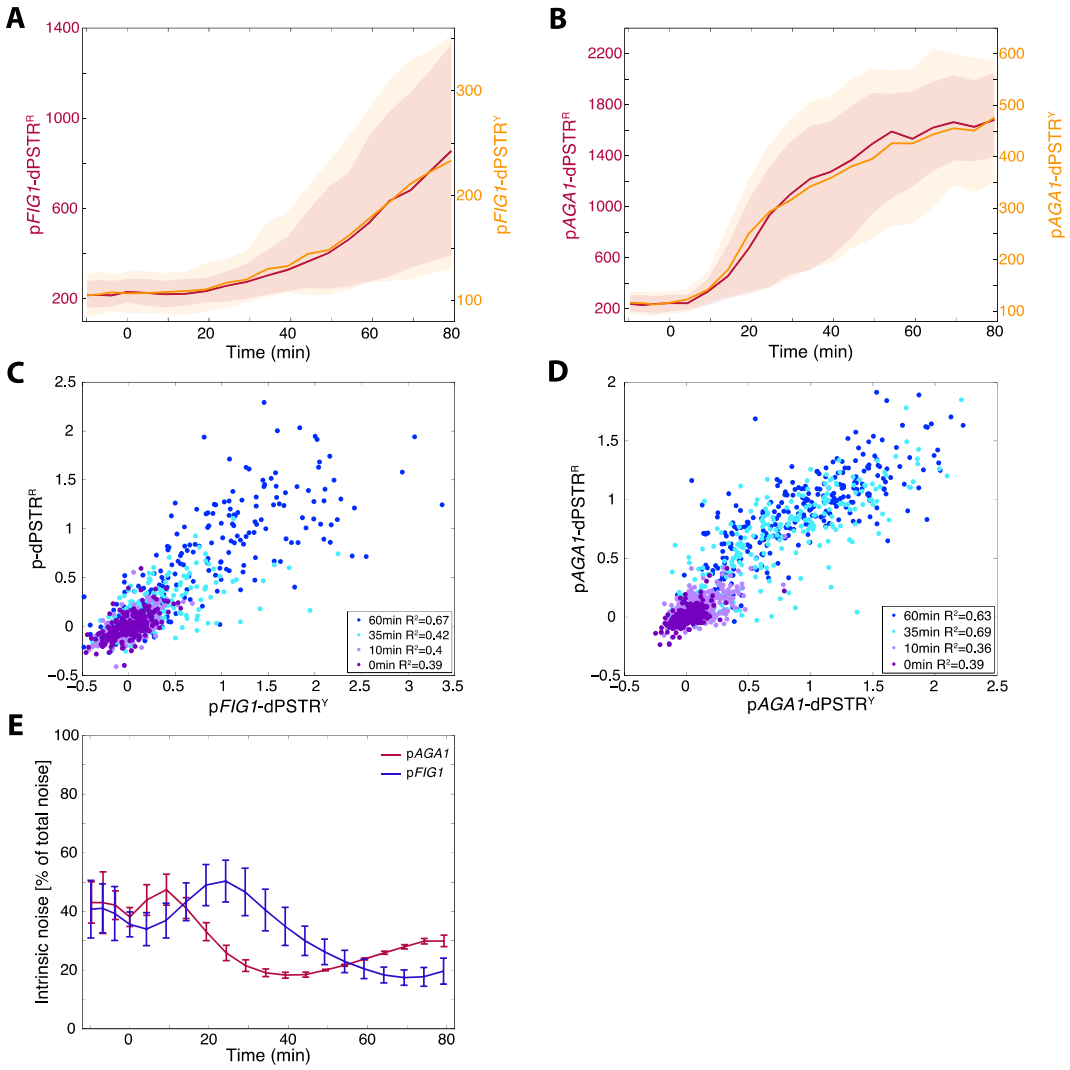
Correlation of the expression output of p*AGA1*-dPSTR^{Y} and p*FIG1*-dPSTR^{R} for all single cells of the experiment, for the indicated strain. Dotted lines represent the threshold of expression (defined as the 20% of the wt mean expression output for each dPSTR). The Venn diagram represents the proportion of cells expressing p*AGA1* (red circle) or p*FIG1* (blue circle) or none of them (black circle).



Supplementary Figure 3

Northern blot quantification of *AGA1* and *FIG1* expression in strains bearing tagged TFs

Northern blot measurements of *AGA1* and *FIG1* transcripts in the strain used for dPSTR (ySP580, left) compared to a strain bearing Ste12-myc (2, left) or Kar4-HA (B, left).
Note that the presence of tagged TF modifies the expression kinetics of both genes.

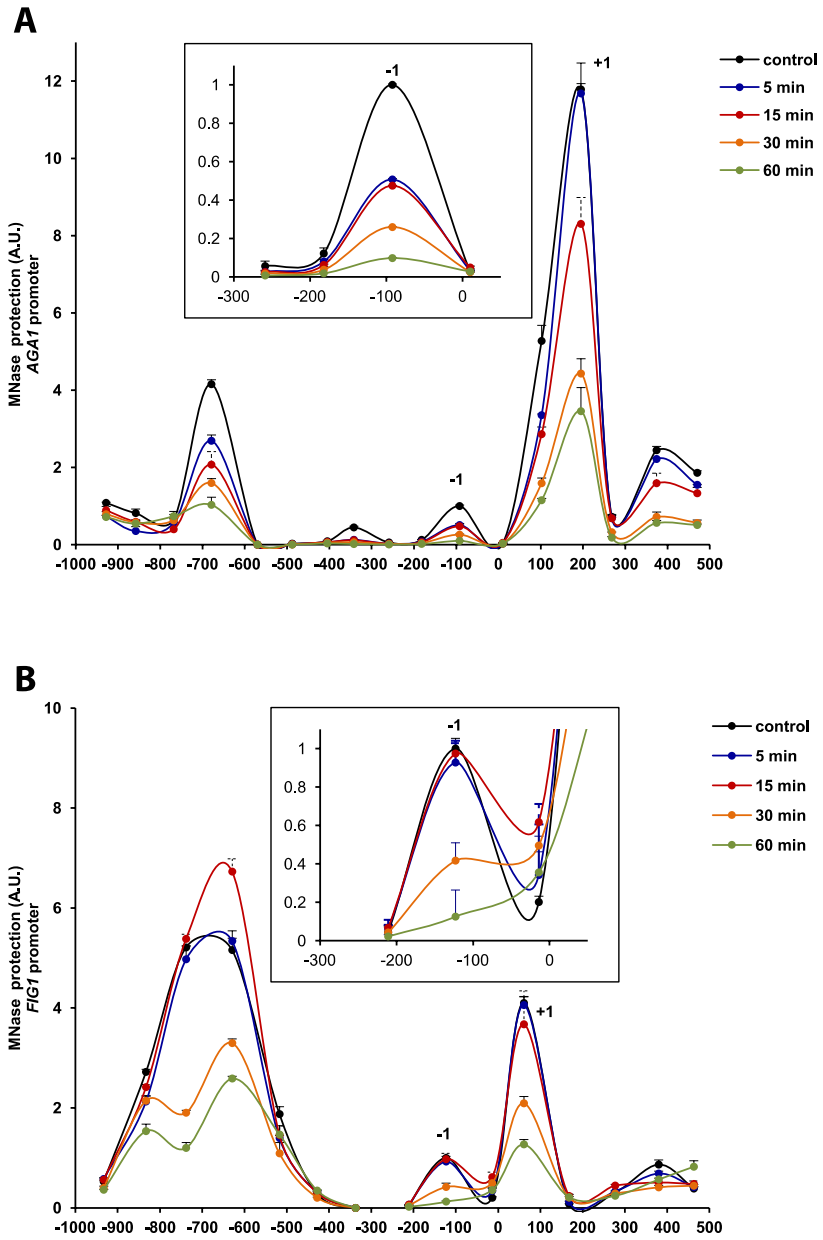


Supplementary Figure 4
Temporal evolution of noise in p*FIG1* and p*AGA1* induction

A. and B. Nuclear enrichment of p*FIG1*-dPSTR^{Y} and p*FIG1*-dPSTR^{R} (A) p*AGA1*-dPSTR^{Y} and p*AGA1*-dPSTR^{R} (B) as function of time. dPSTR^{Y} is plotted in yellow, right axis, and dPSTR^{R} in red, left axis.

C. and D. Correlation of the normalized dPSTR nuclear enrichment from all single cells at diferent time points after stimulation in strains carrying two dPSTRs measuring either p*FIG1* (C) or p*AGA1* (D).

E. Intrinsic noise as percent of total noise calculated for p*FIG1* and p*AGA1* at all time points of the time-lapse movie, using the formula from Elowitz *et al.* 2002 (see Methods). The error bar is the standard deviation calculated from 3 biological replicates.



Supplementary Figure 5

Histone positioning over p*AGA1* and p*FIG1*

A and B. MNase protection assay performed on the full promoters p*AGA1* (A) and p*FIG1* (B).

Chapter 5

Discussion around the dPSTR system

Abstract

Throughout this thesis, we used the dPSTR to describe the dynamics of gene expression in the response to osmotic stress and pheromone signaling. In this chapter, we will compare dPSTR measurements to other gene expression methods in order to validate the use of the dPSTR. Then, we will discuss some limitations ensuing from the dPSTR system regarding quantification and locus insertion. Finally, we will describe some perspectives on the dPSTR method that can be envisioned in the future.

Part I: Comparison of the dPSTR to other assays

Comparison of the dPSTR to classical gene expression reporter

Gene expression is very often quantified by means of fluorescent proteins (FP). Different quantification techniques can be used to measure the fluorescence appearance resulting from the fusion of a promoter of interest with a gene encoding a FP. First of all, fluorescent microscopy, either in time-lapse or static measurements, allows to quantify the fluorescence appearance at the single cell level (Elowitz et al., 2002; Hansen and O'Shea, 2015; Patterson et al., 2010). Time-lapse experiments allow to track the fate of a single cell in course of the experiment (Pelet et al., 2013). However, the slow maturation time of the FP induces a delay in the detection of the fluorescent signal, which hinders the quantification of fast gene expression dynamics (Olenych et al., 2007). One way to counteract this is to block protein production by addition of a drug at various time-points after stimulation. Hence, the presence of cycloheximide prevents ribosomes to initiate or pursue protein synthesis. It will allow to wait until the produced pool of FPs undergo maturation before fluorescence quantification. However, this is not compatible with time-lapse microscopy. Instead, this principle is applied to samples of cells often taken from a liquid culture stimulated at various times, which are then quantified by a flow cytometer. This method has the advantage to access to true kinetics of protein expression, although it precludes the monitoring of single-cell responses in course of time (Pelet et al., 2013; Shapiro, 2003).

Figure 5.1 shows measurements of the promoter p*STL1* expression, induced by hyper-osmotic shock, by fusion with a FP and compares the quantification by either time-lapse microscopy, flow cytometry following protein synthesis block, or by the dPSTR. We can clearly see that the dPSTR reflects the true induction kinetics measured by the flow cytometry experiment. Instead, the time-lapse measurement of fluorescence appearance by microscopy depicts a delay of 10 to 20 minutes compared to the two other methods. This shows that the dPSTR combines the advantages of the two other techniques: real-time quantification of protein expression in living cells and access to single-cells kinetics.

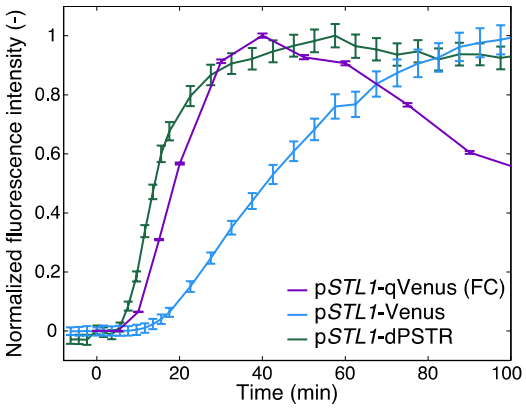


Figure 5.1 Comparison between dPSTR and protein-FP fusion

Measurements of salt-induced expression of p*STL1* (0.2M NaCl) by the dPSTR (green), or promoter-Venus fusion measured by flow cytometry (quadruple-Venus, purple) or time-lapse microscopy (simple venus, cyan).

Comparison of dPSTR measurements to mRNA measurements

We performed Northern Blot quantifications of the mRNA levels for both *AGA1* and *FIG1* in response to pheromone stimulation (Figure 5.2). We can see that the mRNA of *AGA1* appears within the first 10 minutes following stimulation, whereas *FIG1* mRNA appears between 15 and 30 minutes after stimulus (Figure 5.2A and B). An interesting feature that we can observe is the decay in expression of the two genes after 20 minutes for *AGA1* and 30 minutes for *FIG1*. With the dPSTR, we also quantify a delay in the induction of p*FIG1* compared to p*AGA1* (Figure 5.2C). We start to detect a nuclear enrichment of the dPSTR from 5-10 minutes for p*AGA1* and 20 minutes for p*FIG1*, which is comparable to the Northern blot measurements. However, we do not observe the decay visible by mRNA quantifications after 30 minutes of stimulation, probably because we use the stable version of the dPSTR here.

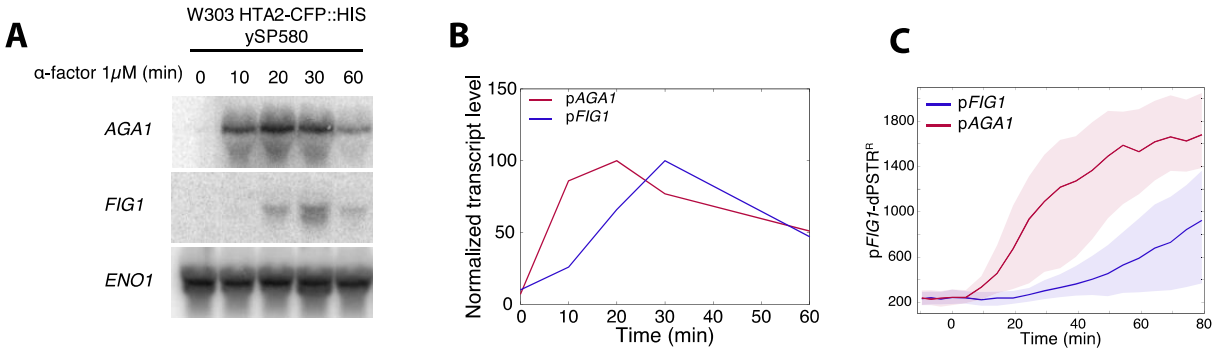


Figure 5.2 mRNA levels of p*AGA1* and p*FIG1*

A. mRNA blot of *aga1* or *fig1* various times before and after pheromone stimulation, in our reference strain ySP580.

B. Corresponding quantifications normalized to the peak of induction.

C. dPSTR measurement of p*AGA1* (red) and p*FIG1* (blue)

Nevertheless, we should note that the Northern blot experiment is performed in slightly different conditions than the dPSTR experiments (see Methods from Chapter 1 to 3). Cells were incubated in rich liquid medium, exposed to pheromone in liquid medium, and samples were taken at various time points. For the dPSTR experiments, we use 96-well plates, and cells are in direct contact with a glass surface. One difference between these two conditions is the pheromone concentration surrounding the cells. In the liquid medium, the cell concentration is much higher than in wells (OD_{600} 0.5 versus 0.04 in wells). Hence, in flasks, the amount of Bar1 protease in the medium is higher, which degrade more pheromone. However, the shaking of the flask may homogenize the concentration of pheromone surrounding cells. In wells, Bar1 is also degrading the pheromone, but diffusion occurs quickly to replenish pheromone levels.

In Chapter 4, we saw that the tagging of Ste12 and Kar4 was changing the mRNA production of *AGA1* and *FIG1* (Figure 5.3A). The induction of the two promoters occurs with similar kinetics, *AGA1* is induced at lower levels, and *FIG1* is induced at higher levels than in a WT strain. We verified that we could also observe this effect with the dPSTRs (Figure 5.3). The p*FIG1*-dPSTR^R does not show drastic difference in both backgrounds (Figure 5.3B). The kinetics are similar and the level of expression as well. *FIG1* mRNA levels decrease 60 minutes after stimulation in a WT strain, which we don't observe with the dPSTR because we use the stable form in this strain. Moreover, the induction of p*AGA1* is strongly delayed in presence of the tagged proteins (Figure 5.3C). There is a delay of 10-30 minutes when

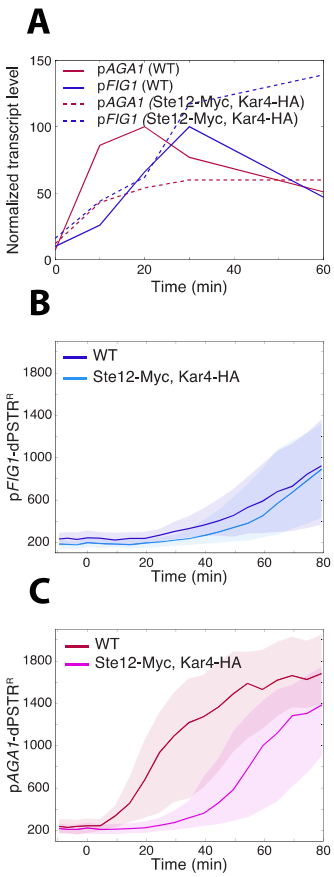


Figure 5.3 Effect of TF tagging on gene expression

A. Normalized transcript levels of *FIG1* (blue) and *AGA1* (red) in either a WT (filled lines) or a strain with Ste12 and Kar4 tagged (dotted lines), quantified by Northern blot experiment.

B-C. Nuclear enrichment of the dPSTR^R measuring p*FIG1* (B) or p*AGA1* (C) in a WT strain or in the tagged strain.

quantified with the dPSTR^R. In addition to this, the expression level measured with the dPSTR^R is lower in the tagged strain (Figure 5.3C). Hence, with the dPSTR, we are able to detect the effect of the presence of tagged TFs on the expression of p*AGA1*.

As an additional control, we also performed mRNA quantifications by using the PP7 assay (Figure 5.4). This method allows direct visualization of an actively transcribed locus in real-time (Larson et al., 2011). Briefly, a phage protein, PP7, can bind to specific mRNA stem loops that are repeatedly inserted downstream of the transcription start site of the gene of interest. For visualization, the PP7 protein is fused to a double GFP. Figure 5.3A shows two pairs of cells in which either *AGA1* or *FIG1* mRNA production is monitored. We can see the dot representing the transcription site of *AGA1* appearing 2 minutes after pheromone addition, whereas *FIG1*'s appears 19 min after pheromone stimulation. Quantifications show that the *AGA1* promoter starts to be transcribed from 2 to 50 minutes after pheromone addition, whereas for *FIG1* promoter, the first dots appear 10-15 minutes following stimulation, although many cell require much longer time to induce transcription (Figure 5.4B).

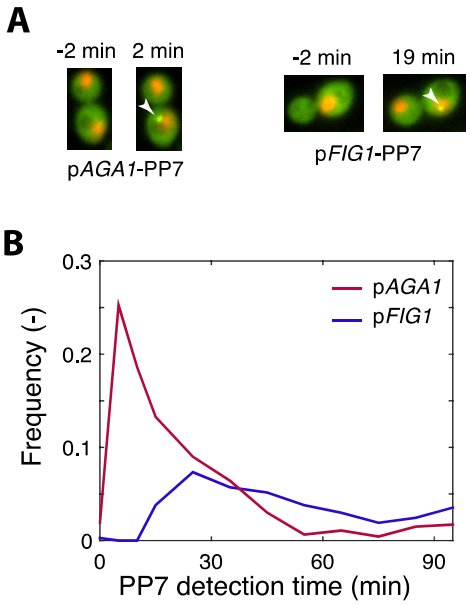


Figure 5.4 Transcription monitoring by live single-cell assay

A. Images of cells expressing a PP7-2xGFP constitutively and PP7 stem loops under the control of the p*AGA1* (left) or p*FIG1* (right) promoters at the indicated time, stimulated by pheromone at time 0. Arrows heads indicate when PP7 foci are observed in a cell for the first time.

B. Histograms of the time of PP7 dot detection in the nucleus.

With the two methods, the delay in the expression of *FIG1* compared to *AGA1* is confirmed. However, the absolute time of detection by these assays differs than those from the dPSTR. Indeed, as the dPSTR relies on the expression of a protein to measure a signal, it takes longer time to obtain a visible output from this assay. The delay between mRNA production and relocation of the dPSTR signal corresponds to the translation step and the recruitment into the nucleus. The elongation rate of the ribosomes is around 9AA/s (Shah et al., 2013). The inducible peptide is 309 AA long, which leads to an overall delay of 35 seconds due to the translation step. The newly synthesized protein is then released, and will be able to bind an mCherry molecule of the dPSTR through SZ interaction, and recruit the FP into the nucleus. These two reactions happen at the sub-minute timescale, which should result in a total delay around one minute (Kaffman and O'Shea, 1999; Reinke et al., 2010).

Nevertheless, the PP7 presents major technical drawbacks at the time being. First, there are only two compatible systems (PP7 and MS2), allowing the measurement of only two genes at the time, whereas we hope to measure more promoters with the dPSTR (see next part). Moreover, the MS2 has a weaker affinity for the loops and hence might be less reliable (Wu et al., 2012). Second, the detection of the transcriptional foci requires the acquisitions of numerous stacks of the cells, in order to localize the focus in a focal plan. The repeated illuminations of cells leads to bleaching of the FPs, and can also cause phototoxicity (Logg et al., 2009). Finally, automatized quantification of transcriptional foci is still difficult to achieve. Indeed, the

detection of actual transcriptional foci is delicate because of the presence of floating dots of mRNA in the cytosol. Moreover, the magnification needs to be high enough to allow clear visualization of the dots (60X or 100X), which reduces the number of cells monitored per field of view. However, this technique offers other advantages, like the ability to monitor the position of the transcription site in the nucleus. This is interesting because locus movement to the nuclear periphery upon transcription has been suggested to be involved in transcriptional memory (Chekanova et al., 2008; Tan-Wong et al., 2012).

To summarize, we showed here that the dPSTR provides an efficient measurement of protein expression in real-time, and in addition allows to access to single-cell behavior, which was not possible with other methods.

Part II: Limitations of the dPSTR

As we established before, the dynamic Protein Synthesis Translocation Reporter is able to report on kinetics of gene expression in real-time and at the single-cell level. As a readout, the dPSTR converts gene expression into the relocation of an already mature FP, which is easily quantifiable by automated time-lapse microscopy and computer analysis. However, the design of the dPSTR required several optimization steps, and the data analysis sometimes requires the application of arbitrarily chosen thresholds. This section will describe some limits of the system.

Quantification of the expression output

In order to quantify the level of promoter induction, we calculate the increase in nuclear enrichment before and after stimulation. To do this, we define the basal nuclear enrichment as the average of all time points preceding stimulation, and we find the maximal nuclear enrichment for the dPSTR after stimulation. The difference between these two values is called expression output. Of course, even in cells that do not express the promoter quantified by the dPSTR, there will be fluctuations in the nuclear fluorescence during the experiment. As such, the expression output can be either positive or negative. Hence, to determine whether a cell was expressing the promoter or not, we had to define a certain threshold of expression output above which we would consider that the variation in nuclear enrichment we observe is significant and reflects gene expression and not measurement noise. We determined this threshold arbitrarily, based on the average expression output. We defined that a cell with an expression output equal or higher to 20% of the population expression output is considered as expressing. As measured promoter can have very different induction levels, we cannot use a fixed value of increase in dPSTR nuclear enrichment, but we rather use this value, which is relative to each measured promoter. This threshold may appear high, but it prevents considering non-expressing cells for expressing cells. This is a first bias imposed by dPSTR measurements, but the choice of a threshold value to consider expression is an arbitrary bias that needs to be applied to any fluorescent measurement.

Response time calculation

Once expressing cells have been identified, we can quantify the time at which the expression starts. To do this, a second threshold has to be chosen for the moment at which the cell is considered as expressing. As illustrated in Chapter 2, cell can have very different expression outputs, even within a population bearing the same dPSTR. Hence, it is not possible to determine an absolute threshold regarding the dPSTR nuclear enrichment. Rather, we defined a threshold on the cell relative nuclear enrichment. To do so, all cell traces are normalized based on their own basal and maximal nuclear enrichment, so that they are comprised between 0 and 1. We defined the response time as the moment at which the nuclear enrichment overcomes 0.2, which is 20% of the cell expression output. The threshold used here is different than the one used to determine expressing cells. Indeed, two cells can have the same expression kinetics but different expression levels, due to different expression capacities. By quantifying the response time on normalized cell traces, we ensure that the expression level of the cell does not bias this measurement. The actual bias here is that we do not measure exactly the starting point of expression, which cannot be precisely quantified. Indeed, as there is a certain amount of unspecific fluctuations in the nuclear fluorescence, it is difficult to precisely determine the expression start. However, the threshold we use prevents o mistakenly take a point of unspecific fluctuations for an expression event. Overall, this method provides a measurement of induction kinetics that is comparable between all cells from a population, but also between populations

bearing dPSTRs measuring different promoters. However, the same problem might appear for other reporter measurements like luciferase or FPs.

Detection of low induced promoters

The nature of the signal detected imposes several constraints on the range of promoter expression that can be measured. Indeed, we faced some problems regarding the sensitivity of the dPSTR during its design. For a same amount of FP present in the cell (arbitrarily 100 molecules, with 60 localized in the nucleus and 40 in the cytosol), a promoter with a high expression level that drives the expression of 20 NLS molecules for instance, leads to a total of 80 molecules in the nucleus, which represents an increase of 30% of the nuclear enrichment and is detectable and quantifiable. However, a low induced promoter may produce only 5 NLS molecules, leading to an increase of only 8% of the nuclear enrichment, which may be considered as unspecific fluctuation of the nuclear fluorescence and fall below the detection limit. To counteract this, low expressed promoters should be measured with dPSTR constructs driving low levels of FP in the cell. If the total number of FP is smaller, 40 molecules per cell for instance, with 24 molecules in the nucleus and 16 in the cytosol, the expression of the same low induced promoter will increase the nuclear enrichment by 20%, which can be detected.

To illustrate this, we measured the induction of the promoter p*CDC6* during vegetative growth using the unstable version of the dPSTR. To measure a signal of expression, we had to decrease the expression of the red FP (Figure 5.5).

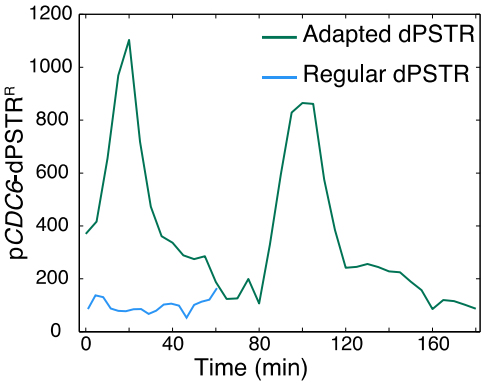


Figure 5.5 Measurement of a low-induced cell cycle regulated promoter

The p*CDC6* is measured by either the dPSTR^R used classically (blue) or a dPSTR with a lower expression of the FP moiety (green).

Alternatively, we can also increase the amount of relocated FP by NLS-SZ molecules produced, by adding more than one SZ sequence to the inducible peptide. For instance, by having 4 SZs by induced molecule, we increased the nuclear accumulation twice, suggesting that the SZ may require longer linkers between them to all be functional (data not shown).

Saturation

The opposite problem can occur: if the amount of FP-SZ in the cell is too low compared to the strength of the promoter of interest, the nuclear enrichment can saturate whereas there is still production of the NLS-SZ part. For instance, we believe that the current construct used for the dPSTR^Y can reach saturation for highly induced promoters. Indeed, in the double mutant *dig1*Δ*dig2*Δ, all Ste12-regulated genes are induced in basal condition (Roberts et al., 2000). However, we observed a pheromone-dependent induction of p*FIG1* measured with the dPSTR^R, but not of p*AGA1* measured by the dPSTR^Y (Chapter 4, Figure 4.7). This is unexpected, as induction of the *dig1*Δ*dig2*Δ mutant leads to expression of *AGA1* (Roberts et al., 2000).

A way to counteract this saturation is to increase the amount of fluorescent protein by choosing a pair of promoter/terminator leading to more protein production (Curran et al., 2013; Knight et al., 2014; Yamanishi et al., 2013). We believe that if the dPSTR^Y has a decreased sensitivity and can saturate more easily than the dPSTR^R, it is due to the amount of MCitrine

molecules synthetized. Indeed, for the induction of the constitutive parts of both dPSTRs, we chose two promoters and terminators that lead to the same expression output (Knight et al., 2014; Yamanishi et al., 2013). We verified this by placing the two combinations of promoter/terminator on identical plasmids, where they drive the induction of the same fluorescent protein. However, the plasmids carrying the two dPSTRs are inserted at different loci, *URA3* for the $dPSTR^R$ and *LEU2* for $dPSTR^Y$. As we will discuss later, the chromatin context of an inserted construct can impact on its expression. It is possible that the *leu2* locus is less favorable to expression than the *URA3* locus. Another way to avoid saturation is to use the destabilized version of the dPSTR, which contains a Ubiquitin sequence that will reduce the stability of the induced peptide (Aymoz et al., 2016; Varshavsky, 1996).

Unfortunately, there is no obvious way to detect a saturation effect of the dPSTR. In our case, we were able to suspect it because we expected a certain induction pattern. But in cases where the user is quantifying a promoter without knowing its outcome, a saturation effect may remain hidden. A good way to address this question would be to compare the dPSTR level to the expression of a fluorescent protein, for instance by using the construct with the inducible part carrying the Venus protein fused to the 2 NLSs (Chapter 2, Figure 2.1)(Aymoz et al., 2016). Indeed, the fluorescence of the Venus protein has a larger dynamic range and a simple correlation with the dPSTR can indicate whether the users need to tune the expression of the FP-SZ part or add a destabilization sequence. This tuning step should avoid having issues on saturation level and detection limit levels.

Part III: Impact of the locus of insertion on gene expression

It has been shown that the position at which a transcriptional unit is inserted can influence its expression level. Indeed, different studies based on integration of the *lacZ* genes at different loci in the yeast genome showed that the resulting B-galactosidase activity varied. Integration at random uncharacterized sites led to a range of 14-fold differences in Miller units (Thompson and Gasson, 2001). Another study that targeted specific integration sites observed an 8.4-fold range of β-galactosidase activity (Flagfeldt et al., 2009). They created two sets of transformants carrying the *lacZ* gene at the very same position, but that differed in the constitutive promoter driving the expression of the enzyme (either p*TEF1* or p*ACT1*, two strong promoters (Mumberg et al., 1995). This allowed them to prove that the expression differences they measured are site-related, and not promoter-dependent. Interestingly, they noticed that integration at two sites separated by only 1.5kb could lead to a 3.5-fold difference in enzymatic activity, illustrating how sensitive a gene can be to its environment.

dPSTR plasmids are inserted at the *URA3* locus for the dPSTR^{R} and at the *LEU2* locus for the dPSTR^{Y} (Figure 5.6). They are built on pSIV backbones, designed in the lab (Wosika et al., 2016). They possess the two transcriptional units for each part of the dPSTR: a constitutive unit for the FP-SZ part, and the inducible unit driving the expression of the 2xNLSs-SZ part, which may carry or not the destabilization sequence (UbiY). The two transcriptional units are in opposite directions, and are separated by the two sequences targeting the construct at the chosen locus. These sequences are 250nt long parts of the terminator or the promoter of the targeted locus, and allow the integration of the plasmid by homologous recombination. On the plasmid, they are separated by the selection marker for bacterial growth, which is cut out of the plasmid by restriction digestion before transformation into yeast. The two dPSTR transcriptional units are separated by the unit encoding the selection marker for the yeast, which is orientated towards the constitutive promoter from the FP unit. Considering how we designed the plasmid, we do not expect to have an effect on the expression of the promoter of interest if transcription arrest from the other transcriptional unit does not occur correctly.

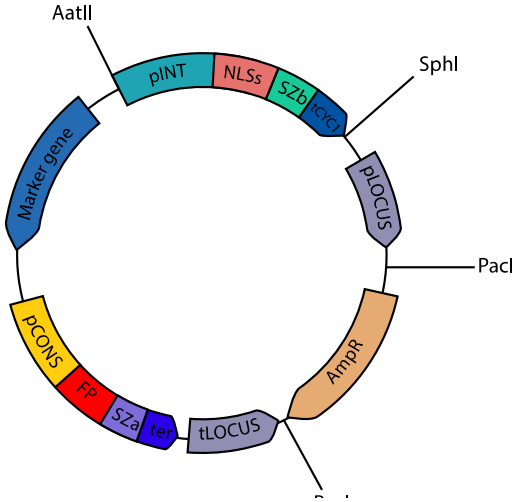


Figure 5.6 Map of a dPSTR plasmid

The dPSTR plasmid is inserted in the *ura3* locus (or the *leu2* locus) for the dPSTR^{R} (resp. dPSTR^{Y}) via homologous recombination between tLOCUS and pLOCUS from the plasmid and from the genome. The dPSTR is composed of two transcriptional units. The FP part is driven by a constitutive promoter (p*RPL24A* or p*RPL24B*), and drives the expression of the mCherry (resp. MCitrine) fused to SZ2 (resp. SZ4). Transcription is ended by the presence of the terminator t*SIF2* (resp. t*NUP53*).
The inducible part of the dPSTR is driven by the promoter of interest, and encodes two Sv40 NLSs fused to SZ1 (resp. SZ3). The terminator t*CYC1* ends the transcription.

As it is carried on a plasmid, the dPSTR quantifies the expression of an additional copy of the promoter of interest. One advantage is that it does not require the deletion of the ORF of the gene of interest to monitor its expression. However, the locus of integration from which the dPSTR quantifies the expression of the promoter is not the endogenous one, which can lead to differential expression.

Endogenously-tagged dPSTR

In order to verify that the dPSTR system reports faithfully on the kinetics and levels of gene expression, we decided to insert the inducible module at endogenous loci of two mating-induced promoters described in Chapter 4, *FIG1* and *AGA1*. Our dPSTR plasmids are inserted either at the *URA3* locus (dPSTR^{R}) or at the *LEU2* locus (dPSTR^{Y}) in single integrations. We designed a plasmid based on the tagging plasmid pGT in order to target the inducible unit of the dPSTR^{R} to any locus, with the *TRP3* selection marker (Wosika et al., 2016). We transformed the PCR product carrying homologous sequences to target the 2xNLS-SZ1 construct in a strain where the constitutive fluorescent part of the dPSTR^{R} had been previously inserted elsewhere. In addition, the corresponding dPSTR^{Y} was also transformed, to directly compare the two outputs.

We can observe that the kinetics of induction differ between the two dPSTRs. Indeed, the endogenously targeted construct leads to faster induction kinetics than the corresponding dPSTR^{Y}, which is true for both p*AGA1* and p*FIG1* (Figure 5.7A and B). Figure 5.7C shows the difference between the response time of the dPSTR^{Y} and the dPSTR^{R} measuring the same promoter, for either the constructs carrying the endogenously targeted dPSTR or the two regular dPSTRs. The strains with the two regular dPSTRs have a distribution centered around 0, reflecting only the intrinsic differences in the expression of two identical copies of p*FIG1*. The distribution of the construct bearing the endogenously tagged dPSTR is shifted towards negative numbers, meaning that the endogenous construct is expressed before the dPSTR^{Y}. Moreover, we observe a difference in the expression output quantified by the two red dPSTRs (Figure 5.7D). Indeed, for both promoters measured the expression output is higher with the endogenously tagged dPSTR^{R}. Another difference lies in the basal level of p*AGA1* induction that we can detect: using the same threshold for expression, with the endogenous construct we can find more cells with a basal induction of p*AGA1* (Figure 5.7E). However, this could reflect a difference in degradation of the pool of inducible peptides rather than in number of expressing cells. Indeed, the number of 2xNLS-SZ1 proteins may be more important when transcription arises from the endogenous locus. Hence, since the temporal window of cell cycle induced expression is similar in all cells, the amount of 2xNLS-SZ1 peptides to be degraded will differ accordingly to the insertion locus. As the degradation rate is expected to be similar for both loci insertions, it takes more time to degrade a bigger pool of protein when the dPSTR is inserted at the endogenous locus. This will lead to a bigger number of cells with a basal nuclear enrichment.

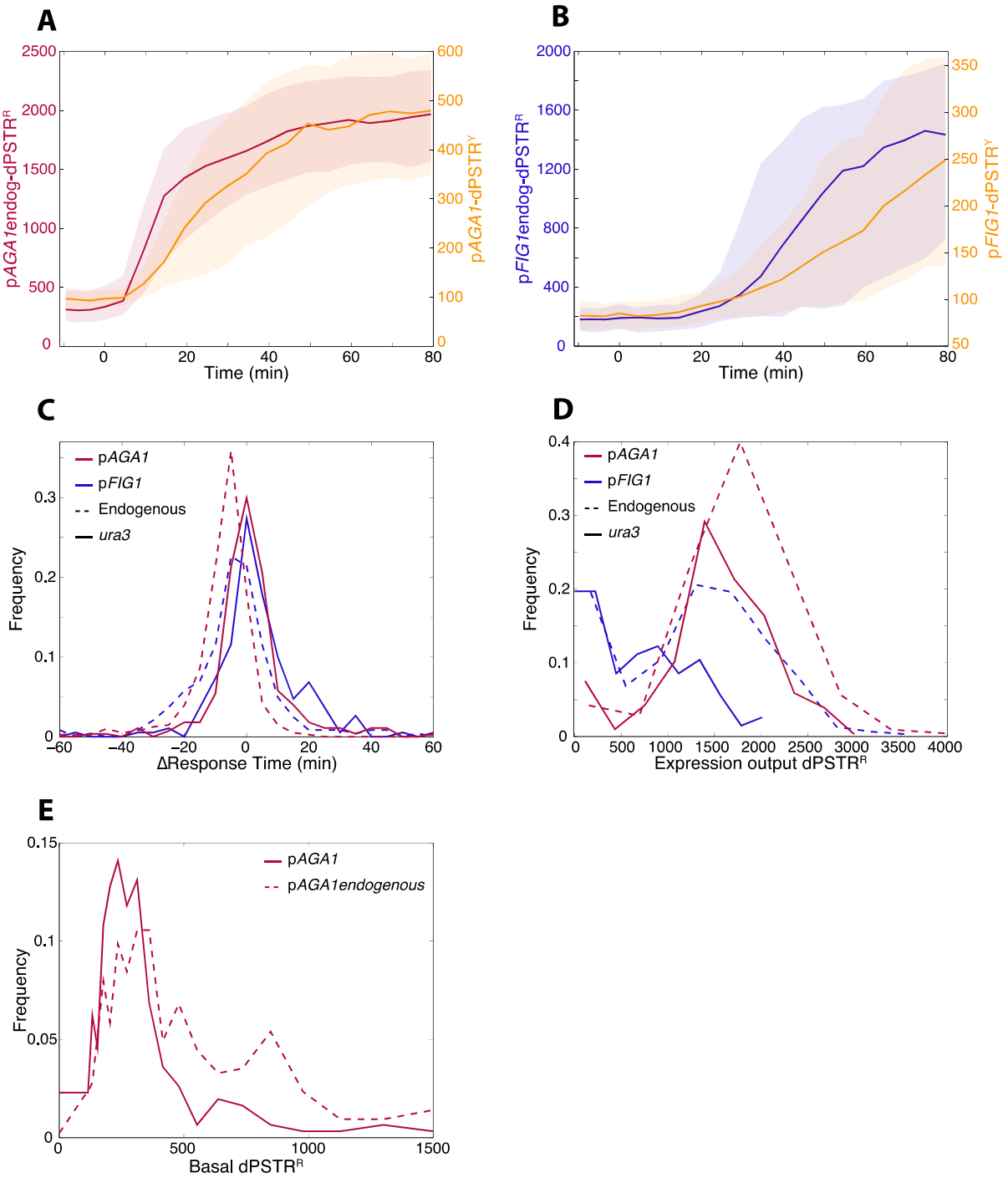


Figure 5.7 Endogenously tagged dPSTR

The inducible part of the dPSTR was inserted downstream of p*FIG1* or p*AGA1*, at their endogenous loci.
A-B. Nuclear enrichment of the endogenously located dPSTR^R compared to the regular dPSTR^Y inserted in the *leu2* locus, measuring the expression of p*AGA1* (A) or p*FIG1* (B), upon pheromone stimulation.

C. Quantification of the difference between the response time of the dPSTR^Y and the dPSTR^R for endogenous constructs (dotted lines) or regular constructs in the *ura3* locus (filled lines), for p*AGA1* (red) and p*FIG1* (blue).

D. Expression output measured by the dPSTR^R in various configurations (same legend as in C).

E. Quantification of basal level of p*AGA1* expression measured from the endogenous locus (dotted line) or the *URA3* locus (filled line).

These differences can also be quantified through expression noise measurements (Figure 5.8). We can see that the noise between the two dPSTRs inserted in the *LEU2* and *URA3* loci is lower than if the endogenous promoter is targeted. For p*AGA1*, there is a higher noise before stimulation because the endogenously tagged construct is more sensitive to low cell cycle dependent expression of this promoter (Figure 5.7E; Figure 5.8A). The behavior of the intrinsic noise following pheromone addition is similar. For p*FIG1* with the endogenous dPSTR, the noise is higher before induction and decreases faster than with the insertion at the *URA3* locus, because we detect induction at the endogenous locus before (Figure 5.8B).

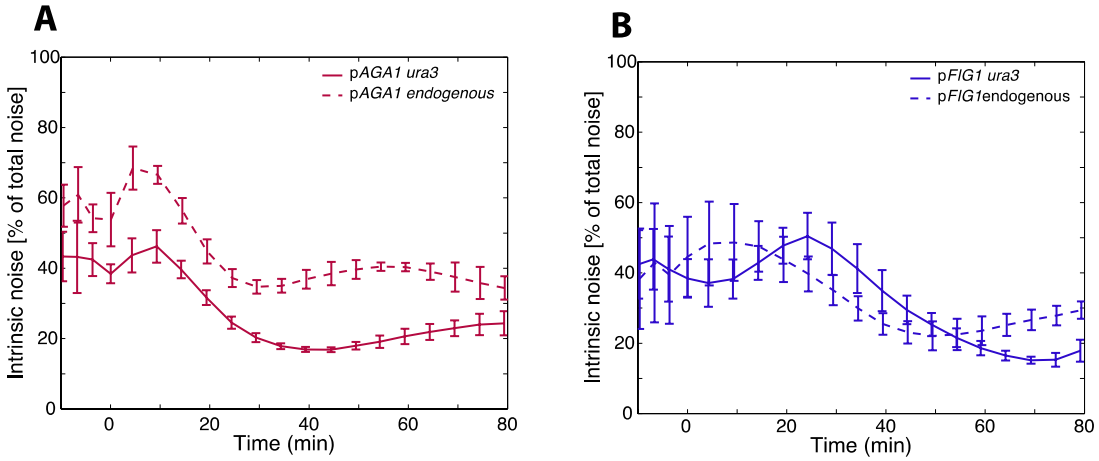


Figure 5.8 Intrinsic noise

Intrinsic noise evolution in response to pheromone stimulation between the dPSTR^{R} in the *URA3* locus (filled lines) or the endogenously located dPSTR^{R} (dotted lines) compared to the dPSTR^{Y} inserted in the *LEU2* locus, measuring the expression of p*AGA1* (A) of p*FIG1* (B), upon pheromone stimulation.

To sum up, it seems that the location of the inducible dPSTR moiety impacts on the expression measurements. However, the construct we used in Figures 5.7 and 5.8 can be used to tag any promoter with the inducible part. Yet, it requires deletion of the target gene, which can be a problem in some instances. Hence, there is here a choice that has to be made. In the studies presented throughout this thesis, we always used the dPSTR targeted at the *URA3* or *LEU2* loci. Hence, all our measurements are comparable because biased in the same direction. Indeed, with measurements at the endogenous loci, we confirmed the kinetics differences between p*FIG1* and p*AGA1*. However, the expression from the endogenous locus is slightly different in terms of kinetics and levels. Nevertheless, the dPSTR allows to compare induction kinetics of promoters integrated in the same locus, which still allows to understand mechanisms of gene expression regulation.

Many features can explain why a unique construct integrated at various loci could be differentially expressed, and will be detailed below.

Nucleosome landscape

One major determinant of the transcription is the nucleosomes position at a promoter. The nucleosome landscape can be different for the same construct integrated at different positions. A study proposed that roughly half of the nucleosome positions were encoded in the yeast genome (Segal et al., 2006). This suggests that a construct inserted at a different locus than the endogenous one may have a different nucleosome organization, which could lead to dramatically different induction kinetics. Indeed, the recruitment of the polymerase may be

easier or harder depending on histones position, which may hide or free some regulatory sequences. Segal *et al.* developed a model for the prediction of nucleosome positioning in the yeast genome, which can also be applied to more complex organisms, like *C. elegans* (Field et al., 2008; Kaplan et al., 2009; Segal et al., 2006). We used this model to predict the probability for each nucleotide of p*AGA1* or p*FIG1* to be covered by a nucleosome, whether it is at its endogenous locius or in the inserted dPSTR plasmid. We also compared these predictions to the histone positions obtained by DNA cleavage around histones coupled with sequencing experiments by Brogaard *et al.* (Brogaard et al., 2012). The model does not predict significant differences between the two loci of insertion (Figure 5.9). This suggests that the nucleosome landscape might be equivalent in the dPSTR plasmid inserted in the *URA3* locus and in the endogenous locus. This will be verified by our collaborators in Barcelona through MNase assay, in a strain containing the p*FIG1*-dPSTR^{R} and deleted for the endogenous p*FIG1*. However, our comparison reveals that the model does not correctly predict the nucleosome landscape on our promoters. Indeed, only the -2 nucleosome of p*AGA1* is strongly predicted by the model. This highlights the need to confirm experimentally nucleosome positions for further analysis.

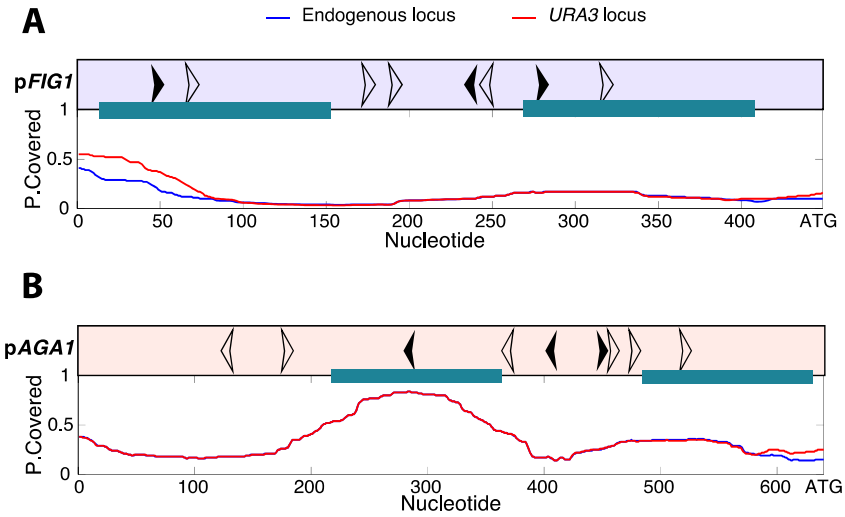


Figure 5.9 Nucleosome prediction at p*FIG1* and p*AGA1*

Model prediction of the nucleosomes positions at p*FIG1* (A) or p*AGA1* (B) located at their endogenous loci (blue) or in the dPSTR plasmid inserted in the *URA3* locus (red) (Model from (Kaplan et al., 2009)). The probability of each nucleotide to be covered by a nucleosome is plotted.

Impact of neighboring genes

The surrounding genes may be more or less transcribed, which could impact on the transcription of the gene of interest. Indeed, the presence of the transcription machinery close to the gene of interest may allow chromatin remodeling that could facilitate its induction. The ORF of *FIG1* is surrounded by two genes oriented in the same direction as *FIG1*, *ATP3* upstream and *FAT1* downstream. They both encode proteins involved in metabolism (resp. ATP and acyl-CoA production). They are strongly expressed in normal growth conditions, as each protein is present in several thousands of copies per cell (Ghaemmaghami et al., 2003) (YeastGFPDatabase). The *AGA1* gene is downstream of the *MVD1* gene and upstream of the *PET494* gene, all in the same direction. Similarly, the function of the encoded proteins suggests a constitutive expression, and Mvd1 proteins are present in 13700 copies per cell, whereas PET494 is present in 450 copies per cell (resp. an essential enzyme involved in sterol synthesis, and a mitochondrial translational activator)(YeastGFPDatabase, (Ghaemmaghami et al., 2003)). We can imagine that if termination of the upstream gene fails, the expression of the gene of interest may occur indirectly. However, in the dPSTR plasmid, the inducible part is facing the promoter of the *URA3* gene. Upstream of the *URA3* gene, there is a more than 4kb-long gene in the same direction, *GEA2*, strongly expressed, as there are 7700 molecules per cell (YeastGFPDatabase, (Ghaemmaghami et al., 2003)). Upstream of the inducible part, there is the selection marker gene, transcribed in the opposite direction. It is really unlikely that in this

locus the transcription of the inducible part is modified by transcription of surrounding genes. However, the context is different between endogenous loci and *URA3* locus, which can explain the changes we observe.

Impact of the terminator on gene expression

The 3'-UTR region of a gene may have a strong impact on its expression levels. Indeed, it was shown that a FP gene driven by a constitutive promoter could lead to various fluorescence intensity levels when different terminators are used (Curran et al., 2013). This study identified 30 terminators that yielded to more protein expression than the *CYC1* terminator, classically used. They showed that this was the result of an increase in the half-life of the mRNA. For an identical transcriptional rate, the amount of mRNA molecules present in the cell is increased if the mRNA is more stable. This leads to more translation events per mRNA molecule, and hence more proteins produced in the end. Another study assessed the effect of all the terminators of the budding yeast genome (Yamanishi et al., 2013). They confirmed that the *CYC1* terminator was leading to low levels of protein expression. In comparison, the *AGA1* terminator leads to 2 times more protein expression, and *FIG1* terminator to 3 times more. However, we use the *CYC1* terminator in all constructs of the inducible part of the dPSTR, so that all our measurements are comparable and truly reflect the variability between promoters.

To sum up, the location of the dPSTR system in the genome may impact on the kinetics and levels of expression measured. This hinders absolute quantifications. However, since the locus of insertion is always identical, comparisons between promoters or in different backgrounds can be done.

Discussion and perspectives

The system we developed allowed us to address multiple questions with more depth than previously. We could show that even though hyper-osmostress signaling and gene expression are correlated at the population level, this is not the case at the single-cell level (Chapter 2). In another topic, we were able to quantify a timeline of genes expression in response to a single stimulus, in the pheromone-driven response (Chapter 3). And finally, we used the dPSTR to understand the transcriptional rules regulating induction kinetics in response to pheromone signaling (Chapter 4). Overall, the dPSTR allowed us to answer questions on a dynamic scale that could not be reached before.

Because of the modularity of the dPSTR system, it is easily adaptable to other organisms. Indeed, NLSs are functional in all eukaryotic cells (Kaffman and O'Shea, 1999). Fluorescent proteins are also functional in any organisms, and SynZips are expected to interact in the same manner. However, the vector carrying the construct needs to be adapted to the target organism, as well as promoters driving the induction of both transcriptional units and their terminators. Also, there may be need for optimization regarding the level of constitutive expression of the FP moiety, in order to avoid problems of saturation and detection limit described before. Moreover, for the dPSTR*, the destabilization sequence needs to be modified accordingly to the target organism. The half-life of the protein changes drastically according to the organism (Gonda et al., 1989). Moreover, in *S. pombe*, the Ubi sequence needs to be replaced by an N-degron sequence (Kitamura and Fujiwara, 2013).

Relocation to other cellular compartments

When designing the dPSTR, we thought to relocalize the fluorescent protein into the nucleus because we knew we would be able to easily quantify this relocation signal with our experimental settings. However, it is possible to target any other cellular compartment, providing that a peptide sequence allowing this targeting exists. For instance, we are developing a dPSTR able to relocalize at the plasma membrane upon gene expression. We replaced the NLSs part by a C-terminal RitC peptide, which attaches to the plasma membrane (Figure 5.10) (Bendezu et al., 2015).

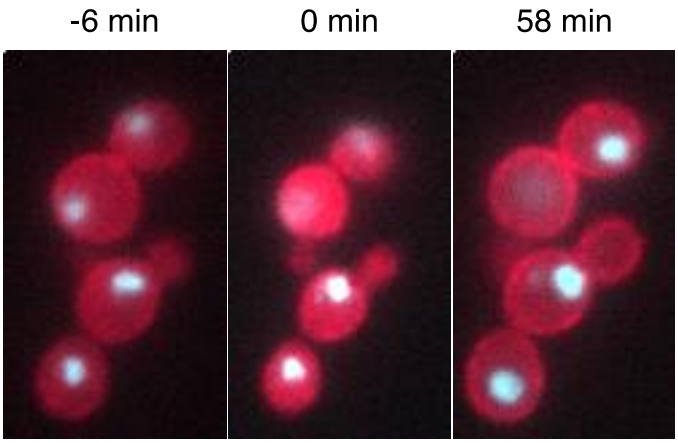


Figure 5.10 Membrane-targeted dPSTR

Membrane-targeted dPSTR^{R} quantifying the expression of p*GPD1* promoter

We can see that the version we built to measure p*GPD1* is relocated at the membrane in basal conditions, as expected since this promoter is constitutively induced. After osmotic stress, we can clearly see that the membranar signal is higher than before stress, indicating that the dPSTR measured the induction of the p*GPD1*. Further optimizations are required to improve the membranar signal and its automatized quantification.

Enlightening dPSTR

An alternative to the relocation signal would be to observe appearance of a fluorescent signal. The ddFPs provide the ability to promote this and exist in three colors, red, yellow or green (Alford et al., 2012a; Alford et al., 2012b; Ding et al., 2015). The ddFP comports two

non-fluorescent monomers that can dimerize into a fluorescent protein. The two halves of the ddFP can be fused to non-interacting SynZips (SZ1 and SZ3 for instance). The promoter of interest can drive the induction of a peptide containing two non-interacting SZs that can bind to those fused to the halves of the ddFP. In basal conditions, the two halves of ddFP are produced but cannot interact, and no fluorescence is measurable. Upon expression of the gene of interest, the SZs interaction promotes the association of the two halves, which leads to the appearance of a fluorescent signal. In HELA cells, this system was used to quantify dynamically the interaction between the calmodulin and an interacting peptide (Ding et al., 2015). The appearance of fluorescent signal occurred in roughly 1 minute, and oscillation behavior could be observed, showing that this system is highly dynamic and reversible. With the appropriate choice of SynZips, this system could be coupled to two classical dPSTRs, increasing the number of promoters that can be dynamically quantified within a same cell.

Measurements of several promoters with one dPSTR

Thanks to the variety of SynZips available, we can imagine many interacting networks for the dPSTR (Reinke et al., 2010). With the help of a first-step master student, Mathieu Clément, we designed a cell cycle sensor based on the dPSTR^R system relocating into different cellular compartments at different cell cycle stages. A promoter induced in G1 drives the relocation of the FP in the nucleus, like the classical dPSTR. A second promoter, induced in G2, drives the relocation of the FP at the membrane. In between, during S and M phase, the FP is cytosolic. This sensor requires three SynZips: one for the FP (SZa), and one for each relocating peptide (SZb and SZc). The SynZip of the FP, SZa, can interact with SZb and SZc, which cannot interact with each other. Unfortunately, our student did not obtain a fully functional cell cycle sensor because of lack of time.

dPSTR in other organisms

In collaboration with Aleksandar Vjestica from the lab of Sophie Martin, we already started to translate the dPSTR to the fission yeast *Schizosaccharomyces pombe* (Figure 5.11). Cells of the two mating types produce each a non-functional TF, and the two proteins interact together after fusion to drive genes involved in sporulation (Kelly et al., 1988; Willer et al., 1995). Interestingly, it seems that these two TFs diffuse with different kinetics throughout the fusion pore. The dPSTR will allow to test whether resulting gene expression arises asymmetrically or from both nuclei at the time (Vjestica *et al*, in prep.).

Figure 5.11 dPSTR in *S. pombe*

Fission yeast cells carrying the p*MAP3*-dPSTR^R were grown on plate in presence of mating partners for one day before imaging.

We also started a collaboration with the group of Jan van der Meer in order to translate the dPSTR to the bacteria *Escherichia coli*. We used the strategy of targeting the fluorescent protein to the plasma membrane upon expression of the promoter of interest. However, we are worried that the quantification of the relocation signal might be a challenge because of the size of the cells. This work is still on-going.

We are also working on translating the dPSTR method to higher eukaryote cells (see general discussion).

To sum up, we believe that the dPSTR may be used in the future as a routine system in various studies and in many organisms.

Methods

Strains and plasmids.

Yeast strains are listed in Supplementary Table 1 and plasmids in Annex 1.

Other methods

Strain construction was performed as described in Chapter 2 to 4.

Sample preparation, microscopy experiments, and data analysis was performed as explained in Chapter 2 to 4.

Northern blot experiments were performed as described in Chapter 4.

PP7 assay was performed as explained in Chapter 2.

Expression noise calculation was performed as in Chapter 2 and 3.

Supplementary table

Supplementary Table 1: List of yeast strains used in Chapter 5

Strain	Background	Genotype	Plasmid
ySP2	W303	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*	
ySP580	W303	*HTA2-CFP:HIS3*	
ySP643	ySP641	*HTA2-CFP* *URA3: pFIG1-dPSTR$^{R2\text{-}1}$* *LEU2: pAGA1-dPSTR$^{Y4\text{-}3}$*	 pSP366 pSP368
yDA347	ySP580	*HTA2-CFP* *Ste12-9myc::NAT* *Kar4-6HA::HPH* *URA3: pAGA1-dPSTR$^{R2\text{-}1}$* *LEU2: pAGA1-dPSTR$^{Y4\text{-}3}$*	 pSP365 pSP368
yDA341	ySP580	*HTA2-CFP* *Ste12-9myc::NAT* *Kar4-6HA::HPH* *URA3: pFIG1-dPSTR$^{R2\text{-}1}$* *LEU2: pAGA1-dPSTR$^{Y4\text{-}3}$*	 pSP366 pSP368
ySP377	W303	*HTA2-mCherry:URA3* *TRP3:pMET25-PP7-2xGFP* *GLT1: HIS3-pFIG1-24xPP7-SL*	 pSP268 pSP267
ySP725	W303	*HTA2-mCherry:URA3* *TRP3:pMET25-PP7-2xGFP* *GLT1: HIS3-pAGA1-24xPP7-SL*	 pSP268 pSP409
yDA353	W303	*HTA2-mCherry:HIS3* *URA3*: p*GPD1*-RitC-dPSTR^{R}	 pDA386
yDA280	W303	*HTA2-mCherry:HIS3* *URA3:* pRPL24A-mCherry-SynZip2 *LEU2: pAGA1-dPSTR$^{Y4\text{-}3}$* *AGA1:*2xNLSs-SynZip1:*TRP*	 pDA157 pSP368
yDA281	W303	*HTA2-mCherry:HIS3* *URA3:* pRPL24A-mCherry-SynZip2 *LEU2: pFIG1-dPSTR$^{Y4\text{-}3}$* *FIG1:*2xNLSs-SynZip1:*TRP*	 pDA157 pSP367
yDA172	W303	*HTA2-mCherry:HIS3* *LEU2:* p*CDC6*-dPSTR^{R2,1}	 pDA251
ySP37	W303	*HTA2-CFP*	
yDA119	ySP37	*HTA2-CFP* *URA3: pSTL1-Venus-dPSTR$^{R2\text{-}1}$* (*)	 pDA176
ySP9	W303	*LEU2: pSTL1-quadrupleVenus*	pSP34
ySP646	ySP644	*HTA2-CFP* *URA3: pFIG1-dPSTR$^{R2\text{-}1}$* *LEU2: pFIG1-dPSTR$^{Y4\text{-}3}$*	 pSP366 pSP367
ySP642	ySP641	*HTA2-CFP* *URA3: pAGA1-dPSTR$^{R2\text{-}1}$* *LEU2: pAGA1-dPSTR$^{Y4\text{-}3}$*	 pSP365 pSP368
yBB10	*S. pombe*	h+ *ADE6*::p^{pom1}:mCherry-linker-SynZip1:term^{sif2} *URA4*+::p^{map3}:2xNLS-linker-SynZyp2:term^{nmt}	 pBB3 pBB7

Summary of the main results

In this thesis, we report on the development of a new class of gene expression reporter. The dynamic Protein Synthesis Translocation Reporter, or dPSTR, allows quantification of levels and dynamics of gene expression in living single cells. Based on the relocation of a fluorescent protein as readout for promoter activation, it reaches a precision at the sub-minute timescale, because it bypasses the maturation time of the fluorophore. Time-lapse microscopy combined to automated image analysis allow the quantification of hundreds of cells per conditions, leading to statistically relevant results. As of today, two compatible versions of the dPSTR exist, with orthogonal pairs of SynZips and different fluorescent proteins, which can be combined within the same cells. This system allows combinatorial measurement of multiple promoters for periods ranging up to several cell cycles. Thanks to the introduction of a destabilization sequence, it is possible to quantify transient expression events.

In the first chapter, we quantified gene expression induced by hyper-osmotic shock in the budding yeast. This stress results in immediate and transient MAPK signaling and gene expression, which we were able to quantify concurrently in the same cells. We showed a lack of correlation between signaling activity and the resulting gene expression in single-cells. Using the two colors of dPSTRs, we quantified for the first time expression noise in real-time in single-cells. We showed that the intrinsic noise of the inducible p*STL1* is higher than this of p*GPD1*, which is induced under basal growth conditions. We propose that this higher noise is due to stochastic events of chromatin remodeling that can arise randomly at either copy of the promoter. Additionally, we were able to show a very short delay of 2 minutes in the induction of p*STL1* compared to p*GPD1* that can be attributed to the chromatin remodeling step. Finally, because of the high turnover rate of the dPSTR, we could quantify different rounds of signaling activity and consecutive gene expression triggered by successive stresses applied in microfluidic chips. Here, we showed a lack of correlation between both signaling and gene expression driven by the two stresses at the single-cell level, suggesting that these genes do not conserve a memory of previous expression events under the condition tested.

In the second chapter, we used the modularity of the dPSTR to dynamically measure the expression of 14 mating-induced promoters. The mating is driven by sensing of pheromone from the opposing mating type, and eventually leads to the formation of a diploid zygote. It is a typical example of cell-fate decision process, where different signals will be integrated by the cells to control the decision to mate or pursue vegetative growth. Pheromone sensing will lead to the activation of the mating MAPKs, which controls the establishment of a transcriptional program via de-repression of the main mating transcription factor Ste12. We showed that despite a signaling activity occurring minutes after addition of exogenous pheromone, resulting gene expression occurs on different timescales. We divided the tested promoters in 3 categories. Early-induced genes are encoding proteins involved in the commitment to the mating process, and are induced between 10 to 20 minutes after stimulation. Some genes have a slightly delayed induction compared to the early genes, and we qualified them as intermediate. They encode proteins involved in the fusion step. Finally, late-induced genes are expressed from 25 minutes on after stimulation, and encode proteins involved in the last steps of the mating, fusion and karyogamy. We showed that cells were able to control the expression of these genes based on the pheromone concentration. Early-induced genes are expressed at low pheromone

concentrations, whereas late-induced genes require high pheromone concentrations. We confirmed that this chronology in gene expression was also occurring in mating mixture

In the next Chapter, we deciphered the regulation of the temporal classes of promoters. We showed that mating-induced transcription requires intact signaling activity, as well as the presence of Ste12. We also found that induction of late genes was dependent on high MAPK activity. We showed a requirement for the presence of an additional transcription factor, Kar4, for induction of late genes and correct temporal induction of the intermediate genes. We showed in Chapter 2 that this TF is among the early-induced genes. We then assessed how the promoters were able to control the timing of induction. We showed that histone remodeling occurs at different moments at early and late promoters, accordingly to their expression time. Moreover, using synthetic constructs, we built an array of variants of two prototypical promoters, the early-induced p*AGA1* and late-induced p*FIG1*. We modified the numbers, orientations and positions of the Ste12 binding sites in order to understand how they control gene induction kinetics. We found that there is an interplay between histone positioning and presence of binding sites. We propose a model in which the presence of pairs of Ste12 binding sites in nucleosome-depleted region leads to induction during early times, whereas pairs containing a non-consensus binding site requires the help of Kar4 to promote induction. We think that the chromatin remodeling step is controlling the expression time, and requires the stabilization of a Ste12 dimer and this may require the presence of Kar4 at late-induced genes.

Overall, we developed and use a method that allowed us to perform dynamic measurements of expression, which offers new mechanistic insights in MAPK pathway-induced gene expression.

General discussion & perspectives

In Chapter 3, we established that pheromone sensing in budding yeast initiates a dynamic transcriptional program with target genes expressed at different moments. In Chapter 4, we tried to understand how these expression kinetics are controlled through genetic studies and systematic mutations of promoting sequences. Overall, we found that temporal regulation is governed by pheromone concentration sensing, and that cells need a prolonged exposure to pheromone to induce the full mating transcriptional program. We also found that the presence of a TF, Kar4, was required for the expression of the latest genes.

Grammar of transcription regulation by Ste12

We designed and built several variants of p*AGA1* and p*FIG1* promoters in order to understand what rules over Ste12 binding-sites organization would control the induction kinetics of the genes. We understood some of these rules, but it would be interesting to get a broader sense of them.

A synthetic approach could be used here, with the fusion of a core promoter with upstream regulatory elements. This approach was already used by Su *et al.*, where they fused a minimal *GAL1* promoter to various combinations of PREs (Su et al., 2010). However, the number of tested constructs was fairly low and did not recapitulate every possible PREs organization. In another study, a library of thousands of synthetic promoters systematically designed to offer a wide range of variety of number, spacing, affinity, types, and accessibility of regulatory elements was designed and built (Raveh-Sadka et al., 2012; Sharon et al., 2012; van Dijk et al., 2017). The library was synthetized using Agilent programmable microarrays and cloned upstream of a core promoter driving a FP-coding gene (Cleary et al., 2004; Sharon et al., 2012). With the high number of promoters obtained, it is possible to find pairs of them with only one varying element at the time, which allows the definition of transcriptional rules. In our context, the design of such a library could be really interesting. We could imagine to test different affinities of PREs, by using different sequences, but also test the impact of the distance between PREs, as well as their orientations, their numbers, their position relative to the transcription start site, and their accessibility via use of poly(dA:dT) tracts, destabilizing nucleosome formation (Segal and Widom, 2009). This would allow a precise understanding of important organizational constraints required for efficient pheromone-induced response.

However, measuring thousands of promoters with the dPSTR is also quite challenging. Indeed, the previous studies were using flow cytometry of mixed population, allowing to group the measurements (Raveh-Sadka et al., 2012; Sharon et al., 2012; van Dijk et al., 2017). The dPSTR requires microscopy analysis, as we need to quantify the difference of fluorescence between the nucleus and the cytosol of the cells, which is not possible with the use of flow cytometer. High-throughput fluorescence microscopy would allow the automated quantification of this promoter library in real-time in response to pheromone stimulation (Falconnet et al., 2011; Pepperkok and Ellenberg, 2006; Styles et al., 2016).

A role for the timeline of mating-induced transcription

We described that mating-induced transcription was temporally controlled following pheromone sensing. We think that the cell controls this timeline through pheromone concentration sensing and MAPK activity levels. As this phenomenon was also observed in real

mating conditions, in presence of a mating partner, we believe that it is required to mate successfully. However, this should be proven experimentally.

One approach we tried to test this was to delay the expression of one of the early-induced genes. However, this approach requires the identification of a good target gene to observe a significant phenotype. Unfortunately, we could not identify a single gene deletion that results in a strong fusion phenotype.

Another approach to illustrate the importance of the mating-induced transcription timeline is to completely disturb it, by inducing global transcription kinetics changes. Hypothetically, there are three ways to perturb the transcription timeline.

One way is to delay the induction of all early genes, to match the induction of late genes. We could imagine that the mating could either occur with a delay, which would be deleterious in terms of fitness in case of a mating competition assay. Alternatively, there is also a chance that the mating could not happen in this condition. Indeed, a delay in the induction of all early genes would mean a delay in the induction of Far1. This could have dramatic impacts on the ability of the cell to mate. The absence of Far1 in the early mating response could lead to misplacement of the polarization site, resulting in failure to find the partner. It can also lead to a delay in the attempt to arrest cell cycle, which may happen at a point when the cell already pursued its cycling.

Another way to disturb the mating-induced transcription timeline would be to accelerate the timing of induction of late genes, so that they are induced early in the mating response. If proteins involved in the fusion process are present before correct establishment of the polarization site, we can imagine that the cell may attempt to fuse at random sites, without any partner. Moreover, a genome-wide study of overexpression found that the overexpression of the *CIK1* gene is lethal (Sopko et al., 2006). This gene encodes a protein involved with Kar3 in the process of nuclei fusion, and is co-regulated with Kar3 (Kurihara et al., 1996; Page et al., 1994; Sproul et al., 2005). This suggests that a premature expression of genes involved in the karyogamy may be deleterious for the cell.

Otherwise, further delay the induction of late-induced genes may also lead to interesting phenotypes. Indeed, if the cell is ready to fuse but cannot because it lacks some key components of the fusion machinery, how will it react? This can lead to a delay in the fusion time or even a failure of fusion, because of resuming of the vegetative growth.

However, to achieve these global genetic manipulations, there are some technical challenges. In order to btain a delay in the expression of early-induced genes it would be necessary to modify one by one all the promoters of these genes. This could technically be achieved by genome editing, at a very costly price. Moreover, this would require the knowledge of all early-induced genes among the more than 200 mating-induced genes. Instead, modifying the induction kinetics of late-genes might be easier to attain. It was reported that overexpression of either the long or the short forms of Kar4 prior to pheromone stimulation leads to earlier appearance of mRNA of the gene *PRM2* (Lahav et al., 2007). We are currently testing whether it also leads to earlier induction of p*FIG1*. If so, we could imagine to induce pheromone signaling in cells that already have a pool of the shorter form of Kar4, driven by the p*GAL1* promoter. This could result in the earlier activation of the set of late-induced Kar4-dependent genes, and it would be possible to observe whether the mating processing is altered. However, the amount of Kar4 present in the cell probably matters for the regulation of its targets, and it is hard to predict the amount of proteins driven by the p*GAL1*. On the opposite, we can also imagine delaying the induction of the short form of Kar4 by placing this gene under the control of one of the slow p*AGA1* variants we obtained, so that the expression level reached remains comparable to the normal induced levels of Kar4. This would lead to a situation similar as in a *kar4*Δ background until the expression of the short form of Kar4, which should delay the

induction of late genes. However, it was observed that *kar4*Δ cells are able to perform all mating steps until the karyogamy process (Kurihara et al., 1996).

Overall, this type of global genetic manipulations could help to understand and prove the importance of the mating-induced transcription chronology.

The dPSTR as a universal gene expression tool

We initially developed the dPSTR system in the budding yeast, a model organism easy to genetically modify. This way, we could optimize the different parts of the dPSTR to get a well-functioning system that we can now envision to translate to other organisms. As mentioned in Chapter 4, we are working on translating the dPSTR system into the fission yeast *Schizosaccharomyces pombe*, in collaboration with Sophie Martin's lab. Meanwhile, a postdoc in the group (Min Ma) is currently working on translating the dPSTR system into mammalian cells. Preliminary results show that the stable version of the dPSTR is functional in Hela cells (Figure 6.1). We built a dPSTR measuring the activity of the tumor suppressor p53, by using a promoter containing an array of p53 binding sites (Purvis et al., 2012; Zilfou and Lowe, 2009). However, in this specific example, the use of the dPSTR over a classic fusion with a FP does not present much advantages, since the kinetics of gene expression are slow enough to not being impaired by the maturation time of the FP. Rather, we envision that the dPSTR could be used for studies in more complex systems.

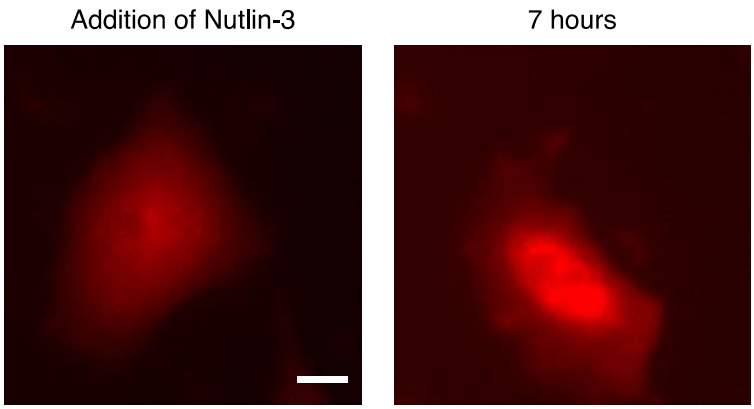


Figure 6.1 Mammalian dPSTR

A mammalian dPSTR^{R} was built and transfected in HELA cells. The promoter measured reports on the activity of p53. Cells were stimulated at the beginning of the experiment with Nutlin-3 and imaged for more than 24 hours. Scale bar is 10µm.

Timeline of gene expression in development

The study we present on the chronology of gene expression induced by yeast mating strongly resonates with the field of developmental biology. During embryos development, gradients of morphogens are controlling the fate chosen by cells that will differentiate into functional organs, and eventually give rise to a functional organism. Morphogens are defined as signaling molecules that are secreted in the embryo, either around the cells or within multinucleated cytoplasm called syncytium (Ashe and Briscoe, 2006; Christian, 2012; Gurdon and Bourillot, 2001; Rogers and Schier, 2011). Gradients arise from a single source and can form through mechanisms of diffusion of the morphogen itself, or by diffusion of the mRNA that encodes it, which, upon translation at different locations, leads to an amplification of the gradient. Active at very low concentrations, morphogens require a signaling cascade to trigger downstream events. The morphogens gradients drive different patterns of gene expression within neighboring cells, by driving activation of TFs depending on the concentration sensed by the cell.

A typical example is the embryonic development of the frog *Xenopus* that requires the formation of gradients of morphogens like activin (Gurdon and Bourillot, 2001; McDowell and Gurdon, 1999). This occurs within hours after fertilization, and is required for mesoderm formation (McDowell and Gurdon, 1999). It was shown that activin gradient can form in less than 2 hours over a distance covering at least 10 cells through passive diffusion (Gurdon et al.,

1994). The cells in the gradient are sensing the activin concentration through evaluating the absolute number of bound receptors (Dyson and Gurdon, 1998). The receptor complexes phosphorylate the target TF, Smad2, which is able to translocate into the nucleus and drive gene expression (Bourillot et al., 2002; Dyson and Gurdon, 1998). The amount of active TF will directly depend on the number of occupied receptor, reflecting the morphogen concentration.

The mechanisms of regulation of the pheromone-induced gene expression timeline in budding yeast can provide new trails to understand how morphogens sensing controls gene expression patterns. However, there are already several hypotheses about how the cell can activate different transcriptional programs according to its position within the gradient (Reviewed in (Ashe and Briscoe, 2006)). First of all, different genes can be regulated by different affinities of the TF for binding sites. At low morphogen concentrations, there is a low amount of TF activated, and they bind only to their highest affinity binding sites. However, in cells exposed to higher morphogen concentration, the amount of active TF is sufficient for the TF to start binding to lower affinity binding sites, hence inducing additional target genes. Another regulation can occur through combinatorial inputs, combining the TF targeted by the morphogen with other TFs specific to the cell type for instance. Alternatively, there can be proteins expressed by a first transcriptional wave at the lowest concentrations that will promote induction of other targets (Ashe and Briscoe, 2006). Our study on the mating-induced transcription can provide a specific example of temporal regulation driven by a single molecule, in a simplified model system. We hope that our results will, in the future, inspire researchers in the field of developmental biology.

Moreover, with the dPSTR system, it is now possible to address gene expression kinetics in real-time. This could be interesting knowing that in the case of the *Xenopus*, in which the gradient formation and resulting expression occur on a short timescale of a few hours (De Robertis et al., 2000; Dyson and Gurdon, 1998). However, *Xenopus* embryos observation by fluorescent microscopy can be challenging due to the yolk auto-fluorescence (Saka et al., 2007). The *Zebrafish* embryo is another example in which morphogen gradient drives patterning within the few hours following fertilization (Harvey and Smith, 2009). In these two examples, fluorescent microscopy could be applied to monitor gene expression in the single cells of an embryo. Alternatively, it is possible to dissociate blastula cells from *Xenopus* embryo and expose them to specific morphogen concentrations (Green et al., 1992; Gurdon et al., 1999). This method allows to culture them at various density on fibronectin substrate, which is more suitable for microscopy experiments.

Ultimately, we can imagine to couple various sensors to measure morphogen gradients and their effect in embryos. Two studies reported the construction of a sensor of the activated form of Smad2, which reflects the gradient of Activin in *Xenopus* or Nodal in *Zebrafish* (Harvey and Smith, 2009; Saka et al., 2007). The combination of these types of sensors to the dPSTR would allow direct correlation of gene expression levels and TF activity.

Conclusion

As we illustrated throughout this thesis, the regulation of gene and protein expression is a very important aspect of cellular life, in all organisms, from yeast to mammals. The mechanisms of regulation of expression are extremely various and sometimes complicated, and often still not understood. We also showed that a dynamic study of the gene expression, and at the single cell level, is crucial to allow a full understanding of the regulation of transcription. We hope that in the future the dPSTR system will provide many insights on the mechanisms of gene expression regulation, and in a variety of model systems.

Annex I: List of the plasmids used in the thesis

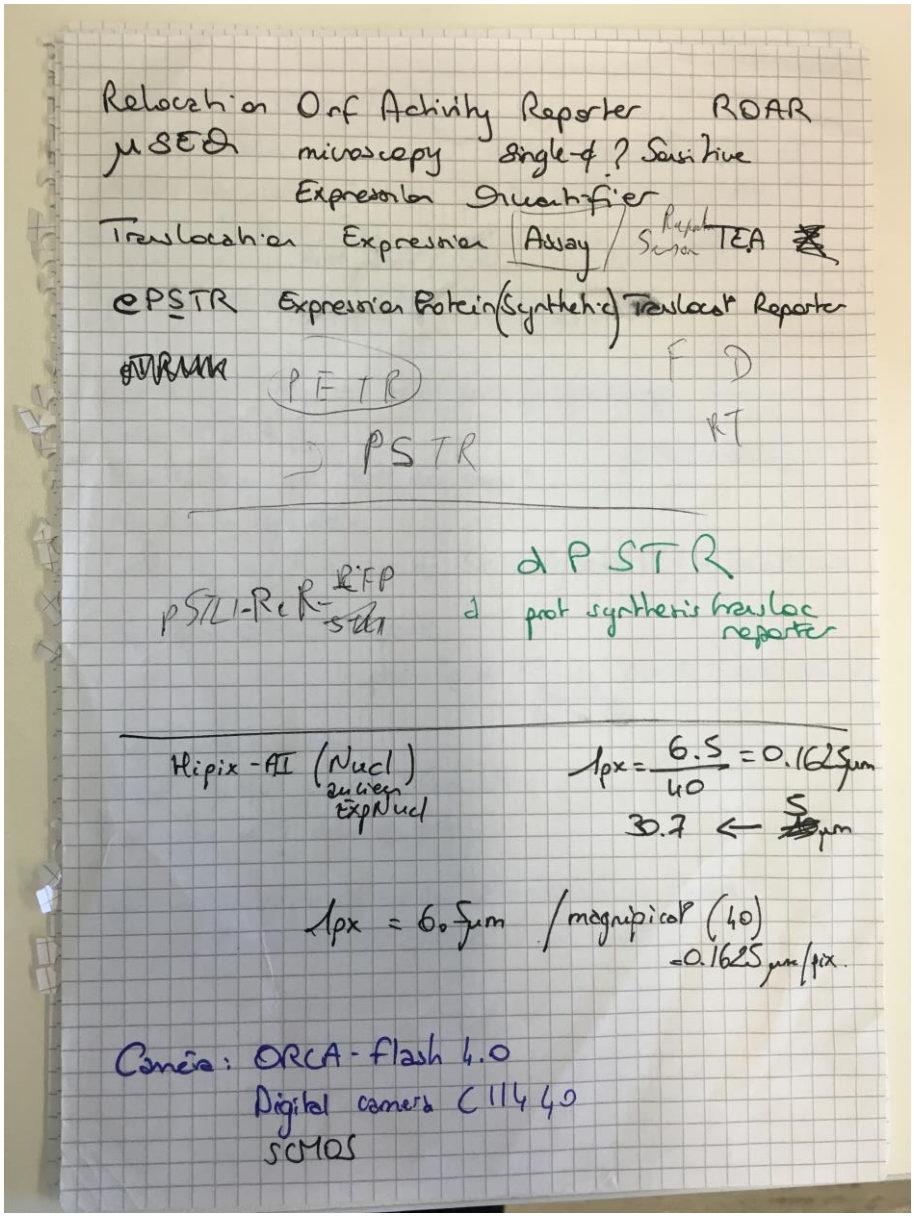
Plasmid	MCS1	MCS2	Backbone	Strain
pDA137	pRPL24A-mCherry-SynZip2	pSTL1-UbiY-2xSv40NLS-SynZip3-tCYC1	pSIVU	yDA85
pDA140	pRPL15A-mCherry-SynZip2	pSTL1-UbiY-2xSv40NLS-SynZip1-tCYC1	pSIVU	yDA93
pDA157	p*RPL24A*-mCherry-SynZip2-tSIF2	-	pSIVU	YDA278
pDA169	pRPL24A-mCherry-SynZip2	pGAL1-UbiY-2xSv40NLS-SynZip1-tCYC1	pSIVU	yDA112
pDA176	pRPL24A-mCherry-SynZip2	pSTL1- 2xSv40NLS-Venus-SynZip1-tCYC1	pSIVU	yDA119
pDA183	pRPL24A-mCherry-SynZip2-tSIF2	pSTL1-UbiY-2xSv40NLS-SynZip1-tCYC1	pSIVU	yDA134, yDA137
pDA193	pRPL24A-mCherry-SynZip2-tSIF2	pGPD1-UbiY-2xSv40NLS-SynZip1-tCYC1	pSIVU	yDA139, yDA155,
pDA199	pRPL24B-MCitrine-SynZip4-tNUP53	pSTL1-UbiY-2xSv40NLS-SynZip3-tCYC1	pSIVL	yDA137
pDA200	pRPL24B-MCitrine-SynZip4-tNUP53	pGPD1-UbiY-2xSv40NLS-SynZip3-tCYC1	pSIVL	yDA139
pDA223	pRPL24A-mCherry-SynZip2-tSIF2	pSTE*12*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	yDA186
pDA228	p*RPL24A*-mCherry-SynZip2-tSIF2	p*KAR4*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	yDA198
pDA229	p*RPL24A*-mCherry-SynZip2-tSIF2	p*FAR1*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	yDA146
pDA239	p*RPL24A*-mCherry-SynZip2-tSIF2	p*FUS1*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	yDA189
pDA251	pRPL24A-mCherry-SynZip2	p*CDC6-UbiY*-2xSv40NLS-SynZip1-t*CYC1*	pSIVL	yDA172
pDA268	p*RPL24B*-MCitrine-SynZip4-t*NUP53*	p*KAR3*-2xSv40NLS-SynZip3-t*CYC1*	pSIVL	ySP641
pDA274	pRPL24B-MCitrine-SynZip4-tNUP53	pKAR4-UbiY-2xSv40NLS-SynZip3-tCYC1	pSIVL	yDA222
pDA282	p*RPL24A*-mCherry-SynZip2-tSIF2	p*AGA1*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	yDA237
pDA283	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1syn* -2xSv40NLS-SynZip1-t*CYC1*	pSIVU	

Plasmid	MCS1	MCS2	Backbone	Strain
pDA284	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1*AGA1* -2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA285	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1*FIG1* -2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA286	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1syn_ΔPREI*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA287	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1syn_ΔPREII*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA288	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1syn_ΔPREIII*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA289	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1syn_ΔPREI*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA290	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1syn_ΔPREII*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA291	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1syn_ΔPREIII*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA300	p*RPL24A*-mCherry-SynZip2-tSIF2	p*BAR1*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	yDA244
pDA301	p*RPL24A*-mCherry-SynZip2-tSIF2	p*SST2*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	yDA245
pDA305	p*RPL24A*-mCherry-SynZip2-tSIF2	p*PRM3*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	yDA249
pDA306	p*RPL24A*-mCherry-SynZip2-tSIF2	p*KAR3*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	yDA247
pDA307	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1syn_4th PRE*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA308	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1syn_PREIIswapped*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA309	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1syn_PREII_36nt_PREIII*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA310	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1syn_ΔPREIΔPREII*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	yDA254
pDA311	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1syn_ΔPREIIΔPREIII*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	yDA255
pDA312	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1syn_PREII_20nt_PREIII*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA318	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1syn_ΔPREIΔPREIII*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA319	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1syn_PREII_10nt_PREIII*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA320	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1syn_PREII_27nt_PREIII*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA321	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1syn_PREII_27nt_PREIII_OoN*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA322	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1syn_ΔPREIΔPREIIΔPREIII*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	

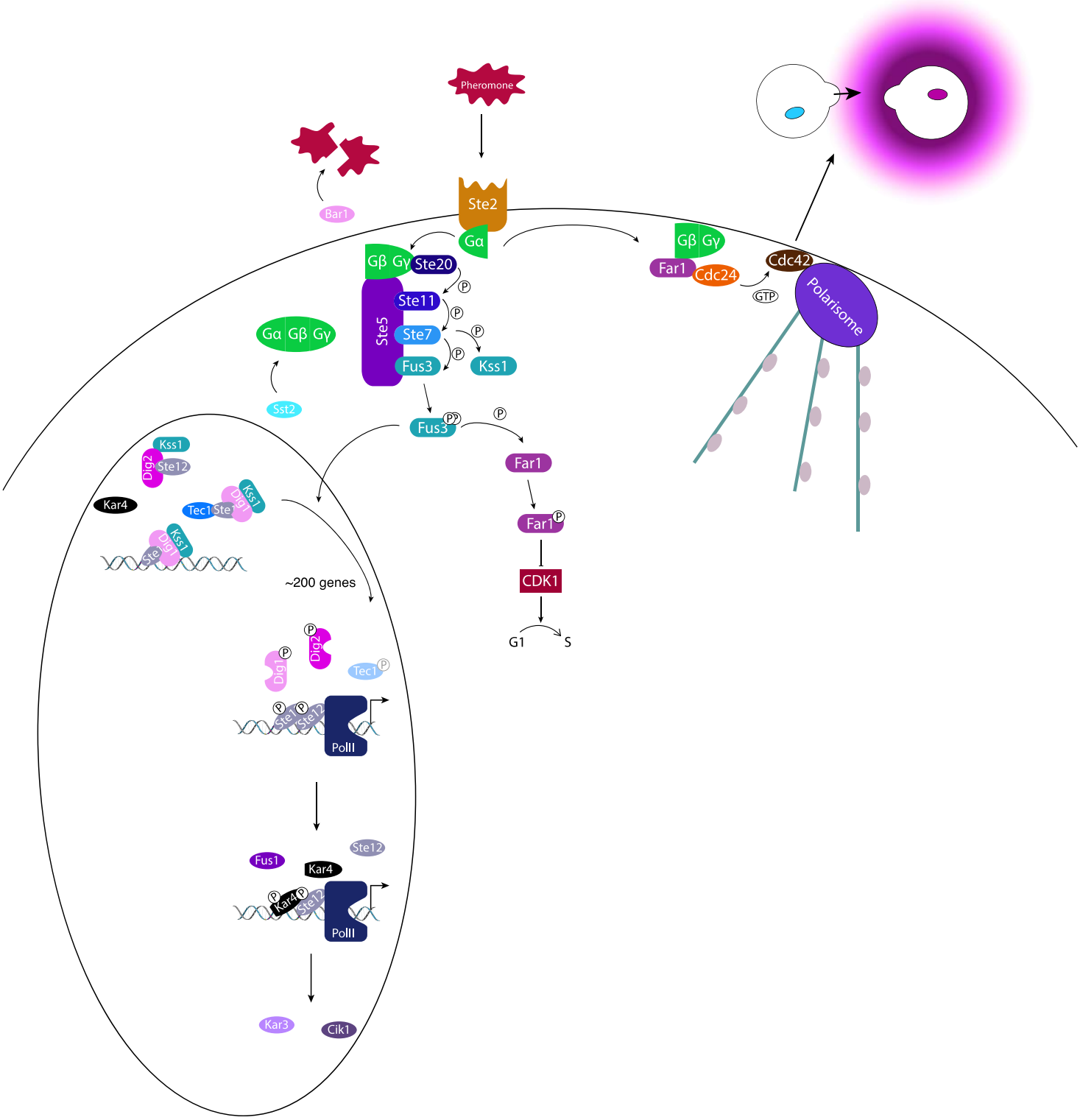
Plasmid	MCS1	MCS2	Backbone	Strain
pDA323	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1syn_ΔPREI_PREIIswapped*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA324	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1syn_ΔPREIΔPREIII*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA329	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1syn_163nt_PREIIΔPREIII*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA330	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1syn_PREII_36nt_PREIII_OoN*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA334	p*RPL24A*-mCherry-SynZip2-tSIF2	p*FUS2*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	yDA296
pDA338	p*RPL24A*-mCherry-SynZip2-tSIF2	p*PRM1*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	yDA306
pDA340	p*RPL24A*-mCherry-SynZip2-tSIF2	p*FIG2*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	yDA302
pDA343	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1*FIG1ΔPREII*AGA1* -2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA344	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1syn_PREII^{nc}_to_consensus*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA345	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1syn_ΔPREII^{nc}* -2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA346	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1syn_ΔPREIΔPREII^{nc}ΔPREIII* -2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA365	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1*FIG1 *AGA1* -2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA386	p*RPL24A*-mCherry-SynZip2-tSIF2	*pGPD1-UbiY*-4xSynZip1-RitC-t*CYC1*	pSIVU	yDA353
pDA387	p*RPL24A*-mCherry-SynZip2-tSIF2	*pFIG1 PREII^{nc}_to_consensus *AGA1* -2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA388	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1ΔPREII*FIG1* -2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pDA389	p*RPL24A*-mCherry-SynZip2-tSIF2	*pAGA1ΔPREIΔPREII *FIG1* -2xSv40NLS-SynZip1-t*CYC1*	pSIVU	
pSP264	pGLT1-pSTL1-24xPP7StemLoop	-	pDZ306 Addgene:35196	ySP374
pSP267	pFIG1 24xPP7SL	-	pHIS	ySP377
pSP268	pMET-PP7-2xGFP	-	pRS304	ySP374
pSP268	pMET PP7-2xGFP	-	pRS304	ySP377, ySP725
pSP34	pSTL1-quadrupleVenus	-	pRS305	ySP9
pSP365	p*RPL24A*-mCherry-SynZip2-tSIF2	p*AGA1*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	ySP642
pSP366	p*RPL24A*-mCherry-SynZip2-tSIF2	p*FIG1*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	ySP643

Plasmid	MCS1	MCS2	Backbone	Strain
pSP367	p*RPL24B*-MCitrine-SynZip4-t*NUP53*	p*FIG1*-2xSv40NLS-SynZip3-t*CYC1*	pSIVL	ySP644 yDA281
pSP368	p*RPL24B*-MCitrine-SynZip4-t*NUP53*	p*AGA1*-2xSv40NLS-SynZip3-t*CYC1*	pSIVL	ySP641 yDA280
pSP384	p*RPL24A*-mCherry-SynZip2-tSIF2	p*PRM5*-2xSv40NLS-SynZip1-t*CYC1*	pSIVU	ySP664
pSP390	pRPL24A-mCherry-SynZip2-tSIF2	pAGA1-UbiY-2xSv40NLS-SynZip1-tCYC1	pSIVU	yDA222
pSP397	pGPD-tdiRFP		pRS306	ySP711
pSP409	pAGA1 24xPP7SL		pHIS	ySP725
pBB3	pAde6^{PmeI}-p^{pom1}-mCherry-linker-SynZip1-term^{cyc1}			yBB10
pBB7	pUra4^{AfeI}-p^{map3}-2xNLS-linker-SynZyp2-term^{nmt}			yBB10

Annex II: Team work leads to fun names



Annex III: Overview of the mating-induced events



References

. Yeast GFP database: https://yeastgfp.yeastgenome.org/

. Yestract: http://www.yeastract.com/

Aguilar, P.S., Engel, A., and Walter, P. (2007). The plasma membrane proteins Prm1 and Fig1 ascertain fidelity of membrane fusion during yeast mating. Molecular biology of the cell *18*, 547-556.

Albertyn, J., Hohmann, S., Thevelein, J.M., and Prior, B.A. (1994). GPD1, which encodes glycerol-3-phosphate dehydrogenase, is essential for growth under osmotic stress in *Saccharomyces cerevisiae*, and its expression is regulated by the high-osmolarity glycerol response pathway. Molecular and cellular biology *14*, 4135-4144.

Alepuz, P.M., de Nadal, E., Zapater, M., Ammerer, G., and Posas, F. (2003). Osmostress-induced transcription by Hot1 depends on a Hog1-mediated recruitment of the RNA Pol II. The EMBO journal *22*, 2433-2442.

Alepuz, P.M., Jovanovic, A., Reiser, V., and Ammerer, G. (2001). Stress-induced map kinase Hog1 is part of transcription activation complexes. Molecular cell *7*, 767-777.

Alford, S.C., Abdelfattah, A.S., Ding, Y., and Campbell, R.E. (2012a). A fluorogenic red fluorescent protein heterodimer. Chem Biol *19*, 353-360.

Alford, S.C., Ding, Y., Simmen, T., and Campbell, R.E. (2012b). Dimerization-dependent green and yellow fluorescent proteins. ACS synthetic biology *1*, 569-575.

Altschuler, S.J., and Wu, L.F. (2010). Cellular heterogeneity: do differences make a difference? Cell *141*, 559-563.

Ashe, H.L., and Briscoe, J. (2006). The interpretation of morphogen gradients. Development *133*, 385-394.

Atay, O., and Skotheim, J.M. (2017). Spatial and temporal signal processing and decision making by MAPK pathways. J Cell Biol.

Atienza, J.M., Suh, M., Xenarios, I., Landgraf, R., and Colicelli, J. (2000). Human ERK1 induces filamentous growth and cell wall remodeling pathways in *Saccharomyces cerevisiae*. The Journal of biological chemistry *275*, 20638-20646.

Aymoz, D., Wosika, V., Durandau, E., and Pelet, S. (2016). Real-time quantification of protein expression at the single-cell level *via* dynamic protein synthesis translocation reporters. Nature communications *7*, 11304.

Ballensiefen, W., and Schmitt, H.D. (1997). Periplasmic Bar1 protease of *Saccharomyces cerevisiae* is active before reaching its extracellular destination. Eur J Biochem *247*, 142-147.

Bardwell, L. (2005). A walk-through of the yeast mating pheromone response pathway. Peptides *26*, 339-350.

Bardwell, L., Cook, J.G., Voora, D., Baggott, D.M., Martinez, A.R., and Thorner, J. (1998a). Repression of yeast Ste12 transcription factor by direct binding of unphosphorylated Kss1 MAPK and its regulation by the Ste7 MEK. Genes & development *12*, 2887-2898.

Bardwell, L., Cook, J.G., Zhu-Shimoni, J.X., Voora, D., and Thorner, J. (1998b). Differential regulation of transcription: repression by unactivated mitogen-activated protein kinase Kss1 requires the Dig1 and Dig2 proteins. Proceedings of the National Academy of Sciences of the United States of America *95*, 15400-15405.

Barkai, N., Rose, M.D., and Wingreen, N.S. (1998). Protease helps yeast find mating partners. Nature *396*, 422-423.

Bendezu, F.O., Vincenzetti, V., Vavylonis, D., Wyss, R., Vogel, H., and Martin, S.G. (2015). Spontaneous Cdc42 polarization independent of GDI-mediated extraction and actin-based trafficking. PLoS biology *13*, e1002097.

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H., and Long, R.M. (1998). Localization of ASH1 mRNA particles in living yeast. Molecular cell *2*, 437-445.

Bickel, K.S., and Morris, D.R. (2006). Role of the transcription activator Ste12p as a repressor of PRY3 expression. Molecular and cellular biology *26*, 7901-7912.

Blackwell, E., Kim, H.J., and Stone, D.E. (2007). The pheromone-induced nuclear accumulation of the Fus3 MAPK in yeast depends on its phosphorylation state and on Dig1 and Dig2. BMC Cell Biol *8*, 44.

Blomberg, A., and Adler, L. (1992). Physiology of osmotolerance in fungi. Adv Microb Physiol *33*, 145-212.

Bourillot, P.Y., Garrett, N., and Gurdon, J.B. (2002). A changing morphogen gradient is interpreted by continuous transduction flow. Development *129*, 2167-2180.

Breitkreutz, A., Boucher, L., Breitkreutz, B.J., Sultan, M., Jurisica, I., and Tyers, M. (2003). Phenotypic and transcriptional plasticity directed by a yeast mitogen-activated protein kinase network. Genetics *165*, 997-1015.

Brewster, J.L., de Valoir, T., Dwyer, N.D., Winter, E., and Gustin, M.C. (1993). An osmosensing signal transduction pathway in yeast. Science *259*, 1760-1763.

Brizzio, V., Gammie, A.E., and Rose, M.D. (1998). Rvs161p interacts with Fus2p to promote cell fusion in *Saccharomyces cerevisiae*. J Cell Biol *141*, 567-584.

Brogaard, K., Xi, L., Wang, J.P., and Widom, J. (2012). A map of nucleosome positions in yeast at base-pair resolution. Nature *486*, 496-501.

Brown, J.L., North, S., and Bussey, H. (1993). SKN7, a yeast multicopy suppressor of a mutation affecting cell wall beta-glucan assembly, encodes a product with domains homologous to prokaryotic two-component regulators and to heat shock transcription factors. J Bacteriol *175*, 6908-6915.

Bruckner, S., Kohler, T., Braus, G.H., Heise, B., Bolte, M., and Mosch, H.U. (2004). Differential regulation of Tec1 by Fus3 and Kss1 confers signaling specificity in yeast development. Current genetics *46*, 331-342.

Buehrer, B.M., and Errede, B. (1997). Coordination of the mating and cell integrity mitogen-activated protein kinase pathways in *Saccharomyces cerevisiae*. Molecular and cellular biology *17*, 6517-6525.

Butty, A.C., Pryciak, P.M., Huang, L.S., Herskowitz, I., and Peter, M. (1998). The role of Far1p in linking the heterotrimeric G protein to polarity establishment proteins during yeast mating. Science *282*, 1511-1516.

Capaldi, A.P., Kaplan, T., Liu, Y., Habib, N., Regev, A., Friedman, N., and O'Shea, E.K. (2008). Structure and function of a transcriptional network activated by the MAPK Hog1. Nat Genet *40*, 1300-1306.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., and Prasher, D.C. (1994). Green fluorescent protein as a marker for gene expression. Science *263*, 802-805.

Chao, M.V., Rajagopal, R., and Lee, F.S. (2006). Neurotrophin signalling in health and disease. Clin Sci (Lond) *110*, 167-173.

Chekanova, J.A., Abruzzi, K.C., Rosbash, M., and Belostotsky, D.A. (2008). Sus1, Sac3, and Thp1 mediate post-transcriptional tethering of active genes to the nuclear rim as well as to non-nascent mRNP. RNA *14*, 66-77.

Chen, R.E., and Thorner, J. (2007). Function and regulation in MAPK signaling pathways: lessons learned from the yeast *Saccharomyces cerevisiae*. Biochimica et biophysica acta *1773*, 1311-1340.

Chen, Z., Gibson, T.B., Robinson, F., Silvestro, L., Pearson, G., Xu, B., Wright, A., Vanderbilt, C., and Cobb, M.H. (2001). MAP kinases. Chem Rev *101*, 2449-2476.

Chou, S., Huang, L., and Liu, H. (2004). Fus3-regulated Tec1 degradation through SCFCdc4 determines MAPK signaling specificity during mating in yeast. Cell *119*, 981-990.

Chou, S., Lane, S., and Liu, H. (2006). Regulation of mating and filamentation genes by two distinct Ste12 complexes in *Saccharomyces cerevisiae*. Molecular and cellular biology *26*, 4794-4805.

Christian, J.L. (2012). Morphogen gradients in development: from form to function. Wiley Interdiscip Rev Dev Biol *1*, 3-15.

Cleary, M.A., Kilian, K., Wang, Y., Bradshaw, J., Cavet, G., Ge, W., Kulkarni, A., Paddison, P.J., Chang, K., Sheth, N.*, et al.* (2004). Production of complex nucleic acid libraries using highly parallel in situ oligonucleotide synthesis. Nature methods *1*, 241-248.

Clotet, J., and Posas, F. (2007). Control of cell cycle in response to osmostress: lessons from yeast. Methods Enzymol *428*, 63-76.

Colby, D.W., Chu, Y., Cassady, J.P., Duennwald, M., Zazulak, H., Webster, J.M., Messer, A., Lindquist, S., Ingram, V.M., and Wittrup, K.D. (2004). Potent inhibition of huntingtin aggregation and cytotoxicity by a disulfide bond-free single-domain intracellular antibody. Proceedings of the National Academy of Sciences of the United States of America *101*, 17616-17621.

Coleman, R.A., and Pugh, B.F. (1995). Evidence for functional binding and stable sliding of the TATA binding protein on nonspecific DNA. The Journal of biological chemistry *270*, 13850-13859.

Colman-Lerner, A., Gordon, A., Serra, E., Chin, T., Resnekov, O., Endy, D., Pesce, C.G., and Brent, R.

(2005). Regulated cell-to-cell variation in a cell-fate decision system. Nature *437*, 699-706.

Cook, J.G., Bardwell, L., Kron, S.J., and Thorner, J. (1996). Two novel targets of the MAP kinase Kss1 are negative regulators of invasive growth in the yeast *Saccharomyces cerevisiae*. Genes & development *10*, 2831-2848.

Coulon, A., Chow, C.C., Singer, R.H., and Larson, D.R. (2013). Eukaryotic transcriptional dynamics: from single molecules to cell populations. Nature reviews Genetics *14*, 572-584.

Cullen, P.J., and Sprague, G.F., Jr. (2012). The regulation of filamentous growth in yeast. Genetics *190*, 23-49.

Curran, K.A., Karim, A.S., Gupta, A., and Alper, H.S. (2013). Use of expression-enhancing terminators in *Saccharomyces cerevisiae* to increase mRNA half-life and improve gene expression control for metabolic engineering applications. Metabolic engineering *19*, 88-97.

Dadiani, M., van Dijk, D., Segal, B., Field, Y., Ben-Artzi, G., Raveh-Sadka, T., Levo, M., Kaplow, I., Weinberger, A., and Segal, E. (2013). Two DNA-encoded strategies for increasing expression with opposing effects on promoter dynamics and transcriptional noise. Genome Res *23*, 966-976.

de Nadal, E., Ammerer, G., and Posas, F. (2011). Controlling gene expression in response to stress. Nature reviews Genetics *12*, 833-845.

de Nadal, E., and Posas, F. (2010). Multilayered control of gene expression by stress-activated protein kinases. The EMBO journal *29*, 4-13.

de Nadal, E., Real, F.X., and Posas, F. (2007). Mucins, osmosensors in eukaryotic cells? Trends in cell biology *17*, 571-574.

De Nadal, E., Zapater, M., Alepuz, P.M., Sumoy, L., Mas, G., and Posas, F. (2004). The MAPK Hog1 recruits Rpd3 histone deacetylase to activate osmoresponsive genes. Nature *427*, 370-374.

De Robertis, E.M., Larrain, J., Oelgeschlager, M., and Wessely, O. (2000). The establishment of Spemann's organizer and patterning of the vertebrate embryo. Nature reviews Genetics *1*, 171-181.

Dhillon, A.S., Hagan, S., Rath, O., and Kolch, W. (2007). MAP kinase signalling pathways in cancer. Oncogene *26*, 3279-3290.

Ding, Y., Li, J., Enterina, J.R., Shen, Y., Zhang, I., Tewson, P.H., Mo, G.C., Zhang, J., Quinn, A.M., Hughes, T.E.*, et al.* (2015). Ratiometric biosensors based on dimerization-dependent fluorescent protein exchange. Nature methods *12*, 195-198.

Dodou, E., and Treisman, R. (1997). The *Saccharomyces cerevisiae* MADS-box transcription factor Rlm1 is a target for the Mpk1 mitogen-activated protein kinase pathway. Molecular and cellular biology *17*, 1848-1859.

Doi, K., Gartner, A., Ammerer, G., Errede, B., Shinkawa, H., Sugimoto, K., and Matsumoto, K. (1994). MSG5, a novel protein phosphatase promotes adaptation to pheromone response in S. cerevisiae. The EMBO journal *13*, 61-70.

Dolan, J.W., Kirkman, C., and Fields, S. (1989). The yeast STE12 protein binds to the DNA sequence mediating pheromone induction. Proceedings of the National Academy of Sciences of the United States of America *86*, 5703-5707.

Durandau, E., Aymoz, D., and Pelet, S. (2015). Dynamic single cell measurements of kinase activity by synthetic kinase activity relocation sensors. BMC Biol *13*, 55.

Dyson, S., and Gurdon, J.B. (1998). The interpretation of position in a morphogen gradient as revealed by occupancy of activin receptors. Cell *93*, 557-568.

Edelstein, A., Amodaj, N., Hoover, K., Vale, R., and Stuurman, N. (2010). Computer control of microscopes using microManager. Curr Protoc Mol Biol *Chapter 14*, Unit14 20.

Elion, E.A. (2001). The Ste5p scaffold. J Cell Sci *114*, 3967-3978.

Elion, E.A., Satterberg, B., and Kranz, J.E. (1993). FUS3 phosphorylates multiple components of the mating signal transduction cascade: evidence for STE12 and FAR1. Molecular biology of the cell *4*, 495-510.

Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene expression in a single cell. Science *297*, 1183-1186.

Engel, S.R., Dietrich, F.S., Fisk, D.G., Binkley, G., Balakrishnan, R., Costanzo, M.C., Dwight, S.S., Hitz, B.C., Karra, K., Nash, R.S.*, et al.* (2014). The reference genome sequence of *Saccharomyces cerevisiae*: then and now. G3 *4*, 389-398.

Erdman, S., Lin, L., Malczynski, M., and Snyder, M. (1998). Pheromone-regulated genes required for yeast mating differentiation. J Cell Biol *140*, 461-483.

Errede, B. (1993). MCM1 binds to a transcriptional control element in Ty1. Molecular and cellular biology *13*, 57-62.

Errede, B., and Ammerer, G. (1989). STE12, a protein involved in cell-type-specific transcription and signal transduction in yeast, is part of protein-DNA complexes. Genes & development *3*, 1349-1361.

Esch, R.K., Wang, Y., and Errede, B. (2006). Pheromone-induced degradation of Ste12 contributes to signal attenuation and the specificity of developmental fate. Eukaryotic cell *5*, 2147-2160.

Falconnet, D., Niemisto, A., Taylor, R.J., Ricicova, M., Galitski, T., Shmulevich, I., and Hansen, C.L. (2011). High-throughput tracking of single yeast cells in a microfluidic imaging matrix. Lab Chip *11*, 466-473.

Femino, A.M. (1998). Visualization of Single RNA Transcripts in Situ. Science *280*, 585-590.

Ferreira, C., van Voorst, F., Martins, A., Neves, L., Oliveira, R., Kielland-Brandt, M.C., Lucas, C., and Brandt, A. (2005). A member of the sugar transporter family, Stl1p is the glycerol/H+ symporter in *Saccharomyces cerevisiae*. Molecular biology of the cell *16*, 2068-2076.

Field, Y., Kaplan, N., Fondufe-Mittendorf, Y., Moore, I.K., Sharon, E., Lubling, Y., Widom, J., and Segal, E. (2008). Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. PLoS Comput Biol *4*, e1000216.

Flagfeldt, D.B., Siewers, V., Huang, L., and Nielsen, J. (2009). Characterization of chromosomal integration sites for heterologous gene expression in *Saccharomyces cerevisiae*. Yeast *26*, 545-551.

Fleming, J., Spinoulas, A., Zheng, M., Cunningham, S.C., Ginn, S.L., McQuilty, R.C., Rowe, P.B., and Alexander, I.E. (2005). Partial correction of sensitivity to oxidant stress in Friedreich ataxia patient fibroblasts by frataxin-encoding adeno-associated virus and lentivirus vectors. Hum Gene Ther *16*, 947-956.

Frydman, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F.U. (1999). Co-translational domain folding as the structural basis for the rapid de novo folding of firefly luciferase. Nat Struct Biol *6*, 697-705.

Gammie, A.E., Brizzio, V., and Rose, M.D. (1998). Distinct morphological phenotypes of cell fusion mutants. Molecular biology of the cell *9*, 1395-1410.

Gammie, A.E., Stewart, B.G., Scott, C.F., and Rose, M.D. (1999). The two forms of karyogamy transcription factor Kar4p are regulated by differential initiation of transcription, translation, and protein turnover. Molecular and cellular biology *19*, 817-825.

Garcia, R., Bermejo, C., Grau, C., Perez, R., Rodriguez-Pena, J.M., Francois, J., Nombela, C., and Arroyo, J. (2004). The global transcriptional response to transient cell wall damage in *Saccharomyces cerevisiae* and its regulation by the cell integrity signaling pathway. The Journal of biological chemistry *279*, 15183-15195.

Gartner, A., Jovanovic, A., Jeoung, D.I., Bourlat, S., Cross, F.R., and Ammerer, G. (1998). Pheromone-dependent G1 cell cycle arrest requires Far1 phosphorylation, but may not involve inhibition of Cdc28-Cln2 kinase, in vivo. Molecular and cellular biology *18*, 3681-3691.

Gartner, A., Nasmyth, K., and Ammerer, G. (1992). Signal transduction in *Saccharomyces cerevisiae* requires tyrosine and threonine phosphorylation of FUS3 and KSS1. Genes & development *6*, 1280-1292.

Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., and Brown, P.O. (2000). Genomic expression programs in the response of yeast cells to environmental changes. Molecular biology of the cell *11*, 4241-4257.

Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. Nature *425*, 737-741.

Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M*., et al.* (1996). Life with 6000 genes. Science *274*, 546, 563-547.

Gonda, D.K., Bachmair, A., Wunning, I., Tobias, J.W., Lane, W.S., and Varshavsky, A. (1989). Universality and structure of the N-end rule. The Journal of biological chemistry *264*, 16700-16712.

Good, M., Tang, G., Singleton, J., Remenyi, A., and Lim, W.A. (2009). The Ste5 scaffold directs mating signaling by catalytically unlocking the Fus3 MAP kinase for activation. Cell *136*, 1085-1097.

Green, J.B., New, H.V., and Smith, J.C. (1992). Responses of embryonic Xenopus cells to activin and FGF are separated by multiple dose thresholds and correspond to distinct axes of the mesoderm. Cell *71*, 731-739.

Greenbaum, D., Colangelo, C., Williams, K., and Gerstein, M. (2003). Comparing protein abundance and mRNA expression levels on a genomic scale. Genome biology *4*, 117.

Grishin, A.V., Rothenberg, M., Downs, M.A., and Blumer, K.J. (1998). Mot3, a Zn finger transcription factor that modulates gene expression and attenuates mating pheromone signaling in *Saccharomyces cerevisiae*. Genetics *149*, 879-892.

Gurdon, J.B., and Bourillot, P.Y. (2001). Morphogen gradient interpretation. Nature *413*, 797-803.

Gurdon, J.B., Harger, P., Mitchell, A., and Lemaire, P. (1994). Activin signalling and response to a morphogen gradient. Nature *371*, 487-492.

Gurdon, J.B., Standley, H., Dyson, S., Butler, K., Langon, T., Ryan, K., Stennard, F., Shimizu, K., and Zorn, A. (1999). Single cells can sense their position in a morphogen gradient. Development *126*, 5309-5317.

Gustin, M.C., Albertyn, J., Alexander, M., and Davenport, K. (1998). MAP kinase pathways in the yeast *Saccharomyces cerevisiae*. Microbiology and molecular biology reviews : MMBR *62*, 1264-1300.

Gygi, S.P., Rochon, Y., Franza, B.R., and Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. Molecular and cellular biology *19*, 1720-1730.

Hagen, D.C., McCaffrey, G., and Sprague, G.F., Jr. (1991). Pheromone response elements are necessary and sufficient for basal and pheromone-induced transcription of the FUS1 gene of *Saccharomyces cerevisiae*. Molecular and cellular biology *11*, 2952-2961.

Hahn, S., and Young, E.T. (2011). Transcriptional regulation in *Saccharomyces cerevisiae*: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. Genetics *189*, 705-736.

Han, J., Lee, J.D., Bibbs, L., and Ulevitch, R.J. (1994). A MAP kinase targeted by endotoxin and hyperosmolarity in mammalian cells. Science *265*, 808-811.

Hansen, A.S., and O'Shea, E.K. (2015). cis Determinants of Promoter Threshold and Activation Timescale. Cell Rep *12*, 1226-1233.

Hartwell, L.H. (1980). Mutants of *Saccharomyces cerevisiae* unresponsive to cell division control by polypeptide mating hormone. J Cell Biol *85*, 811-822.

Harvey, S.A., and Smith, J.C. (2009). Visualisation and quantification of morphogen gradient formation in the zebrafish. PLoS biology *7*, e1000101.

Henchoz, S., Chi, Y., Catarin, B., Herskowitz, I., Deshaies, R.J., and Peter, M. (1997). Phosphorylation- and ubiquitin-dependent degradation of the cyclin-dependent kinase inhibitor Far1p in budding yeast. Genes & development *11*, 3046-3060.

Hersen, P., McClean, M.N., Mahadevan, L., and Ramanathan, S. (2008). Signal processing by the HOG MAP kinase pathway. Proceedings of the National Academy of Sciences of the United States of America *105*, 7165-7170.

Hertel, K.J., Lynch, K.W., and Maniatis, T. (1997). Common themes in the function of transcription and splicing enhancers. Curr Opin Cell Biol *9*, 350-357.

Hinnebusch, A.G., and Natarajan, K. (2002). Gcn4p, a master regulator of gene expression, is controlled at multiple levels by diverse signals of starvation and stress. Eukaryotic cell *1*, 22-32.

Hocine, S., Raymond, P., Zenklusen, D., Chao, J.A., and Singer, R.H. (2013). Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. Nature methods *10*, 119-121.

Hohmann, S. (2002). Osmotic Stress Signaling and Osmoadaptation in Yeasts. Microbiology and Molecular Biology Reviews *66*, 300-372.

Hohmann, S., Krantz, M., and Nordlander, B. (2007). Yeast osmoregulation. Methods Enzymol *428*, 29-45.

Houser, J.R., Ford, E., Chatterjea, S.M., Maleri, S., Elston, T.C., and Errede, B. (2012). An improved short-lived fluorescent protein transcriptional reporter for *Saccharomyces cerevisiae*. Yeast *29*, 519-530.

Hung, W., Olson, K.A., Breitkreutz, A., and Sadowski, I. (1997). Characterization of the basal and pheromone-stimulated phosphorylation states of Ste12p. Eur J Biochem *245*, 241-251.

Igual, J.C., Johnson, A.L., and Johnston, L.H. (1996). Coordinated regulation of gene expression by the cell cycle transcription factor Swi4 and the protein kinase C MAP kinase pathway for yeast cell integrity. The EMBO journal *15*, 5001-5013.

Jacoby, T., Flanagan, H., Faykin, A., Seto, A.G., Mattison, C., and Ota, I. (1997). Two protein-tyrosine phosphatases inactivate the osmotic stress response pathway in yeast by targeting the mitogen-activated protein kinase, Hog1. The Journal of biological chemistry *272*, 17749-17755.

Jiang, C., and Pugh, B.F. (2009). A compiled and systematic reference map of nucleosome positions across the *Saccharomyces cerevisiae* genome. Genome biology *10*, R109.

Jin, H., Carlile, C., Nolan, S., and Grote, E. (2004). Prm1 prevents contact-dependent lysis of yeast mating pairs. Eukaryotic cell *3*, 1664-1673.

Jin, M., Errede, B., Behar, M., Mather, W., Nayak, S., Hasty, J., Dohlman, H.G., and Elston, T.C. (2011). Yeast dynamically modify their environment to achieve better mating efficiency. Science signaling *4*, ra54.

Jorgensen, P., Nishikawa, J.L., Breitkreutz, B.J., and Tyers, M. (2002). Systematic identification of pathways that couple cell growth and division in yeast. Science *297*, 395-400.

Kadonaga, J.T. (2004). Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. Cell *116*, 247-257.

Kaffman, A., and O'Shea, E.K. (1999). Regulation of nuclear localization: a key to a door. Annu Rev Cell Dev Biol *15*, 291-339.

Kalisky, T., Blainey, P., and Quake, S.R. (2011). Genomic analysis at the single-cell level. Annu Rev Genet *45*, 431-445.

Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J.*, et al.* (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. Nature *458*, 362-366.

Kelly, M., Burke, J., Smith, M., Klar, A., and Beach, D. (1988). Four mating-type genes control sexual differentiation in the fission yeast. The EMBO journal *7*, 1537-1547.

Kheradpour, P., Ernst, J., Melnikov, A., Rogov, P., Wang, L., Zhang, X., Alston, J., Mikkelsen, T.S., and Kellis, M. (2013). Systematic dissection of regulatory motifs in 2000 predicted human enhancers using a massively parallel reporter assay. Genome Res *23*, 800-811.

Khmelinskii, A., Keller, P.J., Bartosik, A., Meurer, M., Barry, J.D., Mardin, B.R., Kaufmann, A., Trautmann, S., Wachsmuth, M., Pereira, G.*, et al.* (2012). Tandem fluorescent protein timers for in vivo analysis of protein dynamics. Nature biotechnology *30*, 708-714.

Kirkman-Correia, C., Stroke, I.L., and Fields, S. (1993). Functional domains of the yeast STE12 protein, a pheromone-responsive transcriptional activator. Molecular and cellular biology *13*, 3765-3772.

Kitamura, K., and Fujiwara, H. (2013). The type-2 N-end rule peptide recognition activity of Ubr11 ubiquitin ligase is required for the expression of peptide transporters. FEBS Lett *587*, 214-219.

Klopf, E., Paskova, L., Sole, C., Mas, G., Petryshyn, A., Posas, F., Wintersberger, U., Ammerer, G., and Schuller, C. (2009). Cooperation between the INO80 complex and histone chaperones determines adaptation of stress gene transcription in the yeast *Saccharomyces cerevisiae*. Molecular and cellular biology *29*, 4994-5007.

Knight, B., Kubik, S., Ghosh, B., Bruzzone, M.J., Geertz, M., Martin, V., Denervaud, N., Jacquet, P., Ozkan, B., Rougemont, J.*, et al.* (2014). Two distinct promoter architectures centered on dynamic nucleosomes control ribosomal protein gene transcription. Genes & development *28*, 1695-1709.

Komano, H., Seeger, M., Gandy, S., Wang, G.T., Krafft, G.A., and Fuller, R.S. (1998). Involvement of cell surface glycosyl-phosphatidylinositol-linked aspartyl proteases in alpha-secretase-type cleavage and ectodomain solubilization of human Alzheimer beta-amyloid precursor protein in yeast. The Journal of biological chemistry *273*, 31648-31651.

Krems, B., Charizanis, C., and Entian, K.D. (1996). The response regulator-like protein Pos9/Skn7 of *Saccharomyces cerevisiae* is involved in oxidative stress resistance. Current genetics *29*, 327-334.

Kron, S.J., and Gow, N.A. (1995). Budding yeast morphogenesis: signalling, cytoskeleton and cell cycle. Curr Opin Cell Biol *7*, 845-855.

Kronstad, J.W., Holly, J.A., and MacKay, V.L. (1987). A yeast operator overlaps an upstream activation site. Cell *50*, 369-377.

Kurihara, L.J., Beh, C.T., Latterich, M., Schekman, R., and Rose, M.D. (1994). Nuclear congression and membrane fusion: two distinct events in the yeast karyogamy pathway. J Cell Biol *126*, 911-923.

Kurihara, L.J., Stewart, B.G., Gammie, A.E., and Rose, M.D. (1996). Kar4p, a karyogamy-specific component of the yeast pheromone response pathway. Molecular and cellular biology *16*, 3990-4002.

Kyriakis, J.M., and Avruch, J. (2001). Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation. Physiol Rev *81*, 807-869.

Lahav, R., Gammie, A., Tavazoie, S., and Rose, M.D. (2007). Role of transcription factor Kar4 in regulating downstream events in the *Saccharomyces cerevisiae* pheromone response pathway. Molecular and cellular biology *27*, 818-829.

Lam, F.H., Steger, D.J., and O'Shea, E.K. (2008). Chromatin decouples promoter threshold from dynamic range. Nature *453*, 246-250.

Larson, D.R. (2011). What do expression dynamics tell us about the mechanism of transcription? Current opinion in genetics & development *21*, 591-599.

Larson, D.R., Zenklusen, D., Wu, B., Chao, J.A., and Singer, R.H. (2011). Real-time observation of transcription initiation and elongation on an endogenous yeast gene. Science *332*, 475-478.

Lawrence, M.C., Jivan, A., Shao, C., Duan, L., Goad, D., Zaganjor, E., Osborne, J., McGlynn, K., Stippec, S., Earnest, S., *et al.* (2008). The roles of MAPKs in disease. Cell Res *18*, 436-442.

Lee, Y.J., Jeschke, G.R., Roelants, F.M., Thorner, J., and Turk, B.E. (2012). Reciprocal phosphorylation of yeast glycerol-3-phosphate dehydrogenases in adaptation to distinct types of stress. Molecular and cellular biology *32*, 4705-4717.

Lenstra, T.L., Coulon, A., Chow, C.C., and Larson, D.R. (2015). Single-Molecule Imaging Reveals a Switch between Spurious and Functional ncRNA Transcription. Molecular cell *60*, 597-610.

Lenstra, T.L., Rodriguez, J., Chen, H., and Larson, D.R. (2016). Transcription Dynamics in Living Cells. Annu Rev Biophys *45*, 25-47.

Levin, D.E. (2005). Cell wall integrity signaling in *Saccharomyces cerevisiae*. Microbiology and molecular biology reviews : MMBR *69*, 262-291.

Levin, D.E. (2011). Regulation of cell wall biogenesis in *Saccharomyces cerevisiae*: the cell wall integrity signaling pathway. Genetics *189*, 1145-1175.

Li, S., Ault, A., Malone, C.L., Raitt, D., Dean, S., Johnston, L.H., Deschenes, R.J., and Fassler, J.S. (1998). The yeast histidine protein kinase, Sln1p, mediates phosphotransfer to two response regulators, Ssk1p and Skn7p. The EMBO journal *17*, 6952-6962.

Lindquist, S., Krobitsch, S., Li, L., and Sondheimer, N. (2001). Investigating protein conformation-based inheritance and disease in yeast. Philos Trans R Soc Lond B Biol Sci *356*, 169-176.

Lipke, P.N., Wojciechowicz, D., and Kurjan, J. (1989). AG alpha 1 is the structural gene for the *Saccharomyces cerevisiae* alpha-agglutinin, a cell surface glycoprotein involved in cell-cell interactions during mating. Molecular and cellular biology *9*, 3155-3165.

Loewer, A., Batchelor, E., Gaglia, G., and Lahav, G. (2010). Basal dynamics of p53 reveal transcriptionally attenuated pulses in cycling cells. Cell *142*, 89-100.

Logg, K., Bodvard, K., Blomberg, A., and Kall, M. (2009). Investigations on light-induced stress in fluorescence microscopy using nuclear localization of the transcription factor Msn2p as a reporter. FEMS yeast research *9*, 875-884.

Ma, D., Cook, J.G., and Thorner, J. (1995). Phosphorylation and localization of Kss1, a MAP kinase of the *Saccharomyces cerevisiae* pheromone response pathway. Molecular biology of the cell *6*, 889-909.

Mackiewicz, P., Kowalczuk, M., Mackiewicz, D., Nowicka, A., Dudkiewicz, M., Laszkiewicz, A., Dudek, M.R., and Cebrat, S. (2002). How many protein-coding genes are there in the *Saccharomyces cerevisiae* genome? Yeast *19*, 619-629.

Madhani, H.D. (1997). Combinatorial Control Required for the Specificity of Yeast MAPK Signaling. Science *275*, 1314-1317.

Madhani, H.D., Galitski, T., Lander, E.S., and Fink, G.R. (1999). Effectors of a developmental mitogen-activated protein kinase cascade revealed by expression signatures of signaling mutants. Proceedings of the National Academy of Sciences of the United States of America *96*, 12530-12535.

Maeda, T., Takekawa, M., and Saito, H. (1995). Activation of yeast PBS2 MAPKK by MAPKKKs or by binding of an SH3-containing osmosensor. Science *269*, 554-558.

Maeda, T., Wurgler-Murphy, S.M., and Saito, H. (1994). A two-component system that regulates an osmosensing MAP kinase cascade in yeast. Nature *369*, 242-245.

Maleri, S., Ge, Q., Hackett, E.A., Wang, Y., Dohlman, H.G., and Errede, B. (2004). Persistent activation by constitutive Ste7 promotes Kss1-mediated invasive growth but fails to support Fus3-dependent mating in yeast. Molecular and cellular biology *24*, 9221-9238.

Marshall, C.J. (1994). MAP kinase kinase kinase, MAP kinase kinase and MAP kinase. Current opinion in genetics & development *4*, 82-89.

Mas, G., de Nadal, E., Dechant, R., Rodriguez de la Concepcion, M.L., Logie, C., Jimeno-Gonzalez, S., Chavez, S., Ammerer, G., and Posas, F. (2009). Recruitment of a chromatin remodelling complex by the Hog1 MAP kinase to stress genes. The EMBO journal *28*, 326-336.

Mattison, C.P., Spencer, S.S., Kresge, K.A., Lee, J., and Ota, I.M. (1999). Differential regulation of the cell wall integrity mitogen-activated protein kinase pathway in budding yeast by the protein tyrosine phosphatases Ptp2 and Ptp3. Molecular and cellular biology *19*, 7651-7660.

Mazo-Vargas, A., Park, H., Aydin, M., and Buchler, N.E. (2014). Measuring fast gene dynamics in single cells with time-lapse luminescence microscopy. Molecular biology of the cell *25*, 3699-3708.

McDowell, N., and Gurdon, J.B. (1999). Activin as a morphogen in Xenopus mesoderm induction. Semin Cell Dev Biol *10*, 311-317.

Merlini, L., Dudin, O., and Martin, S.G. (2013). Mate and fuse: how yeast cells do it. Open Biol *3*, 130008.

Mettetal, J.T., Muzzey, D., Gomez-Uribe, C., and van Oudenaarden, A. (2008). The frequency dependence of osmo-adaptation in *Saccharomyces cerevisiae*. Science *319*, 482-484.

Miermont, A., Uhlendorf, J., McClean, M., and Hersen, P. (2011). The Dynamical Systems Properties of the HOG Signaling Cascade. Journal of signal transduction *2011*, 930940.

Miermont, A., Waharte, F., Hu, S., McClean, M.N., Bottani, S., Leon, S., and Hersen, P. (2013). Severe osmotic compression triggers a slowdown of intracellular signaling, which can be explained by molecular crowding. Proceedings of the National Academy of Sciences of the United States of America *110*, 5725-5730.

Miyawaki, A., Nagai, T., and Mizuno, H. (2003). Mechanisms of protein fluorophore formation and engineering. Curr Opin Chem Biol *7*, 557-562.

Molin, C., Jauhiainen, A., Warringer, J., Nerman, O., and Sunnerhagen, P. (2009). mRNA stability changes precede changes in steady-state mRNA amounts during hyperosmotic stress. RNA *15*, 600-614.

Moore, T.I., Chou, C.S., Nie, Q., Jeon, N.L., and Yi, T.M. (2008). Robust spatial sensing of mating pheromone gradients by yeast cells. PloS one *3*, e3865.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nature methods *5*, 621-628.

Muhlenhoff, U., Richhardt, N., Ristow, M., Kispal, G., and Lill, R. (2002). The yeast frataxin homolog Yfh1p plays a specific role in the maturation of cellular Fe/S proteins. Hum Mol Genet *11*, 2025-2036.

Muller, E.M., Mackin, N.A., Erdman, S.E., and Cunningham, K.W. (2003). Fig1p facilitates Ca2+ influx and cell fusion during mating of *Saccharomyces cerevisiae*. The Journal of biological chemistry *278*, 38461-38469.

Mumberg, D., Muller, R., and Funk, M. (1995). Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. Gene *156*, 119-122.

Murphy, L.O., and Blenis, J. (2006). MAPK signal specificity: the right place at the right time. Trends Biochem Sci *31*, 268-275.

Murphy, L.O., MacKeigan, J.P., and Blenis, J. (2004). A network of immediate early gene products propagates subtle differences in mitogen-activated protein kinase signal amplitude and duration. Molecular and cellular biology *24*, 144-153.

Murphy, L.O., Smith, S., Chen, R.H., Fingar, D.C., and Blenis, J. (2002). Molecular interpretation of ERK signal duration by immediate early gene products. Nat Cell Biol *4*, 556-564.

Muslin, A.J. (2008). MAPK signalling in cardiovascular health and disease: molecular mechanisms and therapeutic targets. Clin Sci (Lond) *115*, 203-218.

Nadal-Ribelles, M., Conde, N., Flores, O., Gonzalez-Vallinas, J., Eyras, E., Orozco, M., de Nadal, E., and Posas, F. (2012). Hog1 bypasses stress-mediated down-regulation of transcription by RNA polymerase II redistribution and chromatin remodeling. Genome biology *13*, R106.

Nadal-Ribelles, M., Mas, G., Millan-Zambrano, G., Sole, C., Ammerer, G., Chavez, S., Posas, F., and de Nadal, E. (2015). H3K4 monomethylation dictates nucleosome dynamics and chromatin remodeling at stress-responsive genes. Nucleic Acids Res *43*, 4937-4949.

Nadal-Ribelles, M., Sole, C., Xu, Z., Steinmetz, L.M., de Nadal, E., and Posas, F. (2014). Control of Cdc28 CDK1 by a stress-induced lncRNA. Molecular cell *53*, 549-561.

Nagai, T., Ibata, K., Park, E.S., Kubota, M., Mikoshiba, K., and Miyawaki, A. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. Nature biotechnology *20*, 87-90.

Nagiec, M.J., and Dohlman, H.G. (2012). Checkpoints in a yeast differentiation pathway coordinate signaling during hyperosmotic stress. PLoS genetics *8*, e1002437.

Natarajan, A., Subramanian, S., and Srienc, F. (1998). Comparison of mutant forms of the green fluorescent protein as expression markers in Chinese hamster ovary (CHO) and *Saccharomyces cerevisiae* cells. J Biotechnol *62*, 29-45.

Neuert, G., Munsky, B., Tan, R.Z., Teytelman, L., Khammash, M., and van Oudenaarden, A. (2013). Systematic identification of signal-activated stochastic gene regulation. Science *339*, 584-587.

O'Rourke, S.M., Herskowitz, I., and O'Shea, E.K. (2002). Yeast go the whole HOG for the hyperosmotic response. Trends in genetics : TIG *18*, 405-412.

Oehlen, L.J., McKinney, J.D., and Cross, F.R. (1996). Ste12 and Mcm1 regulate cell cycle-dependent transcription of FAR1. Molecular and cellular biology *16*, 2830-2837.

Olenych, S.G., Claxton, N.S., Ottenberg, G.K., and Davidson, M.W. (2007). The fluorescent protein color palette. Curr Protoc Cell Biol *Chapter 21*, Unit 21 25.

Olson, K.A., Nelson, C., Tai, G., Hung, W., Yong, C., Astell, C., and Sadowski, I. (2000). Two regulators of Ste12p inhibit pheromone-responsive transcription by separate mechanisms. Molecular and cellular biology *20*, 4199-4209.

Page, B.D., Satterwhite, L.L., Rose, M.D., and Snyder, M. (1994). Localization of the Kar3 kinesin heavy chain-related protein requires the Cik1 interacting protein. J Cell Biol *124*, 507-519.

Pasa-Tolic, L., Lipton, M.S., Masselon, C.D., Anderson, G.A., Shen, Y., Tolic, N., and Smith, R.D. (2002). Gene expression profiling using advanced mass spectrometric approaches. J Mass Spectrom *37*, 1185-1198.

Paterson, J.M., Ydenberg, C.A., and Rose, M.D. (2008). Dynamic localization of yeast Fus2p to an expanding ring at the cell fusion junction during mating. J Cell Biol *181*, 697-709.

Patterson, J.C., Klimenko, E.S., and Thorner, J. (2010). Single-cell analysis reveals that insulation maintains signaling specificity between two yeast MAPK pathways with common components. Science signaling *3*, ra75.

Pelet, S., Aymoz, D., and Durandau, E. (2013). Temporal quantification of MAPK induced expression in single yeast cells. Journal of visualized experiments : JoVE.

Pelet, S., Dechant, R., Lee, S.S., van Drogen, F., and Peter, M. (2012). An integrated image analysis platform to quantify signal transduction in single cells. Integrative biology : quantitative biosciences from nano to macro *4*, 1274-1282.

Pelet, S., Rudolf, F., Nadal-Ribelles, M., de Nadal, E., Posas, F., and Peter, M. (2011). Transient activation of the HOG MAPK pathway regulates bimodal gene expression. Science *332*, 732-735.

Peng, J., Elias, J.E., Thoreen, C.C., Licklider, L.J., and Gygi, S.P. (2003). Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. J Proteome Res *2*, 43-50.

Pepperkok, R., and Ellenberg, J. (2006). High-throughput fluorescence microscopy for systems biology. Nat Rev Mol Cell Biol *7*, 690-696.

Peter, M., Gartner, A., Horecka, J., Ammerer, G., and Herskowitz, I. (1993). FAR1 links the signal transduction pathway to the cell cycle machinery in yeast. Cell *73*, 747-760.

Peter, M., and Herskowitz, I. (1994). Direct inhibition of the yeast cyclin-dependent kinase Cdc28-Cln by Far1. Science *265*, 1228-1231.

Pi, H., Chien, C.T., and Fields, S. (1997). Transcriptional activation upon pheromone stimulation mediated by a small domain of *Saccharomyces cerevisiae* Ste12p. Molecular and cellular biology *17*, 6410-6418.

Picotti, P., Clement-Ziza, M., Lam, H., Campbell, D.S., Schmidt, A., Deutsch, E.W., Rost, H., Sun, Z., Rinner, O., Reiter, L., *et al.* (2013). A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. Nature *494*, 266-270.

Pokholok, D.K., Zeitlinger, J., Hannett, N.M., Reynolds, D.B., and Young, R.A. (2006). Activated signal transduction kinases frequently occupy target genes. Science *313*, 533-536.

Posas, F., Chambers, J.R., Heyman, J.A., Hoeffler, J.P., de Nadal, E., and Arino, J. (2000). The transcriptional response of yeast to saline stress. The Journal of biological chemistry *275*, 17249-17255.

Posas, F., and Saito, H. (1998). Activation of the yeast SSK2 MAP kinase kinase kinase by the SSK1 two-component response regulator. The EMBO journal *17*, 1385-1394.

Posas, F., Wurgler-Murphy, S.M., Maeda, T., Witten, E.A., Thai, T.C., and Saito, H. (1996). Yeast HOG1 MAP kinase cascade is regulated by a multistep phosphorelay mechanism in the SLN1-YPD1-SSK1 "two-component" osmosensor. Cell *86*, 865-875.

Primig, M., Winkler, H., and Ammerer, G. (1991). The DNA binding and oligomerization domain of MCM1 is sufficient for its interaction with other regulatory proteins. The EMBO journal *10*, 4209-4218.

Proft, M., Mas, G., de Nadal, E., Vendrell, A., Noriega, N., Struhl, K., and Posas, F. (2006). The stress-activated Hog1 kinase is a selective transcriptional elongation factor for genes responding to osmotic stress. Molecular cell *23*, 241-250.

Purvis, J.E., Karhohs, K.W., Mock, C., Batchelor, E., Loewer, A., and Lahav, G. (2012). p53 dynamics control cell fate. Science *336*, 1440-1444.

Radmaneshfar, E., Kaloriti, D., Gustin, M.C., Gow, N.A., Brown, A.J., Grebogi, C., Romano, M.C., and Thiel, M. (2013). From START to FINISH: the influence of osmotic stress on the cell cycle. PloS one *8*, e68067.

Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. Nature methods *5*, 877-879.

Ralser, M., Kuhl, H., Ralser, M., Werber, M., Lehrach, H., Breitenbach, M., and Timmermann, B. (2012). The *Saccharomyces cerevisiae* W303-K6001 cross-platform genome sequence: insights into ancestry and physiology of a laboratory mutt. Open Biol *2*, 120093.

Raser, J.M., and O'Shea, E.K. (2004). Control of stochasticity in eukaryotic gene expression. Science *304*, 1811-1814.

Raveh-Sadka, T., Levo, M., Shabi, U., Shany, B., Keren, L., Lotan-Pompan, M., Zeevi, D., Sharon, E., Weinberger, A., and Segal, E. (2012). Manipulating nucleosome disfavoring sequences allows fine-tune regulation of gene expression in yeast. Nat Genet *44*, 743-750.

Regot, S., de Nadal, E., Rodriguez-Navarro, S., Gonzalez-Novo, A., Perez-Fernandez, J., Gadal, O., Seisenbacher, G., Ammerer, G., and Posas, F. (2013). The Hog1 stress-activated protein kinase targets nucleoporins to control mRNA export upon stress. The Journal of biological chemistry *288*, 17384-17398.

Reinke, A.W., Grant, R.A., and Keating, A.E. (2010). A Synthetic Coiled-Coil Interactome Provides Heterospecific Modules for Molecular Engineering. J Am Chem Soc *132*, 6025-6031.

Reiser, V., Ruis, H., and Ammerer, G. (1999). Kinase activity-dependent nuclear export opposes stress-induced nuclear accumulation and retention of Hog1 mitogen-activated protein kinase in the budding yeast *Saccharomyces cerevisiae*. Molecular biology of the cell *10*, 1147-1161.

Remenyi, A., Good, M.C., Bhattacharyya, R.P., and Lim, W.A. (2005). The role of docking interactions in mediating signaling input, output, and discrimination in the yeast MAPK network. Molecular cell *20*, 951-962.

Remington, S.J. (2006). Fluorescent proteins: maturation, photochemistry and photophysics. Current opinion in structural biology *16*, 714-721.

Ren, B., Robert, F., Wyrick, J.J., Aparicio, O., Jennings, E.G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin, E.*, et al.* (2000). Genome-wide location and function of DNA binding proteins. Science *290*, 2306-2309.

Rep, M., Albertyn, J., Thevelein, J.M., Prior, B.A., and Hohmann, S. (1999a). Different signalling pathways contribute to the control of GPD1 gene expression by osmotic stress in *Saccharomyces cerevisiae*. Microbiology *145 (Pt 3)*, 715-727.

Rep, M., Reiser, V., Gartner, U., Thevelein, J.M., Hohmann, S., Ammerer, G., and Ruis, H. (1999b). Osmotic stress-induced gene expression in *Saccharomyces cerevisiae* requires Msn1p and the novel nuclear factor Hot1p. Molecular and cellular biology *19*, 5474-5485.

Roberts, C.J., Nelson, B., Marton, M.J., Stoughton, R., Meyer, M.R., Bennett, H.A., He, Y.D., Dai, H., Walker, W.L., Hughes, T.R.*, et al.* (2000). Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. Science *287*, 873-880.

Rogers, K.W., and Schier, A.F. (2011). Morphogen gradients: from generation to interpretation. Annu Rev Cell Dev Biol *27*, 377-407.

Rose, M., and Botstein, D. (1983). Construction and use of gene fusions to lacZ (beta-galactosidase) that are expressed in yeast. Methods Enzymol *101*, 167-180.

Roth, C.M. (2002). Quantifying gene expression. Curr Issues Mol Biol *4*, 93-100.

Roy, A., Lu, C.F., Marykwas, D.L., Lipke, P.N., and Kurjan, J. (1991). The AGA1 product is involved in cell surface attachment of the *Saccharomyces cerevisiae* cell adhesion glycoprotein a-agglutinin. Molecular and cellular biology *11*, 4196-4206.

Saito, H., and Posas, F. (2012). Response to hyperosmotic stress. Genetics *192*, 289-318.

Saka, Y., Hagemann, A.I., Piepenburg, O., and Smith, J.C. (2007). Nuclear accumulation of Smad complexes occurs only after the midblastula transition in Xenopus. Development *134*, 4209-4218.

Salzman, V., Porro, V., Bollati-Fogolin, M., and Aguilar, P.S. (2015). Quantitation of yeast cell-cell

fusion using multicolor flow cytometry. Cytometry A *87*, 843-854.

Santos, S.D., Verveer, P.J., and Bastiaens, P.I. (2007). Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. Nat Cell Biol *9*, 324-330.

Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thastrom, A., Field, Y., Moore, I.K., Wang, J.P., and Widom, J. (2006). A genomic code for nucleosome positioning. Nature *442*, 772-778.

Segal, E., and Widom, J. (2009). Poly(dA:dT) tracts: major determinants of nucleosome organization. Current opinion in structural biology *19*, 65-71.

Shah, P., Ding, Y., Niemczyk, M., Kudla, G., and Plotkin, J.B. (2013). Rate-limiting steps in yeast protein translation. Cell *153*, 1589-1601.

Shahrezaei, V., and Swain, P.S. (2008). The stochastic nature of biochemical networks. Current opinion in biotechnology *19*, 369-374.

Shaner, N.C., Steinbach, P.A., and Tsien, R.Y. (2005). A guide to choosing fluorescent proteins. Nature methods *2*, 905-909.

Shapiro, H.M. (2003).

Sharon, E., Kalma, Y., Sharp, A., Raveh-Sadka, T., Levo, M., Zeevi, D., Keren, L., Yakhini, Z., Weinberger, A., and Segal, E. (2012). Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. Nature biotechnology *30*, 521-530.

Shav-Tal, Y., Singer, R.H., and Darzacq, X. (2004). Imaging gene expression in single living cells. Nat Rev Mol Cell Biol *5*, 855-861.

Sheff, M.A., and Thorn, K.S. (2004). Optimized cassettes for fluorescent protein tagging in *Saccharomyces cerevisiae*. Yeast *21*, 661-670.

Shen, S., Tobery, C.E., and Rose, M.D. (2009). Prm3p is a pheromone-induced peripheral nuclear envelope protein required for yeast nuclear fusion. Molecular biology of the cell *20*, 2438-2450.

Skotheim, J.M., Di Talia, S., Siggia, E.D., and Cross, F.R. (2008). Positive feedback of G1 cyclins ensures coherent cell cycle entry. Nature *454*, 291-296.

Smith, M.G., and Snyder, M. (2006). Yeast as a model for human disease. Curr Protoc Hum Genet *Chapter 15*, Unit 15 16.

Smith, R.D., Anderson, G.A., Lipton, M.S., Masselon, C., Pasa-Tolic, L., Shen, Y., and Udseth, H.R. (2002a). The use of accurate mass tags for high-throughput microbial proteomics. Omics : a journal of integrative biology *6*, 61-90.

Smith, R.D., Anderson, G.A., Lipton, M.S., Pasa-Tolic, L., Shen, Y., Conrads, T.P., Veenstra, T.D., and Udseth, H.R. (2002b). An accurate mass tag strategy for quantitative and high-throughput proteome measurements. Proteomics *2*, 513-523.

Smith, R.D., Pasa-Tolic, L., Lipton, M.S., Jensen, P.K., Anderson, G.A., Shen, Y., Conrads, T.P., Udseth, H.R., Harkewicz, R., Belov, M.E.*, et al.* (2001). Rapid quantitative measurements of proteomes by Fourier transform ion cyclotron resonance mass spectrometry. Electrophoresis *22*, 1652-1668.

Song, D., Dolan, J.W., Yuan, Y.L., and Fields, S. (1991). Pheromone-dependent phosphorylation of the yeast STE12 protein correlates with transcriptional activation. Genes & development *5*, 741-750.

Sopko, R., Huang, D., Preston, N., Chua, G., Papp, B., Kafadar, K., Snyder, M., Oliver, S.G., Cyert, M., Hughes, T.R.*, et al.* (2006). Mapping pathways and phenotypes by systematic gene overexpression. Molecular cell *21*, 319-330.

Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D., and Futcher, B. (1998). Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. Molecular biology of the cell *9*, 3273-3297.

Sproul, L.R., Anderson, D.J., Mackey, A.T., Saunders, W.S., and Gilbert, S.P. (2005). Cik1 targets the minus-end kinesin depolymerase kar3 to microtubule plus ends. Current biology : CB *15*, 1420-1427.

Strickfaden, S.C., Winters, M.J., Ben-Ari, G., Lamson, R.E., Tyers, M., and Pryciak, P.M. (2007). A mechanism for cell-cycle regulation of MAP kinase signaling in a yeast differentiation pathway. Cell *128*, 519-531.

Styles, E.B., Friesen, H., Boone, C., and Andrews, B.J. (2016). High-Throughput Microscopy-Based Screening in *Saccharomyces cerevisiae*. Cold Spring Harb Protoc *2016*, pdb top087593.

Su, T.C., Tamarkina, E., and Sadowski, I. (2010). Organizational constraints on Ste12 cis-elements for a pheromone response in *Saccharomyces cerevisiae*. The FEBS journal *277*, 3235-3248.

Sun, Y., Liu, W.Z., Liu, T., Feng, X., Yang, N., and Zhou, H.F. (2015). Signaling pathway of MAPK/ERK

in cell proliferation, differentiation, migration, senescence and apoptosis. J Recept Signal Transduct Res *35*, 600-604.

Suter, D.M., Molina, N., Gatfield, D., Schneider, K., Schibler, U., and Naef, F. (2011). Mammalian genes are transcribed with widely different bursting kinetics. Science *332*, 472-474.

Takekawa, M., Posas, F., and Saito, H. (1997). A human homolog of the yeast Ssk2/Ssk22 MAP kinase kinase kinases, MTK1, mediates stress-induced activation of the p38 and JNK pathways. The EMBO journal *16*, 4973-4982.

Tamas, M.J., Luyten, K., Sutherland, F.C., Hernandez, A., Albertyn, J., Valadi, H., Li, H., Prior, B.A., Kilian, S.G., Ramos, J.*, et al.* (1999). Fps1p controls the accumulation and release of the compatible solute glycerol in yeast osmoregulation. Mol Microbiol *31*, 1087-1104.

Tan-Wong, S.M., Zaugg, J.B., Camblong, J., Xu, Z., Zhang, D.W., Mischo, H.E., Ansari, A.Z., Luscombe, N.M., Steinmetz, L.M., and Proudfoot, N.J. (2012). Gene loops enhance transcriptional directionality. Science *338*, 671-675.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A.*, et al.* (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nature methods *6*, 377-382.

Tatebayashi, K., Tanaka, K., Yang, H.Y., Yamamoto, K., Matsushita, Y., Tomida, T., Imai, M., and Saito, H. (2007). Transmembrane mucins Hkr1 and Msb2 are putative osmosensors in the SHO1 branch of yeast HOG pathway. The EMBO journal *26*, 3521-3533.

Tedford, K., Kim, S., Sa, D., Stevens, K., and Tyers, M. (1997). Regulation of the mating pheromone and invasive growth responses in yeast by two MAP kinase substrates. Current biology : CB *7*, 228-238.

Teixeira, M.C., Monteiro, P.T., Guerreiro, J.F., Goncalves, J.P., Mira, N.P., dos Santos, S.C., Cabrito, T.R., Palma, M., Costa, C., Francisco, A.P.*, et al.* (2014). The YEASTRACT database: an upgraded information system for the analysis of gene and genomic transcription regulation in *Saccharomyces cerevisiae*. Nucleic Acids Res *42*, D161-166.

Thompson, A., and Gasson, M.J. (2001). Location effects of a reporter gene on expression levels and on native protein synthesis in Lactococcus lactis and *Saccharomyces cerevisiae*. Appl Environ Microbiol *67*, 3434-3439.

Thompson, K.E., Bashor, C.J., Lim, W.A., and Keating, A.E. (2012). SYNZIP protein interaction toolbox: in vitro and in vivo specifications of heterospecific coiled-coil interaction domains. ACS synthetic biology *1*, 118-129.

Timney, B.L., Tetenbaum-Novatt, J., Agate, D.S., Williams, R., Zhang, W., Chait, B.T., and Rout, M.P. (2006). Simple kinetic relationships and nonspecific competition govern nuclear import rates in vivo. J Cell Biol *175*, 579-593.

Trevino, V., Falciani, F., and Barrera-Saldana, H.A. (2007). DNA microarrays: a powerful genomic tool for biomedical and clinical research. Mol Med *13*, 527-541.

Tsien, R.Y. (2009). Constructing and exploiting the fluorescent protein paintbox (Nobel Lecture). Angewandte Chemie *48*, 5612-5626.

Tyson, C.B., Lord, P.G., and Wheals, A.E. (1979). Dependency of size of *Saccharomyces cerevisiae* cells on growth rate. J Bacteriol *138*, 92-98.

Vaga, S., Bernardo-Faura, M., Cokelaer, T., Maiolica, A., Barnes, C.A., Gillet, L.C., Hegemann, B., van Drogen, F., Sharifian, H., Klipp, E.*, et al.* (2014). Phosphoproteomic analyses reveal novel cross-modulation mechanisms between two signaling pathways in yeast. Molecular systems biology *10*, 767.

Valtz, N., Peter, M., and Herskowitz, I. (1995). FAR1 is required for oriented polarization of yeast cells in response to mating pheromones. J Cell Biol *131*, 863-873.

Van Arsdell, S.W., Stetler, G.L., and Thorner, J. (1987). The yeast repeated element sigma contains a hormone-inducible promoter. Molecular and cellular biology *7*, 749-759.

van Dijk, D., Sharon, E., Lotan-Pompan, M., Weinberger, A., Segal, E., and Carey, L.B. (2017). Large-scale mapping of gene regulatory logic reveals context-dependent repression by transcriptional activators. Genome Research *27*, 87-94.

Varshavsky, A. (1996). The N-end rule: functions, mysteries, uses. Proceedings of the National Academy of Sciences of the United States of America *93*, 12142-12149.

Wang, R., and Brattain, M.G. (2007). The maximal size of protein to diffuse through the nuclear pore is larger than 60kDa. FEBS Lett *581*, 3164-3170.

Washburn, M.P., Wolters, D., and Yates, J.R., 3rd (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nature biotechnology *19*, 242-247.

Weake, V.M., and Workman, J.L. (2010). Inducible gene expression: diverse regulatory mechanisms. Nature reviews Genetics *11*, 426-437.

Weingarten-Gabbay, S., and Segal, E. (2014). The grammar of transcriptional regulation. Human genetics *133*, 701-711.

Westfall, P.J., Patterson, J.C., Chen, R.E., and Thorner, J. (2008). Stress resistance and signal fidelity independent of nuclear MAPK function. Proceedings of the National Academy of Sciences of the United States of America *105*, 12212-12217.

Widmann, C., Gibson, S., Jarpe, M.B., and Johnson, G.L. (1999). Mitogen-activated protein kinase: conservation of a three-kinase module from yeast to human. Physiol Rev *79*, 143-180.

Willer, M., Hoffmann, L., Styrkarsdottir, U., Egel, R., Davey, J., and Nielsen, O. (1995). Two-step activation of meiosis by the mat1 locus in Schizosaccharomyces pombe. Molecular and cellular biology *15*, 4964-4970.

Wilson, T., and Hastings, J.W. (1998). Bioluminescence. Annu Rev Cell Dev Biol *14*, 197-230.

Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J.D., Bussey, H.*, et al.* (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science *285*, 901-906.

Wong Sak Hoi, J., and Dumas, B. (2010). Ste12 and Ste12-like proteins, fungal transcription factors regulating development and pathogenicity. Eukaryotic cell *9*, 480-485.

Wosika, V., Durandau, E., Varidel, C., Aymoz, D., Schmitt, M., and Pelet, S. (2016). New families of single integration vectors and gene tagging plasmids for genetic manipulations in budding yeast. Mol Genet Genomics *291*, 2231-2240.

Wu, B., Chao, J.A., and Singer, R.H. (2012). Fluorescence fluctuation spectroscopy enables quantitative imaging of single mRNAs in living cells. Biophysical journal *102*, 2936-2944.

Wu, C., Jansen, G., Zhang, J., Thomas, D.Y., and Whiteway, M. (2006). Adaptor protein Ste50p links the Ste11p MEKK to the HOG pathway through plasma membrane association. Genes & development *20*, 734-746.

Wu, J.Q., and Pollard, T.D. (2005). Counting cytokinesis proteins globally and locally in fission yeast. Science *310*, 310-314.

Yaakov, G., Duch, A., Garcia-Rubio, M., Clotet, J., Jimenez, J., Aguilera, A., and Posas, F. (2009). The stress-activated protein kinase Hog1 mediates S phase delay in response to osmostress. Molecular biology of the cell *20*, 3572-3582.

Yamanishi, M., Ito, Y., Kintaka, R., Imamura, C., Katahira, S., Ikeuchi, A., Moriya, H., and Matsuyama, T. (2013). A genome-wide activity assessment of terminator regions in *Saccharomyces cerevisiae* provides a "terminatome" toolbox. ACS synthetic biology *2*, 337-347.

Yu, R.C., Pesce, C.G., Colman-Lerner, A., Lok, L., Pincus, D., Serra, E., Holl, M., Benjamin, K., Gordon, A., and Brent, R. (2008). Negative feedback that improves information transmission in yeast signalling. Nature *456*, 755-761.

Yuan, Y.L., and Fields, S. (1991). Properties of the DNA-binding domain of the *Saccharomyces cerevisiae* STE12 protein. Molecular and cellular biology *11*, 5910-5918.

Zapater, M., Sohrmann, M., Peter, M., Posas, F., and de Nadal, E. (2007). Selective requirement for SAGA in Hog1-mediated gene expression depending on the severity of the external osmostress conditions. Molecular and cellular biology *27*, 3900-3910.

Zarzov, P., Mazzoni, C., and Mann, C. (1996). The SLT2(MPK1) MAP kinase is activated during periods of polarized cell growth in yeast. The EMBO journal *15*, 83-91.

Zechner, C., Unger, M., Pelet, S., Peter, M., and Koeppl, H. (2014). Scalable inference of heterogeneous reaction kinetics from pooled single-cell recordings. Nature methods *11*, 197-202.

Zeitlinger, J., Simon, I., Harbison, C.T., Hannett, N.M., Volkert, T.L., Fink, G.R., and Young, R.A. (2003). Program-specific distribution of a transcription factor dependent on partner transcription factor and MAPK signaling. Cell *113*, 395-404.

Zenklusen, D., Larson, D.R., and Singer, R.H. (2008). Single-RNA counting reveals alternative modes of gene expression in yeast. Nature structural & molecular biology *15*, 1263-1271.

Zhan, X.L., Deschenes, R.J., and Guan, K.L. (1997). Differential regulation of FUS3 MAP kinase by tyrosine-specific phosphatases PTP2/PTP3 and dual-specificity phosphatase MSG5 in *Saccharomyces cerevisiae*. Genes & development *11*, 1690-1702.

Zhang, W., Espinoza, D., Hines, V., Innis, M., Mehta, P., and Miller, D.L. (1997). Characterization of beta-amyloid peptide precursor processing by the yeast

Yap3 and Mkc7 proteases. Biochimica et biophysica acta *1359*, 110-122.

Zhang, W., and Liu, H.T. (2002). MAPK signal pathways in the regulation of cell proliferation in mammalian cells. Cell Res *12*, 9-18.

Zhao, S., Douglas, N.W., Heine, M.J., Williams, G.M., Winther-Larsen, H.C., and Meaden, P.G. (1994). The STL1 gene of *Saccharomyces cerevisiae* is predicted to encode a sugar transporter-like protein. Gene *146*, 215-219.

Zilfou, J.T., and Lowe, S.W. (2009). Tumor suppressive functions of p53. Cold Spring Harb Perspect Biol *1*, a001883.