Year : 2019

Single-cell analysis of osmostress genes transcription

Victoria Wosika-Witz

UNIL | Université de Lausanne
Faculté de biologie et de médecine

Département de Microbiologie Fondamentale (DMF)

SINGLE-CELL ANALYSIS OF OSMOSTRESS GENES TRANSCRIPTION

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de Biologie et de Médecine
de l'Université de Lausanne

par

Victoria WOSIKA-WITZ

Master de l'Université de Lausanne

Jury

Prof. Laurent Lehmann, Président
Dr. Serge Pelet, Directeur de thèse
Prof. Sophie Martin, Co-directrice
Prof. David Shore, Co-directeur
Prof. Peter Swain, Expert
Prof. David Suter, Expert

Lausanne 2019

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.	Laurent	**Lehmann**
Directeur·trice de thèse	Monsieur	Prof.	Serge	**Pelet**
Co-directeurs·trices	Madame	Prof.	Sophie	**Martin**
	Monsieur	Prof.	David	**Shore**
Expert·e·s	Monsieur	Prof.	Peter	**Swain**
	Monsieur	Prof.	David	**Suter**

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

Madame Victoria Wosika-Witz

Maîtrise universitaire ès Sciences moléculaires du vivant, Université de Lausanne

intitulée

Single-cell analysis of osmostress genes transcription

Lausanne, le 27 septembre 2019

pour le Doyen
de la Faculté de biologie et de médecine

Prof. Laurent Lehmann

[…] *I don't want your hope. I want you to panic.*
– Greta Thunberg, WEF 25th of January 2019.

Pelet lab on its third Climate March, Lausanne on the 17th of January 2020. Credits: Cassandre Kinnaer.

Acknowledgments

First, I would like to thank the members of my committee for accepting to be part of this journey and for their contributions to the completion of my thesis. A special thanks to **David**, for your interest in my numerous projects, which were not on RP genes as originally planned. **Sophie**, thank you for correcting the introduction of my manuscript, and for your support and advices during these five years.

I would like to thank my thesis supervisor **Serge**. Although you always text me in English, je vais continuer à t'écrire en français. Merci d'avoir vu plus loin que ma timidité et mon manque d'assurance lors de mon interview en m'offrant une place dans ton labo. Merci de m'avoir mise sur le projet PP7, je l'ai aimé autant que je l'ai détesté, mais au final on ne s'en est pas trop mal sortis. Merci pour ta confiance et la liberté que tu m'as donné, j'ai pu tester mes (trop?) nombreuses idées. Finalement, merci pour ta patience, ton empathie et ton écoute lors de mes hauts et nombreux bas, autant professionnels que personnels.

Because science does not exist without a lab, I'm grateful to my lab mates. **Eric**, merci pour ton enthousiasme pour la science, ta persévérance, ta pédagogie incomprise et tes goûts musicaux. Merci d'être un ami toujours présent et plein de bons conseils. **Min**, darling, if you weren't there for me during these rough times, I would have quit. Thanks for your support, your advices and for pushing me to be a stronger person. Thank you for all your Sundays spent in the lab starting our experiments, your help was precious. **Gaëlle**, mon rayon de soleil, merci d'avoir été là pour moi alors que c'était à moi de te guider. Merci pour ton excellent travail, ta maturité, ta constante bonne humeur et tes conseils, tu es une magnifique personne avec un cœur en or. Je suis fière d'avoir été ta superviseuse de master et encore plus d'être ton amie. **Marta**, merci pour ta précieuse aide au labo, tes pauses café, notre re-rencontre et ton humour qui ont illuminés ma fin de thèse. **Delphine**, merci pour ton aide avec Matlab, principalement pour tes scriptes. Je ne maitrise toujours pas les *for loops* et les *find* mais t'auras essayé. **Sandrine**, trop peu de temps passé ensemble mais déjà tant à te remercier. Merci pour ton soutient pendant l'écriture de ma thèse, pour les pauses café-glaces et pour voir tant de choses chez moi que je ne vois pas. **Terry** et **Ives**, merci pour vos sourires communicatifs et votre deuxième degré que je ne comprends pas.

Because our lab cannot be dissociated from the other yeast lab of the DMF, I would like to thank all the current and past members from the Martin lab, for always be willing to help your colleagues next door. **Dani** chérie, I cannot even list all I'm thankful for. When you graduated, you left a huge whole. Thanks for your smile, your sense of humor, your support, your hugs. You were the real *Sunshine*. **Laura**, thank you for your advices, on *cerevisiae* and on people, and for your support on tough personal matters. You are the strongest and most intelligent person I know and I admire you. **Laeti**, merci pour ton soutient cette dernière année de thèse, ton humour, ton authenticité, ton écoute, tes conseils et nos retours en train. **Vincent**, désolée de t'avoir si souvent dérangé pendant que tu travaillais, juste pour pouvoir parler pendant l'écriture de la thèse. Merci pour ton aide autant professionnelle que personnelle, tes chocolats aux pauses café et pour ton humour décapant. **Veneta**, thank you for your spontaneity, kindness, beauty and your sense of humor, you are a true inspiration. **Cass**, **Magda, Bita, Ingrid** and **Thais**, thanks for adding a cute cat side to PhD. Thanks to all the Martin ladies for the girls'night out and baby showers. **Felipe**, thanks for giving us our first home.

Parce que je n'aurais clairement pas eu le temps d'accomplir autant de travail sans le soutien technique et administratif du département, je voudrais remercier **Nadine**, **Aline**, **Nassim**, **Vincent**, **Martine**, **Paul**, **Nazife**, **Janette**, **Maria**, **Armindo**, **Clémence**, nos stagiaires HES-Santé et mes étudiants de master Gaëlle et **Gabriele** pour leur précieux travail.

Parce que la science ne fonctionne pas la plupart du temps, il est important d'avoir des amis et de la famille là pour nous soutenir. Je voudrais d'abord remercier les BioLOLgistes : **Camille, Léa, Florian, Maxime, Daniel, Mélanie**, **Théo**, **Noémie**, **Dušan, Benji**. Vous avez été essentiels à la réussite de mes études et je remercie tous les jours l'Université de Lausanne de nous avoir mis dans le même amphi il y a 10 ans. Merci pour votre soutient, nos moments détentes, nos discussions pour changer le monde, nos vacances et nos blagues de biologistes incomprises. Merci pour vous, tout simplement**.** Merci plus particulièrement à Camion et Léon. Ma blonde, **Margaux**, merci d'avoir été là pour moi à tout instant. Ton amitié inconditionnelle, tes conseils (médicaux ou non) et ton calme ont étés des piliers pour moi. Mon autre blonde et plus vieille amie **Stéph**. Merci d'être toujours là pour moi malgré la distance, merci pour ton humour, nos facetimes, tes biscuits par la poste, tes aller-retour en train et ton joli accent jurassien.

Je voudrais remercier ma famille pour leur soutient tout au long de ces années d'Université. Je suis désolée de ne vous avoir que trop peu vus ; grâce à vous j'ai compris que le travail ce n'est pas la vie. **Maman**, merci de toujours vouloir prendre ma défense et de m'aider à relativiser. Merci d'être là pour moi, même quand c'est toi qui en aurais le plus besoin. **Papa**, merci pour ton aide pour nos nombreux déménagements et pour être si spécial. **Nona**, oui ça va bien à l'école, merci d'être aussi fière de moi sans savoir ce que je fais vraiment. **Aaron**, merci de ton oreille attentive, de me forcer à me lever le matin et de ta présence au quotidien.

Last but not least, non je ne vais pas écrire en anglais ne t'inquiète pas, je voudrais remercier mon meilleur ami, mari et coach personnel, **Sébastien**. Partager la vie d'une personne aussi sensible ne doit pas être facile et pourtant tu es toujours là, calme, patient et à mon écoute. Donc merci, merci de me rattraper à chaque fois qu'un projet tombe à l'eau ; tu as appris plus vite que moi à ce jeu-là. Merci d'être venu de trop nombreuses fois avec moi à l'uni le dimanche et désolé que mes « juste dix minutes » ce soient si souvent transformées en quarante-cinq. Merci pour tous ces « je suis là » et d'effectivement toujours être là. Merci de m'aider à lire entre les lignes de la vie et des gens, et de me pousser à être toujours une meilleure personne. Merci de m'avoir soutenue à la fois pendant mon master et mon doctorat, et supportée dans les deux lignes droites intenses, je n'y serais pas arrivée sans toi. Merci de m'avoir dit « oui » pour la vie, je t'aime.

Abstract

A crucial ability of any cellular life is sensing its environment and integrating its changes into intracellular information. Faithful reading and response to environmental inputs is thus essential for cell survival. In all eukaryotic cells, multiple environmental stimuli are sensed via the highly conserved MAPK signaling pathways. Each MAPK pathway has the ability to detect specific extracellular signal and encodes the dedicated response both through post-translational modifications and at the gene expression level, to ensure long-term adaptation and cell fate decision. In the budding yeast *S. cerevisiae*, an increase in the environment osmolarity is sensed by the High Osmolarity Glycerol (HOG) pathway, which is responsible for the accumulation of glycerol and restoration of turgor pressure. The direct output from the response is thus glycerol uptake and synthesis, but also the gene expression of glycogenesis and osmostress enzymes and proteins. Upon osmotic shock, the MAPK of the pathway, Hog1, relocates into the nucleus to trigger gene expression, which are otherwise repressed under non-stressful conditions. Indeed, under basal conditions, osmostress genes are repressed through closed chromatin conformation. Thanks to a multitasking MAPK and to recruiter transcription factors, Hog1 targets chromatin remodeling and modifying complexes to osmostress loci. Hog1 activity thus promotes nucleosomes eviction and RNA polymerase recruitment at stress genes. Induction of growth-repressed genes by HOG MAPK activity is a model for chromatin-based regulation of gene expression. However, quantitative measurements of osmostress gene promoters' gene transcription dynamics, revealing the dynamic regulation occurring at each locus, are still missing.

Gene expression has been shown to occurs in bursts of mRNA production and the bursting kinetics, like bursting time, amplitude, duration or number of bursts, is known to reflect the regulation occurring at the DNA locus. Thanks to live RNA imaging methods based on phage coat proteins and their high affinity RNA binding sequences (like the PP7 and MS2 systems), gene bursting can be visualized and quantified in live single-cells. In this study, we construct PP7 reporter strains for models of osmostress genes and acquired live single cell dynamic data on mRNA production by osmo-stress induced promoters. We demonstrate that transcription initiation happens preferentially during ascending or stable Hog1 activity, highlighting a balance between activating and inhibitory pathways directly acting on each locus. We show

that the transcription initiation time of HOG-induced genes depends on the chromatin environment of the locus, with a positive correlation between the basal expression level and the fraction of transcribing cells upon osmotic shock. We demonstrate that transcription factor dependency is dictated by the requirement for chromatin remodeling, which is translated into a delay transcription initiation, a lower RNA pol II recruitment and a shorter transcriptional window. Finally, we show that transcription termination is dictated by a combination of Hog1 activity and the regulation from the locus.

In this study, we also report on the development of novel tools; for a highly robust genetic manipulation of the budding yeast and for the live single-cell monitoring of transcription and translation from a single allele, that enables to study post transcriptional regulation and local translation of mRNA.

Résumé

Une capacité cruciale de toute vie cellulaire est de percevoir son environnement et d'intégrer ses changements sous forme d'informations intracellulaires. Une lecture fidèle et une réponse appropriée aux stimuli environnementaux sont donc essentielles à la survie de toutes cellules. Dans toutes les cellules eucaryotes, des levures aux cellules végétales, les stimuli environnementaux sont détectés via des voies de signalisation MAPK hautement conservées. Chaque voie de signalisation est dédiée à la détection d'un signal extracellulaire précis et à sa correcte intégration à la fois au niveau de modification post-translationelles de protéines et de l'expression des gènes, afin de garantir une adaptation à long terme. Dans la levure à bourgeonnements *S. cerevisiae*, une augmentation de l'osmolarité de l'environnement est détectée par la voie de signalisation High Osmolarity Glycerol (HOG), qui est responsable de l'accumulation de glycérol et de la restauration de la turgescence. Le résultat de la réponse est donc l'absorption et la synthèse de glycérol, mais également l'expression génique des enzymes et protéines de la glycogénèse. Lors d'un choc osmotique, la MAP kinase de la voie HOG, Hog1, relocalise dans le noyau pour déclencher l'expression des gènes, qui sont normalement réprimés dans des conditions non-stressantes. En effet, en conditions basales, les gènes de réponse au stress osmotique sont réprimés par une conformation fermée de la chromatine. Grâce à une MAPK multitâche et à des facteurs de transcription recruteurs, Hog1 cible les complexes de remodelage et de modification de la chromatine au gènes osmotiques. L'activité de Hog1 favorise donc l'éviction des nucléosomes et le recrutement de l'ARN polymérase au niveau des gènes de réponse au stress osmotique. L'induction des gènes réprimés par la croissance cellulaire par l'activité MAPK est un modèle de régulation génique par la chromatine. Cependant, des données quantitatives sur la dynamique de transcription des gènes de réponse au stress osmotique, révélant la régulation dynamique de chaque locus, manquent toujours.

Il a été démontré que l'expression des gènes se produisait sous forme d'impulsion de production d'ARNm et que la cinétique d'impulsion, comme le temps d'activation, l'amplitude, la durée ou le nombre d'impulsions, est connue pour refléter la régulation se produisant au locus d'ADN. Grâce aux méthodes d'imagerie basées sur les protéines de capsides virales et leurs séquences de liaison de hautes affinités (comme les systèmes PP7 et

le MS2), l'impulsion des gènes peut être visualisée et quantifiées dans des cellules vivantes uniques. Dans cette étude, nous avons construit des souches rapporteuses PP7 pour des gènes modèles de la réponse au stress osmotique et avons acquis des données dynamiques sur la production d'ARNm par des promoteurs induits lors d'un choc osmotique. Nous démontrons que l'initiation de la transcription se produit préférentiellement pendant une activité ascendante ou stable de Hog1, mettant en évidence un équilibre entre les voies d'activation et d'inhibition agissant directement sur chaque locus. Nous montrons que le temps d'initiation de la transcription du gène induit par HOG dépend de l'environnement de la chromatine du locus, avec une corrélation positive entre le niveau basal d'expression et la fraction de cellules transcrivant après un choc osmotique. Nous démontrons que la dépendance au facteur de transcription est dictée par la nécessité d'un remodelage de la chromatine, qui se traduit par une initiation de la transcription retardée, un recrutement plus faible de l'ARN PolII et une fenêtre de transcription plus courte. Enfin, nous montrons que la terminaison de la transcription est dictée pas une combinaison de l'activité de Hog1 et de la régulation par le locus.

Dans cette étude, nous rapportons également le développement de nouveaux outils ; pour une manipulation génétique robustes des levures à bourgeonnement et pour l'imagerie simultanée de la transcription et traduction au niveau d'un allèle unique, qui permet d'étudier la régulation post-transcriptionelle et la traduction locale d'ARNm.

Résumé pour le grand public

Toute cellule vivante, qu'elle fasse partie d'un organisme multicellulaire ou non, doit être capable de percevoir son environnement et ses changements de façon à s'assurer de sa survie ou de celle de l'organisme dont elle fait partie. Pour ce faire, les cellules ont à leur surface des récepteurs pour différentes molécules de l'environnement, que cela soit des signaux positifs, tels que des nutriments ou des phéromones, ou des signaux négatifs, tels qu'une limitation en nutriments ou des facteurs de stress. Dans toutes les cellules avec noyaux, appelées cellules eucaryotes, qui constituent les animaux, les plantes ou les champignons, les signaux permettant de décoder l'environnement sont intégrés sous forme d'information cellulaire par des voies de signalisation très conservées. Celles-ci sont responsables de la réponse cellulaire à ces signaux, qui doit être spécifique au stimulus détecté et hautement régulée en temps et intensité. La plupart du temps, ces cascades finissent dans le noyau des cellules où elles vont moduler l'expression des gènes. Il est donc important que l'information soit transportée correctement du début à la fin, afin d'initier la bonne réponse sur le court terme et l'adaptation sur le long terme.

Pendant ma thèse, je me suis particulièrement intéressée à la cascade de signalisation responsable de la détection de l'augmentation de l'osmolarité de l'environnement dans les levures dites à bourgeon ou levures du boulanger. Ce champignon unicellulaire est un modèle pour l'étude des voies de signalisation et nombre des trouvailles faites dans cet organisme ont pu être démontrées plus tard dans des cellules humaines, grâce notamment à une forte conservation des séquences d'ADN et des mécanismes d'action. Le but de ma thèse était de décrire comment les gènes de réponse au stress osmotique sont exprimé et quelles sont les composantes qui déterminent leur profil d'expression. Pour ceci, j'ai fabriqué des souches produisant un signal fluorescent en réponse au stress osmotique afin de pouvoir mesurer en temps réel leur réponse et au niveau de chaque cellule unique, grâce à un microscope à fluorescence. Grâce à mes données, j'ai pu démontrer que la réponse au niveau des gènes n'est pas bien corrélée avec la signalisation de la cellule et que la façon dont l'ADN est compacté dans le noyau et spécifique à chaque gène va grandement influencer sa dynamique d'expression.

Table of content

List of Figures

List of Tables

Chapter 1 : General introduction

Chapter 1: General introduction

1.1 Environment sensing systems

A crucial ability of any cellular life is sensing its environment and integrating its changes into intracellular information. Whether one is a unicellular or multicellular organism, misreading of extracellular environment can be deleterious to the cell or to the whole organism. Extracellular signals can be growth-promoting, such as nutrients, growth factors and pheromones, or growth-inhibitory, such as oxidative stress, low nutrients or high confluence. Combining the extracellular information with intracellular state, such as the cell-cycle stage, will determine each cell fate decision; from cell death to cell proliferation, cell-cell fusion or differentiation, and cell quiescence to migration. An additional complexity of signal integration, is the necessity for the appropriate response to be initiated at the right time. As an example, two cells expressing pheromones will undergo cell-cell fusion. Although sensing might have occurred for a certain amount of time, cell fusion can only occur at a specific timepoint to avoid cell bursting and signaling should be block to avoid multiple events of cell-cell fusion and aberrant genetic content [1]. Similarly, activation of a response is as important as its deactivation. Taking the stress response as an example, triggering the response is necessary for immediate cell survival. However, stress signaling often leads to cell-cycle arrest [2]. Thus, in order to proliferate, the response has to be downregulated in order to allow cell-cycle progression to resume.

1.1.1 Sensing in prokaryotic and archaeal cells

In prokaryotic cells, environment sensing is performed through a two-component or phosphorelay system. The two components generally consist in a membrane bound sensor histidine kinase that autophosphorylates itself in presence of a stimulus, which then activates a cytosolic transcription factor, that regulates gene expression accordingly [3]. The phosphorelay system thus resembles the eukaryotic MAPK signaling system and consists in a sensor that gets phosphorylated in response to a stimulus and relays its phosphoryl group to a second protein, which in turn activates a Transcription Factor (TF) [3]. Archeal cells have supposedly acquired a two-component-like system from bacteria through horizontal gene transfer [4]. However, unlike bacterial cells, they would only possess a sensor protein on which the input is computed, but lack the TF counterpart, which is responsible for gene expression [5]. Therefore, regulation seem to happen only at the level of post-translational modifications through phosphorylation, rather than by regulating gene transcription [5]. Unfortunately, crucial data are still missing on archeal sensing systems.

1.1.2 Sensing in eukaryotic cells

In all eukaryotic cells (from fungi to animals and plants), Mitogen-Activated Protein Kinases (MAPK) are implicated in environment sensing. They consist in a three-core module of kinase nodes often activated by sensor at the Plasma Membrane (PM) and leading to the activation of the main effector of the pathway and last node of the cascade, the MAP kinase. Upon activation, the latter will phosphorylate other substrates in the cytoplasm but will also eventually transiently relocate into the host cell nucleus to modulate gene expression.

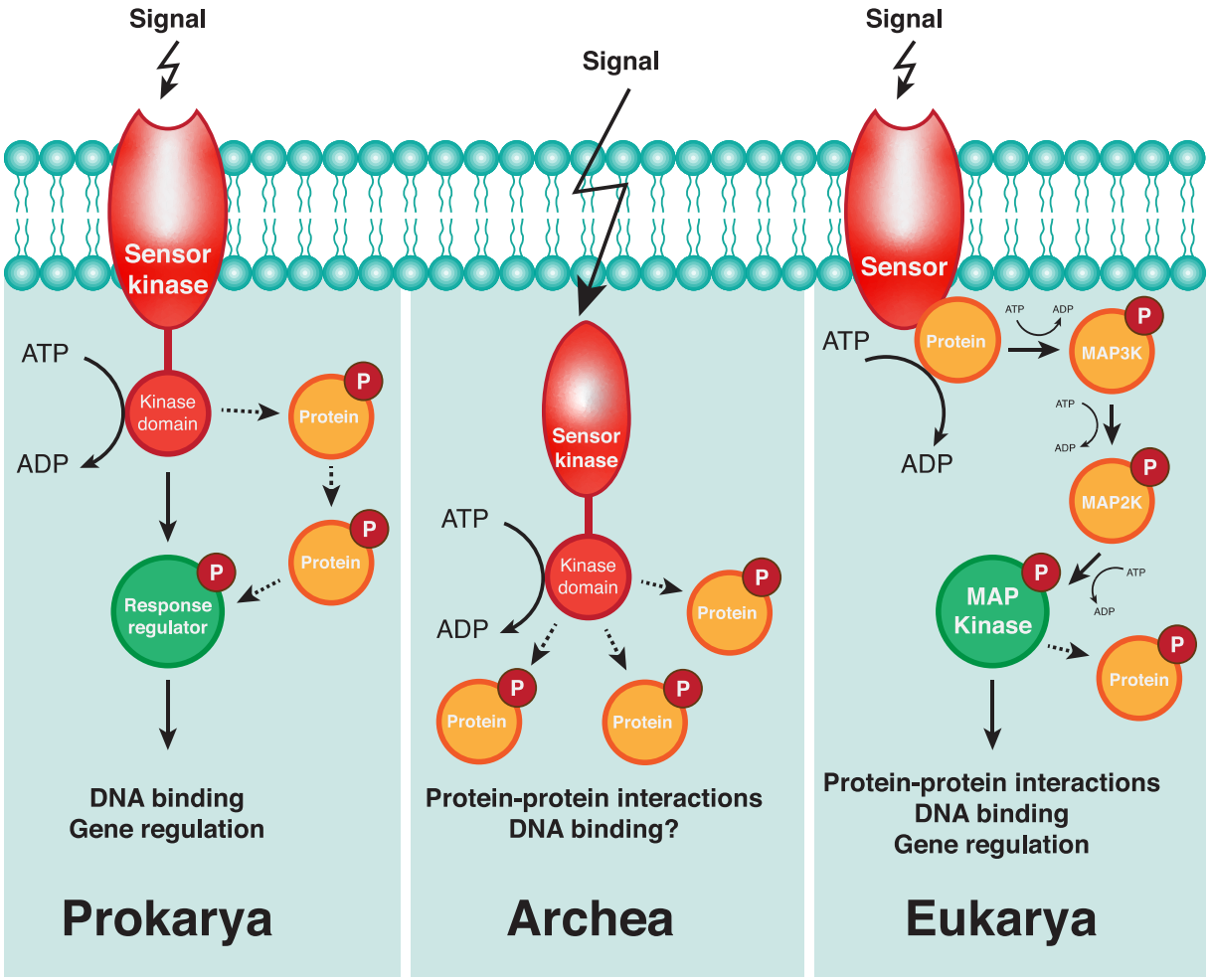


Figure 1: Extracellular sensing systems.

Bacteria sense their environment through a two-component system composed of a plasma membrane sensor kinase and a response regulator, acting as a transcription factor, that will modulate gene expression according to the signal sensed. Communication between this bipartite system is performed through protein-protein interactions and phosphorylation events. Archeal sensing systems are less known but seem to have inherited a prokaryotic-like system. In eukaryotic cells, extracellular signals are sensed at the plasma membrane by dedicated sensors, which will transmit the signal internally to a protein or protein complex, which will activate the MAPK cascade. Eukaryotic signaling pathways are thus composed of a core three-component system, with a MAP kinase kinase kinase (MAP3K), a MAP kinase kinase (MAP2K) and a MAP kinase (MAPK). This latter undergoes cytoplasmic protein-protein interactions and in some cases, relocalizes into the host cell nuclei to trigger signal specific gene expression to modulate cell fate.

1.2 MAPK signaling cascades

In the early 1980, a highly abundant 42 kilodalton (kDa) tyrosine-phosphorylated protein was isolated from insulin-treated mammalian cells [6]. A variety of extracellular signals were shown to lead to its phosphorylation and thus activation. It was the first identified Mitogen-Activated Protein Kinase (MAPK), the Extracelullar signal Regulated Kinase 1 (ERK1) [7]. Since then, many more MAPK pathways proteins were identified. Interestingly, the organism complexity seems to be reflected in the number of MAPK with a higher number in higher eukaryotes, like in humans, where 13 MAPK proteins are found and 20 in an *Arabidopsis thaliana* plants [8, 9]. The unicellular fungus *Saccharomyces cerevisiae* has five MAPKs, four MAP2Ks and five MAP3Ks [10].

1.2.1 MAPK conservation

Alignment of mammalian ERK sequences led to the discovery of *Saccharomyces cerevisiae* Fus3, Kss1 and Mpk1 MAPKs, highlighting a high degree of sequence conservation from budding yeasts to human MAPK proteins [11]. This high sequence conservation was additionally tested for function conservation by expressing mammalian MAPK into budding yeast cells. The results showed a functional activity of the exogenous proteins, that could even complement yeast MAPK mutant sensitivity [12-14]. Thanks to this high degree of conservation, MAPK studies have gain large insight into their functions and regulatory mechanisms by studying the

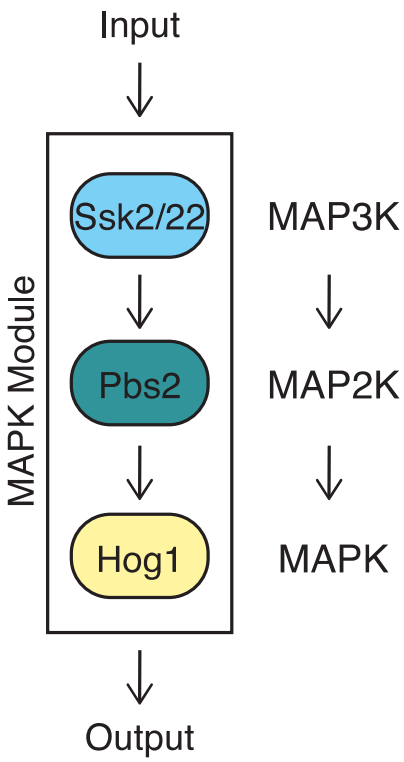


Figure 2: MAPK modules are three-components signaling cascades.
Eukaryotic MAPK pathways have a conserved structure based on a three-component module, composed of a MAP Kinase Kinase Kinase (MAP3K), a MAP Kinase Kinase (MAP2K) and a MAP Kinase. Each MAPK signaling cascade will be activated by a specific input, a phosphorelay between the module nodes and trigger the according output, defining the pathway specificity.

MAPK of low complexity organisms, especially yeasts. In particular, *Saccharomyces cerevisiae* has been extensively used as a model to study the structure, function and regulation of MAPK signaling cascades [15]. Indeed, thanks to being the first fully sequenced eukaryotic genome [16] and with the development of genetic manipulation tools [17], studies using *S. cerevisiae* as a model have thrived since the early nineties.

1.2.2 MAPK signaling cascades

A key finding that has emerged from the study of MAPK in budding yeasts is the three-component modules architecture of MAPK cascades, composed of a MAP3K, a MAP2K and a MAPK (Figure 2). The first component of the cascade is the MAP3K, a serine/threonine kinase activated either by the phosphorylation from a MAP4K or by the interaction with a GTP-binding protein from the Ras or Rho family (yeasts), responsible for the coupling of the external environment to the intracellular signaling pathway [18]. Once activated, the MAP3K phosphorylates the next node of the phosphorelay, the MAP2K. This dual specificity kinase, acting both as a tyrosine and serine/threonine kinase [19], is responsible for the dual phosphorylation of the Thr-X-Thyr motif in the activation loop of the key component of the pathway, the MAP kinase. MAPK are proline directed serine/threonine kinases that have as substrates transcription factors, which they will phosphorylate at S/T-P motifs. Since this motif is quite common among proteins, increased selectivity is performed through a specific “lock-and-key” based three-dimensional interaction domain upstream of the MAPK phosphorylation site and termed “docking groove/D-motif”, with a short linear sequence on the substrate called “docking site/sequence” [20, 21].

1.3 Budding yeast MAPK pathways

Saccharomyces cerevisiae possesses five MAPK pathways: the pheromone response, the filamentous growth, the high osmolarity glycerol, the cell wall integrity and the spore wall assembly pathways [10, 22]. Only the first four are active in haploid yeast cells and are therefore the most studied (Figure 3). Specific activation of each pathway relies on sensors at the plasma membrane [22]. These sensors, which mainly consist in transmembrane receptors, are essential to detect the input signal and transmit it to the intracellular signaling cascade, in order to trigger the appropriate output. All MAPK pathways activation lead to changes in the gene expression program, by the induction of input-specific genes.

In addition to a similar core module architecture, three of the four MAPK pathways have Ste11 as MAP3K (Figure 3). This shared component is itself phosphorylated by the PAK kinase Ste20, therefore the MAP4K of these pathways. In addition to Ste11, the MAPKs Ste7 and Kss1 are shared between the mating and filamentous growth pathways. This raises the question on how signal fidelity is achieved among these pathways? The main hypothesis relies on pathway insulation via subcellular localization, whereby the components of a pathway are brought together by the means of plasma membrane anchored scaffold proteins, thereby preventing their interaction with components from other pathways [23-25]. As an example, upon pheromone treatment, the mating pathway scaffold protein Ste5 relocates to the plasma membrane and its activation of the pathway can be mimicked by adding a PM targeting domain to the protein [23].

1.3.1 Mating pathway

Haploid yeasts exist in two different mating types: the MATa and MATα. Communication between these two cell types is performed by the means of secreted pheromones that will be sensed by the cells through the pheromone receptors Ste2 (α-factor receptor) and Ste3 (a-factor receptor) present on the cells' surfaces [26, 27]. The intracellular domain of the receptor is bound to a G protein and binding of pheromone leads to its dissociation [28, 29], which in turn activates the downstream MAPK cascade. The mating pathway [15] is composed of Ste20 (MAP4K), Ste11 (MAP3K), Ste7 (MAP2K) and finally the MAPKs Fus3 and Kss1 (Figure 3), the yeast orthologs of the mammalian ERK1 and ERK2 [30]. Fus3 is responsible for the majority of the outputs from the cascade, which are cell-cycle arrest [31], polarized growth (shmoo formation) [32] and the fusion to an opposite mating type partner cell [33]. Downregulation of the pathway is performed through the internalization and degradation of the receptors, the degradation of pheromones and the resuming of cell-cycle.

1.3.2 Invasive growth pathway

When haploid cells encounter glucose-limiting conditions, they change their budding pattern from axial to a unipolar one, leading to a chain of cells and altered colony shape [34]. Aside from this morphological change, the cell-cell and cell-substrate adhesion is increased, which can cause the cell to penetrate surfaces [34]. This mode of proliferation, called "pseudohyphal growth", is triggered by the filamentous growth pathway, composed of Ste11, Ste7 and Kss1 (Figure 3) [22, 35, 36]. Although being shared between the mating and filamentous growth pathways, Kss1 is dispensable for mating but decisive to filamentous growth. In addition to the MAPK pathway, two other inputs from two different glucose-sensing pathways are triggered by low glucose levels, the cyclic adenosine monophosphate (cAMP)-dependent

protein kinase (PKA) pathway, that will be activated by the Ras2 protein [36, 37] and the Snf1 protein kinase pathway, regulating repressor proteins at filamentous genomic loci [36, 38]. In addition to glucose-limiting conditions, diploid cells will undergo filamentation under nitrogen starvation [34, 36] through the activation of the filamentous pathway but with inputs from the Target of Rapamycin (TOR) pathway, acting on a transcription factor of filamentous genes [36, 39].

1.3.3 Cell Wall Integrity (CWI) pathway

When cells encounter a cell wall stress, like hypotonic conditions, cell-wall drugs, bud growth or pheromone-induced morphogenesis (shmoo formation), they activate the Cell Wall Integrity (CWI) pathway composed of Bck1, Mkk1/Mkk2 and Mpk1 (Figure 3). Indeed, since cell-cell fusion requires shmoo formation, mutant in the CWI can therefore also be defective in the mating pathway [40]. Activation of the CWI occurs through the sensing of the cell's actin cytoskeleton depolymerization, perceived as a cell wall stress. Since actin-based processes are regulated by TORC2 [41], the CWI receives inputs from the TOR pathway as well [22]. The output from the cascade is the homeostasis of the cell wall, through the synthesis or modifications of the cell wall components, like glucan, mannan and chitin [22].

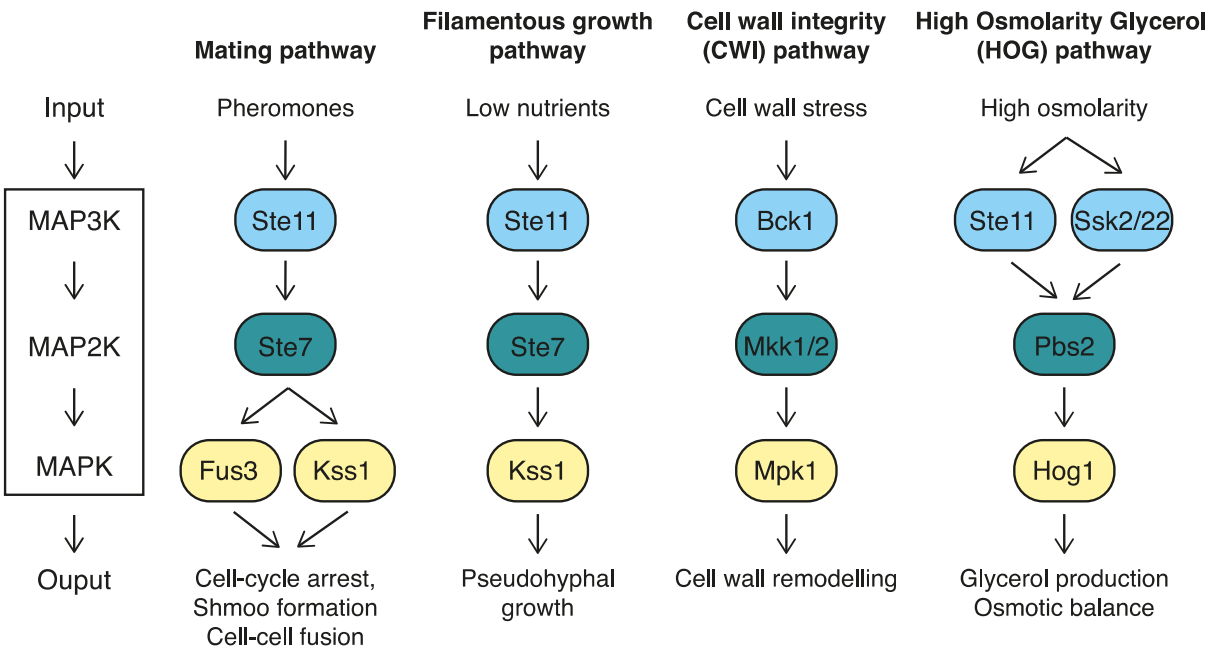


Figure 3: *Saccharomyces cerevisiae* MAPK pathways.

Haploid yeast cells have four MAPK signaling cascades activated in response to environmental stimuli and leading to pathway-specific cellular responses. Pheromone sensing leads to the activation of the mating pathway, which will trigger a cell-cycle arrest, promote the formation of a mating projection called "Shmoo" and mediate cell-cell fusion to produce a zygote with the opposite mating type cell. Starved cells will activate the filamentous growth pathway, which will cause a change in the polarity and growing phenotype called "pseudohyphal growth". This is in order for the cells to escape the harsh environmental conditions and find another carbon source. Cell wall stress, causes by bud growth or hypo-osmotic stress for instance, activates the Cell Wall Integrity (CWI) pathway, which will remodel the cell wall to restore its integrity and tension. A sudden increase in the environment osmolarity leads to the activation of the High Osmolarity Glycerol (HOG) pathway, which will restore cell turgor by producing and accumulating intracellular glycerol.

1.4 High Osmolarity Glycerol (HOG) pathway

Budding yeast cells maintain higher internal osmolyte concentrations than their environment, enabling them to absorb water and create turgor pressure. A sudden increase in the extracellular osmolarity, for example through a drying grape, which is their natural habitat, causes a cell shrinkage due to water loss and a drop in the cell turgor pressure. This threatens essential cellular processes and thus the cell's integrity. To restore osmolyte homeostasis, cells trigger the HOG pathway [42], which will lead to the accumulation and synthesis of osmolytes, mainly glycerol (Figure 3) [43]. Osmotic shock also triggers a cell-cycle arrest, due to the depolymerization of the actin cables necessary for bud formation [44], which recovery depends on HOG activity [45].

1.4.1 Signaling in the HOG pathway

When the HOG pathway was discovered in 1993, it was first described as a prokaryote-like two-component system, due to the characterization of only one of the two sensing branches of the pathway, the Sln1 branch [42, 46], with the second branch originating from Sho1 being discovered only few months later (Figure 4) [47]. The two branches have been shown to be redundant to high stress exposure, however, low stress has been shown to be primarily signaled through the Sln1 branch, highlighting a broader sensitivity range [47].

Sln1 is a two-transmembrane segments protein similar to bacterial two-component systems, with a histidine kinase domain on its intracellular part [42]. Upon osmotic shock, Sln1 autophosphorylates itself and passes on the phosphoryl group to the Ypd1 protein, which in turn transfers it to Ssk1. This prevents its interaction with two of the MAP3K from the pathway

Ssk2 and Ssk22, which are redundant and thus often termed as “Ssk2/22” [22]. The Sho1 branch, through which the filamentous growth pathway is also signaling, leads to the activation of their other shared component Ste11, the third MAP3K of the HOG pathway [22]. Ste11 is bound by the plasma membrane-anchored Pbs2, the MAP2K of the pathway and intersection node of the two branches, preventing its association to Ste7 and ensuring faithful pathway signaling [22, 48]. Pbs2 acts as a crucial scaffold protein because of its binding to the osmosensor Sho1, the MAP3K Ssk2/22, Ste11 and the MAPK Hog1 [48], homologs of the mammalian p38 and JNK [14]. Because of its central role in the HOG pathway, catalytically dead Pbs2 leads to osmosensitivity from the cells [48]. Recruitment of Ste11 to the plasma membrane and to the HOG pathway components is further achieved through the Ste50 protein, which both binds Ste11 and the plasma membrane-anchored Cdc42-Ste20 complex and Opy2 protein [22, 49].

The HOG pathway is thus composed of the MAP3K Ste11 and Ssk2/22, the MAP2K Pbs2 and the MAPK Hog1 (Figure 2 and 3). Phosphorylation of Hog1 by Pbs2 leads to its rapid relocation into the cell nucleus [50], where it induces osmostress genes expression [51]. Indeed, to increase intracellular glycerol concentration, cells diminish their permeability by closing the glycerol export channel Fps1 [52] and trigger the gene expression of glycerol synthesis enzymes [53, 54]. Similarly to *HOG1* deletion, deletion of the *GPD1* gene, encoding a crucial enzyme of the glycerol synthesis, leads to sensitivity to hyperosmotic stress [53]. Interestingly, nuclear relocation of Hog1 is not essential to cell adaptation to mild osmotic pressure, highlighting a higher dependency on its cytoplasmic function than its nuclear one for short-term adaptation [55].

HOG activation is transient and its activity window correspond to the time necessary for the cell to adapt [56]. Downregulation of the HOG pathway occurs through Hog1 phosphatases Ptp2 [42, 57], Ptp3 [58] and Ptc1 [57], which bind the MAPK and dephosphorylate it [58]. This is crucial to limit Hog1 activity both under high osmolarity and basal conditions [42, 46]. Both Ptp2 and Ptp3 genes expression are induced upon osmotic stress, ensuring efficient deactivation of the cascade [58]. Ptp2 appears to play a more important role in the deactivation of the HOG pathway, since its deletion leads to high basal Hog1 activity, which is not the case for Ptp3 deletion [58]. This greater effect of Ptp2 can be explained by the fact that it is enriched in the nucleus, where the main pool of active Hog1 is, while Ptp3 is mainly distributed into the cytoplasm [59]. It was shown that Ptc1 regulates Hog1 basal level, whereas Ptp2 and Ptp3 regulate Hog1 maximal activity [60]. Thus, the double mutants for Ptp2 and Ptp3 shows a hyperactivation of the HOG pathway [58], whereas Ptc1 single mutant displays a higher Hog1 basal level [61]. Interestingly, deletion of the Sln1 osmosensor leads to a constitutively active HOG pathway and is lethal to the cells due to cell-cycle arrest [46].

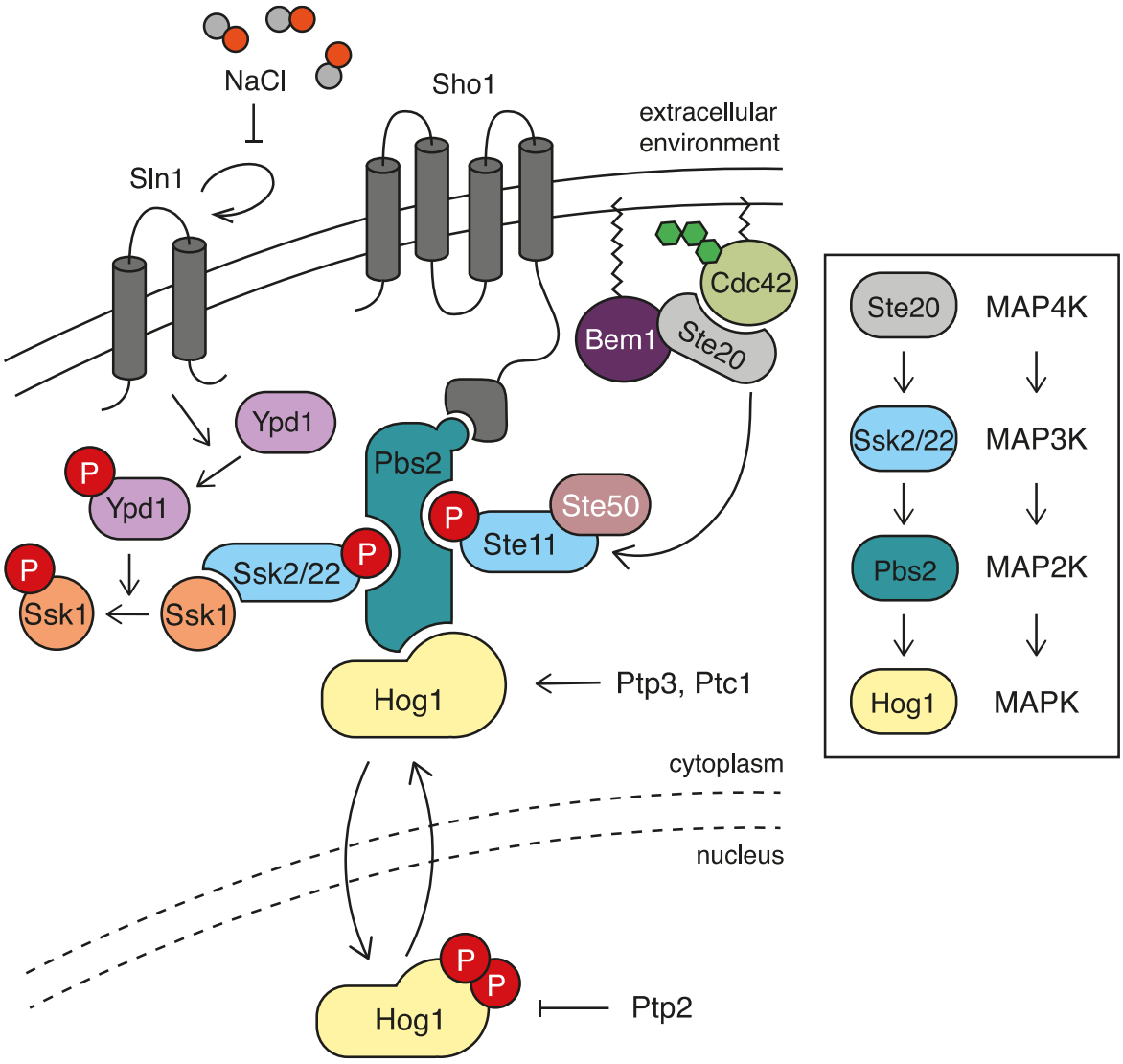


Figure 4: The High Osmolarity Glycerol (HOG) pathway.

A sudden increase in the extracellular osmolarity is sensed by the HOG pathway through two separated branches routing from two different sensors at the plasma membrane, Sln1 and Sho1. The Sln1 branch is a two-component-like system, which autophosphorylates itself in present of osmotic shock and activate the Ypd1 protein, in turn activating the Ssk1 protein, therefore inhibiting its interaction with the Ssk2/22 MAP3K. The later will bind Pbs2, which serves both as a scaffold and as MAP2K and will thus activate the MAPK Hog1 by phosphorylation. The Sho1 branch activates the third MAP3K of the pathway, Ste11, which in turn activates Pbs2 and lead to more Hog1 activation. Inactivation of the pathway is performed by the MAPK phosphatases Ptp2/3 and Ptc1, both in the cytoplasm and in the nucleus of the cell.

1.4.2 HOG TFs regulation of gene expression

A hyperosmotic shock leads to the induction of about five hundred genes, causing a global reallocation of the cellular transcription machinery from housekeeping to osmostress genes [62]. Hog1 is implicated in osmostress genes' expression at all stages of the process, from transcription initiation, elongation, mRNA export and translation [63]. However, because the MAPK lacks DNA binding capacity, modulation of transcription at gene regulatory elements is mediated through the interaction with Transcription Factors (TFs) bound to HOG regulated genes. Artificial tethering of Hog1 to DNA has been shown to be sufficient to induced transcription in an osmostress-dependent manner, highlighting a crucial requirement of TFs for recruiting Hog1 to genomic locus and transcriptional activation [64]. There are six TFs regulating osmostress genes: Hot1 [65, 66], Msn1 [65], Msn2 [66], Msn4 [66], Smp1 [67] and Sko1 [68], which operate through two distinct regulatory mechanisms (Figure 5).

Smp1 is a transcriptional activator that required Hog1 phosphorylation to fulfill its gene activator function [67]. On the contrary, Hot1 TF was shown to be phosphorylated by Hog1, however, this is not required to serve its activator function [65]. Msn2 and Msn4 are redundant TFs regulated by the general stress response pathway [66], which are activated upon any type of stress, including osmotic stress, and bind specific DNA sequences called STress Response Elements (STRE) [69]. Interestingly, these two TFs show a decrease in their activation of gene expression in cells lacking Hog1, highlighting a regulation from the HOG pathway on another signaling cascade [66]. Msn1 is also controlled by the general stress response pathway and is structurally similar to Hot1 [65]. The last TF Sko1 is a particular case since it is a native repressor of transcription bound to DNA and to the general Ssn6-Tup1 repressor complex under basal conditions [70]. Phosphorylation by Hog1 converts it into an

activator, although it remains bound to the Tup1 repressor [71]. Because Sko1 down-regulates a large subset of osmostress genes, *SSN6* deleted cells show an increased tolerance toward salt stress [72]. Since osmostress genes differ in their binding sites' presences and numbers for each of the TFs, a great variability is observed in their gene expression patterns [73].

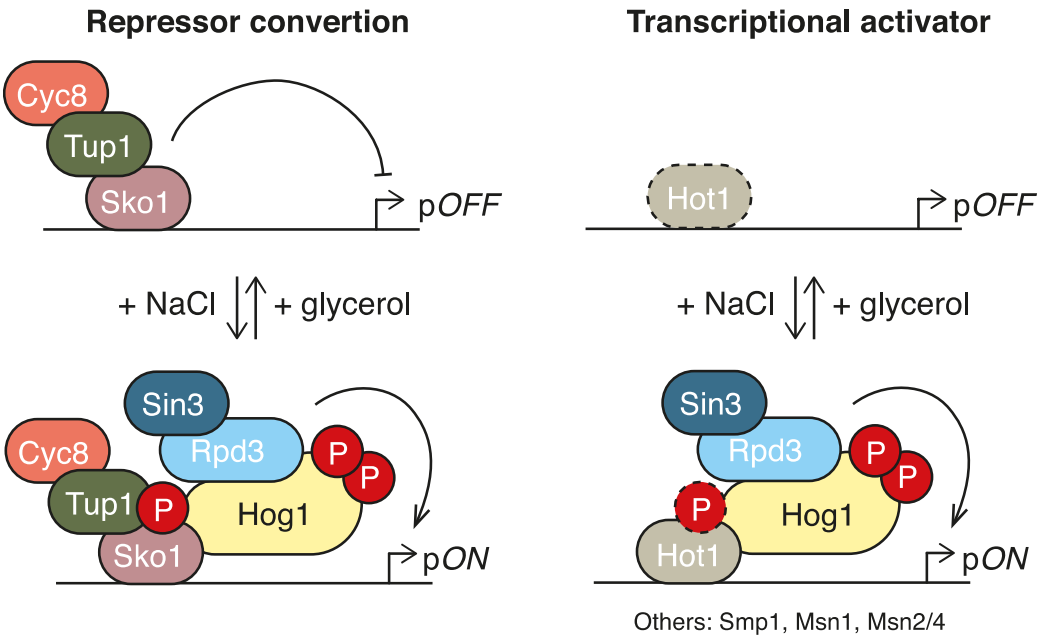


Figure 5: Osmostress genes activation by HOG transcription factors.
Osmostress genes transcription is triggered by specific Transcription Factors (TFs) with two different general modes of action: the repressor to activator conversion or transcriptional activator. Sko1 transcription factor is a native repressor of gene expression bound to the Cyc8-Tup1 repressor complex. Phosphorylation by Hog1 converts it into an activator, which will recruit more Hog1 and the chromatin remodeling and modifying complexes, although its interaction with the repressor complex is not abolished and is even necessary for its full gene activation. Hot1 is a transcriptional activator, that is supposedly constitutively bound to osmostress genes' regulatory sequences (dashed protein). It gets phosphorylated by Hog1, however, this is not essential to fulfill its activator function (dashed phosphoryl group) and can thus trigger gene expression in both methylation states.

1.4.3 Chromatin regulation of HOG genes

Under non-stressful conditions, osmoresponsive genes' regulatory sequences are masked from the transcriptional machinery through a compact chromatin environment [74]. Hog1 activity at osmostress-induced genes is required for the recruitment of chromatin remodeling and modifying complexes to enable access from the transcriptional machinery and subsequent transcription [63]. Because of this requirement for chromatin remodeling to express osmotic genes, chromatin remodeling mutants display sensitivity towards osmotic stress [74, 75]. Higher salt stress leads to higher and longer periods of Hog1 activity, enabling longer nucleosomes evictions from osmostress genes, which is not occurring at non-targeted genomic locations [76]. On the contrary, low salt stress only induces partial chromatin remodeling when measured by population average measurements [63]. This is translated into a bimodality at the single-cell level, with transcribing cells that overcome the Hog1 activity threshold necessary to remodel chromatin and non-transcribing cells that didn't [77, 78].

To remodel chromatin, Hog1 recruits several chromatin remodeling and modifying complexes to osmotic stress loci [74], including Remodel the Structure of Chromatin (RSC), the Swr1 complex, the Spt-Ada-Gcn5-acetyl transferase (SAGA) and the Rpd3 histone deacetylase. Surprisingly, histone deacetylation has been attributed to gene silencing. However, cells lacking the Rpd3 histone deacetylase presence and activity are osmosensitive and show an impaired induction of stress genes, highlighting a positive regulation from the enzyme [79]. In presence of another chromatin mark, H3K4 methylation, RSC remodels the chromatin to form a Nucleosome Depleted Region (NDR) at the promoter region, enabling transcriptional activators binding [74, 80]. On the contrary, unmethylated chromatin will be modified by the Swr1 complex, which catalyzes the exchange of Histone 2A (H2A) against the H2AZ variant,

encoded by the *HTZ1* gene, which has a decreased affinity for DNA [81]. Once chromatin is more permissive, SAGA recruit its co-factor RNA polymerase II and assembles the Pre-Initiation Complex (PIC) [75], which precedes osmostress gene transcription activation.

RSC and Swr1 thus act in parallel and will remodel chromatin depending on its methylation state. Indeed, the COMPASS subunit Set1 histone methyltransferase is recruited to osmostress loci by actively transcribing RNA pol II to methylate Histone 3 Lysine 4 on histones' tails, not only on promoters but on the gene bodies and the 3'UTRs of osmostress genes [81]. H3K4 methylation has been previously defined as an activating mark or to transcriptional memory [82], however, deletion of Set1 rescues partially RSC loss of function in HOG gene transcription. Thus, HOG-induced genes are repressed both by chromatin compaction and methylation states [81].

Like the negative regulator Set1, the INO80 and Asf1/Rtt109 complexes are recruited by the RNA pol II and act in parallel to re-establish chromatin during osmostress adaption and after acute transcription of other genes as well [83]. Deletion of the Arp8 subunit of INO80 or Asf1 leads to a delayed chromatin closure after osmotic shock, causing an extended transcriptional activity [84], and deletion of both leads to an additive phenotype [83].

1.4.4 Osmostress genes as gene expression models

Osmostress genes are well-suited to gene expression studies. Indeed, since they are repressed under basal conditions and highly induced, their signal-to-noise ratio is nicely quantifiable. Second, since they are only transiently induced, one can monitor the entire process from transcriptional activation to termination. Third, since they were a model for MAPK signaling cascade regulation and because HOG genes were a model from chromatin remodeling regulated genes, their regulatory network is well established and candidates for biological/genetic studies are easily identifiable. Finally, since they are in the budding yeast, a fully sequenced, easily handled and genetically modifiable organism, one is not limited on the technical handling side.

1.5 Modes of gene expression

Gene expression is the process of transforming the information encoded in DNA into a final product, generally proteins. Since DNA is at the source of gene expression, it is a target of choice to regulate the genes' products. Promoters are thus major regulatory platforms, where inputs are computed into transcriptional programs. Although gene expression had been studied for a long time, the discovery of alternatives modes of transcription, thanks to the arise of single-cell reporter assays and fluorescent proteins, has opened new fields of study and challenged established gene expression models [85, 86].

1.5.1 One-state model

Since proteins are cellular effectors affecting cellular fate, their presence and abundance must be tightly regulated and should show only little variation from cell-to-cell in a clonal cell population. Therefore, the model to describe gene expression that was first described depicted an active ground state of gene promoters, going through stochastic events of transcription (Figure 6). This "one-state" or "constitutive" model implies constitutive expression from gene promoters, leading to a distribution of the number of messenger RNA (mRNA) in the cells among a cell population that can be fitted by a Poison distribution (Figure 6). Although constitutive gene expression has been observed for some housekeeping genes in the budding yeast [85], it does not seem to be the predominant mode of expression.

1.5.2 Two-state model

With the advent of single-cell reporter assays and even more with single-molecule visualization assays (detailed in the next chapter), a never before seen variability was observed for numerous genes. The number of mRNA molecules per cell within a clonal cell population could no longer be explained by a Poisson distribution but by an inverted exponential, with cells displaying a high number of mRNA and cells were none were detected (Figure 6). This “two-state” or “bursty” model implied that promoters were switching from an ON state, characterized by a high production of mRNA within a short time window, to an OFF state, where no transcript was transcribed (Figure 6).

Bursty mRNA production has been consistently observed from bacteria [87], to yeasts [88] and drosophila embryos [89]. Since this mode of expression would in theory favor a high cell-to-cell variability, it was first proposed to have as purpose to create high phenotypic variability, to increase cell fitness or ensure survival [90, 91]. However, this mode of expression was repeatedly observed from housekeeping [88], to stress response [92] and developmental genes [89], highlighting a conserved mode of mRNA production. In addition, bursting parameters (like burst size and frequency) have been monitored for environmental responsive genes and shown to be actively regulated depending on the stimuli [93], highlighting a both conserved and controlled mechanism. Interestingly, bursting kinetics have been presented to be gene-specific rather than following a general pattern [85, 88, 94]. Note that high-mRNA heterogeneity may not be conserved at the protein level, due to buffering mechanisms [95].

1.5.3 Three-state model

The two-state model implies that promoter activity is not affected by previous transcriptional events, with all transcriptional activations being independent. However, it has been observed for certain genes, that a memory of previous transcriptional activity was encoded at the locus and that a period following transcriptional termination was refractory to any new initiation event. This additional promoter state thus defined the three-state model, with an ON, an OFF and a refractory state, with gene-specific transition rates between these states [94].

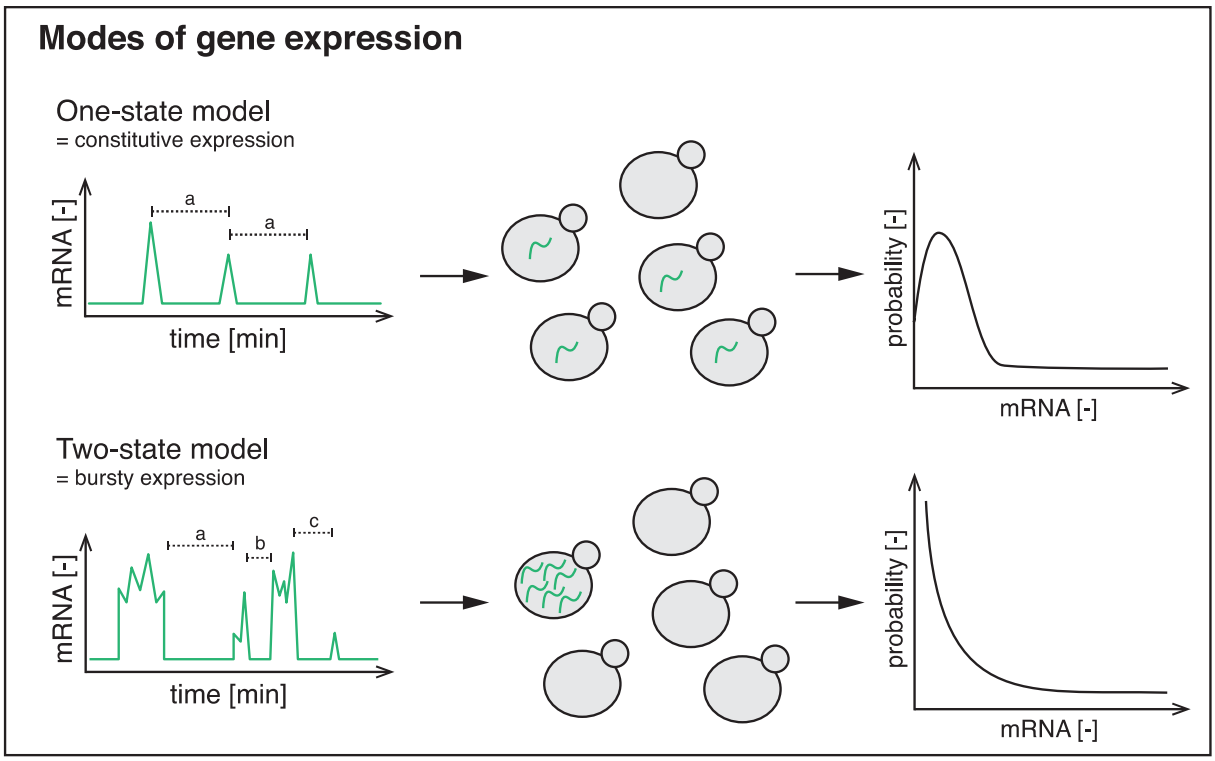


Figure 6: Modes of gene expression.

The ground state of promoter was originally thought to be active and always transcriptionally potent, thus undergoing stochastic events of transcription with a constant probability over time ("a"). This mode of expression leads to a low number of mRNA molecules per cell and thus low cell-to-cell variability in the molecule number, which can then be explained by a Poisson distribution. However, although some promoters were clearly shown to be following this one-state model, single-cell and single-molecule analyses have revealed a greater variability among a cell population, which could no longer be explained by a Poisson distribution, but rather by an inverted exponential. This revealed a new mode of transcription, whereby promoters display distinct periods of activity, where numerous mRNA produced, followed by clear periods of inactivity. These uncoordinated periods ("a", "b" and "c") of transcriptional activity defined the bursty nature of promoters.

1.6 Mechanisms of genes bursting

Although gene bursting has been observed in numerous studies, the mechanisms behind transcriptional bursting remain unclear. Gene bursting can be differentiated as “extrinsic” and “intrinsic” bursting; “extrinsic or transmitted bursting” being the induction of bursts through changes in upstream signaling and “intrinsic or spontaneous bursting”, being the results of promoter activity changes [96]. However, the distinction between these two categories may not be as straightforward, since promoter ground state in eukaryotic cells is to be inactive, activation requires transcriptional activator binding. Therefore, uncoupling of promoter innate bursting activity versus TF binding dynamics can be challenging.

1.6.1 Sources of extrinsic/transmitted bursting

Pulse of signaling

Transcriptional bursting is defined by cycles of transcription interspaced by period of promoter inactivity. If the transcription factor responsible for the transcription undergoes itself cycles of activity, it can lead to transcriptional bursting. The period between bursts of transcription will then depend on the TF ON/OFF switching rates. That is the case for the Crz1 budding yeast transcription factor, which upon calcium stress will relocate into the cell nucleus with a calcium stress level-dependent periodicity, but constant burst duration, to trigger waves of gene expression [97]. Similarly, budding yeast growth-repressed osmostress genes are transiently induced upon Hog1 MAPK activity, which follows a pulse-like dynamic [51]. Further examples may be found in cell-cycle or circadian regulated genes.

1.6.2 Intrinsic bursting hypotheses

Supercoiling

Studies in bacterial cells suggest that the topology of the DNA at a genomic location would be involved in transcriptional bursting [98]. This hypothesis relies on the fact that when RNA is transcribed from a template DNA, this causes a discrepancy in the torsion level of the DNA helix between the DNA upstream and downstream the transcribing polymerase. DNA torsion-specialized enzymes, namely topoisomerases and gyrases, act then together to release the constraint on the helix as the RNA polymerase transcribes [99]. In the case where this activity would not affect all the genes similarly, for instance due to differential expression level, and of gyrase limiting conditions, this could potentially lead to bursting [100]. Although DNA supercoiling has been shown to be involved in gene regulation in higher eukaryotic cells [101], the timescale of eukaryotic transcriptional bursts together with the evidences from other components being implicated in bursting, make this hypothesis a supposedly less prominent mechanism of eukaryotic transcriptional regulation [102].

Chromatin loops

In higher eukaryotes, this is well established that cis-acting regulatory elements, like enhancers, can influence a distant promoter element through intergenic loops [103, 104]. As an example, it was shown that the transcriptional output of the *sna* promoter in the drosophila embryo is regulated by the *sna* enhancer in a burst frequency-dependent manner, where the stronger enhancer induces a higher bursting frequency from the targeted promoter [105]. In simpler eukaryotic systems, it's mainly intragenic loops that have been observed. These loops are dependent on the transcription machinery and formed by the interaction between the promoter and the terminator of a DNA locus and are called "DNA looping" [106, 107]. This

conformation of the locus would lead to the recycling of transcribing RNA polymerase II from the terminator back to the promoter, improving RNA pol II recruitment rate and accordingly the transcription rate, leading to transcriptional bursts. This has been observed for two very long genes in yeasts [109], but is proposed to be a general phenomenon because of the compact genome of *S. cerevisiae* and the general dependency on the RNA polymerase II and cofactors of gene looping.

Chromatin state

Since eukaryotic genome is well packed into nucleosomes, which require complex activity to be removed, chromatin remodeling has been suggested by numerous studies as a prominent source of gene bursting [110, 111]. In addition to nucleosome position, chromatin marks like acetylation have been shown to modulate specific bursting parameters [112]. Since chromatin remodeling requires the recruitment of chromatin remodeling complexes via transcription factors or activators, uncoupling of both to determine the origin of bursting might be challenging, especially if the transcription factor activity cannot be assessed from its cellular localization, like the Crz1 transcription factor [97]. Although the advent of single-molecule tracking might provide a good opportunity to distinguish between these two contributions [113].

Pol II pause and release

Transcriptional bursting corresponds to actively transcribing RNA polymerases. Upstream of active transcription lies transcription initiation, which encompasses several distinct steps. Indeed, right after having transcribed a few base pairs, Pol II pauses and requires signals to undergo elongation causing stalling [114]. This maturation time of the transcriptional

machinery and subsequent release of polymerases is called “promoter proximal pause-release of Pol II” [115, 116]. Stalled RNA pol II has not yet been observed in yeasts but it has been in metazoans [114], highlighting a postrecruitement regulation of gene expression in higher eukaryotic systems. Through this mechanism, promoters should be able to initiate faster and more efficient transcription by bypassing a number of steps implicated in PIC reassembly [115, 117].

Phase separation

Assemblies of regulatory sequences like enhancers and promoters, or just paired alleles, have been proposed to lead to the local accumulation and retention of transcriptional activators and effectors like transcription factors and RNA polymerases. This local accumulation of transcriptional effectors would cause a liquid-liquid phase separation from the cytosol and lead to a higher probability to initiate transcription from promoters within this macromolecular assembly [118, 119]. This “transcriptional hub” or “phase-separation of the transcriptional machinery” would thus be comparable to a local accumulation of TF bindings sites on a promoter or to gene looping on macromolecular scale; it increases the chance of recruiting recycled RNA pol II and increase transcription rate, but in a trans-acting manner [119]. Indeed, it has been shown that several parts of the transcriptional machinery could lead to phase separation, like the human and yeast RNA polymerase II CTD domain [120]. This model supports a pre-existing one on “transcriptional hot spots” or “transcriptional factories” [121], where it was observed in several cells and organisms that, rather than being homogenously distributed into a cell nuclei, RNA pol II or TF molecules were clustering into distinct nuclear puncta [122-124]. Discrete temporal association of these macromolecular complexes would then lead to gene bursting.

1.7 Monitoring gene expression

Since gene bursting reflects the underlying regulatory mechanisms occurring at the transcriptional level, the ability to measure bursting parameters allows to unveil dynamic information on subsequent transcriptional regulation, which could otherwise not be extracted from population-averaged or static single-cell measurements. Thanks to the development of fluorescent proteins, the field of gene expression and single-cell studies has thrived. However, all methods do not measure the same biological phenomenon and each method has its advantages and drawbacks. In addition, apart from active post-transcriptional regulation, it is still debated whether mRNA and proteins levels correlate, therefore the use of protein-based expression reporter assays may not reflect the true dynamics of mRNA production at the single-allele level. Below we describe the relevant gene expression systems that were developed over the years, described in order of increasing dynamicity in their transcriptional readouts.

1.7.1 Static measurements

Population-averaged measurement of transcription enable to retrieve both qualitative and quantitative measurement of mRNA production.

Northern blot

The oldest gene expression assay is the Norther blot [125]. With a gel-separated bulk RNA extraction and transcript-specific probes, as well as an internal control transcript, this method gives snapshot quantitative relative measurement of population transcriptional profile.

Real-Time quantitative Polymerase Chain Reaction (RT-qPCR)

RT-qPCR is a method more sensitive than Northern blot to low abundance transcripts and requires less amount of RNA extract [126]. This technique measures the amount of DNA reverse amplified from extracted RNA templates. Thanks to the use of fluorescent dyes, it can inform on the quantity of initial template, based on standard references.

Both of these techniques rely on population-averaged measurements and discrete timepoints, which does not allow single-cell analysis, nor time-course experiments. In addition, they both depend on the quality of the RNA extraction, the target transcript abundance and on the specificity of a probe from its target transcript, which requires some steps of optimization for every new transcript to be quantified.

Single-cell RNA sequencing

Based on the isolation of single cells from a population or a tissue, RNA extraction followed by sequencing enables to capture the instantaneous transcriptome of a single cell. This technique, called single-cell RNA sequencing, bypasses population averaging and enables to gain insight into transcript variability at the single-cell level [127]. The major advantage of this assay is the amount of information that can be extracted in a single experiment. Indeed, since it relies on sequencing and transcript non-specific probes for the amplification, one can capture the entire single-cell transcriptional profile at a glimpse. However, because it starts from a small amount of material that is massively amplified, internal controls are crucial to rely on the output information gathered, especially with transcript of low relative abundance. In addition, although this technique enables to reach single-cell level, these are still snapshots measurement, that do not enable dynamic data acquisition. Finally, due to cell lysis, any information on the spatial localization of the sequenced RNA is lost.

single-molecule Fluorescence In Situ Hybridization (smFISH)

Base on sequence complementarity, Fluorescent In Situ Hybridization (FISH) can be used to labelled DNA and RNA, thanks to small DNA probes labelled with fluorescent dyes [128, 129]. This technique requires cell fixation, permeabilization, labelling, imaging and image quantification. To improve the signal-to-noise ratio, an improved version of FISH, single-molecule FISH (smFISH), has been developed [130]. Thanks to probes multiplication on the targeted transcript, the signal-to-noise ratio is greatly improved and enables low abundance or single cytoplasmic transcript to be visualized. Although it enables single-cell and single-transcript measurement, this method relies on fixed cells and thus does not enable real-time imaging and thus dynamic measurement of gene expression. However, unlike the previously presented assays, this technique enables to retrieve both temporal and spatial information on the transcript simultaneously, which makes it nowadays still a widely used method. This method was further improve ten years ago, by multiplying shorter single-labelled probes, going from five 50-nucleotides-long probes labelled with five fluorophores to forty-eight 20-nucleotides-long probes labelled with a single fluorophore [131].

1.7.2 Dynamic measurements

Fluorescent reporter assays

The Green Fluorescent Protein (GFP), isolated from the jellyfish *Aequorea victoria* [132], was the first fluorescent protein cloned and expressed in different cell types. It has been extensively used then as a gene expression reporter since [133]. Indeed, because it only requires oxygen to be fully functional, it makes it a low-constraint reporter assay to use [134]. From this initial GFP, the family tree has largely expended and gave rise not only to GFP variants, but Yellow (YFP), Red (RFP) and Cyan (CFP) Fluorescent Proteins (FPs). The major

disadvantage of FPs are their maturation times; indeed, the fastest maturating FP are the Venus (40min) [135] and the recently developed superfolder GFP (sfGFP) (6min) [136]. This drawback is also a caveat for split-FP, which were engineered to measure dynamic molecular interactions or to improve the signal-to-noise ratio. The major advantage of FP, but which can also be a drawback, is their stability. Indeed, the half-life of FP is typically in the order of hours, which permits long periods of acquisition. However, they are not suited for genes of short cell-cycle organism or transiently expressed genes, which typically have a smaller half-life than the FP, therefore masking part of the dynamics. To overcome the maturation time issue, studies started to block translation with Cycloheximide and analyze gene expression typically by flow cytometry after few hours, to let the pool of synthetized fluorescent proteins maturate [77]. Unfortunately, although it bypasses the maturation time of FP and still gives single-cell measurement, it removes the ability to track single cells over time.

β- Galactosidase

This bacterial enzyme has been long used as a colorimetric reporter method [137]. The *lacZ* gene used in the reporter assay, encodes for an hydrolase that converts lactose to glucose and galactose [138]. Because this enzyme can also hydrolyze X-gal into a blue-colored product, it provides a visual readout to promoter activity in a broad specimen range [139]. However, It suffers however from a poor sensitivity, small narrow dynamic range and may show some endogenous activity in mammalian cells, it was therefore quickly replaced by more sensitive assays for the quantification of gene expression. It is nowadays mainly used as a qualitative control, for instance during mammalian cell transfection.

Luciferases

A special class of reporter fusion assays relies on bioluminescence instead of fluorescence. These reporter proteins typically encode a luciferase gene, which catalyzes light-emitting reactions [140, 141]. The photon emitted is then used as a readout for gene expression. Luciferase genes have been isolated from different organisms, the most familiar being the firefly *Photinus pyralis* (Fluc) or the sea pansy *Renilla* (Rluc) [142]. In practice, the enzyme is encoded downstream the promoter of interest and luciferin is added to the medium. When the promoter is activated, the luciferase is synthetized and oxidizes luciferin in oxyluciferin, emitting a detectable photon [142]. This enables to bypass the maturation time of fluorescent proteins, and like any other protein, requires only the translation and folding times [143]. Because of the high sensitivity of the system, absolute absence of any light has to be ensured, which make the experimental handling challenging. In addition, since luciferin is used as substrate, it needs to be in excess in the cell environment or freshly perfused during acquisition, which also complicates the use of other inducer or chemicals.

dynamic Protein Synthesis Translocation Reporter (dPSTR)

Since the major drawback of the use of Fluorescent Proteins (FPs) is their maturation time, we developed a maturation-free fluorescent protein-based assay called “dynamic Protein Synthesis Translocation Reporter” or dPSTR [144]. This bipartite system consists in a constitutively-expressed fluorescent protein fused to a synthetic peptide (SynZip, [145]), which has a high affinity for its partner peptide fused to a double Nuclear Localization Sequence (NLS) and a degradation tag (UbiY, [146]) placed under the control of a promoter to be monitored. The constitutively expressed FP is distributed in the whole cell in absence of transcription from the targeted promoter. Upon activation, the monitored promoter will

induce a proportional nuclear accumulation of the fluorescent signal thanks to the strong affinity of the SynZips associating the FP and the NLS [144]. Because it relies on the relocation of an already matured fluorescent protein and since it is reversible thanks to the degradation tag, this system enables to follow gene expression dynamics from induction to deactivation and repression [144]. It is thus a powerful tool to measure single-cell single-promoter expression. However, since it is based on signal relocation, this assay cannot infer on the transcript discrete spatial localization.

All the previously described systems are based on protein expression, which involves translation and folding, and thus may mask, slow down or not reflect the true dynamics of transcription. To get closer to promoter dynamics, mRNA-based reporters have been developed.

RNA aptamers: Spinach

Based on an 80-nucleotides long RNA aptamer called “Spinach”, that becomes green fluorescent upon 3,5-difluoro- 4-hydroxybenzylidene imidazolinone (DFHBI) binding, modified mRNA can be labelled and monitored after an initial pulse of DFHBI [147]. The advantages of this method over the later described phage coat proteins are the low photobleaching, low background fluorescence and smaller DNA sequence integration compared to phage coat proteins’ binding sites [148]. However, in addition to a low brightness and thus low signal-to-noise ratio, the question on the stability of the chemical, its absorption from the cells and the quantities to use to label all mRNA makes this system challenging to standardize and rely on.

1.8 Phage coat proteins

RNA aptamer-protein based systems are used for RNA imaging since 1998, with the development of the MS2 system [149]. These natural RNA-protein interaction systems are found in RNA phages, which form their viral particles inside a host cell by assembling their viral capsid proteins and genomic RNA. Indeed, the coat protein can oligomerize but can also bind the phage RNA genome through specific RNA encoded sequences forming secondary hairpin-like structures. The first-described so-called "MS2 system" derived from this natural interaction consists in a constitutively expressed MS2 protein, isolated from an MS2 bacteriophage and fused to a Fluorescent Protein (FP), and a DNA-encoded array of typically twenty-four MS2 binding sites or MS2 Stem-Loops (MSL), placed under the control of a promoter of interest [149]. Thanks to the multiplicity of MS2 binding sites and to the dual binding of each stem-loop, a labelled transcript will produce a fluorescent signal above the background of free floating MS2-FP upon promoter transcription.

There are four variants of these phage isolated RNA-protein interaction assays: the MS2 just described [149], the PP7 [88], the λN22 [150] and the Qβ [151] systems [152, 153]. They all rely on naturally-occurring RNA-protein interactions, composed of DNA-encoded hairpins and fluorescent proteins. They differ from each other in the sequence of their binding sites, the number of nucleotides per hairpin/stem loop and the number of loops that composed the array [152, 153]. As an example, the two very similar MS2 and PP7 systems only share 15% of sequence identity of their coat proteins [154].

1.8.1 Principle

In absence of transcription from the targeted promoter, the cell only displays the background fluorescence from the constitutively-expressed Phage Coat Protein (PCP) (Figure 7A). However, upon monitored promoter transcriptional activation and PCP binding sites transcription, PCP-FP proteins will bind strongly to these secondary structures and lead to the local accumulation of the PCP-FP protein, forming a bright focus at the Transcription Site (TS) (Figure 7B). The background-corrected intensity at the TS is then proportional to the number of PCP-tagged transcripts and thus to the number of actively transcribing polymerases at the DNA locus. Therefore, this system both gives qualitative, promoter ON or OFF states, and quantitative outputs. Note that this system enables to follow the nascent mRNA production but can also be used to visualize cytoplasm-exported mRNA, especially with the sequestering of the majority of the PCP pool inside the host cell nucleus with the addition of NLS sequences to the PCP-FP fusion protein.

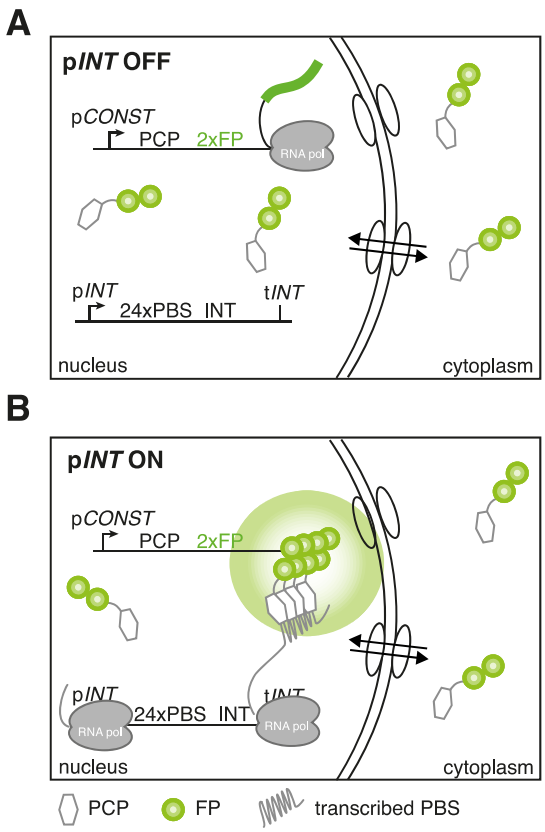


Figure 7: Phage coat proteins- based reporter assays.

(A) A Phage Coat Protein (PCP) fused to a Fluorescent Protein (FP) is constitutively expressed and diffuses freely into the whole cell. An array of 24 Phage coat protein Binding Sites (PBS) are encoded downstream the promoter of a gene of interest (p*INT*). (B) Upon transcription of the PBS by the promoter of the targeted ORF, these sequences form stem-loops that are recognized by the PCP, leading to a local enrichment of the PCP-FP at the site of transcription, which is proportional to the amount of mRNA transcribed at that time and to the number of actively transcribing RNA polymerases II (Pol II).

1.8.2 Advantages

These powerful assays hold a handful of advantages. The most important impact of the phage coat protein reporter assays is the ability to monitor transcription dynamics in live single cells. Indeed, these assays are the closest measurements to promoter mRNA production since they are at the RNA level and thus don't suffer from translational or maturation lag time, and the dynamics observed are not buffered through translational kinetics discrepancies. Second, thanks to the exogeneous origin of both parts of the system, these assays can be used in a wide range of specimen, from bacterial cells [87] to drosophila embryos [119]. Third, since all of the variants of PCP-based assays are orthogonal, they can be used simultaneously to probe more than one promoter transcription dynamic at a time in a single cell, like previous FP-based reporter assays [155], but in real-time. They can thus be used to measure intrinsic and extrinsic noises in gene expression [156] or to access quantitative kinetic data on fundamental processes like transcription elongation or splicing kinetics [88, 157]. Finally, since the readout of the system does not involve an artificial change of localization from the reporter system, this assay enables RNA localizations studies [149].

1.8.3 Drawbacks

A major drawback from this technique is the requirement for a constitutive expression of the coat protein, which causes a high background fluorescence lowering the signal-to-noise ratio. In an attempt to optimize this, split Fluorescent Proteins (sFP) were used [158]. Indeed, thanks to a high affinity of the coat protein for the stem loops and to the orthogonality of the existing systems, the two halves of the sFP can be fused to two different phage coat proteins, like MS2 and PP7. A mixed array of alternated MS2sl and PP7sl will lead to the reunion the two sFP moieties that will form a complete fluorescent protein [159]. However, this system requires

the use of two phage coat systems for the labeling of only one transcript, which limits the experimental design. In addition, the assembled FP requires some maturation time to become fluorescent [159]. A split-FP assay is thus no longer suitable for dynamics measurements.

1.8.4 Limitations

Use of the phage coat proteins-based reporter assay requires to acknowledge the possible limitations from the system, like any other method. Here are presented the reported possible sources of loss-of-function or artefacts from the system.

Imaging conditions

When using phage coat protein-based reporter assays, we monitor actively transcribing RNA polymerases. Therefore, the time window to image a polymerase will be dictated by the transcription rate of the polymerase and the targeted promoter's transcript length. There is thus a close link between the integration site (namely the downstream transcript size) and the imaging conditions. As illustrated in Figure 8, the 5'UTR integration in long Open Reading Frames (ORFs) gives a larger window of imaging time than shorter ORFs, thanks to the transcription time required by RNA polymerases to complete elongation. A longer imaging time enables longer time-course measurements thanks to a decrease time resolution requirement. This 5'UTR integration strategy also increases the probability of having multiple transcribing polymerases on the locus and thus a higher signal-to-noise ratio. Thus, integration into the 5'UTR will increase the imaging window compared to 3'UTR integration and give brighter transcription sites [160]. Since the integration site impacts both the imaging conditions and the outputs, analysis of multiple transcript at their endogenous location will

thus require different imaging settings, which increases the complexity of data comparison.

An alternative integration mean for these examples is discussed in the following section.

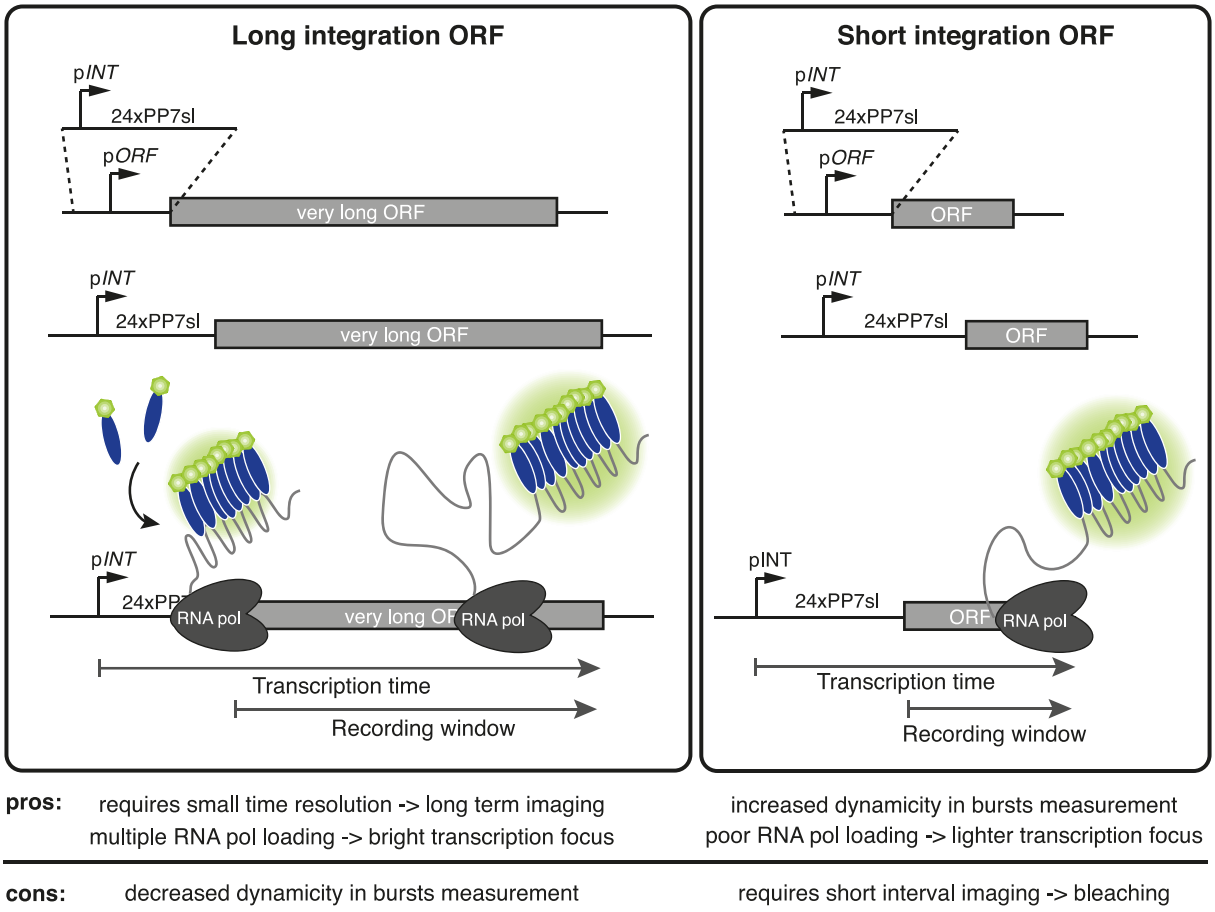


Figure 8: Effect of phage coat protein binding site integration on imaging conditions.

The readout of phage coat protein-based assays, like the PP7 system depicted here, relies on the accumulation of phage coat protein on an ORF, for the signal to be higher than the background expression signal. Therefore, the stem-loops are typically integrated into the 5'UTR of the targeted gene, or with a second copy of the gene promoter of interest into a non-related gene due to gene essentiality, like in the *GLT1* gene in Larson *et al.* 2011. Indeed, since a longer ORF will require a longer time to be fully transcribed by the RNA polymerase, it will have a higher probability of accumulating more RNA polymerases than a shorter ORF. This long ORF will therefore give a brighter signal and bigger half-life signal at the transcription site, which will enable longer time-interval for imaging than a shorter ORF. On the other hand, a short ORF will require a time-resolution during imaging, to capture all transcriptional events and thus will limit the imaging time due to bleaching of the phage coat protein and/or phototoxicity from the cells.

Integration site

In addition to dictating the imaging conditions, integration site selection has a major impact on the biology of the system. Indeed, because the stem-loops' sequence contains multiple translational STOP codons, nothing downstream the binding sites' cassette is translated. Therefore, they cannot be integrated in the 5'UTR of a single copy essential gene. It is thus generally put in the 3'UTR of endogenously tagged ORF. An alternative strategy coming from the mammalian cells studies, is to insert the binding sites into an intron, which does not alter the protein function, nor leads to cytoplasmic dots thanks to splicing of the reporter part co-transcriptionally. Although this alternative does not enable localization studies, it enables the quantification of splicing kinetic [161]. Since budding yeast rarely have introns, Larson and colleagues developed a targeting plasmid that integrates a second copy of the promoter of interest and the PP7sl in the 5'UTR of the non-essential 7kb-long *GLT1* gene [88]. This enables to standardize the measurements between different imaged promoters by exchanging only the cloned promoter sequence and keeping the downstream ORF and thus imaging conditions identical between promoters to be compared.

Phage coat proteins titration

The labelling of newly synthetized transcripts relies on the existing cellular pool of free Phage Coat Proteins (PCPs). The level of PCP expression should thus be high enough not to be depleted or cause partial or no labelling of transcripts, and lower than the signal to have a quantifiable signal-to-noise ratio [162]. As a mean to monitor depletion and ensure labelling of all transcripts, NLS sequences were fused to the fluorescently-tagged coat proteins to restrict them to the cell nucleus and ensure better labelling through local confinement of both interacting partners [162]. Depletion is then arbitrarily assessed based on the disappearance

of the nuclear enrichment. However, this may not be sufficient to ensure the absence of depletion and proper controls have to be made for all promoters tested. With this possible downside, the use of smFISH as a general method to control for labelling and to translate Relative Fluorescence Units (RFU) into numbers of mRNA molecules became quite prominent.

Localization artefacts

Recently, it was shown that the high affinity between the coat proteins and the stem-loops may alter the degradation of the protein-RNA multiplex and lead to localization artefacts, notably under glucose-starved conditions [163, 164]. However, this does not seem to be an ubiquitous observation and to depend greatly both on the experimental conditions, the expression level of the labelled transcript and the stem-loops insertion site. Therefore this should be analyzed on case-by-case, generally by means of smFISH controls [165].

1.8.5 PP7 system

First described by Larson and colleagues in 2009 [88], the PP7 system was purified from *Pseudomonas aeruginosa* bacteriophage PP7. This phage coat protein-based assay has the highest affinity between the coat proteins and its binding sites, with a K_d of 1nM [166] and thus offers the most efficient labelling of transcripts [167]. The binding occurs between one PP7 stem-loop and two PP7 coat proteins, leading to 48 bound PP7 proteins per tagged transcript [167]. The PP7 protein consists in a 128 amino acids polypeptide, forming an N-terminal β-hairpin, a five-stranded antiparallel β-sheet and two α-helices in C-terminus [154]. Recognition of the loop is performed by the beta-sheet [154]. The binding sites of PP7 are hairpins containing a 6 nucleotides (nt) loop and a 8 nt stem, with a purine bulge on its 5' side [166]. Interestingly, removal of the bulge leads to a perfect hairpin but complete loss of

binding activity *in vitro* [166]. Any change in the loop and stem size or sequence leads to higher K_d, depicting the best existing PP7-PP7sl combination possible [166].

PP7ΔFG variant

Since the native function of the PP7 affinity for the PP7sl is to form viral capsids containing the genomic RNA of the bacteriophage, the PP7 protein has the intrinsic property to oligomerize and form particles inside its host cell. To avoid particles formation and imaging artifacts, the residues 67 to 75 (CSTSVCGE) responsible for the capsid assembly were truncated from the original PP7 sequence, leading to a truncated monomeric phage coat protein called "PP7ΔFG", with an intact RNA binding property [154]. A similar approach has been used for the MS2 protein and dedicated system.

1.9 Aim of the study

The general aim of this study was to gain insight on promoter transcription dynamics regulation, from signaling, to chromatin, transcription factors and promoter DNA sequence.

Since osmostress gene transcription is repressed in normal growth-conditions and transiently induced upon a pulse of MAPK activity, this system is ideal to answer these questions. Previous population-averaged biochemical analyses on osmostress gene transcription seemed to highlight a similar gene expression pattern. Although some dynamic single-cell studies were performed, they were mostly realized at the protein level, with the use of long-maturing fluorescent proteins, which causes a delay and a buffering to the true promoter transcription dynamics. Building PP7 reporter strains thus enabled us to assess the diversity of transcriptional profiles and acquire single-cell single-allele data that were not provided before for osmostress-induced genes of the budding yeast.

The highly dynamic data gathered would thus give a partial response on how promoter sequences dictate transcription dynamics, using native and not synthetic sequences, by comparing the different osmostress promoters' transcription dynamics in similar conditions. In addition, combining the PP7 reporter to other available assays enabled us to assess the correlation between signaling, mRNA and proteins, at the single-cell level.

To fulfill our study's aim, we have (1) developed new vectors for the genetic manipulation of budding yeasts, that enabled us to have robust and reliable single-integration of our reporter constructs, (2) studied the biology of the PP7 system to highlight its limitations and (3) optimized the PP7 system to apply it to osmostress genes and (4) improved an *in silico*

quantitative approach to extract bursting kinetics parameters from imaging data. All of these enabled us to reach high confidence single-promoter transcription dynamics data acquired in a semi-automated manner. During these process, new gene expression reporters have been developed and are described in this thesis.

1.10 Thesis chapters presentation

This study is divided into six chapters, comprising the general introduction, three results chapters, the discussion and the conclusion.

The first chapter is on the development of a novel type of yeast shuttling vectors, which enabled us to faithfully and efficiently modify genetically *S. cerevisiae*. The story has been published as a method article in a peer review journal in 2016 [168]. In the second chapter, we coupled the dPSTR and the PP7 system into a unique assay, that enabled us to monitor transcription and translation from a single allele in live single cells. The results presented are currently in a manuscript preparation.

In the third chapter, we used the PP7 system to monitored osmostress genes in a semi-automated manner to decipher the regulatory mechanisms dictating the HOG induced gene promoter transcription dynamics. The results are currently presented in a PDF version of a preprint under review in a peer-review journal as "*Single-particle view of stress-promoters induction dynamics: an interplay between MAPK signaling, chromatin and transcription factors*", Wosika V and Pelet S, bioRxiv 2020. A previous version of the preprint can be found as "*Single-cell analysis of osmostress promoters reveals the dynamics of transcription initiation and shutoff*", Wosika V and Pelet S, bioRxiv 2019.

Chapter 2: Single-Integration Vector and Gene Tagging plasmids development

2.1 Background

In the field of gene expression, reporter assays are highly used and often require the expression of exogenous proteins. To have a homogenous expression of the reporter protein among the single cells, genomic integrations are preferred to extra-genomic plasmids, which copy number and stability are not certain. In the budding yeasts, integration plasmids have been used for decades [169]. The most used vectors, the pRS, take advantages of the high recombination power of yeast, to integrate a plasmid at a genomic location by homology with a non-functional endogenous auxotrophy marker gene [17]. This integration leads to a duplication of the marker cassette leading to a non-functional and a functional allele. Unfortunately, this genome-integrated plasmidic copy of the marker gene can also serve as an integration site, along with the endogenous one, for more integrations of the transformed plasmids, therefore leading to multiple integrations of the construct and possible artifact observations. In addition, since the vectors for gene deletion and gene tagging are often derived from the same building blocks, each transformation can lead to a plethora of undesired results, with the combinations number depending on the number of already integrated cassette in the transformed strain. Although reporter assays often rely on fluorescent proteins that can be used to screen for the integration number or genotyping, the process is time-consuming and the result not certain.

2.2 Results

The content of this chapter has been published as a Methods Paper in Molecular Genetics and Genomic on the 16th of September 2016.

Wosika, V., Durandau, E., Varidel, C., Aymoz D., Schmitt M. & Pelet S., “New families of single integration vectors and gene tagging plasmids for genetic manipulations in budding yeast”, *Molecular Genetics and Genomics* 2016 Dec;291(6):2231-2240.

Author contributions:

Victoria Wosika, Eric Durandau, Delphine Aymoz and Serge Pelet designed the tools. Victoria Wosika, Clémence Varidel and Marta Schmitt constructed the vectors. Victoria Wosika constructed the strains and performed the experiments. Victoria Wosika and Serge Pelet wrote the manuscript.

2.3 Summary

In this paper, we report on the construction of a set of Single Integration Vectors (pSIV) and Gene Tagging plasmids (pGT) for the genetic manipulation of budding yeasts. Thanks to the use of completely exogenous marker cassettes (promoters, marker genes and terminators) and promoter/terminator recombination sites instead of selection marker open reading frame to integrate, these plasmids show a robust unique integration and no off-target effect. Finally, the pGT presented were built with the current best FP, thus updating available tags.

2.4 Conclusion

We developed highly robust and efficient tools for the genetic manipulation of *Saccharomyces cerevisiae*, which are user-friendly and can be of use to the entire budding yeast community.

Title: New families of single integration vectors and gene tagging plasmids for genetic manipulations in budding yeast.
Authors: Wosika V, Durandau E, Varidel C, Aymoz D, Schmitt M, Pelet S
Journal: Molecular genetics and genomics : MGG
Year: 2016 Dec
Volume: 291
Issue: 6
Pages: 2231-2240
DOI: 10.1007/s00438-016-1249-1

New families of Single Integration Vectors and Gene Tagging plasmids for genetic manipulations in budding yeast

Victoria Wosika, Eric Durandau, Clémence Varidel, Delphine Aymoz, Marta Schmitt and Serge Pelet*

Department of Fundamental Microbiology, University of Lausanne, Switzerland

*Corresponding Author: serge.pelet@unil.ch; phone +41 21 692 5621

Abstract

The tractability of the budding yeast genome has provided many insights into the fundamental mechanisms regulating cellular life. With the advent of synthetic biology and single-cell measurements, novel tools are required to manipulate the yeast genome in a more controlled manner. We present here a new family of yeast shuttle vectors called Single Integration Vectors (pSIV). Upon transformation in yeast, these plasmids replace the entire deficient auxotrophy marker locus by a cassette containing an exogenous marker. As shown using flow cytometry, this complete replacement results in a unique integration of the desired DNA fragment at the marker locus. In addition, a second transcriptional unit can be inserted to achieve the simultaneous integration of two constructs. The selection marker cassettes, present in the pSIV, were also used to generate a complete set of gene tagging plasmids encompassing a large palette of fluorescent proteins, from a Cyan Fluorescent Protein (CFP) to a near-infrared tandem dimer Red Fluorescent Protein (tdiRFP). These tagging cassettes are orthogonal to each other thanks to the use of different TEF promoter and terminator couples thereby avoiding marker cassette switching and favoring integration in the desired locus. In summary, we have created two sets of robust molecular tools for the precise genetic manipulation of the budding yeast.

Keywords: plasmid, integration, gene tagging, fluorescent protein, transformation, molecular biology, genetic modification, yeast expression

Introduction

Pioneer model organism in the study of eukaryotic cells, *Saccharomyces cerevisiae* has contributed to the understanding of many cellular mechanisms such as cell-cycle regulation, cell polarity, signal transduction, metabolism or aging (Drubin and Nelson 1996; Nasmyth 1996; Chen and Thorner 2007; Cai and Tu 2012; Longo et al. 2012). This important role in fundamental biology can be attributed in a large part to the variety of tools developed to manipulate the yeast genome (Botstein et al. 1997; Gietz and Woods 2006; Botstein and Fink 2011). The generation of genome-wide collections of ORF deletions or ORF-GFP tagged strains are two excellent illustrations of this ability (Winzeler et al. 1999; Huh et al. 2003).

Shuttle vectors, allowing the transfer of cloning plasmids directly from a bacterial host to a yeast strain, have also participated in the success of *S.cerevisiae* as a eukaryotic model system. A standard set of plasmids (pRS) with four auxotrophy markers for *HIS3*, *LEU2*, *TRP1* and *URA3* is widely used in the yeast community (Sikorski and Hieter 1989). Three variants of these plasmids are available: two-micron, centromeric and integrative. Two-micron plasmids are high copy number plasmids used for the over-expression of proteins. Centromeric plasmids possess an autonomously replicating sequence (ARS) allowing them to be replicated at each cell cycle and segregated between mother and daughter cells. For these constructs, the average number of plasmids per cell in the population is around one. However, at the single-cell level, some yeast contains two copies of the plasmid while others have lost it. Finally, integrative plasmids are directly inserted into the yeast genome via homologous recombination at a specific locus. This generates an evident uniformity between sister cells bearing the plasmids since they are isogenic. Moreover, the insert can be maintained at its locus even in absence of selection.

Centromeric plasmids have been used successfully in countless studies, where responses were measured at the population level. In the last decade, with the development of single-cell studies and synthetic biology, a need has emerged to control precisely the genetic information in each individual cell. The integrative plasmids have thus gained in importance. The stable integration in the host genome allows a more uniform expression of the exogenous construct in each cell of the population. This can be crucial for synthetic gene regulatory circuits where each part of the network has to be expressed at a stable and defined level (Cantone et al. 2009). Studies in the transcription field also require the stable integration of expression reporters in order to assess noise in gene expression. Indeed, the presence of multiple copies will alter the characterization of the variability in expression or will change the measured rate of protein expression.

Unfortunately, the standard pRS integrative plasmids have a large tendency to integrate multiple times at the same locus. In this paper, we have designed a new family of shuttle vectors that integrate only once thanks to a complete replacement of the deficient auxotrophy locus. We assessed the efficiency of unique integration of our plasmids against the standard pRS vectors by yeast transformation and flow cytometry analysis of the transformants fluorescence. Additionally, using the same auxotrophy cassettes, we generated a complete set of gene tagging plasmids comprising five commonly used fluorescent proteins, thus covering a large spectral range from blue to near-infrared.

Materials and methods

Strains and plasmids
All experiments were performed in W303 background. Yeast strains are listed in Table S1. The Single Integration Vector URA (pSIVu) was constructed by Gibson Assembly (Gibson Assembly Master Mix, New England BioLabs) of PCR amplified fragments (Microsynth) and synthetized DNA geneblocks (Integrated DNA technologies) with overlapping compatible tails. The Single Integration Vectors LEU (pSIVl), HIS (pSIVh) and TRP (pSIVt) were constructed by restriction digest of pSIVu with PacI to retrieve the bacterial part (Fig. 2a white part) and ligating it with synthetic DNA fragments (Biomatik) bearing the synthetized yeast part sequences, either coming from a different yeast species or codon shuffled *in silico* (Table 1).

For the FACS analysis, the same construction (p*PRS2*-mCherry-t*SIF2*) was inserted into a standard pRS306 plasmid and the newly designed Single Integration Vector URA (pSIVu). Additionally, in the pSIV vector, we inserted a second transcriptional unit in the MCS2 encoding for a double green fluorescent protein (p*TEF*-PP7-2xGFP-t*CYC1*). This construct was first cloned in the minimal plasmid (pMCV) and subsequently cloned with AatII-SphI into the pSIV. Thes plasmids are listed in Table S2

Transformants generation and flow cytometry analysis
Chemically competent cells were generated from a culture of W303 wild type strain (ySP2) following standard protocol (Gietz and Woods 2006) and transformed with either 1.5µg of EcoRV linearized pRS vector (pVW110) or 1.5µg of PacI linearized pSIV vector (pVW169) and plated on SD-URA selection plates. Four replicates were performed in parallel, leading to four transformation plates per vector. After three days of growth at 30 °C, twenty-four transformants were randomly picked from each transformation plate and streaked on a new selection plate for one day of growth. All isolated transformants were then grown overnight in 200µL selective medium to saturation in 96 well plates (Greiner CELLSTAR 96 M9311-100EA). The next morning, cells were diluted 40-fold into 200 µl of non-selective medium (SD-Full) and analyzed by flow cytometry after at least four hours of growth.

Samples were loaded as 96 well plates with the High Throughput Sampler into an LSRFortessa flow cytometer (BD Biosciences). Red and green fluorescently tagged histones strains were used as fluorescence positive controls. Isolated versus clustered cells were separated by a gating based on FSC and SSC values of a sonicated WT culture. Measurements were acquired for 10'000 recorded events in the defined gate or 100 seconds. For the red fluorescence, samples were excited with a 561 nm yellow-green laser and the fluorescence emission was detected through a 610/20 nm filter. Green fluorescence was measured by exciting with a 488 nm blue laser and detected with a 530/30 nm filter.

Flow cytometry data were analyzed with Matlab (The MathWorks). A gating on the FSC and SSC was applied to select the single cell events. Only 6 samples out of 192 had fewer than 8'000 events in the gate. The fluorescence of these samples was within the expected distribution of the population suggesting that they were less concentrated due to low inoculation or small dilution issues during sample preparation. Definition of the single integrant status was performed by setting low and high thresholds for the median fluorescence value of the sample. For the pSIV, the difference between the lowest and highest single integration transformant is 750. For the pRS, we added 500 to the lowest transformant intensity above background to set the high fluorescence threshold.

Gene tagging and transformation
Gene tagging primers were designed with consensus forward and reverse sequences annealing to all pGT vectors (Fig. 4a blue primers): 5'-GCGGCCGCTCTAGAACTA-3' and 5'-ATGGAAAAACGCCAGCAACG-3'respectivelly, to which a gene specific sequence of 40 bp was flanked and used for the homologous recombination with the C-terminus of the targeted ORF. Table S3 lists all the primers used in this study. Gene tagging PCRs were run as 50 µl reaction at an annealing temperature of 65 °C (30s), with an extension time of 2 min (72 °) for all pGTs, using a high-fidelity enzyme (Q5 High-Fidelity polymerase, New England Biolabs). After a gel electrophoresis control of the PCR specificity, 20µl of PCR reaction were used to transform the cells following a standard lithium acetate protocol (Gietz and Woods 2006).

Microscopy
Images were acquired on an inverted epi-fluorescence microscope (Ti-Eclipse, Nikon) controlled by micro-manager (Edelstein et al. 2010), with a 60X oil objective and excitation and emission filters: CFP, YFP, RFP and iRFP. The excitation light is provided by a solid-state light source (SpectraX, Lumencor) and the images were recorded with an sCMOS camera (Flash4.0, Hamamatsu) with the following exposure times: CFP (100ms), YFP (300ms) RFP (100 ms) iRFP (100 ms). Log-phase yeast cultures were diluted to OD 0.04, briefly sonicated and 200 µl of culture were loaded in a well coated with ConcanavalinA (0.5 mg/ml, C2010-250MG, Sigma-Aldrich). Cells were stressed by addition of 100 µl of SD-full NaCl 1.2M in the well to a final concentration of 0.4M NaCl.

All pSIV and pGT plasmids presented in this study will be made available on Addgene. Their accession numbers are provided in Table 1 and Table 2.

Results

pRS integration
The integration of the vectors from the pRS family requires a linearization of the backbone by enzymatic restriction in the selection marker. In the W303 background, auxotrophy for all selection markers is due to deficient marker genes, with one or two point mutations (*leu2-3,112; trp1-1; ura3-1; his3-11,15*) (Ralser et al. 2012). Homologous recombination can be obtained by digestion of the pRS plasmid within the ORF of the marker. In BY4741, there is a complete deletion of the auxotrophy ORFs (*his3*Δ, *leu2*Δ and *ura3*Δ) (Brachmann et al. 1998). Integration of the pRS vectors is performed by cutting either in the promoter or the terminator of the marker to obtain homology regions with the corresponding genomic sequence. In either case, the integration of a wild type sequence restores the prototrophy of the targeted gene thanks to a duplication of the genomic region upon integration (Fig. 1a). These two sequences can in turn serve as recombination sites for other transformed DNA molecules, possibly leading to multiple integrations. Therefore, transformants have to be screened for single integration of the vector. This screening process can be relatively straightforward in the case of a fluorescent construct. However, genotyping by PCR is typically difficult due to the large size of the plasmid backbone. Alternatively, single versus multiple integrations can be verified by qPCR or Southern blots.

Single Integration Vector architecture
The solution to this multiple integration issue is rather simple and can be achieved by having two sites of homology at the 5' and 3'-ends of the integrating cassette such that the entire

locus is replaced at each integration event (Fig. 1b). This implies, however, that the auxotrophy marker should not have any homology with the endogenous locus to prevent recombination at wrong sites. To this purpose, we synthetized four exogenous auxotrophy cassettes (Table 1). For the *URA3* and *HIS3* loci, we used the *C. albicans* and *S. pombe* homologs, respectively, which were already used in the gene tagging plasmids (pKT, (Sheff and Thorn 2004)). For *LEU2* and *TRP1*, we scrambled the endogenous *S. cerevisiae* sequence to preserve the amino acid order while randomizing the codon usage and considering the codon bias of the budding yeast. We also optimized the DNA sequence to remove the most commonly used restriction sites. Similarly to a large family of gene deletion plasmids (Wach et al. 1997; Longtine et al. 1998), we used TEF promoters and terminators to control the expression of these four markers (Steiner and Philippsen 1994). In order to avoid marker switch when transforming successively multiple pSIV plasmids in the same strain, we used four different combinations of p*TEF* and t*TEF* from close relative yeast species (Table 1). We verified by growth curve and spot assays that these new markers can sustain growth in selective and non-selective media (Supp. Fig. 1). To have the ability to recycle some selection markers, we inserted inverted *loxP* sites (Fig. 2a) at each side of the markers to offer the possibility to loop out the marker using the Cre recombinase (Sauer 1987).

The homology with the genomic DNA is based on two 250 bp regions in the promoter and terminator of the marker gene (Table 1). The sequence in the promoter region was selected upstream of the transcription start site based on tiling array data (David et al. 2006), in order to prevent the production of transcripts arising from the genomic promoter inside the plasmid. Upstream of the promoter and downstream of the terminator homology regions, two restriction sites for PacI and BstBI have been inserted. As it can be seen on the map of the pSIV vector on Figure 2a, digestion of the plasmid with either of one of these restriction enzymes will generate two fragments. One half contains the yeast integrative element (green elements in Fig. 2a) and the other contains all the sequences required for the plasmid amplification in bacteria. In opposition to what happens with the pRS plasmid, this bacterial part will not be integrated in the yeast genome, which reduces the size of the integrated DNA.

A multiple cloning site (MCS) is present downstream of the TEF terminator (Fig. 2a). This MCS contains all the sites present in the MCS of the standard pBLUESCRIPT or pRS vectors (Sikorski and Hieter 1989). The sequences of the vector have been optimized in order to ensure that each one of these sites is unique. Because it is often necessary to introduce multiple constructs in one strain, we created the possibility to integrate a second MCS, between the unique restriction sites AatII and SphI, which are positioned upstream of the TEF promoter. For this purpose, we generated a MCS2 cloning vector (pMCV) containing the bacterial part from the pSIV and a standard MCS flanked by AatII and SphI that we call MCS2 (Fig. 2b). The second construct can be assembled in this minimal plasmid, sequence verified and sub-cloned into the pSIV. Note that since the MCS1 and MCS2 contain the same restriction sites, insertion of the second MCS often leads to a duplication of many sites. Sequences annealing to the standard sequencing primers T3/T7 and M13forward/reverse were inserted on each side of the MCS1 and MCS2, respectively, allowing independent sequencing of each MCS.

Verification of single insertion into the genome

To compare the efficiency of the pSIV and pRS plasmids to generate single integration transformants, we cloned a red fluorescent protein (RFP) variant mCherry under the control of the constitutive promoter p*RPS2* in a pSIV URA3 (pSIVu) and in a pRS URA3 (pRS306). A

similar amount of these two backbones was linearized by restriction digest (respectively PacI and EcoRV). Without purification, the entire restriction digestion mix was added to chemically competent W303 wild type cells following standard transformation protocol (Giesecke et al. 2006).

After three days of growth on SD-URA plates, twenty-four single colonies from each transformation were streaked on selective plates. They were then grown overnight in 96-well plates in liquid medium to saturation, diluted 40 fold and grown for 4 hours before the fluorescence of individual cells was measured by flow cytometry. Figure 3a represents the histogram of the fluorescence of more than eight thousand cells for ten pSIV transformants. All the curves show an almost complete overlap. In comparison, the ten transformants bearing the pRS plasmid display a large variability in fluorescence levels (Fig. 3b). The median fluorescence intensities of the 24 selected colonies are plotted in increasing order for the four replicates (Fig. 3c and 3d). Out of the total 96 colonies screened for the pSIV, only two of them display an aberrant fluorescence level. One is non-fluorescent, the other one is 6-times brighter than the other clones (Figure 3c). The picture is strikingly different with the pRS plasmid, where the clones cover a large range of fluorescence (Figure 3d). In both cases, we defined a range of fluorescence intensity that we consider as single integration transformants. For the pSIV, 98% of the clones fell in this range, while only 12.5% can be considered to have a single integration for the pRS transformation. The rest of them are false positives (23%) or multiple integrations (64.5%) (Fig. 3e).

Correlation of MCS1 and MCS2 integration

In order to verify if the MCS1 and the MCS2 present on the pSIV plasmid integrate with the same efficiency in the genome, the pSIV vector with the p*RPS2*-mCherry construct also contained a p*TEF* 2xGFP inserted in the MCS2. The fluorescence intensity of the pSIV transformants was scored simultaneously in the red and green channels of the flow cytometer (Fig. 3f). Out of the 96 clones, three outliers were found: the two previously measured clones with no RFP and very high RFP intensities, plus one additional strain that has a normal RFP level but displays a two-fold higher GFP intensity. Overall this analysis suggests that the simultaneous transformation of both MCSs is the most likely outcome.

Development of a complete set of gene tagging plasmids

In parallel to plasmid integration, the ability to tag genes directly with fluorescent proteins has provided many insights into the localization and function of yeast proteins (van Drogen et al. 2001; Maeder et al. 2007; Cai et al. 2008; Pelet et al. 2011). Moreover, it is often desirable to tag simultaneously multiple proteins in the same cells to correlate their location. The pKT plasmids are widely used in the yeast community to tag proteins with various fluorescent spectral variants. However, only two auxotrophy markers (HIS and URA) and one antibiotic resistance (KAN) are available for selection of the positive transformants. In addition, the same TEF promoter and terminator are used for all these cassettes. This increases the chances of inducing a marker exchange rather than the correct insertion of a second fluorescent tag at the desired locus.

In order to allow the tagging of four different proteins, each with a different fluorescent tag, we decided to use the auxotrophy marker generated for the pSIV plasmids and create a family of gene tagging plasmids (pGT) that covers the spectrum of commonly used FPs from CFP to iRFP (Table 2). For the CFP channel, we used the yeast enhanced monomeric variant of CFP (yemCFP) (Sheff and Thorn 2004). In the YFP channel, the mCitrine variant was used because of its higher photostability compared to the faster maturing Venus. Note that two

mutations (A206K, L221K) are used to render it monomeric (Zacharias et al. 2002; Sheff and Thorn 2004). The superfolder GFP (sfGFP) was chosen due to its brightness, fast maturation and a recognized ability to tag proteins that are usually difficult to render fluorescent (Pédelacq et al. 2006). We selected the RFP variant mCherry due to its brightness and photostability (Shaner et al. 2004). In the far-red region of the spectrum, we use the tandem dimer infrared FP (tdiRFP) (Filonov et al. 2011). This protein has a relatively low brightness compared to other FPs, and should mostly be used to tag abundant proteins. In order to facilitate the PCR amplification of this tandem dimer protein, we codon shuffled the coding sequence of the second copy of the iRFP.

All monomeric fluorescent proteins are cloned between XbaI and XhoI sites, allowing an easy exchange of the tagging peptide, which could be newer versions of FP or protein affinity tags for biochemistry experiments (Fig. 4a). An *ADH* terminator is placed downstream of the fluorescent protein sequence. As in the pSIV, *lox*P sites flank the marker cassette to excise and recycle it for other genetic manipulations.

To test the efficient tagging of pGT plasmids and their orthogonality, we used all four markers to tag four proteins inside the same cell: Pma1-yemCFP:*URA3* (membranes), Dcp2-mCitrine:*LEU2* (P-bodies), Hog1-mCherry:*TRP1* (cytoplasm and nucleus) and Hta2-tdiRFP:*HIS3* (nucleus). Cells were imaged before and after a hyper-osmotic shock. This stress leads to the accumulation of Hog1 in the nucleus (Reiser et al. 1999) and the formation of P-bodies (Teixeira et al. 2005) (Fig. 4b).

Discussion

In this study, we developed two sets of plasmids for precise genetic manipulation of the budding yeast based on optimized auxotrophy markers. Theses non-endogenous cassettes were then used to build four shuttle vectors that only integrate once into the genome (pSIVs, Table 1) and new gene tagging plasmids (pGTs, Table 2). The presence of a different set of promoter and terminator for each one of the markers allows an orthogonality within the pGTs and pSIVs within a single cell and increases the efficiency of correct genetic integrations during yeast transformation.

To demonstrate the unique genomic integration of the Single Integration Vectors (pSIVs), we scored the number of single integrations against the standard pRS vectors. As expected, pSIV transformation led to a clearly improved homogeneity between the transformants, compared to pRS transformants. Surprisingly, the low RFP fluorescence intensity range that we attributed to single integration events was not the same for both types of transformants. The single integration threshold for the pSIV plasmids was set higher than for the pRS vectors (Fig. 3c and 3d dashed blue lines). The fact that virtually all pSIV transformants displayed the same fluorescence intensity strongly hinted that they were single integration clones. In order to verify that single integration was the predominant form of integration of the pSIV plasmid, we transformed an empty plasmid and verified by PCR on genomic locus the size of the inserted cassette using forward and reverse primers outside of the integration region (Supp Fig. 2). The eleven transformants tested displayed the expected band at 2.2kb versus 1.5kb for the endogenous URA3 locus in the mother strain. Thus, the discrepancy in fluorescence intensity between pSIV and pRS plasmids is not due to multiple integrations of the pSIV. A more likely explanation of this difference resides in the backbone composition. Terminators are known to play a key role in mRNA stability and therefore in controlling the steady-state

level of protein expression. Our red fluorescent construct carries a *SIF2* terminator, which has been described to be of intermediate strength (Yamanishi et al. 2013). Thus, if not all transcription is terminated within the 500bp of the terminator, the sequence directly following this terminator could influence the stability of the mRNA and thereby explain the difference between pSIV and pRS transformants intensities.

Because there is only a handful of available selection markers, a second insert can be cloned between AatII and SphI in the pSIVs, with an intermediate cloning step into a specifically designed MCS2 cloning vector (Fig. 1b). The two transcriptional units can be expressed and integrated within the same plasmid. Care must be taken, however, when designing the plasmid, to avoid large homology regions between the two inserts. These similar DNA sequences could perturb the whole replacement of the cassette when multiple successive homologous recombinations happen in the same cell.

In summary, pSIVs transformation provides a clearly improved reliability in generating single integration transformants. Combined integration of the two MCSs happens in an overwhelming majority of cases. Despite this great fidelity, individual transformants have to be screened. However, the number of transformants to screen is reduced to a minimum, allowing to perform directly a deeper phenotypic analysis of each one of these few clones. In comparison, the pRS plasmid transformation generates a great diversity in fluorescence expression levels, which renders careful screening processes essential and time-consuming. It has to be noted that the amount of plasmid transformed and the competency of the cells will strongly influence the output of this transformation. Reducing the amount of plasmid tends to decrease the number of multiple integrations in the final clones. In each one of the four pRS transformations performed, we identified at least one single integration clone, but the large diversity in fluorescence intensity makes it difficult to set the threshold with high precision.

To conclude, we described here four new non-endogenous transcriptional units expressing auxotrophy markers for the design of a set of Single Integration Vectors (pSIV) and their derivative Gene Tagging plasmids (pGT). In comparison to the standard yeast shuffle pRS vectors, which integrate inside the coding sequence of selection markers, pSIV plasmids integrate into the host genome thanks to two homology sequences: the first one in the promoter and the second one in the terminator of the auxotrophy locus. This complete exchange of the deficient gene by an exogenous DNA fragment leads to a single integration of the vector. In comparison, the pRS plasmids integrate multiple times because of the duplication of the marker locus. We demonstrated these behaviors by quantifying the integration of a fluorescent protein construct using flow cytometry. Moreover, thanks to this mechanism of integration, pSIVs can be applied identically for transformation in the two most commonly used *S. cerevisiae* backgrounds W303 and BY4741. In addition, using the same four auxotrophy markers, we constructed a complete family of gene tagging plasmids for the fluorescent labeling of endogenous proteins. We tagged four different proteins in the same cell with different FP spectral variants demonstrating that pGTs can be used in parallel to visualize multiple proteins in the same strain, thanks to their orthogonality. All the plasmids described here are available on Addgene together with their sequences.

Acknowledgements

We thank all members of the Pelet and Martin labs for helpful discussions, Barbara Brandani for technical assistance and Paul Majcherczyk for the support with the FACS machine.

Author contributions

SP, VW, DA and ED designed the experiments. VW, DA and MS constructed and tested the pSIV plasmids. ED and CV constructed and tested the pGT library. VW performed the flow cytometry and microscopy experiments. SP and VW analyzed the data and wrote the manuscript.

Compliance with ethical standards

Funding. This study was supported by SystemsX.ch (IPhD 51PHP0_157354) and Swiss National Science Foundation grants (PP00P3_139121) and the University of Lausanne.

Conflict of interest. All authors declare that they have no conflict of interest.

Ethical approval. This article does not contain any studies with human participants or animals performed by any of the authors.

References

Botstein D, Chervitz SA, Cherry JM (1997) Yeast as a model organism. Science 277:1259–1260.

Botstein D, Fink GR (2011) Yeast: an experimental organism for 21st Century biology. Genetics 189:695–704.

Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, Boeke JD (1998) Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast 14:115–132.

Cai L, Dalal CK, Elowitz MB (2008) Frequency-modulated nuclear localization bursts coordinate gene regulation. Nature 455:485–490.

Cai L, Tu BP (2012) Driving the cell cycle through metabolism. Annu Rev Cell Dev Biol 28:59–87.

Cantone I, Marucci L, Iorio F, Ricci MA, Belcastro V, Bansal M, Santini S, di Bernardo M, di Bernardo D, Cosma MP (2009) A yeast synthetic network for in vivo assessment of reverse-engineering and modeling approaches. Cell 137:172–181.

Chen RE, Thorner J (2007) Function and regulation in MAPK signaling pathways: lessons learned from the yeast Saccharomyces cerevisiae. Biochim Biophys Acta 1773:1311–1340.

David L, Huber W, Granovskaia M, Toedling J, Palm CJ, Bofkin L, Jones T, Davis RW, Steinmetz LM (2006) A high-resolution map of transcription in the yeast genome. Proc Natl Acad Sci USA 103:5320–5325.

Drubin DG, Nelson WJ (1996) Origins of cell polarity. Cell 84:335–344.

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

Filonov GS, Piatkevich KD, Ting L-M, Zhang J, Kim K, Verkhusha VV (2011) Bright and stable near-infrared fluorescent protein for in vivo imaging. Nat Biotechnol 29:759–763.

Giesecke AV, Fang R, Joung JK (2006) Synthetic protein–protein interaction domains created by shuffling Cys2His2 zinc-fingers. Molecular Systems Biology. doi: 10.1038/msb4100053

Gietz RD, Woods RA (2006) Yeast transformation by the LiAc/SS Carrier DNA/PEG method. Methods Mol Biol 313:107–120.

Huh W-K, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK (2003) Global analysis of protein localization in budding yeast. Nature 425:686–691.

Longo VD, Shadel GS, Kaeberlein M, Kennedy B (2012) Replicative and chronological aging in Saccharomyces cerevisiae. Cell Metab 16:18–31.

Longtine MS, McKenzie A, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast 14:953–961.

Maeder CI, Hink MA, Kinkhabwala A, Mayr R, Bastiaens PIH, Knop M (2007) Spatial regulation of Fus3 MAP kinase activity through a reaction-diffusion mechanism in yeast pheromone signalling. Nat Cell Biol 9:1319–1326.

Nasmyth K (1996) At the heart of the budding yeast cell cycle. Trends in Genetics 12:405–412.

Pelet S, Rudolf F, Nadal-Ribelles M, de Nadal E, Posas F, Peter M (2011) Transient activation of the HOG MAPK pathway regulates bimodal gene expression. Science 332:732–735.

Pédelacq J-D, Cabantous S, Tran T, Terwilliger TC, Waldo GS (2006) Engineering and characterization of a superfolder green fluorescent protein. Nat Biotechnol 24:79–88.

Ralser M, Kuhl H, Ralser M, Werber M, Lehrach H, Breitenbach M, 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.

Reiser V, Ruis H, 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. Mol Biol Cell 10:1147–1161.

Sauer B (1987) Functional expression of the cre-lox site-specific recombination system in the yeast Saccharomyces cerevisiae. Mol Cell Biol 7:2087–2096.

Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, Tsien RY (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat Biotechnol 22:1567–1572.

Sheff MA, Thorn KS (2004) Optimized cassettes for fluorescent protein tagging in Saccharomyces cerevisiae. Yeast 21:661–670.

Sikorski RS, Hieter P (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics 122:19–27.

Steiner S, Philippsen P (1994) Sequence and promoter analysis of the highly expressed TEF gene of the filamentous fungus Ashbya gossypii. Mol Gen Genet 242:263–271.

Teixeira D, Sheth U, Valencia-Sanchez MA, Brengues M, Parker R (2005) Processing bodies require RNA for assembly and contain nontranslating mRNAs. RNA 11:371–382.

van Drogen F, Stucke VM, Jorritsma G, Peter M (2001) MAP kinase dynamics in response to pheromones in budding yeast. Nat Cell Biol 3:1051–1059.

Wach A, Brachat A, Alberti-Segui C, Rebischung C, Philippsen P (1997) Heterologous HIS3 marker and GFP reporter modules for PCR-targeting in Saccharomyces cerevisiae. Yeast 13:1065–1075.

Winzeler EA, Shoemaker DD, Astromoff A, et al (1999) Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science 285:901–906.

Yamanishi M, Ito Y, Kintaka R, Imamura C, Katahira S, Ikeuchi A, Moriya H, Matsuyama T (2013) A Genome-Wide Activity Assessment of Terminator Regions in Saccharomyces cerevisiaeProvides a ″Terminatome″ Toolbox. ACS Synth Biol 2:337–347.

Zacharias DA, Violin JD, Newton AC, Tsien RY (2002) Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells.

Figure Legends

Figure 1. Integration mechanism of pRS vectors compared to Single Integration Vectors a. Scheme describing the integration of a pRS plasmid at the marker locus. The first integration event duplicates the marker generating one functional and one non-functional copy of the auxotrophy marker. Subsequent homologous recombination events will result in further multiplication of the marker and MCSs. **b.** Integration of the pSIV in the auxotrophy locus by homologous recombination with the promoter and terminator regions (pLOC, tLOC). The complete replacement of the locus results in a single MCS being integrated even if multiple successive homologous recombination events take place.

Figure 2. Map of the Single Integration Vector. a. The vector is composed of two parts, the bacterial part (white) and the yeast part (green). The expression of the marker is controlled by a *TEF* promoter and terminator couple. The homologous recombination occurs via the homology regions pLOCUS and tLOCUS. The MCS1 is composed of a set of unique restriction sites. A second MCS can be inserted between the AatII and SphI sites. **b.** Map of the MCS2 cloning vector composed only of the bacterial part and an MCS1 flanked by AatII and SphI for sub-cloning into the pSIV.

Figure 3. Comparison of the single integration efficiency for pSIV and pRS vectors. a. and **b.** Histograms of the RFP fluorescence of 10 transformants bearing the pSIV (a) and the pRS backbone (b). The insets in each panel describe the integrated construct. **c.** and **d.** Median RFP fluorescence intensities of 24 transformants obtained from four independent transformations of the pSIV (c) or the pRS (d) vectors. The colonies were sorted according to their median intensity level. The dashed lines delimit the intensity intervals considered as single integration transformants. **e.** Percentages of false positives, single integrations and multiple integrations for the transformation of the two different backbones. **f.** Correlation between RFP and GFP intensities in the 96 transformants quantified by flow cytometry. The inset is a blown-up of the region where the large majority of the clones are found.

Figure 4. Four colour tagging of a yeast strain using the pGT cassettes. a. Schematic of the gene tagging cassette. **b.** Images of cells bearing four fluorescent proteins imaged before and after (20 min) stress with 0.4M NaCl. Hta2-tdiRFP stains the nucleus. Pma1-CFP localizes at the plasma membrane and the vacuole. Hog1-mCherry translocates in the nucleus after stimulus. The aggregation of P-bodies induced by the stress can be visualized by local enrichments of Dcp2-mCitrine.

ables

able 1: pSIV markers information

larker	Gene origin	pTEF species	tTEF species	Marker coordinates	pMARKER coordinates	tMARKER coordinates	Marker size (bp)	Addgene ID
URA3	*C. albicans* Ura3	*A. gossypii*	*A.gossypii*	116167-116970 (chr. V)	116067-115815	116791-117041	1420	**81089**
LEU2	*S. cerevisiae* Leu2 scr.	*S. kudriavzevii*	*C.glabrata*	91324-92418 (chr.III)	90980-91230	92450-92700	1845	**81090**
HIS3	*S. pombe* His5	*S. kudriavzevii*	*A.gossypii*	721946-722608 (chr.XV)	721650-721900	722630-722880	1382	**81091**
TRP1	*S. cerevisiae* Trp1 scr.	*A. gossypii*	*C.glabrata*	461842-462516 (chr.IV)	461350-461600	462530-462780	1306	**81092**

: scrambled, chr : chromosome, p: promoter, t: terminator
rker size includes the pTEF, the marker and the tTEF sequences.

able 2: Gene tagging plasmids (pGT) set

arker	yemCFP	sfGFP	mCitrine	mCherry	tdiRFP
RA3	GTU-c (81093)	GTU-g (81094)	GTU-y (81095)	GTU-r (81096)	GTU-ir (81097)
EU2	GTL-c (81098)	GTL-g (81099)	GTL-y (81100)	GTL-r (81195)	GTL-ir (81102)
S3	GTH-c (81103)	GTH-g (81104)	GTH-y (81105)	GTH-r (81106)	GTH-ir (81107)
RP1	GTT-c (81108)	GTT-g (81109)	GTT-y (81110)	GTT-r (81111)	GTT-ir (81112)

': gene tagging, c: cyan, g: green, y: yellow, r: red, ir: infrared, (Addgene ID).

Figure 1:

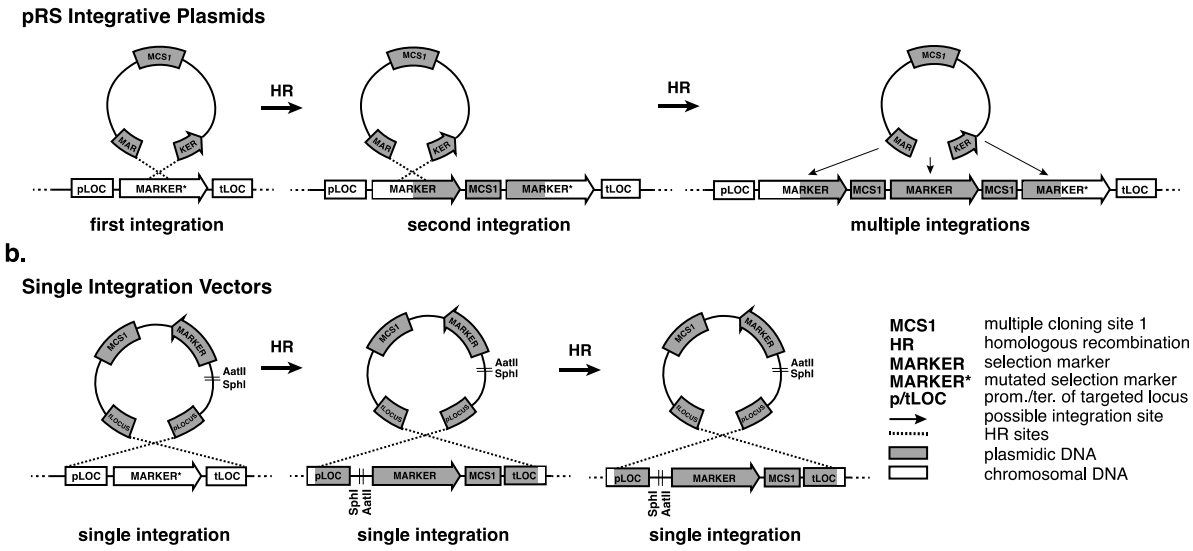

a.
pRS Integrative Plasmids
MCS1
HR
MAR
KER
pLOC
MARKER*
tLOC
MARKER
first integration
second integration
multiple integrations
b.
Single Integration Vectors
AatII
SphI
pLocus
tLocus
single integration
MCS1 multiple cloning site 1
HR homologous recombination
MARKER selection marker
MARKER* mutated selection marker
p/tLOC prom./ter. of targeted locus
possible integration site
HR sites
plasmidic DNA
chromosomal DNA

Figure 2:

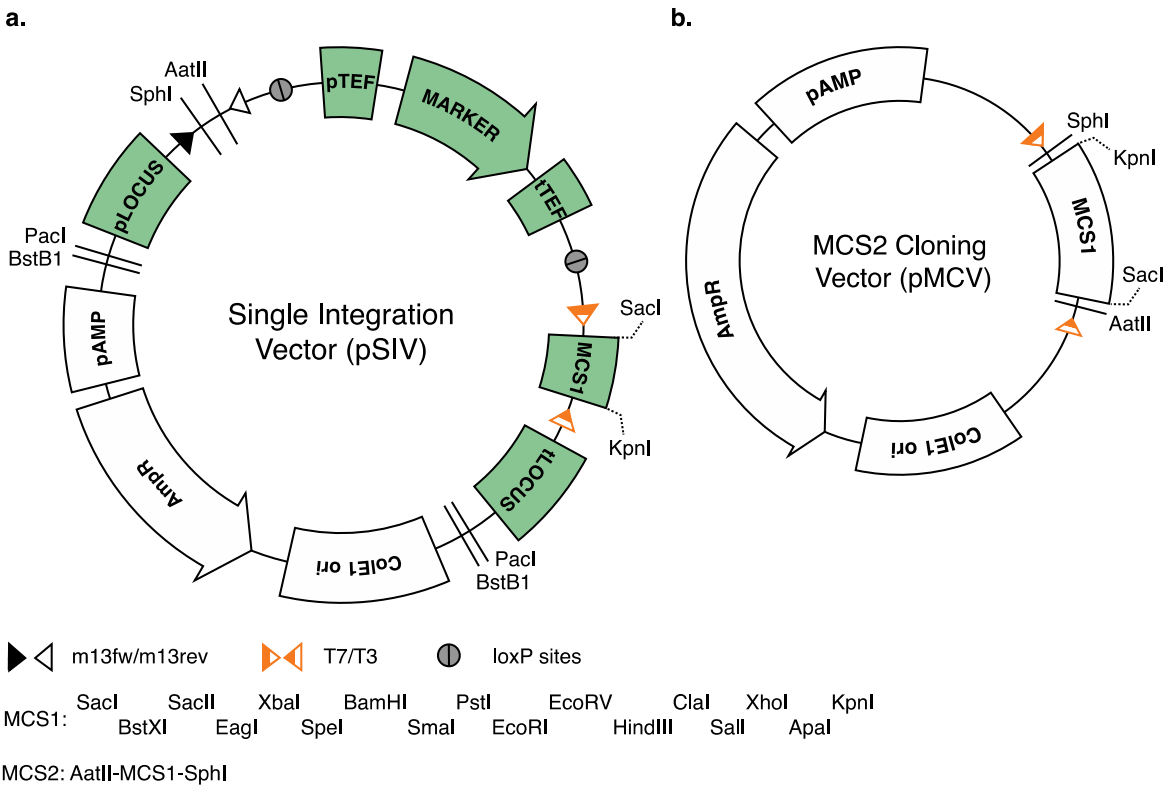

a.
b.
AatII
SphI
pTEF
MARKER
tTEF
pLOCUS
PacI
BstB1
pAMP
Single Integration
Vector (pSIV)
SacI
MCS1
KpnI
tLOCUS
AmpR
ColE1 ori
PacI
BstB1
pAMP
SphI
KpnI
MCS1
MCS2 Cloning
Vector (pMCV)
SacI
AatII
AmpR
ColE1 ori
m13fw/m13rev
T7/T3
loxP sites
MCS1: SacI SacII XbaI BamHI PstI EcoRV ClaI XhoI KpnI
BstXI EagI SpeI SmaI EcoRI HindIII SalI ApaI
MCS2: AatII-MCS1-SphI

Figure 3:

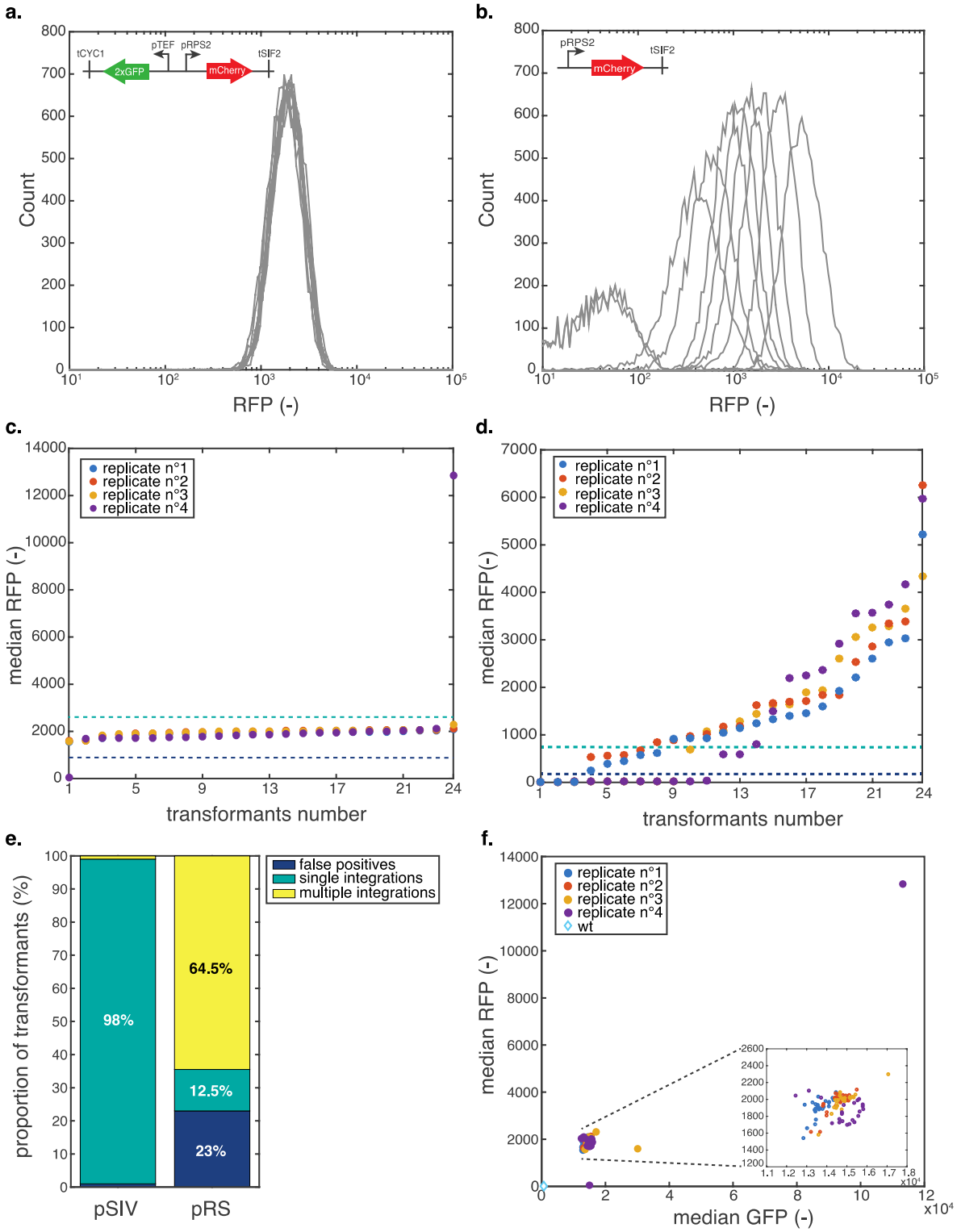
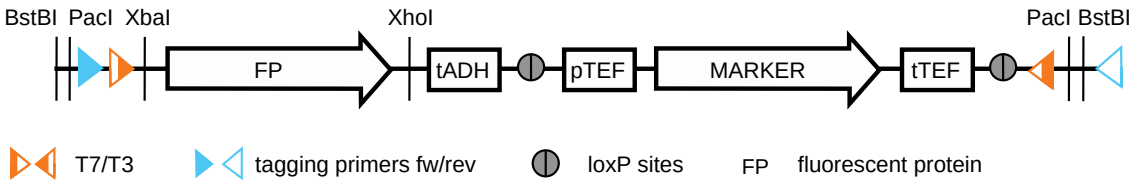
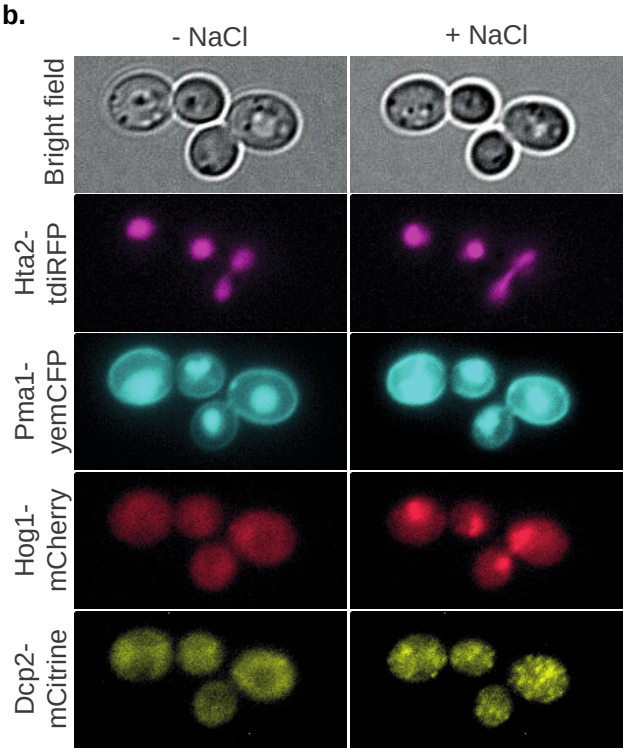

a.
tCYC1
2xGFP
pTEF
pRPS2
mCherry
tSIF2
Count
RFP (-)
b.
pRPS2
mCherry
tSIF2
Count
RFP (-)
c.
replicate n°1
replicate n°2
replicate n°3
replicate n°4
median RFP (-)
transformants number
d.
replicate n°1
replicate n°2
replicate n°3
replicate n°4
median RFP(-)
transformants number
e.
false positives
single integrations
multiple integrations
proportion of transformants (%)
98%
64.5%
12.5%
23%
pSIV
pRS
f.
replicate n°1
replicate n°2
replicate n°3
replicate n°4
wt
median RFP (-)
median GFP (-)
x10⁴

Figure 4:

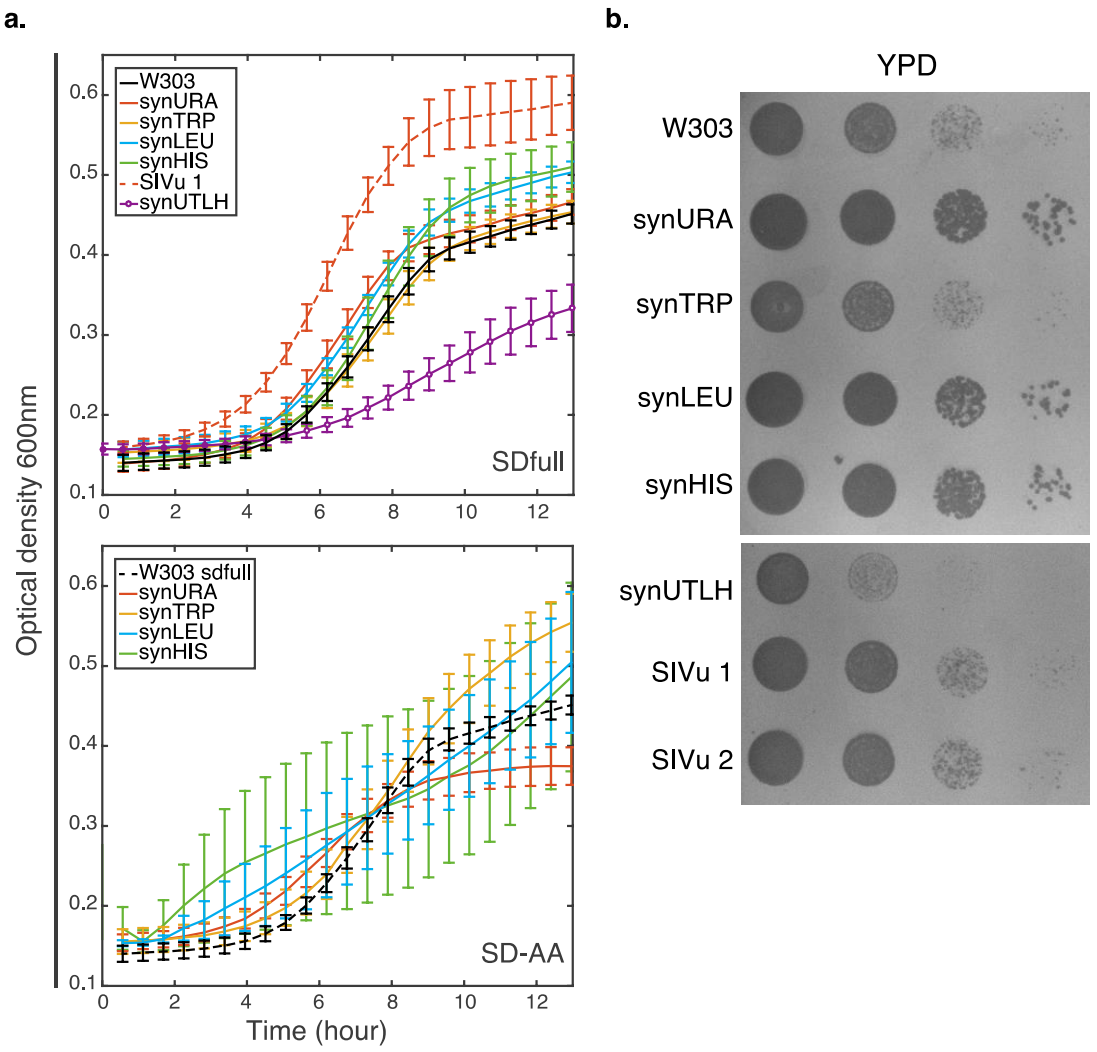
a.


BstBI
PacI
XbaI
XhoI
PacI
BstBI
FP
tADH
pTEF
MARKER
tTEF
T7/T3
tagging primers fw/rev
loxP sites
FP
fluorescent protein

b.

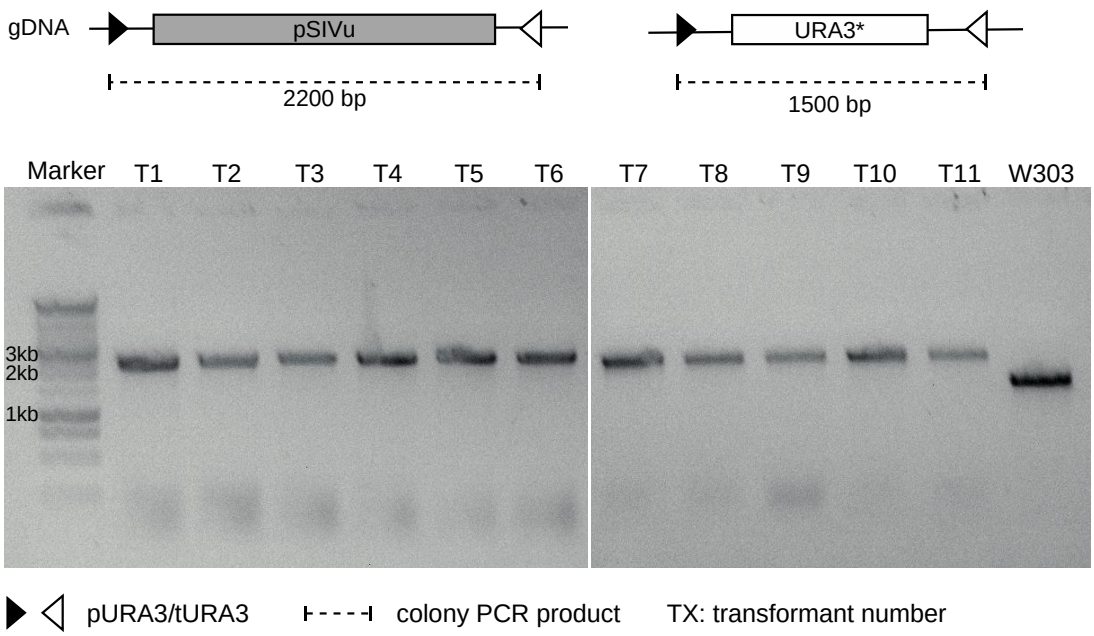

- NaCl
+ NaCl
Bright field
Hta2-tdiRFP
Pma1-yemCFP
Hog1-mCherry
Dcp2-mCitrine

Supplementary Figure 1:



Supplementary Figure 1: Integration of the synthetic marker cassettes in the genome has little influence growth rate. a) Growth curves of *hta2*-mCherry pGT tagged strains bearing each of the four marker cassettes, compared to the wt W303 background, the four colors tagged strain from Fig. 4b and a pSIVu clone from Fig 3, in SDfull and selective medium (SD-AA, except WT). The curves represent the mean and standart deviation of 8 replicates automatically measured every 30 min in a 96 wells plate by a Tecan Infinite microplate reader (Tecan Group Ltd, Männedorf, Switzerland), starting from an OD600 0.1 diluted overnight culture. **(b)** Spot assay performed with the same cultures as the growth assay. Cells were initially diluted to $1.26x10^7$ cells (first column) and then followed by 1/10 serial dilutions.

Supplementary Figure 2: Genotyping pSIV transformants for single integration



Supplementary Figure 2: Colony PCR to confirm the single integration of the pSIVu plasmid. Scheme showing the location of the primers (sequence found in Table S3) on the genomic DNA used for genotyping the presence and number of integrated plasmids (grey box) versus the endogenous locus (white box), and the expected PCR product size. Compared to the mother strain, all eleven transformants returned a band of 2.2 kb corresponding to the single integration of the empty plasmid, as shown in the agarose gel.

Supplementary Table S1: Strains used in this study

Yeasts	Genotype	Resistance
ySP2	*MAT*a {*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*}	-
yVW81	*w303 pRPS2 mCherry tSIF2*	U
yVW82	*w303 pRPS2 mCherry tSIF2 pTEF PP7 2xGFP tCYC1*	U
yVW140	*w303 hta2-iRFP:HIS dcp2-mCitrine:LEU hog1-RFP:TRP pma1-yemCFP:URA*	UHTL

ySP2: W303 naked background. U: URA3 resistance. UHLT: URA3, TRP1, LEU2 and HIS3 resistance

Supplementary Table S2: Plasmids used in this study

Plasmids	Insert	Selection
pVW110	pRS306 *pRPS2 mCherry tSIF2*	URA
pVW169	pSIVu *pRPS2 mCherry tSIF2 pTEF PP7 2xGFP tCYC1*	URA

Supplementary Table S3: Primer sequences

Primer name	5'-3' Sequence
pURA3	agaaaaggattaaagatgctaagagatag
tURA3	actcttgttgttctttggagttca
T3	AATTAACCCTCACTAAAGGG
T7	TAATACGACTCACTATAGGG
M13fw	TGTAAAACGACGGCCAGT
M13rev	CAGGAAACAGCTATGACCATG
pGT fw	GCGGCCGCTCTAGAACTA
pGT rev	ATGGAAAAACGCCAGCAACG
Dcp2 fw	cgaatggaacttcagggtctaatgaattattaagcattttgcataggaagGCGGCCGCTCTAGAACTA
Dcp2 rev	tcaaattgtgttatggttgtttaatcttattgaataccagtatcaaggatATGGAAAAACGCCAGCAACG
Hog1 fw	cggtaaccaggccatacagtacgctaatgagttccaacagGCGGCCGCTCTAGAACTA
Hog1 rev	gctgataaacaaacaatacgccataagtgacggttcttggATGGAAAAACGCCAGCAACG
Pma1 fw	acttcatggctgctatgcaaagagtctctactcaacacgaaaaggaaaccGCGGCCGCTCTAGAACTA
Pma1 rev	agttgattaaaatgtgacaaaaattatgattaaatgctacttcaacaggaATGGAAAAACGCCAGCAACG
Hta2 fw	acttgttgccaaagaagtctgccaagactgccaaagcttctcaagaactgGCGGCCGCTCTAGAACTA
Hta2 rev	cgtaacaaaagaaagagagcctagctgtaatatatctttataacatgtatATGGAAAAACGCCAGCAACG

Capital letter: plasmid-annealing sequence. Lower case: sequence annealing to genomic DNA.

Chapter 3: Simultaneous monitoring of transcription and translation

3.1 Background

3.1.1 mRNA and protein correlation

The central dogma in the field of gene expression is that DNA is transcribed into mRNA, which in turn is translated into proteins (James Watson, *Molecular Biology of the Gene* 1965). Since mRNAs and proteins require complex and abundant cellular machines to be synthetized, transcription and translation represent a cellular cost and should therefore be strictly controlled to maximize cell fitness [170]. Therefore, the central thought is that the mRNA and protein levels of a gene are correlated at the population level. Nonetheless, most studies have reported a poor correlation between these two entities, with a gene-specific correlation rather than a general trend [171]. However, studies often contradict themselves due to differential methodological approaches [172].

In the field of gene expression, live-cell reporter assays only measure mRNA or proteins and use it as a proxy for promoter output. In the past years, Tanenbaum and colleagues have developed a combined reporter enabling the visualization of mRNA transcription and polypeptide synthesis from a transcript in live single cells [173]. This assay relies on the coupling of the PP7 and the SunTag systems [174]. The later functions like a PP7 system but at the protein level, with a DNA-encoded array of peptide epitopes downstream the promoter of interest, recognized by a constitutively expressed single-chain antibody fused to a Fluorescent Protein (FP). Upon transcription, the PP7 signal will appear in the host nuclei and

later on in the cytoplasm, where a second signal will appear as a PP7-colocalizing dot upon translation of the SunTag binding sites [174]. Since translation occurs in the entire cytoplasm, following a single transcript requires its plasma membrane targeting to restrict the diffusion to two dimensions and enable time-lapse recording [174]. Thus, although this coupled reporter enables to measure crucial data on translational kinetics and posttranscriptional regulation, it does not allow to measure the total translational output and thus cannot be used as an accurate gene expression reporter. In addition, single-chain antibodies were to date not functionally expressed in a budding yeast cells and are thus restricted to their use in mammalian systems.

3.1.3 Aim of the project

In this project, we aimed at the simultaneous monitoring of transcription and translation arising from a single promoter in live single cells. This aim thus required dynamic measurements for both entities. At the transcriptional level, the best-established assays to monitor real-time mRNA production are the phage coat protein-based reporters, like the PP7 system. At the translational level, dynamic quantification of protein synthesis can be achieved with a fast-maturing fluorescent protein like the superfolder GFP (sfGFP) or a dPSTR reporter [144]. Because the relocation of the dPSTR is faster than the sfGFP maturation time and because its stability is decreased thanks to its degradation tag, the dPSTR assay appeared as the most dynamic readout. We thus aimed at developing a coupled PP7-dPSTR gene expression reporter assay. All intermediate systems and strains not described in this study are listed in Annex 1, at the very end of the manuscript.

3.2 Results: assay development

3.2.1 Stem loops and coat protein binding effects on transcript translation

To first assess the effect of the twenty-four repeated PP7 stem-loops on downstream gene product synthesis, we designed a simplified version of the desired coupled mRNA-protein reporter by replacing the dPSTR with a Venus translational reporter [135]. We then increased in complexity and moved from a Venus reporter to a stable dPSTR, to finally an unstable or dynamics dPSTR assay [144].

Stem loops and translation

Translation of mRNA into proteins is influenced by the mRNA structure. Indeed, it was shown that the presence of a hairpin in the 5'UTR of a yeast transcript was decreasing its translation [175]. This extent of the negative effect was shown to be dependent on the sequence from the loop and its distance from the ribosomal START codon [175, 176]. Secondary structures in the mRNA were then used to tune the translational level of transcripts [176].

The PP7 system, like the other aptamer-based RNA imaging reporter assays, is based on an array of stem-loops [88]. The higher the loops' number, the better the signal-to-noise ratio from the PP7-tagged transcripts, the better the detection [162]. In addition to the number of loops, the signal can be increased thanks to a better labelling of the mRNAs with a higher affinity of the phage coat protein for its binding sites, by changing the loop sequence [177] or by choosing the strongest coat protein-stem-loop affinity couple (PP7 over the MS2).

Results

Since the array of PP7 binding sites is composed of ATG and STOP codons, nothing downstream can be translated. Therefore, we placed the Venus protein upstream of the PP7sl cassette in our constructs (Figure 9A). To assess the effect of the stem-loops' presence on a transcript, we built two variants of the initial Venus-PP7sl construct: a loop-free version with only the Venus reporter protein, and a construct with a fifty base pair spacer between the Venus STOP codon and the first PP7 stem-loop, as distance between these two entities has been shown to influence the translational output (Figure 9A) [175, 176]. All the constructions were placed under the control of a salt inducible *STL1* promoter and integrated into the 5'UTR of the *GLT1* Open Reading Frame (ORF), as previously published [88, 144], to standardize the biological and experimental systems between the different constructions (Figure 9A). A wild type *S. cerevisiae* strain bearing a histone-tagged with cyan FP (CFP) was used as carrier for all the constructions, to allow for an automated segmentation and tracking of the cells and their nuclei through YeastQuant [178] Strains and plasmids are listed at the end of the chapter (Table 1 and Table 2). To test the effect of the PP7-bound and unbound PP7 stem-loops on the tagged transcript, strains were further transformed with a Single Integration Vector plasmid (pSIV) [168] expressing a PP7-mCherry allele (Figure 9, B to D).

As shown in Figure 9B, although the twenty-four PP7 stems-loops were placed downstream the Venus ORF, their presence led to a dramatic reduction in the Venus translation, with a two-third decrease in the final outputs. The addition of a fifty base pair spacer did not moderate this negative effect (Figure 9B). Interestingly, expression of the coat proteins restored the translation of the PP7sl labelled transcripts to even better outputs than the stem-loops-free variants strains, highlighting a stabilization of the mRNA through the binding of the

loops by the coat protein (Figure 9, B and C). Surprisingly, expression of PP7p in the stem-loop-free construct strain also led to a mild but significant increase in the Venus production (Figure 9, B and C). This could highlight a partial digestion of the loops during the cloning of the construct or to an increased stress sensitivity from the cells expressing the phage coat protein. Note that the PP7-RFP allele used in this chapter still has the ability to form viral particles inside the yeast cells, since it was not yet truncated for the coat protein oligomerization domain [154], unlike the PP7ΔFG-GFP variant used in the final PP7-dPSTR construct and in chapter four.

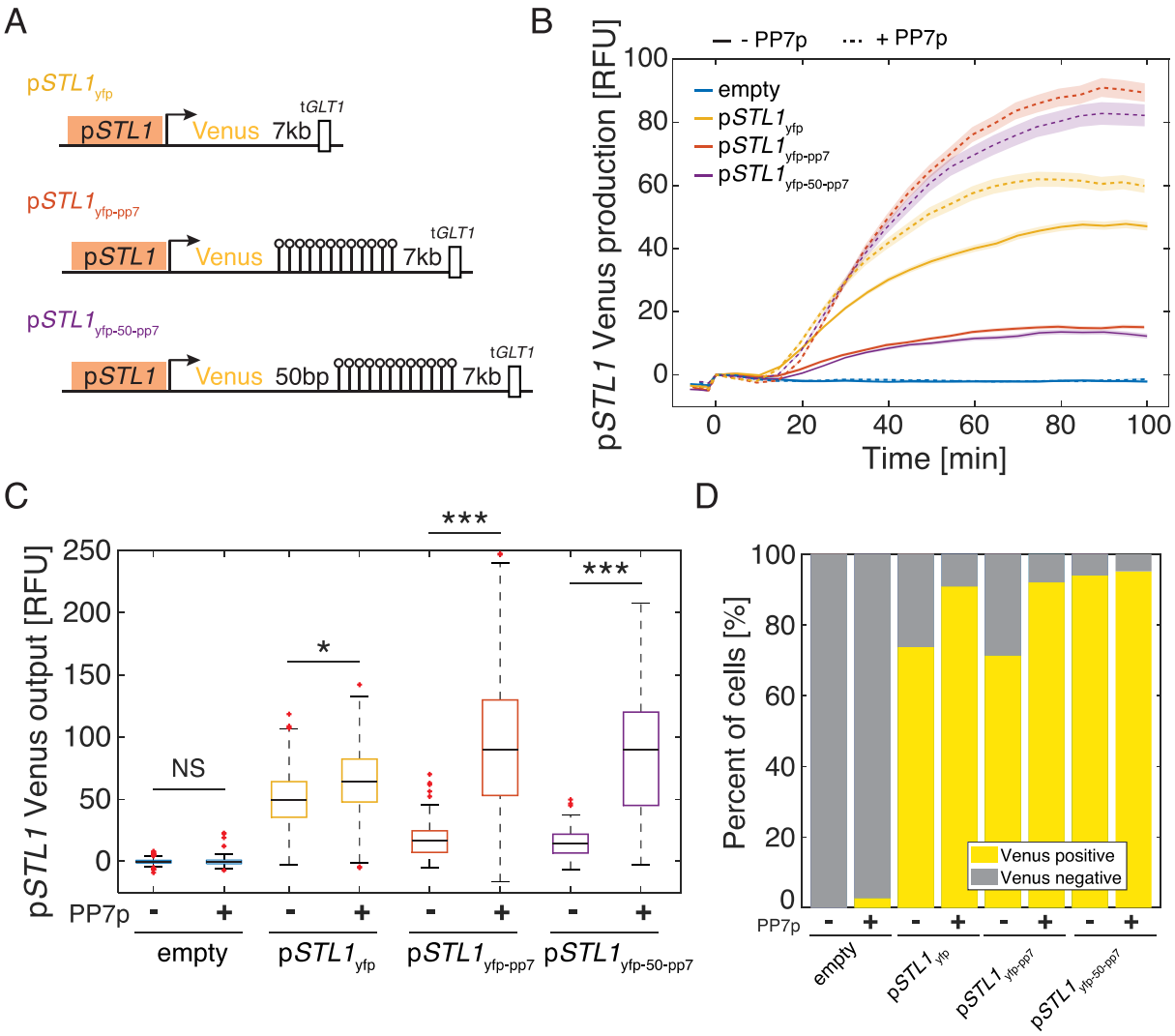


Figure 9: Tagging a transcript with PP7sl in 3'UTR decreases its translation but is rescued by the expression of the coat CP. (A) Schematics representing the three different constructions tested: a Venus (p*STL1*$_{yfp}$), a Venus and twenty-four PP7 stem-loops (p*STL1*$_{yfp\text{-}PP7}$), or a Venus separated by a fifty-base pair (50bp) spacer from the twenty-four PP7 stem-loops (p*STL1*$_{yfp\text{-}50\text{-}PP7}$). All the constructions are expressed under the control of an osmostress inducible *STL1* promoter and integrated in the

5'UTR of the *GLT1* gene. (B) The Venus production was monitored over time after addition of 0.2M NaCl at time zero. Strains harboring a histone tagged CFP ("*empty*") and expressing the different constructions were imaged for 100 min after a 0.2M NaCl induction at tp3, corresponding to time zero (full lines). Each single-cell trace was corrected for its background fluorescence and cell shrinkage-induced fluorescence increase by subtracting the first value after the timepoint of inducer addition (tp4), defined as time zero. Strains were imaged similarly after transformation and expression of the PP7-mCherry allele to test the effect of the PP7 bound and unbound stem-loops on Venus production (dashed lines). N>140 cells (C) Distribution of the maxima of the smoothed single-cell traces extracted from B. Statistically different Venus outputs were determined with a two-samples *t*-test with a 95% confidence interval. (D) Percent of Venus positive cells from the experiment in B, with the expression threshold allowing 5% of Venus positive cells in the empty strain with the coat protein.

3.2.2 Coupled PP7-Venus reporter

A long-standing question in the field of gene expression is the correlation between a gene's transcripts and proteins levels, which are supposed to be correlated in accordance with the central dogma of DNA>mRNA>proteins. To this hypothesis, we used a strain expressing the p*STL1*-Venus-PP7sl construct and a PP7-mCherry to record both the mRNA and proteins production arising from a single promoter, the osmoresponsive p*STL1* (yVW447, Figure 10). Because the *STL1* promoter has no basal expression level and is highly and rapidly induced upon NaCl treatment, it is an ideal candidate for quantitative gene expression studies [77, 144].

Results

As shown in the Figure 10, induction of cells expressing the p*STL1* PP7 Venus reporter with 0.2M NaCl gave rise to a peak of mRNA production in the fifteen minutes following the stress at the population level (Figure 10, A and B). Venus fluorescence was detected from the same population from 30 minutes after stress, due to its maturation time, and quantified as final level after 120 minutes (Figure 10, A and B). For the PP7 signal, the Transcription Site (TS) fluorescence was plotted as the 10 brightest RFP pixels in each cell's expended nucleus, which

corresponds to a five pixels augmented nucleus, that were divided for their whole cell median RFP expression and subtracted for their post-shrink level, to correct for bleaching and basal levels (Figure 10B). For the Venus, the whole cell median yellow fluorescence subtracted for each cell's basal level is plotted. As shown in Figure 10B, we observed an increased in both reporters' readouts during the time course of the experiments that were induced with salt, which are absent from the unstimulated control.

To compare the reporters'outputs in a quantitative manner for all the single cells, we defined the "p*STL1* PP7 output" and "p*STL1* Venus output" parameters. To do so, we smoothed with a moving average the Venus single-cell traces that were corrected for the background and shrink, and extracted each cell's maximum Venus fluorescence (Figure 10, B to D). The transcriptional output on the other hand was defined as each cell's sum of segmented PP7 foci (ConnectedHiPix) fluorescences (see Material and Methods at the end of the chapter). As expected for the *STL1* promoter, no Venus signal was detected under basal conditions and increasing salt concentrations led to a gradual increase in Venus production from p*STL1* (Figure 10C) [77].

To assess the correlation between the mRNA and protein outputs from the p*STL1* reporter assay, p*STL1* Venus and PP7 outputs were compared at the single-cell level (Figure 10D). To calculate a correlation coefficient, only the cell with both signals were considered. We therefore identified expressing versus non-expressing cells for each assay. For the PP7 system, expressing cells were identified by the segmentation of a PP7 focus. For the Venus reporter, positive cells were defined with an arbitrary threshold of 3.15 RFU, which allowed for 5% of Venus positive cells in the non-induced control (Figure 10E, SDfull). An R of 0.39 was calculated

for the 0.1M NaCl data, which decreased to 0.06 at 0.2M NaCl, suggesting a poor correlation between p*STL1* mRNA and protein outputs. Since this low correlation could arise from a poor detection of one or both signals, we analyzed the data further and split each experiment's entire cell population into four sub-populations, corresponding to all the possible combinations of PP7 and Venus signals' occurrences (Figure 10E).

As shown in Figure 10E, the biggest sub-population consisted in cells positive for both signals, which suggest a good detection and appropriate imaging conditions in general. As expected, only few cells showed a PP7 signal without a Venus signal (Figure 10E, PP7+Venus-). Indeed, since the Venus signal is in the whole cell, its measurement only requires a single plane imaging and is quite accurate due to the size of the object quantified. Unfortunately, we detected 20% of the population of cells with a Venus fluorescence but no corresponding PP7 signal at 0.1M NaCl (Figure 10E, PP7-Venus+). This population then decreased to 8% at 0.2M, which strongly suggests that the lack of detection came from lower transcriptional output and thus PP7 fluorescence (Figure 10E). Indeed, when differentiating the Venus signal from the PP7 positive and negative cells for both induced experiments, we observed a strong bias for low Venus fluorescence for the PP7 signal-less cells, which is more likely to be represented by low intensity PP7 readout cells, despite the poor general correlation between the two metrics (Figure 10F).

A possible explanation for this lack of correspondence could thus be the imaging conditions, either with a too low illumination power or z-distance coverage. Since RFP are in general less bright than GFP, fluorophore switch may improve the detection. Indeed, budding yeast cells have a higher RFP background than GFP, which renders the signal-to-noise ratio lower for this

RFP version of the system. Therefore, in the following coupled reporter assays, we switched to a GFP variant of the PP7.

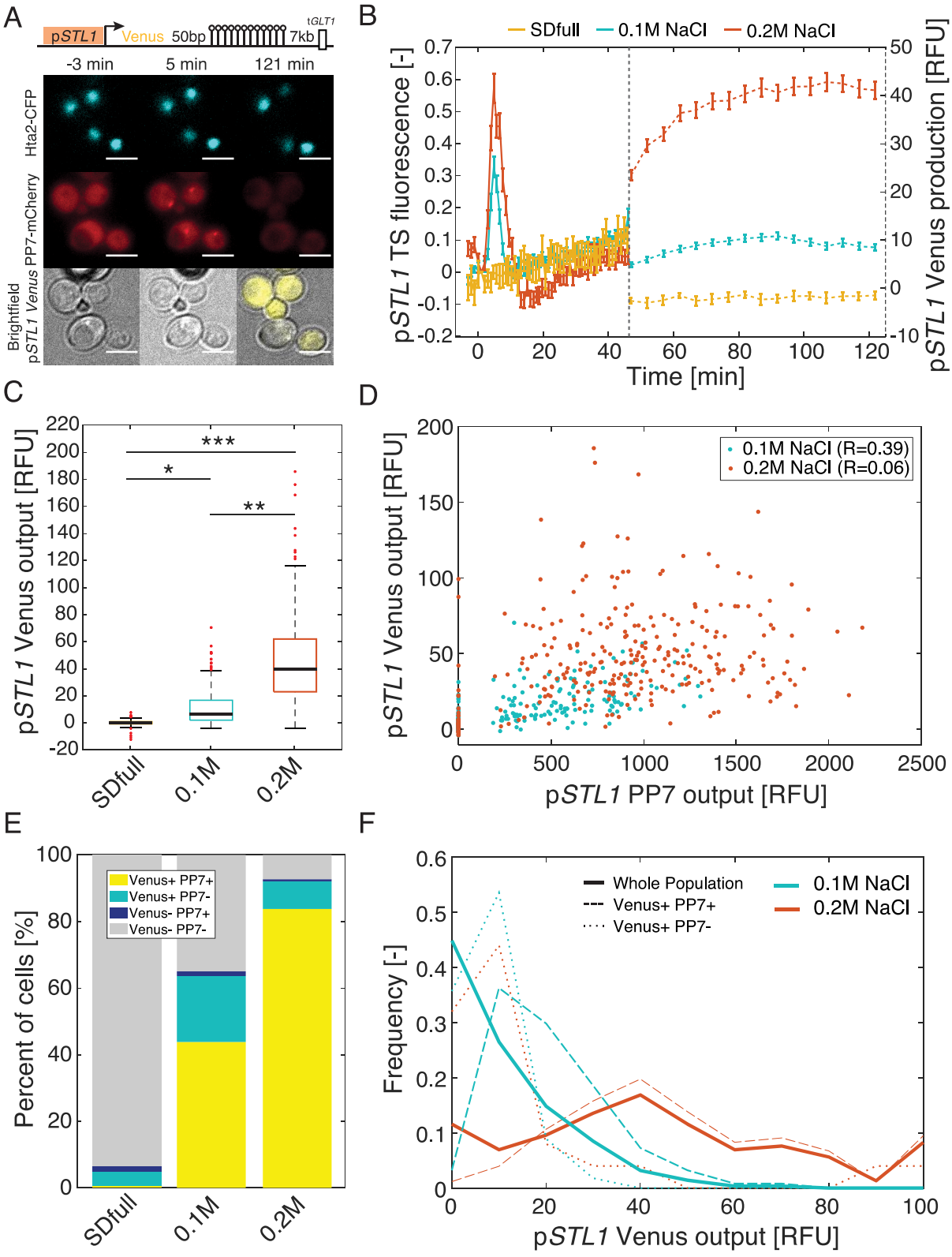


Figure 10: Single-promoter measurement of mRNA and protein outputs.
(A) Microscopy images from a strain with a histone-tagged CFP expressing a p*STL1-Venus-50bp-24xPP7sl* construction schematized above and integrated at the *GLT1* locus. Cells were stressed with 0.2M NaCl at time zero and imaged for 120 minutes. (B) Experiment from A was performed along with a 0.1M NaCl and a non-induced experiment. The PP7 signal is plotted as the ratio between the 10 brightest pixels over the RFP background to correct for bleaching and subtracted for the post inducer timepoint level, to correct for shrinking artifacts (full lines). The Venus signal corresponds to the cell median YFP level subtracted for each cell's signal at first timepoint after salt addition as well (dashed lines). The Venus LED was turned off after the first timepoint post-induction for 40 minutes to minimize bleaching (see Methods at the end of chapter) (C) Boxplots of the maxima Venus signal for each single-cell smoothed trace. Statistically different means were tested with a two-samples *t*-test with a 95% confidence interval. *P*-values for *= $6x10^{-31}$, **= $2x10^{-46}$ and ***= $1x10^{-71}$. (D) Scatter plot of the maxima Venus signal plotted in C against the sum of segmented PP7 signal for all the single cells. (E) Bar plot of each experiment's cell population composition for the PP7 and Venus signal occurrences. (F) Histograms of the Venus production for the induced experiments, differentiated between cells with (dashed line) and without (dotted line) a corresponding PP7 signal, and the whole Venus positive population as reference (full line).

3.2.3 Coupled PP7-dPSTR reporter

The coupled PP7-Venus reporter enabled us to get the mRNA and protein outputs from a single promoter in single cells. However, only the PP7 moiety displays a live readout, the Venus protein requiring a maturation time to become fluorescent [135]. To improve this coupled reporter, we swapped the Venus FP for an almost live gene expression reporter assay, the dPSTR [144]. Indeed, because it relies on the relocation of an already matured fluorescent protein, this reporter enables to bypass this downside of FPs. This assay was originally published in two different versions: with and without a degradation tag [144, 146]. In absence of this tag, the dPSTR accumulates in the nucleus in a protein expression-dependent manner and remains stable, enabling precise measurement of translational output [144]. With the degradation tag, the system is in a constant equilibrium between synthesis and degradation, therefore as soon as transcriptional activity ceases, the NLS-SZ1/mCherry-SZ2 heterodimers are quickly degraded, enabling dynamic measurements of promoter activation and deactivation [144].

In addition to the translational reporter improvement, we shifted from a PP7-RFP to a PP7-GFP allele to increase the transcription site brightness and to decrease the probability of having a protein output without a PP7 signal like in the previous PP7-RFP/Venus reporter (Figure 10E). In accordance, we used an RFP variant of the dPSTR to combined with the PP7-GFP allele and a histone-tagged with a tandem-dimer infrared RFP for spectral compatibility [179].

Development challenge

The challenge of this project was to couple both reporters' inducible parts: the PP7 stem-loops and the dPSTR peptide responsible for the promoter specific relocation of the fluorescent protein (Figure 11). The first concern was the impact of the PP7 stem-loops on the dPSTR translation, which we have addressed in our previous PP7-Venus reporter settings. The second concern was the impact of the total length from each transcriptional units' addition on the dPSTR readouts (Figure 11). Indeed, upstream PP7 stem-loops sequences supposedly do not affect the system's labelling.

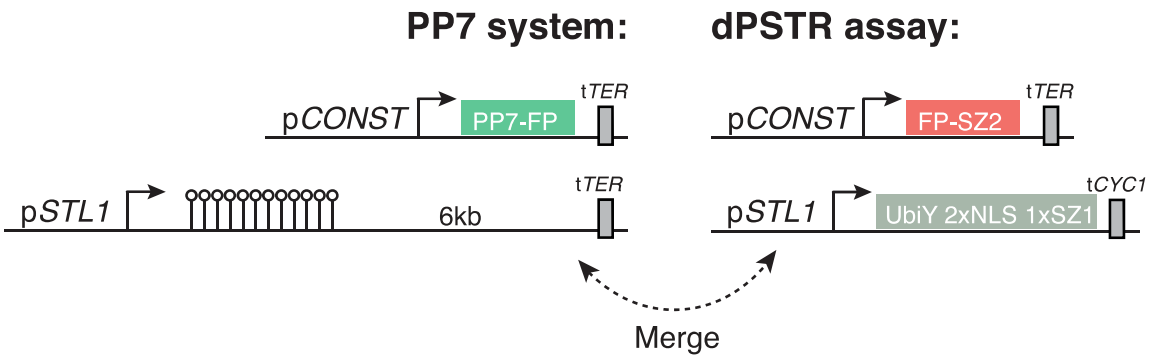


Figure 11: Merging the PP7 and dPSTR transcriptional units.

Both the PP7 and the dPSTR reporter are bipartite assays, composed of a constitutively expressed (p*CONST*) transcriptional unit and an inducible promoter-specific transcriptional unit, here under the control of an *STL1* promoter (p*STL1*). The PP7 system consists in a constitutively expressed PP7ΔFG protein fused to a fluorescent protein (PP7-FP) and the dPSTR in a fluorescent protein fused to a SynZip 2 (FP-SZ2). The PP7 promoter-specific transcriptional unit is composed of an array of

24xPP7sl and a downstream transcribed but not translated ORF, here the 6kb represent the length of the *GLT1* gene, integration site of our original PP7 construct. The dPSTR promoter-specific transcriptional unit is composed of a degradation tag (UbiY), two Nuclear Localization Sequences (2xNLS) and a SynZip 1 (SZ1) and a weak *CYC1* terminator. Creation of a coupled pSTL1-PP7-dPSTR requires the merging of the two inducible transcriptional units of the two systems (arrows).

Reporter development

Exchange of the Venus protein for the stable dPSTR promoter-specific moiety and integration at the *GLT1* ORF as performed for the PP7-Venus coupled reporter, did not lead to any nuclear accumulation of the mCherry-SZ2 part upon salt addition (data not shown). Two hypotheses were drawn from this preliminary result: first, the stabilization of the transcript by the binding of the PP7 stem-loops by the coat proteins was not sufficient to increase the signal-to-noise ratio and second, the length of the transcript is affecting the dPSTR relocation ability. Since these two hypotheses could potentially be linked, we designed different synthetic coupled reporter constructs and scored their abilities to relocate the fluorescent moiety during a salt challenge time-lapse experiment to assess the source of the dPSTR loss-of-function (Figure 12). Strains and plasmids described here are listed at the end of the chapter. Additional strains resulting from the developmental phase of the coupled reporters can be found in Annex 1.

Figure 12 summarizes the results from all the variants designed and tested. As expected, the constructs with the best and worst relocation abilities were the original dPSTR and the Venus swapped constructs respectively (Figure 12, n°1 and n°8). The addition of a spacer in between the dPSTR STOP codon and the first PP7 hairpin had a small beneficial effect, highlighting a putative sterical hindrance from the loops on the translational machinery (Figure 12, n°2 and n°3, n°5 and n°6). However, this was not sufficient to recover a relocation from the reporter. Decreasing the downstream DNA tail of the PP7sl, on the other hand, led to a greater nuclear accumulation of the dPSTR. These results either suggested a transcript length effect on the

dPSTR transcriptional or translational rate, or a sequence dictated increased stability of the transcript (Figure 12, n°6 to n°8). These two hypotheses could also explain the constructions lacking the PP7 stem-loops or the one where they are replaced by a 3kb long downstream transcribed tail, both of which were in the top scores (Figure 12, n°1 and n°4). Indeed, a lower dPSTR production rate might not lead to a detectable nuclear signal accumulation, due to the imaging conditions or to bleaching.

The hypothesis of the decreased stability seemed, however, to account for much of the defect since the exchange of the *CYC1* for the *SIF2* terminator led to a better nuclear signal (Figure 12, n°3 and n°6) [180]. By combining all the positive effect changes, we designed the best score construction, which bears a 50bp-spacer, a 1kb downstream transcribed DNA and a strong *SIF2* terminator (Figure 12, n°2).

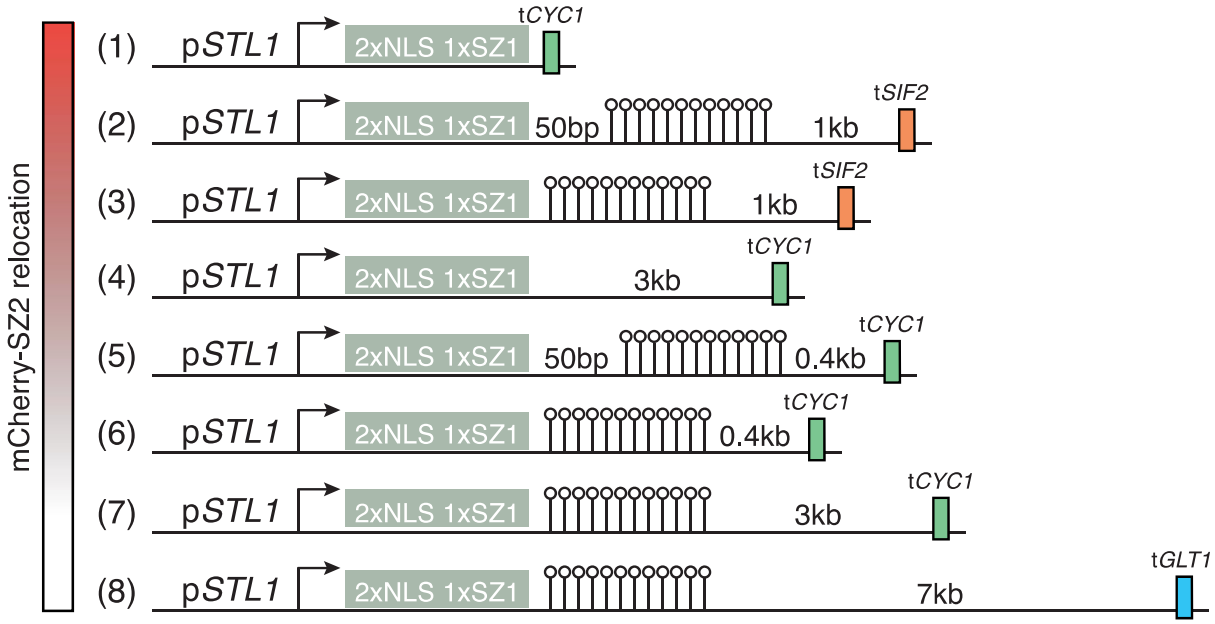


Figure 12: Summary of the coupled pSTL1-PP7-dPSTR reporter constructs.

Different variants of the combination of the two systems were designed and assessed for their relocation scores during an 0.4M NaCl time-lapse experiment. The constructs are here ranked depending on their relocation abilities relative to the dPSTR construct alone (n°1). As shown here, the best synthetic system was the dPSTR alone (1) and the worst was the coupling of dPSTR and PP7 at the PP7 usual integration site, the 7kb long *GLT1* gene (8). Variants were originally tested in a strain

expressing a dPSTR CFP, before the PP7-GFP allele shift, but for clarity with the rest of the chapter, mCherry is indicated here. Strains and plasmids are listed at the end of the chapter.

Coupled PP7-dPSTR reporter with stable dPSTR

In a final effort to further improve the dPSTR signal from the coupled reporter and in the optic of upgrading to an unstable allele, the ratio between the interacting partners SynZip 1 and SynZip 2 was biased toward SynZip 2, to increase the dPSTR signal-to-noise ratio (Figure 13A). For each synthetized NLS-NLS-SZ1 peptide, four mCherry-SZ2 peptides could potentially bind, instead of one in the original construct [144]. In addition, the Nuclear Localization Signals (NLS) and their linkers were mutated to alanine for all negatively charged residues and possible phosphorylation sites that might decrease the nuclear enrichment and which may be influenced by cellular state, like the cell-cycle or glucose levels (Figure 13A, npNLS: non-phosphorylatable NLS) [181].

dPSTR 2xNLS-linker: MRSEPKKKRKVGAGAEPKKKRKVGGSSVDGGSNQTSLYKKAGSAAAPFTMEFK

npNLS and linker: MRAAPKKKRKVGAGAAPKKKRKVGGAAVAGGANQAALYKKAGAAAAPFAMAFK

As shown in Figure 13, we observed nicely quantifiable signals from both moieties of the coupled reporter in response to various salt concentrations and flat signals from both in absence of induction (Figure 13, B and C). The p*STL1* transcription site fluorescence showed a transient increase in fluorescence, proportional to the stress level, as quantified by the intensity of the 10 brightest pixels subtracted for their median cell GFP and post-induction levels (Figure 13B). The p*STL1*-dPSTR, here in its stable version, led to a signal increase and then a plateau, proportional to the stress as well (Figure 13C) [144]. A small decrease in the final dPSTR readout was actually observed, which was attributed to the 40% of bleaching observed. Thanks to the tuning of the expression levels of both fluorescent proteins, we

reached a sensitivity range that enabled us to measure responses from 0.1M to 0.3M salt stress, which correspond to our usual working concentrations.

We next defined the cells with positive signals from both, either or none of the reporters. For the PP7 system, expressing cells were defined by the segmentation of a PP7 focus, as done previously. The time at which the first dot was segmented was further defined as the "PP7 response time" of that particular cell (Figure 13D). This enabled us to bypass the need for arbitrary threshold definition. For the dPSTR assay, on the contrary, we defined an arbitrary threshold at 0.45 of nuclear over cytoplasmic fluorescences ratio, subtracted for their value at the first time-point after induction for each single-cell trace, to remove artifacts from cell shrinkage (Figure 13C, dashed line). This threshold was set to have 5% of dPSTR positive cells under non-stressful conditions, as done for the PP7-Venus reporter. The first timepoint at which a single-cell overcame that threshold was defined as its response time (Figure 13E).

As expected from our previous data on the *STL1* promoter, we observed different response times for the different salt concentrations (Figure 13, D and E) [144]. Indeed, with increasing stress level, the cell shrinkage decreases molecule diffusion through molecular crowding, which causes a delay in transcription activation [182]. When quantifying the time difference between both signals' appearances at the single-cell level, we observed some differences between the different salt concentrations experiments (Figure 13F). Indeed, we observed a mean of 8'30'', 9'35'' and 11'15'' time difference between both readouts at 0.1M, 0.2M and 0.3M respectively. This could either highlight a differential effect of cell shrinkage or bleaching, which would impact later response times or in this case, higher salt concentrations experiments. However, more replicate experiments should be performed in order to confirm the significance of these discrepancies. On average, we estimated a ten minutes delay

between transcription initiation (appearance of PP7 focus) and properly folded proteins (accumulation of nuclear dPSTR over expression threshold), which includes transcription elongation, mRNA export and translation, and protein folding.

Using these same activation thresholds, we defined the PP7 and/or dPSTR positive and negative cells. As shown in Figure 13G, we confirmed the bimodal behavior of p*STL1*, with 55% of responding cells (PP7+ dPSTR+) at 0.1M NaCl, which was further increased to at least 85% at higher concentrations [77]. To test our ability to detect both signals simultaneously, we extracted the sub-population of cells which displayed only one of the two signals. Similarly to our previous PP7-Venus coupled reporter, we observed a small population of cells positive for the translational reporter and lacking a transcriptional readout (Figure 13G, PP7- dPSTR+). This cluster represented 8% of the total cell population at 0.1M NaCl. Interestingly, this sub-population almost disappeared at higher salt stresses, which suggested an intensity-driven lack of segmentation, as previously. This hypothesis was indeed confirmed when comparing the dPSTR signal of these PP7 negative cells, for which the average maximum intensity was lower than the whole dPSTR positive cell population average (Figure 14A).

Finally, a third population suggested a too stringent gene expression threshold on the dPSTR signal (Figure 13G, PP7+ dPSTR-), which could explain the large time-difference calculated between both signal (Figure 13F). Indeed, segregation of the PP7 signal of this sub-population versus the entire PP7 positive one depicted a tendency for lower intensity foci (Figure 14B). However, unlike the previous populations, the proportion was lower at low salt stress, or even inexistent in absence of stress, and a mild decrease was observed with increasing

concentration of salt (Figure 13G). This is contradictory to a loose threshold and may suggest the presence of some post-transcriptional regulation of the p*STL1*-PP7-dPSTR transcript.

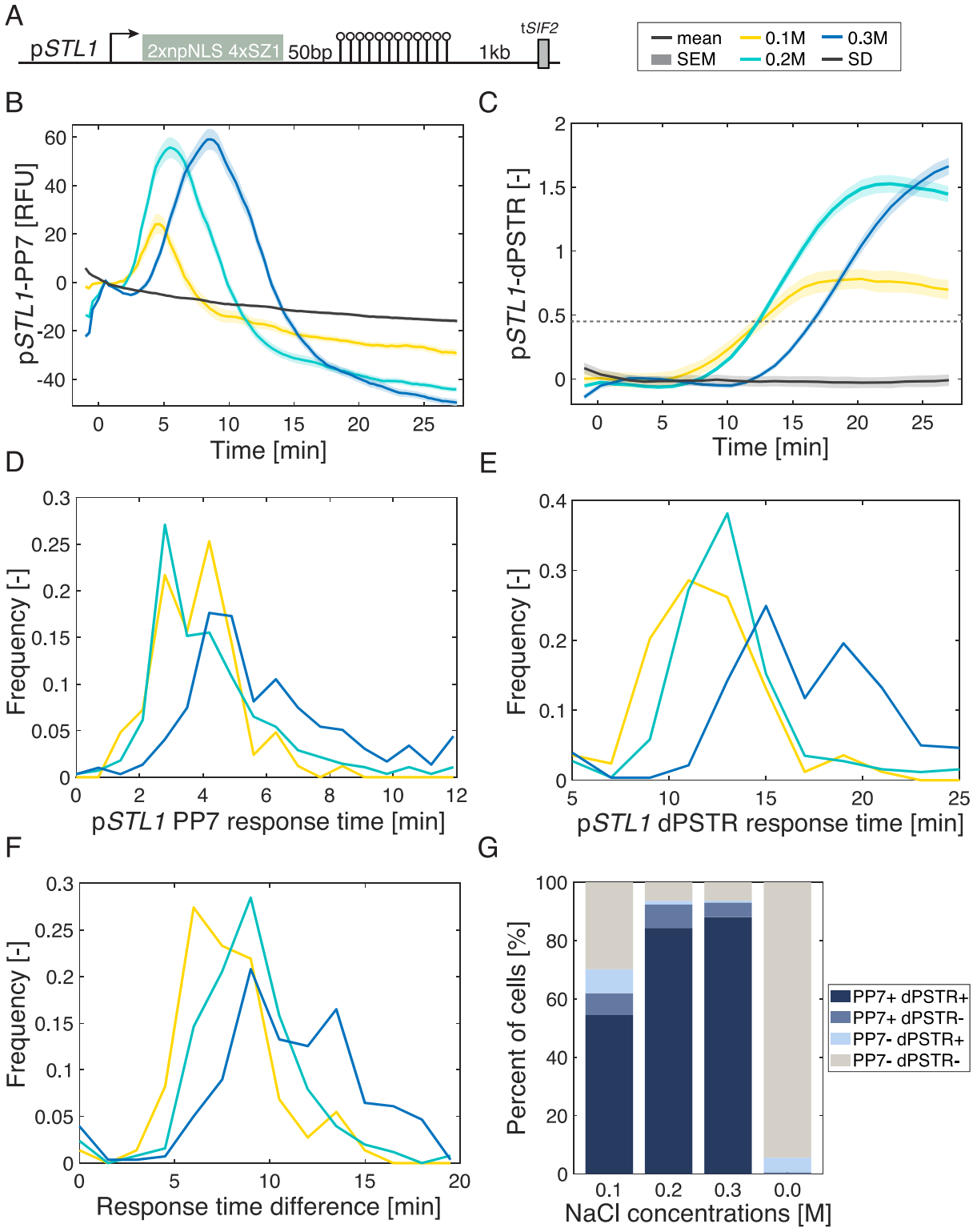


Figure 13: Coupled PP7-dPSTR stable.

(A) Schematic representation of the p*STL1*-dPSTR-PP7 coupled assay. The dPSTR part consists in two non-phosphorylatable Nuclear Localization Signals (npNLS) and four SynZip 1 (4xSZ1). The PP7 part consists in 24xPP7sl, 1kb of downstream DNA and a strong *SIF2* terminator. (B) p*STL1* transcription site fluorescence over time after osmotic stress (0.1M, 0.2M or 0.3M) or non-inducing conditions (SDfull) applied at time zero. Mean and SEM of the PP7 signal is plotted as the 10 brightest pixels in the expended nucleus, to which the average cell GFP background fluorescence and the first value after induction were subtracted, to normalize for background and cell shrinkage. Single-cell traces were smoothed with a moving average. (C) p*STL1* dPSTR nuclear over cytoplasmic fluorescence ratio corresponding to the cells in the experiments in B. Data were subtracted for the first value after induction to correct for cell shrinkage. Single-cell traces were smoothed with a moving average. Dashed line represents the expression threshold set to 0.45 to discriminate the responding cells and their response time, which is set to 5% of dPSTR positive cells in SD-full conditions. (D) Histograms of p*STL1* PP7sl response times calculated as the first timepoint were a PP7 dot was segmented for each single-cell (see Methods). Only the cells where a dot was segmented were considered (PP7+ cells). (E) Histograms of p*STL1* dPSTR response times corresponding to the overcome of the expression threshold set to 0.45 (dashed line in C). Only the dPSTR positive cells were plotted. (F) Histograms of the time difference between the PP7 and dPSTR signals at the single-cell level, for the PP7+ dPSTR+ cells. (G) Percent of cells considered as PP7 positives or negative, and dPSTR positive or negatives.

To test this hypothesis, we analyzed further the individual responses and assessed the correlation between the PP7 and dPSTR outputs at the single-cell level (Figure 14E). Like for the previous PP7-Venus coupled reporter assay (Figure 10D), we did not observe a correlation between p*STL1* transcriptional and translational outputs (Figure 14C).

Because transcription and translation outputs may not be correlated but both timings should, we assessed the correlation between p*STL1* PP7 and dPSTR response times (Figure 14D). Indeed, we observed a positive correlation between the timing of transcription and translation of the reporter assay, even with the previously calculated ten minutes delay between both readouts. These results suggested that, first, transcriptional level does not correlate with translational level, and second, that transcription initiation time does not dictate translational output. Extracted examples single-cell traces indeed illustrate the correlative trend in both systems' response times, but the poor of correlation from a quantitative point of view (Figure 14, E and F).

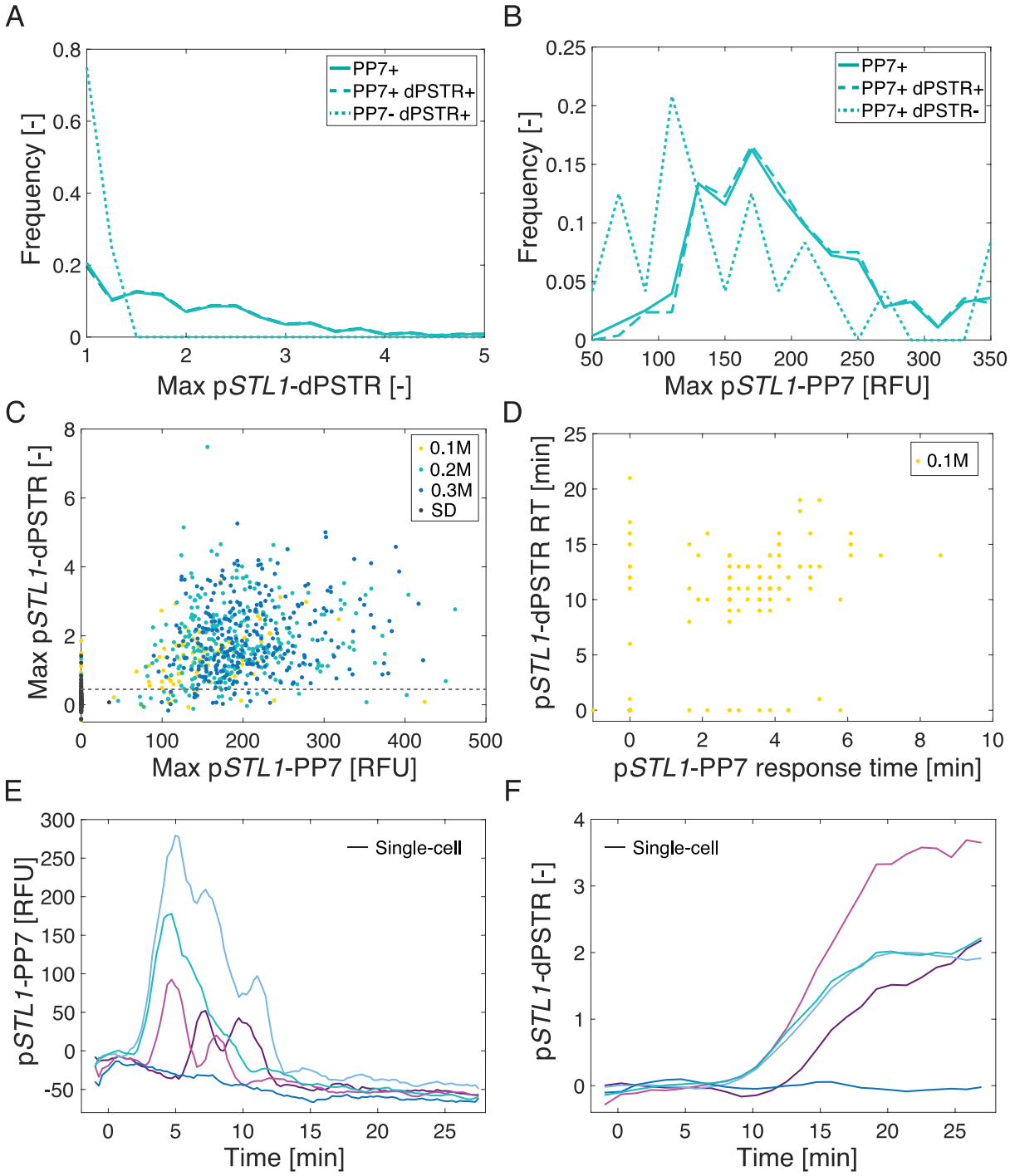


Figure 14: Single-cell analysis of the coupled pSTL1 PP7 dPSTR reporter.

(A) Histograms of the maximum signal of the dPSTR positive cells from the 0.2M NaCl experiment in Figure 13. From the total population (full line), two sub-populations are further defined as PP7 positive or negative based the segmentation of a PP7 focus. (B) Same analysis as done in A but for the PP7+ cell population and their dPSTR signal, with the 0.45 threshold for the dPSTR cells identification. (C) Scatter plot of the maximum intensity of ConnectedHiPix versus the maximum dPSTR accumulation for the experiments in Figure 13. Dashed line represents the 0.45 threshold set to define the dPSTR positive cells. (D) Scatter of the Responses Times (RT) of both readouts for the PP7+ dPSTR+ cells at 0.1M. (E) Five single-cell traces from the 0.2M NaCl experiment from Figure 13 were plotted as a function of time. The 10 brightest pixels minus the basal and post-induction levels are plotted here. (F) Corresponding dPSTR signal from the single-cells in C. Data represent the nucleus over cytoplasmic signal subtracted for the post-induction level.

3.3 Results: evaluation of the established coupled reporter

Coupled PP7-dPSTR reporter: unstable dPSTR

Because the final aim of the project was to develop an assay that enables to monitor promoter activity, including activation and deactivation, both at the transcriptional and translational levels, we added the degradation tag of the original dPSTR assay (Figure 15A and Figure 16) [144]. This tag consists in a leading Ubiquitin, which upon translation of the peptide is cleaved off and exposes the next amino acid, Tyrosine (Y). Based on the identity of this leading amino acid, different half-lives can be obtained, following the "N-end rule" [146]. In our case, the UbiY tag decreases the half-life to few minutes [144, 146]. Because both of our reporter strains were imaged identically, side-by-side comparison are feasible (Figure 15, dashed lines).

As shown in Figure 15 and 16, addition of the degradation tag did not affect the coupled reporter sensitivity; we detected signals for both systems from 0.1M to 0.3M NaCl salt stresses (Figure 15, B and C). As expected, addition of the degradation tag led to a transient dPSTR nuclear accumulation, now reflecting p*STL1* transient transcriptional activity (Figure 15C and Figure 16). We extracted the response times as done previously for the stable coupled reporter using the segmented PP7 foci and the 0.45 threshold on the dPSTR nuclear over cytoplasmic ratio corrected for the basal level (Figure 15C, dashed line). As shown in Figure 15D, the response time of the p*STL1*-PP7 was not affected by the dPSTR peptide modification; we found comparable response times for all concentrations (Figure 13D and Figure 15D). For the dPSTR part on the contrary, we found noticeable differences between the two strains. As expected from an unstable construct, we did not reach the same maxima and drop to approximately two-third of the stable allele's plateau (Figure 15C). However, although the

expression thresholds were kept identical between both analyses, we extracted significantly different response times, with about four minutes faster responses from the unstable construct for all concentrations tested (Figure 15E). These results suggested that the sequence modification caused a differential translational rate from the construct, either through a stabilization of the mRNA half-life, which would not affect the PP7 signal at the transcription site, or through ribosome elongation velocity directly. However, more replicate experiments should be performed in order to confirm these discrepancies.

Faster dPSTR signal accumulation led to a decrease time difference between the p*STL1* PP7 and dPSTR response times (Figure 15F). Interestingly, unlike the stable allele, the unstable reporter showed a constant stress level-independent time difference between both signals' appearances, as expected at these salt concentrations (Figure 15F). On average, we estimated a five minutes time requirement between transcription and protein maturation (Figure 15F).

As done previously, we divided the whole cell population of each experiment into positive and negative cells for both readouts (Figure 15G). As shown in Figure 15G, addition of the degradation tag increased the proportion of PP7+ dPSTR- cells. Indeed, since we lowered the maximum nuclear relocation of the dPSTR, low expressing cells were less likely to overcome the threshold set from the stable allele outputs. This was then rescue partially with increasing salt stress, which in turn increases the cells' response (Figure 15G), confirming that lower threshold would potentially detect more dPSTR positive cells. Together these data suggested a trade-off between the system dynamicity and the imaging conditions, whereby lower dynamics would require less heavy light illuminations and lower time resolution. This actually holds true for any type of reporters, whether it is a Venus, a dPSTR or a PP7 assay.

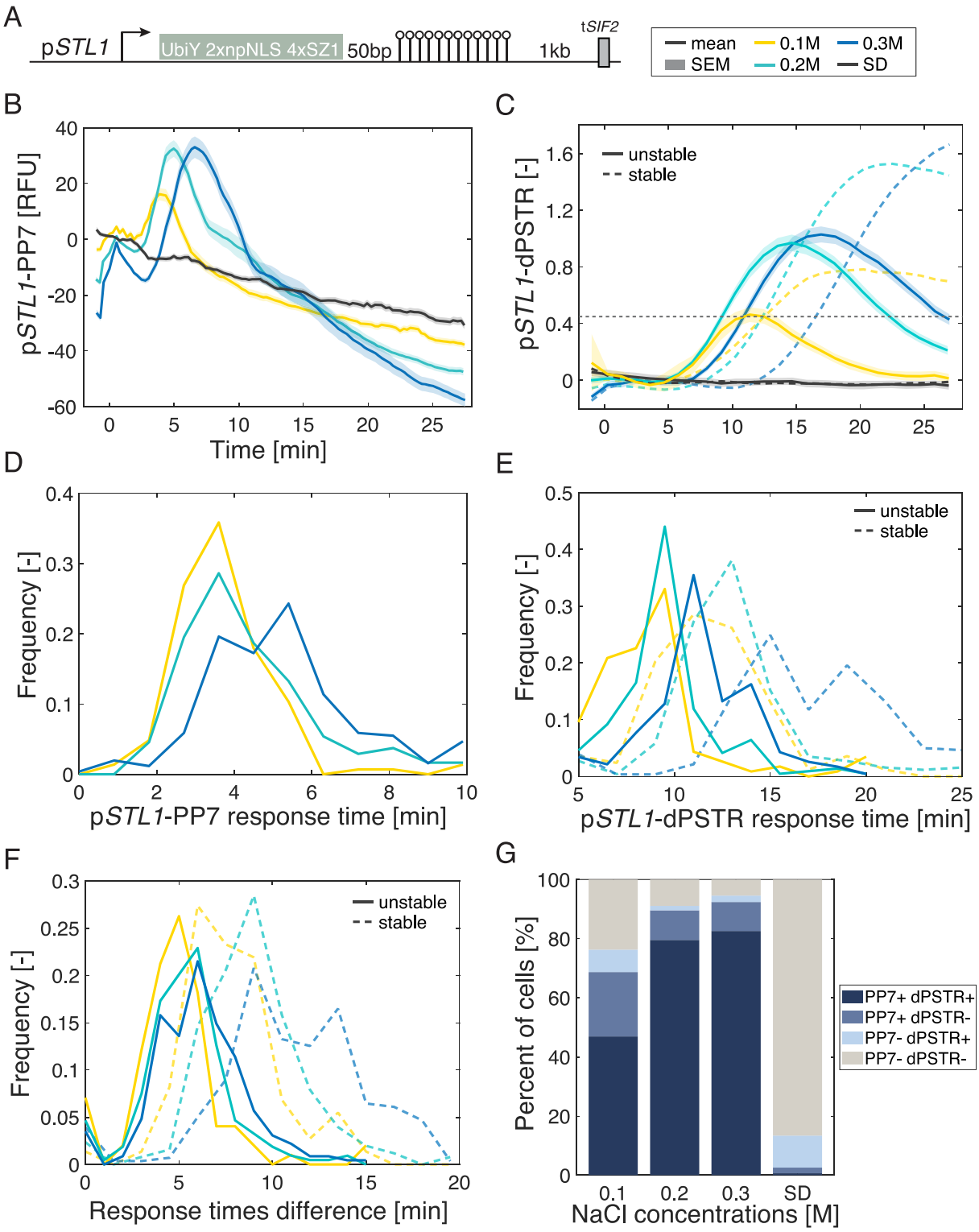


Figure 15: pSTL1 coupled PP7-dPSTR unstable.

(A) Schematic representation of the coupled p*STL1*-PP7-dPSTR. (B and C) Dose response population mean traces and SEM for the PP7 (B) and the corresponding dPSTR (C). Dashed colored lines represent the data from the stable construct on Figure 13. Single-cell traces were smoothed with a moving average. The horizontal grey dashed line corresponds to the 0.45 expression threshold used to determine the dPSTR positive cells and their response time. (D and E) Histograms of response times for the p*STL1* PP7 (D) and dPSTR (E) positive cells. Dashed lines represent the data from the stable allele on Figure 13. (F) Histograms of the time difference between the PP7 and dPSTR readouts at the single-cell levels, only the PP7+ dPSTR + cells were plotted.

Dashed lines represent the data from the stable construct on Figure 13. (G) Percentage of cells categorized as PP7 and/or dPSTR positive and negative based on the segmentation of a focus and overcome of 0.45 expression threshold respectively.

As shown in Figure 16, addition of the degradation tag did not lead to a better correlation between the p*STL1* PP7 and dPSTR outputs, compared to the stable coupled reporter allele (Figure 16). This was expected due to the decreased detection of dPSTR positive cells in this reporter settings.

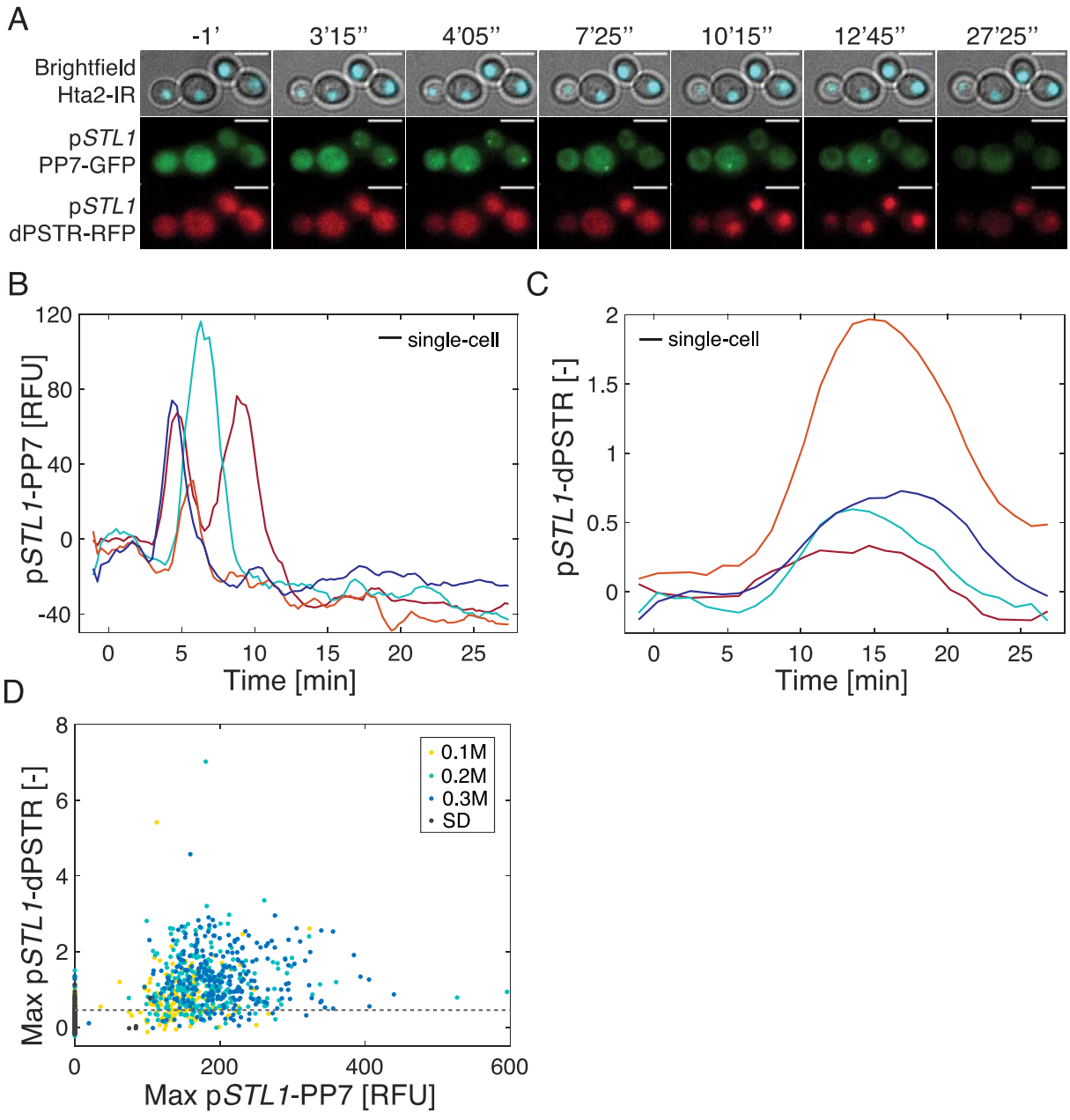


Figure 16: Single-cell traces of unstable coupled reporter.
(A) Representative images of the 0.2M NaCl experiment from Figure 15A. The yVW447 strain bears a histone tagged with CFP for nucleus segmentation and cell tracking and expresses a PP7ΔFG-GFP allele and an dPSTR-RFP reporting on p*STL1*. Cells were stressed with 0.2M NaCl at time zero and imaged every 15 seconds for the PP7 and every minute for the dPSTR, nucleus and brightfield images to 30 minutes. Scale bar represents five microns. (B and C) Five representatives single-cell traces from the same 0.2M NaCl experiment in Figure 15. Traces were smoothed with a moving average. (D) Scatter plot of the maximum intensity and maximum nuclear over cytoplasmic fluorescences ratio corrected for basal level from the experiment in Figure 15B and C. Dashed line represents the 0.45 expression threshold set for the p*STL1*-dPSTR.

Comparison between coupled and single dPSTR assay: unstable version

We have shown that our coupled reporter stable and unstable alleles behave differently (Figure 15). Next, we wanted to compare our coupled reporter dynamics to the original published dPSTR assay, therefore we compared the data from the published paper to our coupled reporter assay (Figure 17) [144]. Because both strains were imaged differently, we made general comparisons. As shown in Figure 17 panels A and B, both dPSTR-based assays produced transient responses to salt concentrations from 0.1 to 0.2M. Note that the coupled reporter is plotted as the ratio of the nuclear and cytoplasmic fluorescences (Figure 17A), whereas the single dPSTR is displayed as the difference. From the single-cell traces, we defined thresholds for both systems and extracted the response time and number of expressing cells. Both assays showed similar responses times for all the salt concentrations tested (Figure 17, C and D). These results suggested a similar dPSTR relocation ability in both reporter settings and putative similar sensitivity ranges. These will however have to be confirmed under identical experimental and imaging conditions. In addition, these results indicated that the levels of expression of the PP7 and dPSTR were suited to the detection of expression from the *STL1* promoter at 0.1M to 0.3M NaCl salt stresses. More generally, these results proposed that we built a functional variant of the original dPSTR assay, which can provide a transcriptional readout in addition to the translational, thanks to the coupling to a PP7 system.

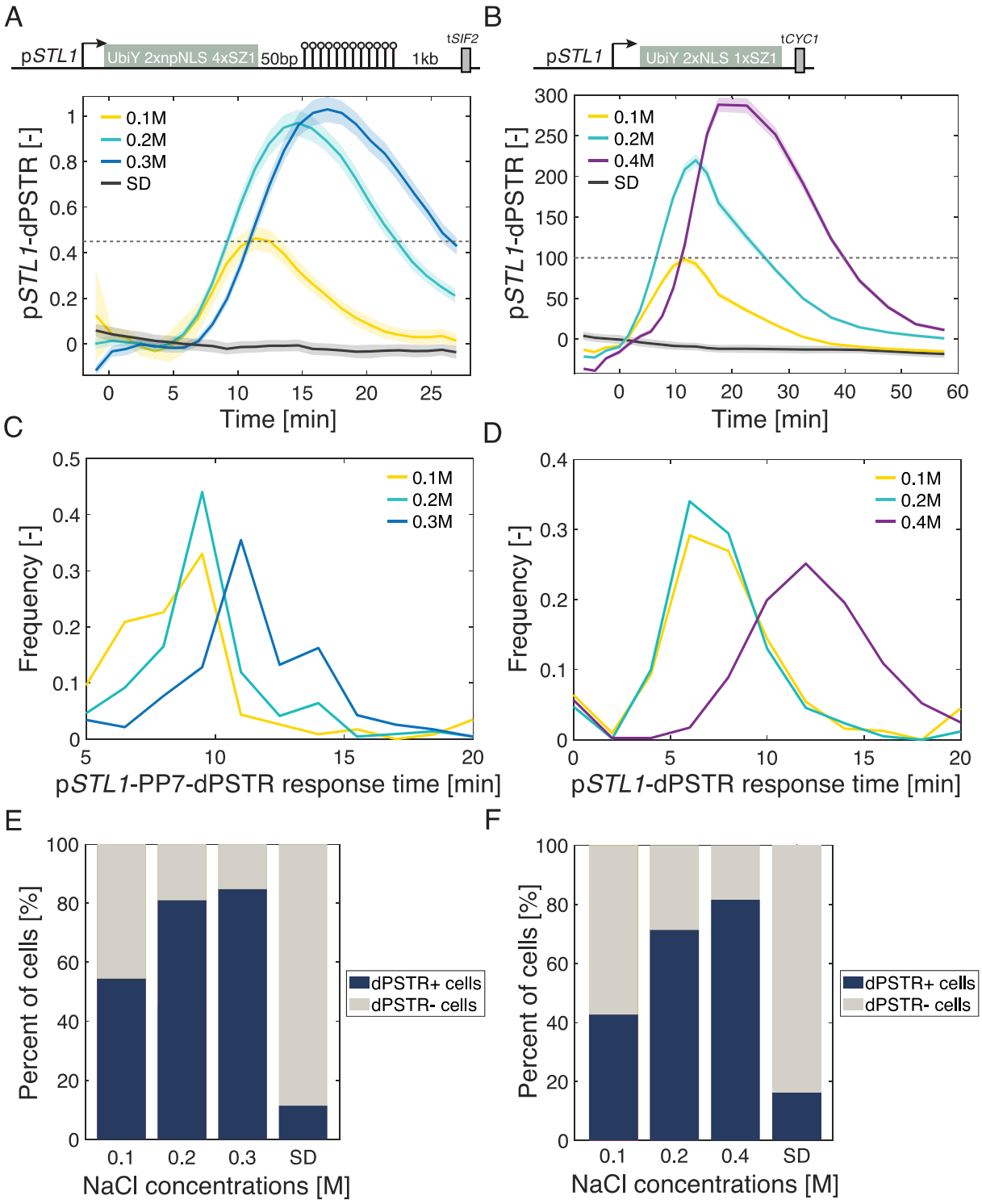


Figure 17: comparing single and PP7-coupled unstable dPSTR alleles.

(A) Schematic representation and dose response of cells expressing the coupled p*STL1*-PP7-dPSTR unstable allele in response to 0.1M, 0.2M and 0.3M NaCl induction. Dashed line represents the expression threshold at 0.45, set to identify the dPSTR positive cells and their response time. Single-cell traces were smoothed with a moving average. (B) Schematic representation and dose response of cells expressing the original p*STL1*-dPSTR unstable construct in response to 0.1M, 0.2M and 0.4M NaCl induction. Dashed line represents the expression threshold at 100 RFU set to identify the responding cells and their response time. Single-cell traces were smoothed with a moving average. Data originated from the dPSTR paper published in Aymoz *et al.* 2016. (C and D) Histograms of the response time of the dPSTR positive cells, defined as the first timepoint of the overcome

threshold. (E and F) Percentages of responding cells as defined with the expression threshold for the different reporter constructs, as marked with dashed lines in A and B.

Chemical uncoupling of PP7 and dPSTR

Although being coupled in a single reporter assay, the PP7 and dPSTR systems report on different biological processes [88, 144]. As part of the assay validation process, we tested whether we could affect translation without affecting transcription. To do so, we performed an experiment with a translational inhibitor, cycloheximide (CHX). This compound is naturally secreted by *Streptomyces griseus* bacteria and inhibits translation elongation through ribosomes binding and prevention of their elongation factor mediated-translocation [183]. In the following experiment, cells were treated for three minutes prior time-lapse imaging with 100x water-diluted cycloheximide at a final well concentration of 0.1mg/ml or just with SDfull medium, and then induced with a 0.2M NaCl step (Figure 18).

As shown in Figure 18, CHX treatment abolished p*STL1* dPSTR nuclear relocation to the untreated cells' levels, whereas maintaining a PP7 output (Figure 18, A and B). We confirmed the correct induction of the cells by plotting the mean cell area of the cells subtracted for their basal cell size, prior salt addition (Figure 18C). Similar shrinkages and adaptation times were observed for both the CHX- and SDfull-incubated cells, depicting a similar induction medium osmolarity and thus induction of the cells (Figure 18C). Using the previous 0.45 expressing threshold, we extracted the number of p*STL1*-dPSTR expressing cells. As shown in Figure 18D, the percentage of PP7+dPSTR+ cells dropped from 80% to 5% and was replaced by a dominant PP7+ dPSTR- sub-population (Figure 18D). These results indicated an uncoupling of the PP7 and dPSTR outputs. Indeed, when correlating the maxima of both readouts, data majoritarily

spread into a single dimension and depicted a complete lack of correlation (Figure 18E). Unlike previous analyzes on our coupled reporter assay, this population was constituted of equal PP7 intensity cells, to the level of the control experiment (Figure 18F). Therefore, there was no correlation between the PP7 intensity and the lack of dPSTR detection, unlike in our previous reporter assay.

Surprisingly, the p*STL1*-PP7 population response differed from the control experiment both in amplitude and duration, but not in response time (Figure 18A). These differences were observed in replicate experiments and could not be attributed to a different salt induction (Figure 18C). The small decrease in the population response's amplitude could be attributed to the small decrease in the number of PP7 expressing cells, since the maximum intensity did not vary (Figure 18F). The increase in population duration on the other hand, could partially be attributed to the lack Gpd1p synthesis upon normal salt stress, which contributes to the cell adaptation [53].

Known CHX side effects on cell fitness, like DNA damage, were ruled out due to the short incubation time and cell-cycle arrest generated by the following osmotic stress. Thus, two hypotheses were drawn: first, cycloheximide treatment is an additional stress, which thus causes a higher or longer response from stress-induced genes; second, CHX impacts the reporter in a stress-unspecific manner. Since we did not observe a change in the basal GFP level from the cells, we concluded that cycloheximide does not affect the properties of the fluorescent protein itself, unlike previously observed during ethanol treatment (data not shown). To confirm the first hypothesis on CHX induction of stress response, or draw an

alternative one, Hog1 and Msn2/4 relocation should be tested under these conditions, which was not assess at that time.

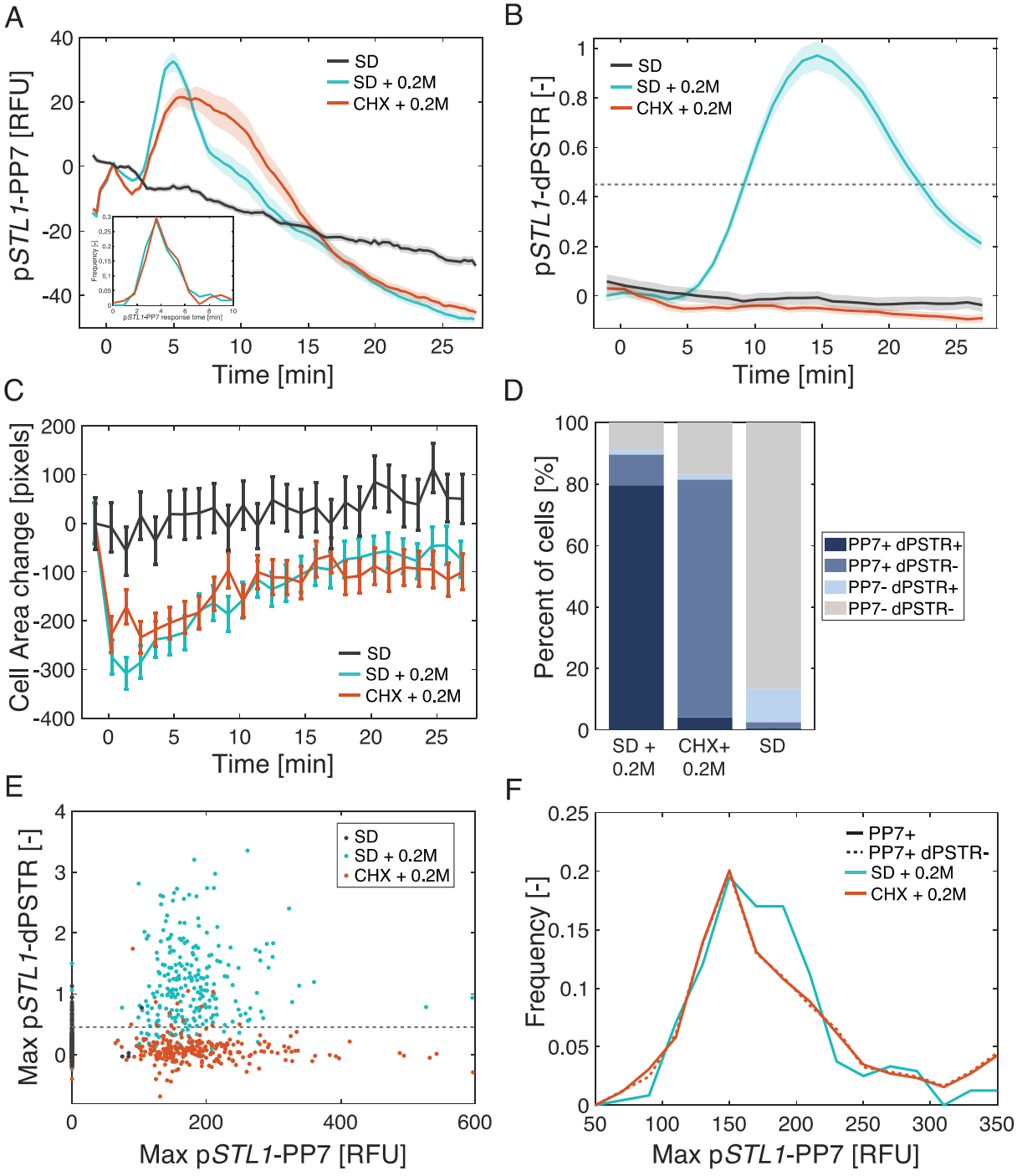


Figure 18: Uncoupling the coupled PP7-dPSTR reporter.

(A and B) Population traces for the coupled p*STL1* PP7-dPSTR reporter. Cells were induced with SDfull or CHX for three minutes before a 0.2M NaCl induction. Single-cell traces were smoothed with a moving average. (A) The 10 brightest pixels of the expended nucleus subtracted for their mean cell fluorescence and post-induction levels. Inset are the histograms of the p*STL1*-PP7 response times, only PP7+ cells were considered. (B) Corresponding p*STL1*-dPSTR response. Plotted is the

nuclear over cytoplasmic fluorescences ratio subtracted for the post-induction level. Dashed line represents the expression threshold at 0.45 defined to discriminate dPSTR positive cells. (C) Mean Cell area and SEM over time. Single-cell area traces were subtracted for their basal values. (D) Percentage of cells positive or negative for PP7 and/or dPSTR signals. (E) Scatter plot of the maximum p*STL1* PP7 fluorescence and dPSTR nuclear over cytoplasmic fluorescence ratio from A and B. (F) Histograms of the maximum PP7 signal for all the PP7 positive cells (full line) and the PP7+ dPSTR- cells (dashed line). The values of the PP7+ cells of the control experiment are plotted as reference.

Recapitulation of mRNA localization

As a second validation step of the coupled reporter assay, we testes whether our coupled reporter could reproduce previously published mRNA localization data. Therefore, we applied it to the well-described *ASH1* transcript asymmetrical localization to the bud tip of anaphase cells [149]. Indeed, it was shown that cloning of the first 250 bp after *ASH1* STOP codon on a transcript is sufficient to recapitulate the endogenous transcript active transport into the mother bud tip [149]. Following the original publication strategy, we cloned a middle strength promoter, here a Ribosomal Protein Gene promoter (pRPG, p*RPL15A*) [184], in front of the coupled reporter, and the exact same *ASH1* 3'UTR minimal sequence as Bertrand *et al.* in the 50bp spacer between the coupled reporter peptide and the PP7 stem-loops (Figure 19A).

As shown in Figure 19B, p*RPL15A* constitutive expression led to a strong nuclear enrichment of the dPSTR part [184]. However, upon bud growth, a PP7-labelled nuclear transcription site can be observed, quickly followed by the accumulation of the signal at the bud tip (Figure 19B, arrows). Therefore, these results suggest that our coupled reporter assay can reproduce faithfully previously published mRNA transport and localization data acquired with single phage coat protein reporter assay and can thus be applied to a broad range of studies.

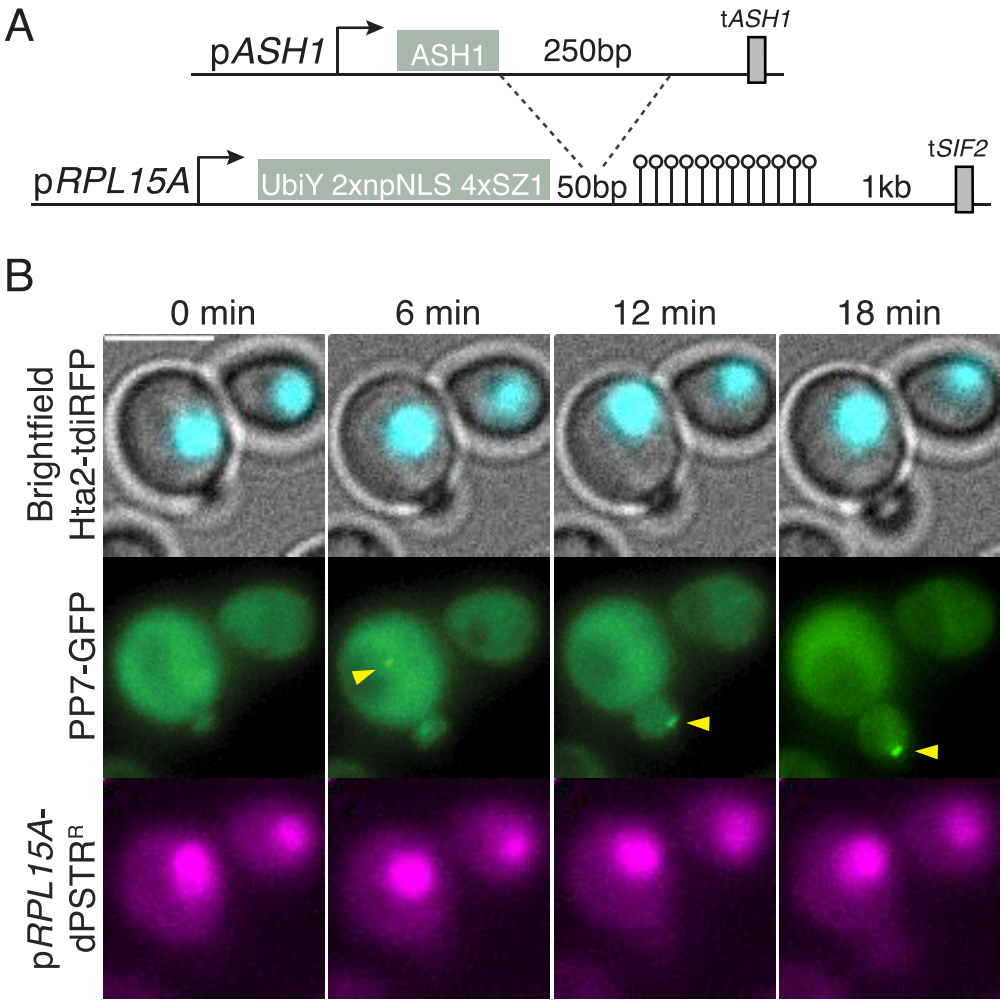


Figure 19: Application of the coupled reporter to ASH1 mRNA bud tip localization.

(A) Schematic representation of the cloning of the 250bp after *ASH1* STOP codon into the coupled reporter assay 50bp spacer, where the *STL1* promoter was replaced by the constitutively expressed middle strength *RPL15A* promoter. (B) Representative microscopy images of the *ASH1* coupled reporter strain. The strain bears a histone tagged with tdiRFP, a PP7-GFP and the p*RPL15A* coupled dPSTR in RFP, here false colored in magenta. Scale bar represents five microns. Arrows point at PP7 signal accumulation, first at the transcription site during active transcription and then at the bud tip, due to *ASH1* 3'UTR sequence addition and recognition by the endogenous transcript transport machinery.

3.3 Discussion

Sensitivity versus dynamicity trade-off

In this chapter, we have developed a live-cell imaging assay for the simultaneous recording of transcription and translation dynamics from a single promoter at the single-cell level. Starting with a slow readout translational Venus reporter, we have upgraded to a stable relocation-based dPSTR assay, to finally reach the most dynamic readout with a degradable dPSTR reporter. Although we reached higher dynamicity with the latest version of the system, we lost some sensitivity during the process. Indeed, the proportion of cells with both readouts increased to the expense of the one displaying only the PP7. Although more dPSTR positive cells could be detected with a higher light illumination, this would lead to a higher bleaching of the system, which is already quite prominent. There thus seems to be a trade-off between readout accuracy and dynamicity that has to be taken under consideration for the choice of the reporter assay, depending on the specific question to be answer.

Improvement on dPSTR bleaching

In all our coupled reporters with the dPSTR system, we observed a noticeable amount of bleaching. Although we have tried different settings acquisition, the signal loss was consistently higher than 40%, unlike the PP7 part, which in its GFP version did not exceed 37%. Therefore, a good improvement of the system would be to exchange the mCherry fluorescent protein for a more photostable one and/or brighter FP, like the mScarlet RFP [185]. This would enable us to use to diminish the excitation light intensity and increase our signal-to-noise ratio, and thus improve further the coupled reporter assay.

Application to other promoters

In this study, we have developed a coupled p*STL1*-PP7-dPSTR assay. This assay gave us detectable readouts for the *STL1* promoter under non-stressful conditions and exposed to stress levels ranging from 0.1M to 0.3M NaCl (Figure 15). Detection of a signal, either for the PP7 or dPSTR reporter, depends on the availability of already synthetized fluorescent proteins to be relocated either on a transcript or into the nucleus. Thus, the level expression of the constitutively expressed parts of the coupled p*STL1*-PP7-dPSTR will have to be adapted for both reporter systems for each promoter, and possibility for each experimental condition, to be tested to ensure proper signal detection and the absence of titration. An easy way to control for this is to increase the levels of the fluorescent protein expression and compare it to the previous strain to assess the discrepancies in the results, like we did in chapter 4 for the PP7 reporter strains reporting on the two strongest promoters monitored, p*HSP12* and p*GPD1* (Supplementary Figure 5). Similarly, application of the PP7 reporter system to constitutively expressed promoters instead of inducible ones may be more challenging and require more precise controls like smFISH. The same actually holds true for the dPSTR assay, which quantification of a nuclear signal increase instead of its apparition, in the case of promoters with basal level, may be more challenging, especially if the basal level promoter induction is relatively low, unlike for the *GPD1* promoter (Chapter 4, figure 2) [144].

Another consideration to the application of the reporter system would be the reporter construct itself. Thanks to the exogenous proteins parts constituting the PP7 and dPSTR assays, namely the phage coat protein and its binding sites, and the synthetic SynZip and its viral Nuclear Localization Sequences (NLS), application to other organisms should be feasible and was already performed for the PP7 system [87, 186-190]. Indeed, budding yeast

promoters are typically 1kb long, therefore, amplification and cloning of the promoter region and integration into an exogenous location is generally sufficient to recapitulate the chromatin environment (nucleosomes position and histone marks) and thus the endogenous promoter expression profile [77, 191]. However, in higher eukaryotes, like mammalian cells, regulatory elements generally not only consist in promoter regions but may be temporally and spatially regulated by distant cis-regulatory elements, like enhancers and transposable elements [192, 193]. Thus, amplification and sub-cloning of the promoter region would likely not recapitulate the full regulatory network and not reflect the endogenous promoter gene expression profile. For this reason, in the mammalian gene expression field, reporter assays are generally integrated at the endogenous location, either in the Open Reading Frame, or in untranslated regions like introns or 3'UTRs, depending on the reporter system used [157, 160]. Therefore, direct comparison of different promoters' expression data with the coupled reporter may be more challenging than in yeast, as presented in Figure 8. This is not specific to our assay but to any gene expression reporter assay used in mammalian systems.

Absence of correlation between mRNA and protein readouts

In all of our reporter settings, we found only poor correlation between the p*STL1* mRNA and protein measurements. This absence of correlation might either reflect a true mRNA dosage-independent protein output, or a limited detection from both readouts, or additional cellular variables to take into consideration. To date, data on gene transcript and protein correlation is still debated and seems to depend on every aspect of the gene tested; from its identity to the measurement method [172]. It would thus be interesting to test if our coupled reporter can report on different promoters' transcripts-proteins correlations by testing other promoters and compare the results side-by-side to data acquired with different assays.

In addition, all PP7-labelled transcript may not be translated or fully transcribed. Indeed, it was shown *in vitro*, but remains to be presented *in vivo*, that RNA pol II could produce abortive transcripts due to a break in its interaction with the targeted promoter [194]. However, it was also shown *in vitro*, that abortive transcripts' sizes were not exceeding eight to fifteen base pair, which would not be sufficient to be detected with the loops. Indeed, each loop is about 15bp and more than ten stem-loops have to be transcribed to have a detectable signal, in general [162]. Therefore, mRNA buffering mechanisms are more likely to occur and could explain part of the lack of correlation. Buffering corresponds to a modulation of the differential rate between synthesis and degradation of a transcript [95]. In particular, studies on the export of the mRNA from the nucleus to the cytoplasm have highlighted a buffering of transcript by the nuclear envelop [95]. This hypothesis is further explored in the general discussion.

3.4 Conclusion

In this chapter, we report on the development of three reporter assays for the quantification of mRNA and protein expressed from a single *STL1* promoter at the single-cell level: a p*STL1* PP7-Venus, a p*STL1* PP7-dPSTR stable and a p*STL1* PP7-dPSTR unstable. We have estimated the dynamic ranges of each of these assays to provide a readout with salt concentrations ranging from 0M to 0.3M NaCl experiments. We have compared our final and most dynamics coupled reporter to the published dPSTR system and to previous stable construct, and assessed the effects of reporter design on the reporter readout. Together, our data demonstrated that our system in its final version (p*STL1* PP7-dPSTR unstable) can reproduce the dPSTR data, by giving simultaneously a dynamic readout of transcription thanks to the coupling of a PP7 system. This reporter can now be applied to other gene expression studies.

3.5 Material and methods

3.5.1 Strain handling

All strains were constructed in W303 background and are listed in Supplementary Table 1. Yeast cells were grown in YPD medium (YEP Broth: CCM0405, ForMedium, Norfolk, UK) for transformation or in Synthetic Defined (SD) medium for imaging (YNB:CYN3801/CSM:DCS0521, ForMedium, Norfolk, UK). Transformants were selected with auxotrophy markers (Uracil, Histidine, Leucine and Tryptophan) and gene deletions were performed with antibiotic resistance to Nourseothricin (NAT), to a concentration of 100µg/ml. For microscopy experiments, cells were grown to log phase, starting with an overnight in SD medium diluted into fresh SDfull medium (YNB:CYN3801/CSM:DCS0521, ForMedium, Norfolk, UK) to OD_{600} 0.025 in the first morning and grown for 8 hours to OD_{600} 0.3-0.5 in the evening, followed by a last dilution of 0.5/OD_{600} before overnight growth, to reach an OD_{600} of 0.1-0.3 in the second morning, before imaging. For time-lapse experiment, a culture of yeast cells grown to OD_{600} 0.1-0.4 was diluted to OD_{600} 0.05 and sonicated twice 45 seconds before welling 200µL into a 96-wells glass bottom plate (MGB096-1-2LG, Matrical Bioscience) coated with a filtered solution of Concanavallin A diluted to 0.5mg ml^{-1} in water (C2010-250MG, Sigma-Aldrich). Cells were let to settle for 35-45 minutes before imaging. Osmotic shock was performed under the microscope into the 96 wells plate, by adding three times concentrated SDfull+NaCl stock solutions to 200µL of cells, to reach the desired final desired salt concentration.

3.5.2 Strains construction

PP7 Venus reporter strains construction

The p*STL1 Venus-24PP7sl* (pVW11) was constructed by cloning a Venus *BamH*1/*BamH*1 from pSP7 plasmid. The p*STL1 Venus-linker-24PP7sl* (pVW97) was built by cloning oVW512 GenBlock fragment *BamH*1/*BamH*1 into pSP264 plasmid. The p*STL1 Venus* construction was obtained by overnight digestion of the pVW11 plasmid with *Kpn*I and self-ligation back.

pSTL1-PP7-dPSTR stable coupled reporter strains construction

From the empty pMS9 (pMCV, Chapter 2), oVW1069 genblock was inserted by *BamH*I/*Spe*I digestion, giving pVW203 plasmid. pVW204 was obtained by the cloning of oVW1070 genblock into pVW203 with *BamH*I/*Mfe*I. Digestion of pVW183 with SpeI/NheI and cloning of pVW204 transcriptional unit, gave pVW207. The entire MSC of pVW207 was cloned AatII/SphI into pDA133, giving pVW208 plasmid. yED212 strain was transformed with pVW208, giving yVW210 strain, which was further transformed with the PP7-GFP from pVW297, resulting in the final strain yVW447.

pSTL1-PP7-dPSTR unstable coupled reporter strains construction

From pVW204 plasmid in the previous section, were extracted the UbiY npNLS 4xSZ1 spacer by *Spe*I/*Nhe*I digestion and inserted into pVW125, giving pVW206. *Mlu*I digestion and self-ligation of pVW206 gave pVW209. Cloning of pVW209 into pVW183 by *Spe*I/*Nhe*I digestion, gave pVW218. pVW219 was cloned from pDA133 *Aat*II/*Sph*I digestion and insertion from pVW218. yED212 strain was transformed with pVW219 plasmid, which resulted in yVW232, which was further transformed with the PP7 from pVW297 plasmid, giving the final yVW489 strain.

PP7 dPSTR ASH1 3'UTR reporter strains construction

From the coupled reporter plasmid pVW207, which was digested *Sac*I/*Spe*I to insert pRPL15A from pNM19, giving pVW303 plasmid. This resulting plasmid was further digested *Nhe*I/*BamH*I to insert the 250bp of ASH1 3'UTR amplified from gDNA with oligos 1608/1609. The resulting plasmid pVW306 was transformed into yED212, giving yVW479. The PP7-GFP where finally transformed into yVW479 with pVW297 plasmid, giving the final yVW480 strain.

3.5.3 Time-lapse imaging

PP7 Venus strains (loop effects, Figure 9)

Strains were imaged with a 40x objective every min for three minutes, stimulated with 0.6M NaCl (0.2M final concentration in well) after timepoint 3 and imaged every five minutes for 1h40, with RFP 100ms and YFP 300ms illuminations for all the strains, even the "empty" control without RFP expression (ySP269).

Coupled PP7-Venus reporter strain (Figure 10)

Strains were imaged with a 60x objective every 30s for 20min and then every 5min to 1h40, with the following illumination settings: CFP 20ms, RFP 50ms 15%LED z-stacks [-1.2; 0.4; +1.2], YFP 400ms. Cells were stressed at timepoint 3, with SDfull, 0.3M or 0.6M NaCl (final concentration in well 0M, 0.1M and 0.2M). The YFP LED was turned OFF after timepoint 6 and turned back on at timepoint 54 to avoid Venus bleaching. Two wells were acquired in parallel, with four positions per well. An average of 45% of bleaching was measured for the RFP channel and 32% for the CFP channel. The ConnectedHiPix of each time-lapse were filtered for single timepoint dot segmentation and the filtered matrix was used to define PP7 positive cells as cells with a non-null ConnectedHiPix sum over the time-lapse.

pSTL1-PP7-dPSTR strains (Figure 13 to 18)

Strains were imaged with a 60x objective every 15s for 25min with the following illumination settings: Cy5p5 300ms, RFP 100ms and GFP 60ms with 6 z-stacks [-1.2; 0.4; +1.2]. A frame skip of three was applied on all channels, except the GFP, to limit bleaching and increase acquisition speed. Cells were stressed at timepoint 3 with different inducer solutions. One well was acquired at a time, with three positions. An average of 45% of bleaching was measured for the GFP channel and 48% for the RFP channel. The ConnectedHiPix of each time-lapse were filtered for single timepoint dot segmentation and the filtered matrix was used to define PP7 positive cells as cells with a non-null ConnectedHiPix sum over the time-lapse. The reference strain for the comparison between the coupled and the dPSTR-Venus strain was imaged as described in Aymoz *et al.* 2016 (yDA119).

ASH1 3'UTR coupled reporter strain (Figure 19)

The yVW480 strain was imaged every 3 minutes for 1h30 with the following illumination settings: Cy5p5 300ms, RFP 100ms and GFP 40ms with 6 z-stacks [-1.2; 0.4; +1.2]. A frame skip of three was applied on brightfield channels.

3.5.4 Statistical analysis

Significant differences between means of replicates were tested by two-samples *t*-tests with 95% confidence intervals.

3.5.5 Supplementary tables

Supplementary Table 1: strains of chapter 3

Strain name	Ancestor strain	Plasmid	Genotype	Reference
ySP2	-	-	MATa/MATα leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 [phi+]	Ralser et al 2012
ySP269	ySP2	-	Hta2-mCherry (URA3)	This study
yVW112	ySP269	pVW11	Hta2-mCherry (URA3) pSTL1 Venus 24xPP7sl::GLT1 (HIS)	This study
yVW113	ySP269	pVW97	Hta2-mCherry (URA3) pSTL1 50bp Venus 24xPP7sl::GLT1 (HIS)	This study
yVW114	ySP269	pVW163	Hta2-mCherry (URA3) pSTL1 Venus::GLT1 (HIS)	This study
yVW119	ySP269	pVW171	Hta2-mCherry (URA3) pCYC1 PP7-dCherry tCYC1::TRP	This study
yVW120	yVW112	pVW171	Hta2-mCherry (URA3) pSTL1 Venus 24xPP7sl (His) pCYC1 PP7-dCherry tCYC1::TRP	This study
yVW121	yVW113	pVW171	Hta2-mCherry (URA3) pSTL1 Venus 50bp 24xPP7sl (His) pCYC1 PP7-dCherry tCYC1::TRP	This study
yVW122	yVW114	pVW171	Hta2-mCherry (URA3) pSTL1 Venus::GLT1 (HIS) pCYC1 PP7-dCherry tCYC1::TRP	This study
ySP261	ySP2	-	Hta2-mCherry (HIS)	This study
yVW63	ySP261	pVW71	Hta2-mCherry (HIS) pSIVu pRPS20 CFP SZ2 tCYC1 pSTL1 2xNLS SZ1 tCYC1::URA	This study
yVW65	ySP261	pVW83	Hta2-mCherry (HIS) pSIVu pRPS20 CFP SZ2 tCYC1 pSTL1 2xNLS SZ1 24xPP7sl 1kb tSIF2::URA	This study
yVW68	ySP261	pVW85	Hta2-mCherry (HIS) pSIVu pRPS20 CFP SZ2 tCYC1 pSTL1 2xNLS SZ1 3kb tSIF2::URA	This study
yVW67	ySP261	pVW95	Hta2-mCherry (HIS) pSIVu pRPS20 CFP SZ2 tCYC1 pSTL1 2xNLS SZ1 50bp 24xPP7sl 400bp tSIF2::URA	This study
yVW58	ySP261	pVW74	Hta2-mCherry (HIS) pSIVu pRPS20 CFP SZ2 tCYC1 pSTL1 2xNLS SZ1 24xPP7sl 400bp tCYC1::URA	This study
yVW59	ySP261	pVW76	Hta2-mCherry (HIS) pSIVu pRPS20 CFP SZ2 tCYC1 pSTL1 2xNLS SZ1 24xPP7sl 3kb tCYC1::URA	This study
yVW75	ySP261	pVW103	Hta2-mCherry (HIS) pSIVu pRPS20 CFP SZ2 tCYC1 pSTL1 2xNLS SZ1 50bp 24xPP7sl 1kb tSIF2::URA	This study
yED216	ySP2	-	Hta2-CFP (LEU)	Wosika et al. 2016
yED212	ySP2	-	Hta2-tdiRFP (TRP)	Wosika et al. 2016
yED159	ySP2	-	Hta2-tdiRFP (HIS)	Wosika et al. 2016
yED212	ySP2	-	Hta2-tdiRFP (TRP)	Wosika et al. 2016
yVW94	yED159	pVW125	Hta2-tdiRFP (TRP) pSIVu pRPL24 mCherry SZ2 tCYC1 pSTL1 2xNLS SZ1 50bp 24xPP7sl 1kb tSIF2::URA	This study
yVW210	yED212	pVW208	Hta2-tdiRFP (TRP) pSIVu pRPL24 mCherry SZ2 tCYC1 pSTL1 UbiY 2xNP-NLS 4xSZ1 50bp 24xPP7sl 1kb tSIF2::URA	This study
yVW426	yED216	pVW296	Hta2-CFP (LEU) pSIVu pADH1 PP7 mCherry tCYC1::URA3	This study
yVW427	yVW426	pVW97	Hta2-CFP (LEU) pSIVu pADH1 PP7 mCherry tCYC1::URA3 pSTL1 Venus 50bp 24xPP7sl::GLT1	This study
yVW232	yED212	pVW219	Hta2-tdiRFP (TRP) pSIVu pRPL24 mCherry SZ2 tCYC1 pSTL1 2xNP-NLS 4xSZ1 50bp 24xPP7sl 1kb tSIF2::URA	This study
yVW447	yVW210	pVW297	Hta2-tdiRFP (TRP) pSIVu pRPL24 mCherry SZ2 tCYC1 pSTL1 UbiY 2xNP-NLS 4xSZ1 50bp 24xPP7sl 1kb tSIF2::URA pSIVh pADH1 PP7ΔFG-GFPenvy tCYC1	This study
yVW489	yVW232	pVW297	Hta2-tdiRFP (TRP) pSIVu pRPL24 mCherry SZ2 tCYC1 pSTL1 2xNP-NLS 4xSZ1 50bp 24xPP7sl 1kb tSIF2::URA3 pSIVh pADH1 PP7ΔFG-GFPenvy tCYC1::HIS	This study
yVW479	yED212	pVW306	Hta2-tdiRFP (TRP) pSIVu pRPL24A mCherry SZ2 pRPL15A UbiY 2xnpNLS 4xSZ1 ASH1 250bp 24xPP7sl 1kb tSIF2::URA3	This study

yVW480	yVW479	pVW297	Hta2-tdiRFP (TRP) pSIVu pRPL24A mCherry SZ2 pRPL15A UbiY 2xnpNLS 4xSZ1 ASH1 250bp 24xPP7sl 1kb tSIF2::URA3 pSIVh pADH1 PP7ΔFG-GFPenvy tCYC1::HIS	This study
ySP37			Hta2-CFP (marker looped out)	Aymoz et al. 2016
yDA123	ySP37	pDA183	Hta2-CFP pRPL24A mCherry SZ2 pSTL1 UbiY NLS SZ1 tCYC1::URA3	Aymoz et al. 2016
yDA134	yDA123		Hta2-CFP pRPL24A mCherry SZ2 pSTL1 NLS Venus SZ1 tCYC1::URA3 Hog1-mCitrine (HIS)	Aymoz et al. 2016

Supplementary Table 2: plasmids of chapter 3

Plasmid name	Backbone	Insert	Construction	Reference
pSP7	-	-	pBS Venus	pRF70 Fabian Rudolf
pSP219	-	-	pRS416	This study
pSP223	-	-	pRS305 pCYC1 mCherry	This study
pSP226	-	-	pCEN pMET25 PP7-2xGFP	pDZ276 Addgene 35194
pSP227	-	-	pRS303 pPOL1 24xPP7sl	pDZ306 Addgene 35196
pSP264	pSP227	-	pRS303 pSTL1 24xPP7sl	Aymoz et al. 2016
pSP268	pSP226	-	pRS304 pMET25 PP7-2xGFP	Aymoz et al. 2016
pDA13	pSP219	pSP223	pRS416 pCYC1 mCherry	This study
pVW1	pSP125	pSP268	pMET25 PP7	This study
pVW3	pVW1	pSP97	pMET25 PP7-dCherry	This study
pVW9	pVW3	pDA13	pCYC1 PP7-dCherry	This study
pVW11	pSP264	pSP7	pRS303 pSTL1 Venus 24xPP7sl	This study
pVW71	pVW67	pVW70	pSIVu pRS20 CFP SZ2 tCYC1 pSTL1 2xNLS SZ1 tCYC1	This study
pVW74	pVW67	pVW61	pSIVu pRS20 CFP SZ2 tCYC1 pSTL1 2xNLS SZ1 24xPP7sl 400bp tCYC1	This study
pVW76	pVW67	pVW63	pSIVu pRS20 CFP SZ2 tCYC1 pSTL1 2xNLS SZ1 24xPP7sl 3kb tCYC1	This study
pVW81	pVW78	pDA13	pRS304 pCYC1 PP7-2xGFP tCYC1	This study
pVW83	pVW67	pVW82	pSIVu pRS20 CFP SZ2 tCYC1 pSTL1 2xNLS SZ1 24xPP7sl 1kb tSIF2	This study
pVW85	pVW67	pVW84	pSIVu pRS20 CFP SZ2 tCYC1 pSTL1 2xNLS SZ1 3kb tSIF2	This study
pVW95	pVW67	pVW87	pSIVu pRS20 CFP SZ2 tCYC1 pSTL1 2xNLS SZ1 24xPP7sl 400bp tSIF2	This study
pVW97	pSP264	-	pRS303 pSTL1 Venus 50bp 24xPP7sl	This study
pVW163	pVW11	-	pRS303 pSTL1 Venus	This study
pVW171	pVW81	pVW9	pRS304 pCYC1 PP7-dCherry tCYC1	This study
pMS9	-	-	pMCV	Wosika et al. 2016
pVW203	pMS9	-	pMCV UbiY 2xnpNLS SZ1	This study
pVW204	pVW203	-	pMCV UbiY 2xnpNLS 4xSZ1 50bp	This tudy
pDA125	pDA99	-	pRS305 pSTL1 UbiY 2xNLS SZ1	Delphine thesis
pVW206	pDA125	pVW204	pRS305 pSTL1 UbiY 2xnpNLS 4xSZ1 50bp	This study
pVW207	pVW183	pVW204	pRS305 pSTL1 UbiY 2xnpNLS 4xSZ1 50bp 24xPP7sl 1kb tSIF2	This study
pVW183	pVW100	pVW49	pRS305 pSTL1 2xNLS 4xSZ3 50bp 24xPP7sl 1kb tSIF2	This study
pDA133	pSIVu	pSP329	pSIVu pRPL24A mCherry SZ2	Aymoz et al. 2016
pVW208	pDA133	pVW207	pSIVu pRPL24A mCherry SZ2 pSTL1 UbiY 2xnpNLS 4xSZ1 50bp 24xPP7sl 1kb tSIF2	This study
pVW209	pVW206	-	pRS305 pSTL1 2xnpNLS 4xSZ1 50bp	This study
pVW219	pDA133	pVW218	pSIVu pRPL24A mCherry SZ2 pSTL1 2xnpNLS 4xSZ1 50bp 24xPP7sl 1kb tSIF2	This study
pVW296	pVW284	pVW212	pSIVu pADH1 PP7 mCherry tCYC1	This study
pVW297	pVW101	pVW284	pSIVu pADH1 PP7 mCherry tCYC1	This study
pNM19	pSP68	-	pBS pRPL15A	This study
pVW303	pVW207	pNM19	pRS305 PRPL15A UbiY 2xnpNLS 4xSZ1 50bp 24xPP7sl 1kb tSIF2	Aymoz et al 2016
pVW306	pDA133	pVW303	pSIVu pRPL24A mCherry SZ2 pRS305 PRPL15A UbiY 2xnpNLS 4xSZ1 50bp 24xPP7sl 1kb tSIF2	This study

Chapter 4: HOG genes transcription

4.1 Background

In response to a sudden increase in extracellular osmolarity, budding yeast cells trigger the High Osmolarity Glycerol (HOG) pathway (Figure 4), which cumulates in the activation of the main effector of the pathway, the MAPK Hog1. The immediate response of the cell aims at increasing the intracellular glycerol concentration by closing glycerol channels and increasing the synthesis of this internal osmolyte. However, a transient change in the transcriptional profile is also initiated to ensure long term adaptation. This transient gene activation is mediated by the transient activation and nuclear accumulation of the MAPK. Indeed, osmostress genes are repressed under basal conditions through closed chromatin conformation and histone modifications. To initiate transcription, Hog1 is targeted to osmostress genes via DNA-bound transcription factors and in turn recruits chromatin remodeling and modifying complexes, as well as the transcription machinery. Although more than 250 genes are governed by the same transient MAPK activity, initial single-cell experiment have shown significant differences in transcription dynamics. However, these data were generated with protein-based reporter assay and thus do not reflect precisely promoter activity. Thus, dynamic data on osmostress gene transcription dynamics and the crucial determinant of inter-promoter variability in transcriptional profile is still lacking.

4.2 Results

The results of this chapter are presented as a second version of a preprint deposited on bioRxiv (Wosika et al 2019, Wosika et al 2020) under revision at *Nature Communications*.

Author contributions:

Victoria Wosika and Serge Pelet designed the experiments, analyzed the data and wrote the manuscript. Victoria wrote the established the conditions for PP7 imaging. Victoria and Serge performed the experiments.

4.3 Conclusion

In this study, we have characterized the six most studied promoters reporting on the HOG pathway transcriptional activity at an unprecedent resolution by engineering PP7 reporter strains. Using gene deletions and differential experimental conditions, we have dissected the regulation of chromatin, promoter sequence and transcription factor on osmostress genes transcription dynamics.

4.4 Supplementary controls

All the strains in chapter 4 were genotyped for their proper integration at the *GLT1* locus and for the integrity of their PP7 stem-loops array. Most of the PCR results are displayed in Annex 2. The summary Table of the strains built and their controls are found in Annex 3 and all their microscopy replicate experiments in Annex 4.

Single-particle view of stress-promoters induction dynamics: an interplay between MAPK signaling, chromatin and transcription factors

Victoria Wosika and Serge Pelet*

Department of Fundamental Microbiology
University of Lausanne
1015 Lausanne, Switzerland.

*Correspondance : serge.pelet@unil.ch

Abstract

Precise regulation of gene expression in response to environmental changes is crucial for cell survival, adaptation and proliferation. In eukaryotic cells, extracellular signal integration is often carried out by Mitogen-Activated Protein Kinases (MAPK). Despite a robust MAPK signaling activity, downstream gene expression can display a great variability between single cells. Using a live mRNA reporter, we monitored the dynamics of transcription in *Saccharomyces cerevisiae* upon hyper-osmotic shock. The transient activity of the MAPK Hog1 opens a temporal window where stress-response genes can be activated. Here we show that the first minutes of Hog1 activity are essential to control the activation of a promoter. The chromatin repression on a locus slows down this transition and contributes to the variability in gene expression, while binding of transcription factors increases the level of transcription. However, soon after Hog1 activity peaks, negative regulators promote chromatin closure of the locus and transcription progressively stops.

Introduction

A crucial function of all cellular life is the ability to sense its surroundings and adapt to its variations. These changes in the extracellular environment will induce specific cellular responses, orchestrated by signal transduction cascades which receive cues from plasma membrane sensors. This information is turned into a biological response by inducing complex transcriptional programs implicating hundreds of genes[1–3]. Tight regulation of signaling is thus crucial to ensure the correct temporal modulation of gene expression, which can otherwise alter the cell physiology[4–6]. Interestingly, single-cell analyses have revealed that genes regulated by an identical signaling activity can display a high variability in their transcriptional responses[7–10]. This noise in transcriptional output questions how signal transduction can faithfully induce different loci and which molecular mechanisms contribute to the variability in gene expression.

In eukaryotic cells, various environmental stimuli are transduced by the highly conserved Mitogen-Activated Protein Kinases (MAPK) cascades[11,12]. They control a wide range of cellular responses from cell proliferation, differentiation or apoptosis. In *Saccharomyces cerevisiae*, a sudden increase in the osmolarity of the medium is sensed by the High Osmolarity Glycerol (HOG) pathway, which leads to the activation of the MAPK Hog1, a homolog of p38 in mammals[13,14]. Upon hyper-osmotic stress, the kinase activity of Hog1 promotes the adaptation of the cells to their new environment by driving an increase in the internal glycerol concentration, thereby allowing to balance the internal and the external osmotic pressures. In parallel to its cytoplasmic activity, Hog1 also transiently accumulates into the nucleus to induce the expression of hundreds of osmostress-responsive genes (Fig. 1a). The MAPK is recruited to promoter regions by Transcription Factors (TFs) and, in turn, Hog1 recruits chromatin remodeling complexes, the Pre-Initiation Complex and the RNA Polymerase II (PolII) to trigger gene expression[15,16]. Once cells have adapted, Hog1 is inactivated and exits the cell nucleus, transcription stops and chromatin is rapidly reassembled at HOG-induced gene loci.

Biochemical analyses of this pathway have identified the key players implicated in gene expression and the central role played by the MAPK in all these steps[15]. In

parallel, single-cell measurements have uncovered the large variability present in their expression. In particular, translational reporters and RNA-FISH measurements have identified that slow chromatin remodeling promoted by the MAPK at each individual locus is generating strong intrinsic noise in the activation of many stress-responsive genes[9,17].

In order to get deeper insights into the regulation of osmostress-genes expression kinetics, we aimed at monitoring the dynamics of mRNA production in live single cells. Phage coat protein-based assays, like the MS2 or PP7 systems, have been used to visualize mRNA in live single cells[18–20]. These experiments contributed to revealing the bursty nature of transcription, whereby a set of polymerases simultaneously transcribing a gene generates a burst in mRNA production, which is followed by a pause in transcription[21–23].

In this study, we dissect the kinetics of transcription of osmostress-genes. The production of mRNA is monitored using the PP7 phage coat protein assay. This reporter allows us to measure with high temporal resolution and in a fully automated manner, the fluctuations in transcription arising in hundreds of live single cells. This analysis enables to dissect the contribution of various players to the overall transcriptional output. We show that the first few minutes of MAPK activity will determine if a gene is transcribed. We also demonstrate that the chromatin state of a promoter will control the timing of activation and thus the variability of the transcription, while the TF binding will influence the level and duration of the mRNA production.

Results

High osmotic pressure is sensed and transduced in the budding yeast *Saccharomyces cerevisiae* via the HOG signaling cascade, which culminates in the activation of the MAPK Hog1 (Fig. 1a). Upon activation, this key regulator accumulates in the nucleus to trigger gene expression in a stress level-dependent manner (Supplementary Figure 1a). The activity of the kinase can be monitored by following its own nuclear enrichment[24,25]. In parallel to Hog1, the general stress response pathway is induced by the hyper-osmotic shock and the transcription factor Msn2 also relocate into the nucleus with dynamics highly similar to the ones observed for Hog1 (Supplementary

Figure 1b and c) [26,27]. Nuclear Hog1 and Msn2 (together with its paralog Msn4) induce osmostress-genes expression, with approximately 250 genes being up-regulated upon osmotic shock[1,28,29]. The activity of the pathway is limited to the cellular adaptation time, which conincide with the nuclear exit of Hog1 and the recovery of the cell size (Supplementary Figure 1a and d). The fast and transient activity of the osmostress response as well as the homogenous activation of the MAPK within the population[9,25] (Supplementary Figure 1e), make this signaling pathway an excellent model to understand the induction of eukaryotic stress-responsive genes, which are often accompanied by important chromatin remodeling.

Monitoring the dynamics of osmostress-genes transcription

In order to quantify the dynamics of transcription in live single cells, we use the PP7 system to label the production of messenger RNAs (mRNA)[30]. Briefly, constitutively expressed and fluorescently labeled PP7 phage coat proteins strongly associate to a binding partner: an array of twenty-four PP7 mRNA stem-loops (PP7sl). In our settings, this PP7 reporter construct is placed under the control of a promoter of interest and integrated in the genome at the *GLT1* locus (Fig. 1b)[30], in a strain bearing a nuclear tag (Hta2-mCherry) and expressing a fluorescently tagged PP7 protein (PP7ΔFG-GFPenvy[31,32], abbreviated PP7-GFP, Methods). Upon activation of the promoter, local accumulation of newly synthesized transcripts at the Transcription Site (TS) leads to the formation of a bright fluorescent focus due to the enrichment in PP7-GFP fluorescence above the background signal (Fig. 1c and Supplementary Movie 1). The fluorescence intensity at the TS is proportional to the number of mRNA being transcribed and thus to the instantaneous load of RNA polymerases. After termination, single mRNAs are exported out of the nucleus and their fast diffusion in the cytoplasm prevents their detection under the selected illumination conditions.

Typically, time-lapses with fifteen-second intervals for twenty-five minutes with six Z-planes for the PP7-GFP channel on four fields of view were performed. Image segmentation and quantification were performed automatically, allowing to extract up to four hundred single-cell traces for each experiment[33]. The mean intensity of the 20 brightest pixels in the nucleus, from which the average cell fluorescence was

subtracted, was used as a measurement of TS intensity and thus as proxy for transcriptional activity (Fig. 1d, Methods).

Fig. 1d displays the average TS fluorescence from more than 200 cells bearing the p*STL1*-PP7sl reporter, following the activation of the HOG pathway by various NaCl concentrations. The HOG-induced *STL1* promoter has been extensively studied at the population and single cell level[9,17,34,35]. As expected, increasing salt concentrations lead to a proportionally increasing transcriptional output from the cell population, whereas no change in TS fluorescence is detected in the control medium.

The hundreds of dynamic measurements acquired with the PP7 reporter form a rich dataset where multiple features can be extracted from each single-cell trace (Fig. 1e, Methods). Our automated image segmentation and analysis allow to reliably quantify the appearance (Start Time) and disappearance (End Time) of the TS (Supplementary Figure 2 and Method). The maximum intensity of the trace and the integral under the curve provide estimates of the transcriptional output from each promoter (Fig. 1e). In addition, transcriptional bursts can be identified by monitoring of strong fluctuations in the TS intensity.

Validation of the live mRNA reporter assay

The mRNA dynamics measured with the PP7 assay are in close agreement with previously reported data set[34,36]. Nonetheless, we also verified with a dynamic protein expression reporter that comparable results can be obtained (Supplementary Figure 3a). The dynamic Protein Synthesis Translocation Reporter (dPSTR) is an assay that allows the kinetics of gene expression from a promoter of interest. It by-passes the slow maturation time of fluorescent proteins (FP) by monitoring the relocation of the fluorescent signal in the nucleus of the cell[37].

For the PP7 assay, as well as the dPSTR and many other expression reporters, an additional copy of the promoter of interest is inserted in a non-native locus. In order to address if this modified genomic environment alters the dynamics of gene expression, we used CRISPR-Cas9 to integrate the PP7sl downstream of the endogenous *STL1* promoter (Supplementary Figure 4). Interestingly, we observe only minor differences between the p*STL1* at its endogenous location and at the *GLT1* locus. This observation

strongly suggests that the *STL1* promoter sequence placed at a non-endogenous locus replicates many of the properties of the endogenous promoter.

Intrinsic noise in osmostress-gene activation

The microscopy images presented in Fig. 1c illustrate the noise that can be observed in the activation of the p*STL1* promoter upon osmotic stress and which has been previously reported[9,17]. In order to verify that this noise is not due to a lack of activation of the MAPK Hog1 in the non-responding cells, we combined the p*STL1*-PP7sl reporter and the Hog1-mCherry relocation assay in the same strain. As expected, we observe an absence of correlation between the two measurements (Supplementary Figure 5). Indeed, cells with similar Hog1 relocation behaviors can display highly variable transcriptional outputs.

An additional assay to observe this heterogeneity is to monitor the activation of two *STL1* promoters within the same cell. Using a diploid strain where both *GLT1* loci were modified with either a p*STL1*-24xPP7sl or a p*STL1*-24xMS2sl and expressing PP7-mCherry and MS2-GFP proteins, we observe an uncorrelated activation of both loci within each single cell (Supplementary Figure 6 and Supplementary Movie 2). This observation confirms the high intrinsic noise generated by the *STL1* promoter upon osmotic stress[9,37]. The highly dynamic measurements provided by the PP7 reporter allows us to decipher some of the parameters that contribute to this large variability.

High variability in osmostress-genes transcription dynamics

In addition to p*STL1*, five other stress-responsive promoters often used in the literature to report on the HOG pathway transcriptional activity were selected for this study[34,38]. Each reporter strain differs only by the one thousand base pairs of the promoter present in front of the PP7sl (800bp for p*STL1*, 660 for p*ALD3*[9,39]); however, each strain displays a different transcriptional response following a 0.2M NaCl stimulus (Fig. 2a). Because the level of accumulation of the PP7 signal at the transcription site and the timing of the appearance and disappearance of the TS is different for each tested promoter, it implies that the promoter sequence dictates multiple properties of the transcription dynamics. These dynamic measurements are in general agreement with control experiments performed with the dPSTR assay (Supplementary Figure 3b) and previously published population-averaged data[34,37]. Importantly, expressing three

times more phage coat proteins did not change the parameters extracted from the PP7 measurements for the two strongest promoters, denoting the absence of titration of PP7-GFP reporter proteins in our experimental settings (Supplementary Figure 7).

The automated analysis allows to identify the presence or absence of a transcription site in each single cell, and thus the fraction of cells that induce the promoter of interest (Fig. 2b). Interestingly, even in absence of stimulus, some promoters display a basal level of transcription. In the p*GRE2*, p*HSP12* and p*GPD1* reporter strains, an active transcription site can be detected in 5 to 20% of the cells in the few time points before the stimulus (Fig. 2c, Supplementary Movie 3). If the period of observation is extended to a twenty-five-minute time lapse without stimulus, this fraction increases 2 to 3-folds (Supplementary Figure 8). Upon activation by 0.2M NaCl, the fractions of responding cells for the three promoters that display basal expression overcomes 85%, while it remains below 65% for the three promoters without basal induction.

Chromatin state sets the timing of transcription initiation

A key parameter controlled by the promoter sequence is the timing of induction. In Fig. 2d, the time when cells become transcriptionally active (Start Time) is plotted as a Cumulative Distribution Function (CDF) only for the cells where a TS is detected after the stimulus, thereby excluding basal expressing cells and non-responding cells. Treatment with 0.2M NaCl results in a sudden activation of transcription (Fig. 2d). This contrasts with non-induced samples, where the CDF of the promoters displaying basal activity rises almost linearly due to stochastic activation during the recording window (Supplementary Figure 8c).

Upon stress, the promoters displaying basal activity are induced faster than the promoters that are repressed under log-phase growth, with p*GPD1* being activated the fastest (~1 min), while p*ALD3* and p*STL1* require more than 4 minutes for activation (Fig. 2e). However, there is a great variability in transcription initiation between cells of the same population, since we generally observe 3 to 4 minutes delay between the 10^{th} and 90^{th} percentiles, with the exception of p*GPD1* where the induction is more uniform and less than 2 min delay is observed (Fig. 2e). Comparison between individual replicates demonstrates the reliability of our measurement strategy. Interestingly, we observe a positive correlation between faster transcriptional activation

183 from p*GPD1*, p*HSP12* and p*GRE2* and the presence of basal expression level. These
184 promoters also display the highest numbers of responding cells upon a 0.2M NaCl
185 shock. These results suggest that basal expression is associated with a more
186 permissive chromatin state, which enables a faster activation and higher probability of
187 transcription among the cell population.

188 To test this hypothesis, we disrupted the function of the SAGA chromatin remodeling
189 complex by deleting *GCN5*[40]. As expected, we observe fewer transcribing cells and a
190 slower induction of the p*STL1* promoter in this background (Fig. 2f). Less remarkably,
191 abolishing histone H2AZ variants exchange at +1 and -1 nucleosomes by deleting
192 *HTZ1*[41] only results in a reduced percentage of transcribing cells. Conversely,
193 chromatin state at the *STL1* promoter can be loosened by relieving the glucose
194 repression using raffinose as a C-source[42]. Interestingly, a fraction of the cells grown
195 in these conditions displays basal expression from the p*STL1*-PP7 reporter and the
196 Start Time in raffinose is accelerated by 1 min compared to glucose (Fig. 2g).

197 The link between the chromatin state under log-phase growth and the ability to induce
198 stress-responsive genes is cofirmed by these results. A promoter that is tightly
199 repressed will need more Hog1 activity and thus more time to become transcriptionally
200 active, therefore displaying a lower fraction of responding cells.

201 *Early Hog1 activity dictates transcriptional competence*

202 The period of Hog1 activity provides a temporal window where transcription can
203 potentially be initiated. However, the switch to a transcriptionally active state takes
204 place almost exclusively within the first few minutes after the stimulus. When
205 comparing the characteristic timing of Hog1 nuclear enrichment to the CDF of Start
206 Times for cells bearing the p*STL1* reporter (Fig. 3a and Supplementary Figure 9a), we
207 observe that 90% of the transcribing cells initiate transcription during the first few
208 minutes of the stress response, while Hog1 nuclear accumulation rises and before it
209 drops below 80% of its maximum (decay time). A similar behavior is observed for all
210 the promoters tested, independently of the presence of basal levels (Fig. 3a and b).
211 For p*ALD3*, which is the slowest promoter tested, 87% of the Start Times are detected
212 before the decay of Hog1 activity (7 min) while the full adaptation time takes 14 min at
213 0.2M NaCl.

Interestingly, promoter output also decreases with the time after stimulus. Cells that start transcribing p*STL1* early display a larger integral over the PP7 signal and a higher maximum intensity compared to cells that initiate transcription later (Fig. 3c and Supplementary Figure 9b). A similar behavior is quantified for all tested promoters (Supplementary Figures 9c and d). These measurements demonstrate that the high Hog1 activity present in the first minutes of the response is key to determine both the transcriptional state and overall output of the promoters.

TFs control the dynamics and level of mRNA production

Promoters dictate the timing of transcriptional activation of the ORF and the level at which the mRNA is produced. To extract the transcriptional level of each promoter, we use as a proxy the maximum of the PP7 trace of each single cell where a transcription event could be detected (Fig. 4a). This value represents the maximal loading of polymerases on the locus during the period of transcription. Similar results are obtained when comparing the integral below the PP7 trace, which represents the total transcriptional output from a promoter (Supplementary Figure 10a). As shown in Fig. 4a, each promoter has an intrinsic capability to induce a given level of transcription, which is independent from the presence of basal transcription or the locus activation time. Indeed, p*GRE2* displays the lowest level of induction among all tested promoters, despite the presence of basal transcription and being the second-fastest promoter activated.

As expected, the recruitment of the RNA polymerases is stimulated by the stress; the three promoters with basal activities display a higher transcriptional level upon a 0.2M NaCl stress than in normal growth conditions (Fig. 4b and Supplementary Figure 10b). Both the general stress transcription factors Msn2 and Msn4 and the TFs activated by the MAPK Hog1 (Hot1, Sko1, Smp1) contribute to the transcriptional up-regulation[29,38,43]. Based on studies on synthetic promoters, it has been established that binding site number and distance from the Transcription Start Site (TSS) influence the promoter output[44]. Unfortunately, osmostress promoters display a wide diversity in number and affinity of TF binding sites and no obvious prediction of the transcriptional activity can be drawn (Supplementary Figure 11). While multiple Msn2/4 binding sites

can be found on the *GPD1* and *STL1* promoter sequences, their activation are only mildly affected by deletions of these two TFs (Supplementary Figure 12).

Both *GPD1* and *STL1* are primarily Hog1 targets[28,29,38]. However, their requirements for Hog1 activity is strikingly different. In strains where the MAPK has been anchored to the plasma membrane to limit its nuclear enrichment[45], p*STL1* induction is virtually abolished (only 1.5% transcribing cells) while p*GPD1* activity is barely affected (Supplementary Figure 13). Similarly, deletion of either TF Sko1 or Hot1 profoundly alter the capacity of p*STL1* to be induced (Fig. 4c -f) while these same mutations have a weaker effect on the *GPD1* promoter.

Because the induction of the *STL1* promoter requires an efficient chromatin remodeling, every defect (TF deletion or absence of Hog1 in the nucleus) strongly alters its capability to induce transcription. In comparison, the p*GPD1* is less perturbed by these same defects. We postulate that TFs act in a cooperative manner on p*STL1*, while they act independently of each other on p*GPD1*.

Bursts of PolII transcription in osmostress-gene activation

The PP7 and MS2 systems have allowed to directly visualize transcriptional bursting. In order to identify bursts arising from osmostress promoters, we sought to detect strong fluctuations in each single-cell trace. Fluctuations in TS intensities were filtered to retain only peaks separated by pronounced troughs (Methods). In 20 to 30% of the traces, two or more peaks are identified (Fig. 5a and b). The total length of the transcript downstream of the promoter is 8kb (1.5kb for the stem loops + 6.5kb for *GLT1*). Based on a transcription speed of 20bp/s[30], the expected lifetime of a transcript at the TS is 6.6 min. This corresponds well to the mean duration observed for the p*ALD3*, p*CTT1*, p*STL1* and p*GRE2* reporters (Fig. 5c). However, it is unlikely that the strong TS intensities recorded are generated by a single transcript, but rather by a group of RNA PolII that simultaneously transcribe the locus, probably forming convoys of polymerases[46]. Indeed, single mRNA FISH experiments have shown that following a 0.2M NaCl stress, the endogenous *STL1* locus produces on average 20 mRNAs per cell, with some cells producing up to 100[36].

For p*HSP12* and p*GPD1*, the average peak duration is longer than 11 min (Fig. 5c), suggesting that multiple convoys of polymerases are traveling consecutively through

the ORF. Unfortunately, the long half-lives of the transcripts on the locus prevent a separation of individual groups of polymerases. However, when we achieve to isolate individual peaks in the single cell traces, their duration becomes closer to the expected value of 6.6 min (Fig. 5d). In addition, the output of the transcription estimated by the maximum intensity of the trace or the integral under the whole curve is equal or lower for traces with multiple pulses compared to traces where only a single peak is present (Fig. 5e and Supplementary Figure 14). Together these data strengthen the notion that these stress-responsive promoters are highly processive, displaying an elevated rate of transcription once activated. Only brief pauses in the transcription can be observed in a small fraction of the responding cells.

MAPK activity opens an opportunity window for transcription

We have shown that transcription initiation is dictated by early Hog1 activity. Next, we want to assess what the determinants of transcription shutoff are and by extension, the duration of transcriptional activity. In the HOG pathway, the duration of transcription has been reported to be limited by the cellular adaptation time[34,38]. Therefore, the duration of transcription is shorter after a 0.1M NaCl stress and longer after a 0.3M stress, compared to a 0.2M stress (Fig. 6a). For the p*STL1* promoter, the last time point where a PP7 signal is detected at the TS matches the timing of nuclear exit of the MAPK at all concentrations tested (Fig. 6b).

In order to challenge this link between Hog1 activity and transcriptional arrest, we sought to modulate the MAPK activity pattern by controlling the cellular environment in a dynamic manner. Using a flow channel set-up, we generated a step, a pulse, or a ramp in NaCl concentrations (Fig. 6c, Methods). These experiments were performed in a strain carrying the p*STL1*-PP7sl reporter in conjunction with Hog1-mCherry, allowing to monitor kinase activity and the downstream transcriptional response in the same cell.

The step stimulus at 0.2M NaCl mimics the experiments performed in wells, where the concentration of the osmolyte is suddenly increased at time zero and remains constant throughout the experiments (Supplementary Movie 4). The mean responses at the population level (Fig. 6c) confirm this relationship between Hog1 adaptation time and transcription shutoff time. However, at the single cell level, no direct correlation is

observed between these two measurements due to important single cell variability (Fig. 6f).

In the pulse assay, 7 min after the initial 0.2M step, the NaCl concentration is set back to 0M (Supplementary Movie 5). This shortens the MAPK activity period, as Hog1 leaves the nucleus immediately when cells are brought back in the normal growth medium. Removing the kinase from the nucleus has a direct impact on the transcriptional process. First, fewer cells become transcriptionally active. Second, the active TS sites disappear within a few minutes after the end of the pulse (Fig. 6d and e). Therefore in this context, we observe a direct correlation between Hog1 activity and transcription, which is in line with the known role played by MAPKs, and Hog1 in particular, on multiple steps of the transcriptional process[47,48].

The ramp experiment starts with a pulse at 0.2M NaCl followed by a slow increase of the NaCl concentration up to 0.6M over the next 20 min (Supplementary Movie 6). This constant rise in external osmolarity extends the Hog1 activity window by preventing the adaptation of the cells. More cells can become transcriptionally active and the transcription shut off is delayed (Fig. 6d and e). However, in these conditions, there is a clear lack of correlation between Hog1 activity, which is sustained in many cells over the 30min of the time-lapse, and the transcription output of the p*STL1* that stops much earlier. Taken together, these experiments demonstrate that the MAPK activity is required but not sufficient to sustain the transcriptional process. In the ramp experiment, transcription cannot be sustained throughout the whole Hog1 activity window, demonstrating that other factors contribute to limiting the duration of the transcription.

Promoter identity influences the transcription shutoff time

In order to test whether the promoter identity plays a role in the process of transcription shutoff, we quantified the duration of the transcriptional period for the six promoters and plotted the cumulative distribution of End Times following a 0.2M NaCl stress (Fig. 7a and b). Interestingly, despite similar cell volume adaptation time for all the experiments, the promoters display substantially different kinetics of inactivation. Promoters transcribing at a lower level (p*CTT1* and p*GRE2*) terminate transcription earlier. This shorter transcriptional window may reflect an inferior recruitment of

transcriptional activators to the promoter, enabling an earlier inhibition of transcription due to chromatin closure. In addition, promoters with basal activity display an extended period of transcription after adaptation (Fig. 7a and b). For p*GPD1* and p*GRE2*, this results in a biphasic decay, where the first phase corresponds to the arrest of Hog1-induced transcription and the second phase can be associated to the basal transcription arising from these promoters (Fig. 7b). Note that basal transcription may even be increased due to a higher basal Hog1 signaling activity post high osmolarity conditions[49].

Remarkably, p*HSP12* transcription persists beyond the adaptation time, with nearly 30% of the cells displaying an active TS at the end of the experiment. This suggests that basal expression from this promoter is strongly increased post-stimulus. In contrast to p*GPD1* and p*GRE2*, p*HSP12* possesses numerous Msn2/4 binding sites. Although the relocation dynamics of Hog1 and Msn2 are highly similar during the adaptation phase, Msn2 displays some stochastic secondary pulses[27], that are not correlated to Hog1 relocation events. This could explain the stronger basal expression arising from this promoter post-adaptation (Supplementary Figure 1e and f).

To summarize, these measurements demonstrate that the pattern of MAPK activity provides a temporal window where transcription can take place. When the signaling cascade is shut off, transcription ceases soon afterward. However, the promoter identity, and probably its propensity to recruit positive activators, will determine for how long it can sustain an open chromatin environment favorable to transcription before Hog1 activity decreases due to cellular adaptation.

Discussion

In this study, we have constructed PP7 reporter strains to monitor the transcription dynamics of osmostress promoters. The second exogenous copy of the promoter is integrated at the *GLT1* locus. This strategy provides a similar genomic environment for all the promoters, in order to compare their specific characteristics. Interestingly, we saw only minor differences in CDF of Start Times of the p*STL1* when integrated at its endogenous locus compared to the *GLT1* locus. This observation provides a strong evidence that TF binding and chromatin state of the duplicated promoter sequences

mimic closely the ones at the native environment of the gene. Note that the signal at the TS is expected to be proportional to the length of the transcribed mRNA. The *GLT1* locus with its 6.5kB length was expected to provide a signal four times stronger than the endogenous *STL1* ORF (1.7kB). The unexpectedly high signal obtained from the PP7 reporter at the endogenous locus may be indicative of global difference in transcription rates between the *GLT1* and *STL1* ORFs alternatively, the smaller *STL1* ORF might enhance transcription efficiency via gene-looping[50,51].

Our data illustrate the complex balance that exists between positive and negative regulators taking place at the stress induced loci. At each locus, positive and negative regulators will control the level and duration of transcription. We have shown that the first few minutes of Hog1 activity are essential to initiate the transcription. Transcription factors, chromatin remodelers such as the SAGA and RSC complexes, together with Hog1 will contribute to open and maintain an accessible chromatin environment at the stress-response loci[35,40]. Once initiated, transcription seems highly processive and only in a small fraction of traces, we are able to detect a pause in transcription. However, it has been shown that PolII recruits additional chromatin remodelers, including the Ino80 complex and Asf1 that will redeposit nucleosomes after acute transcription[52]. These conflicting activities will determine the overall duration of transcription at a locus. Indeed, promoters with lower transcriptional activity, such as p*GRE2* and p*CTT1*, recruit fewer positive activators and will be repressed faster by the negative regulators.

The repression level of a promoter during log-phase growth will determine the speed and the noise of the transcription activation process. Thus, for each promoter, a trade-off has to be found between these two contradictory requirements. For instance, *GPD1*, which is essential for survival to osmotic stress, has an important basal expression level and can thus be rapidly induced upon stress. Interestingly, the chromatin state, encoded in part by the promoter sequence, can be tuned by external growth conditions. Thus, the noise generated by a promoter is not rigidly set by its DNA sequence but fluctuate based on the environment.

In higher eukaryotes, the stress response MAPKs p38 and JNK relocate to the nucleus upon activation[53,54]. Early genes such as c-Fos or c-Jun, are induced within minutes

after activation of signaling[3,55]. Interestingly, these loci display basal expression and require minimal chromatin modification for their induction[56,57]. Conversely, delayed primary response and secondary response genes require more chromatin remodeling to induce their activation[55,58]. These similarities with Hog1-gene transcriptional regulation suggest a high conservation in the mechanisms used by MAPK in eukaryotes to regulate the dynamics of gene expression.

Methods

Plasmids and yeast strains

Plasmids and yeast strains used in this study are listed in Supplementary Tables 1 and 2. All strains were constructed in the W303 background. Transformations were performed with standard lithium-acetate protocols. Gene deletions and gene tagging were performed with either pFA6a cassettes[59,60] or pGT cassettes[61]. Transformants were selected with auxotrophy markers (Uracil, Histidine, Leucine, Tryptophan, Adenine) and gene deletions were performed with antibiotic resistance to Nourseothricin (NAT) or Kanamycin (KAN). In order to generate the membrane anchored Hog1, the pGTT-mCherry vector was modified by inserting annealed oligos encoding the last 30bp of the Ras2 sequence to obtain the pGTT-mCherry-CaaX plasmid. A strain possessing the Hog1-mCherry:*LEU2* modification was transformed with the pGTT-mCherry-CaaX plasmid linearized with *Xba*I and *Sac*I to induce a marker switch and introduce the membrane anchoring motif.

PP7 and MS2 strains construction

The PP7-GFP plasmids are based on the bright and photostable GFPenvy fluorescent protein[32]. The PP7 protein was derived from Larson et al.[30] (Addgene# 35194) with a truncation in the capsid assembly domain (PP7ΔFG residues 67–75: CSTSVCGE[31]). The expression of the PP7 construct is controlled by an *ADH1* promoter and a *CYC1* terminator. The final construct pVW284 is cloned in a Single Integration Vector URA3 (pSIVu[61]). The PP7-mCherry was cloned by replacing the GFP by the mCherry sequence. The MS2-GFP was generated by using the original MS2 sequence from Betrand et al.[18], which also lacks the capsid assembly domain (Addgene# 27117), inserted into the pVW284. The PP7 stem-loops plasmids are based on the previously

published p*POL1* 24x*PP7sl* integrative plasmid[30] (Addgene #35196). The stress responsive promoters replace the *POL1* promoter in the original construct using 1kb (0.8kb for p*STL1*, 0.66kb for p*ALD3*) upstream of the start codon. The p*STL1*-24xMS2sl was generated by replacing the PP7sl with the MS2 stem loops obtained from Betrand et al.[18] (Addgene# 31865).

A strain bearing the Hta2-mCherry nuclear marker and expressing the PP7-GFP was transformed with plasmids containing the different osmostress promoters driving the PP7sl expression, linearized with a *Not*I digestion and integrated upstream of the *GLT1* ORF, as previously published[30]. Correct integration into the *GLT1* locus was screened by colony PCR with primers in the *GLT1* ORF (+600 bp) and in the *TEF* terminator of the selection marker on genomic DNA extractions. The integrity of the PP7 stem-loops array was assessed with primers within the *TEF* terminator and in *GLT1* ORF (+250 bp) for all the promoters and deletions, after each transformation performed. For all the strains used in the study, at least two clones with correct genotypes were isolated and tested during a salt challenge time-lapse experiment. From the data analysis, the most frequent phenotype was isolated and the strain selected.

To tag the endogenous locus of *STL1* with the 24xPP7sl, the plasmid pSP264 with the *STL1* promoter was modified by replacing the *GLT1* ORF sequence by a 500bp sequence starting 100 bp after the start codon of *STL1*. The plasmid was digested *Sac*I-*Not*I and purified over a gel to isolate a fragment that contains the p*STL1*-24xPP7sl-*STL1*$_{100\text{-}600}$. A double-strand break was generated in the *STL1* ORF using Cas9 and a sgRNA targeting the PAM motif GGG 62 bp upstream of the start codon. The Cas9 and sgRNA are expressed from a 2*μ* plasmid (Addgene # 35464[62]) slightly modified from the work from Laughery et al. (Addgene# 67639[63]). The purified DNA fragment containing the stem-loops was used as repair DNA to promote homologous recombination at the *STL1* locus (Supplementary Figure 4a). The correct size of the inserted fragment was verified by colony PCR around the PP7sl insert. Multiple positive clones were screened by microscopy. The results from two transformants are presented in this work to ensure that potentially undesired DNA alterations by Cas9 do not affect the response in the HOG pathway.

458 In order to generate the diploid reporter strain, a MAT**a** strain containing the PP7-
459 mCherry::*URA3*, p*STL1* 24xPP7sl:*GLT1* and Hta2-tdiRFP:*TRP1* was crossed to a
460 MATα strain bearing the MS2-GFP::*URA3*, p*STL1* 24xMS2sl:*GLT1* and Hta2-
461 tdiRFP:*NAT*. Haploid cells were mixed on a YPD plate for a few hours before cells
462 were resuspended in water and spread with beads on a selection plate (SD-TRP
463 +NAT).

464 The plasmids generated for this study are available on Addgene.

465 *Yeast culture*

466 Yeast cells were grown in YPD medium (YEP Broth: CCM0405, ForMedium) for
467 transformation or in Synthetic Defined (SD) medium (YNB:CYN3801/CSM : DCS0521,
468 ForMedium) for microscopy experiments. Before time-lapse experiments, cells were
469 grown at least 24 hours in log-phase. A saturated overnight culture in SD medium was
470 diluted into fresh SD-full medium to OD_{600} 0.025 in the morning and grown for roughly
471 8 hours to reach OD_{600} 0.3-0.5. In the evening, cultures were diluted by adding
472 $(0.5/OD_{600})\mu l$ of cultures in 5ml SD-full for an overnight growth that kept cells in log-
473 phase conditions. Cultures reached an OD_{600} of 0.1-0.3 in the morning of the second
474 day and were further diluted when necessary to remain below an OD_{600} of 0.4 during
475 the day. To prepare the samples, these log-phase cultures were further diluted to an
476 OD_{600} 0.05 and sonicated twice 1min (diploids were not sonicated) before placing 200
477 μl of culture into the well of a 96-well glass bottom plate (MGB096-1-2LG, Matrical
478 Bioscience) previously coated with a filtered solution of Concanavalin A diluted to
479 0.5mg/ml in water (C2010, Sigma-Aldrich)[64]. Cells were let to settle for 30–45 minutes
480 before imaging. Osmotic shock was performed under the microscope, by adding 100
481 μl of a three times concentrated SD-full+NaCl stock solutions to the 200 μl of medium
482 already in the well, to reach the final desired salt concentration.

483 *Microscopy*

484 Images were acquired on a fully automated inverted epi-fluorescence microscope (Ti2-
485 Eclipse, Nikon) placed in an incubation chamber set at 30°C. Excitation was provided
486 by a solid-state light source (SpectraX, Lumencor) and dedicated filter sets were used
487 to excite and detect the proper fluorescence wavelengths with a sCMOS camera
488 (Flash 4.0, Hamamatsu). A motorized XY-stage was used to acquire multiple fields of

views in parallel and a piezo Z-stage (Nano-Z200, Mad City Labs) allowed fast Z-dimension scanning. Micro-manager was used to control the multidimensional acquisitions[65].

Experiments with PP7sl were acquired with a 60X oil objective. For strains with PP7-GFP and Hta2-mCherry, GFP (40ms, 3% LED power) and RFP (20ms), along with two bright field images were recorded every 15 seconds for the GFP and every minute for the other channels, for a total duration of 25 minutes. Six z-stacks were performed on the GFP channels covering ±1.2 μm from the central plane with 0.4 μm steps. An average bleaching of 32% for the GFP and 26% for the RFP for the whole time-lapse was quantified in a strain without the PP7 stem-loops, to avoid artifacts from the appearance of bright fluorescent foci. For all time-lapse experiments, media addition was performed before time point 4, defined as time zero. All microscopy experiments were performed in duplicate for non-induced control experiment and at least triplicate for the NaCl induced experiments.

Flow chamber experiments

The flow experiments were performed in Ibidi chambers (μ-Slide VI 0.4, Ibidi). Two 50ml Falcon tube reservoirs containing SD-full + 0.5 μg/ml fluorescein-dextran (D3305, ThermoFischer) and SD-full + 0.6 M NaCl were put under a pressure of 30mbar (FlowEZ, Fluigent). The media coming from each reservoir were connected using FEP tubing (1/16" OD x 0.020" ID, Fluigent) to a 3-way valve (2-switch, Fluigent). The concentration of NaCl in the medium was controlled using a Pulse-Width Modulation strategy[66,67]. Periods of 4 seconds were used and within this time, the valve controlled the fraction of time when SD-full versus SD-full + NaCl was flowing. TTL signals generated by an Arduino Uno board and dedicated scripts were used to control precisely the switching of the valve. The fluorescein present in the SD-full medium quantified outside the Cell object provided an estimate of the NaCl concentration in the medium. Some strong fluctuations in this signal were probably generated by dust particles in the imaging oil or FLSN-dextran aggregates in the flow chamber. Following 24hrs log-phase growth, cells bearing the p*STL1*-PP7sl reporter, Hog1-mCherry and Hta2-tdiRFP tags were diluted to OD 0.2, briefly sonicated and loaded in the ibidi

519 channel previously coated by Concanavalin A. Cells were left to settle in the channel
520 for 10 minutes before SD-full flow was started.

521 *Raffinose experiment*

522 For the experiments comparing p*STL1*-PP7sl induction in glucose versus raffinose,
523 cells were grown overnight to saturation in SD-full medium. The cultures were diluted
524 to OD 0.025 (Glucose) or 0.05 (Raffinose) and grown at 30°C for at least four hours.
525 In the raffinose medium, the expression level of the PP7-GFP was 2-fold lower than in
526 glucose. Because of this low fluorescence intensity, cells were imaged with a 40X
527 objective, and a single Z-plane was acquired. Manual curation of the images was
528 performed to define the Start Time in more than 250 cells. This experiment was
529 performed in duplicate.

530 *Data analysis*

531 Time-lapse movies were analyzed in an automated way: cell segmentation, tracking
532 and feature measurements were performed by the YeastQuant platform[33]. Summary
533 of the dataset, strains and cell numbers are provided in Supplementary Table 3. All
534 PP7 experiments were realized in at least two or three fully independent replicate
535 experiments. A representative experiment was selected for each strain and inducing
536 conditions, based on cell size and cell adaptation dynamics. The replicates which did
537 not pass one of these controls were discarded from the replicate analyses. Individual
538 single-cell traces were filtered based on cell shape and GFP intensity to remove
539 segmentation errors or other experimental artifacts. In addition, cells in mitosis were
540 removed from the analysis with a 0.95 filter on the nuclei eccentricity, to remove
541 artifacts from locus and PP7 signal duplication.

542 The Hta2 signal combined with the two bright field images allowed to define the
543 nucleus and cell borders. The GFP z-stacks were converted by a maximum intensity
544 projection in a single image that was used for quantification. In order to avoid improper
545 quantification of transcription sites at the nuclear periphery, the Nucleus object defined
546 by the histone fluorescence was expanded by 5 pixels within the Cell object to define
547 the ExpNucl object. The transcription site intensity was quantified by the difference
548 between the mean intensity of the 20 brightest pixels in the ExpNucl (HiPix) and the
549 median intensity from the whole cell. This provides a continuous trace which is close

to zero in absence of TS and increases by up to few hundred counts when a TS is present. To identify the presence of a transcription site, a second feature named ConnectedHiPix was used (Supplementary Figure 2a). Starting from the 20 HiPix, a morphological opening of the image was performed to remove isolated pixels and retaining only the ones that clustered together which correspond to the transcription site. The ConnectedHiPix value was set to the mean intensity of the pixel present in the largest object remaining after the morphological operation. If no pixel remained after the morphological operation, the ConnectedHiPix was set to NaN. In each single-cell trace, ConnectedHiPix values only detected for a single time point were removed. After this filtering, the first and last time points where a ConnectedHiPix was measured were defined as transcription initiation (Start Time) and shutoff (End Time), respectively. Manual curation of Start and End Times from raw microscopy images was performed to validate this transcription site detection strategy (Supplementary Figure 2b and c). In order to detect individual transcriptional bursts in the HiPix traces, the *findpeak* algorithm was used to identify in the trace all the peaks larger than a threshold of 7 counts within the Start and End Times. Following this first process, a set of conditions were defined to retain only the more reliable fluctuations: the drop following the peak has to be larger the fourth of the peak intensity; the intensity of the following peak has to rise by more than a third of the value at the trough. In addition, the value of the peak has to be at least one fifth of the maximum intensity of the trace in order to remove small intensity fluctuations being considered as peaks.

Data Availability

Source data for Figures 1d, 2a, 2f,2g, 3a, 4b, 4c, and 6c and Supplementary Figures 1, 4, 5, 6, 8, 12, 13 are provided with the paper. The raw images and additional features measurements that support the findings of this study are available from the corresponding author upon reasonable request.

Code Availability

The image analysis platform has been published previously[33]. A more recent version of the code can be obtained from the corresponding author. A script to extract measured parameters from the data is provided as a supplementary file.

581

582 ## References

583 1. Gasch, A. P. *et al.* Genomic expression programs in the response of yeast cells to
584 environmental changes. *Mol Biol Cell* **11**, 4241–4257 (2000).
585 2. Roberts, C. J. *et al.* Signaling and circuitry of multiple MAPK pathways revealed by a matrix
586 of global gene expression profiles. *Science* **287**, 873–880 (2000).
587 3. Ferreiro, I. *et al.* Whole genome analysis of p38 SAPK-mediated gene expression upon
588 stress. *BMC Genomics* **11**, 144 (2010).
589 4. Berry, D. B. & Gasch, A. P. Stress-activated genomic expression changes serve a
590 preparative role for impending stress in yeast. *Mol Biol Cell* **19**, 4580–4587 (2008).
591 5. Chen, R. E., Patterson, J. C., Goupil, L. S. & Thorner, J. Dynamic localization of Fus3
592 mitogen-activated protein kinase is necessary to evoke appropriate responses and avoid
593 cytotoxic effects. *Mol Cell Biol* **30**, 4293–4307 (2010).
594 6. Formstecher, E. *et al.* PEA-15 Mediates Cytoplasmic Sequestration of ERK MAP Kinase.
595 *Developmental Cell* **1**, 239–250 (2001).
596 7. Raser, J. M. & O'Shea, Erin K. Control of stochasticity in eukaryotic gene expression.
597 *Science* **304**, 1811–1814 (2004).
598 8. Colman-Lerner, A. *et al.* Regulated cell-to-cell variation in a cell-fate decision system.
599 *Nature* **437**, 699–706 (2005).
600 9. Pelet, S. *et al.* Transient activation of the HOG MAPK pathway regulates bimodal gene
601 expression. *Science* **332**, 732–735 (2011).
602 10. Corrigan, A. M. & Chubb, J. R. Regulation of Transcriptional Bursting by a Naturally
603 Oscillating Signal. *Current Biology* **24**, 205–211 (2014).
604 11. Roux, P. P. & Blenis, J. ERK and p38 MAPK-Activated Protein Kinases: a Family of Protein
605 Kinases with Diverse Biological Functions. *Microbiology and Molecular Biology Reviews*
606 **68**, 320–344 (2004).
607 12. Chen, R. E. & Thorner, J. Function and regulation in MAPK signaling pathways: lessons
608 learned from the yeast Saccharomyces cerevisiae. *Biochim Biophys Acta* **1773**, 1311–
609 1340 (2007).
610 13. Saito, H. & Posas, F. Response to hyperosmotic stress. *Genetics* **192**, 289–318 (2012).
611 14. Hohmann, S., Krantz, M. & Nordlander, B. Yeast osmoregulation. *Meth Enzymol* **428**, 29–
612 45 (2007).
613 15. de Nadal, E. & Posas, F. Multilayered control of gene expression by stress-activated
614 protein kinases. *EMBO J* **29**, 4–13 (2010).
615 16. de Nadal, E., Ammerer, G. & Posas, F. Controlling gene expression in response to stress.
616 *Nat Rev Genet* **12**, 833–845 (2011).
617 17. Neuert, G. *et al.* Systematic Identification of Signal-Activated Stochastic Gene Regulation.
618 *Science* **339**, 584–587 (2013).
619 18. Bertrand, E. *et al.* Localization of ASH1 mRNA particles in living yeast. *Mol Cell* **2**, 437–
620 445 (1998).
621 19. Buxbaum, A. R., Haimovich, G. & Singer, R. H. In the right place at the right time:
622 visualizing and understanding mRNA localization. *Nat Rev Mol Cell Biol* **16**, 95–109
623 (2015).
624 20. Urbanek, M. O., Galka-Marciniak, P., Olejniczak, M. & Krzyzosiak, W. J. RNA imaging in
625 living cells – methods and applications. *RNA Biology* **11**, 1083–1095 (2014).
626 21. Lionnet, T. & Singer, R. H. Transcription goes digital. *EMBO Rep* **13**, 313–321 (2012).
627 22. Fritzsch, C. *et al.* Estrogen-dependent control and cell-to-cell variability of transcriptional
628 bursting. *Molecular Systems Biology* **14**, e7678 (2018).

629 23. Zoller, B., Little, S. C. & Gregor, T. Diverse Spatial Expression Patterns Emerge from
630 Unified Kinetics of Transcriptional Bursting. *Cell* **175**, 835-847.e25 (2018).
631 24. Reiser, V., Ruis, H. & Ammerer, G. Kinase activity-dependent nuclear export opposes
632 stress-induced nuclear accumulation and retention of Hog1 mitogen-activated protein
633 kinase in the budding yeast Saccharomyces cerevisiae. *Mol Biol Cell* **10**, 1147–1161
634 (1999).
635 25. Muzzey, D., Gómez-Uribe, C. A., Mettetal, J. T. & van Oudenaarden, A. A systems-level
636 analysis of perfect adaptation in yeast osmoregulation. *Cell* **138**, 160–171 (2009).
637 26. Görner, W. *et al.* Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated
638 by stress and protein kinase A activity. *Genes Dev.* **12**, 586–597 (1998).
639 27. Hao, N. & O'Shea, E. K. Signal-dependent dynamics of transcription factor translocation
640 controls gene expression. *Nature Structural & Molecular Biology* **19**, 31–39 (2012).
641 28. O'Rourke, S. M. & Herskowitz, I. Unique and Redundant Roles for HOG MAPK Pathway
642 Components as Revealed by Whole-Genome Expression Analysis. *Mol Biol Cell* **15**, 532–
643 542 (2003).
644 29. Capaldi, A. P. *et al.* Structure and function of a transcriptional network activated by the
645 MAPK Hog1. *Nat Genet* **40**, 1300–1306 (2008).
646 30. Larson, D. R., Zenklusen, D., Wu, B., Chao, J. A. & Singer, R. H. Real-Time Observation
647 of Transcription Initiation and Elongation on an Endogenous Yeast Gene. *Science* **332**,
648 475–478 (2011).
649 31. Chao, J. A., Patskovsky, Y., Almo, S. C. & Singer, R. H. Structural basis for the coevolution
650 of a viral RNA–protein complex. *Nat. Struct. Mol. Biol.* **15**, 103–105 (2007).
651 32. Slubowski, C. J., Funk, A. D., Roesner, J. M., Paulissen, S. M. & Huang, L. S. Plasmids
652 for C-terminal tagging in Saccharomyces cerevisiae that contain improved GFP proteins,
653 Envy and Ivy. *Yeast* **32**, 379–387 (2015).
654 33. Pelet, S., Dechant, R., Lee, S. S., van Drogen, F. & Peter, M. An integrated image analysis
655 platform to quantify signal transduction in single cells. *Integrative biology : quantitative*
656 *biosciences from nano to macro* **4**, 1274–1282 (2012).
657 34. de Nadal, E. *et al.* The MAPK Hog1 recruits Rpd3 histone deacetylase to activate
658 osmoresponsive genes. *Nature* **427**, 370–374 (2004).
659 35. Mas, G. *et al.* Recruitment of a chromatin remodelling complex by the Hog1 MAP kinase
660 to stress genes. *EMBO J* **28**, 326–336 (2009).
661 36. Li, G. & Neuert, G. Multiplex RNA single molecule FISH of inducible mRNAs in single yeast
662 cells. *Sci Data* **6**, 94 (2019).
663 37. Aymoz, D., Wosika, V., Durandau, E. & Pelet, S. Real-time quantification of protein
664 expression at the single-cell level via dynamic protein synthesis translocation reporters.
665 *Nature Communications* **7**, 11304 (2016).
666 38. Rep, M. *et al.* Osmotic stress-induced gene expression in Saccharomyces cerevisiae
667 requires Msn1p and the novel nuclear factor Hot1p. *Mol Cell Biol* **19**, 5474–5485 (1999).
668 39. Alepuz, P. M., de Nadal, E., Zapater, M., Ammerer, G. & Posas, F. Osmostress-induced
669 transcription by Hot1 depends on a Hog1-mediated recruitment of the RNA Pol II. *EMBO*
670 *J* **22**, 2433–2442 (2003).
671 40. Zapater, M., Sohrmann, M., Peter, M., Posas, F. & de Nadal, E. Selective requirement for
672 SAGA in Hog1-mediated gene expression depending on the severity of the external
673 osmostress conditions. *Mol Cell Biol* **27**, 3900–3910 (2007).
674 41. Wan, Y. *et al.* Role of the histone variant H2A.Z/Htz1p in TBP recruitment, chromatin
675 dynamics, and regulated expression of oleate-responsive genes. *Mol Cell Biol* **29**, 2346–
676 2358 (2009).
677 42. Ferreira, C. & Lucas, C. Glucose repression over Saccharomyces cerevisiae glycerol/H+
678 symporter gene STL1 is overcome by high temperature. *FEBS Lett* **581**, 1923–1927
679 (2007).
680 43. de Nadal, E. & Posas, F. Regulation of gene expression in response to osmostress by the
681 yeast stress-activated protein kinase Hog1. *Topics in Current Genetics* **20**, 81 (2008).

682 44. Sharon, E. *et al.* Inferring gene regulatory logic from high-throughput measurements of
683 thousands of systematically designed promoters. *Nat Biotechnol* **30**, 521–530 (2012).
684 45. Westfall, P. J., Patterson, J. C., Chen, R. E. & Thorner, J. Stress resistance and signal
685 fidelity independent of nuclear MAPK function. *Proc Natl Acad Sci USA* **105**, 12212–12217
686 (2008).
687 46. Tantale, K. *et al.* A single-molecule view of transcription reveals convoys of RNA
688 polymerases and multi-scale bursting. *Nat Commun* **7**, 12248 (2016).
689 47. Pokholok, D. K., Zeitlinger, J., Hannett, N. M., Reynolds, D. B. & Young, R. A. Activated
690 signal transduction kinases frequently occupy target genes. *Science* **313**, 533–536 (2006).
691 48. Proft, M. *et al.* The stress-activated Hog1 kinase is a selective transcriptional elongation
692 factor for genes responding to osmotic stress. *Mol Cell* **23**, 241–250 (2006).
693 49. Macia, J. *et al.* Dynamic signaling in the Hog1 MAPK pathway relies on high basal signal
694 transduction. *Science Signaling* **2**, ra13 (2009).
695 50. O'Sullivan, J. M. *et al.* Gene loops juxtapose promoters and terminators in yeast. *Nat*
696 *Genet* **36**, 1014–1018 (2004).
697 51. Ansari, A. A role for the CPF 3'-end processing machinery in RNAP II-dependent gene
698 looping. *Genes & Development* **19**, 2969–2978 (2005).
699 52. Klopf, E. *et al.* Cooperation between the INO80 complex and histone chaperones
700 determines adaptation of stress gene transcription in the yeast Saccharomyces cerevisiae.
701 *Mol Cell Biol* **29**, 4994–5007 (2009).
702 53. Cavigelli, M., Dolfi, F., Claret, F. X. & Karin, M. Induction of c-fos expression through JNK-
703 mediated TCF/Elk-1 phosphorylation. *The EMBO Journal* **14**, 5957–5964 (1995).
704 54. Wood, C. D., Thornton, T. M., Sabio, G., Davis, R. A. & Rincon, M. Nuclear Localization of
705 p38 MAPK in Response to DNA Damage. *Int. J. Biol. Sci.* 428–437 (2009)
706 doi:10.7150/ijbs.5.428.
707 55. Tullai, J. W. *et al.* Immediate-Early and Delayed Primary Response Genes Are Distinct in
708 Function and Genomic Architecture. *J. Biol. Chem.* **282**, 23981–23995 (2007).
709 56. Fowler, T., Sen, R. & Roy, A. L. Regulation of Primary Response Genes. *Molecular Cell*
710 **44**, 348–360 (2011).
711 57. O'Donnell, A., Odrowaz, Z. A. & Sharrocks, A. D. Immediate-early gene activation by the
712 MAPK pathways: what do and don't we know? *Biochemical Society transactions* **40**, 58–
713 66 (2012).
714 58. Ramirez-Carrozzi, V. R. *et al.* Selective and antagonistic functions of SWI/SNF and Mi-2β
715 nucleosome remodeling complexes during an inflammatory response. *Genes Dev.* **20**,
716 282–296 (2006).
717 59. Goldstein, A. L. & McCusker, J. H. Three new dominant drug resistance cassettes for gene
718 disruption in Saccharomyces cerevisiae. *Yeast* **15**, 1541–1553 (1999).
719 60. Sheff, M. A. & Thorn, K. S. Optimized cassettes for fluorescent protein tagging in
720 Saccharomyces cerevisiae. *Yeast* **21**, 661–670 (2004).
721 61. Wosika, V. *et al.* New families of single integration vectors and gene tagging plasmids for
722 genetic manipulations in budding yeast. *Molecular Genetics and Genomics* **291**, 2231–
723 2240 (2016).
724 62. Chee, M. K. & Haase, S. B. New and Redesigned pRS Plasmid Shuttle Vectors for Genetic
725 Manipulation of Saccharomyces cerevisiae. *G3 (Bethesda)* **2**, 515–526 (2012).
726 63. Laughery, M. F. *et al.* New vectors for simple and streamlined CRISPR-Cas9 genome
727 editing in Saccharomyces cerevisiae. *Yeast* **32**, 711–720 (2015).
728 64. Pelet, S., Aymoz, D. & Durandau, E. Temporal quantification of MAPK induced expression
729 in single yeast cells. *J Vis Exp* (2013) doi:10.3791/50637.
730 65. Edelstein, A., Amodaj, N., Hoover, K., Vale, R. & Stuurman, N. Computer control of
731 microscopes using *μ*Manager. *Curr Protoc Mol Biol* **Chapter 14**, Unit14.20 (2010).
732 66. Unger, M., Lee, S.-S., Peter, M. & Koeppl, H. Pulse Width Modulation of Liquid Flows:
733 Towards Dynamic Control of Cell Microenvironments. in *15th International Conference on*

miniaturized systems for chemistry and life sciences : Microtas 2011 1567–1569 (Chemical and Biological Microsystems Society, 2011).
67. Unger, M. P. Interrogating the single cell: computational and experimental methods for optimal live cell experiments. (ETH Zurich, 2014). doi:10.3929/ethz-a-010350761.

End Notes

Acknowledgment

We thank members the Pelet lab and Martin lab and for helpful discussions. Marta Schmitt, Yves Dusserre, Gaëlle Spack, Joan Jordan and Clémence Varidel for technical assistance. David Shore and his lab for helpful discussions and reagents. Marie-Pierre Peli-Gulli and Claudio de Virgilio for plasmids, Tineke Lenstra for suggesting the PP7ΔFG allele. Agathe Pelet for manual curation of microscopy images. Eulalia de Nadal, Mariona Nadal-Ribelles and Veneta Gerganova for critically reading the manuscript. Work in the Pelet lab is supported by SystemsX.ch (IPhD 51PHP0_157354), the Swiss National Science Foundation (SNSF, PP00P3_172900 and 31003A_182431) and the University of Lausanne.

Author Contributions

VW and SP designed the experiments, analyzed the data and wrote the manuscript. VW established the condition for the PP7 imaging. VW and SP performed the experiments.

Declaration of Interests

The authors declare no competing interests.

Figure Legends

Fig. 1. Monitoring the dynamics of osmostress-genes transcription.

a. Schematics of the transcriptional response induced by the MAPK Hog1 upon osmotic stress. Under normal growth conditions, the genomic locus is repressed by histones set in place by the Ino80 complex and Asf1/Rtt109. In addition, H3K4 methylated histones mediated by Set1 contribute to the further repression of the locus

763 (upper panel). When Hog1 is active (lower panel), it accumulates in the nucleus with
764 the transcription factors Msn2/4. Hog1 binds to the transcription factors Hot1 and Sko1,
765 allowing the remodeling of the chromatin by Rpd3 and the SAGA complex. The
766 polymerases can be recruited to the locus and the RSC and SWR complexes evict
767 nucleosomes on the ORF. **b.** Construction of the transcriptional reporter. The promoter
768 of interest (p*PROM*) is cloned in front of 24 stem-loops (*24xPP7sl*). This construct is
769 transformed in yeast and integrated in the *GLT1* locus 5'UTR, replacing the
770 endogenous promoter. Upon induction of the promoter, the mRNA stem-loops are
771 transcribed and recognized by the fluorescently-tagged PP7 phage coat protein. **c.**
772 Maximum intensity projections of Z-stacks of microscopy images from the p*STL1*-
773 PP7sl reporter system in a 0.2M NaCl osmotic stress time-lapse experiment**.** The
774 appearance of bright foci (arrow heads) in the nucleus of the cells denotes the active
775 transcription arising from the promoter**.** Scale bar represents 5 μm. **d.** Dynamics of the
776 p*STL1*-PP7sl transcription site intensity (20 brightest pixels in the nucleus minus the
777 median fluorescence of the cell) following hyperosmotic stress. The mean from 200 to
778 400 cells is represented by the solid line. The shaded areas represent the s.e.m**. e.**
779 Analysis of one representative single-cell trace. The raw trace (gray) is smoothed with
780 a moving average (dark blue) and normalized by subtracting the intensity of the first
781 time point after the stimulus. Multiple quantitative values can be extracted from this
782 trace (see Methods). Source data are provided for d.

783

784 **Fig. 2. Chromatin state dictates the transcription initiation of stress-induced**
785 **promoters**

786 **a.** Dynamics of the transcription site intensity from six different promoters following a
787 0.2M NaCl stress. The mean of at least 140 cells is represented by the solid line. The
788 shaded areas represent the s.e.m. **b.** Percentage of cells where a PP7 TS site was
789 detected. The light shaded area represents the percentage of PP7 positive cells before
790 the stimulus was added (basal transcription). **c.** The microscopy thumbnails display
791 cells bearing the p*GPD1*-PP7sl reporter system, where transcription sites (arrow
792 heads) can be detected before and after the stress of 0.2M NaCl. Scale bar represents
793 5 μm. **d.** Cumulative distribution function (CDF) of the Start Time for each promoter

794 considering only the cells that induce transcription after time zero. **e.** 10th, 50th and 90th
795 percentiles of the Start Times shown for the two to three replicates measured for each
796 promoter. The number of stars next to each measurement corresponds to the number
797 of promoters without basal level that are significantly different from the promoter with
798 basal level (two-sample *t*-test, p<0.05). **f.** Cumulative distribution functions of Start
799 Times for the p*STL1*-PP7sl strain in wild type, *htz1*Δ or *gcn5*Δ backgrounds. The inset
800 shows the percentage of PP7 positive cells in each background. **g.** Cumulative
801 distribution functions of Start Times for the p*STL1*-PP7sl strain grown in glucose or
802 raffinose. The inset shows the percentage of PP7 positive cells, the light blue bar the
803 basal positive PP7 cells. Source data are provided for a, f, g.

804

805 **Fig. 3. Early Hog1 activity dictates promoter activation and output**

806 **a.** In Hog1 nuclear relocation traces obtained from single cells, the timing of Hog1
807 nuclear entry (▬), maximum enrichment ($), start of the decay in nuclear enrichment
808 (◆) and Hog1 adaptation (▲) can be identified (upper panel). The median (marker)
809 and 25th to 75th percentiles (lines) for these measurements are plotted for three
810 different osmotic stresses (central panel). The cumulative distribution functions of Start
811 Times for the p*STL1*-PP7sl reporter for these same three concentrations are plotted
812 (lower panel). **b.** Histograms of Start Times following a 0.2M stress for the five other
813 promoters tested. The vertical dashed line represents the median decay time of Hog1
814 measured at 0.2M. The number in the legend indicates the percentage of cells which
815 have initiated transcription before the median Hog1 decay time. **c.** The population of
816 p*STL1*-PP7sl positive cells is split in four quartiles based on their Start Time. The
817 median ($) and 25th to 75th percentiles (line) of the integral of the PP7 trace is plotted
818 for each quartile. Source data are provided for a.

819

820 **Fig. 4. Transcription factors control the dynamics and level of mRNA production**

821 **a.** Violin plots of the trace intensity (maximum of the TS during the transcription period)
822 for the six promoters after stimulation by 0.2M NaCl. Each dot represents the value
823 calculated from a single cell. The solid line is the median and the dashed line the mean

824 of the population. **b.** Comparison between the trace intensity in stimulated (0.2M NaCl)
825 and unstimulated conditions (0.0M) for the three promoters displaying basal
826 expression. **c. - f.** Effect of the deletions of the *HOT1* and *SKO1* transcription factor
827 genes on p*STL1* and p*GPD1* dynamics of transcription (c), cumulative distribution
828 functions of Start Times (d), the trace intensity (e) and the percentage of responding
829 cells (f) for the p*STL1*-PP7sl and p*GPD1*-PP7sl reporter strains following a 0.2M NaCl
830 stress for at least 200 cells. For the p*STL1*-PP7sl *hot1Δ* sample, 349 cells were
831 analyzed with only 9 displaying a PP7 positive signal. This low number does not allow
832 to draw a meaningful CDF curve in panel d. Source data are provided for b and c.

833

834 **Fig. 5. Identification of transcription bursts in stress-induced transcription.**

835 **a.** Percentage of cells where 1, 2 and 3 or more peaks are identified among the
836 population of responding cells for the different promoters following a 0.2M NaCl stress.
837 **b.** Examples of single-cell traces displaying 1 or 2 peaks for the p*STL1*-PP7sl and the
838 p*GPD1*-PP7sl reporter strains. **c.** Violin plots representing the Peak Duration. Each dot
839 represents the value calculated for a single peak. The solid line is the median and the
840 dashed line the mean of all the peaks measured. **d. - e.** The population of cells was
841 split between cells displaying one peak and two or more peaks. The Peak Duration (d)
842 and Trace Intensity (e) are plotted for the p*HSP12*-PP7sl and p*GPD1*-PP7sl strains.
843 Each dot represents the value calculated for a single peak (d) or a single cell (e). The
844 solid line is the median and the dashed line the mean of the population.

845

846 **Fig. 6. Hog1 activity and promoter identity control the shutoff of transcription**

847 **a.** Violin plots representing the Transcription Period (time difference between End Time
848 and Start Time) measured for the p*STL1*-PP7sl reporter following 0.1, 0.2 and 0.3M
849 NaCl stresses. Each dot represents the value calculated from a single cell. The solid
850 line is the median and the dashed line the mean of the population. **b.** One minus the
851 cumulative distribution function of End Times for the p*STL1*-PP7sl reporter. The
852 vertical dashed lines represent the median adaptation time of Hog1 for the three
853 different stress levels. **c.** Dynamics of the estimated NaCl concentrations in the
854 medium for the pulse, step and ramp experiment protocols (upper panel, Methods).

855 Corresponding Hog1 relocation dynamics (middle panel) and p*STL1*-PP7sl
856 transcription site intensity (lower panel). The mean of at least 180 cells is represented
857 by the solid line. The shaded areas represent the s.e.m. **d.** Cumulative distribution
858 function (CDF) of the Start Time for all cells in the pulse, step and ramp experiments.
859 The CDF at 15 min represents the fraction of responding cells for each condition. **e.**
860 One minus the cumulative distribution function of End Times only for the responding
861 cells in the pulse, step and ramp experiments. **f.** Correlation between the Hog1
862 adaptation time and the PP7 End Time measured in the same cells in the pulse, step
863 and ramp experiments. The open markers indicate cells where Hog1 has not adapted
864 at the end of the time lapse. Adaptation time is arbitrarily set to 35 min for this sub-
865 population. Source data are provided for c.

866

867 **Fig. 7. Transcription shutoff from different promoters**

868 **a.** Violin plots representing the Transcription Period measured for the six different
869 promoters following a 0.2 NaCl stress. Each dot represents the value calculated from
870 a single cell. The solid line is the median and the dashed line the mean of the
871 population. **b.** One minus the cumulative distribution function of End Times for the
872 different promoters. The vertical dashed line represents the median adaptation time of
873 Hog1 at 0.2M NaCl.

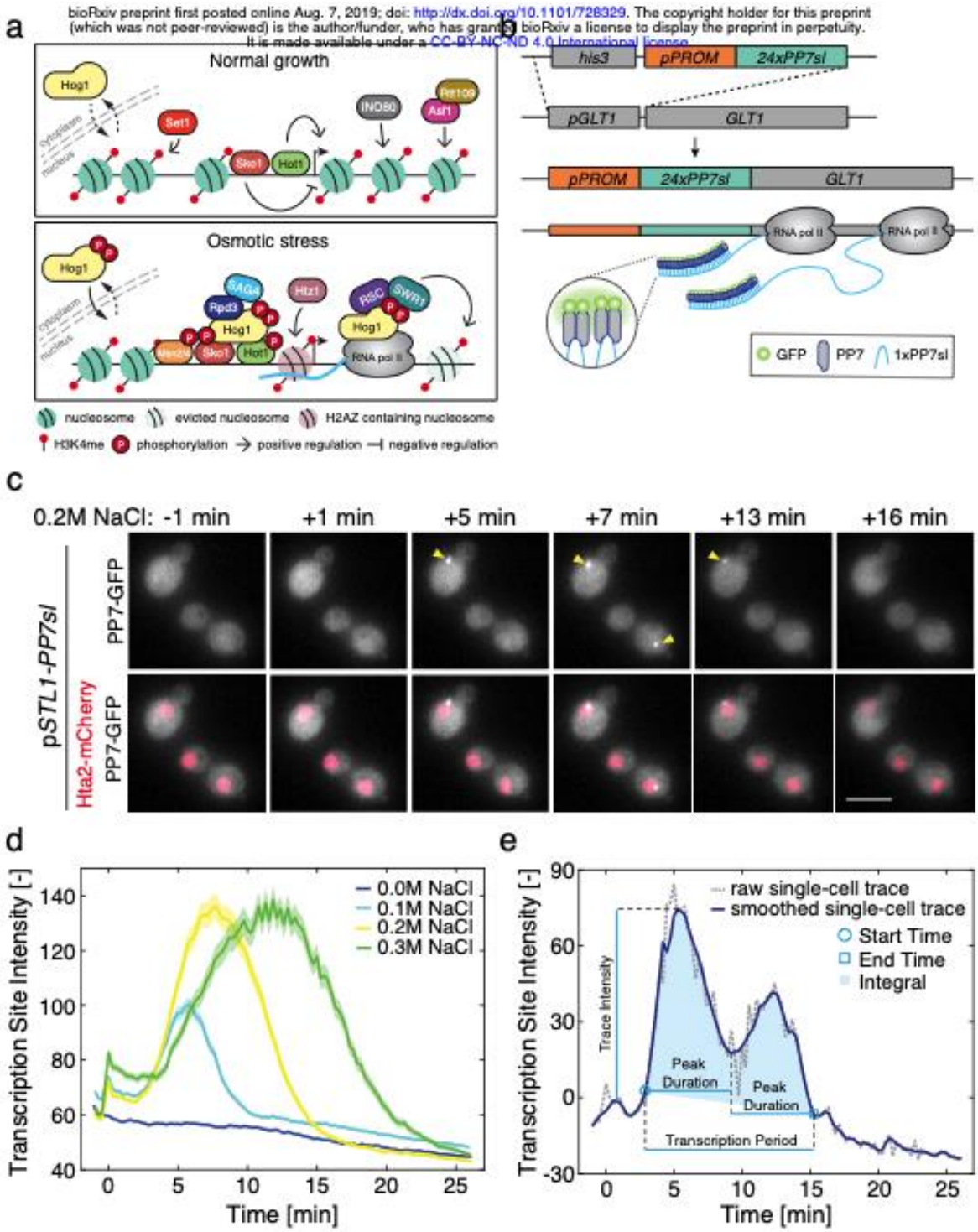


Figure 1

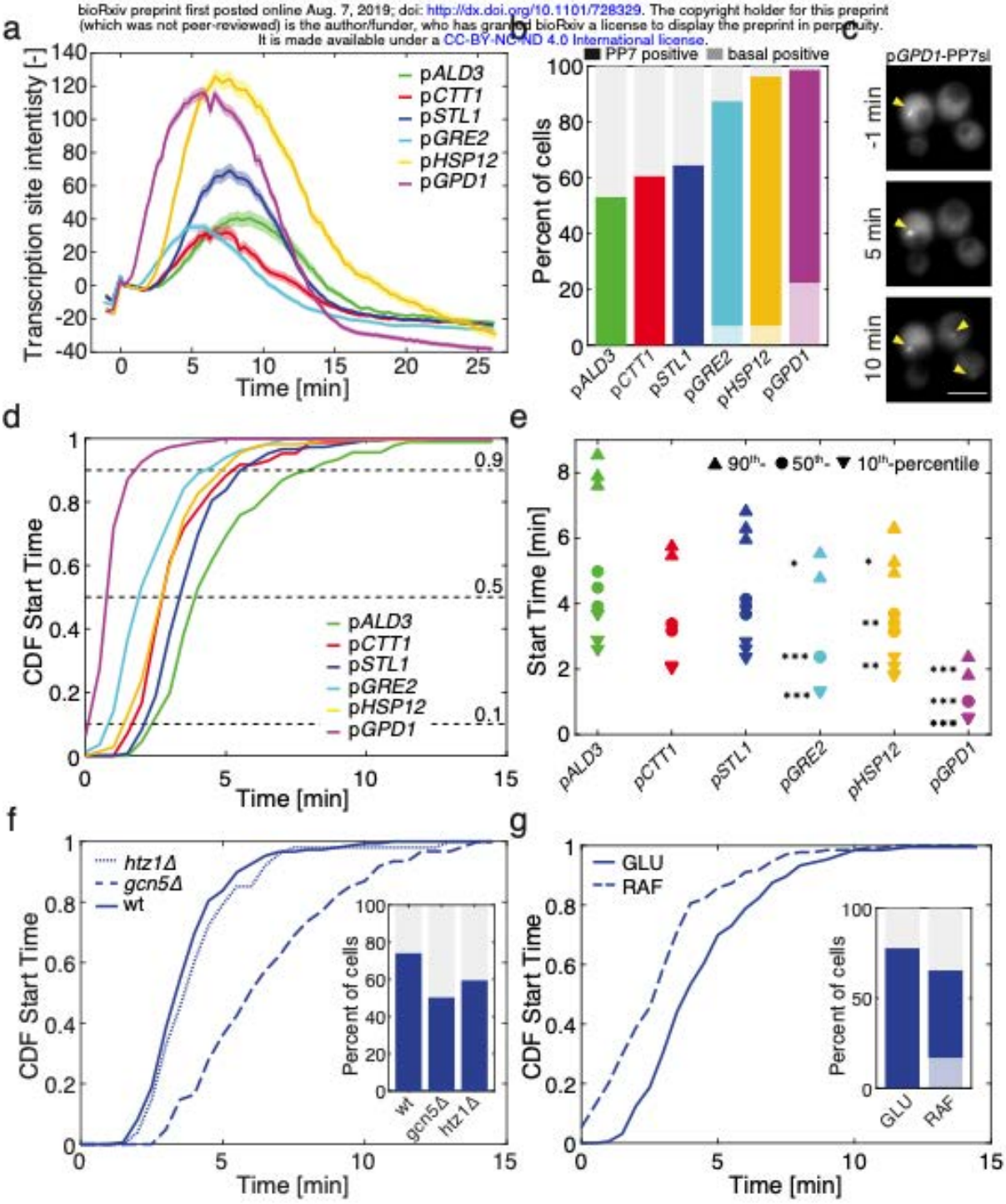


Figure 2

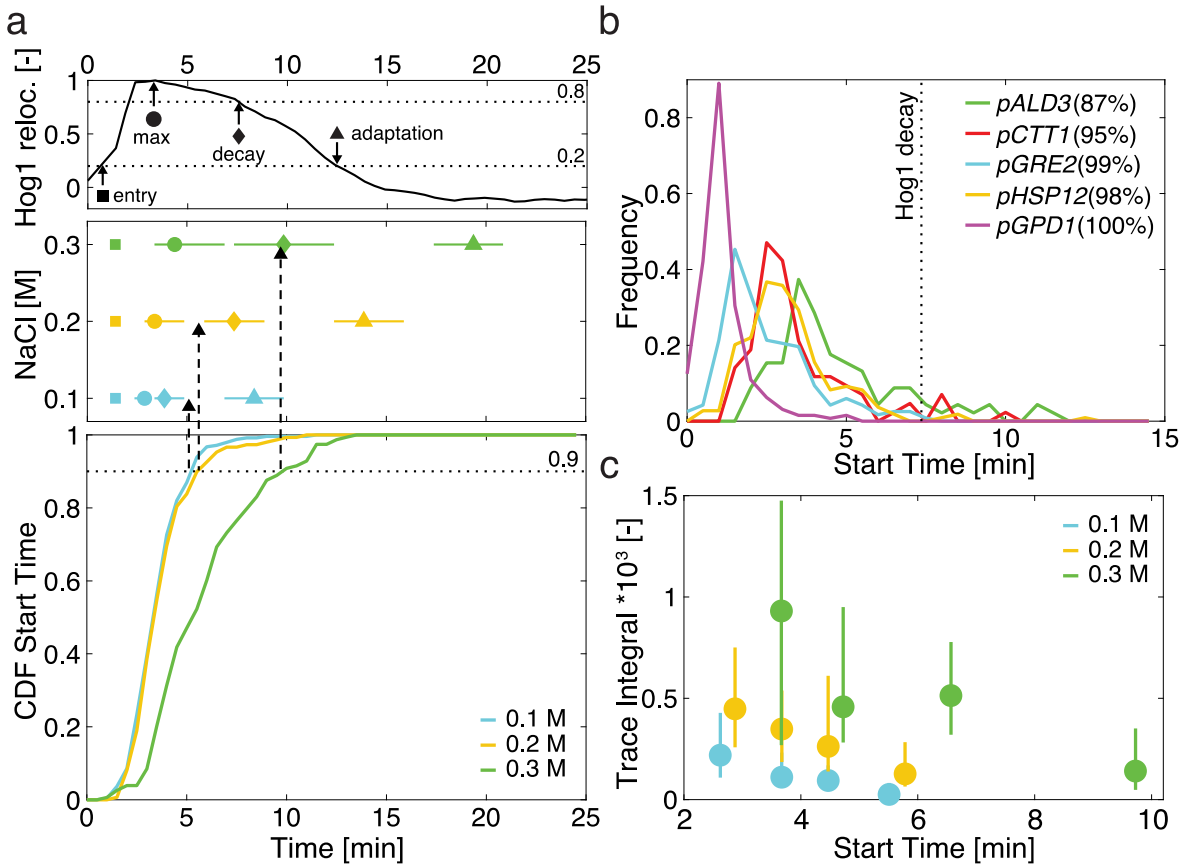

a
b
c
Hog1 reloc. [-]
max
decay
adaptation
entry
0.8
0.2
NaCl [M]
CDF Start Time
0.9
0.1 M
0.2 M
0.3 M
Time [min]
Frequency
Hog1 decay
pALD3(87%)
pCTT1(95%)
pGRE2(99%)
pHSP12(98%)
pGPD1(100%)
Start Time [min]
Trace Integral *10³ [-]
0.1 M
0.2 M
0.3 M
Start Time [min]

Figure 3

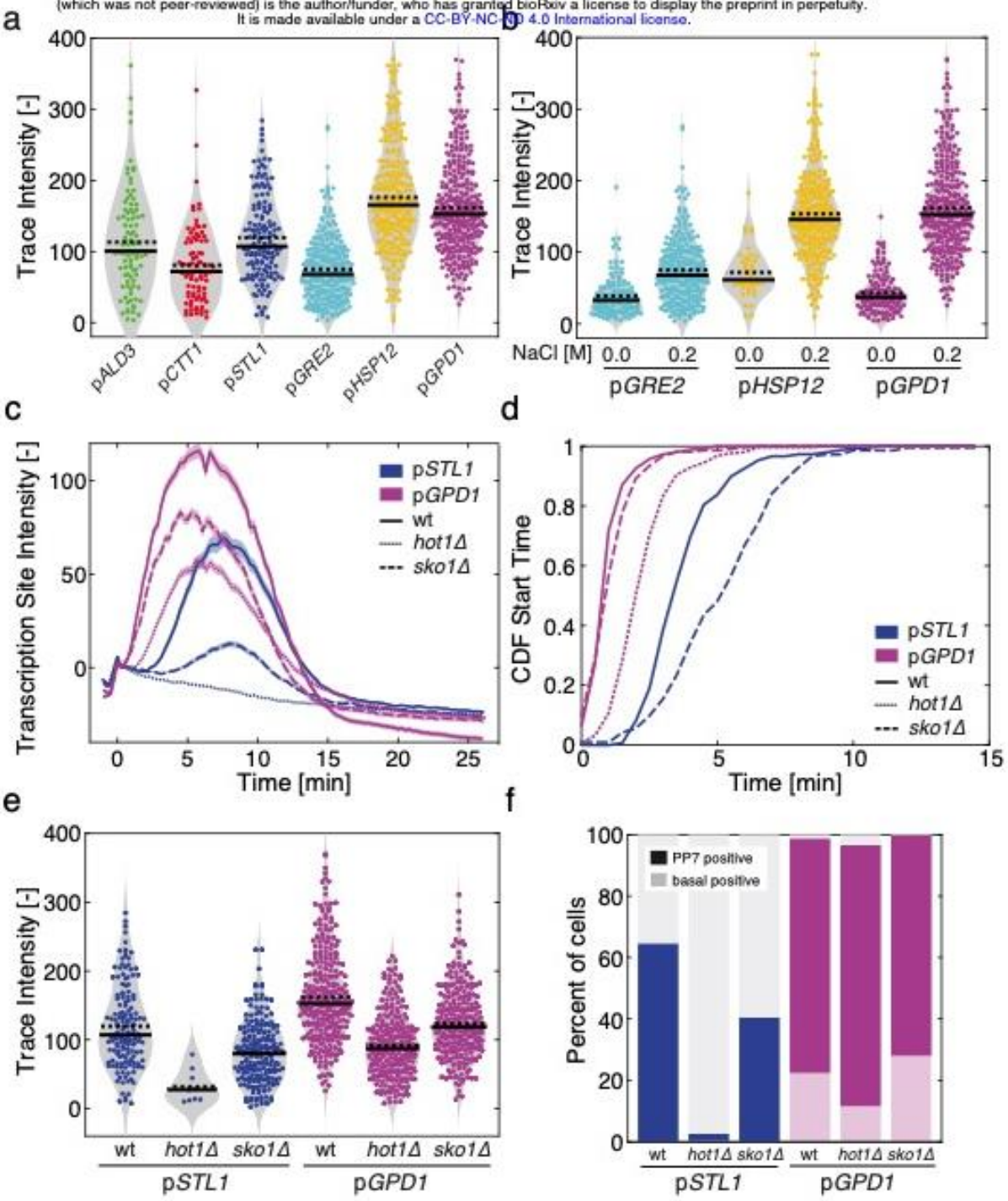


Figure 4

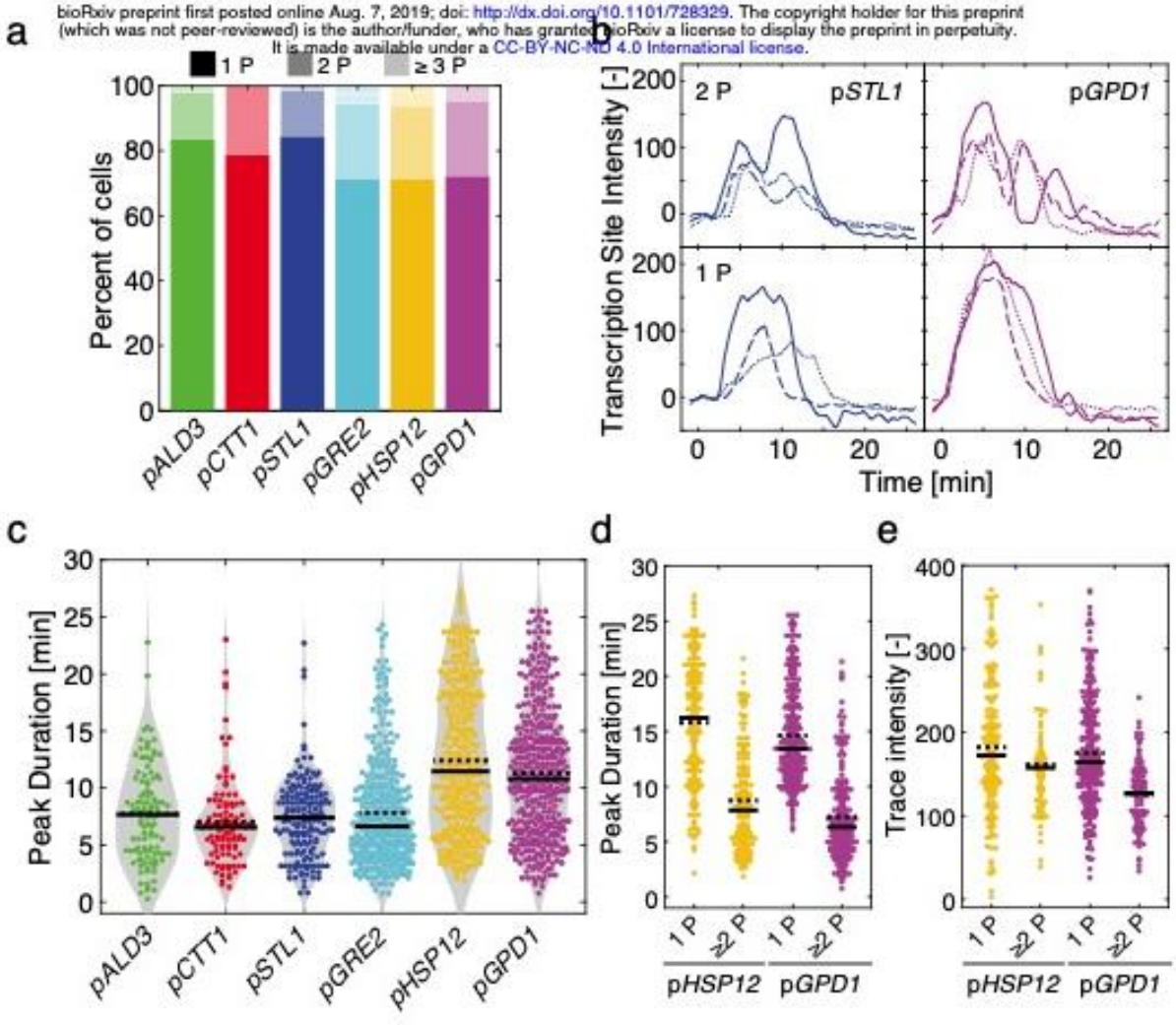


Figure 5

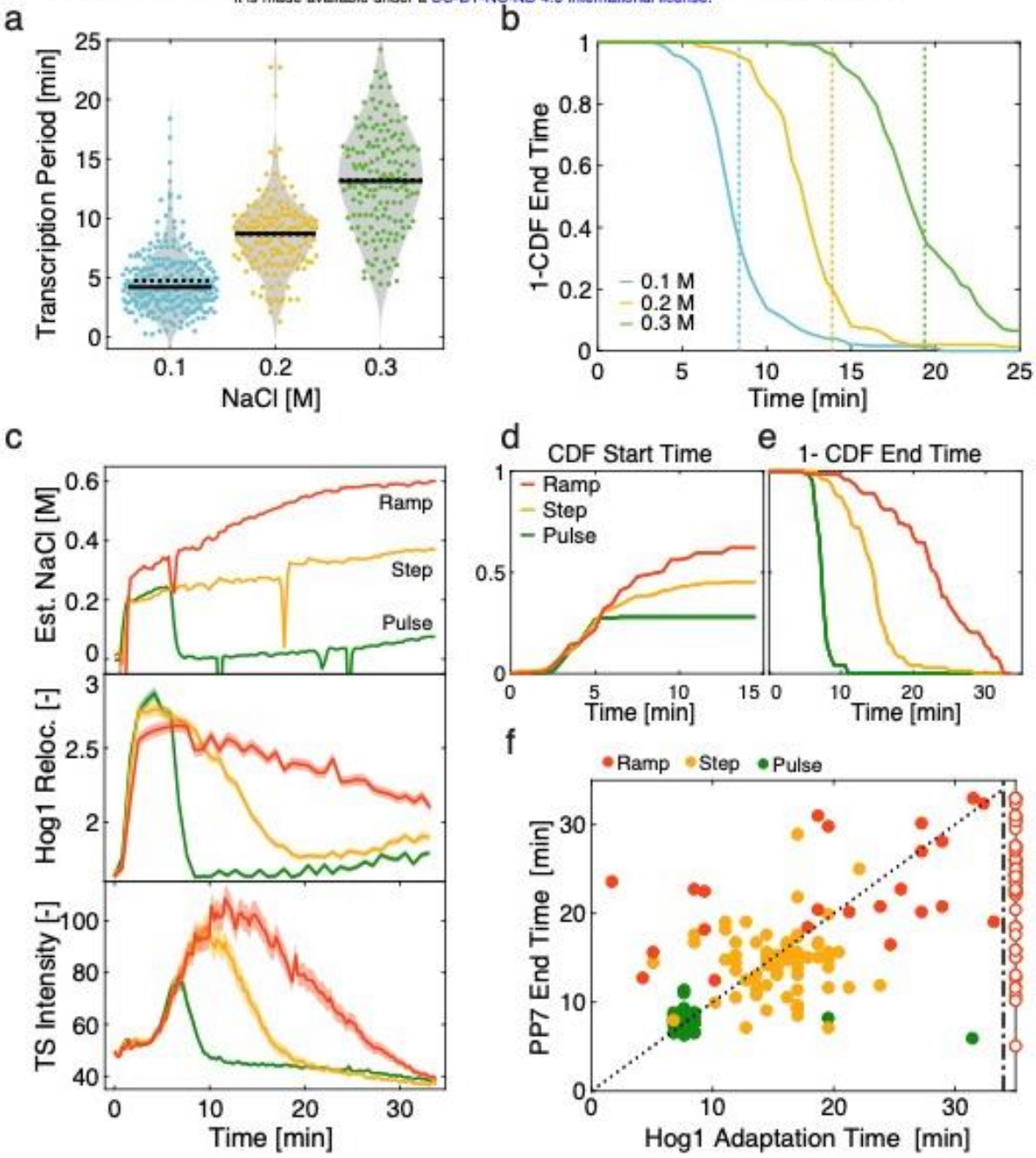


Figure 6

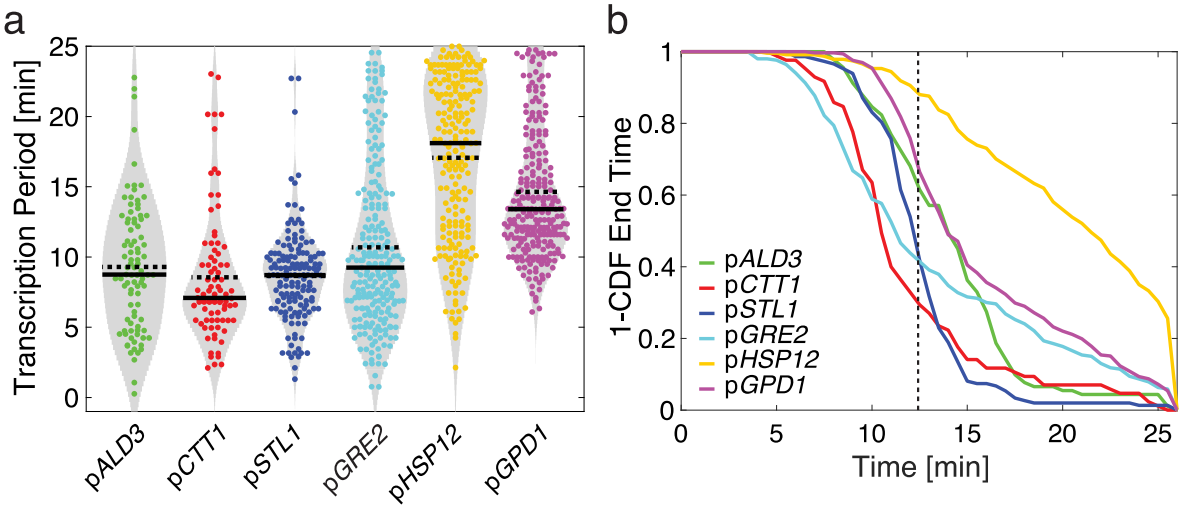

a
Transcription Period [min]
25
20
15
10
5
0
pALD3
pCTT1
pSTL1
pGRE2
pHSP12
pGPD1
b
1-CDF End Time
1
0.8
0.6
0.4
0.2
0
pALD3
pCTT1
pSTL1
pGRE2
pHSP12
pGPD1
0
5
10
15
20
25
Time [min]

Figure 7

Supplementary Materials to:

Single-particle view of stress-promoters induction dynamics: an interplay between MAPK signaling, chromatin and transcription factors

Victoria Wosika and Serge Pelet
Department of Fundamental Microbiology,
University of Lausanne

Supplementary Tables

Supplementary Table 1. Plasmids used in this study

Plasmid name	Description	Backbone
pSP264	pSTL1 24xPP7sl	pHIS [a]
pSP266	pCTT1 24xPP7sl	pHIS [a]
pVW200	pGPD1 24xPP7sl	pHIS [a]
pVW293	pHSP12 24xPP7sl	pHIS [a]
pVW294	pGRE2 24xPP7sl	pHIS [a]
pVW295	pALD3 24xPP7sl	pHIS [a]
pSP568	pSTL1 24xMS2sl	pHIS [a]
pVW284	pADH1 PP7ΔFG-GFPenvy tCYC1	pSIV URA3 [b]
pVW300	pTEF PP7ΔFG-GFPenvy tCYC1	pSIV URA3 [b]
pVW296	pADH1 PP7ΔFG-mCherry tCYC1	pSIV URA3 [b]
pSP561	pADH1 MS2-GFPenvy tCYC1	pSIV URA3 [b]
pSP571	pGPD Cas9 / sgRNA ($STL1_{62}$)	pRS423II [c]
pSP569	pSTL1 24xPP7sl $STL1_{100-600}$	pHIS [d]
pSP566	mCherry-CaaX	pGT TRP1 [e]

[a] Modified from Larson *et al.*[1] Integrates in the *GLT1* locus.
[b] pSIV vector from Wosika *et al.*[2].
[c] Modified from Laughery *et al.*[3]
[d] Used as repair DNA for CRISPR transformation.
[e] Modified from Wosika *et al.*[2].

Supplementary Table 2. List of yeast strains

All strains were constructed in the W303 background (ySP2) *MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*. [4]

Strain name	Relevant Genotype	Ancestor strain
ySP269	HTA2-mCherry:URA3	ySP2
ySP329	HTA2-mCherry:URA3 Hog1-GFP:HIS3	ySP269
yED215	HTA2-mCherry:URA3	ySP2
yVW401	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3	yED215
yVW403	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pSTL1 24xPP7sl GLT1 tGLT1::GLT1:HIS3	yVW401
yVW428	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pCTT1 24xPP7sl GLT1 tGLT1::GLT1:HIS3	yVW401
yVW429	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pHSP12 24xPP7sl GLT1 tGLT1::GLT1:HIS3	yVW401
yVW430	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pGRE2 24xPP7sl GLT1 tGLT1::GLT1:HIS3	yVW401
yVW431	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pALD3 24xPP7sl GLT1 tGLT1::GLT1:HIS3	yVW401
yVW432	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pGPD1 24xPP7sl GLT1 tGLT1::GLT1:HIS3	yVW401
yVW409	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pSTL1 24xPP7sl GLT1 tGLT1::GLT1:HIS3 GCN5::NAT	yVW403
yVW416	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pSTL1 24xPP7sl GLT1 tGLT1::GLT1:HIS3 HTZ1::NAT	yVW403
yVW405	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pSTL1 24xPP7sl GLT1 tGLT1::GLT1:HIS3 HOT1::NAT	yVW403
yVW407	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pSTL1 24xPP7sl GLT1 tGLT1::GLT1:HIS3 SKO1::NAT	yVW403
ySP915	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pSTL1 24xPP7sl GLT1 tGLT1::GLT1:HIS3 MSN4::KAN, MSN4::NAT	yVW403
yVW471	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pGPD1 24xPP7sl GLT1 tGLT1::GLT1:HIS3 HOT1::NAT	yVW432
yVW472	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pGPD1 24xPP7sl GLT1 tGLT1::GLT1:HIS3 SKO1::NAT	yVW432
ySP918	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pGPD1 24xPP7sl GLT1 tGLT1::GLT1:HIS3 MSN4::KAN, MSN4::NAT	yVW432

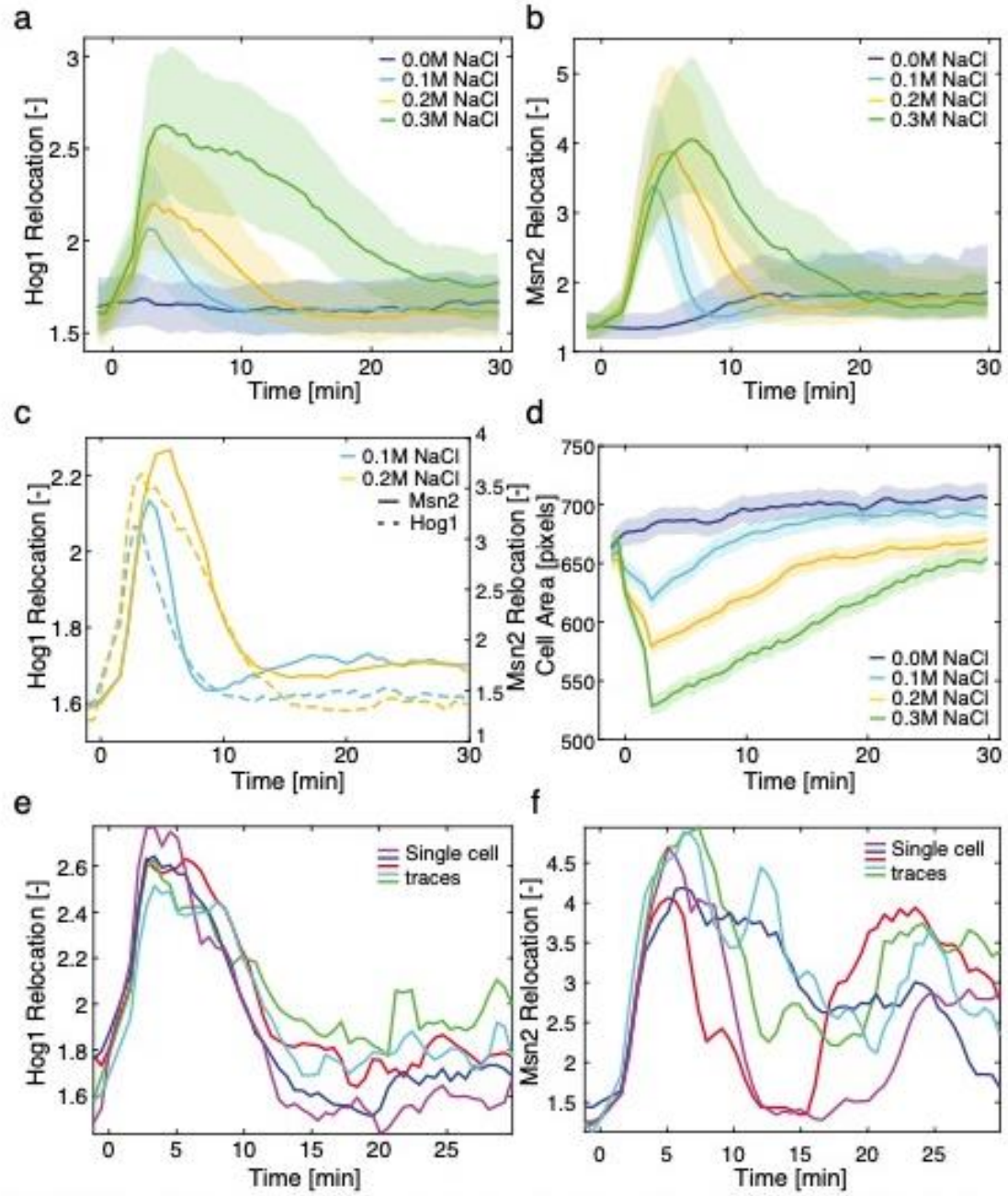
yVW476	HTA2-mCherry:URA3 pSIVu pTEF PP7ΔFG-GFPenvy tCYC1::URA3 pGPD1 24xPP7sl GLT1 tGLT1::GLT1:HIS3	yVW454
yVW477	HTA2-mCherry:URA3 pSIVu pTEF PP7ΔFG-GFPenvy tCYC1::URA3 pHSP12 24xPP7sl GLT1 tGLT1::GLT1:HIS3	yVW454
yVW474	HTA2-tdiRFP:ADE pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pSTL1 24xPP7sl GLT1 tGLT1::GLT1:HIS3 HOG1-mCherry: LEU2	
ySP919	HTA2-tdiRFP:ADE pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pSTL1 24xPP7sl GLT1 tGLT1::GLT1:HIS3 HOG1-mCherry-CaaX: TRP1	
ySP921	HTA2-tdiRFP:NAT pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pGPD1 24xPP7sl GLT1 tGLT1::GLT1:HIS3 HOG1-mCherry: LEU2	
ySP922	HTA2-tdiRFP:NAT pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pGPD1 24xPP7sl GLT1 tGLT1::GLT1:HIS3 HOG1-mCherry-CaaX: TRP1	
ySP929	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pSTL1 24xPP7sl :STL1 (clone 8)	yVW401
ySP930	HTA2-mCherry:URA3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 pSTL1 24xPP7sl :STL1 (clone 10)	yVW401
ySP927	MAT**a** / MATα Hta2-tdiRFP:NAT / Hta2-tdiRFP :TRP pSTL1 24xPP7sl GLT1 tGLT1::GLT1:HIS3 / pSTL1 24xMS2sl GLT1 tGLT1::GLT1:HIS3 pSIVu pADH1 PP7ΔFG-GFPenvy tCYC1::URA3 / pSIVu pADH1 MS2-GFPenvy tCYC1::URA3	
ySP884	HTA2-CFP:HIS3 HOG1-mCitrine:LEU2 MSN2-mCherry:URA3	
ySP763	HTA2-CFP:HIS3 HOG1-mCitrine:LEU2 pSIVu pSTL1-dPSTR-mCherry::URA3	
ySP764	HTA2-CFP:HIS3 HOG1-mCitrine:LEU2 pSIVu pHSP12-dPSTR-mCherry::URA3	
yVW418	HTA2-CFP:HIS3 HOG1-mCitrine:LEU2 pSIVu pALD3-dPSTR-mCherry::URA3	
ySP766	HTA2-CFP:HIS3 HOG1-mCitrine:LEU2 pSIVu pCTT1-dPSTR-mCherry::URA3	

Supplementary Table 3. Summary of source data, strains and cell numbers

	Strain	Stress	Replicates	Date yymmdd	Exp Num	Nb Cells
Figure 1 d	yVW403	0.0 M NaCl	rep1	181120	3739	313
	yVW403	0.1M NaCl	rep1	181120	3745	404
	yVW403	0.2M NaCl	rep2	181127	3762	229
	yVW403	0.3M NaCl	rep3	190124	3849	201
Figure 2 a	yVW431	0.2M NaCl	rep1	181127	3752	171
	yVW428	0.2M NaCl	rep3	190118	3827	140
	yVW403	0.2M NaCl	rep2	181127	3762	229
	yVW430	0.2M NaCl	rep1	181127	3754	289
	yVW429	0.2M NaCl	rep4	190329	3951	243
	yVW432	0.2M NaCl	rep2	181127	3758	335
Figure 2 f	yVW403	0.2M NaCl	rep2	181127	3762	229
	yVW409	0.2M NaCl	rep3	190124	3852	148
	yVW416	0.2M NaCl	rep1	181207	3796	175
Figure 2 g	yVW403	0.2M NaCl Glucose	B	190625	4076	248
	yVW403	0.2M NaCl Raffinose	B	190625	4076	275
Figure 3 a middle	ySP329	0.1M NaCl	C	190412	3989	327
	ySP329	0.2M NaCl	C	190412	3989	341
	ySP329	0.3M NaCl	C	190412	3989	311
Figure 3 a bottom	Data from Figure 1d					
Figure 3 b	Data from Figure 2a					
Figure 3 b	Data from Figure 1d					
Figure 4a	Data from Figure 2a					
Figure 4 b	yVW430	0.0 M NaCl	rep1	181127	3756	248
	yVW429	0.0 M NaCl	rep1	190222	3873	216
	yVW432	0.0 M NaCl	rep1	181120	3743	214
Figure 4 c	yVW405	0.2M NaCl	rep3	181129	3772	349
	yVW407	0.2M NaCl	rep1	181120	3742	529
	yVW471	0.2M NaCl	rep1	190322	3920	293
	yVW472	0.2M NaCl	rep3	190405	3968	297
Figure 5	Data from Figure 2a					
Figure 6 a	Data from Figure 1d					
Figure 6 c	yVW474	Pulse	D	190619	4056	168
	yVW474	Step	D	190619	4054	159
	yVW474	Ramp	D	190619	4052	119
Figure 7 a	Data from Figure 2a					
Sup Fig 1	ySP884	0.0 M NaCl	B	190625	4074	266
	ySP884	0.1M NaCl	B	190625	4074	370
	ySP884	0.2M NaCl	B	190625	4074	396
	ySP884	0.3M NaCl	B	190625	4074	291

Sup Fig 4	yVW403	0.2M NaCl	B	200117	4205	436
	ySP929	0.2M NaCl	A	200117	4200	408
	ySP930	0.2M NaCl	B	200117	4203	483
Sup Fig 5	yVW474	0.0 M NaCl	rep2	190503	4009	65
	yVW474	0.1M NaCl	rep2	190503	4010	143
	yVW474	0.2M NaCl	rep2	190503	4009	195
	yVW474	0.3M NaCl	rep2	190503	4010	306
Sup Fig 6	ySP927	0.2M NaCl	A	191219	4191	89
	ySP927	0.2M NaCl	B	191219	4192	91
	ySP927	0.2M NaCl	C	191219	4193	77
Sup Fig 8	yVW431	0.0 M NaCl	rep2	181129	3769	183
	yVW428	0.0 M NaCl	rep2	181213	3820	120
	yVW403	0.0 M NaCl	rep1	181120	3739	313
	yVW430	0.0 M NaCl	rep1	181127	3756	248
	yVW429	0.0 M NaCl	rep1	190222	3873	216
	yVW432	0.0 M NaCl	rep1	181120	3743	214
Sup Fig 9 a	Data from Figure 3a					
Sup Fig 9 b	Data from Figure 1d					
Sup Fig 9 c	Data from Figure 2a					
Sup Fig 10 a	Data from Figure 2a					
Sup Fig 10 b	Data from Figure 4b					
Sup Fig 12	yVW403	0.2M NaCl	rep1	191206	4164	269
	ySP915	0.2M NaCl	B	191206	4166	282
	yVW432	0.2M NaCl	rep1	191206	4168	293
	ySP918	0.2M NaCl	A	191206	4169	310
Sup Fig 13	yVW474	0.2M NaCl	C	191217	4187	366
	ySP919	0.2M NaCl	E	191217	4188	257
	ySP899	0.2M NaCl	B	191217	4189	319
	ySP922	0.2M NaCl	D	191217	4190	367
Sup Fig 14	Data from Figure 2a					

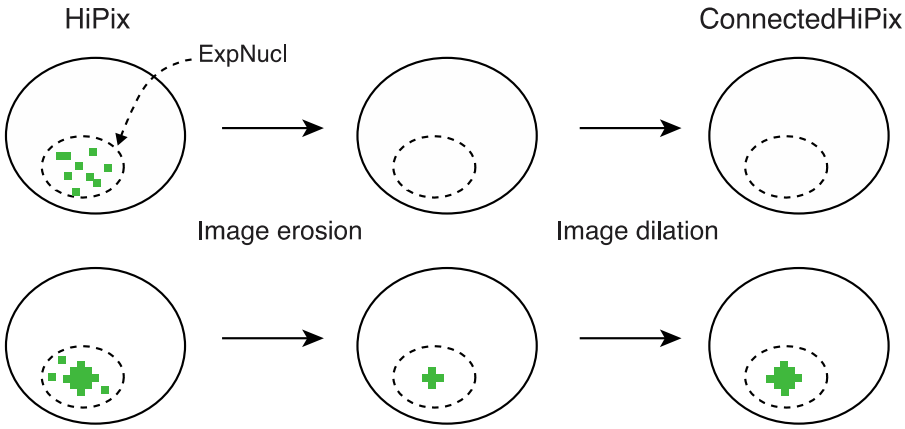
Supplementary Figures



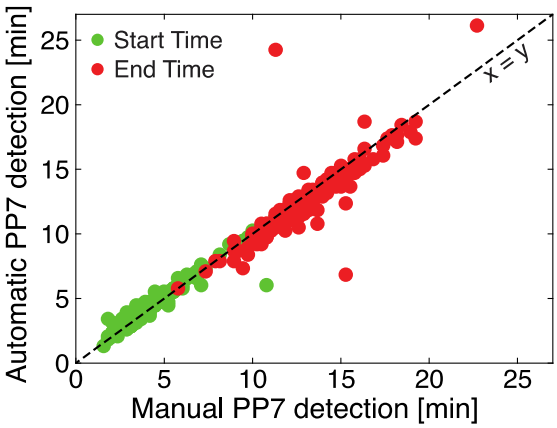
Supplementary Figure 1. Comparison between the nuclear relocation of Msn2 and Hog1.

a. - c. Strains bearing a Hta2-CFP, Hog1-mCitrine and Msn2-mCherry were stressed with various concentrations of NaCl. The nuclear relocation was quantified by the ratio in nuclear over cytoplasmic fluorescence. The median (solid line) and 25th -75th percentiles (shaded areas) are plotted for Hog1 in the yellow channel (a), Msn2 in the red channel (b) and directly compared on the same graph (c) for at least 260 cells. **d.** Change in cell size following hyper-osmotic stresses. **e. - f.** The nuclear relocation traces from the same single cells for Hog1 (e) and Msn2 (f) are plotted. These cells were selected in the population because they display a strong re-entry of Msn2 in the nucleus following the first pulse of activity. This second phase in the response is absent from the Hog1 dynamics.

a



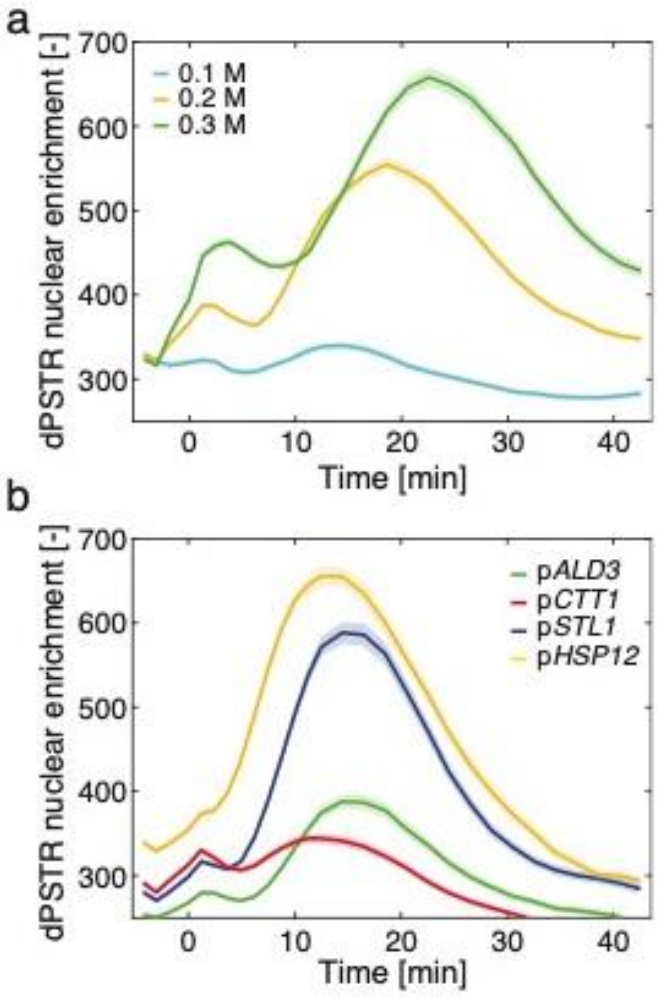
b



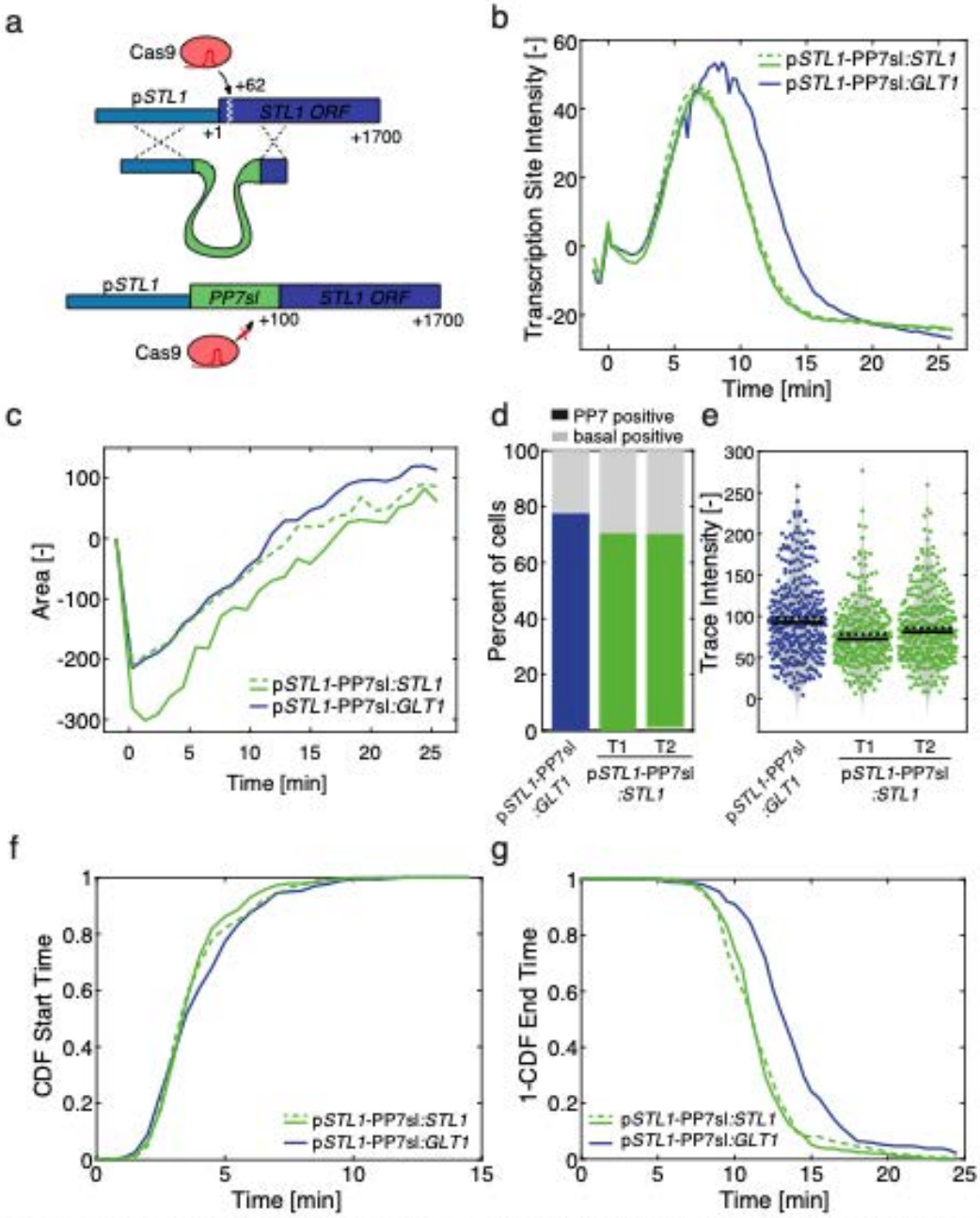
c

		Manual Curation	
		PP7 Positive	PP7 Negative
Automated measurement	PP7 Positive	193	1
	PP7 Negative	1	43

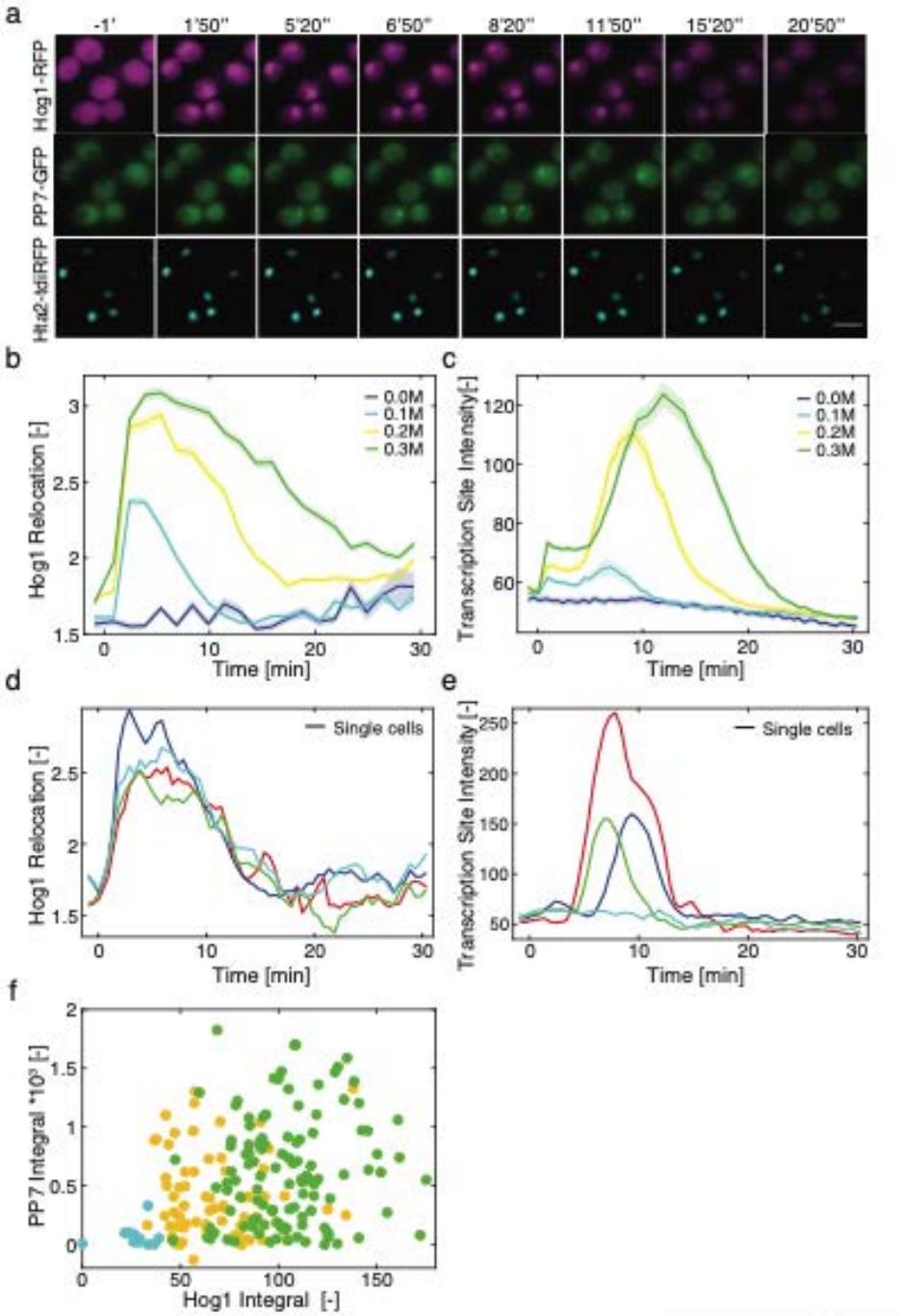
Supplementary Figure 2. Quantification of the PP7 traces with the ConnectedHiPix feature.
a. Scheme describing the process used to generate the ConnectedHiPix feature from the 20 highest intensity pixels (HiPix) in the ExpNucl object (Nucleus expanded by 5 pixels). **b.** Comparison between a manual curation of Start Times (green) and End Times (red) by visual inspection of the cells and automated quantification by the ConnectedHiPix measurement. **c.** Table displaying the number of PP7 positive and negative cells obtained by manual versus automated quantification.



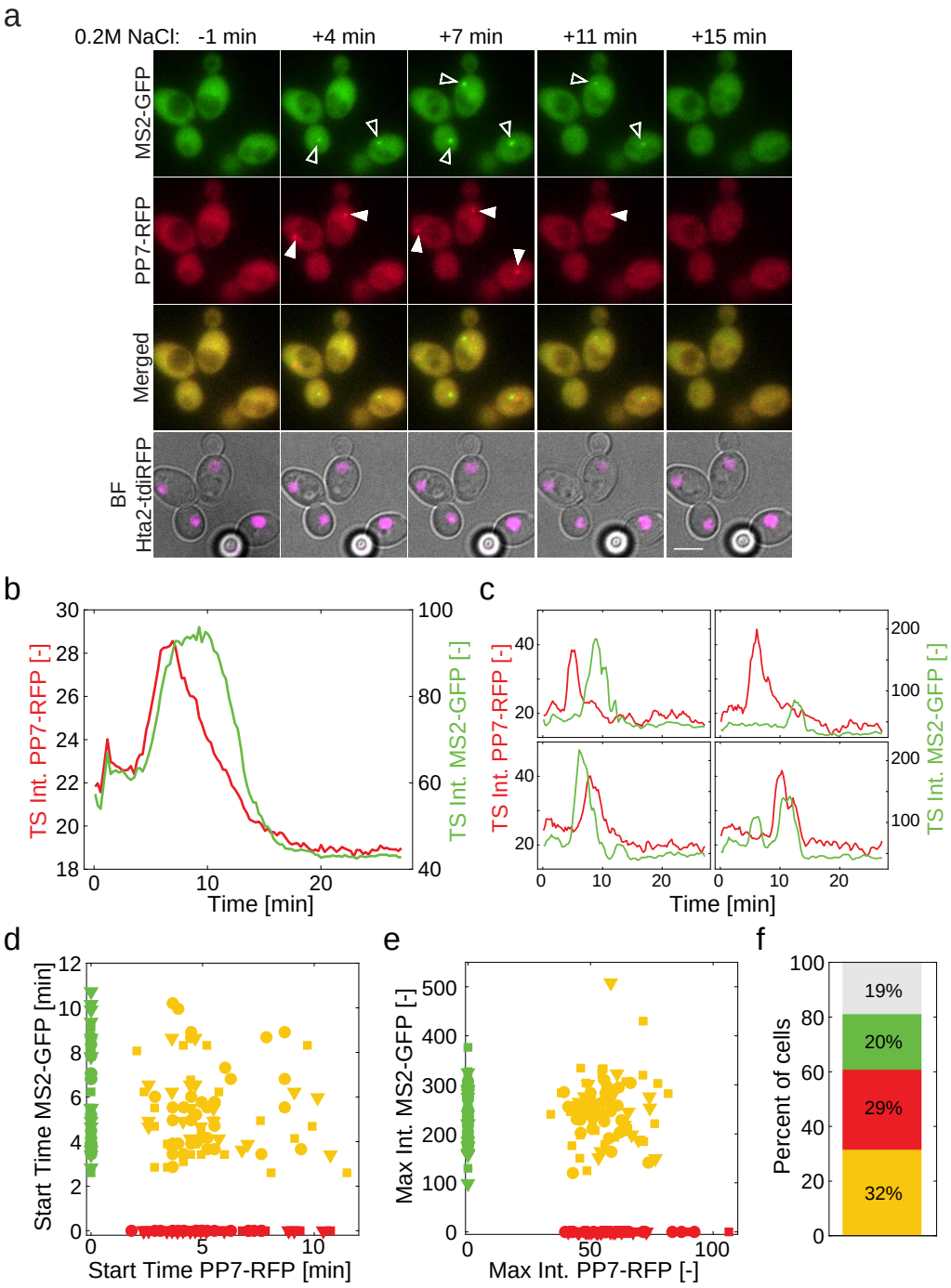
Supplementary Figure 3. Dynamic protein synthesis translocation reporter (dPSTR) induction upon osmotic stress.
The dPSTR reporter allows to by-pass the slow maturation time of protein expression reporters. A constitutively expressed FP is functionalized by a leucine zipper. The compatible zipper is under the control of the inducible promoter of interest and coupled to two strong Nuclear Localization Signals (NLS) motifs. When the pair of leucine zippers interact, the enrichment of the FP in the nucleus allows to quantify the level of induction of the promoter[5]. **a.** p*STL1*-dPSTR^R nuclear enrichment (nuclear fluorescence minus cytoplasmic fluorescence) for three different stress levels. **b.** dPSTR^R nuclear enrichment for 4 different promoters following a 0.2M NaCl stress. In all graphs, the solid line represents the mean difference and the shaded area represents the s.e.m.



Supplementary Figure 4. Monitoring transcription from the endogenous *STL1* locus
a. Scheme describing the integration of the 24xPP7sl at the *STL1* locus by CRISPR-Cas9. The sgRNA recognizes the PAM motif 62bp upstream of the start codon. The DNA break is repaired by homologous recombination from a DNA fragment that contains homology with the *STL1* promoter and 500 bp from the *STL1* ORF starting at position 100. **b.** Transcription site intensity arising from the p*STL1* promoter upon 0.2M NaCl stress. **c.** Cell size adaptation dynamics following hyper-osmotic for the experiment presented in b. **d.** Percent of PP7 positive cells. **e.** Maximum intensity of single-cell traces. **f.** and **g.** Dynamics of transcription onset and shut off represented with the CDF of Start Times (f) and 1-CDF of End Times (g). In all these graphs, the response arising from the *GLT1* locus (blue) is compared to the one monitored at the native *STL1* locus (green). Two different transformants were measured in order to verify that undesired Cas9 activity has not disrupted the HOG response.

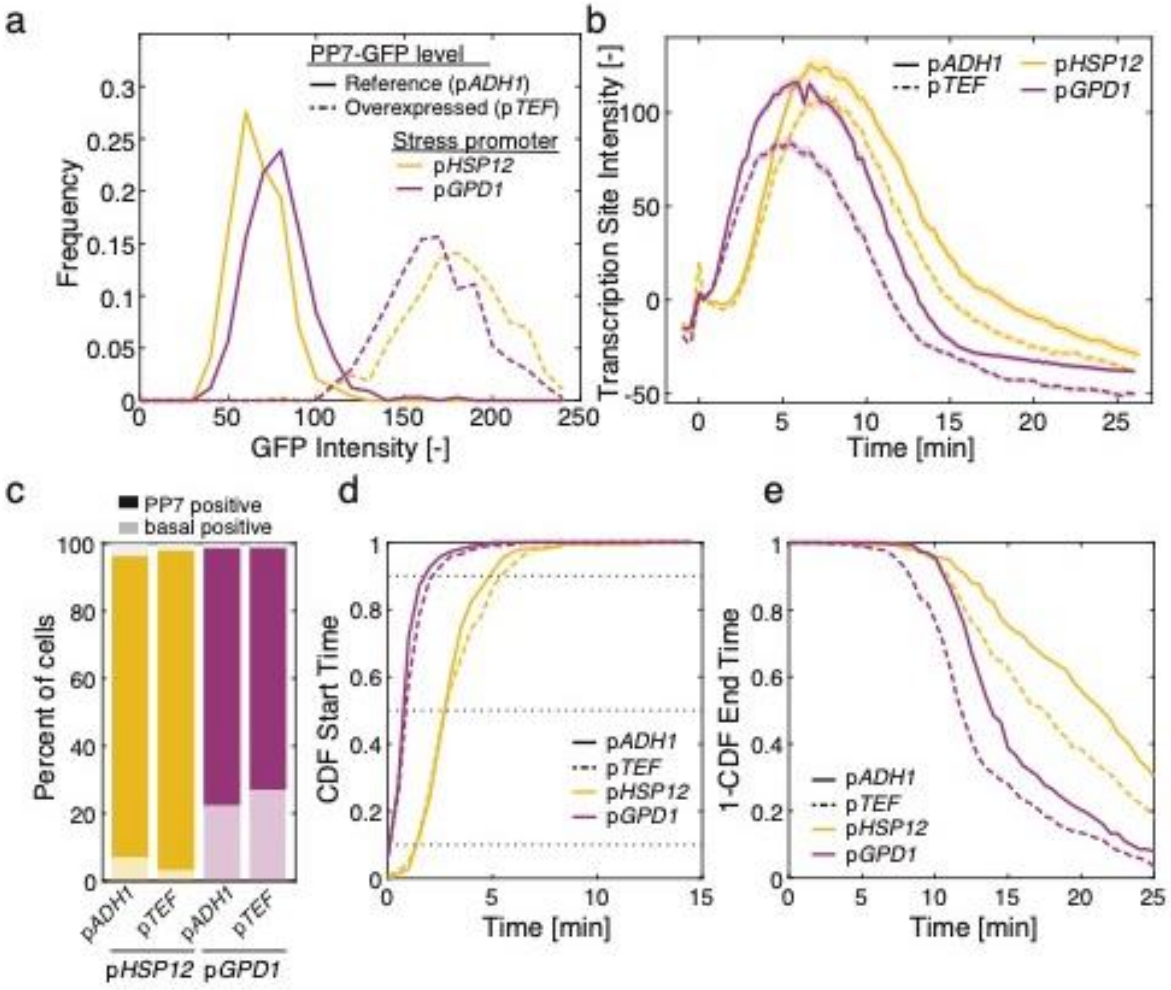


Supplementary Figure 5. Correlating Hog1 activity and downstream transcription in the same cell.
a. Thumbnails images of the strain combining a Hta2-tdiRFP nuclear marker, Hog1-mCherry and the p*STL1*-PP7 reporter. Cells were stressed with 0.2M NaCl at time 0. Scale bar is 5 μm **b.** - **c.** Mean dynamics of Hog1 nuclear enrichment (b) and PP7 transcription site fluorescence intensity (c) following different osmotic stresses. More than 140 cells are quantified for the inducing conditions and only 65 in the SD-full experiment. **d.** - **e.** Examples of single-cell traces that display similar Hog1 relocation dynamics (d) and different transcriptional responses as quantified by the p*STL1*-PP7 transcription site intensity (e). **f.** Correlation between Hog1 relocation and the PP7 output measured by their integrals. Each dot corresponds to a single-cell measurement.

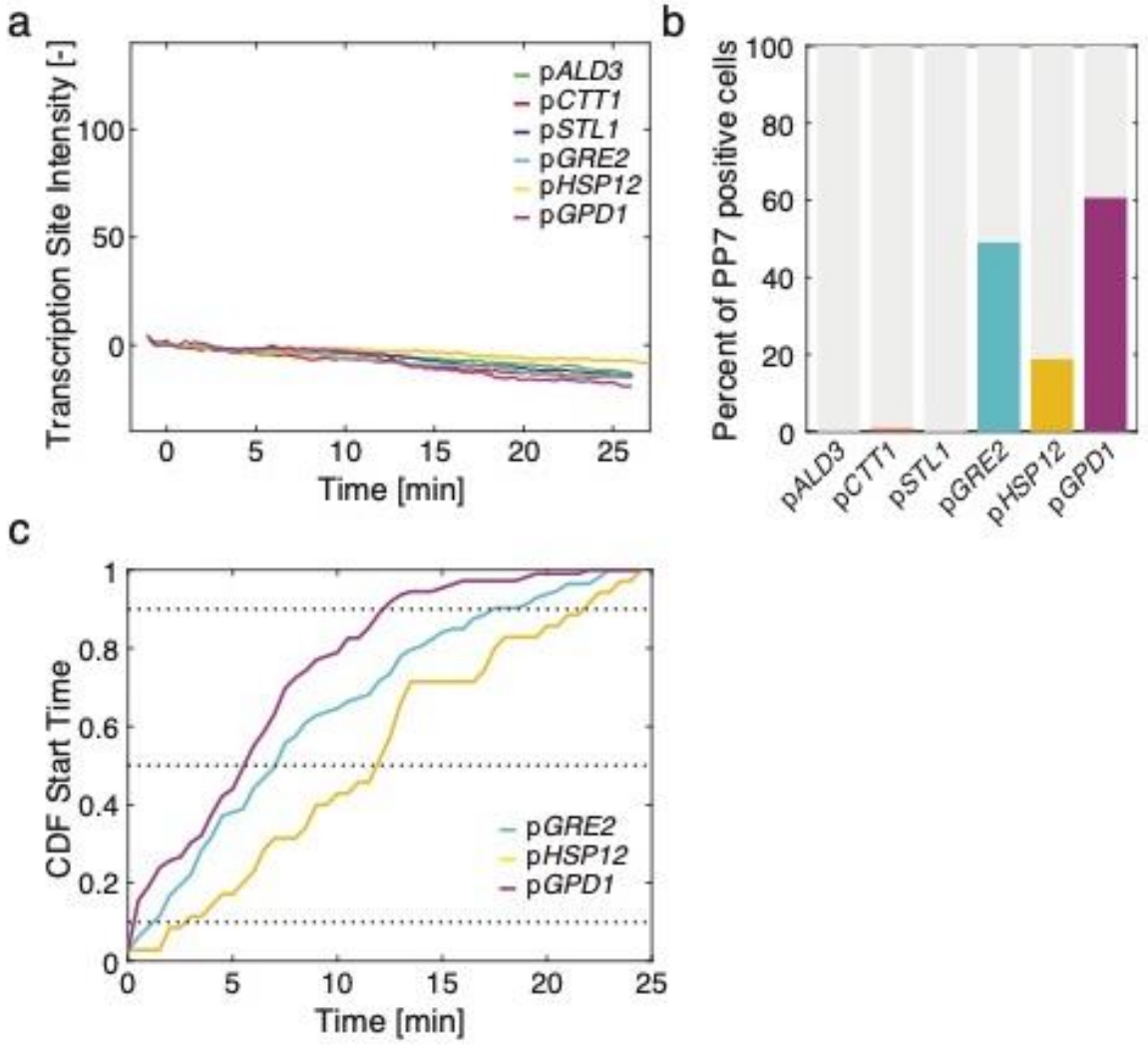


Supplementary Figure 6. Monitoring p*STL1*-induced transcription from two identical loci in diploids.
a. Thumbnails of diploid cells bearing the MS2-GFP and PP7-mCherry reporter systems monitoring the induction of two p*STL1* in the same cell following 0.2M NaCl stress. Open arrowheads (MS2sl) and closed arrowheads (PP7sl) highlight the stochastic activation of the transcription within a cell. Scale bar 5µm. **b.** Transcription site intensity from the p*STL1* promoter monitored with the MS2 (green) or the PP7 (red) systems. The low signal provided by the PP7-mCherry assay and bleaching of this FP can explain the discrepancy between the two reporter systems. **c.** Examples of single cell traces where the activation of

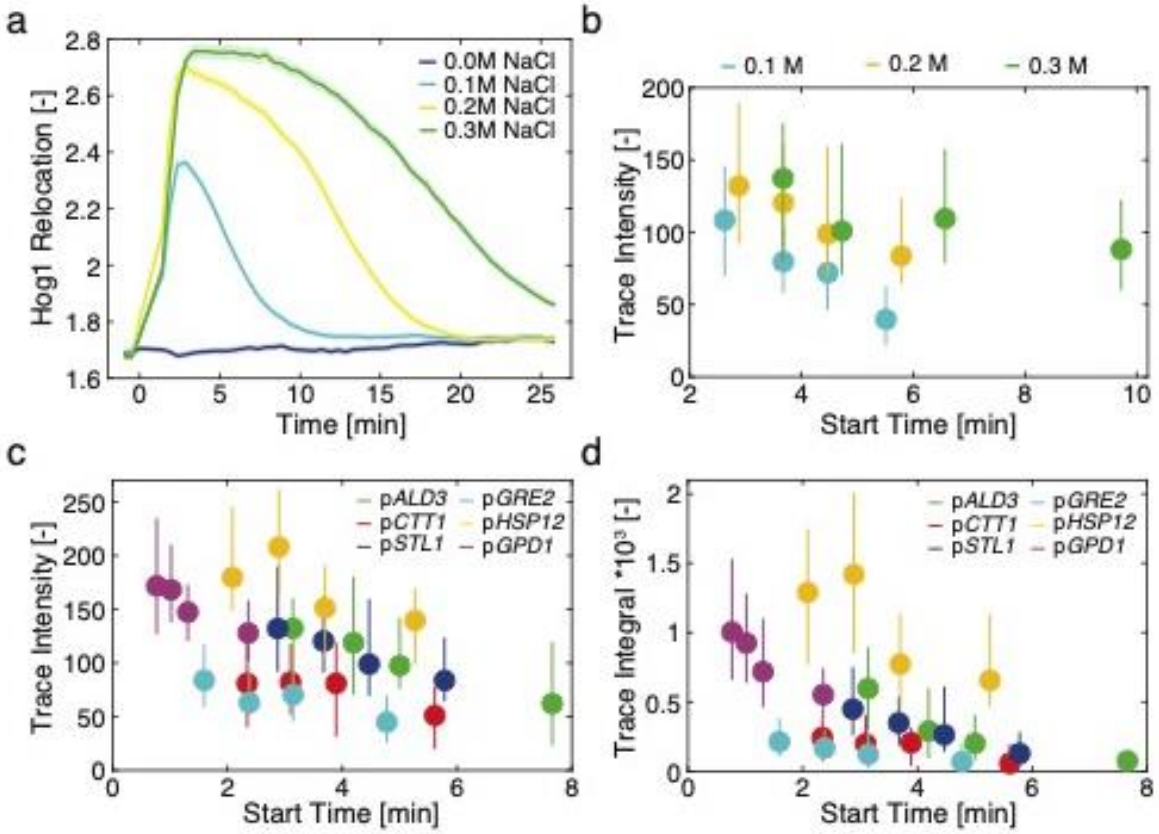
both *STL1* promoters was detected. **d.** and **e.** Scatter plots representing the Start Times (d) and the Maximum Intensity (e) of the PP7 and MS2 assays. The data from three different experiments (rounds, squares and triangles) are combined. Cells where only the MS2 system activation was detected are plotted in green, while cells with only PP7 TS are in red. Cells where both systems where detected are in yellow. **f.** Mean percentage of cells over the three experiments where both promoters (yellow) or only one promoter (green / MS2 and red / PP7) or no transcription (gray) was detected.



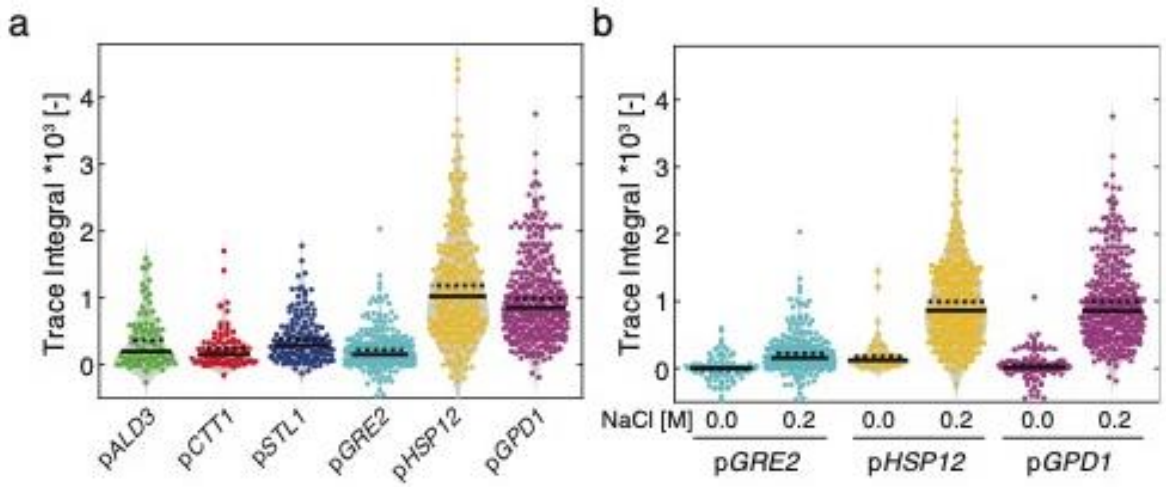
Supplementary Figure 7. Testing the effect of the overexpression of PP7-GFP on the transcription site measurements.
a. Comparison of the initial GFP fluorescence for the PP7-GFPenvy expressed from the p*ADH1* (reference) or p*TEF* promoter (overexpression) which leads to a 3-fold higher fluorescence. **b.** Mean transcription site intensity following a 0.2M NaCl stress. **c.** Percentages of PP7 positive cells. Cells displaying a TS before time point 4 (basal positive) are displayed in a lighter color. **d.** Cumulative distribution of Start Times. **e.** One minus the cumulative distribution function of End Times. The lower expression level of the PP7-GFP by the reference p*ADH1* promoter improves the detection efficiency of the TS.



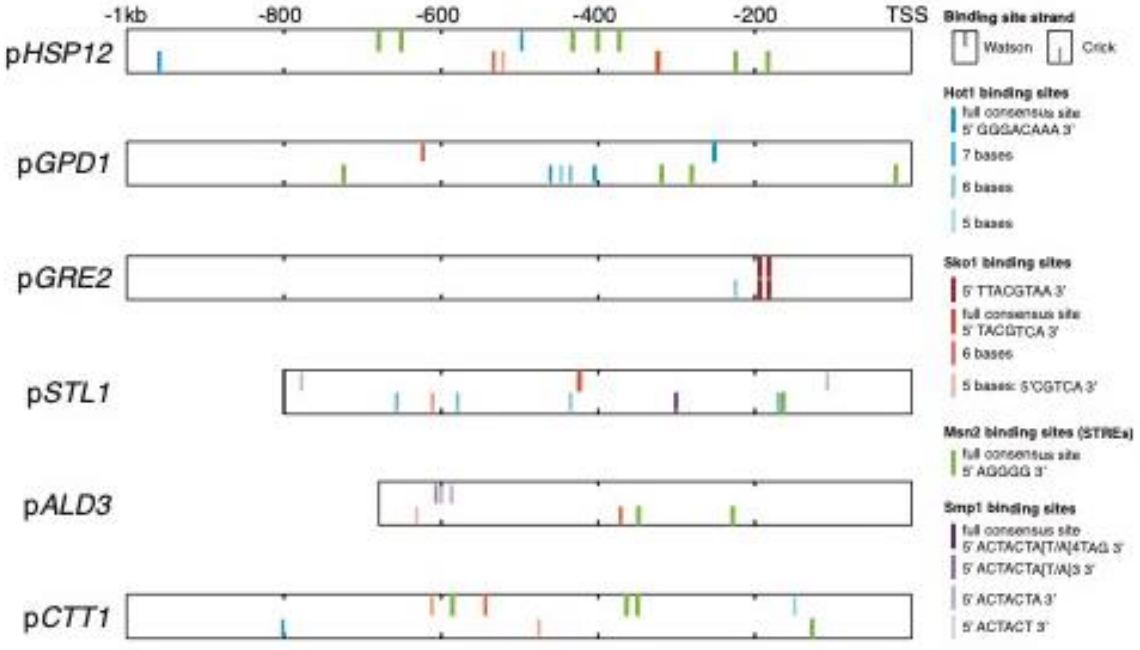
Supplementary Figure 8. HOG promoters with basal expression level.
a. Average transcription site intensity following an SD-full addition. The solid line represents the mean ratio and the shaded area represents the s.e.m. of at least 120 cells.
b. Percentages of PP7 positive cells during the entire SD-full time-lapse experiment in a.
c. Cumulative distribution function of Start Times for the promoters displaying more than 10% of expressing cells in the SD-full time-lapse experiment.



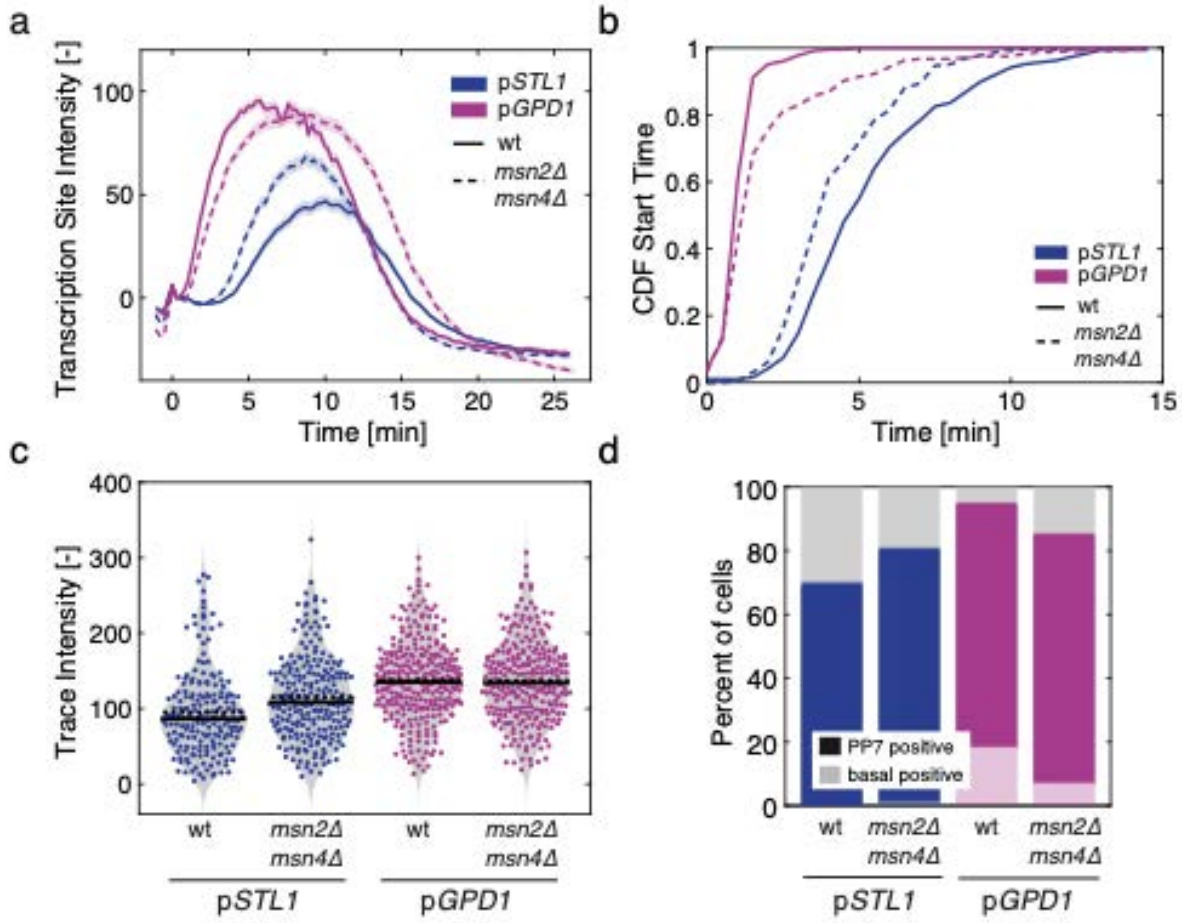
Supplementary Figure 9. Negative correlation between Start Time and transcriptional output for all HOG promoters.
a. Dynamics of Hog1 nuclear enrichment following hyper-osmotic stress. The mean ratio of nuclear over cytoplasmic fluorescence of Hog1-GFP for more than 250 cells is plotted as function of time. The shaded area represents the s.e.m. **b.** The population of p*STL1*-PP7 positive cells is split in four quartiles based on their Start Time. The median (•) and 25th to 75th percentiles (line) of the intensity of the PP7 trace is plotted for each quartile. **c.- d.** Plot of the Start Time versus the Trace intensity (c) or Trace integral (d) for all the PP7 reporter strains following a 0.2M NaCl stress. The population of responding cells is split in four quartiles based on their Start Time. The median (•) and 25th to 75th percentiles (line) of the Trace Intensity (c) or Trace Integral (d) are plotted for each quartile.



Supplementary Figure 10. Trace integral of HOG promoters.
a. Violin plots of the Trace integral of osmostress promoters response after 0.2M NaCl treatment. **b.** Violin plots of the Trace integral of basal level positive osmostress promoters response after SD-full (0.0M) or 0.2M NaCl treatment. For both graphs, each dot represents the data of a single cell, the full line the median of the response and the dashed line the mean of the response.

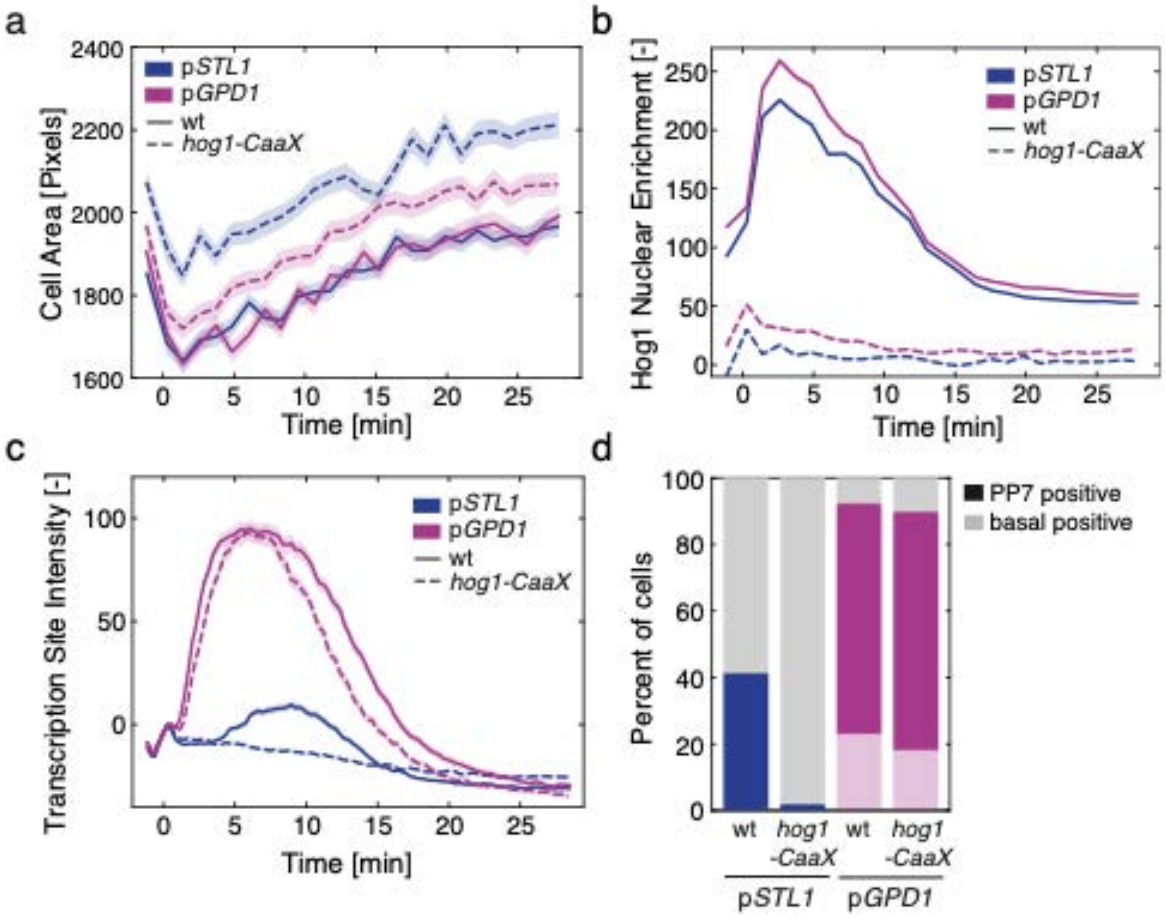


Supplementary Figure 11. Stress promoter architecture.
Consensus binding sites of Hot1[6], Sko1[7], Msn2/4[8], and Smp1[9] and some deviations from these consensus sequences have been mapped on the six promoters used in this study.

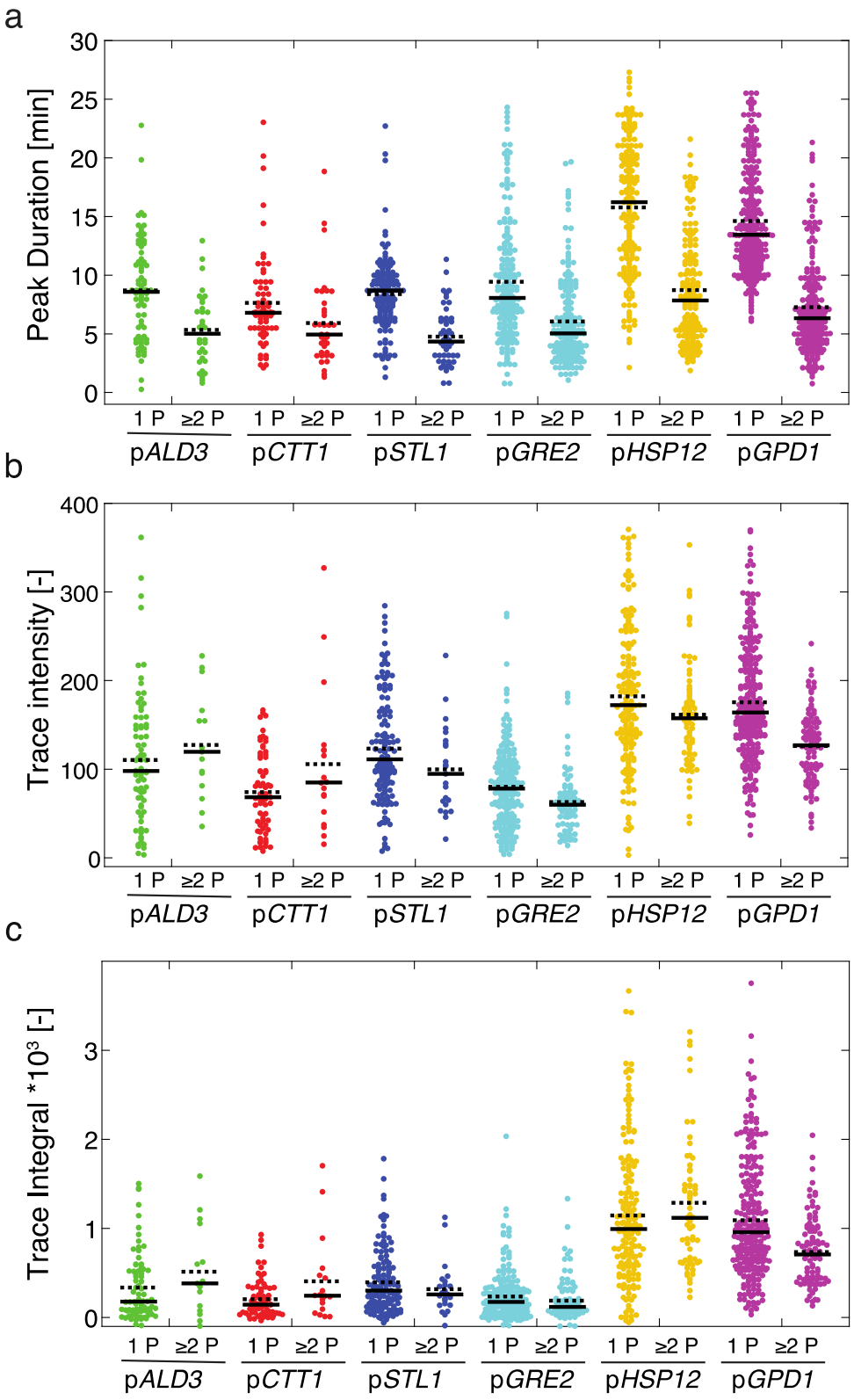


Supplementary Figure 12. Analysis of transcription dynamics in the *msn2Δmsn4Δ* mutant

a. Transcription site intensity of WT (solid line) and *msn2Δmsn4Δ* (dashed line) bearing the p*STL1*-PP7 (blue) or the p*GPD1*-PP7 (magenta) reporters following a 0.2M NaCl stress. **b.** Cumulative distribution of Start Time for the two promoters in the WT and mutant backgrounds for cells that induce transcription after time zero. **c.** Violin plots of the trace intensity (maximum of the TS during the transcription period) after stimulation by 0.2M NaCl. Each dot represents the value calculated from a single cell. The solid line is the median and the dashed line the mean of the population. **d.** Percentages of cells where a PP7 TS site was detected. The light shaded area represents the percentage of PP7 positive cells before the stimulus was added (basal transcription).



Supplementary Figure 13. Impact of Hog1 anchoring at the plasma membrane of p*STL1* and p*GPD1* activation.
a. - c. Dynamics of cell size adaptation (a), Hog1-mCherry nuclear enrichment (b) and Transcription Site intensity (c) in cells expressing either freely diffusing (wt, solid line) or membrane anchored Hog1 via a CaaX motif (dashed line) for the two transcriptional reporters p*STL1*-PP7 (blue) and p*GPD1*-PP7 (magenta) upon 0.2M NaCl stress. **d.** Percentage of cells where a PP7 TS site was detected. The light shaded area represents the percentage of PP7 positive cells before the stimulus was added (basal transcription).



Supplementary Figure 14. Peak analysis of osmostress-promoter in response to 0.2M NaCl.

a. - c. Violin plots of the Peak Duration (a), Trace Intensity (b) and Trace Integral (c) for the different PP7 reporter strains stresses with 0.2 M NaCl. The population of cells was split between cells displaying one peak and cells where two peaks or more were detected. Each dot represents the value calculated for a single peak (a) or single cell (b and c). The solid line is the median and the dashed line the mean of the population.

Supplementary Movies Legends

Supplementary Movie 1: Time lapse movie of the p*STL1*-PP7 reporter strain.
The left image is a maximum intensity projection of a Z-stack in the green channel allowing to image the PP7-GFP and visualizing the presence of transcription sites as bright foci. The central image is the red channel image, representing the fluorescence of the Hta2-mCherry nuclear tag. The right image is a merged image between the green and red channels. The number in the upper right corner indicates the time in minutes. Cells are stressed with 0.2M NaCl at time 0.

Supplementary Movie 2: Time lapse movie of a diploid p*STL1*-PP7 / p*STL1*-MS2 reporter strain.
The leftmost image is a maximum intensity projection of a Z-stack in the green channel allowing to image the MS2-GFP and visualizing the presence of transcription sites as bright foci. The center-left image is a maximum intensity projection of a Z-stack in the red channel allowing to image the PP7-mCherry and visualizing the presence of transcription sites as bright foci. The center-right image is the far-red channel image representing the fluorescence of the Hta2-tdiRFP nuclear tag. The rightmost image is a merged image between the green and red channels (PP7 and MS2) allowing to observe to which extent the induction of two pSTL1 correlate in the same cell. The number in the lower right corner indicates the time in minutes. Cells are stressed with 0.2M NaCl at time 0.

Supplementary Movie 3: Time lapse movie of the p*GPD1*-PP7 reporter strain.
The left image is a maximum intensity projection of a Z-stack in the green channel allowing to image the PP7-GFP and visualizing the presence of transcription sites as bright foci. Note the presence of some transcription site in absence of stimulus in the first frames of the movie. The central image is the red channel image, representing the fluorescence of the Hta2-mCherry nuclear tag. The right image is a merged image between the green and red channels. The number in the upper right corner indicates the time in minutes. Cells are stressed with 0.2M NaCl at time 0.

Supplementary Movie 4: Time lapse movie of the p*STL1*-PP7 in a step experiment.
The leftmost image is a maximum intensity projection of a Z-stack in the green channel allowing to image the PP7-GFP and visualizing the presence of transcription sites as bright foci. The background fluorescence in the image allows to follow NaCl concentration changes in the flow channel. Higher fluorescence is indicative of lower NaCl concentrations. The center-left image is the red channel image and allows to follow the changes in Hog1 nuclear localization. The center-right image is the far-red channel image representing the fluorescence of the Hta2-tdiRFP nuclear tag. The rightmost image is a merged image between the green and red channels (PP7 and Hog1). The number in the lower right corner indicates the time in minutes from the start of the experiment.

Supplementary Movie 5: Time lapse movie of the p*STL1*-PP7 in a pulse experiment.
The leftmost image is a maximum intensity projection of a Z-stack in the green channel allowing to image the PP7-GFP and visualizing the presence of transcription sites as bright foci. The background fluorescence in the image allows to follow NaCl concentration changes in the flow channel. Higher fluorescence is indicative of lower NaCl concentrations. The center-left image is the red channel image and allows to follow the changes in Hog1 nuclear localization. The center-right image is the far-red channel image representing the fluorescence of the Hta2-tdiRFP nuclear tag. The rightmost image is a merged image between the green and red channels (PP7 and Hog1). The number in the upper right corner indicates the time in minutes from the start of the experiment.

Supplementary Movie 6: Time lapse movie of the p*STL1*-PP7 in a ramp experiment.
The leftmost image is a maximum intensity projection of a Z-stack in the green channel allowing to image the PP7-GFP and visualizing the presence of transcription sites as bright foci. The background fluorescence in the image allows to follow NaCl concentration changes in the flow channel. Higher fluorescence is indicative of lower NaCl concentrations. The center-left image is the red channel image and allows to follow the changes in Hog1 nuclear localization. The center-right image is the far-red channel image representing the fluorescence of the Hta2-tdiRFP nuclear tag. The rightmost image is a merged image between the green and red channels (PP7 and Hog1). The number in the lower right corner indicates the time in minutes from the start of the experiment.

Supplementary References

1. Larson, D. R., Zenklusen, D., Wu, B., Chao, J. A. & Singer, R. H. Real-Time Observation of Transcription Initiation and Elongation on an Endogenous Yeast Gene. *Science* **332**, 475–478 (2011).
2. Wosika, V. *et al.* New families of single integration vectors and gene tagging plasmids for genetic manipulations in budding yeast. *Molecular Genetics and Genomics* **291**, 2231–2240 (2016).
3. Laughery, M. F. *et al.* New vectors for simple and streamlined CRISPR-Cas9 genome editing in Saccharomyces cerevisiae. *Yeast* **32**, 711–720 (2015).
4. Ralser, M. *et al.* The Saccharomyces cerevisiae W303-K6001 cross-platform genome sequence: insights into ancestry and physiology of a laboratory mutt. *Open Biol* **2**, 120093 (2012).
5. 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 (2016).
6. Gomar-Alba, M., Alepuz, P. & del Olmo, M. lí. Dissection of the elements of osmotic stress response transcription factor Hot1 involved in the interaction with MAPK Hog1 and in the activation of transcription. *Biochim Biophys Acta* **1829**, 1111–1125 (2013).
7. Proft, M., Gibbons, F. D., Copeland, M., Roth, F. P. & Struhl, K. Genomewide identification of Sko1 target promoters reveals a regulatory network that operates in response to osmotic stress in Saccharomyces cerevisiae. *Eukaryotic Cell* **4**, 1343–1352 (2005).
8. Martínez-Pastor, M. T. *et al.* The Saccharomyces cerevisiae zinc finger proteins Msn2p and Msn4p are required for transcriptional induction through the stress response element (STRE). *EMBO J* **15**, 2227–2235 (1996).
9. Dodou, E. & Treisman, R. The Saccharomyces cerevisiae MADS-box transcription factor Rlm1 is a target for the Mpk1 mitogen-activated protein kinase pathway. *Mol Cell Biol* **17**, 1848–1859 (1997).

Chapter 5: Main results summary

In this thesis, we report on the development and optimization of fluorescent single-cell reporter strains for the monitoring of osmostress gene expression. Acquisition of dynamic data on osmostress gene expression dynamics through an automated cell segmentation and quantification platform (YeastQuant, [178]) enabled us to gathered large datasets, which could be used for further studies.

5.1 Construction of a live mRNA and protein synthesis reporter

In this study, we have developed novel reporter assays, based on the coupling and adaptation of existing ones, for the simultaneous quantification of transcription and translation from a single-promoter in live single-cells. Our initial construct, coupled the PP7 system with the Venus reporter assay, which enables us to have a dynamic readout of transcriptional activity and an accurate but delayed translational readout. Using this initial assay, we could assess the correlation between the mRNA and proteins productions from a single *STL1* promoter and observed a poor correlation between these two readouts. With this allele, we could demonstrate that the presence of the PP7 stem-loops on a transcript diminishes its stability, but can be rescue by the binding from the corresponding coat protein. We then designed our second assay, based on the PP7 and the stable dPSTR assay, enabled us to bypass the maturation time of the Venus and have stable readout for an easy quantification. With this second assay, we still did not observe any correlation between transcripts and proteins from the p*STL1* promoter. Our third and final assay, coupled the PP7 to the unstable dPSTR reporter. With this allele, we obtained dynamics readout that enables us to monitor p*STL1* activation and deactivation in both readouts.

5.2 Single-molecule analysis of osmostress genes transcription

In this study, we have optimized the PP7 system to the monitoring of osmostress gene transcription dynamics, which had not been done before. We have fine-tuned the phage coat protein expression level to enable transcript labelling and quantification of a set of osmostress promoters induced with working salt concentrations ranging from 0 to 0.2M NaCl. We have demonstrated that these strains indeed had no titration effect, by constructing alternative reporter strains expressing three times more coat proteins. Thanks to this optimization, we acquired trustful data on their dynamic of mRNA production. We could uncover the diversity of transcriptional profile upon an identical salt challenge at the single-allele level. We have demonstrated that transcription initiation is not permitted throughout MAPK activity, but only during increasing or stable Hog1 activity. Similarly, we have highlighted a positive correlation between transcription initiation time and the transcriptional output of a cell, illustrating the heterogeneity in MAPK activity through the full activity window. In addition to signaling, we have shown that chromatin compaction at the locus negatively correlates with the transcriptional response time and the number of expressing cells. We have demonstrated that this requirement for chromatin remodeling will dictate the dependency toward transcription factor, with low compaction promoter exhibiting only small loss-of-function. Finally, we have illustrated the promoter-specific component dictating the transcriptional termination, which cannot solely be explained by HOG deactivation.

Chapter 6: General discussion and perspectives

6.1 Discussion

6.1.1 On the use of the PP7 system for mRNA labelling

Start and end times

Detection of a PP7 focus occurs as soon as a certain fluorescence above background is reached at the transcription site, when enough PP7 stem-loops are transcribed and thus bound by fluorescently labelled PP7 proteins. Similarly, signal disappearance occurs supposedly when polymerases detach from the DNA locus and/or the transcript is exported from the nucleus to the cytoplasm. Thus, transcriptional activation is not detected immediately and transcription stopped before signal disappearance. One could thus argue that the PP7 system does not enable to truly measure promoter activity. However, unlike most of the available reporter assay which are protein-based, the PP7 system is RNA-based and thus is not delayed, nor buffered by export or translation.

Transcription site fluorescence

We discussed previously the importance of the level of synthetized pool of phage coat proteins to ensure labelling of all transcript and avoid titration of the system. Fine tuning of the coat protein levels is necessary for the detection of lowly expressed promoters and to avoid depletion from highly expressed ones. In addition, this protein pool is composed of fully mature and thus fluorescent PCP-FP fusion proteins and some that are folded but not yet

fluorescent. Yet these latter ones are capable of binding transcripts and could thus potentially lower the transcript fluorescent due to a mixed of fluorescent and non-fluorescent bound PP7-FP. Highly induced promoters should be the more sensitive to this effect than less induced or less expressed transcripts.

mRNA localization

Because they give both spatial and temporal readouts, phage coat protein-based reporter assay have been used for mRNA localization and nuclear export studies. Although the presence of forty-eight bound PP7 protein and their fluorescent protein on a transcript may impact the quality and kinetic of nuclear export or cytoplasmic transport, comparison between strains identically tagged is feasible. Thus, the PP7 system can be used to study mRNA export but could also lead to transcription site residence time artifact in studies of transcription dynamics in mutant backgrounds. As an example, the nuclear envelop protein Nup60p, has been shown to be implicated in HOG gene regulation [195]. Gene deletion of *NUP60* leads to an ever-lasting PP7 nuclear signal, which does not represent continuous active transcription but nuclear envelop staling of mRNA export (data not shown).

Finally, because the presence of the stem-loops on a transcript were shown to alter its degradation and lead to its accumulation into Processing bodies (P-bodies) upon glucose starvation, utilization of phage coat proteins during harsh cellular conditions may not reflect the true transcript localization [163], although some improved reporter system were developed to remove this artifacts [177].

6.1.2 On the use of gene deletions to study regulation

In chapter 4, we report on transcription factor and chromatin implication on osmostress genes transcription. To do so, we performed total gene deletions, as traditionally performed in yeasts. However, proteins of chromatin complexes are often part of other complexes, like the Gcn5 protein, which is a catalytic subunit of both the ADA and SAGA histone acetyltransferase complexes and a subunit of RSC chromatin-remodeling complex [196, 197]. Additionally, chromatin remodeling complexes are not only regulating osmostress genes, thus the deletion can have multiple side effects that may contribute to the mutant phenotype. This is already visible by the morphological differences of the cells, on their shapes (*arp8*Δ, *asf1*Δ) or size (*set1*Δ, *gcn5*Δ), on their growth rates (decreased in all mutants) and on their Hog1 MAPK signaling activity (data not shown). Similarly, deleting a transcription factor will affect all its regulon and possibly affect the signaling and/or adaptation itself, through the transcription of signaling proteins or glycogenic enzymes [73]. However, at low salt stress, we did not observe an effect on the adaptation time from Hot1 or Sko1 deletion, which suggest that it may not play a role for short adaptation times.

Instead of complete knock out of protein from chromatin remodeling and modifying complexes, conditional mutants could be used. In addition to temperature sensitive alleles, a fast loss-of-function can be achieved either by targeting the protein to another sub-cellular location or by inducing its specific degradation [198, 199]. Either of these two following methods would enable us to limit the prominent side effects of chromatin complexes proteins deletion and should be integrated in futures studies, especially when using single-cell level assay, on the contrary to population-averaged measurements where differences may be less detectable.

Anchor-away of nuclear proteins

Ten years ago was developed the Anchor-Away (AA) technique, which enables to deplete quite rapidly a targeted protein from a yeast nucleus [198]. Thanks to the addition of a tag to the protein of interest and to the cytoplasmic protein chosen as anchor, rapid heterodimerization is achieved by addition of Rapamycin, which regulates the interaction, leading to nuclear depletion of the target. AA strains have been recently designed for Gcn5 and INO80, which show a complete depletion from the nucleus within an hour of rapamycin treatment [200, 201]. The clear advantage of this method is that it enables to growth yeast cells with a functional protein before its depletion. However, this assay suffers from three major drawbacks: first, since Rapamycin is used as inducer for the system, strains have to be made Rapamycin insensitive through the point mutation of *TOR1* into *tor1-1* and deletion of Fpr1p [198], which does not allow to work in wild type background. Second, the interaction between the target and anchor is not reversible, which does not enable transient inactivation and thus limits experimental conditions. Third, the depletion is still relatively long compared to the budding yeast cell-cycle.

Inducible degradation of nuclear proteins

A faster method derived from plant cells for the conditional loss-of-function of specific protein is the Auxin Inducible Degron (AID) [199]. Tagging of the targeted protein with IAA and constitutive expression of the Tir1p, leads to the polyubiquitination and proteasomal degradation of the protein upon auxin mediated IAA-Tir1p interaction. The advantage of AID over the AA is the faster depletion, with a thirty minutes requirement, and fewer genetic modifications of the carrier strain. However, recent studies suggest an auxin-independent

depletion of endogenously tagged proteins, which then has to be tested and can be rescue by decreasing the IAA tag size [202].

d-Cas9 mediated sterical hindrance of TF binding sites

Similarly to chromatin remodeling complexes proteins, Transcription Factors (TFs) regulate a set of genes, which can even vary depending on the experimental conditions. In the HOG pathway, five TFs activate more than three hundred genes upon osmotic shock. Deletion of one of them will affects more than the gene of interest and possibly the cell responses. To avoid TFs depletion, regulatory sequence mutations have been extensively used. However, this could affect neighboring binding sites and/or chromatin conformation. Thanks to the recent development of the CRISPR/Cas9 technology, many novel applications have been developed. The Skotheim lab developed an assay for the reversible disruption of specific transcription factor-DNA interactions using the dead Cas9 (dCas9) to hinder TF binding [203].

6.2 Perspectives

6.2.1 Reveal the mechanism behind osmostress gene bursting

In this study, we have constructed PP7 strains reporting on six different osmostress induced promoters. Because all the promoters we tested show a bursty phenotype, independently of their basal level and thus chromatin compaction level, we can question the source of transcriptional bursts. Indeed, from this we would exclude the implication of chromatin remodeling, since the p*GPD1* promoter, which displays the larger number of bursts, also displays the higher basal level (Chapter 4, Figure 2). In addition, promotor proximal pausing has still not been shown in yeasts and seems to appear only in higher eukaryotic cells [114]. Therefore, the two most probable explanations, which may not exclude each other, are the signaling activity and gene looping. Note that, this also questions whether there is a unique mechanism of transcriptional bursting for all osmostress genes? An attempt to dissect the source of pSTL1 bursting through promoter sequence truncation is shown in Annex 2.

Transcription factor activation

Because most osmostress genes are repressed under basal conditions and induced upon osmotic shock, the source of HOG gene transcription dynamics should be “transmitted” or “extrinsic” bursting, as described in the introduction section. Indeed, we observed a tight correlation between Hog1 dynamics and transcription initiation in Chapter 2, whereby the onset of the MAPK activity is dictating the promoter’s transcriptional profile. Since osmostress transcription factors Hot1 and Sko1 are constitutively bound to the chromatin, it is not their DNA binding dynamics per se dictating the transcriptional bursts. However, their activation through Hog1 recruitment may responsible for the transmitted transcriptional bursts. A

possible way to test this hypothesis would be to simultaneously follow single Hog1 molecule attachment to an osmostress promoter and the PP7 signal apparition from this promoter transcription activity in live-single cells. This imaging feat was recently achieved by Donovan and colleagues' study, in which they monitor transcription factor binding dynamics influence on downstream transcriptional activity [113]. A similar experiment could thus be performed with a Hog1-Halo tag [204], however, due to the high number of Hog1 molecules, only partial labelling of the pool should be realized to enable detection.

However, since osmostress genes receive both inputs from the HOG and from the general stress response pathway, it is then likely that Msn1, Msn2 and Msn4 activation also transmits downstream gene transcriptional activation. Therefore, choice of the reporter gene promoter to be monitored should be based on its dependency toward both pathways or in TF deletion background, to simplify the inputs acting on the system. For instance, *HSP12* and *STL1* could be good candidates for Msn2 and Hog1 single-molecule imaging experiment respectively, since they majoritarily depend on one of the two pathways and do not display basal expression, which would otherwise increase the data analysis complexity.

Gene looping maintains a locus transcriptionally potent

Although transmitted activity to transcription factor may explain the initiation of bursting, the question on how are multiple bursts generated remains. It was shown in yeast that some genes were associated with the nuclear enveloped upon transcription to couple mRNA production and export, phenomenon named "gene gating" [205]. This interaction at the periphery would be mediated by protein from the Nuclear Pore Complex (NPC) and from the transcriptional machinery in the format of NPC-mediated gene looping, as observed for the

galactose inducible gene *GAL1* [108]. Indeed, it was previously shown that osmostress genomic loci were targeted to the nuclear periphery in a Hog1-dependent manner and that gene expression was affected by nuclear envelop protein deletions [148, 195]. Because NPC deletion induces a defect in the mRNA export, the best way to test for this hypothesis would be to perform a Hi-C chromosome interaction assay [206], to capture all physical 3D conformational changes happening upon osmostress with a one kilobase resolution [207].

Cell state influences bursting

A third explanation which may account for either the entire or partial mechanism of gene bursting was recently reported from mammalian cells studies, in which the high transcriptional single-cell variation attributed to gene bursting was abolished by taking into consideration the cellular parameters like cell-cycle stage or cell size. The remaining variability could be explained by a Poisson distribution, therefore suggesting that bursting arises from cellular condition (Forman and Wollman, *bioRxiv* 2019) [208]. A similar analysis could be performed by addition a cell-cycle reporter or chemically synchronizing the cells prior salt stress or *in silico* clustering based on cell-size, to assess whether osmostress gene transcription dynamics are partially (or totally but less likely) explained by cellular state. However, HOG signaling is independent of cell-cycle, thus this variable does likely not come into play [209].

6.2.2 d-Cas9 alteration of nucleosomes positioning in vivo

We have shown in Chapter 4 that promoters with basal activity had a faster transcriptional response than promoters without basal level. These results highlighted a stronger requirement for chromatin remodeling activity at the later one. Thus, we have tried to modify the chromatin compaction by performing chromatin remodeling complexes mutants or

growing cells in a different carbon sources to reduce the locus repression. Alternatively, one could have a more specific effect by targeting chromatin remodeling complexes to osmostress genomic loci in an osmotic stress-independent manner. Indeed, it was shown that dCas9 targeting of a catalytic domain of a chromatin remodeling complex either through TF-fusion, tag and epitope or gRNA induced a selective repositioning of target site neighboring nucleosomes *in vivo* (Donovan *et al.* bioRxiv 2018). Thanks to a *GAL* inducible promoter expressing the catalytic subunit, nucleosomes sliding can be induced at will with carbon source shift. With this assay, we could envision to slide nucleosomes around TSS from non-basal expressing promoter to induce a faster transcriptional activation. An alternative idea could be to recruit INO80 or Set1 catalytic domain to osmostress prior stress instead of post-stress to see how competing forces, like nucleosomes repositioning and epigenetic marks would balance between activation and repression of the targeted locus.

6.2.3 Identifying rate-limiting factors for transcriptional activation

In this study, we use the *STL1* promoter to developed our coupled reporter assay and in the study of HOG gene transcription dynamics. However, we did not investigate further the previously demonstrated high extrinsic noise of this promoter [77, 144]. The activation threshold set by chromatin compaction in a locus-specific fashion has been proposed as the main source of gene expression noise in HOG gene expression [77, 78]. Indeed, we and others have shown that osmostress genes are only poorly correlated to single-cell MAPK activity at the single-cell level. In addition, we have shown that two copies of an *STL1* promoter are poorly correlated within a cell population (extrinsic noise) and within a single-cell (intrinsic noise) [77, 144].

In this study, we have performed some initial experiments in diploids PP7 reporter strains to assess the effect of signaling protein numbers on p*STL1* transcriptional activity. To test this hypothesis, we generated diploid budding yeast strains of our p*STL1*-PP7 reporter at the identical genomic location in both parental strains and remove one allele of *HOT1*, *SKO1* or *HOG1* and compared them to the homozygote wild type in an osmotic stress challenge experiment (Figure 20 and 21). As a measure of transcriptional activity, we counted the maximum number of transcription sites per cell and defined sub-population based on these numbers (Figure 20A). Indeed, our current analysis platform does not yet allow us to monitor more than one transcription site at a time. To ensure labelling of all transcripts, both parental strains were expressing a PP7-GFP allele, thus maintaining the ratio between p*STL1*-PP7sl/PP7-GFP equal to one (Figure 20A). In addition, strains were subjected to a 0.1M NaCl stress, to make differences between the strains more visible than at 0.2M NaCl where almost all the cells are expressing in haploid background.

As shown in Figure 20B, almost all the diploid homozygotes cell population transcribed after a 0.1M NaCl induction, with only 11% of non-responding cells. In addition, we observed equal probabilities of having one or two transcription sites and very few cells with three dots or more, which corresponds to actively dividing cells. Interestingly, cells with half less Hog1 molecules lost a high transcriptional power, with the loss of 30% of responding cells, to the expense of the two-dot cell population (Figure 20B). We observed an even more severe phenotype in Hot1 heterozygotes, with a 40% loss of responding cells. This demonstrated the epistatic effect of Hot1 on Hog1 for the *STL1* promoter, whereby *HOT1* deletion haploid strain is transcriptionally dead. Interestingly, Sko1 heterozygotes showed the weaker loss-of-function, which goes in line with the repressor innate function of the transcription factor.

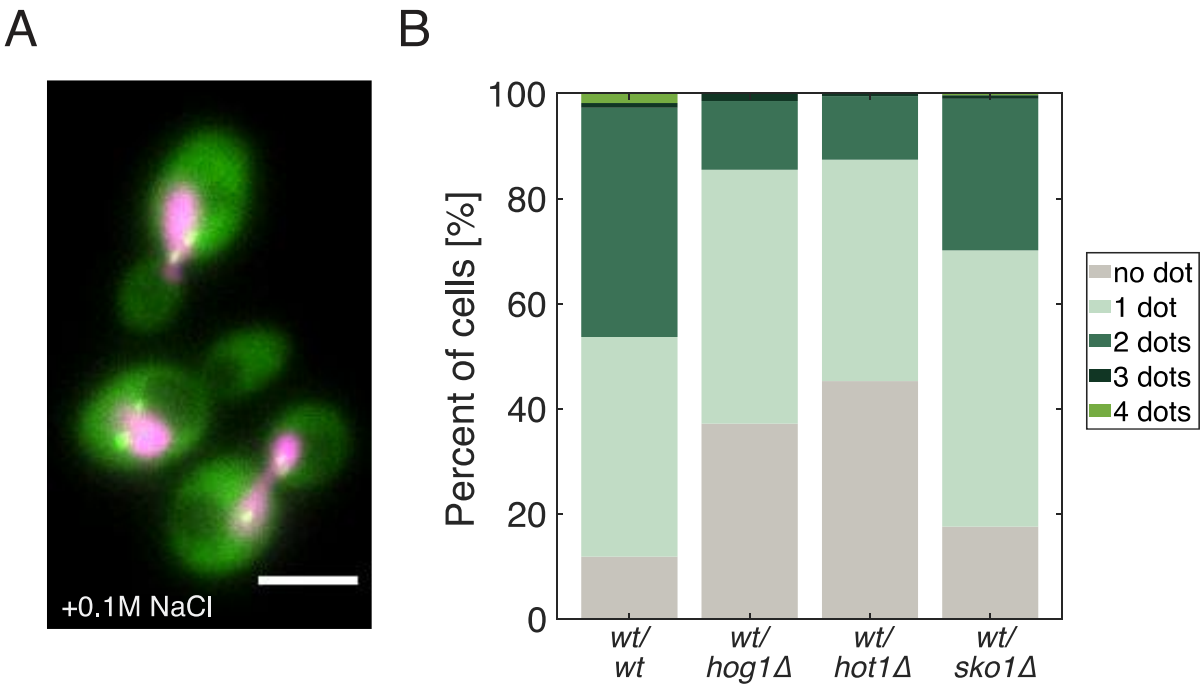


Figure 20: Diploids pSTL1-PP7 reporter strains and their dose-dependency toward transcriptional effectors.
(A) Representative microscopy images of homozygote diploids reporter strains carrying two p*STL1*-PP7 alleles integrated into both *GLT1* loci, two allele expressing the phage coat proteins and the two HTA2 allele tagged with tandem dimer infrared RFP (tdiRFP) fluorophores, here false colored in magenta. (B) Bar plot of the percentages of cells in each sub-population defined based on the maximum number of actively transcribing transcription site per cell. Since the cells carry two copies, there can be up to four simultaneous transcription foci per nuclei of dividing cell.

Hog1 sub-cellular localization is tightly regulated. To assess whether deletion of half of the pool of protein affected the dynamics of the remaining ones, we tagged one Hog1 allele with an mCherry to monitor the dynamics of MAPK activity in the heterozygote and compare it to the homozygote. We subjected the strains to various salt concentrations and tested their ability to relocate the MAPK and trigger p*STL1*-PP7 response, independently of the transcription site number (Figure 21). As shown in Figure 21A, Hog1 nuclear relocation dynamics upon osmotic shock, and thus cells' adaptation, is not affected by the MAPK's protein level, since we observed comparable dynamics. The only noticeable difference being a slight increase at 0.1M NaCl (Figure 21A). However, we indeed observed a loss of transcriptional readout from the heterozygotes at all salt concentrations (Figure 21B). These

results are in line with our previous quantification of the number of alleles being transcriptionally activated (Figure 20B). Note that here we did not make z-stacks in this acquisition, nor used the same imaging conditions, therefore the absolute percentages of responding cells cannot be directly compared to the previous figure.

Interestingly, the number of responding cells was more affected at low salt stress with a more than 50% and 70% loss at 0.1M and 0.2M respectively, than at higher salt stress, where the decrease if of 15% (Figure 21C). We confirmed that these results were not due to Hog1 nuclear relocation at the single-cell level, since more than 95% of the cells overcome the Hog1 nuclear relocation threshold set at 20 RFU of nuclear minus cytoplasmic fluorescences (Figure 21, A and D).

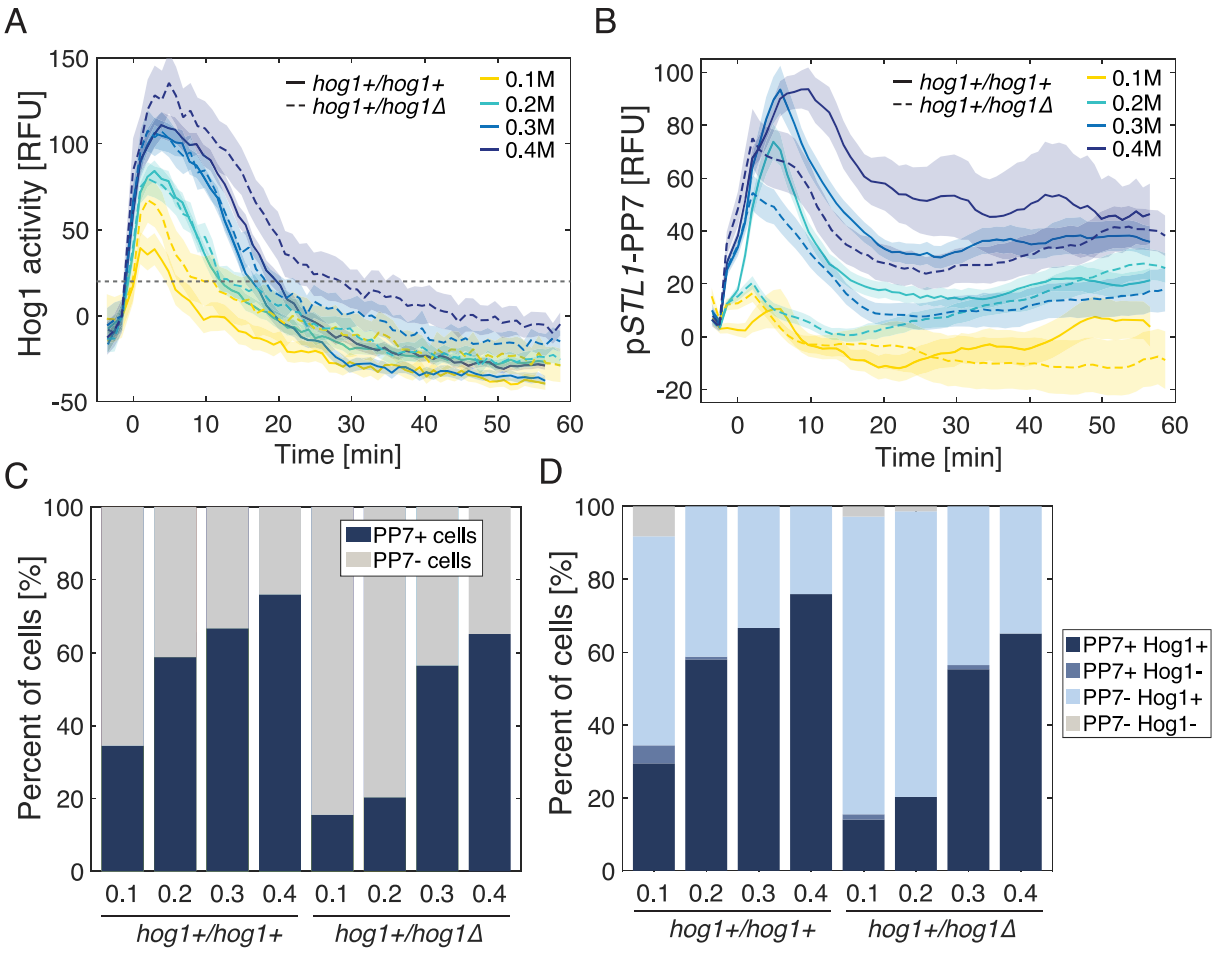


Figure 21: pSTL1 expression is limited by Hog1 molecules number.
(A) Hog1-mCherry nuclear minus cytoplasmic fluorescence subtracted for the basal level in diploids heterozygotes (dashed lines) and homozygotes (full lines) in response to increasing salt concentrations from 0.1M to 0.4M NaCl. (B) Corresponding p*STL1*-PP7 population traces, where the intensity of the transcription site minus the background fluorescence and the basal level are subtracted. No distinction between the transcription site number, only the higher intensity one is quantified at each time point. (C) Bar plot of the percentage of PP7 positive and negative cells for each strain and experiment. Expressing cells were defined thanks to the segmentation of at least one transcription focus. (D) Bar plot of the percentage of each sub-population corresponding positive and negatives for both PP7 and Hog1 nuclear accumulation, as defined by the arbitrary threshold of 20 RFU on the nuclear minus cytoplasmic fluorescence subtracted for the basal level, as plotted in A.

These initial data strongly suggest a system in which the number of TF molecules would be limiting for the recruitment of chromatin remodeling enzymes. Thus, noise from different osmostress genes would be the side effect of low protein number and affect predominantly the locus which are highly Hog1-dependent, like *STL1*, on the contrary to less dependent ones, like p*GPD1*, which interestingly displays a lower intrinsic noise [144]. Indeed, according to the Yeast GFP Fusion Localization Database [210], there is only be 149 and 504 molecules per cell of Hot1p and Sko1p respectively, compared to the 6780 Hog1p/cell and more than 300 genes induced upon osmotic shock [210]. Since we would not be able to resolve the presence of a single or double allele within a single PP7 focus, these data should be repeated in a dual color diploid strain or a strain with a single-labelled allele for more accurate measurement of each phenotype.

Increase PP7 assay throughput

The current biggest drawback of phage coat protein assay is their low throughput. Indeed, due to the dynamic nature of the readout, precision is lost with smaller time-resolution and thus limits the simultaneously monitoring of multiple reporter strains. A possible improvement to the method could make use of a microfluidic system. The first possibility would be to use the device published by Dénervaud and colleagues [211]. This parallel microchemostat array

enables to grow more than a thousand of yeast strains independently directly into the microfluidic device and image them either simultaneously or consecutively [211]. However, cells are neither trapped nor attached to the chamber's surface, it would thus require a highly performant tracking and segmentation software to follow single-cell over time. In addition, since cells will fill the entire chamber before being waste into the outlet, the confluence of the cells might not allow to make a homogenous induction with exogenous medium like osmotic solution. This system may thus more suitable to housekeeping genes monitoring.

The second possibility, would be to use a multiplied Alcatraz system [212], which enables to trap single yeast cells for long term imaging. Unlike the previous system, this device does not lead to cellular crowding and traps the cells at a defined position, it would thus simplify the segmentation and tracking parts, together with increasing the accuracy of medium perception by all the single cells. However, since this assay was initially designed to study aging, the same original single-cell is image throughout the generation, generating a bias in the single-cell population age over time. A possible was to circumvent this would be to flow cells one chamber after the other through an additional inlet.

6.3 General conclusion

In this study, we illustrate the versatile application of mRNA-based reporter assays, which are up to now the closest assays to report on promoter activity in living cells and that can be applied to a broad range of specimen, from single-cell to whole animal. Therefore, the use of phage coat proteins in the field of mRNA studies will continue to thrive. Together with the advent of single-cell sequencing, targets of interest will be identified and dynamics temporal and localization studies could be performed with the phage coat protein-based reporter assays.

References

References

1. Vjestica, A., et al., *Gamete fusion triggers bipartite transcription factor assembly to block re-fertilization.* Nature, 2018. **560**(7718): p. 397-400.
2. Rowley, A., et al., *Heat shock-mediated cell cycle blockage and G1 cyclin expression in the yeast Saccharomyces cerevisiae.* Mol Cell Biol, 1993. **13**(2): p. 1034-41.
3. Hoch, J.A. and K.I. Varughese, *Keeping signals straight in phosphorelay signal transduction.* J Bacteriol, 2001. **183**(17): p. 4941-9.
4. Wolf, Y.I., et al., *Updated clusters of orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of horizontal gene transfer.* Biol Direct, 2012. **7**: p. 46.
5. Galperin, M.Y., et al., *Phyletic Distribution and Lineage-Specific Domain Architectures of Archaeal Two-Component Signal Transduction Systems.* J Bacteriol, 2018. **200**(7).
6. Cooper, J.A., et al., *Similar effects of platelet-derived growth factor and epidermal growth factor on the phosphorylation of tyrosine in cellular proteins.* Cell, 1982. **31**(1): p. 263-73.
7. Boulton, T.G., et al., *An insulin-stimulated protein kinase similar to yeast kinases involved in cell cycle control.* Science, 1990. **249**(4964): p. 64-7.
8. Zaman, N., et al., *A Forster resonance energy transfer sensor for live-cell imaging of mitogen-activated protein kinase activity in Arabidopsis.* Plant J, 2019. **97**(5): p. 970-983.
9. Li, M., J. Liu, and C. Zhang, *Evolutionary history of the vertebrate mitogen activated protein kinases family.* PLoS One, 2011. **6**(10): p. e26999.
10. Widmann, C., et al., *Mitogen-activated protein kinase: conservation of a three-kinase module from yeast to human.* Physiol Rev, 1999. **79**(1): p. 143-80.
11. Boulton, T.G., et al., *ERKs: a family of protein-serine/threonine kinases that are activated and tyrosine phosphorylated in response to insulin and NGF.* Cell, 1991. **65**(4): p. 663-75.
12. Atienza, J.M., et al., *Human ERK1 induces filamentous growth and cell wall remodeling pathways in Saccharomyces cerevisiae.* J Biol Chem, 2000. **275**(27): p. 20638-46.
13. Levin-Salomon, V., et al., *When expressed in yeast, mammalian mitogen-activated protein kinases lose proper regulation and become spontaneously phosphorylated.* Biochem J, 2009. **417**(1): p. 331-40.
14. Galcheva-Gargova, Z., et al., *An osmosensing signal transduction pathway in mammalian cells.* Science, 1994. **265**(5173): p. 806-8.
15. Herskowitz, I., *MAP kinase pathways in yeast: for mating and more.* Cell, 1995. **80**(2): p. 187-97.
16. Goffeau, A., et al., *Life with 6000 genes.* Science, 1996. **274**(5287): p. 546, 563-7.

17. Sikorski, R.S. and P. Hieter, *A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae.* Genetics, 1989. **122**(1): p. 19-27.

18. Bar-Sagi, D. and A. Hall, *Ras and Rho GTPases: a family reunion.* Cell, 2000. **103**(2): p. 227-38.

19. Lindberg, R.A., A.M. Quinn, and T. Hunter, *Dual-specificity protein kinases: will any hydroxyl do?* Trends Biochem Sci, 1992. **17**(3): p. 114-9.

20. Bardwell, L., *Mechanisms of MAPK signalling specificity.* Biochem Soc Trans, 2006. **34**(Pt 5): p. 837-41.

21. Cargnello, M. and P.P. Roux, *Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases.* Microbiol Mol Biol Rev, 2011. **75**(1): p. 50-83.

22. Chen, R.E. and J. Thorner, *Function and regulation in MAPK signaling pathways: lessons learned from the yeast Saccharomyces cerevisiae.* Biochim Biophys Acta, 2007. **1773**(8): p. 1311-40.

23. Pryciak, P.M. and F.A. Huntress, *Membrane recruitment of the kinase cascade scaffold protein Ste5 by the Gbetagamma complex underlies activation of the yeast pheromone response pathway.* Genes Dev, 1998. **12**(17): p. 2684-97.

24. Zarrinpar, A., et al., *Sho1 and Pbs2 act as coscaffolds linking components in the yeast high osmolarity MAP kinase pathway.* Mol Cell, 2004. **14**(6): p. 825-32.

25. Faux, M.C. and J.D. Scott, *More on target with protein phosphorylation: conferring specificity by location.* Trends Biochem Sci, 1996. **21**(8): p. 312-5.

26. Jackson, C.L. and L.H. Hartwell, *Courtship in Saccharomyces cerevisiae: an early cell-cell interaction during mating.* Mol Cell Biol, 1990. **10**(5): p. 2202-13.

27. Nakayama, N., A. Miyajima, and K. Arai, *Common signal transduction system shared by STE2 and STE3 in haploid cells of Saccharomyces cerevisiae: autocrine cell-cycle arrest results from forced expression of STE2.* EMBO J, 1987. **6**(1): p. 249-54.

28. Whiteway, M., et al., *The STE4 and STE18 genes of yeast encode potential beta and gamma subunits of the mating factor receptor-coupled G protein.* Cell, 1989. **56**(3): p. 467-77.

29. Versele, M., K. Lemaire, and J.M. Thevelein, *Sex and sugar in yeast: two distinct GPCR systems.* EMBO Rep, 2001. **2**(7): p. 574-9.

30. Roskoski, R., Jr., *ERK1/2 MAP kinases: structure, function, and regulation.* Pharmacol Res, 2012. **66**(2): p. 105-43.

31. Elion, E.A., J.A. Brill, and G.R. Fink, *FUS3 represses CLN1 and CLN2 and in concert with KSS1 promotes signal transduction.* Proc Natl Acad Sci U S A, 1991. **88**(21): p. 9392-6.

32. Tkacz, J.S. and V.L. MacKay, *Sexual conjugation in yeast. Cell surface changes in response to the action of mating hormones.* J Cell Biol, 1979. **80**(2): p. 326-33.

33. Cross, F., et al., *Conjugation in Saccharomyces cerevisiae.* Annu Rev Cell Biol, 1988. **4**: p. 429-57.

34. Gimeno, C.J., et al., *Unipolar cell divisions in the yeast S. cerevisiae lead to filamentous growth: regulation by starvation and RAS.* Cell, 1992. **68**(6): p. 1077-90.

35. Liu, H., C.A. Styles, and G.R. Fink, *Elements of the yeast pheromone response pathway required for filamentous growth of diploids.* Science, 1993. **262**(5140): p. 1741-4.

36. Cullen, P.J. and G.F. Sprague, Jr., *The regulation of filamentous growth in yeast.* Genetics, 2012. **190**(1): p. 23-49.

37. Mosch, H.U., R.L. Roberts, and G.R. Fink, *Ras2 signals via the Cdc42/Ste20/mitogen-activated protein kinase module to induce filamentous growth in Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 1996. **93**(11): p. 5352-6.

38. Kuchin, S., V.K. Vyas, and M. Carlson, *Snf1 protein kinase and the repressors Nrg1 and Nrg2 regulate FLO11, haploid invasive growth, and diploid pseudohyphal differentiation.* Mol Cell Biol, 2002. **22**(12): p. 3994-4000.

39. Braus, G.H., et al., *Amino acid starvation and Gcn4p regulate adhesive growth and FLO11 gene expression in Saccharomyces cerevisiae.* Mol Biol Cell, 2003. **14**(10): p. 4272-84.

40. Costigan, C., S. Gehrung, and M. Snyder, *A synthetic lethal screen identifies SLK1, a novel protein kinase homolog implicated in yeast cell morphogenesis and cell growth.* Mol Cell Biol, 1992. **12**(3): p. 1162-78.

41. Ho, H.L., Y.S. Shiau, and M.Y. Chen, *Saccharomyces cerevisiaeTSC11/AVO3 participates in regulating cell integrity and functionally interacts with components of the Tor2 complex.* Curr Genet, 2005. **47**(5): p. 273-88.

42. Maeda, T., S.M. Wurgler-Murphy, and H. Saito, *A two-component system that regulates an osmosensing MAP kinase cascade in yeast.* Nature, 1994. **369**(6477): p. 242-5.

43. Blomberg, A. and L. Adler, *Roles of glycerol and glycerol-3-phosphate dehydrogenase (NAD+) in acquired osmotolerance of Saccharomyces cerevisiae.* J Bacteriol, 1989. **171**(2): p. 1087-92.

44. Chowdhury, S., K.W. Smith, and M.C. Gustin, *Osmotic stress and the yeast cytoskeleton: phenotype-specific suppression of an actin mutation.* J Cell Biol, 1992. **118**(3): p. 561-71.

45. Belli, G., et al., *Osmotic stress causes a G1 cell cycle delay and downregulation of Cln3/Cdc28 activity in Saccharomyces cerevisiae.* Mol Microbiol, 2001. **39**(4): p. 1022-35.

46. Ota, I.M. and A. Varshavsky, *A yeast protein similar to bacterial two-component regulators.* Science, 1993. **262**(5133): p. 566-9.

47. Maeda, T., M. Takekawa, and H. Saito, *Activation of yeast PBS2 MAPKK by MAPKKKs or by binding of an SH3-containing osmosensor.* Science, 1995. **269**(5223): p. 554-8.

48. Posas, F. and H. Saito, *Osmotic activation of the HOG MAPK pathway via Ste11p MAPKKK: scaffold role of Pbs2p MAPKK.* Science, 1997. **276**(5319): p. 1702-5.

49. Wu, C., et al., *Adaptor protein Ste50p links the Ste11p MEKK to the HOG pathway through plasma membrane association.* Genes Dev, 2006. **20**(6): p. 734-46.

50. Ferrigno, P., et al., *Regulated nucleo/cytoplasmic exchange of HOG1 MAPK requires the importin beta homologs NMD5 and XPO1.* EMBO J, 1998. **17**(19): p. 5606-14.

51. Posas, F., et al., *The transcriptional response of yeast to saline stress.* J Biol Chem, 2000. **275**(23): p. 17249-55.

52. Tamas, M.J., et al., *Fps1p controls the accumulation and release of the compatible solute glycerol in yeast osmoregulation.* Mol Microbiol, 1999. **31**(4): p. 1087-104.

53. Albertyn, J., et al., *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.* Mol Cell Biol, 1994. **14**(6): p. 4135-44.

54. Gustin, M.C., et al., *MAP kinase pathways in the yeast Saccharomyces cerevisiae.* Microbiol Mol Biol Rev, 1998. **62**(4): p. 1264-300.

55. Westfall, P.J., et al., *Stress resistance and signal fidelity independent of nuclear MAPK function.* Proc Natl Acad Sci U S A, 2008. **105**(34): p. 12212-7.

56. Hohmann, S., *Control of high osmolarity signalling in the yeast Saccharomyces cerevisiae.* FEBS Lett, 2009. **583**(24): p. 4025-9.

57. Maeda, T., A.Y. Tsai, and H. Saito, *Mutations in a protein tyrosine phosphatase gene (PTP2) and a protein serine/threonine phosphatase gene (PTC1) cause a synthetic growth defect in Saccharomyces cerevisiae.* Mol Cell Biol, 1993. **13**(9): p. 5408-17.

58. Jacoby, T., et al., *Two protein-tyrosine phosphatases inactivate the osmotic stress response pathway in yeast by targeting the mitogen-activated protein kinase, Hog1.* J Biol Chem, 1997. **272**(28): p. 17749-55.

59. Mattison, C.P., et al., *Differential regulation of the cell wall integrity mitogen-activated protein kinase pathway in budding yeast by the protein tyrosine phosphatases Ptp2 and Ptp3.* Mol Cell Biol, 1999. **19**(11): p. 7651-60.

60. Young, C., et al., *Role of Ptc2 type 2C Ser/Thr phosphatase in yeast high-osmolarity glycerol pathway inactivation.* Eukaryot Cell, 2002. **1**(6): p. 1032-40.

61. Warmka, J., et al., *Ptc1, a type 2C Ser/Thr phosphatase, inactivates the HOG pathway by dephosphorylating the mitogen-activated protein kinase Hog1.* Mol Cell Biol, 2001. **21**(1): p. 51-60.

62. Cook, K.E. and E.K. O'Shea, *Hog1 controls global reallocation of RNA Pol II upon osmotic shock in Saccharomyces cerevisiae.* G3 (Bethesda), 2012. **2**(9): p. 1129-36.

63. de Nadal, E. and F. Posas, *Osmostress-induced gene expression--a model to understand how stress-activated protein kinases (SAPKs) regulate transcription.* FEBS J, 2015. **282**(17): p. 3275-85.

64. Alepuz, P.M., et al., *Osmostress-induced transcription by Hot1 depends on a Hog1-mediated recruitment of the RNA Pol II.* EMBO J, 2003. **22**(10): p. 2433-42.

65. Rep, M., et al., *Osmotic stress-induced gene expression in Saccharomyces cerevisiae requires Msn1p and the novel nuclear factor Hot1p.* Mol Cell Biol, 1999. **19**(8): p. 5474-85.

66. Rep, M., et al., *The transcriptional response of Saccharomyces cerevisiae to osmotic shock. Hot1p and Msn2p/Msn4p are required for the induction of subsets of high osmolarity glycerol pathway-dependent genes.* J Biol Chem, 2000. **275**(12): p. 8290-300.

67. de Nadal, E., L. Casadome, and F. Posas, *Targeting the MEF2-like transcription factor Smp1 by the stress-activated Hog1 mitogen-activated protein kinase.* Mol Cell Biol, 2003. **23**(1): p. 229-37.

68. Proft, M. and R. Serrano, *Repressors and upstream repressing sequences of the stress-regulated ENA1 gene in Saccharomyces cerevisiae: bZIP protein Sko1p confers HOG-dependent osmotic regulation.* Mol Cell Biol, 1999. **19**(1): p. 537-46.

69. Marchler, G., et al., *A Saccharomyces cerevisiae UAS element controlled by protein kinase A activates transcription in response to a variety of stress conditions.* EMBO J, 1993. **12**(5): p. 1997-2003.

70. Proft, M., et al., *Regulation of the Sko1 transcriptional repressor by the Hog1 MAP kinase in response to osmotic stress.* EMBO J, 2001. **20**(5): p. 1123-33.

71. Proft, M. and K. Struhl, *Hog1 kinase converts the Sko1-Cyc8-Tup1 repressor complex into an activator that recruits SAGA and SWI/SNF in response to osmotic stress.* Mol Cell, 2002. **9**(6): p. 1307-17.

72. Marquez, J.A., et al., *The Ssn6-Tup1 repressor complex of Saccharomyces cerevisiae is involved in the osmotic induction of HOG-dependent and -independent genes.* EMBO J, 1998. **17**(9): p. 2543-53.

73. Capaldi, A.P., et al., *Structure and function of a transcriptional network activated by the MAPK Hog1.* Nat Genet, 2008. **40**(11): p. 1300-6.

74. Mas, G., et al., *Recruitment of a chromatin remodelling complex by the Hog1 MAP kinase to stress genes.* EMBO J, 2009. **28**(4): p. 326-36.

75. Zapater, M., et al., *Selective requirement for SAGA in Hog1-mediated gene expression depending on the severity of the external osmostress conditions.* Mol Cell Biol, 2007. **27**(11): p. 3900-10.

76. Nadal-Ribelles, M., et al., *Hog1 bypasses stress-mediated down-regulation of transcription by RNA polymerase II redistribution and chromatin remodeling.* Genome Biol, 2012. **13**(11): p. R106.

77. Pelet, S., et al., *Transient activation of the HOG MAPK pathway regulates bimodal gene expression.* Science, 2011. **332**(6030): p. 732-5.

78. Pelet, S. and M. Peter, *Dynamic processes at stress promoters regulate the bimodal expression of HOG response genes.* Commun Integr Biol, 2011. **4**(6): p. 699-702.

79. De Nadal, E., et al., *The MAPK Hog1 recruits Rpd3 histone deacetylase to activate osmoresponsive genes.* Nature, 2004. **427**(6972): p. 370-4.

80. Cairns, B.R., et al., *RSC, an essential, abundant chromatin-remodeling complex.* Cell, 1996. **87**(7): p. 1249-60.

81. Nadal-Ribelles, M., et al., *H3K4 monomethylation dictates nucleosome dynamics and chromatin remodeling at stress-responsive genes.* Nucleic Acids Res, 2015. **43**(10): p. 4937-49.

82. Soares, L.M., et al., *Determinants of Histone H3K4 Methylation Patterns.* Mol Cell, 2017. **68**(4): p. 773-785 e6.

83. Klopf, E., et al., *Cooperation between the INO80 complex and histone chaperones determines adaptation of stress gene transcription in the yeast Saccharomyces cerevisiae.* Mol Cell Biol, 2009. **29**(18): p. 4994-5007.

84. Klopf, E., et al., *INO80 represses osmostress induced gene expression by resetting promoter proximal nucleosomes.* Nucleic Acids Res, 2017. **45**(7): p. 3752-3766.

85. Zenklusen, D., D.R. Larson, and R.H. Singer, *Single-RNA counting reveals alternative modes of gene expression in yeast.* Nat Struct Mol Biol, 2008. **15**(12): p. 1263-71.

86. Elowitz, M.B., et al., *Stochastic gene expression in a single cell.* Science, 2002. **297**(5584): p. 1183-1186.

87. Golding, I., et al., *Real-time kinetics of gene activity in individual bacteria.* Cell, 2005. **123**(6): p. 1025-36.

88. Larson, D.R., et al., *Real-time observation of transcription initiation and elongation on an endogenous yeast gene.* Science, 2011. **332**(6028): p. 475-8.

89. Garcia, H.G., et al., *Quantitative imaging of transcription in living Drosophila embryos links polymerase activity to patterning.* Curr Biol, 2013. **23**(21): p. 2140-5.

90. Maamar, H., A. Raj, and D. Dubnau, *Noise in gene expression determines cell fate in Bacillus subtilis.* Science, 2007. **317**(5837): p. 526-9.

91. Raj, A., et al., *Variability in gene expression underlies incomplete penetrance.* Nature, 2010. **463**(7283): p. 913-8.

92. Locke, J.C., et al., *Stochastic pulse regulation in bacterial stress response.* Science, 2011. **334**(6054): p. 366-9.

93. Molina, N., et al., *Stimulus-induced modulation of transcriptional bursting in a single mammalian gene.* Proc Natl Acad Sci U S A, 2013. **110**(51): p. 20563-8.

94. Suter, D.M., et al., *Mammalian genes are transcribed with widely different bursting kinetics.* Science, 2011. **332**(6028): p. 472-4.

95. Timmers, H.T.M. and L. Tora, *Transcript Buffering: A Balancing Act between mRNA Synthesis and mRNA Degradation.* Mol Cell, 2018. **72**(1): p. 10-17.

96. Lionnet, T. and R.H. Singer, *Transcription goes digital.* EMBO Rep, 2012. **13**(4): p. 313-21.

97. Cai, L., C.K. Dalal, and M.B. Elowitz, *Frequency-modulated nuclear localization bursts coordinate gene regulation.* Nature, 2008. **455**(7212): p. 485-90.

98. Chong, S., et al., *Mechanism of transcriptional bursting in bacteria.* Cell, 2014. **158**(2): p. 314-26.

99. Drlica, K., *Control of bacterial DNA supercoiling.* Mol Microbiol, 1992. **6**(4): p. 425-33.

100. Lenstra, T.L., et al., *Transcription Dynamics in Living Cells.* Annu Rev Biophys, 2016. **45**: p. 25-47.

101. Kouzine, F., et al., *Transcription-dependent dynamic supercoiling is a short-range genomic force.* Nat Struct Mol Biol, 2013. **20**(3): p. 396-403.

102. Levens, D. and D.R. Larson, *A new twist on transcriptional bursting.* Cell, 2014. **158**(2): p. 241-2.

103. Banerji, J., S. Rusconi, and W. Schaffner, *Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences.* Cell, 1981. **27**(2 Pt 1): p. 299-308.

104. Shlyueva, D., G. Stampfel, and A. Stark, *Transcriptional enhancers: from properties to genome-wide predictions.* Nat Rev Genet, 2014. **15**(4): p. 272-86.

105. Fukaya, T., B. Lim, and M. Levine, *Enhancer Control of Transcriptional Bursting.* Cell, 2016. **166**(2): p. 358-368.

106. Hampsey, M., et al., *Control of eukaryotic gene expression: gene loops and transcriptional memory.* Adv Enzyme Regul, 2011. **51**(1): p. 118-25.

107. Randise-Hinchliff, C.E. and J.H. Brickner, *A new direction for gene looping.* Dev Cell, 2012. **23**(5): p. 919-21.

108. Brickner, D.G., et al., *H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state.* PLoS Biol, 2007. **5**(4): p. e81.

109. O'Sullivan, J.M., et al., *Gene loops juxtapose promoters and terminators in yeast.* Nat Genet, 2004. **36**(9): p. 1014-8.

110. Hansen, A.S. and E.K. O'Shea, *cis Determinants of Promoter Threshold and Activation Timescale.* Cell Rep, 2015. **12**(8): p. 1226-33.

111. Kim, H.D. and E.K. O'Shea, *A quantitative model of transcription factor-activated gene expression.* Nat Struct Mol Biol, 2008. **15**(11): p. 1192-8.

112. Nicolas, D., et al., *Modulation of transcriptional burst frequency by histone acetylation.* Proc Natl Acad Sci U S A, 2018. **115**(27): p. 7153-7158.

113. Donovan, B.T., et al., *Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting.* EMBO J, 2019. **38**(12).

114. Adelman, K. and J.T. Lis, *Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans.* Nat Rev Genet, 2012. **13**(10): p. 720-31.

115. Margaritis, T. and F.C. Holstege, *Poised RNA polymerase II gives pause for thought.* Cell, 2008. **133**(4): p. 581-4.

116. Rougvie, A.E. and J.T. Lis, *The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged.* Cell, 1988. **54**(6): p. 795-804.

117. Bartman, C.R., et al., *Transcriptional Burst Initiation and Polymerase Pause Release Are Key Control Points of Transcriptional Regulation.* Mol Cell, 2019. **73**(3): p. 519-532 e4.

118. Hnisz, D., et al., *A Phase Separation Model for Transcriptional Control.* Cell, 2017. **169**(1): p. 13-23.

119. Lim, B., et al., *Visualization of Transvection in Living Drosophila Embryos.* Mol Cell, 2018. **70**(2): p. 287-296 e6.

120. Boehning, M., et al., *RNA polymerase II clustering through carboxy-terminal domain phase separation.* Nat Struct Mol Biol, 2018. **25**(9): p. 833-840.

121. Rieder, D., Z. Trajanoski, and J.G. McNally, *Transcription factories.* Front Genet, 2012. **3**: p. 221.

122. Jackson, D.A., et al., *Visualization of focal sites of transcription within human nuclei.* EMBO J, 1993. **12**(3): p. 1059-65.

123. Wollman, A.J., et al., *Transcription factor clusters regulate genes in eukaryotic cells.* Elife, 2017. **6**.

124. Moorman, C., et al., *Hotspots of transcription factor colocalization in the genome of Drosophila melanogaster.* Proc Natl Acad Sci U S A, 2006. **103**(32): p. 12027-32.

125. Krumlauf, R., *Analysis of gene expression by northern blot.* Mol Biotechnol, 1994. **2**(3): p. 227-42.

126. Heid, C.A., et al., *Real time quantitative PCR.* Genome Res, 1996. **6**(10): p. 986-94.

127. Tang, F., et al., *mRNA-Seq whole-transcriptome analysis of a single cell.* Nat Methods, 2009. **6**(5): p. 377-82.

128. Rudkin, G.T. and B.D. Stollar, *High resolution detection of DNA-RNA hybrids in situ by indirect immunofluorescence.* Nature, 1977. **265**(5593): p. 472-3.

129. Bauman, J.G., et al., *A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridization of fluorochromelabelled RNA.* Exp Cell Res, 1980. **128**(2): p. 485-90.

130. Femino, A.M., et al., *Visualization of single RNA transcripts in situ.* Science, 1998. **280**(5363): p. 585-90.

131. Raj, A., et al., *Imaging individual mRNA molecules using multiple singly labeled probes.* Nat Methods, 2008. **5**(10): p. 877-9.

132. Prasher, D.C., et al., *Primary structure of the Aequorea victoria green-fluorescent protein.* Gene, 1992. **111**(2): p. 229-33.

133. Chalfie, M., et al., *Green fluorescent protein as a marker for gene expression.* Science, 1994. **263**(5148): p. 802-5.

134. Heim, R., D.C. Prasher, and R.Y. Tsien, *Wavelength mutations and posttranslational autoxidation of green fluorescent protein.* Proc Natl Acad Sci U S A, 1994. **91**(26): p. 12501-4.

135. Nagai, T., et al., *A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications.* Nat Biotechnol, 2002. **20**(1): p. 87-90.

136. Pedelacq, J.D., et al., *Engineering and characterization of a superfolder green fluorescent protein.* Nat Biotechnol, 2006. **24**(1): p. 79-88.

137. Silhavy, T.J. and J.R. Beckwith, *Uses of lac fusions for the study of biological problems.* Microbiol Rev, 1985. **49**(4): p. 398-418.

138. Kalnins, A., et al., *Sequence of the lacZ gene of Escherichia coli.* EMBO J, 1983. **2**(4): p. 593-7.

139. Horwitz, J.P., et al., *Substrates for Cytochemical Demonstration of Enzyme Activity. I. Some Substituted 3-Indolyl-Beta-D-Glycopyranosides.* J Med Chem, 1964. **7**: p. 574-5.

140. Contag, C.H., et al., *Photonic detection of bacterial pathogens in living hosts.* Mol Microbiol, 1995. **18**(4): p. 593-603.

141. Leskinen, P., M. Virta, and M. Karp, *One-step measurement of firefly luciferase activity in yeast.* Yeast, 2003. **20**(13): p. 1109-13.

142. Thorne, N., J. Inglese, and D.S. Auld, *Illuminating insights into firefly luciferase and other bioluminescent reporters used in chemical biology.* Chem Biol, 2010. **17**(6): p. 646-57.

143. Mazo-Vargas, A., et al., *Measuring fast gene dynamics in single cells with time-lapse luminescence microscopy.* Mol Biol Cell, 2014. **25**(22): p. 3699-708.

144. Aymoz, D., et al., *Real-time quantification of protein expression at the single-cell level via dynamic protein synthesis translocation reporters.* Nat Commun, 2016. **7**: p. 11304.

145. Reinke, A.W., R.A. Grant, and A.E. Keating, *A synthetic coiled-coil interactome provides heterospecific modules for molecular engineering.* J Am Chem Soc, 2010. **132**(17): p. 6025-31.

146. Varshavsky, A., *The N-end rule: functions, mysteries, uses.* Proc Natl Acad Sci U S A, 1996. **93**(22): p. 12142-9.

147. Paige, J.S., K.Y. Wu, and S.R. Jaffrey, *RNA mimics of green fluorescent protein.* Science, 2011. **333**(6042): p. 642-6.

148. Guet, D., et al., *Combining Spinach-tagged RNA and gene localization to image gene expression in live yeast.* Nat Commun, 2015. **6**: p. 8882.

149. Bertrand, E., et al., *Localization of ASH1 mRNA particles in living yeast.* Mol Cell, 1998. **2**(4): p. 437-45.

150. Daigle, N. and J. Ellenberg, *LambdaN-GFP: an RNA reporter system for live-cell imaging.* Nat Methods, 2007. **4**(8): p. 633-6.

151. Lim, F., M. Spingola, and D.S. Peabody, *The RNA-binding site of bacteriophage Qbeta coat protein.* J Biol Chem, 1996. **271**(50): p. 31839-45.

152. Urbanek, M.O., et al., *RNA imaging in living cells - methods and applications.* RNA Biol, 2014. **11**(8): p. 1083-95.

153. Buxbaum, A.R., G. Haimovich, and R.H. Singer, *In the right place at the right time: visualizing and understanding mRNA localization.* Nat Rev Mol Cell Biol, 2015. **16**(2): p. 95-109.

154. Chao, J.A., et al., *Structural basis for the coevolution of a viral RNA-protein complex.* Nat Struct Mol Biol, 2008. **15**(1): p. 103-5.

155. Lenstra, T.L., et al., *Single-Molecule Imaging Reveals a Switch between Spurious and Functional ncRNA Transcription.* Mol Cell, 2015. **60**(4): p. 597-610.

156. Hocine, S., et al., *Single-molecule analysis of gene expression using two-color RNA labeling in live yeast.* Nat Methods, 2013. **10**(2): p. 119-21.

157. Carmo-Fonseca, M. and T. Kirchhausen, *The timing of pre-mRNA splicing visualized in real-time.* Nucleus, 2014. **5**(1): p. 11-4.

158. Koker, T., A. Fernandez, and F. Pinaud, *Characterization of Split Fluorescent Protein Variants and Quantitative Analyses of Their Self-Assembly Process.* Sci Rep, 2018. **8**(1): p. 5344.

159. Wu, B., J. Chen, and R.H. Singer, *Background free imaging of single mRNAs in live cells using split fluorescent proteins.* Sci Rep, 2014. **4**: p. 3615.

160. Ferraro, T., et al., *New methods to image transcription in living fly embryos: the insights so far, and the prospects.* Wiley Interdiscip Rev Dev Biol, 2016. **5**(3): p. 296-310.

161. Martin, R.M., et al., *Live-cell visualization of pre-mRNA splicing with single-molecule sensitivity.* Cell Rep, 2013. **4**(6): p. 1144-55.

162. Lenstra, T.L. and D.R. Larson, *Single-Molecule mRNA Detection in Live Yeast.* Curr Protoc Mol Biol, 2016. **113**: p. 14 24 1-14 24 15.

163. Garcia, J.F. and R. Parker, *MS2 coat proteins bound to yeast mRNAs block 5' to 3' degradation and trap mRNA decay products: implications for the localization of mRNAs by MS2-MCP system.* RNA, 2015. **21**(8): p. 1393-5.

164. Heinrich, S., et al., *Stem-loop RNA labeling can affect nuclear and cytoplasmic mRNA processing.* RNA, 2017. **23**(2): p. 134-141.

165. Haimovich, G., et al., *Use of the MS2 aptamer and coat protein for RNA localization in yeast: A response to "MS2 coat proteins bound to yeast mRNAs block 5' to 3' degradation and trap mRNA decay products: implications for the localization of mRNAs by MS2-MCP system".* RNA, 2016. **22**(5): p. 660-6.

166. Lim, F. and D.S. Peabody, *RNA recognition site of PP7 coat protein.* Nucleic Acids Res, 2002. **30**(19): p. 4138-44.

167. Wu, B., J.A. Chao, and R.H. Singer, *Fluorescence fluctuation spectroscopy enables quantitative imaging of single mRNAs in living cells.* Biophys J, 2012. **102**(12): p. 2936-44.

168. Wosika, V., et al., *New families of single integration vectors and gene tagging plasmids for genetic manipulations in budding yeast.* Mol Genet Genomics, 2016. **291**(6): p. 2231-2240.

169. Hieter, R.S.S.a.P., *A System of Shuttle Vectors and Yeast Host Strains Designed for Efficient Manipulation of DNA in Saccharomyces ceratisiae.* Genetics, 1989. **122**: p. 19-27.

170. Lang, G.I., A.W. Murray, and D. Botstein, *The cost of gene expression underlies a fitness trade-off in yeast.* Proc Natl Acad Sci U S A, 2009. **106**(14): p. 5755-60.

171. Gygi, S.P., et al., *Correlation between protein and mRNA abundance in yeast.* Mol Cell Biol, 1999. **19**(3): p. 1720-30.

172. Liu, Y., A. Beyer, and R. Aebersold, *On the Dependency of Cellular Protein Levels on mRNA Abundance.* Cell, 2016. **165**(3): p. 535-50.

173. Yan, X., et al., *Dynamics of Translation of Single mRNA Molecules In Vivo.* Cell, 2016. **165**(4): p. 976-89.

174. Tanenbaum, M.E., et al., *A protein-tagging system for signal amplification in gene expression and fluorescence imaging.* Cell, 2014. **159**(3): p. 635-46.

175. Baim, S.B. and F. Sherman, *mRNA structures influencing translation in the yeast Saccharomyces cerevisiae.* Mol Cell Biol, 1988. **8**(4): p. 1591-601.

176. Lamping, E., M. Niimi, and R.D. Cannon, *Small, synthetic, GC-rich mRNA stem-loop modules 5' proximal to the AUG start-codon predictably tune gene expression in yeast.* Microb Cell Fact, 2013. **12**: p. 74.

177. Tutucci, E., et al., *An improved MS2 system for accurate reporting of the mRNA life cycle.* Nat Methods, 2018. **15**(1): p. 81-89.

178. Pelet, S., et al., *An integrated image analysis platform to quantify signal transduction in single cells.* Integr Biol (Camb), 2012. **4**(10): p. 1274-82.

179. Filonov, G.S., et al., *Bright and stable near-infrared fluorescent protein for in vivo imaging.* Nat Biotechnol, 2011. **29**(8): p. 757-61.

180. Yamanishi, M., et al., *A genome-wide activity assessment of terminator regions in Saccharomyces cerevisiae provides a ''terminatome'' toolbox.* ACS Synth Biol, 2013. **2**(6): p. 337-47.

181. Gorner, W., et al., *Acute glucose starvation activates the nuclear localization signal of a stress-specific yeast transcription factor.* EMBO J, 2002. **21**(1-2): p. 135-44.

182. Miermont, A., et al., *Severe osmotic compression triggers a slowdown of intracellular signaling, which can be explained by molecular crowding.* Proc Natl Acad Sci U S A, 2013. **110**(14): p. 5725-30.

183. Obrig, T.G., et al., *The mechanism by which cycloheximide and related glutarimide antibiotics inhibit peptide synthesis on reticulocyte ribosomes.* J Biol Chem, 1971. **246**(1): p. 174-81.

184. Knight, B., et al., *Two distinct promoter architectures centered on dynamic nucleosomes control ribosomal protein gene transcription.* Genes Dev, 2014. **28**(15): p. 1695-709.

185. Bindels, D.S., et al., *mScarlet: a bright monomeric red fluorescent protein for cellular imaging.* Nat Methods, 2017. **14**(1): p. 53-56.

186. Chubb, J.R., et al., *Developmental timing in Dictyostelium is regulated by the Set1 histone methyltransferase.* Dev Biol, 2006. **292**(2): p. 519-32.

187. Bothma, J.P., et al., *Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living Drosophila embryos.* Proc Natl Acad Sci U S A, 2014. **111**(29): p. 10598-603.

188. Park, H.Y., et al., *Visualization of dynamics of single endogenous mRNA labeled in live mouse.* Science, 2014. **343**(6169): p. 422-4.

189. Campbell, P.D., et al., *Dynamic visualization of transcription and RNA subcellular localization in zebrafish.* Development, 2015. **142**(7): p. 1368-74.

190. Buxbaum, A.R., B. Wu, and R.H. Singer, *Single beta-actin mRNA detection in neurons reveals a mechanism for regulating its translatability.* Science, 2014. **343**(6169): p. 419-22.

191. Aymoz, D., et al., *Timing of gene expression in a cell-fate decision system.* Mol Syst Biol, 2018. **14**(4): p. e8024.

192. Medstrand, P., et al., *Impact of transposable elements on the evolution of mammalian gene regulation.* Cytogenet Genome Res, 2005. **110**(1-4): p. 342-52.

193. Kim, T.K. and R. Shiekhattar, *Architectural and Functional Commonalities between Enhancers and Promoters.* Cell, 2015. **162**(5): p. 948-59.

194. Goldman, S.R., R.H. Ebright, and B.E. Nickels, *Direct detection of abortive RNA transcripts in vivo.* Science, 2009. **324**(5929): p. 927-8.

195. Regot, S., et al., *The Hog1 stress-activated protein kinase targets nucleoporins to control mRNA export upon stress.* J Biol Chem, 2013. **288**(24): p. 17384-98.

196. Grant, P.A., et al., *Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex.* Genes Dev, 1997. **11**(13): p. 1640-50.

197. VanDemark, A.P., et al., *Autoregulation of the rsc4 tandem bromodomain by gcn5 acetylation.* Mol Cell, 2007. **27**(5): p. 817-28.

198. Haruki, H., J. Nishikawa, and U.K. Laemmli, *The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes.* Mol Cell, 2008. **31**(6): p. 925-32.

199. Nishimura, K., et al., *An auxin-based degron system for the rapid depletion of proteins in nonplant cells.* Nat Methods, 2009. **6**(12): p. 917-22.

200. Bruzzone, M.J., et al., *Distinct patterns of histone acetyltransferase and Mediator deployment at yeast protein-coding genes.* Genes Dev, 2018. **32**(17-18): p. 1252-1265.

201. Tramantano, M., et al., *Constitutive turnover of histone H2A.Z at yeast promoters requires the preinitiation complex.* Elife, 2016. **5**.

202. Morawska, M. and H.D. Ulrich, *An expanded tool kit for the auxin-inducible degron system in budding yeast.* Yeast, 2013. **30**(9): p. 341-51.

203. Shariati, S.A., et al., *Reversible Disruption of Specific Transcription Factor-DNA Interactions Using CRISPR/Cas9.* Mol Cell, 2019. **74**(3): p. 622-633 e4.

204. Los, G.V., et al., *HaloTag: a novel protein labeling technology for cell imaging and protein analysis.* ACS Chem Biol, 2008. **3**(6): p. 373-82.

205. Blobel, G., *Gene gating: a hypothesis.* Proc Natl Acad Sci U S A, 1985. **82**(24): p. 8527-9.

206. van Berkum, N.L., et al., *Hi-C: a method to study the three-dimensional architecture of genomes.* J Vis Exp, 2010(39).

207. Rao, S.S., et al., *A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping.* Cell, 2014. **159**(7): p. 1665-80.

208. Battich, N., T. Stoeger, and L. Pelkmans, *Control of Transcript Variability in Single Mammalian Cells.* Cell, 2015. **163**(7): p. 1596-610.

209. Escote, X., et al., *Hog1 mediates cell-cycle arrest in G1 phase by the dual targeting of Sic1.* Nat Cell Biol, 2004. **6**(10): p. 997-1002.

210. Huh, W.K., et al., *Global analysis of protein localization in budding yeast.* Nature, 2003. **425**(6959): p. 686-91.

211. Denervaud, N., et al., *A chemostat array enables the spatio-temporal analysis of the yeast proteome.* Proc Natl Acad Sci U S A, 2013. **110**(39): p. 15842-7.

212. Crane, M.M., et al., *A microfluidic system for studying ageing and dynamic single-cell responses in budding yeast.* PLoS One, 2014. **9**(6): p. e100042.

Annex 1: additional reporter strains

Supplementary Table 3: coupled PP7-dPSTR reporter strains

Strain name	Ancestor strain	Plasmid	dPSTR part	Coupled system	Integration site	PP7 expression	YQ + 0.4M *0.2M*
yVW61	yVW58	pVW81	pSIVu pRPS20 CFP SZ2 tCYC1	pSTL1 2xNLS SZ1 PP7sl 400bp tCYC1	URA3	pCYC1 PP7 2xGFP tCYC1	1745 1618
yVW62	yVW59	pVW81	pSIVu pRPS20 CFP SZ2 tCYC1	pSTL1 2xNLS SZ1 PP7sl 3kb tCYC1	URA3	pCYC1 PP7 2xGFP tCYC1	1746
yVW63	ySP261	pVW71	pSIVu pRPS20 CFP SZ2 tCYC1	pSTL1 2xNLS SZ1 tCYC1	URA3	None	1767 1788 1822
yVW65	ySP261	pVW83	pSIVu pRPS20 CFP SZ2 tCYC1	pSTL1 2xNLS SZ1 PP7sl 1kb tSIF2	URA3	None	1748 1789
yVW72	yVW65	pVW81	pSIVu pRPS20 CFP SZ2 tCYC1	pSTL1 2xNLS SZ1 PP7sl 1kb tSIF2	URA3	pCYC1 PP7 2xGFP tCYC1	-
yVW67	ySP261	pVW95	pSIVu pRPS20 CFP SZ2 tCYC1	pSTL1 2xNLS SZ1 50bp PP7sl 400bp tCYC1	URA3	None	1747 1794
yVW73	yVW67	pVW81	pSIVu pRPS20 CFP SZ2 tCYC1	pSTL1 2xNLS SZ1 50bp PP7sl 400bp tCYC1	URA3	pCYC1 PP7 2xGFP tCYC1	-
yVW68	ySP261	pVW85	pSIVu pRPS20 CFP SZ2 tCYC1	pSTL1 2xNLS SZ1 3kb tSIF2	URA3	None	1749
yVW74	yVW68	pVW81	pSIVu pRPS20 CFP SZ2 tCYC1	pSTL1 2xNLS SZ1 PP7sl 3kb tSIF2	URA3	pCYC1 PP7 2xGFP tCYC1	1804
yVW75 (=yVW76)	ySP261	pVW103	pSIVu pRPS20 CFP SZ2 tCYC1	pSTL1 2xNLS SZ1 50bp PP7sl 1kb tSIF2	URA3	None	1819 1820
yVW78 (=yVW79)	yVW75	pVW81	pSIVu pRPS20 CFP SZ2 tCYC1	pSTL1 2xNLS SZ1 50bp PP7sl 1kb tSIF2	URA3	pCYC1 PP7 2xGFP tCYC1	1826 1881 1827 1882
*yVW95 (=*wt *mid thesis)*	yVW94	pVW81	pSIVu pRPL24A mCherry SZ2	pSTL1 2xNLS SZ1 50bp PP7sl 1kb tSIF2	URA3	pCYC1 PP7 2xGFP tCYC1	2029 2150 2179 2498
yVW110	yED159	pVW162	pSIVu pRPS6B mCherry SZ2	pSTL1 2xNLS SZ1 50bp PP7sl 1kb tSIF2	URA3	None	2387
yVW111	yVW110	pVW81	pSIVu pRPS6B mCherry SZ2	pSTL1 2xNLS SZ1 50bp PP7sl 1kb tSIF2	URA3	pCYC1 PP7 2xGFP tCYC1	-
yVW191	yED212	pVW199	pSIVu pRPL24A mCherry SZ4	pSTL1 4xSZ3 PP7sl 1kb tSIF2	URA3	None	*2824*
yVW193	yVW191	pVW201	pSIVu pRPL24A mCherry SZ4	pSTL1 4xSZ3 PP7sl 1kb tSIF2	URA3	pCYC1mut PP7 sfGFP tCYC1	2848 *2849*
yVW205	yED212	pVW211	pSIVu pRPL24A mCherry SZ4	pSTL1 2xNLS 4xSZ3 PP7sl 50bp 1kb tSIF2	URA3	None	2955 *2955* *2962* *2963*
yVW210	yED212	pVW208	pSIVu pRPL24A mCherry SZ2	pSTL1 UbiY 2xnpNLS 4xSZ1 50bp PP7sl 1kb tSIF2	URA3	None	*2982* *3059*
yVW211	yVW205	pVW212	pSIVu pRPL24A mCitrine SZ4	pSTL1 2xNLS 4xSZ3 PP7sl 50bp 1kb tSIF2	URA3	pSIVh pCYC1 PP7 mCherry tCYC1	-
yVW213	yVW210	pVW201	pSIVu pRPL24A mCherry SZ2	pSTL1 UbiY 2xnpNLS 4xSZ1 50bp PP7sl 1kb tSIF2	URA3	pSIVh pCYC1mut PP7 sfGFP	2981 *2980*
yVW220	yVW210	pVW220	pSIVu pRPL24A mCherry SZ2	pSTL1 UbiY 2xnpNLS 4xSZ1 50bp PP7sl 1kb tSIF2	URA3	pSIVh pCYC1mut PP7 GFPenvy tCYC1	*3007* 3008 *3037*
yVW223	yVW210	pVW221	pSIVu pRPL24A mCherry SZ2	pSTL1 UbiY 2xnpNLS 4xSZ1 50bp PP7sl 1kb tSIF2	URA3	pSIVh pCYC1mut PP7 GFPivy tCYC1	*3028* 3028
yVW232 (=210 stable)	yED212	pVW219	pSIVu pRPL24A mCherry SZ2	pSTL1 2xnpNLS 4xSZ1 50bp PP7sl 1kb tSIF2	URA3	None	
yVW278	yVW210	pVW127	pSIVu pRPL24A mCherry SZ2	pSTL1 UbiY 2xnpNLS 4xSZ1 50bp PP7sl 1kb tSIF2	URA3	pSIVh pCYC1 PP7-2xGFP tCYC1	
yVW447	yVW210	pVW297	pSIVu pRPL24A mCherry SZ2	pSTL1 UbiY 2xnpNLS 4xSZ1 50bp PP7sl 1kb tSIF2	URA3	pSIVh pADH1 PP7dFG-GFPenvy tCYC1	*3803* *4011* *4017*

						4018 *4019* *4020* *4021* *4022*

Supplementary Table 4: coupled PP7-Venus reporter strains

Strain name	Ancestor strain	Plasmid	Venus part	Coupled system	Integration site	PP7 expression	YQ + 0.4M *0.2M*
yVW24	yVW5	pVW11	Venus	pSTL1 Venus PP7sl	GLT1	pRS pCYC1 PP7 mCherry	*1279* *1344* 1345 1346 1498 *1535* 1769 *1775* 1883 *3518* *3519*
yVW51	yVW49	pVW11	Venus	pSTL1 Venus PP7sl	GLT1	pRS pCYC1 PP7 qCherry tCYC1	*1518*
yVW376	yVW190	pVW97	Venus	pSTL1 Venus 50bp PP7sl	GLT1	pSIVu pCYC1mut PP7-dCherry tCYC1	*3559*
yVW427	yVW446	pVW97	Venus	pSTL1 Venus 50bp PP7sl	GLT1	pSIVu pADH1 PP7 mCherry tCYC1	*3874* *3884* *3956* *3957* *3973*

Supplementary Table 5: pSTL1 Venus strains for PP7 stem-loops effects

Strain name	Ancestor strain	Plasmid	Coupled part	PP7 part	Integration site	YQ + 0.2M NaCl
ySP269	ySP2	-	-	-	-	2297 2336 2337 3137
yVW119	ySP269	pVW171	-	pCYC1 PP7-dCherry tCYC1	-	2389 2425 2427
yVW112	ySP269	pVW11	pSTL1 Venus PP7sl		GLT1	2297 2336 2337
yVW120	yVW112	pVW171	pSTL1 Venus PP7sl	pCYC1 PP7-dCherry tCYC1	GLT1	2389 2425 2427
yVW113	ySP269	pVW97	pSTL1 Venus 50bp PP7		GLT1	2297 2336 2337
yVW121	yVW113	pVW171	pSTL1 Venus 50bp PP7	pCYC1 PP7-dCherry tCYC1	GLT1	2389 2425 2427
yVW114	ySP269	pVW97	pSTL1 Venus		GLT1	2336 2337 3137 3149
yVW122	yVW114	pVW171	pSTL1 Venus	pCYC1 PP7-dCherry tCYC1	GLT1	2389 2425 2427

Supplementary Table 6: pSTL1 (1-800) truncations

Strain name	**Ancestor strain**	**Plasmid**	**pSTL1 variant**	**Venus**	**Integration site**	**Nickname**	**YQ + 0.4M *0.2M***
ySP269	ySP2					« background »	*3137*
yVW114	ySP269	pVW97				« wt »	*3137* *3149*
yVW259	ySP269	pVW238	164-800	Venus	GLT1	« 1xSTRE »	*3137*
yVW260	ySP269	pVW239	1-771	Venus	GLT1	« gaaaa less »	*3137* *3162*
yVW261	ySP269	pVW240	1-689/701-800	Venus	GLT1	« TATA less »	*3149*
yVW262	ySP269	pVW241	1-174/690-800	Venus	GLT1	« minimal »	*3149*
yVW263	ySP269	pVW243	1-174	Venus	GLT1		*3162*
yVW264	yVW262	Hog1d					*3205*
yVW265	yVW262	Hot1d					*3189*
yVW271	ySP269	pVW249	1-174/mCherry/690-800	Venus	GLT1	« sandwich »	*3234* *3261*
	yVW271	Hot1d					*3261*
	yVW271	Hog1d					*3261*

Annex 2: Chapter 4's strains genotyping

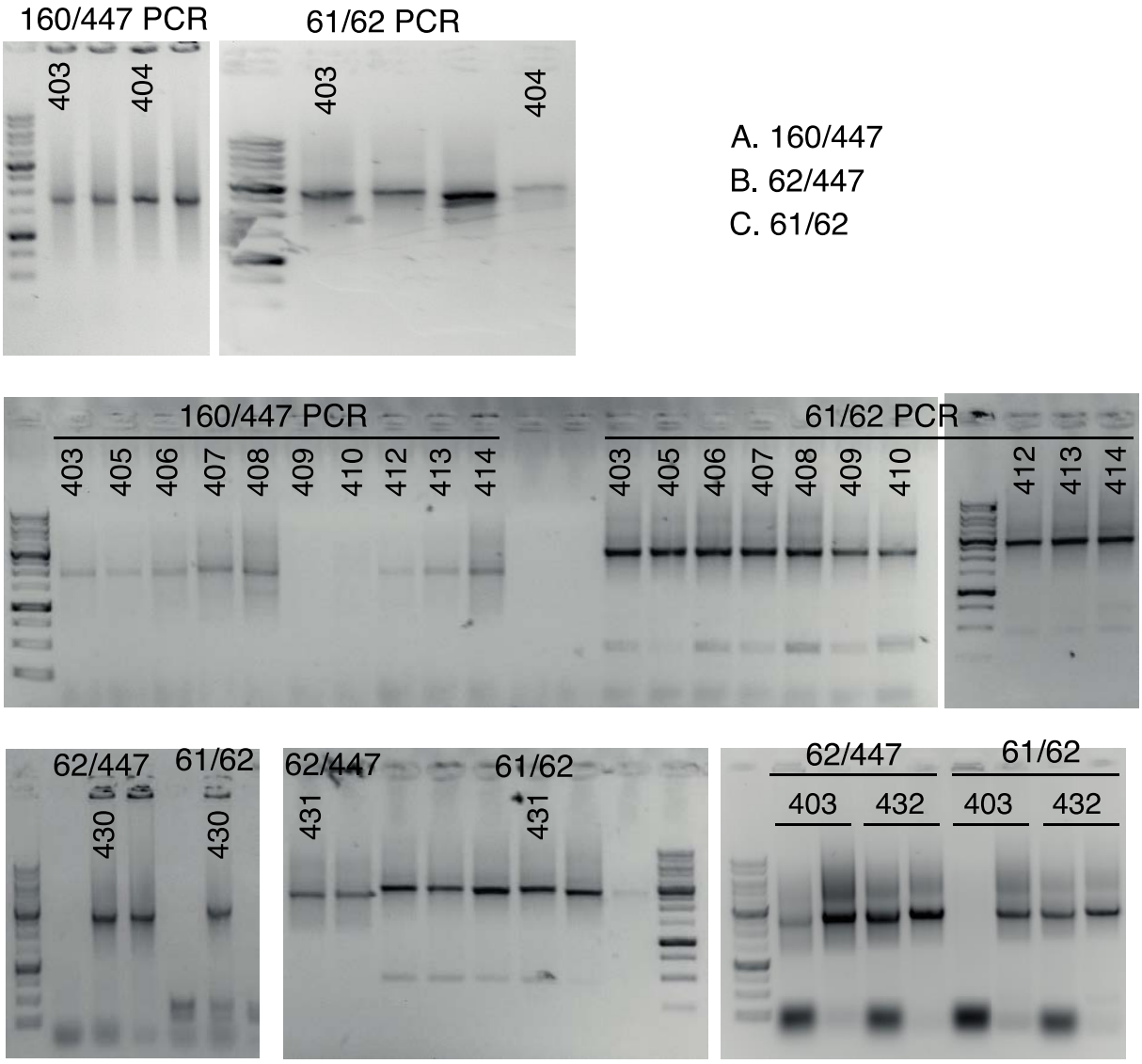


Figure 22: Colony PCR on strains from chapter 4.

Annex 3: chapter 4 strains construction and controls

pSTL1 PP7 strains- yVW403 background - Hta2-RFP+ PP7dFG-GFPenvy + STL1 PP7::GLT1

deletion	strain name	colony PCR date	PCR 62 or 390?	clones isolated	screening TL date	screening TL results
hot1	yVW405	180717	both	2, 3	180723	clones identical
hog1	yVW406	180627	both	1, 2	180711	clones identical
sko1	yVW407	180627	both	1, 2	180711	clones identical
spt3	yVW408	180627	390	1, 2	180711	clones identical
gcn5	yVW409	180627	both	2, 3	180710	clones identical
arp8	yVW410	180627	both	1, 2	180710	clones identical
asf1	yVW411	180627 + 181121	both	1A, 2A + 1,2,3	180711 + 181206	c1A and 2A different, 1 and 2 similar, 3 slightly different
set1	yVW412	180627	390	1, 2	180710	clones very similar
sus1	yVW413	180627	both	1, 2, 3	180723 + 180727	clone 2 sick, 1 and 3 similar
mlp1	yVW414	180627	both	1, 2	180710	clones identical
ada2	yVW415					
htz1	yVW416	181010	fw	2,3,4	181121	clones similar

Annex 4: chapter 4 strains microscopy experiments

Replicate experiments were the cells were further diluted are shown in yellow, but are not exhaustive. Replicate selected as representative of the strain and condition are indicated in green. Time-lapse experiments that were discarded due to cell physiology and/or imaging acquisition errors are indicated in red.

Supplementary Table 7: pSTL1 PP7 mutant strains

Strain	[NaCl]	Replicate n°	TL date	YQ n°	OD600	Cell number	Comments
yVW403	SD	1	181120	3739	0.18	313	identical, take higher cell number
pSTL1		2	181127	3765	0.31	307	identical, take higher cell number
wt	0.1M	1	181120	3745	0.19	404	much more PP7 positive cells and higher amplitude
		2	181127	3763	0.20	196	delayed compared to three others
		3	181129	3776	0.08	258	same hight as 2
		4	190118	3834	0.30	190	similar to rep 1
	0.15M	1	181127	3756	0.13	259	clearly higher than others,
		2	181129	3774	0.32	177	same as rep4 but more cells
		3	181204	3781	0.12	210	delayed compared to two others
		4	190118	3833	0.30	74	similar to rep 2
	0.2M	1	181120	3737	0.12	392	higher/wrong NaCl, highly delayed
		2	181127	3762	0.16	229	same dynamics as other concentrations, similar height to other replicate
		3	181129	3775	0.38	224	slightly faster than other rep and concentrations
		4	190118	3832	0.25		slightly delayed but similar to rep2 and rep3
	0.3M	1	190118	3831	0.20	153	delayed and lower compared to rep3 and to 0.2M
		2	190122	3839	0.09	176	shrinks more than two others, LED weird
		3	190124	3849	0.21	201	faster and higher than rep1
yVW405	0.1M	1	190222	3872	0.26	230	slightly slower
		2	190312	3898	0.31	276	dynamics like rep3
		3	190326	3936	0.11	334	dynamics like rep2, peaks and respcells number like rep1
hot1	0.2M	1	181120	3738	0.21	221	identical, take higher cell number
		2	181127	3757	0.21	63	shrink a bit less and adapt faster, much less cells
		3	181129	3772	0.32	349	identical, take higher cell number
	0.3M	1	190222	3871	0.21	316	all quite similar
		2	190312	3897	0.22	242	all quite similar
		3	190503	4004	0.17	240	all quite similar
yVW406	0.2M	1	181120	3740	0.20	297	shrink much more than others, only that does not adapt
hog1		2	181127	3761	0.24	307	adapts? Smaller cells than two others
		3	190122	3840	0.06	143	adapts?
yVW407	0.1M	1	190312	3896	0.15	416	identical, has more cells
sko1		2	190314	3914	0.15	384	identical
	0.2M	1	181120	3742	0.36	529	identical, sligthly slower than two others
		2	181127	3764	0.26	222	identical

		3	181129	3773	0.35	246	slightls different PP7 kinetics, led issue?
	0.3M	1	190312	3895	0.15	261	similar to rep2
		2	190314	3914	0.15	215	similar except end of response andstartime, shrinks more
		3	190503	4005	0.17		lot of dead cells + weird led + less respcells
yVW408	0.2M	1	181204	3788	0.14	94	identical kinetics, take highest cell number
spt3		2	181207	3795	0.14	187	identical kinetics, adapts faster because smaller cells
yVW409	0.2M	1	181120	3746	0.06	197	response different
gcn5		2	181204	3789	0.04	88	shrinks a lot more, delayed, response lower
		3	190124	3852	0.06	148	looks like rep1
yVW410	0.2M	1	181120	3747	0.03	124	slightly lower than two others, cells bigger
arp8		2	181127	3766	0.05	82	similar to rep3
		3	190124	3853	0.03	95	similar to rep2, bact contaminated so take rep2
yVW411	0.2M	1	181213	3821	0.04	70	very similar, more cells
asf1		2	190122	3846	0.03	38	very similar, not enough cells, cells a bit bigger
yVW412	SD	1	181127	3760	0.38		identical
set1		2	181129	3777	0.08	166	identical
	0.1M	1	181120	3744	0.40	196	slightly faster but lower than rep2 and rep3, OD very high
		2	181129	3779	0.12	188	slightly bigger cells that shrink more, same height as rep3 and rep1
		3	181207	3792	0.05	149	slightly delayed, similar to rep2
		4	190122	3845	0.21	121	lower than three others
	0.15M	1	181129	3778	0.09	218	identical, weird GFP bleaching curve
		2	181204	3785	0.12	94	identical
		3	181207	3793	0.05	202	overlaps with 0.2M data
	0.2M	1	181120	3741	0.14	253	identical
		2	181129	3780	0.14	224	identical
		3	181204	3784	0.11	149	identical
yVW413	0.2M	1	181204	3787	0.13	259	identical
sus1		2	181207	3800	0.20	129	identical
yVW414	0.2M	1	181204	3786	0.12	489	identical
mlp1		2	181207	3794	0.05	207	identical
yVW416	SD	1	181207	3799	0.27	133	slightly different than two others
htz1		2	190118	3838	0.16	141	similar, more cells
		3	190122	3841	0.05	110	similar
	0.1M	1	181207	3798	0.27	133	quite variable population trace at 0.1M, but similar PP7 peaks/respcells
		2	190118	3836	0.12	171	shrinks more than two others
		3	190122	3845	0.08	108	much lower response, less respcells
	0.15M	1	181207	3797	0.16	197	very similar
		2	190118	3837	0.16	201	very similar

		3	190122	3844	0.05	188	shrinks more than two others
	0.2M	1	181207	3796	0.16	175	similar, more cells
		2	190118	3835	0.10	81	not enough cells
		3	190122	3842	0.08	112	slightly delayed compare to two others

Supplementary Table 8: pHOG PP7 strains

Strain	[NaCl]	Replicate n°	TL date	YQ n°	OD600	Cell number	
yVW428	SD	1	181211	3814	0.08	89	identical
pCTT1		2	181213	3820	0.18	120	identical
wt	0.1M	1	190118	3830	0.17	127	cells bigger and faster response
		2	190124	3851	0.22	428	similar to rep3 but higher
		3	190222	3870	0.11	327	similar to rep2 but lower
		4	190314	3917	0.20	324	cell bigger, responds more than all the others
	0.2M	1	181211	3813	0.08	85	shrinks much more, cells much bigger
		2	181213	3819	0.18	206	faster than two others but similar to rep3
		3	190118	3827	0.07	140	similar to rep2
yVW429	SD	1	190222	3873	0.23	291	
pHSP12		2	190312	3903	0.17	193	has less basal level than others TL
wt	0.1M	1	190314	3913	0.21	306	faster and shorter, more respcells
		2	190326	3933	0.09	215	delayed and longer, called yVW426 in export by mistake, bigger cells
		3	190405	3967	0.11	217	faster to two others but similar to rep1, way more basal level
		4	190412	3977	0.13	320	similar to rep2
	0.2M	1	190222	3868	0.16	186	has less basal level than others TL, led oscillating
		2	190226	3875	0.09	96	similar to rep3, but not so many cells, rep4?
		3	190312			192	responds less than two others
		4	190329		0.12	283	similar to rep2 but way more cells
		5	190402	3962	0.23	311	slightly delayed
yVW430	SD	1	181127	3755	0.15	248	identical, more cells
pGRE2		2	181129	3767	0.09	178	identical
wt	0.1M	1	190118	3826	0.08	201	identical
		2	190222	3867	0.10	309	identical
	0.2M	1	181127	3754	0.10	289	identical
		2	181129	3768	0.09	197	cells bigger, shrinks more, response different
		3	181204	3783	0.12	233	identical
yVW431	SD	1	181127	3753	0.16	114	identical
pALD3		2	181129	3769	0.13	183	identical, highest cell number
wt	0.1	1	190118	3829	0.13	218	lower than two others
		2	190124	3850	0.18	334	similar to rep3
		3	190222	3869	0.10	258	similar to rep2
	0.2M	1	181127	3752	0.13	171	identical
		2	181129	3770	0.16	158	slightly faster and shorter
		3	181204	3782	0.09	51	identical, not enough cells
yVW432	SD	1	181120	3743	0.13	214	identical
pGPD1		2	181127	3759	0.31	196	identical
wt	0.1M	1	190118	3828	0.11	229	identical to rep3

		2	190222	3866	0.11	269	30% higher and bit delayed
		3	190314	3916	0.34	402	identical to rep1, a bit faster
	0.2M	1	181120	3736	0.13	259	delayed
		2	181127	3758	0.24	335	identical
		3	181129	3771	0.20	281	identical

Supplementary Table 9: pHOG PP7 mutant strains

yVW458	0.2M	1	190226	3882	0.22	326	
y432 set1d		2	190322	3924	0.08	353	higher and longer response than rep1
		3	190329	3955	0.09	137	similar to rep1
yVW456	0.2M	1	190226	3883	0.22	132	identical, cells a bit bigger
y428 set1d		2	190314	3918	0.20	162	identical, more cells
		3	190503	4008	0.32	198	identical, the most cells
yVW457 yVW430 set1d	0.2M	1	190503	4007	0.37	366	led oscillating, results like wt
		2					
yVW459 yVW431 set1	0.2M	1					
		2					
yVW460 yVW429 set1d	0.2M	1	190405	3974	0.11	376	similar
		2	190503	4006	0.27	362	similar, less basal, shrink less nice
yVW471	0.2M	1	190322	3920	0.06	334	similar
yVW432 hot1d		2	190326	3935	0.12	247	similar slightly faster and bigger cells
		3	190503	4003	0.13		delayed and lower, more basal pos cells
yVW472 yVW432 sko1d	0.2M	1	190322	3925	0.06		led power 100% need to image again
		2	190326	3934	0.16	235	30% higer and longer response, cells bigger
		3	190405	3968	0.10	360	30% lower and shorter response
		4	190412	3981	0.13	310	identical to rep 3, except way more basal cells
yVW486 yVW430 sko1d	0.2M	1	190620	4058	0.18	383	identical
		2	190627	4078	0.07	292	identical
	Sdfull	1	190627	4077	0.07		

Supplementary Table 10: pHOG PP7 depletion control strains

yVW476	0.2M	1	190405	3975	0.41	505	identical, more cells
y432 pTEF PP7	clone 1	2	190412	3983	0.19	288	identical
		3	190503	4001	0.06	198	higher than two others, less basal
yVW477	0.2M	1	190405	3976	0.38	440	
y429 pTEF PP7	clone 1	2	190412	3979	0.14	378	bleaches more than others and more basal
		3	190503	4002	0.06	280	similar to rep2

Annex 5: *pSTL1* truncations

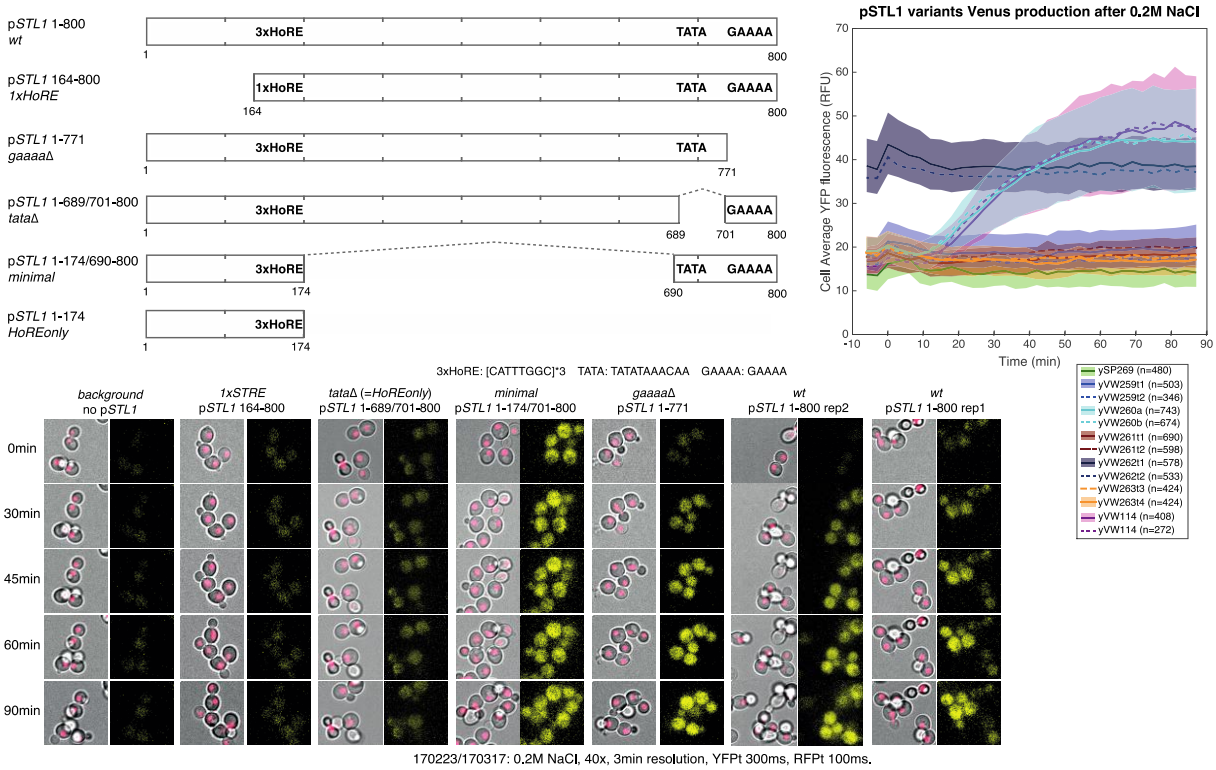


Figure 23: Dissecting p*STL1* bursty behavior through sequence truncations. (A) From the original -800 bp to TSS cloned in all p*STL1* reporters, we generated truncation variants based on documented Hot1 binding sites (Bai et al 2015). Constructs are listed in Annex 3.

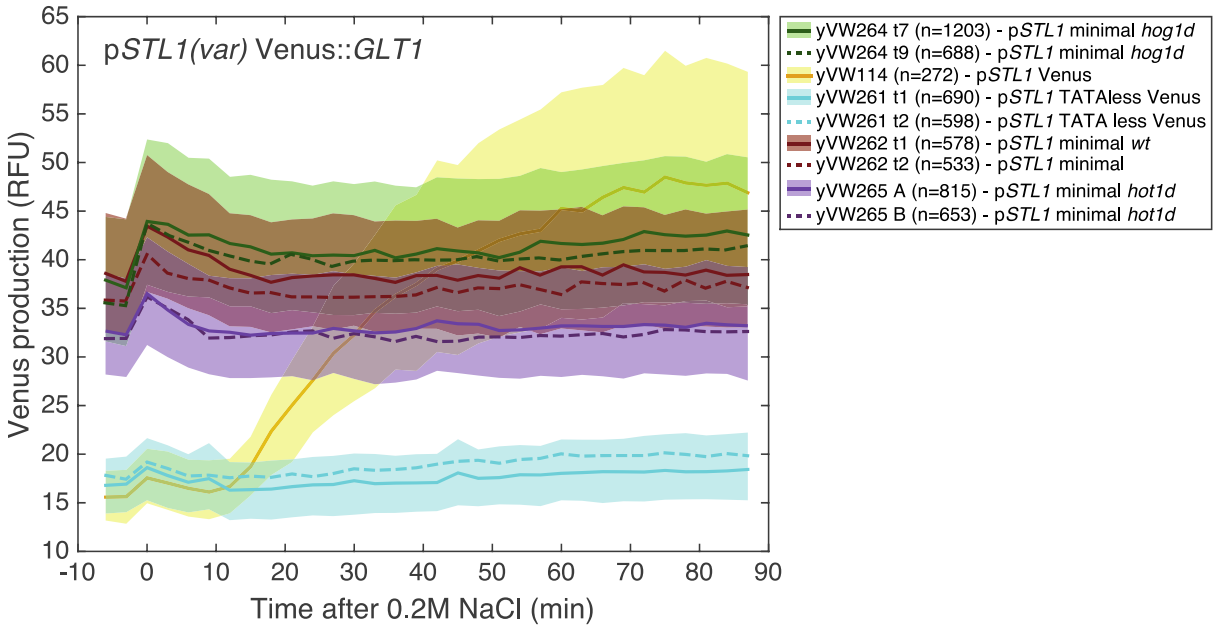


Figure 24: Gene deletions time-lapse experiment of the p*STL1* minimal truncation variant with constitutive expression. Deletion of Hot1p and Hog1p were performed in the constitutive pSTL1 promoter variant to understand the source of endogenous expression.