Coordination and integration of signaling and resources allocation in the yeast stress response

By

Yi-Hsuan Elisha Ho

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Cellular and Molecular Biology)

at the

UNIVERSITY OF WISCONSIN-MADISON

2015

Date of final oral examination: 7/20/2015

The dissertation is approved by the following members of the Final Oral Committee:

Audrey Gasch, Associate Professor, Genetics
Aseem Ansari, Professor, Biochemistry
David Eide, Professor, Nutritional Science
William Sugden, Professor, Oncology
Michael Thomas, Professor, Bacteriology

Acknowledgements

I would like to thank everyone who directly or indirectly helped and supported me over the past eight years.

I would particularly like to express my gratitude to my mentor and thesis advisor, Dr. Audrey Gasch, for the opportunity to join her lab, work on incredibly interesting science, and for her continuous support. I would also like to thank the members of the Gasch lab, James, Katie, Kevin, Maria, Matt and Nikolay, for providing a friendly and supportive work environment.

I would especially like to thank my committee members, Dr. Bill Sugden for his generosity and support during my lab transition, Dr. David Eide for extensive expertise and unwavering support, Dr. Aseem Ansari for fruitful collaborations and helpful comments, and Dr. Michael Thomas for kindly joining my committee for the thesis defense.

My project would not have been successful without our wonderful collaborators. I would like to thank Dr. Josh Coon and his group for proteomic data, Dr. Mark Craven and Dr. Debbie Chasman for computational modeling, and Dr. Aseem Ansari's lab for help with the CTD aspect of my project.

Table of Contents

Chapter 1

Introduction

Free–living microbes living in nature must adjust their internal system precisely and agilely in response to sudden and drastic environmental changes. Failure to properly respond to stress (due to mutations or misregulation of upstream signaling pathways) may cause slow growth, disease, or cell death. Defense responses include orchestrated changes to transcription, translation, and post-translational protein modifications that mediate changes in protein abundance, localization, and function. These responses are often coordinated with delay or arrest of cell-cycle progression. Although much is known about the upstream signaling pathways that control individual physiological responses to stress, how signaling is integrated into a single cellular network that coordinates a multi-faceted response is only beginning to emerge.

Part of this integration influences how resources are used in the cell. Microbial cells must allocate their limited resources to balance conflicting physiological demands to achieve optimal growth in a given environment. A key consumer of cellular resources is translation, since production of ribosomes and translation itself require substantial amounts of energy. Mounting a stress response usually comes at a cost, since it takes cellular energy to dramatically alter the cell. This usually comes at the expense of growth. For example, extremely stress resistant clones are slow-growing (ELLIOTT and FUTCHER 1993; LU *et al.* 2009), whereas the cells most sensitive to environmental insults are rapidly growing (ZAKRZEWSKA *et al.* 2011). Therefore, understanding how cells allocate cellular resources when balancing stress defense and growth is important to understand how cells achieve optimal growth and adapt to environmental change. In actively growing microbial cells, a significant amount of cellular resources are devoted to maintain high translational capacity to support division, in part by promoting ribosome biogenesis. Under optimal conditions, yeast cells produce ~2000 ribosomes per minute (WARNER 1999). Such high demand for ribosomes requires active transcription of the rDNA locus and nearly 60% of cellular transcription is devoted to rRNA production by RNA

polymerase (Pol) I. Furthermore, over half of the activity of Pol II transcription occurs at genes encoding ribosomal proteins (RPs), while Pol III is dedicated to transcribing 5S rRNA and tRNAs (RUDRA and WARNER 2004; WARNER 1999). Additionally, a significant fraction of Pol II is also dedicated to transcribing genes encoding ribosome assembly factors (commonly called the ribosome biogenesis - RiBi regulon) (JORGENSEN *et al.* 2004). In addition to consuming transcriptional capacity, transcripts of the RP and RiBi genes are also highly translated and thus monopolize a large fraction of the cell's ribosomes (90% of which are actively involved in translation during rapid growth) (ARAVA *et al.* 2003; VON DER HAAR 2008). Thus, a significant portion of resources is used to support ribosome biogenesis and translational capacity in cells actively growing under optimal conditions.

The Environmental Stress Response: a common yeast response to diverse stresses.

Upon a sudden shift to stressful conditions, cells must rapidly adjust resource allocation to mount cellular defense strategies (Fig 1). Much has been gleaned from studying transcriptome changes in response to stress. Stressed yeast cells activate condition-specific transcript changes at specialized genes that specifically address the particular stress condition. Concurrently, cells activate the environmental stress response (ESR) (CAUSTON *et al.* 2001; GASCH *et al.* 2000), a common response to diverse types of stress. The ESR includes ~300 induced ESR (iESR) genes that are broadly involved in stress defense and ~600 repressed (rESR) genes that encode RPs and RiBi factors (GASCH 2002; GASCH *et al.* 2000). The expression patterns of the iESR and rESR genes are strikingly anti-correlated across diverse types of stress, indicating that they are likely regulated by the same upstream signaling systems (GASCH 2002).

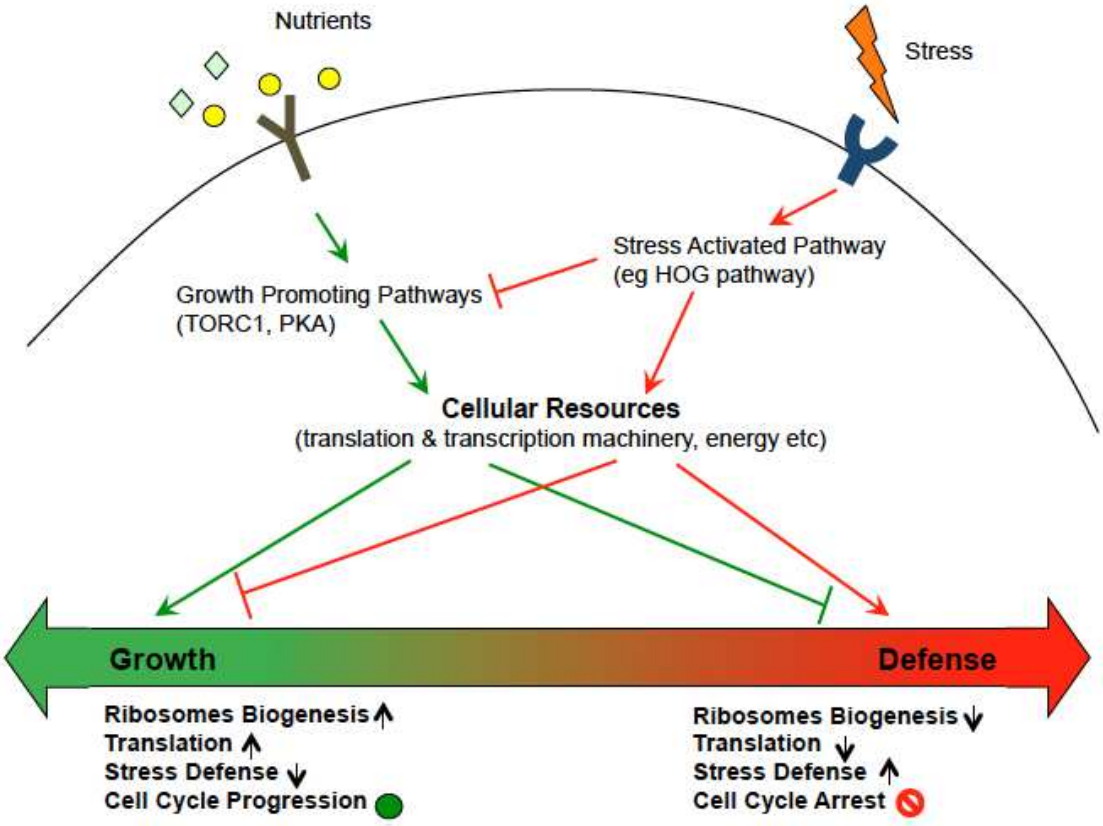


Figure 1. Growth and stress defense compete for limited resources in the cell. An overview schematic of the signaling processes responding to growth promoting signals (green arrows) versus stress signals (red arrows). In nutrient replete conditions, cells activate growth promoting pathways to allocate cellular resources for growth (green). Upon stress, stress activated pathways are activated and re-allocate limited resources for stress defense (red).

Function of the induced-ESR module

Several lines of evidence suggest the ESR expression changes represent a stress defense strategy. iESR genes are broadly involved in diverse processes like carbohydrate metabolism, oxidative stress defense, protein folding and degradation, trehalose biosynthesis and glycogen biosynthesis (GASCH 2002). Nonetheless, there are still about 50% of iESR genes that are functionally unannotated. Originally, transcript changes of the iESR genes were postulated to defend cells against the initial stress treatment, given the functions of the encoded proteins. But surprisingly, the genes whose expression changes in response to acute stress do not correlate well with genes required to survive that stress (GIAEVER *et al.* 2002). Furthermore, Berry *et al* showed that neither the transcript changes nor the translation of new proteins is required to survive a single dose of acute stress (BERRY and GASCH 2008). This indicates that activation of the ESR – and of stress-activated transcript changes in general – is not required to survive the initial stress treatment, at least in yeast (BERRY and GASCH 2008; BERRY *et al.* 2011; GIAEVER *et al.* 2002; MITCHELL *et al.* 2009; WESTFALL *et al.* 2008) . However, a later study found that transcript changes are critical for surviving subsequent stressful insults, through a phenomenon known as acquired stress resistance, where cells exposed to a mild dose of one stress can subsequently survive what would otherwise be a lethal dose of the same or another stress (BERRY and GASCH 2008). The acquired stress resistance is dependent on protein synthesis during mild-stress treatment and requires the ”general-stress” transcription factors Msn2 and Msn4, that regulate induction of many ESR genes (BERRY and GASCH 2008) (see below). In fact, further study showed that there is high correlation between the induction of iESR transcripts and increased abundance of the encoded proteins, which most likely serve to defend against stress (BERRY and GASCH 2008; LEE *et al.* 2011; VOGEL *et al.* 2011). Thus, instead of protecting cells against the original stress insults, iESR plays a preparative role for impending stress. This agrees with the well-studied phenomenon “cross-protection”, where cells develop

better stress tolerance not only to the original stressor but also other stresses given after exposure to a mild dose of stress

Function of the repressed-ESR gene response & its relation to growth and cell cycle progression

Unlike iESR genes, the role of the rESR transcript repression has remained somewhat elusive. Previous observations found that ESR activation is correlated with reduced growth rate in nutrient restricted chemostats (BRAUER *et al.* 2008; CASTRILLO *et al.* 2007; REGENBERG *et al.* 2006). For example, Brauer *et al* observed expression of more than a quarter of all yeast genes is linearly correlated with growth rate, independent of the identity of the different limited nutrients (BRAUER *et al.* 2008). The positively correlated genes mainly encode translational functions. This leads to the suggestion that the stress-dependent decrease in rESR expression is a secondary effect of growth arrest upon stress, instead of a direct stress response (BRAUER *et al.* 2008). This model is an extension of the bacterial model, where production of ribosomes is directly related to growth rate (NEIDHARDT and MAGASANIK 1960). However, hints from a more detailed analysis of gene expression upon stress in chemostats already suggest that this model is too simple to explain the purpose of rESR repression. Lu *et al* performed experiments in which cells growing at different growth rates in chemostats were treated with a short heat shock (LU *et al.* 2009). They observed that the expression of about half of stress-responsive genes are both growth rate and stress-dependent, while the other half are only stress dependent (LU *et al.* 2009). This suggests that the observed repression of rESR transcripts is not solely a byproduct of slow growth. Besides the above observations, one thing to point out is that limiting nutrient availability to control cellular growth rate in those chemostats experiments may itself serve as a stress treatment, thus the induced ESR activation is not due to growth reduction but because of inflicted stress. Therefore, the above experiments are insufficient to distinguish if ESR activation is caused by external stress stimuli or internal slow cell growth.

A recent study by O'Duibhir et al. instead proposed that ESR activation may be merely a byproduct of prolonged G1 progression in slow-growing cells (O'DUIBHIR *et al.* 2014). They extended the correlation of rESR and slow growth to cell cycle populations. At the level of cell cycle progression in *S. cerevisiae*, slower growing cells spend a larger fraction of their cell cycle in G1 phase (while the duration of the other cell cycle phases is generally the same (JOHNSTON *et al.* 1980). In fact, previous chemostat studies have observed an increased population of unbudded (G1) cells in slow glowing cultures (LU *et al.* 2009). O'Duibhir *et al* analyzed a transcriptome dataset measuring the transcriptome in each of 1,484 functionally unrelated deletion strains growing in batch culture (O'DUIBHIR *et al.* 2014). They observed that slow growing strains display a common transcriptome signature, which is similar to the ESR. They further extend this link to cell cycle population and claim that the common expression signature in slow growing strains, namely the ESR, mainly reflects an increase in the G1 population in those cultures (O'DUIBHIR *et al.* 2014). As with previously described studies, this work proposed that the ESR is not a stress response but rather an indirect signature of G1 phase (O'DUIBHIR *et al.* 2014). However, this work lacks evidence to directly indicate which factor is causal for the correlations observed here, and therefore several key questions are left unanswered.

Our lab put forth an alternate model for the function of rESR repression during stress. A previous study from our lab showed that stress-dependent repression of RP and RiBi genes is not required for stress-dependent translational arrest or growth reduction, at least during salt stress (LEE *et al.* 2011). Furthermore, reduced abundance of rESR transcripts does not produce a drop in the corresponding proteins, as evidenced by mass-spectrometric proteome measurements (LEE *et al.* 2011). This led to a separate hypothesis that stress-dependent rESR transcript reduction serves to release engaged ribosomes, thereby redirecting limited translational capacity to newly made stress-defense mRNAs (LEE *et al.* 2011). Consistent with this model, failure to repress rESR transcripts caused those transcripts to remain engaged with

ribosomes after stress (LEE *et al.* 2011). A more recent work from our lab in collaboration with the Ansari lab also implies competition for polymerase by rESR and iESR genes, leading to their anti-correlated expression. As outlined in Chapter 2, we showed that regulation of the C-terminal domain (CTD) of the RNA Pol II subunit Rpb1 plays a crucial role in the iESR/rESR transcriptional balance, by reallocating RNA pol II from rESR to iESR genes (CHASMAN *et al.* 2014). Thus, we proposed that that reduced transcription of rESR genes is required to reallocate RNA Pol II to iESR and other defense genes (CHASMAN *et al.* 2014). The anti-correlation between rESR and iESR gene expression during stress is thereby at least partly explained by competition of these gene modules for limited transcription and translation machinery.

Signaling Pathways Regulating Stress Responses

Cells activate condition-specific signaling pathways upon specific stress treatments. For example, the MAP kinase Hog1 (orthologous to human kinase p38) is activated by osmotic shock as well as related stresses to coordinate myriad responses to cope with external environment change (see below). Alternate signaling pathways are specifically activated by other stresses, for example PKC in the case of cell wall stress (HEINISCH *et al.* 1999), ATM/ATR in response to DNA damage (ABRAHAM 2001), and the AMPK ortholog Snf1 during starvation (CONRAD *et al.* 2014).

Although specific signaling pathways are activated upon specific stimuli, the “general stress” transcription factors, Msn2/Msn4, can be regulated via different upstream pathways upon different stimuli, such as oxidative stress, osmotic stress, heat shock, and low pH (MARTÍNEZ-PASTOR *et al.* 1996). Msn2 and Msn4 share high sequence similarity and are considered functionally redundant based on genetic studies (ESTRUCH and CARLSON 1993). Both Msn2 and Msn4 contain a Zinc-finger domain and regulate ~200 genes by binding to the stress responsive

element (STRE) at the promoter region (MARTÍNEZ-PASTOR *et al.* 1996). Upon stress, Msn2/4 localize to the nucleus via an unclear mechanism and induce downstream gene transcription (GÖRNER *et al.* 1998). This reveals a common stress defense mechanism by activating "general stress" transcription factors. Besides the activation of particular signaling pathways and downstream transcription factors, other signaling pathways must be repressed, for example repression of growth-promoting pathways is another general but important means of stress response (see below), although the repression mechanism upon stress still remains unclear. Thus, further understanding the coordination and integration of these pathways will help delineate the cell's general stress defense strategy. Below I summarize several regulatory systems that are particularly relevant to my own work.

HOG pathway: response to hyperosmotic stress

Much of this thesis focuses on osmotic/salt shock, for which much is already known in terms of upstream signaling. The HOG (high osmolarity glycerol) pathway is a relatively well-studied signaling pathway. It is the major signaling pathway to cope with hyperosmolarity in yeast. The Hog1 MAPK cascade is highly conserved even in higher eukaryotes, including humans. The stress-responsive p38 MAPK (human ortholog) can even rescue the osmosensitivity of *hog1Δ* mutations in response to hyperosmotic challenge (HAN *et al.* 1994). The HOG pathway regulates a range of downstream adaptive responses including induction and repression of ~2000 transcripts (including iESR and rESR genes) (CHASMAN *et al.* 2014; O'ROURKE and HERSKOWITZ 2004), reallocation of RNA Pol II distribution (CHASMAN *et al.* 2014; COOK and O'SHEA 2012; NADAL-RIBELLES *et al.* 2012), alterations in translational capacity (MELAMED *et al.* 2008; UESONO and TOH-E 2002), altered metabolism including osmolyte production (ALBERTYN *et al.* 1994; PROFT and STRUHL 2004), and transient cell-cycle arrest (BELLÍ *et al.* 2001; CLOTET *et al.* 2006; YAAKOV *et al.* 2009).

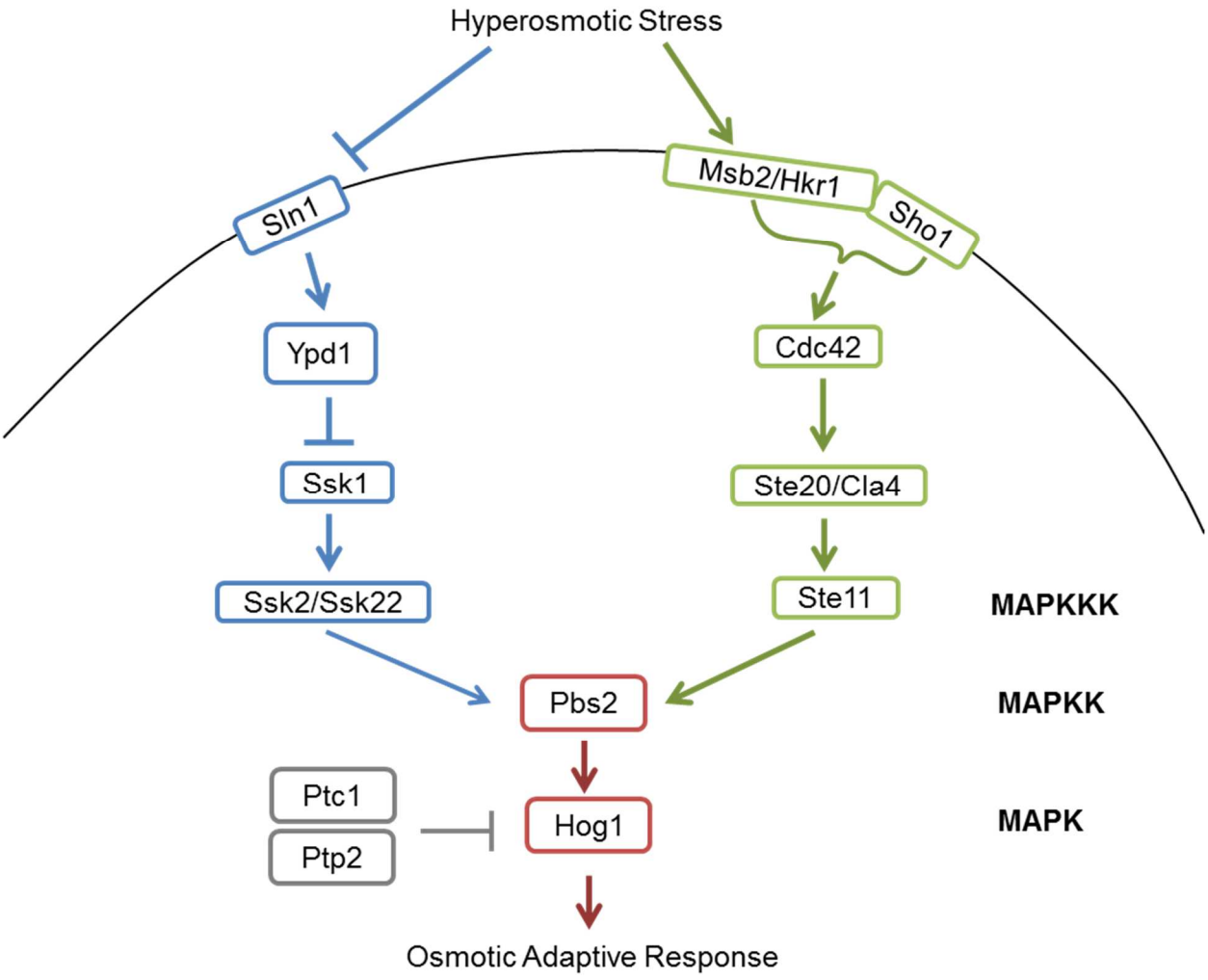


Figure 2. The HOG signaling pathway in yeast. Two functionally redundant upstream signaling branches activate the HOG pathway: the Sln1 branch (blue) and the Sho1 branch (green). The two branches converge on the MAPKK, Pbs2, which further activates the downstream MAPK, Hog1. Inactivation of Hog1 is regulated by two phosphatases, Ptc1 and Ptp2.

The HOG pathway has two functionally redundant upstream activating branches, via either Sln1 or Sho1 regulators (Fig 2). The Sln1 branch activates the redundant Ssk2 and Ssk22 MAPKKKs, while the Sho1 branch activates the Ste11 MAPKKK. The Ssk2/Ssk2 and Ste11 branches converge on the Pbs2 MAPKK, which lies upstream of Hog1. Although either Sln1 or Sho1 branch can almost fully activate the HOG pathway on high osmolality, the Sln1 branch possess higher osmolality sensitivity than Sho1 branch and thus is activated by lower doses of osmotic stress (O'ROURKE and HERSKOWITZ 2004). The Sln1 branch of the HOG pathway is a variation of the so-called two–component regulatory system, which is ubiquitous in prokaryotes, plants and fungi (STOCK *et al.* 2000). The prototypical two-component system is comprised of a His-Asp phosphorelay between two proteins that the phosphate group is transfered from a His residue of a sensor histidine kinase protein to an Asp residue of a receiver protein (STOCK *et al.* 2000). Sln1 seems to respond to changes in turgor pressure and remains activated under normal osmotic conditions (TAMÁS *et al.* 2000). Sln1 activates the Sln1-Ypd1-Ssk1 multistep phosphorelay through a His-Asp-His-Asp multi-step (POSAS *et al.* 1996). Constitutively phosphorylated Ssk1 ("Ssk1~p") cannot interact with downstream Ssk2/Ssk22 (HORIE *et al.* 2008). However, under hyperosmotic conditions, unphosphorylated Ssk1 ("Ssk1-OH") will accumulate, since cells shrink and Sln1 is inactivated (HORIE *et al.* 2008). Ssk1-OH binds and activates Ssk2/Ssk22 which further activates the downstream Hog1 MAPK cascade (HORIE *et al.* 2008). Unlike the Sln1 branch with the well-studied two-component system, the Sho1 branch is only vaguely defined. The signaling response is initiated by the putative osmosensors Msb2 and Hrk1 (TATEBAYASHI *et al.* 2007). An interaction between the Msb2/Hkr1 osmosensors and the Sho1 co-osmosensor seems to induce association of the small G-protein Cdc42 and Ste20/Cla4 through an undefined mechanism (LAMSON *et al.* 2002). The interaction with Cdc42 activates Ste20/Cla4 which results in phosphorylating Ste11 MAPKKK and further activating downstream HOG pathway (DROGEN *et al.* 2000; RAITT *et al.* 2000). The inactivation of Hog1 kinase is primarily regulated by two phosphatases, Ptc1 and Ptp2. Ptc1 dephosphorylates

Hog1 at Thr-174 and its interaction with Hog1 is conferred with the adaptor protein Nbp2, which binds to Pbs2 and Ptc1 (MAPES and OTA 2004). Thus, Ptc1 is indirectly recruited to Hog1 by the Nbp2-Pbs2 complex. Ptp2 is found in the nucleus and dephosphorylates Hog1 at Tyr-176 (MATTISON and OTA 2000). Because dephosphorylation of either Thr-174 or Tyr-176 inactivates Hog1, the *ptc1Δptp2Δ* double-mutant strain is lethal because of Hog1 hyperactivation (MAEDA *et al.* 1993).

Once activated, a substantial fraction of Hog1 protein is transported into the nucleus where it regulates transcription and the cell cycle (see below), although there are also Hog1 targets in the cytoplasm (DIHAZI *et al.* 2004; REISER *et al.* 1999). Surprisingly, a membrane-tethered Hog1 construct is still able to suppress the osmosensitivity of *hog1Δ* mutant. This indicates the cytoplasmic events caused by Hog1, such as glycerol production by direct modulation of metabolic enzymes, is sufficient to maintain the osmotic balance under those conditions (WESTFALL *et al.* 2008). Nevertheless, Hog1 still plays a key role in regulating the profound gene expression changes upon osmostress, via several mechanisms. ~2000 genes (including iESR and rESR genes) require Hog1 for full induction and repression in response to osmo/salt stress (CHASMAN *et al.* 2014). Hog1 can control gene expression by regulating transcription factors, *e.g*. Hot1, Smp1, Msn2, Msn4 and Sko1 (ALEPUZ *et al.* 2001; DE NADAL and POSAS 2010; PROFT *et al.* 2001). Besides the well-known mechanism of regulating activity of transcription factors by phosphorylation (*e.g.* Smp1 and Sko1) (DE NADAL *et al.* 2003; PROFT *et al.* 2001), Hog1 was found to associate with chromatin by physical interactions with transcription factors: recruitment of Hog1 to the *CTT1* promoter (cytosolic catalase T) requires the transcription factors Msn2 and Msn4, whereas recruitment of Hog1 to *STL1* promoter (glycerol proton symporter) depends on the transcription activator Hot1 (ALEPUZ *et al.* 2001). Hog1 also modulates histone deacetylation of induced genes via an interaction with Rpd3, a histone deacetylase regulating expression of a large number of genes. Rpd3, interacting with Hog1, is

recruited to specific promoters leading to histone deaceytylation, RNA Pol II entry, and induction of gene expression (DE NADAL *et al.* 2004). Besides regulating transcription initiation, Hog1 is also present on the coding regions of these genes and travels with elongating RNA pol II, suggesting Hog1 directly affects the process of elongation (POKHOLOK *et al.* 2006; PROFT *et al.* 2006). Hog1 is also postulated to regulate transcription by mediating assembly of the pre-initiation complex ('PIC') at stress-responsive promoters, remodeling chromatin, and controlling of mRNA processing and transport (DE NADAL and POSAS 2010). Overall, Hog1 modulates profound gene expression via multi-mechanisms.

Hog1 is also involved in arresting cell-cycle progression at different phases upon stress. Hog1-mediated G1 arrest is partially attributed to down-regulation of gene expression of the G1 cyclins Cln1, 2 and the S-cyclin Clb5 (BELLÍ *et al.* 2001). In addition, Hog1 delays G1/S transition through direct phosphorylation of the inhibitor of the cyclin-dependent kinase Cdc28, Sic1, to repress its ubiquitination for degradation (ESCOTÉ *et al.* 2004). Hog1 is also proposed to trigger S-phase delay, although not much is known. Hog1-dependent phosphorylation of the Dpb2 subunit of DNA polymerase has been proposed to be involved in S-phase delay (YAAKOV *et al.* 2009). G2 phase arrest by Hog1 is proposed to be accomplished by accumulation of the inhibitor Swe1 kinase, which phosphorylates and inhibits Clb2/Cdc28 activity (CLOTET *et al.* 2006). Hog1 interacts and phosphorylates Hsl1, a septin-dependent protein kinase binding to Hsl7 at the septin ring on the neck between mother and daughter cells (CLOTET *et al.* 2006). The phosphorylation of Hsl1 causes Hsl7 to delocalize from the neck, which prevents Swe1 degradation initiated by interacting with Hsl7 (CLOTET *et al.* 2006). Thus, Swe1 is accumulated and prolongs the G2/M transition. In all, Hog1 mediates cell cycle arrest at different phases in response to stress to ensure cells to rebuild homeostasis and restore before further replication. Taken together, the HOG pathway plays a pivotal role in response to osmotic stress. The multi-level regulatory mechanisms in various physiological processes not only showcase the

complexity of cells' response upon stress but also provide a great model on how cells may adapt to environmental changes.

Regulation of the growth-versus-defense decision in yeast

While the activation of defense-promoting pathways plays a key role during stress, the suppression of growth-promoting pathways is also important. The two central growth promoting signaling pathways are the target of rapamycin (TOR) and the RAS/cAMP/Protein Kinase A (PKA) signaling pathways, which promote growth and synthesis in response to nitrogen and carbon availability, respectively (BROACH 2012; NEUMAN-SILBERBERG *et al.* 1995; WANG *et al.* 2004; XIAO and GROVE 2009). Both pathways are largely conserved from yeast to mammals and activate high rESR gene expression and mediate low iESR expression in the absence of stress (CARDENAS *et al.* 1999; POWERS and WALTER 1999). Yeast encodes two TOR proteins: Tor1 and Tor2, either of which may become part of the TOR complex I (TORC1) (LOEWITH *et al.* 2002). TORC1 controls cell growth via a rapamycin–sensitive pathway and is located at the membrane of the vacuole, a major yeast nutrient reservoir (BINDA *et al.* 2009). The PKA signaling pathway is the other major nutrient responsive pathway. PKA is modulated by cellular cAMP levels and is composed of a regulatory subunit, Bcy1, and three catalytic subunits: Tpk1, Tpk2, and Tpk3 (BROACH 2012). PKA is activated through the dissociation of Bcy1, which is triggered by the cAMP binding under nutrient replete condition (THEVELEIN and DE WINDE 1999). Cellular cAMP concentration is determined by the synthesis from ATP via adenylyl cyclase (Cyr1) and degradation to AMP by the two phosphodiesteraese, Pde1 and Pde2 (BROACH 2012). Adenylyl cyclase activity is activated by two GTP binding proteins, Ras1 and Ras2 (UNO *et al.* 1987). Despite intensive study, a detailed mechanism for how glucose affects Ras-GTP levels still remains elusive. Since both TORC1 and PKA pathways are the two major nutrient sensing pathways, cross-talk between TORC1 and PKA pathways is required to ensure cells effectively

utilize nutrients to achieve optimal growth (BROACH 2012; CHEN and POWERS 2006; RAMACHANDRAN and HERMAN 2011).

Both the TORC1 and PKA pathways are implicated in regulating ribosome biogenesis through transcriptional regulation of rDNA, the RP genes and the RiBi regulon. TORC1 directly binds the rDNA locus under optimal conditions (LI *et al.* 2006) and promotes rDNA transcription, by promoting activity of the rDNA transcription factor Rrn3 and, indirectly, by inhibiting the Pol III repressor Maf1 (CLAYPOOL *et al.* 2004; LEE *et al.* 2009; PHILIPPI *et al.* 2010). Maf1 is also directly regulated by PKA, which suppresses the inhibitory activity of Maf1 during rapid growth (MOIR *et al.* 2006; WILLIS and MOIR 2007).

Several Pol II-dependent transcription factors have been implicated in regulating ribosome biogenesis, and these lie downstream of TORC1 and PKA (Fig 3). Fhl1 may function either as a repressor or activator of RP gene expression, depending on its association with the corepressor Crf1 or the coactivator Ifh1 (MARTIN *et al.* 2004; WADE *et al.* 2004). In actively growing cells, Crf1 is kept in the cytoplasm, while Ifh1 binds RP gene promoters and activates gene expression. The localization of Crf1 is regulated by both the TORC1 and PKA pathways via Yak1-dependent phosphorylation. The kinase activity of Yak1 is negatively regulated by PKA, which is proposed to be regulated by Sch9, a kinase downstream of TORC1 (MARTIN *et al.* 2004; URBAN *et al.* 2007). While in nutrient deplete condition or when TORC1 is inhibited, repression of Yak1 is released. Activated Yak1 phosphorylates Crf1, which translocates to the nucleus and competes with Ifh1 for binding to Fhl1 at RP gene promoters, and thereby inhibits transcription of RP genes (MARTIN *et al.* 2004). Sfp1 is another transcription factor involved in regulating RP gene expression (MARION *et al.* 2004). Under optimal growth condition, Sfp1 is phosphorylated directly by TORC1 and localized to the nucleus, bound to the promoters of RP genes (MARION *et al.* 2004). In response to inhibition of TORC1 signaling, Sfp1 is released from

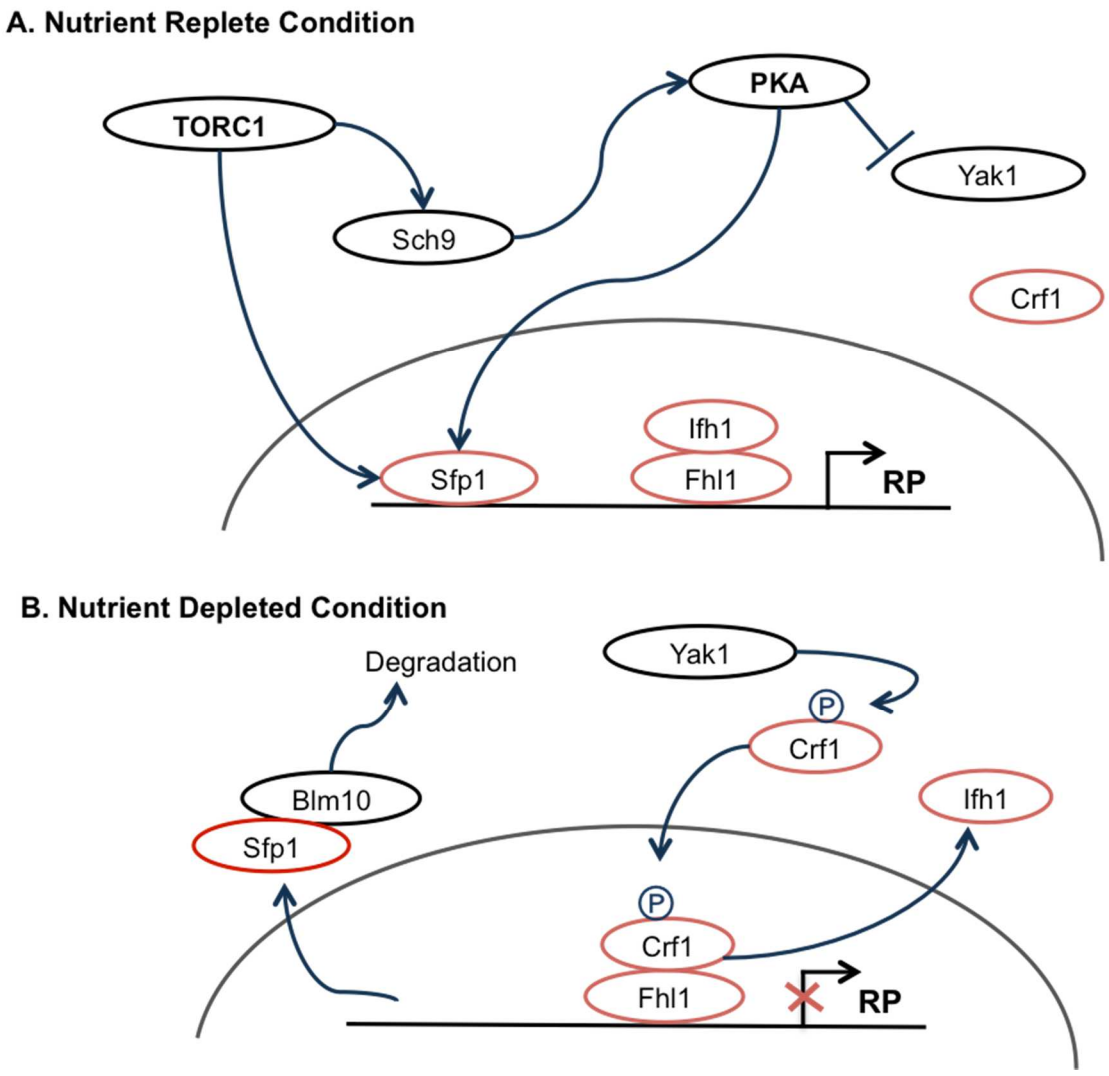


Figure 3. Transcriptional regulation of the RP genes by TORC1 and PKA pathways. TORC1 and PKA pathways regulate the expression of the RP genes by modulating the localization of key transcription factors (red). **A.** Under nutrient replete conditions, Sfp1 and Ifh-Fhl1 binds to the promoters of RP genes, while the transcription inhibitor Crf1 remains in the cytoplasm. **B.** In nutrient depleted conditions, phosphorylated Crf1 translocates into the nucleus and competes with Ifh1 for binding to the promoters of the RP genes. Meanwhile, Sfp1 is exported from the nucleus to the cytoplasm and undergoes proteasome-dependent degradation.

RP promoters and degraded via interacting with proteasome activator Blm10 (LOPEZ *et al.* 2011). Sfp1 is regulated by both TORC1 and PKA pathways, since Sfp1 remains localized to the nucleus in a strain constitutively expressing high PKA activity in spite of rapamycin treatment (which inhibits TORC1 activity) (MARION *et al.* 2004).

The RiBi regulon and RP genes show similar but not identical kinetics of regulation due to different promoter structures, suggesting they are coordinated but regulated by distinct factors. RiBi gene promoters contain two motifs, the RNA polymerases A and C (PAC) and/or ribosomal RNA-processing element (RRPE) motifs (HUGHES *et al.* 2000; JORGENSEN *et al.* 2004; WADE *et al.* 2006). Binding of the PAC element by the repressor Dot6 and its homolog Tod6 represses RiBi regulon expression in response to stress (LIPPMAN and BROACH 2009). In contrast, PKA and TOR pathways inactivate Dot6/Tod6 repressors in the absence of stress, thereby promoting high expression of RiBi genes (LIPPMAN and BROACH 2009). The TORC1 and PKA pathways are clearly interconnected with regard to RP regulation, even though the precise relationships in regulating ribosomal biogenesis are still controversial. Although some studies have suggested PKA as a downstream effector of the TORC1 pathway, possibly via the TORC1 downstream target Sch9 (MARTIN *et al.* 2004; SCHMELZLE *et al.* 2004; SOULARD *et al.* 2010), other evidence has indicated that PKA and TOR are two parallel pathways in activating RP gene expression and promoting cell growth (RAMACHANDRAN and HERMAN 2011; ZURITA-MARTINEZ and CARDENAS 2005).

Along with promoting biogenesis of ribosomes and translational machinery, PKA and TORC1 act in parallel to antagonize ESR activation, in part by suppressing Msn2 and Msn4, the 'general stress' transcription factors (BECK and HALL 1999; GÖRNER *et al.* 1998; SMITH *et al.* 1998). Inhibition of TORC1 with rapamycin induces Msn2/4 accumulation in the nucleus (SANTHANAM *et al.* 2004). Msn2/4 nuclear localization is inhibited by PKA-dependent phosphorylation

(GÖRNER *et al.* 1998; SMITH *et al.* 1998). Thus, under stressful conditions TORC1 and PKA pathways must be suppressed – although the mechanisms remain poorly understood – while stress-activated signaling pathways are mobilized. How activation of these stress-regulated pathways is coordinated with suppression of the TORC1 and PKA pathways remains unclear. A current study proposed that the HOG pathway may inhibit TORC1 activity by mediating Sch9 activity via an unknown mechanism (HUGHES HALLETT *et al.* 2014), hinting that activated stress signaling pathways may suppress the nutrient signaling pathways upon stress. Overall, coordination and integration of signaling pathways is likely to be key for precise allocation of limited resources and adaptation to a new environment.

Overview of the thesis

Despite intensive studies, many questions still remain unresolved in cell stress defense. How do cells coordinate and integrate these pathways into a single signaling network to modulate a multi-faceted physiological process? Does the coordination of signaling pathways play a pivotal role in allocating resources to defend against stress and restore cellular function and homeostasis? Is the anti-correlation between iESR and rESR caused by competition for resource allocation and can that explain the role of reduced expression of rESR genes?

In Chapter 2, I present an experimental and computational pipeline in budding yeast that integrates stress-activated transcriptome alterations, phospho-proteomic changes, and gene-fitness contributions with large-scale protein interaction data, to implicate a single network coordinating the response to salt stress (CHASMAN *et al.* 2014). The resulting network identified known and novel salt-regulated signaling proteins, uncovered previously unrecognized cross-connections between signaling pathways, and implicated important decision points in the growth versus defense decision. In Chapter 3, I discuss the relation between ESR activation, growth, and cell cycle progression via testing stress responses in cells arrested at different cell cycle

phases. The activation of the ESR upon stress in these arrested cells indicates that the ESR is not just a reflection of cell cycle population, but instead is a true response to stress. Finally, in Chapter 4, I address remaining challenges and potential future directions in this area of research.

References

ABRAHAM, R. T., 2001 Cell cycle checkpoint signaling through the ATM and ATR kinases. Genes Dev **15:** 2177-2196.

ALBERTYN, J., S. HOHMANN, J. M. THEVELEIN and B. A. PRIOR, 1994 GPD1, which encodes glycerol-3-phosphate dehydrogenase, is essential for growth under osmotic stress in Saccharomyces cerevisiae, and its expression is regulated by the high-osmolarity glycerol response pathway. Mol Cell Biol **14:** 4135-4144.

ALEPUZ, P. M., A. JOVANOVIC, V. REISER and G. AMMERER, 2001 Stress-induced map kinase Hog1 is part of transcription activation complexes. Mol Cell **7:** 767-777.

ARAVA, Y., Y. WANG, J. D. STOREY, C. L. LIU, P. O. BROWN *et al.*, 2003 Genome-wide analysis of mRNA translation profiles in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A **100:** 3889-3894.

BECK, T., and M. N. HALL, 1999 The TOR signalling pathway controls nuclear localization of nutrient-regulated transcription factors. Nature **402:** 689-692.

BELLÍ, G., E. GARÍ, M. ALDEA and E. HERRERO, 2001 Osmotic stress causes a G1 cell cycle delay and downregulation of Cln3/Cdc28 activity in Saccharomyces cerevisiae. Mol Microbiol **39:** 1022-1035.

BERRY, D. B., and A. P. GASCH, 2008 Stress-activated genomic expression changes serve a preparative role for impending stress in yeast. Mol Biol Cell **19:** 4580-4587.

BERRY, D. B., Q. GUAN, J. HOSE, S. HAROON, M. GEBBIA *et al.*, 2011 Multiple means to the same end: the genetic basis of acquired stress resistance in yeast. PLoS Genet **7:** e1002353.

BINDA, M., M. P. PÉLI-GULLI, G. BONFILS, N. PANCHAUD, J. URBAN *et al.*, 2009 The Vam6 GEF controls TORC1 by activating the EGO complex. Mol Cell **35:** 563-573.

BRAUER, M. J., C. HUTTENHOWER, E. M. AIROLDI, R. ROSENSTEIN, J. C. MATESE *et al.*, 2008 Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast. Mol Biol Cell **19:** 352-367.

BROACH, J. R., 2012 Nutritional control of growth and development in yeast. Genetics **192:** 73-105.

CARDENAS, M. E., N. S. CUTLER, M. C. LORENZ, C. J. DI COMO and J. HEITMAN, 1999 The TOR signaling cascade regulates gene expression in response to nutrients. Genes Dev **13:** 3271-3279.

CASTRILLO, J. I., L. A. ZEEF, D. C. HOYLE, N. ZHANG, A. HAYES *et al.*, 2007 Growth control of the eukaryote cell: a systems biology study in yeast. J Biol **6:** 4.

CAUSTON, H. C., B. REN, S. S. KOH, C. T. HARBISON, E. KANIN *et al.*, 2001 Remodeling of yeast genome expression in response to environmental changes. Mol Biol Cell **12:** 323-337.

CHASMAN, D., Y.-H. HO, D. B. BERRY, C. M. NEMEC, M. E. MACGILVRAY *et al.*, 2014 Pathway connectivity and signaling coordination in the yeast stress-activated signaling network. Mol Syst Biol **19:** 759.

CHEN, J. C., and T. POWERS, 2006 Coordinate regulation of multiple and distinct biosynthetic pathways by TOR and PKA kinases in S. cerevisiae. Curr Genet **49:** 281-293.

CLAYPOOL, J. A., S. L. FRENCH, K. JOHZUKA, K. ELIASON, L. VU *et al.*, 2004 Tor pathway regulates Rrn3p-dependent recruitment of yeast RNA polymerase I to the promoter but does not participate in alteration of the number of active genes. Mol Biol Cell **15:** 946-956.

CLOTET, J., X. ESCOTÉ, M. A. ADROVER, G. YAAKOV, E. GARÍ *et al.*, 2006 Phosphorylation of Hsl1 by Hog1 leads to a G2 arrest essential for cell survival at high osmolarity. EMBO J **25:** 2338-2346.

CONRAD, M., J. SCHOTHORST, H. N. KANKIPATI, G. VAN ZEEBROECK, M. RUBIO-TEXEIRA *et al.*, 2014 Nutrient sensing and signaling in the yeast Saccharomyces cerevisiae. FEMS Microbiol Rev **38:** 254-299.

COOK, K. E., and E. K. O'SHEA, 2012 Hog1 controls global reallocation of RNA Pol II upon osmotic shock in Saccharomyces cerevisiae. G3 (Bethesda) **2:** 1129-1136.

DE NADAL, E., L. CASADOMÉ and F. POSAS, 2003 Targeting the MEF2-like transcription factor Smp1 by the stress-activated Hog1 mitogen-activated protein kinase. Mol Cell Biol **23:** 229-237.

DE NADAL, E., and F. POSAS, 2010 Multilayered control of gene expression by stress-activated protein kinases. EMBO J **29:** 4-13.

DE NADAL, E., M. ZAPATER, P. M. ALEPUZ, L. SUMOY, G. MAS *et al.*, 2004 The MAPK Hog1 recruits Rpd3 histone deacetylase to activate osmoresponsive genes. Nature **427:** 370-374.

DIHAZI, H., R. KESSLER and K. ESCHRICH, 2004 High osmolarity glycerol (HOG) pathway-induced phosphorylation and activation of 6-phosphofructo-2-kinase are essential for glycerol accumulation and yeast cell proliferation under hyperosmotic stress. J Biol Chem **279:** 23961-23968.

DROGEN, F., S. M. O'ROURKE, V. M. STUCKE, M. JAQUENOUD, A. M. NEIMAN *et al.*, 2000 Phosphorylation of the MEKK Ste11p by the PAK-like kinase Ste20p is required for MAP kinase signaling in vivo. Curr Biol **10:** 630-639.

ELLIOTT, B., and B. FUTCHER, 1993 Stress resistance of yeast cells is largely independent of cell cycle phase. Yeast **9:** 33-42.

ESCOTÉ, X., M. ZAPATER, J. CLOTET and F. POSAS, 2004 Hog1 mediates cell-cycle arrest in G1 phase by the dual targeting of Sic1. Nat Cell Biol **6:** 997-1002.

ESTRUCH, F., and M. CARLSON, 1993 Two homologous zinc finger genes identified by multicopy suppression in a SNF1 protein kinase mutant of Saccharomyces cerevisiae. Mol Cell Biol **13:** 3872-3881.

GASCH, A. P., 2002 The Environmental Stress Response: a Common Yeast Response to Environmental Stresses. Springer-Verlag: Heidelberg.

GASCH, A. P., P. T. SPELLMAN, C. M. KAO, O. CARMEL-HAREL, M. B. EISEN *et al.*, 2000 Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell **11:** 4241-4257.

GIAEVER, G., A. M. CHU, L. NI, C. CONNELLY, L. RILES *et al.*, 2002 Functional profiling of the Saccharomyces cerevisiae genome. Nature **418:** 387-391.

GÖRNER, W., E. DURCHSCHLAG, M. T. MARTINEZ-PASTOR, F. ESTRUCH, G. AMMERER *et al.*, 1998 Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity. Genes Dev **12:** 586-597.

HAN, J., J. D. LEE, L. BIBBS and R. J. ULEVITCH, 1994 A MAP kinase targeted by endotoxin and hyperosmolarity in mammalian cells. Science **265:** 808-811.

HEINISCH, J. J., A. LORBERG, H. P. SCHMITZ and J. J. JACOBY, 1999 The protein kinase C-mediated MAP kinase pathway involved in the maintenance of cellular integrity in Saccharomyces cerevisiae. Mol Microbiol **32:** 671-680.

HORIE, T., K. TATEBAYASHI, R. YAMADA and H. SAITO, 2008 Phosphorylated Ssk1 prevents unphosphorylated Ssk1 from activating the Ssk2 mitogen-activated protein kinase kinase kinase in the yeast high-osmolarity glycerol osmoregulatory pathway. Mol Cell Biol **28:** 5172-5183.

HUGHES HALLETT, J. E., X. LUO and A. P. CAPALDI, 2014 State transitions in the TORC1 signaling pathway and information processing in Saccharomyces cerevisiae. Genetics **198:** 773-786.

HUGHES, J. D., P. W. ESTEP, S. TAVAZOIE and G. M. CHURCH, 2000 Computational identification of cis-regulatory elements associated with groups of functionally related genes in Saccharomyces cerevisiae. J Mol Biol **296:** 1205-1214.

JOHNSTON, G. C., R. A. SINGER, S. O. SHARROW and M. L. SLATER, 1980 Cell division in the yeast Saccharomyces cerevisiae growing at different rates. Microbiology **118:** 479-484.

JORGENSEN, P., I. RUPES, J. R. SHAROM, L. SCHNEPER, J. R. BROACH *et al.*, 2004 A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. Genes Dev **18:** 2491-2505.

LAMSON, R. E., M. J. WINTERS and P. M. PRYCIAK, 2002 Cdc42 regulation of kinase activity and signaling by the yeast p21-activated kinase Ste20. Mol Cell Biol **22:** 2939-2951.

LEE, J., R. D. MOIR and I. M. WILLIS, 2009 Regulation of RNA polymerase III transcription involves SCH9-dependent and SCH9-independent branches of the target of rapamycin (TOR) pathway. J Biol Chem **8:** 12604-12608.

LEE, M. V., S. E. TOPPER, S. L. HUBLER, J. HOSE, C. D. WENGER *et al.*, 2011 A dynamic model of proteome changes reveals new roles for transcript alteration in yeast. Mol Syst Biol **19:** 514.

LI, H., C. K. TSANG, M. WATKINS, P. G. BERTRAM and X. F. ZHENG, 2006 Nutrient regulates Tor1 nuclear localization and association with rDNA promoter. Nature **442:** 1058-1061.

LIPPMAN, S. I., and J. R. BROACH, 2009 Protein kinase A and TORC1 activate genes for ribosomal biogenesis by inactivating repressors encoded by Dot6 and its homolog Tod6. Proc Natl Acad Sci U S A **106:** 19928-19933.

LOEWITH, R., E. JACINTO, S. WULLSCHLEGER, A. LORBERG, J. L. CRESPO *et al.*, 2002 Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth contro. Mol Cell **10:** 457-468.

LOPEZ, A. D., K. TAR, U. KRÜGEL, T. DANGE, I. G. ROS *et al.*, 2011 Proteasomal degradation of Sfp1 contributes to the repression of ribosome biogenesis during starvation and is mediated by the proteasome activator Blm10. Mol Biol Cell **22:** 528-540.

LU, C., M. J. BRAUER and D. BOTSTEIN, 2009 Slow growth induces heat-shock resistance in normal and respiratory-deficient yeast. Mol Biol Cell **20:** 891-903.

MAEDA, T., A. Y. TSAI and H. SAITO, 1993 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 **13:** 5408-5417.

MAPES, J., and I. M. OTA, 2004 Nbp2 targets the Ptc1-type 2C Ser/Thr phosphatase to the HOG MAPK pathway. EMBO J **23:** 302-311.

MARION, R. M., A. REGEV, E. SEGAL, Y. BARASH, D. KOLLER *et al.*, 2004 Sfp1 is a stress- and nutrient-sensitive regulator of ribosomal protein gene expression. Proc Natl Acad Sci U S A **101:** 14315-14322.

MARTIN, D. E., A. SOULARD and M. N. HALL, 2004 TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1. Cell **119:** 969-979.

MARTÍNEZ-PASTOR, M. T., G. MARCHLER, C. SCHÜLLER, A. MARCHLER-BAUER, H. RUIS *et al.*, 1996 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.

MATTISON, C. P., and I. M. OTA, 2000 Two protein tyrosine phosphatases, Ptp2 and Ptp3, modulate the subcellular localization of the Hog1 MAP kinase in yeast. Genes Dev **14:** 1229-1235.

MELAMED, D., L. PNUELI and Y. ARAVA, 2008 Yeast translational response to high salinity: global analysis reveals regulation at multiple levels. RNA **14:** 1337-1351.

MITCHELL, A., G. H. ROMANO, B. GROISMAN, A. YONA, E. DEKEL *et al.*, 2009 Adaptive prediction of environmental changes by microorganisms. Nature **460:** 220-224.

MOIR, R. D., J. LEE, R. A. HAEUSLER, N. DESAI, D. R. ENGELKE *et al.*, 2006 Protein kinase A regulates RNA polymerase III transcription through the nuclear localization of Maf1. Proc Natl Acad Sci U S A **103:** 15044-15049.

NADAL-RIBELLES, M., N. CONDE, O. FLORES, J. GONZÁLEZ-VALLINAS, E. EYRAS *et al.*, 2012 Hog1 bypasses stress-mediated down-regulation of transcription by RNA polymerase II redistribution and chromatin remodeling. Genome Biol **13:** R106.

NEIDHARDT, F. C., and B. MAGASANIK, 1960 Studies on the role of ribonucleic acid in the growth of bacteria. Biochim Biophys Acta **42:** 99-116.

NEUMAN-SILBERBERG, F. S., S. BHATTACHARYA and J. R. BROACH, 1995 Nutrient availability and the RAS/cyclic AMP pathway both induce expression of ribosomal protein genes in Saccharomyces cerevisiae but by different mechanisms. Mol Biol Cell **15:** 3187-3196.

O'DUIBHIR, E., P. LIJNZAAD, J. J. BENSCHOP, T. L. LENSTRA, D. VAN LEENEN *et al.*, 2014 Cell cycle population effects in perturbation studies. Mol Syst Biol **10:** 732.

O'ROURKE, S. M., and I. HERSKOWITZ, 2004 Unique and redundant roles for HOG MAPK pathway components as revealed by whole-genome expression analysis. Mol Biol Cell **15:** 532-542.

PHILIPPI, A., R. STEINBAUER, A. REITER, S. FATH, I. LEGER-SILVESTRE *et al.*, 2010 TOR-dependent reduction in the expression level of Rrn3p lowers the activity of the yeast RNA Pol I machinery, but does not account for the strong inhibition of rRNA production. Nucleic Acids Res **38:** 5315-5326.

POKHOLOK, D. K., J. ZEITLINGER, N. M. HANNETT, D. B. REYNOLDS and R. A. YOUNG, 2006 Activated signal transduction kinases frequently occupy target genes. Science **313:** 533-536.

POSAS, F., S. M. WURGLER-MURPHY, T. MAEDA, E. A. WITTEN, T. C. THAI *et al.*, 1996 Yeast HOG1 MAP kinase cascade is regulated by a multistep phosphorelay mechanism in the SLN1-YPD1-SSK1 "two-component" osmosensor. Cell **86:** 865-875.

POWERS, T., and P. WALTER, 1999 Regulation of ribosome biogenesis by the rapamycin-sensitive TOR-signaling pathway in Saccharomyces cerevisiae. Mol Biol Cell **10:** 987-1000.

PROFT, M., G. MAS, E. DE NADAL, A. VENDRELL, N. NORIEGA *et al.*, 2006 The stress-activated Hog1 kinase is a selective transcriptional elongation factor for genes responding to osmotic stress. Mol Cell **23:** 241-250.

PROFT, M., A. PASCUAL-AHUIR, E. DE NADAL, J. ARIÑO, R. SERRANO *et al.*, 2001 Regulation of the Sko1 transcriptional repressor by the Hog1 MAP kinase in response to osmotic stress. EMBO J **20:** 1123-1133.

PROFT, M., and K. STRUHL, 2004 MAP kinase-mediated stress relief that precedes and regulates the timing of transcriptional induction. Cell **118:** 351-361.

RAITT, D. C., F. POSAS and H. SAITO, 2000 Yeast Cdc42 GTPase and Ste20 PAK-like kinase regulate Sho1-dependent activation of the Hog1 MAPK pathway. EMBO J **19:** 4623-4631.

RAMACHANDRAN, V., and P. K. HERMAN, 2011 Antagonistic Interactions Between the cAMP-Dependent Protein Kinase and Tor Signaling Pathways Modulate Cell Growth in Saccharomyces cerevisiae. Genetics **187:** 441-454.

REGENBERG, B., T. GROTKJAER, O. WINTHER, A. FAUSBØLL, M. AKESSON *et al.*, 2006 Growth-rate regulated genes have profound impact on interpretation of transcriptome profiling in Saccharomyces cerevisiae. Genome Biol **7:** R107.

REISER, V., H. RUIS and G. AMMERER, 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.

RUDRA, D., and J. R. WARNER, 2004 What better measure than ribosome synthesis? Genes Dev **18:** 2431-2436.

SANTHANAM, A., A. HARTLEY, K. DÜVEL, J. R. BROACH and S. GARRETT, 2004 PP2A phosphatase activity is required for stress and Tor kinase regulation of yeast stress response factor Msn2p. Eukaryot Cell **3:** 1261-1271.

SCHMELZLE, T., T. BECK, D. E. MARTIN and M. N. HALL, 2004 Activation of the RAS/cyclic AMP pathway suppresses a TOR deficiency in yeast. Mol Biol Cell **24:** 338-351.

SMITH, A., M. P. WARD and S. GARRETT, 1998 Yeast PKA represses Msn2p/Msn4p-dependent gene expression to regulate growth, stress response and glycogen accumulation. EMBO J **17:** 3556-3564.

SOULARD, A., A. CREMONESI, S. MOES, F. SCHÜTZ, P. JENÖ *et al.*, 2010 The rapamycin-sensitive phosphoproteome reveals that TOR controls PKA toward some but not all substrates. Mol Biol Cell **21:** 3475-3486.

STOCK, A. M., V. L. ROBINSON and P. N. GOUDREAU, 2000 Two-component signal transduction. Annu Rev Biochem **69:** 183-215.

TAMÁS, M. J., M. REP, J. M. THEVELEIN and S. HOHMANN, 2000 Stimulation of the yeast high osmolarity glycerol (HOG) pathway: evidence for a signal generated by a change in turgor rather than by water stress. FEBS Lett **472:** 159-165.

TATEBAYASHI, K., K. TANAKA, H. Y. YANG, K. YAMAMOTO, Y. MATSUSHITA *et al.*, 2007 Transmembrane mucins Hkr1 and Msb2 are putative osmosensors in the SHO1 branch of yeast HOG pathway. EMBO J **26:** 3521-3533.

THEVELEIN, J. M., and J. H. DE WINDE, 1999 Novel sensing mechanisms and targets for the cAMP-protein kinase A pathway in the yeast Saccharomyces cerevisiae. Mol Microbiol **33:** 904-918.

UESONO, Y., and A. TOH-E, 2002 Transient inhibition of translation initiation by osmotic stress. J Biol Chem **277:** 13848-13855.

UNO, I., H. MITSUZAWA, K. TANAKA, T. OSHIMA and T. ISHIKAWA, 1987 Identification of the domain of Saccharomyces cerevisiae adenylate cyclase associated with the regulatory function of RAS products. Mol Gen Genet **210:** 187-194.

URBAN, J., A. SOULARD, A. HUBER, S. LIPPMAN, D. MUKHOPADHYAY *et al.*, 2007 Sch9 is a major target of TORC1 in Saccharomyces cerevisiae. Mol Cell **26:** 663-674.

VOGEL, C., G. M. SILVA and E. M. MARCOTTE, 2011 Protein expression regulation under oxidative stress. Mol Cell Proteomics **10:** M111.009217.

VON DER HAAR, T., 2008 A quantitative estimation of the global translational activity in logarithmically growing yeast cells. BMC Syst Biol **2**.

WADE, C. H., M. A. UMBARGER and M. A. MCALEAR, 2006 The budding yeast rRNA and ribosome biosynthesis (RRB) regulon contains over 200 genes. Yeast **23:** 293-306.

WADE, J. T., D. B. HALL and K. STRUHL, 2004 The transcription factor Ifh1 is a key regulator of yeast ribosomal protein genes. Nature **432:** 1054-1058.

WANG, Y., M. PIERCE, L. SCHNEPER, C. G. GÜLDAL, X. ZHANG *et al.*, 2004 Ras and Gpa2 mediate one branch of a redundant glucose signaling pathway in yeast. PLoS Biol **2:** E128.

WARNER, J. R., 1999 The economics of ribosome biosynthesis in yeast. Trends Biochem Sci **24:** 437-440.

WESTFALL, P. J., J. C. PATTERSON, R. E. CHEN and J. THORNER, 2008 Stress resistance and signal fidelity independent of nuclear MAPK function. Proc Natl Acad Sci U S A **105:** 12212-12217.

WILLIS, I. M., and R. D. MOIR, 2007 Integration of nutritional and stress signaling pathways by Maf1. Trends Biochem Sci **32:** 51-53.

XIAO, L., and A. GROVE, 2009 Coordination of Ribosomal Protein and Ribosomal RNA Gene Expression in Response to TOR Signaling. Curr Genomics **10:** 198-205.

YAAKOV, G., A. DUCH, M. GARCÍA-RUBIO, J. CLOTET, J. JIMENEZ *et al.*, 2009 The stress-activated protein kinase Hog1 mediates S phase delay in response to osmostress. Mol Biol Cell **20:** 3572-3582.

ZAKRZEWSKA, A., G. VAN EIKENHORST, J. E. BURGGRAAFF, D. J. VIS, H. HOEFSLOOT *et al.*, 2011 Genome-wide analysis of yeast stress survival and tolerance acquisition to analyze the central trade-off between growth rate and cellular robustness. Mol Biol Cell **22:** 4435-4446.

ZURITA-MARTINEZ, S. A., and M. E. CARDENAS, 2005 Tor and Cyclic AMP-Protein Kinase A: Two Parallel Pathways Regulating Expression of Genes Required for Cell Growth. Eukaryot Cell **4:** 63-71.

Chapter 2

Pathway connectivity and signaling coordination in the yeast stress-activated signaling network

Deborah Chasman*, Yi-Hsuan Ho*, David B Berry, Corey M Nemec, Matthew E MacGilvray, James Hose, Anna E Merrill, M Violet Lee, Jessica L Will, Joshua J Coon, Aseem Z Ansari, Mark Craven, Audrey P Gasch

* These authors share first authorship

This Work is published on *Molecular Systems Biology* (2014) 10: 759

I contributed to all the biological investigation of Cdc14 and CTD analysis and participated in writing the manuscript.

.

Abstract

Stressed cells coordinate a multi-faceted response spanning many levels of physiology. Yet knowledge of the complete stress-activated regulatory network as well as design principles for signal integration remain incomplete. We developed an experimental and computational approach to integrate available protein interaction data with gene fitness contributions, mutant transcriptome profiles, and phospho-proteome changes in cells responding to salt stress, to infer the salt-responsive signaling network in yeast. The inferred subnetwork presented many novel predictions by implicating new regulators, uncovering unrecognized crosstalk between known pathways, and pointing to previously unknown 'hubs' of signal integration. We exploited these predictions to show that Cdc14 phosphatase is a central hub in the network and that modification of RNA Polymerase II coordinates induction of stress-defense genes with reduction of growth-related transcripts. We find that the orthologous human network is enriched for cancer-causing genes, underscoring the importance of the subnetwork's predictions in understanding stress biology.

Introduction

All cells respond to stress by orchestrating complex responses customized for each situation. When grown in optimal conditions, *Saccharomyces cerevisiae* maintains high expression of growth-related genes and low transcription of stress-defense genes, in part via nutrient responsive TOR and RAS-regulated Protein Kinase A (PKA) signaling (SMETS *et al.* 2010; BROACH 2012). Suboptimal conditions suppress these pathways in an unknown manner while activating stress-specific signaling networks that coordinate changes in transcription and translation, protein function, and metabolic fluxes with transient arrest of growth and cell cycle progression. How these disparate physiological processes are coordinated is poorly understood but likely critical for surviving and acclimating to stressful conditions.

At the level of gene expression, stressed yeast activate condition-specific transcript changes that provide specialized stress defenses. These responses are typically regulated by condition-specific transcription factors (TFs) and upstream signaling pathways that are activated under limited circumstances (HOHMANN and MAGER 2003). Concurrently, stressed yeast activate the common environmental stress response (ESR) (GASCH *et al.* 2000; CAUSTON *et al.* 2001). The ESR includes ~300 induced (iESR) genes that are broadly involved in stress defense and ~600 repressed-ESR (rESR) genes that together encode ribosomal proteins (RPs) and proteins involved in ribosome biogenesis/protein synthesis (RiBi). While the complete set of ESR regulators remains elusive, it is clear that the program is regulated by different upstream signaling factors under different situations (GASCH *et al.* 2000; GASCH *et al.* 2001; GASCH 2002). Activation of the ESR, and of transcript changes more broadly, is in fact not required to survive the initial stressor, but rather is necessary for acquired resistance to subsequent stress (BERRY and GASCH 2008; WESTFALL *et al.* 2008; MITCHELL *et al.* 2009; BERRY *et al.* 2011). Therefore, screens for mutants sensitive to a single dose of stress have likely missed many signaling proteins, rendering stress-dependent signaling networks incomplete. Although several isolated 'pathways' are well characterized, how signaling is integrated through a single cellular system is poorly understood.

Here we present an experimental and computational pipeline to infer the complete sodium chloride (NaCl)-activated signaling network from a combination of data types. A key feature of our approach is that we generated several large-scale datasets (including mutant transcriptome profiles, phospho-proteome changes, and gene fitness contributions) under the same culture system in cells responding to acute NaCl stress. Because stress responses are highly context dependent (VAN WUYTSWINKEL *et al.* 2000; O'ROURKE and HERSKOWITZ 2004; BERRY and GASCH 2008), we restrict our analysis to datasets generated in our own lab, despite many insightful prior studies characterizing the salt response in yeast (e.g. (CAUSTON *et al.* 2001; HIRASAWA *et al.* 2006; CAPALDI *et al.* 2008; MELAMED *et al.* 2008; WESTFALL *et al.* 2008;

HALBEISEN and GERBER 2009; SOUFI *et al.* 2009; MARTINEZ-MONTANES *et al.* 2010; WARRINGER *et al.* 2010; MILLER *et al.* 2011)).

We wished to develop a computational method to integrate these datasets and infer the stress-activated signaling subnetwork, both to implicate missing regulators and to understand their connections. Prior approaches tackling the challenge of network inference have leveraged large-scale biological datasets, most commonly transcriptome data (see (FRIEDMAN 2004; SCHADT *et al.* 2005)). Extensions focusing on the osmotic response include the work of Gat-Viks *et al.,* whose probabilistic method described regulatory relationships between known regulators of the Hog pathway, assuming a known network topology (GAT-VIKS *et al.* 2006; GAT-VIKS and SHAMIR 2007). Several approaches leverage protein-protein and protein-nucleic-acid interactions to infer relevant connections between regulators and their downstream gene targets (LIANG *et al.* 1998; IDEKER *et al.* 2000; YEANG *et al.* 2004a; YEUNG *et al.* 2004; MARKOWETZ *et al.* 2005; TU *et al.* 2006; SUTHRAM *et al.* 2008; HUANG and FRAENKEL 2009; VASKE *et al.* 2009; YEGER-LOTEM *et al.* 2009; NOVERSHTERN *et al.* 2011; HUANG and FRAENKEL 2012). The method we present here is most closely related to methods that infer subnetworks by solving an integer linear program (IP) (OURFALI *et al.* 2007; GITTER *et al.* 2011; SILVERBUSH *et al.* 2011). In particular, Gitter *et al* (2013) developed a combined probabilistic/IP method to discern signaling in the potassium chloride-responsive subnetwork from time-series expression data (GITTER *et al.* 2013). However, their approach incorporated transcriptome data only, whereas we were interested in incorporating other data types. Methods that integrate disparate datasets are emerging, for example the work of Huang *et al* (2013) that considered existing transcriptomic and proteomic data to study oncogene-induced signaling (HUANG *et al.* 2013). In our case, we wanted to design a method that could also take mutant transcriptome profiles generated in our own lab.

We therefore designed an integer linear programming (IP) approach to integrate and interpret

our disparate datasets by inferring a signaling subnetwork. The novel facets of our computational approach include a means to integrate these varied data sources, using new types of input paths to the IP, and a multi-part objective function. The resulting subnetwork generated many new insights into stress signaling, by implicating new regulators, unveiling the connections between them, and presenting organization principles that shed light on stress biology.

Results

We previously identified 225 genes important for acquired stress resistance after NaCl pretreatment (BERRY *et al.* 2011), including a subset of the known signaling proteins activated by NaCl (Appendix A Fig. 1). Because only a fraction of NaCl-dependent transcript changes are important for acquired stress resistance, the selection misses many of the upstream transcriptome regulators. Therefore, to implicate the complete upstream signaling subnetwork, we began by profiling NaCl-dependent expression changes in 16 mutants implicated in NaCl-induced acquired stress tolerance (Fig. 1, see Methods). Together, this generated a matrix of regulator-gene target predictions that encompassed 3,300 genes (Appendix A Fig. 2 and Table 1). A third of the affected genes were dependent on ≥2 regulators, and there was significant overlap in several target-gene sets (hypergeometric test, Fig. 1). These results hint at the complex upstream signaling that controls the NaCl-responsive transcriptome.

Because much of signal transduction occurs post-translationally, we next measured changes to the phospho-proteome before and at 5 and 15 min after NaCl treatment, using chemical isobaric tags for phospho-peptide quantification (see Methods). Nearly 600 of 1937 identified phospho-sites (mapping to 973 proteins) showed a ≥ 2-fold change in phosphorylation, roughly split

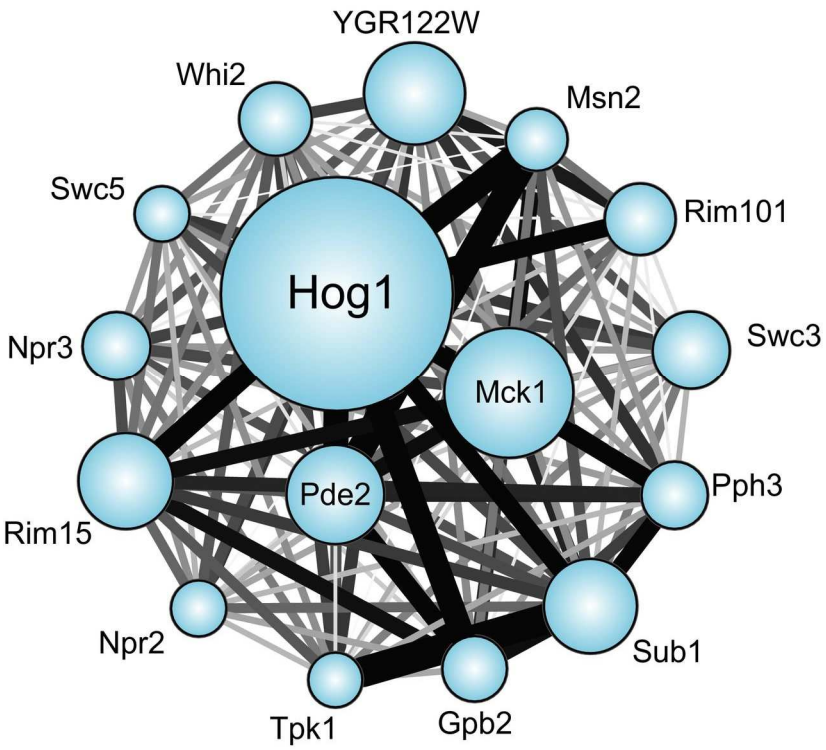


Fig. 1. Overlapping targets of interrogated 'source' regulators. The number of genes whose osmotic response was defective in each of 16 mutants is represented by the size of each circle. Edge thickness represents the fraction of the smaller node's targets that overlap between two nodes. Edge color is proportional to significance of the overlap (hypergeometric test), where black represents a –log(p-value) of 5 or greater.

Mutant[a]	Defective[b]	Amplified[b]
hog1Δ (3)□□	1378	565
pde2Δ (3)□□□□□	517	59
mck1Δ (3)□□□□□	794	101
msn2Δ (3)□□□□	184	26
rim101Δ (3)□□□□□	75	227
gpb2Δ (2)□	202	37
rim15Δ (2)□	438	106
npr2Δ (2)□	75	69
npr3Δ (2)□	184	89
swc3Δ (2)□	108	257
swc5Δ (2)□	84	55
whi2Δ (2)□	118	201
pph3Δ (2)□	235	21
sub1Δ (2)□	431	97
tpk1Δ (2)□	35	96
ygr122wΔ(2)□	106	502

Table 1: Lists the number of targets identified from the originally interrogated 'source' regulators.

[a] Mutant and number of replicates in parentheses.

[b] Number of genes with smaller ('defective') or larger ('amplified') expression changes compared to the wild type strain. Note this table includes non-coding RNAs that were excluded from the inference.

between sites with increased and decreased modification (Appendix A Fig. 3). Over 10% of the altered phospho-proteins represented kinases and phosphatases (including regulators of cell-cycle progression, actin organization, and signal transduction) as well as transcriptional regulators (such as activators Hot1, Sko1, and Sub1 and repressors Mot2, Dot6, and Dig1). Proteins affected at the later time point were involved in cytokinesis, bud-site selection, and actin reorganization (Bonferroni-corrected $p < 0.01$, hypergeometric test), implying downstream physiological effects on these processes.

This analysis generated a rich source of datasets (outlined in Fig 2). To integrate and interpret these disparate datasets, we designed an integer linear programming-based (IP) approach (Fig. 3 and Methods). Using a *background network* of physical or chemical protein interactions, the method infers a *subnetwork* that predicts the *paths* by which each interrogated *source* regulator is connected to its downstream *targets* (identified as dysregulated genes in the *source* mutant responding to NaCl treatment). Each path is a directed, linear chain of interactions between yeast proteins, where the terminal protein node represents a sequence-specific transcription factor (TF) or RNA-binding protein (RBP) known to bind the downstream promoters or transcripts, respectively. The IP's objective function favors the inclusion of salt-responsive proteins, *i.e.* those with differential phosphorylation or required for acquired-stress fitness after NaCl treatment, and allows the sparing inclusion of additional proteins.

Specifically, we start with a background network of directed and undirected intracellular interactions representing protein-protein, kinase-substrate, and gene regulatory interactions

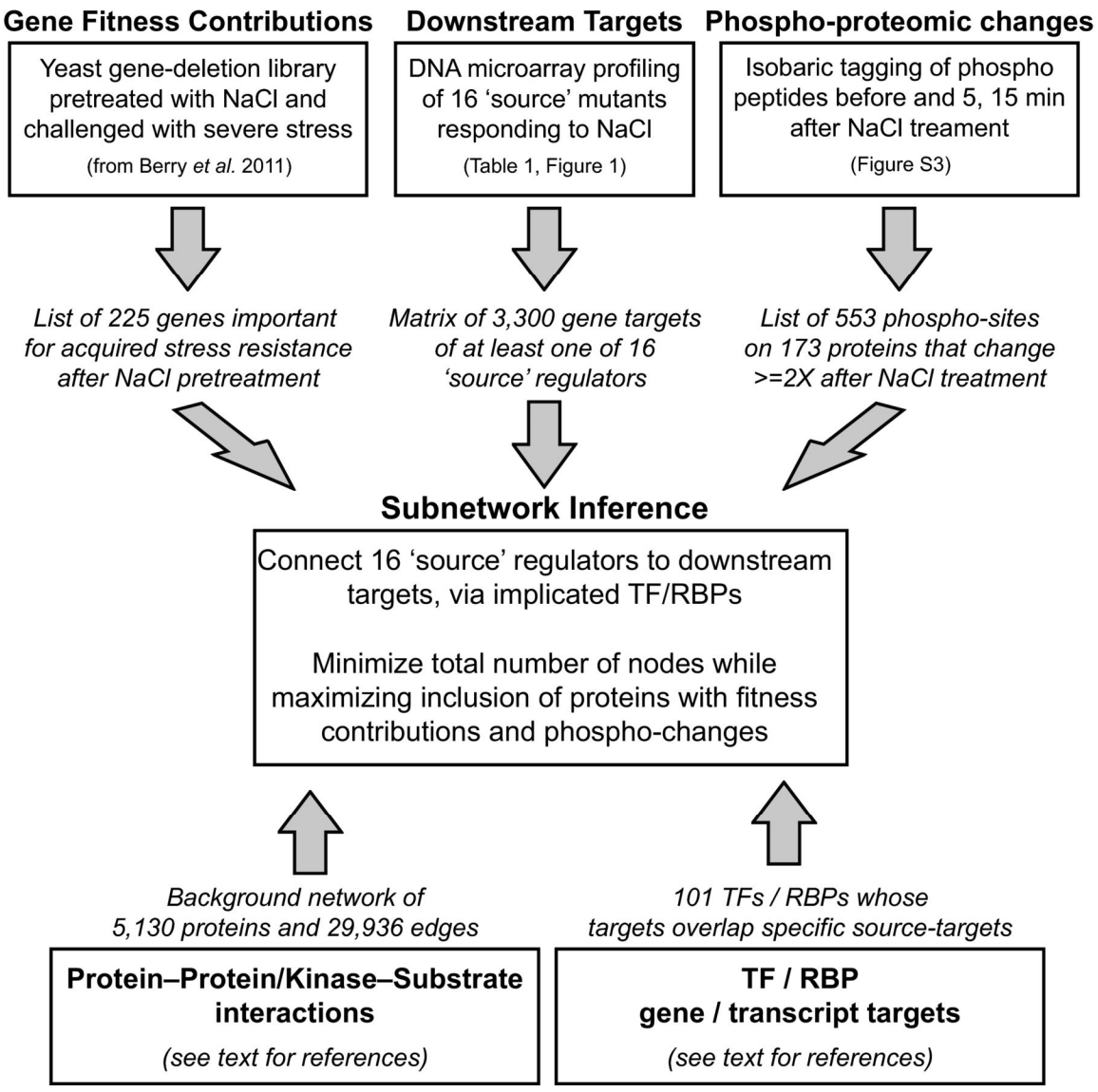


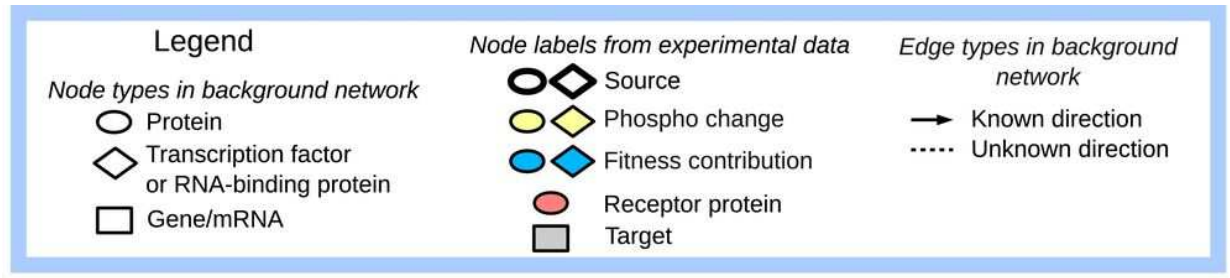
Fig 2. Overview of the experimental data collection and analysis to generate IP input. See text for details.

Fig. 3. Overview of the subnetwork inference method.

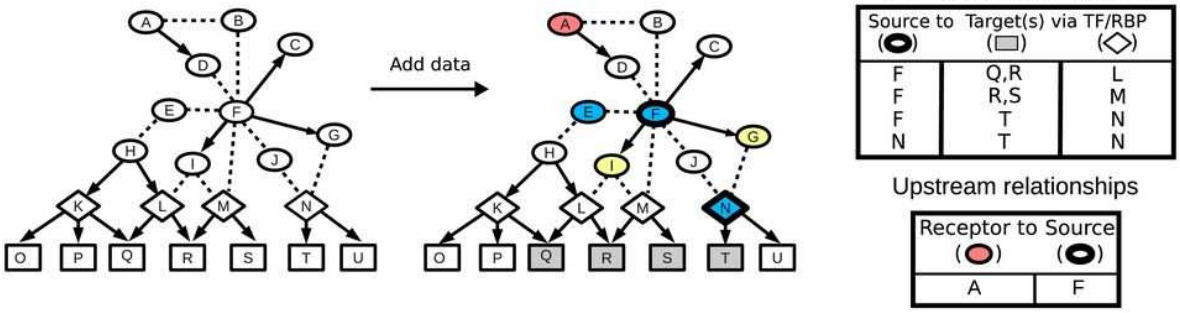
A. The input to the method includes a background network of yeast interactions combined with experimental data that describes the yeast salt stress response, including proteins with phospho-changes (yellow), fitness contribution (blue), or two known upstream regulators (pink), as described in the key.

B. The three different types of paths that we enumerate using the background network and experimental data, where 'hit' refers to proteins identified in the original fitness screen or with significant changes in phosphorylation.

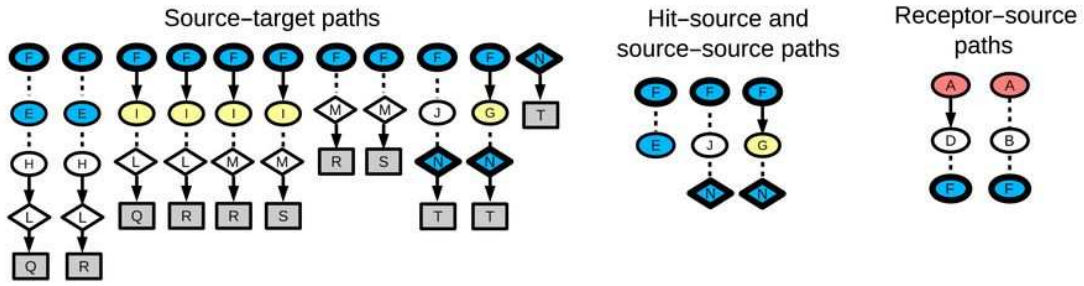
C. The IP for subnetwork inference and the output ensemble of inferred subnetworks.



A Combine background network and experimental data



B Enumerate candidate paths

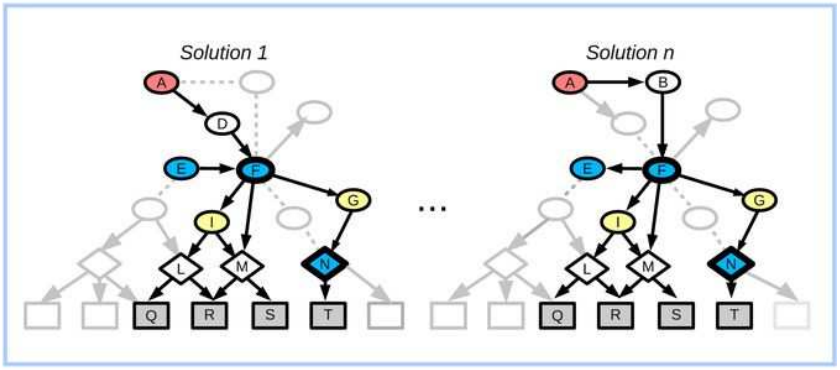


C Solve IP to infer ensemble of subnetworks

IP constraints:
- Connect source–target pairs, hit–source pairs, and receptor–source pairs
- Assign a direction to each undirected edge

IP objectives:
- Favor proteins with
  - Phospho changes
  - Fitness contributions
- Sparingly include proteins without evidence of relevance

Subnetwork ensemble:



between proteins and genes/mRNAs (GUELZIM *et al.* 2002; PTACEK *et al.* 2005; MACISAAC *et al.* 2006; STARK *et al.* 2006; HOGAN *et al.* 2008; EVERETT *et al.* 2009; PU *et al.* 2009; BREITKREUTZ *et al.* 2010; SCHERRER *et al.* 2010; TSVETANOVA *et al.* 2010; ABDULREHMAN *et al.* 2011; FASOLO *et al.* 2011; SHARIFPOOR *et al.* 2011; VENTERS *et al.* 2011; HEAVNER *et al.* 2012; HUEBERT *et al.* 2012). For each interrogated source regulator, we identify candidate TFs and RBPs whose known binding targets significantly overlap with the source's targets (Fig. 3A). We then enumerate all possible directed candidate paths (using an iterative deepening search up to a given length) that connect each of the 16 interrogated *source* regulators to the majority of their *targets*, through candidate TFs or RBPs (Fig. 3B). Other candidate paths connect proteins required for fitness (Fig. 3B, blue nodes), proteins with NaCl-dependent phosphorylation changes (yellow nodes), and two known upstream sensors (pink nodes). The candidate paths serve as input to the IP, which encodes the relevance of each network element as a binary variable and characterizes possible subnetworks using a set of linear constraints over these variables (Fig. 3C). Subnetwork inference is performed by choosing a union of relevant, directed paths that optimize a series of successively applied objective functions that aim to connect experimentally implicated proteins while minimally including proteins not currently supported by experimental evidence. Because many distinct subnetworks may score equally well, we use the IP to identify an ensemble of high-scoring subnetworks. In turn, each protein, interaction, and path is assigned a confidence value based on its frequency across the ensemble.

Validation analysis provides strong support for the inferred subnetwork

Using the datasets described above, the method identified a consensus subnetwork encompassing 380 nodes (predicted regulators) and 1131 edges (relevant interactions) present at 75% confidence (Fig. 4A). To assess the inferred subnetwork's predictive accuracy, we performed precision-recall analysis using an assembled list of known NaCl regulators and

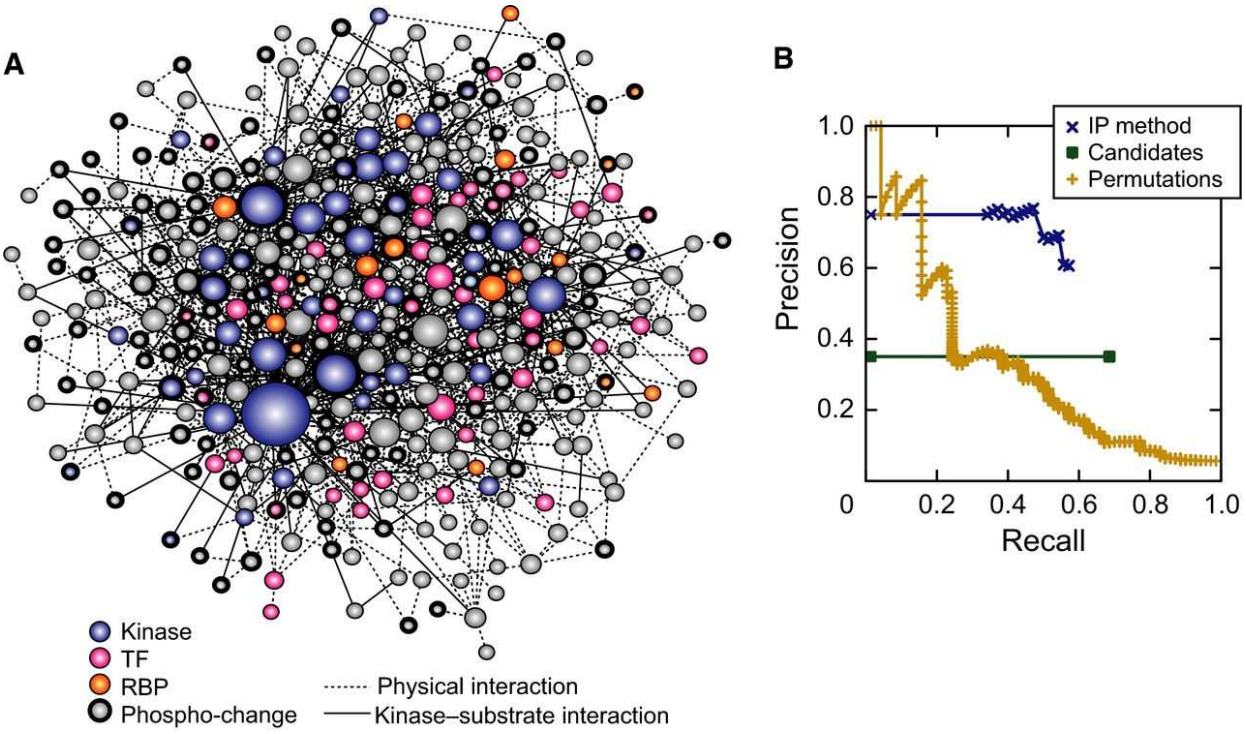


Fig. 4. Inferred NaCl-activated signaling network.

A. Inferred consensus subnetwork at 75% confidence, where node size indicates degree (number of connections) and color is according to the key. Nodes representing proteins with phospho-changes are outlined in bold.

B. Precision-recall of the inferred consensus network was calculated using a list of true positives from the literature and a list of likely negatives, after excluding proteins with phospho-changes and those required for fitness (see Methods). *Precision* is the fraction of predicted nodes known to be involved in the osmo response, and *recall* is the fraction of true positives that are above the threshold. The curves represent the performance of the IP method on the real data (blue), of the method on randomized permutations of the input network (yellow), and of the candidate enumerated pathways used as input (green, see Methods).

another list of unlikely regulators that included metabolic enzymes and exclusively subcellular proteins. We excluded from consideration proteins with phospho-changes or fitness contributions (since they are preferentially included by the inference) and plotted the precision and recall over varying node-confidence thresholds (Fig. 4B). The inferred ensemble achieved substantially higher accuracy than the enumerated candidate paths provided as input to the IP method, highlighting the power of the inference step (Fig. 4B, green line). To assess the effects of the topological properties in the background network, we ran the method on permuted source-target pairs (maintaining the degree distribution from the real data; see Methods). This permuted baseline achieved high accuracy in the low-recall range, suggesting that some regulators are highly central in the background network. However, our inferred ensemble significantly outperformed the permuted baseline at higher levels of recall; thus, our method's accuracy is not simply due to properties of the background network's topology. To understand the contribution of each component of our method, we also performed additional enrichment analyses and other computational evaluations, with results available in Appendix A Information Section 2.

We found additional support for the inferred subnetwork in the non-random inclusion of specific protein functional groups. When compared to the background network, to the enumerated candidate pathways used as input to the IP, and to the permuted subnetworks, the inferred consensus subnetwork was enriched for proteins annotated as 'stress' proteins (background, p = 5e-21; candidates, p = 2e-6; permutations, p = 0.007) and for proteins encoded by genes with genetic interactions (background and candidates, $p \approx 0$; permutations, p = 0.003)(STARK *et al.* 2006), which suggests functional dependencies. The consensus subnetwork was also slightly enriched for kinases (relative to the candidate paths and background network) and for essential genes (relative to the background network), but not relative to the permuted subnetworks (suggesting its bias toward kinases and essential genes).

Mutant	Defective	Amplified
cdc14-3(3)*	929	346
nnk1Δ (1)□	94	278
bck1Δ (1)□	107	169
yak1Δ (1)□	226	248
kin2Δ (1)□	52	266
pho85Δ (1)□	614	342
cka2Δ (2)□	155	63
cka1Δ (2)□	58	133
*ckb1Δckb2Δ(*2)□	129	176
arf3Δ (2)□	466	331
scd6Δ(2)	0	0

Table 2. Lists targets identified for validation mutants.

* *cdc14-3* was compared to its isogenic and identically treated wild type.

The inferred subnetwork included many regulators not previously linked to the NaCl response. To test some of the novel predictions, we analyzed osmo-dependent transcriptome changes in 14 mutants lacking predicted regulators, with preferences for kinases and phosphatases (see Table 2, Appendix A Fig. 4 and Appendix A Information). The results provided strong support overall for the inferred subnetwork. All but one of the mutants (93%) displayed a defect in osmo-responsive expression. Furthermore, the predicted targets of 80% of these regulators overlapped significantly (p < 1e-3) with their measured targets, highlighting the accuracy of regulator-target predictions. To garner support for the subnetwork's structure, we investigated the overlap in targets of each interrogated mutant and the known or measured targets of proteins predicted to lie in the interrogated regulator's paths. Using stringent scoring, we found support for 30-100% of nodes in most paths (53% on average, Appendix A Table 1). Together these results provide strong support for the validity of the inferred consensus subnetwork.

Known and new players captured in the NaCl-responsive signaling subnetwork

We therefore explored the consensus subnetwork for new insights into stress signaling. Many expected pathways were captured, including the canonical HOG, PKA, and TOR pathways. The inferred subnetwork included other stress-activated pathways not previously linked to the NaCl response, such as PKC, Pho85, Rim15 pathways and GSK-3 kinase Mck1 (Fig. 5A). We tested the involvement of these pathways by analyzing our phospho-proteomic data and mutant transcriptome profiles: we found that members of all of these pathways showed NaCl-dependent phospho-changes, and cells lacking specific pathway members (including *BCK1*, *YAK1*, *PHO85*, *RIM15*, and *MCK1*) had defects in NaCl-dependent expression changes (Appendix A Fig. 4 and Appendix A text). The subnetwork also included the 'STE' mating pathway, which shares upstream components with the Hog network and is known to be suppressed by Hog1 signaling (O'ROURKE and HERSKOWITZ 1998; MARLES *et al.* 2004; ZARRINPAR *et al.* 2004; MCCLEAN *et al.* 2007; SHOCK *et al.* 2009; PATTERSON *et al.* 2010; NAGIEC and DOHLMAN 2012). The inclusion of

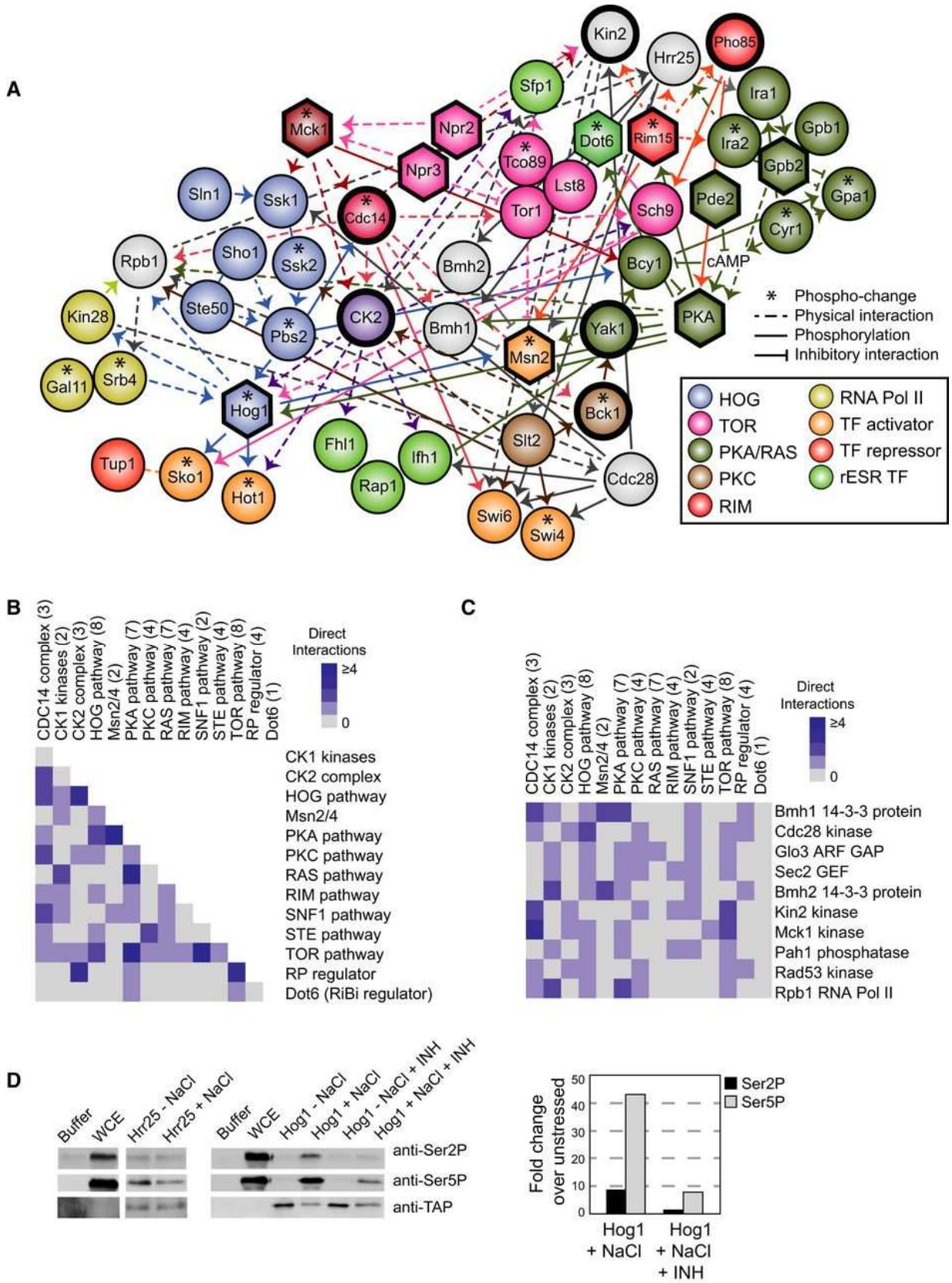
Fig. 5. Connectivity between known pathways and hubs of signal integration.

A. A subregion of the inferred subnetwork, highlighting proteins in known pathways according to the key. Hexagons represent interrogated '*source*' regulators, nodes outlined in bold indicate validated players in the NaCl response, and asterisks represent proteins with phospho-changes upon NaCl treatment. Dashed edges represent physical interactions and solid arrows indicate kinase-substrate relationships. Edge directionality is as predicted by the inference, and edge color is according to the edge's source node. Inhibitory edges were taken from the literature.

B. Connectivity between known pathways, where blue boxes represent the number of interactions between any members of two pathways. Pathway membership is indicated in parentheses.

C. The top 15-ranked 'integrator' nodes with connections to the greatest number of different pathways, as shown in B.

D. A purified CTD peptide was incubated with Hrr25-TAP or Hog1-TAP purified from cells with and without NaCl treatment for 10 min, incubated with and without the reversible p38-specific inhibitor SB203580 (INH) added *in vitro*. Reactions with buffer or yeast whole-cell extract (WCE) served as negative and positive controls, respectively. CTD phosphorylated on serine 2 (Ser2) or Ser5 was detected by immunoblotting (see Methods). TAP-tagged proteins were subsequently quantified on the same blot with the anti-TAP antibody. Quantification of Hog1 phosphorylation, shown to the right, was normalized to Hog1-TAP abundance and then to the corresponding unstressed sample.


A
Kin2
Hrr25
Pho85
Sfp1
Ira1
Mck1
Npr2
Dot6
Rim15
Gpb1
Ira2
Tco89
Npr3
Gpb2
Gpa1
Sln1
Ssk1
Cdc14
Lst8
Tor1
Sch9
Pde2
Cyr1
Rpb1
Sho1
Ssk2
Bmh2
Bcy1
cAMP
Kin28
Ste50
CK2
Bmh1
Yak1
PKA
Pbs2
Msn2
Gal11
Srb4
Hog1
Bck1
Fhl1
Slt2
Tup1
Sko1
Ifh1
Cdc28
Rap1
Hot1
Swi6
Swi4
Phospho-change
Physical interaction
Phosphorylation
Inhibitory interaction
HOG
TOR
PKA/RAS
PKC
RIM
RNA Pol II
TF activator
TF repressor
rESR TF
B
CDC14 complex (3)
CK1 kinases (2)
CK2 complex (3)
HOG pathway (8)
Msn2/4 (2)
PKA pathway (7)
PKC pathway (4)
RAS pathway (7)
RIM pathway (4)
SNF1 pathway (2)
STE pathway (4)
TOR pathway (8)
RP regulator (4)
Dot6 (1)
Direct Interactions
≥4
0
CK1 kinases
CK2 complex
HOG pathway
Msn2/4
PKA pathway
PKC pathway
RAS pathway
RIM pathway
SNF1 pathway
STE pathway
TOR pathway
RP regulator
Dot6 (RiBi regulator)
C
CDC14 complex (3)
CK1 kinases (2)
CK2 complex (3)
HOG pathway (8)
Msn2/4 (2)
PKA pathway (7)
PKC pathway (4)
RAS pathway (7)
RIM pathway (4)
SNF1 pathway (2)
STE pathway (4)
TOR pathway (8)
RP regulator (4)
Dot6 (1)
Direct Interactions
≥4
0
Bmh1 14-3-3 protein
Cdc28 kinase
Glo3 ARF GAP
Sec2 GEF
Bmh2 14-3-3 protein
Kin2 kinase
Mck1 kinase
Pah1 phosphatase
Rad53 kinase
Rpb1 RNA Pol II
D
Buffer
WCE
Hrr25 - NaCl
Hrr25 + NaCl
Buffer
WCE
Hog1 - NaCl
Hog1 + NaCl
Hog1 - NaCl + INH
Hog1 + NaCl + INH
anti-Ser2P
anti-Ser5P
anti-TAP
Fold change over unstressed
50
40
30
20
10
0
Ser2P
Ser5P
Hog1 + NaCl
Hog1 + NaCl + INH

the mating pathway indicates that some connections in the consensus subnetwork represent signaling suppression that prevents crosstalk to other pathways. We also validated several newly implicated regulators, including the CK2 kinase complex (see Appendix A text) and the Cdc14 phosphatase (see below).

Interconnectivity in the inferred signaling subnetwork

The structure of the subnetwork revealed surprising cross-connectivity between previously defined pathways. We defined stress-activated 'pathways' based on the literature, and then summed the number of direct connections between members of those pathways (Fig. 5B). Many of the pathways were intricately connected, with Tor1 and PKA pathways linked to the greatest number of other pathways. We also identified individual subnetwork nodes as 'integration' points, defined as nodes with the greatest number of connections to distinct pathways (Fig. 5C). Nearly half of the top ten integration nodes were kinases or phosphatases, including Mck1 and cell-cycle regulator Cdc28, which regulates RP genes under optimal conditions (CHYMKOWITCH *et al.* 2012) but is suppressed during osmotic shock (ALEXANDER *et al.* 2001; BELLI *et al.* 2001; ADROVER *et al.* 2011). 14-3-3 proteins Bmh1 and Bmh2 were also identified as integration points, confirming their known role as signaling cofactors.

Several of the integration points are also hubs of high connectivity in the consensus subnetwork. While 11 of the top 15 most connected nodes are kinases or phosphatases, the remaining four are known regulatory cofactors – including stress-activated ubiquitin (Ubi4), Sumo (Smt3), and Bmh1 – and the core subunit of RNA polymerase (Pol II), Rpb1. Modification of the Rpb1 carboxyl-terminal domain (CTD) is the basis for the so-called CTD code of transcriptional regulation (BURATOWSKI 2003; ZHANG *et al.* 2012), making it a logical downstream integration point for complex upstream signaling. Consistent with the predictions of the subnetwork, we found that two of the Rpb1-interacting kinases – Hrr25 and Hog1 – phosphorylate the Rpb1-CTD

in vitro. TAP-tag-purified Hrr25-TAP phosphorylated Rpb1-CTD serine 5 (Ser5), regardless of prior NaCl treatment (Fig. 5D). In contrast, TAP-purified Hog1-TAP phosphorylated both Ser2 and Ser5, but only after cellular NaCl treatment and in a manner inhibited by a Hog1-specific inhibitor added *in vitro*. Both Hrr25 and Hog1 are known to interact with Pol II and influence transcriptional processes (ALEPUZ *et al.* 2003; PHATNANI *et al.* 2004; PROFT *et al.* 2006; COOK and O'SHEA 2012; NADAL-RIBELLES *et al.* 2012) but neither had been implicated in direct Rpb1-CTD phosphorylation. These results are consistent with the model that the Rpb1-CTD is a direct target of the signaling network and plays a central role in signaling (see more below).

We also dissected the regulatory connections surrounding a second hub, Cdc14. In the process, we found that Cdc14 is critical for coordinating distinct facets of the NaCl response. First, the defect in NaCl transcriptome changes evident in *cdc14-3* cells overlapped significantly with the Hog1 response, raising the possibility that Cdc14 is important for Hog1 regulation (Appendix A Fig. 4). The subnetwork predicts that Cdc14 is activated in part by the Hog1 regulator Pbs2 (reminiscent of Pbs2 control of Cdc14 localization during the cell cycle (REISER *et al.* 2006)) and that Cdc14 affects Hog1 function via the nuclear exporter, Crm1. This prompted us to follow Hog1 localization in the *cdc14-3* mutant. Indeed, Hog1 nuclear localization was defective in the NaCl-treated *cdc14-3* mutant (Fig. 6A and Appendix A Information), despite Hog1 hyper-phosphorylation under these conditions (Fig. 6B). We found no direct interaction between Cdc14 and Hog1, suggesting that the hyper-phosphorylation of Hog1 is a secondary response to the defect in nuclear localization rather than a deficit of direct Hog1 dephosphorylation by Cdc14. The aberrant Hog1 localization was not a side effect of cell-cycle arrest, since we found no defect in wild-type cells progressing through G2/M phase or in nocodazole-arrested cells (Appendix A Fig. 6). Instead, these results suggest a direct connection between Cdc14 activity and signaling through the Hog pathway.

Fig 6. Cdc14 is a central regulator in the NaCl response.

A. Wild-type (WT) and *cdc14-3* cells were shifted to 35°C for 90 min and then exposed to 0.7M NaCl for up to 20 min. Images represent nuclear DNA (DAPI, left) and Hog1-GFP (right) before and at 10 min after NaCl treatment. The plot below quantifies the fraction of cells (n = >75) with nuclear Hog1-GFP signal that overlapped the DAPI signal, in WT and *cdc14-3* cells.

B. Levels of phospho-Hog1 normalized to total Hog1 in WT and *cdc14-3* cells responding to NaCl at 35°C. Data represent the average of biological duplicates.

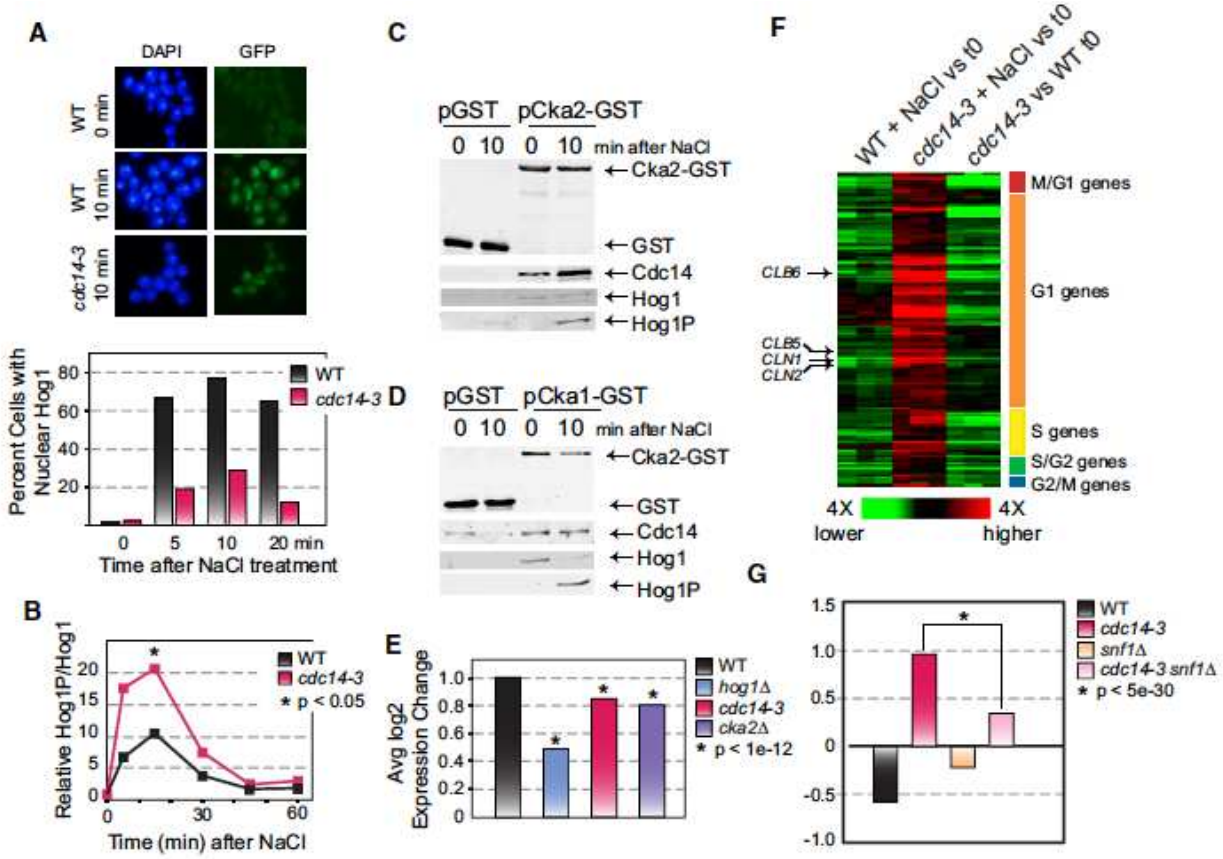
C. GST-tagged Cka2 was immunoprecipitated and blotted for Cdc14 and total or phospho-Hog1.

D. Same as C) but for immunoprecipitated Cka1-GST.

E. The average $\log_2$ fold-change of 67 Hot1 targets in replicated WT, *hog1Δ*, *cdc14-3*, and *cka2Δ* strains responding to NaCl. Data for each mutant and its paired WT were scaled to the plotted WT so as to accurately represent the mutant defect. Asterisks represent a significant difference in the mutant versus its paired WT (paired T-test).

F. Expression data in WT or *cdc14-3* cells responding to NaCl at the nonpermissive temperature and in *cdc14-3* cells versus WT at the non-permissive temperature before NaCl addition. Each column represents one of three triplicated expression responses, and each row represents one of 131 cell cycle genes aberrantly induced in *cdc14-3* after NaCl treatment (FDR<0.05). Red represents higher and green represents lower expression in response to NaCl (or in the *cdc14-3* mutant in the case of the last columns), according to the key. Cell-cycle classification of the genes (SPELLMAN *et al.* 1998) is shown to the right; cyclins are annotated to the left.

G. Average $\log_2$ expression change of genes shown in F, as described in E.


A
DAPI
GFP
WT 0 min
WT 10 min
cdc14-3 10 min
Percent Cells with Nuclear Hog1
WT
cdc14-3
Time after NaCl treatment
B
Relative Hog1P/Hog1
* p < 0.05
Time (min) after NaCl
C
pGST pCka2-GST
0 10 0 10 min after NaCl
Cka2-GST
GST
Cdc14
Hog1
Hog1P
D
pGST pCka1-GST
0 10 0 10 min after NaCl
Cka2-GST
GST
Cdc14
Hog1
Hog1P
E
Avg log2 Expression Change
WT
hog1Δ
cdc14-3
cka2Δ
* p < 1e-12
F
WT + NaCl vs t0
cdc14-3 + NaCl vs t0
cdc14-3 vs WT t0
M/G1 genes
G1 genes
S genes
S/G2 genes
G2/M genes
CLB6
CLB5
CLN1
CLN2
4X lower
4X higher
G
WT
cdc14-3
snf1Δ
cdc14-3 snf1Δ
* p < 5e-30

Second, the subnetwork predicts that Cdc14 regulates CK2 subunits to modulate the Hog1-regulated TF, Hot1. We uncovered a salt-enhanced interaction between Cdc14 and CK2 subunit Cka2 (Fig. 6C), and uncovered a constitutive association between CK2 subunits Cka1/Cka2 and Hog1 (Fig. 6C-D). Although the connection between CK2 and Hog1 was not known in yeast, our results are reminiscent of regulation in mammalian systems, in which CK2 is regulated by the human ortholog of Hog1, p38 (SAYED *et al.* 2000; HILDESHEIM *et al.* 2005; DE AMICIS *et al.* 2011; ISAEVA and MITEV 2011). As predicted by the subnetwork, we found that Cdc14, Cka2, and Hog1 were all required for normal induction of Hot1 targets (Fig. 6E).

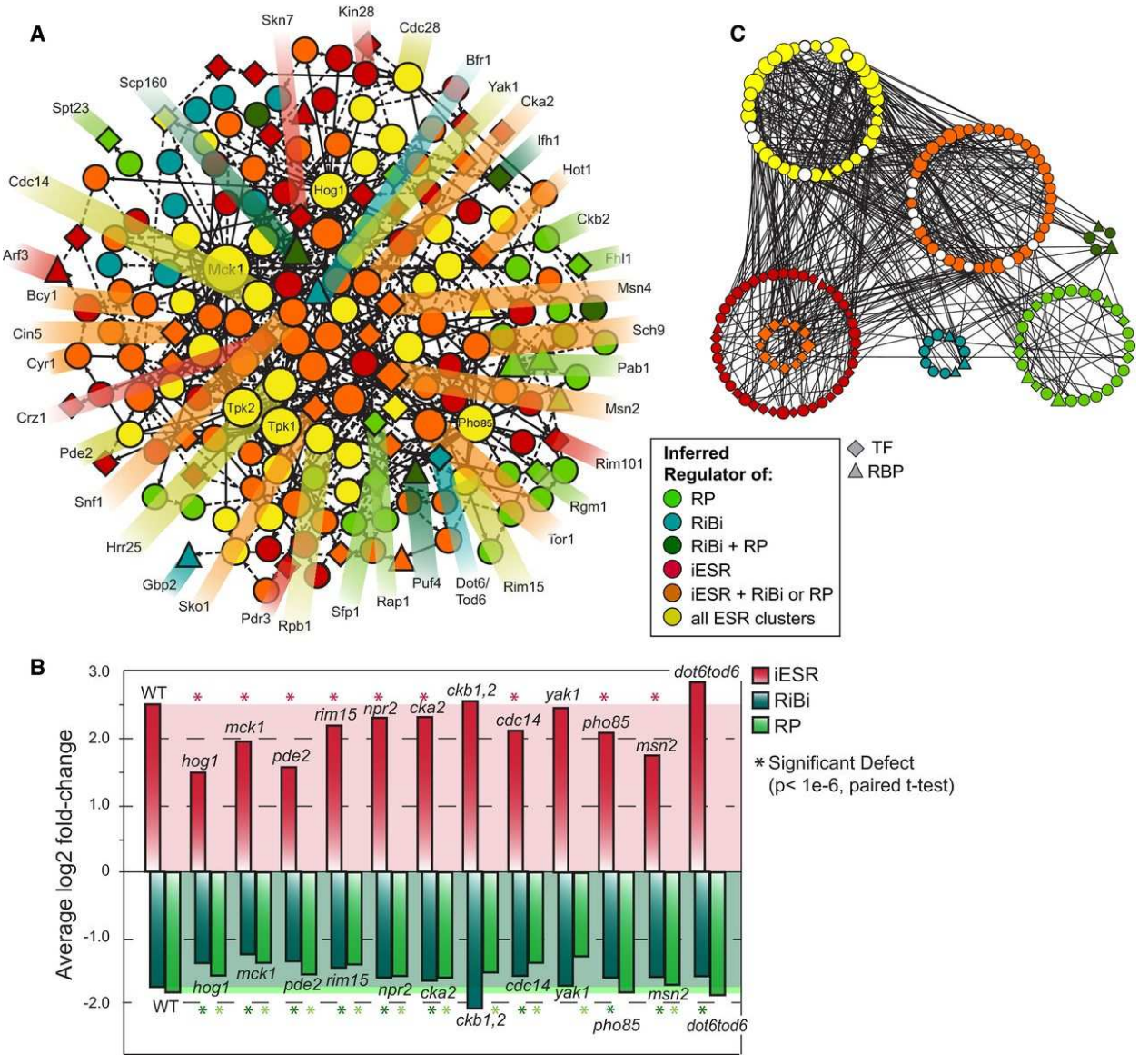
Finally, and surprisingly, we discovered that Cdc14 suppresses NaCl-dependent crosstalk to the cell-cycle network: the *cdc14-3* mutant at the nonpermissive temperature strongly and aberrantly induced G1 and S-phase genes upon NaCl treatment (Fig. 6F), even though cells were completely arrested in M-phase for the duration of the treatment. This included genes encoding G1 and S-phase cyclins *CLN1/2* and *CLB5/6*, respectively. To further understand this effect, we turned to the subnetwork: Cdc14 is predicted to affect these genes via direct interaction with the carbon-responsive kinase Snf1, which is known to be activated by NaCl (HONG and CARLSON 2007; YE *et al.* 2008). Snf1 is also required for proper timing of cell-cycle entry in standard conditions (PESSINA *et al.* 2010; BUSNELLI *et al.* 2013), raising the possibility that it is responsible for the inappropriate G1/S-gene induction in the absence of Cdc14. We found that deletion of *SNF1* in the *cdc14-3* background largely abrogated the hyper-activation of G1 and S genes in the *cdc14-3* mutant (Fig. 6G). This presents a model for future dissection, in which Cdc14 helps to suppress the cell-cycle effect of Snf1 activation, thereby funneling Snf1 activity toward its stress-specific gene targets. Together, our results demonstrate the remarkable and central role of Cdc14 in coordinating cellular signaling upon osmotic shock, while showcasing the predictive power of our inferred subnetwork.

Fig 7. Inferred ESR regulatory subnetwork.

A. Regulators predicted to lie upstream of iESR, RP, and/or RiBi ESR modules are color-coded according to the key and sized according to degree (number of connections). Diamonds and triangles represent TFs and RBPs, respectively.

B. Average $\log_2$ expression changes of iESR, RP, or RiBi genes in mutants responding to salt. Wild-type (WT) levels are highlighted by shaded areas; genes with a significant defect ($p<1e-6$, paired T-test) are indicated with an asterisk.

C. Same as A) but organized by ESR regulatory potential. Top-ranked bifurcation nodes discussed in the text are colored white. Nested orange nodes represent iESR TFs that are also predicted to lie upstream of RP paths.


A
Skn7
Kin28
Cdc28
Bfr1
Scp160
Yak1
Spt23
Cka2
Ifh1
Cdc14
Hot1
Hog1
Ckb2
Arf3
Mck1
Fhl1
Bcy1
Msn4
Cin5
Sch9
Cyr1
Pab1
Tpk2
Tpk1
Msn2
Crz1
Pho85
Pde2
Rim101
Snf1
Rgm1
Tor1
Hrr25
Gbp2
Puf4
Dot6/
Tod6
Rim15
Sko1
Pdr3
Rpb1
Sfp1
Rap1
C
Inferred
Regulator of:
RP
RiBi
RiBi + RP
iESR
iESR + RiBi or RP
all ESR clusters
TF
RBP
B
Average log2 fold-change
3.0
2.0
1.0
0
-1.0
-2.0
WT
hog1
mck1
pde2
rim15
npr2
cka2
ckb1,2
cdc14
yak1
pho85
msn2
dot6tod6
iESR
RiBi
RP
* Significant Defect
(p< 1e-6, paired t-test)

New insights into ESR regulation and coordination

We were especially interested in how distinct modules in the ESR – including iESR genes important for stress defense and RP/RiBi modules required for rapid growth – are regulated and coordinated. Of the 178 nodes implicated in ESR regulation, over half were predicted (Fig 7A) – and several confirmed (Fig 7B) – to lie upstream of all three ESR modules. In contrast to common upstream nodes that were enriched for kinases compared to the consensus subnetwork (p = 2.6e-7), nodes exclusive to iESR regulation were enriched for TFs (p = 5e-5), while rESR regulators showed a preponderance of RBPs (p = 1e-5), implicating regulated RNA stability for these genes. Many more regulators and regulatory connections were unique to the iESR versus RP and RiBi modules (the latter being the largest group) (Fig. 7C). This is consistent with the extensive redundancy in iESR control (GASCH 2002) and hints at a more monolithic regulation of rESR expression during times of adversity.

To better understand how cells coordinate repression of growth-related genes with induction of stress-defense genes in the ESR, we devised a bifurcation score based on information theory, to rank nodes that (a) are upstream of many genes from both modules but (b) have outgoing paths that relatively cleanly divide iESR and rESR genes. A third of top-15 ranked bifurcating nodes are linked to cAMP signaling (including adenylate cyclase Cyr1, cAMP response regulator Bcy1, and phosphodiesterase Pde2). Indeed, we found that the *pde2Δ* mutant has a defect in both iESR induction and rESR repression (Fig 7B), confirming the role of cAMP in the growth/stress-defense decision (see Discussion). Nearly half of the remaining top-ranked bifurcation proteins associate with RNA Pol II (including Pol II core subunit Rpb3, Pol II associated Sub1 and Ask10, transcription elongation factor Spt5, as well as Sds3 of the Rpd3L chromatin remodeling complex). Together with the identification of Pol II subunit Rpb1 as a hub in the subnetwork, these results implicate RNA Pol II at a key decision point in ESR coordination.

To investigate this, we started by checking *in vivo,* bulk modification of Rpb1-CTD in wild-type and *hog1Δ* cells responding to NaCl. The *hog1Δ* mutant showed an initial drop in Ser5 and Ser2 phosphorylation similar to the wild type, but displayed a reproducible defect in the normal subsequent transient increase in Ser5 and Ser2 phosphorylation (Fig 8A). The timing of the transient peaks in bulk Rpb1-CTD phosphorylation correlates with the timing of transcription initiation and elongation upon osmotic stress (BERRY and GASCH 2008; LEE *et al.* 2011; MILLER *et al.* 2011), consistent with the known roles of Hog1 as well as Ser 5 and Ser2 phosphorylation in these processes (ALEPUZ *et al.* 2003; PROFT *et al.* 2006; ZHANG *et al.* 2012).

To test our hypothesis that direct modification of Rpb1-CTD is important for ESR regulation, we measured transcriptomic changes upon salt stress in Rpb1-CTD mutant strains that could not be phosphorylated normally on CTD-Ser2 or -Ser5 (S2A and S5A mutants, respectively). Since S2A or S5A substitution in all CTD repeats is lethal, the mutant cells expressed chimeric CTD sequences with half mutant and half wild-type repeat sequences ((WEST and CORDEN 1995) see Methods). Neither mutant showed significant expression differences in the absence of stress; furthermore, the S2A mutant showed only a subtle defect in NaCl-dependent expression changes (Fig 8B). In contrast, the S5A mutant had a significant defect in iESR induction and even more so in rESR repression, comparable to the defect seen in the *hog1Δ* mutant.

We reasoned that aberrant ESR coordination may be caused by an inability of polymerase to re-localize from rESR genes, which are highly transcribed before stress, to stress-induced iESR genes. To test this, we measured chromatin occupancy of RNA Pol II subunit Rpb3 in both wild type and S5A mutant strains responding to NaCl stress, using ChIP-chip. In wild-type cells responding to stress, Rpb3 occupancy increased at iESR genes but decreased in the body of RP

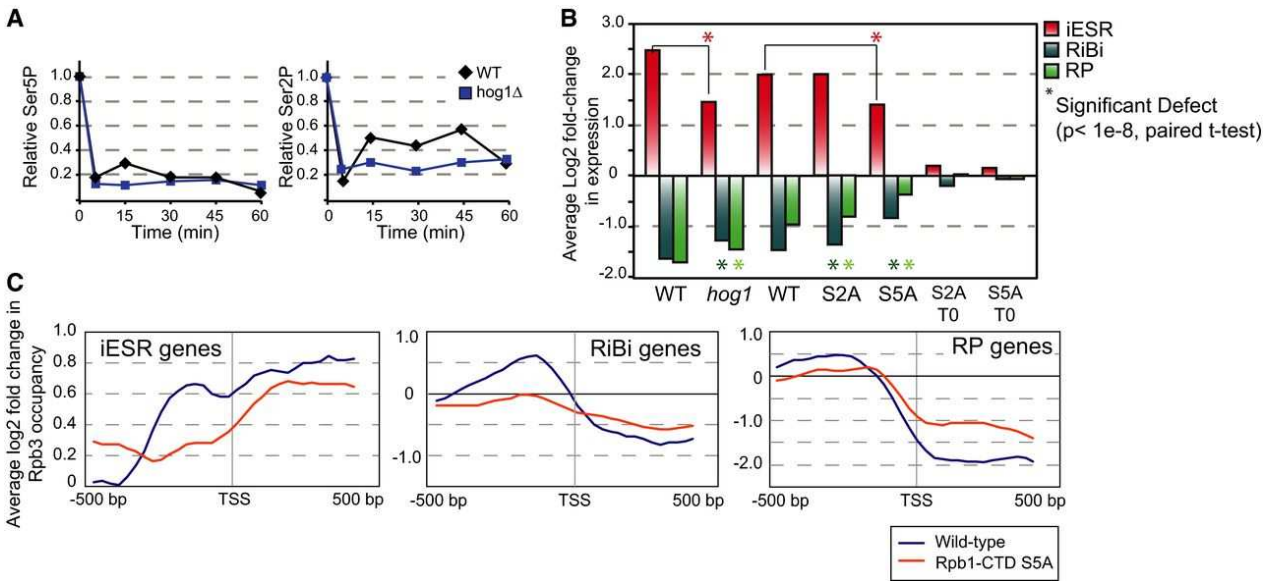


Fig 8. Pol II CTD modification coordinates ESR regulation.

A. Relative abundance of bulk Ser5P (left) and Ser2P (right) normalized to an internal Rpb3 loading control, from yeast cells treated with NaCl for the denoted times. Data are representative of several replicates.

B. The average log2(fold-change) expression of iESR, RP, and RiBi genes is shown for paired wild-type and hog1D strains (as in Fig 7) and paired wild-type, S2A, and S5 strains. There was little expression difference in the S2A and S5A mutants versus wild type before stress (right bars).

C. Average $\log_2$(fold-change) in Rpb3 occupancy 20 min after NaCl treatment at +/- 500bp around the transcription start site (TSS) of iESR (left), RiBi (middle), and RP (right) genes. Profiles were reproducible across biological replicates (see Appendix A Figure 7).

and RiBi genes, with slight accumulation in the promoter regions of specific repressed genes (Fig 8C and Appendix A Figure 7). In contrast, the S5A mutant showed a reproducible defect in Rpb3 recruitment to iESR genes and a concomitant defect in Rpb3 release from rESR genes (Fig 8C and Appendix A Figure 7). These results show that direct modification of the Rpb1-CTD is required for normal regulation of iESR and rESR genes (see Discussion).

The orthologous mammalian networks are enriched for growth-regulating and disease-causing genes

Striking the correct balance between growth rate and stress defense is fundamental for proper cellular function, and improper balance is thought to be a critical driver in diseases such as cancer (JONES and THOMPSON 2009). We therefore interrogated the set of human genes orthologous to the yeast NaCl subnetwork. We found that this set is enriched for genes linked to cancer, mostly through somatic mutation, according to the COSMIC database (FORBES *et al.* 2011): of the 35 human genes in the COSMIC dataset with yeast orthologs, 8 were orthologous to nodes in the consensus-node network, representing a 2.5-fold enrichment above chance (p = 0.0068, Dataset 4A). We also compared the yeast network to Mendelian disease genes in the OMIM database (HAMOSH *et al.* 2005). We identified 25 additional yeast genes whose orthologs are linked to heritable disease (Dataset 4B), with weak enrichment for genes associated with prostate cancer (p = 6e-3, (WOODS *et al.* 2013)). The network was also enriched for yeast proteins whose mouse orthologs are required for pre/perinatal viability, normal growth rate and body size, and male and female fertility (FDR < 5%) (WOODS *et al.* 2013). These results highlight that stress-responsive signaling is likely important for proper regulation of growth rate, and thus may provide insights into cancer biology (see Discussion).

Discussion

A major challenge in network biology remains integrating disparate large-scale datasets in a manner that reveals new insights to biology. The approach we developed here provides a new route to identifying the extensive set of players activated during a response, as well as the connections between them and the flow of information toward the processes they regulate.

The computational approach we developed provides several contributions. First, we provide a means to selectively integrate disparate datasets via four types of paths between proteins in the background network. Each dataset is prioritized separately by a series of objective functions, whereas related approaches for inferring signaling networks use a single objective function. One class of related methods essentially maximizes the number of paths between sources and targets (YEANG *et al.* 2004b; YEANG *et al.* 2005; OURFALI *et al.* 2007; GITTER *et al.* 2011). In contrast, our method's preference for sparse inferred subnetworks is also employed by approaches based on the prize-collecting Steiner Tree algorithm (HUANG and FRAENKEL 2009; YOSEF *et al.* 2009; HUANG and FRAENKEL 2012; HUANG *et al.* 2013) and flow-based algorithms (YEGER-LOTEM *et al.* 2009; LAN *et al.* 2011). However, those methods require the use of a weight parameter to trade off between subnetwork sparsity and the inclusion of known relevant proteins. Another contribution of our approach is the representation of uncertainty in the underlying network. We assign a confidence value to each protein and interaction according to its frequency in an ensemble of optimal inferred subnetworks. This is similar to the score used by Yeang et al. (2005), who actually enumerate all optimal solutions; doing so is practically intractable for our input and our model. In contrast, Ourfali et al (2007) assign confidence values based on the change in objective value when each protein or interaction is individually excluded, and Gitter et al. (2011) present several methods for ranking paths based on input experimental data and local topological features of the inferred subnetworks. The confidence values generated by our approach provide useful guidance for subsequent biological examination.

The resulting subnetwork put forward by our approach identified new regulators in the NaCl response and provides a glimpse of their connections in a single cellular signaling system. The extensive physical connectivity between what are traditionally considered 'distinct' pathways suggests much greater signaling integration than previously realized. The cross-connectivity between pathways, either direct or through apparent 'integration' points, may coordinate the magnitude or timing of signaling through distinct branches, prevent signaling crosstalk, and/or provide important feedback to dampen signaling as cells acclimate to new conditions (SCHWARTZ and MADHANI 2004; WALTERMANN and KLIPP 2010). Our results implicate Cdc14 as a critical integrator that bridges HOG and CK2 signaling, suppresses inappropriate activation of the cell-cycle network, and connects to several other pathways including Tor1, which is reportedly suppressed by Cdc14 (BREITKREUTZ *et al.* 2010). Members of the growth-regulating TOR1 pathway as well as the RAS/PKA pathway show the greatest connectivity to other stress-activated pathways, suggesting that growth regulation is sensitively tuned according to stress conditions.

Our results also shed new light on how growth control and stress defense are related. Optimal growth and maximal stress tolerance are competing interests in the cell: the fastest growing cells are typically the least tolerant of adversity, whereas stress-resistant cells are frequently slow growing or arrested (ELLIOTT and FUTCHER 1993; SUMNER and AVERY 2002; LU *et al.* 2009; ZAKRZEWSKA *et al.* 2011; LEVY *et al.* 2012). Our results here, along with prior studies, suggest that competition for cellular resources – namely those related to transcription and translation – drive the anti-correlated expression of genes involved in stress defense *versus* growth promotion. Under optimal growth conditions, rESR transcripts are among the most highly transcribed and the most highly translated (INGOLIA *et al.* 2009; LIPSON *et al.* 2009), consuming the bulk of cellular ribosomes (WARNER *et al.* 2001). We previously proposed that the drop in rESR transcripts helps to direct translational capacity to iESR genes by releasing sequestered

ribosomes (LEE *et al.* 2011). Work by You *et al.* suggests that cAMP abundance dictates whether translational capacity is directed to growth versus other processes such as stress defense (YOU *et al.* 2013). That our results implicate cAMP in the iESR/rESR regulatory balance is consistent with these models.

In addition to implicating cAMP metabolism, our results show that direct regulation of the RNA Pol II CTD plays a crucial role in the iESR/rESR transcriptional balance, by triggering redistribution of polymerase from highly transcribed growth-related genes to stress-induced defense genes. The ability to fully phosphorylate Ser5 of the Rpb1-CTD is required for normal repression of rESR genes, as indicated by the defect in transcript repression and Pol II redistribution, and is also required for normal recruitment of Rpb3 to iESR promoters for gene induction (Fig 8). Ser5 phosphorylation has been implicated in both gene repression and induction (HENGARTNER *et al.* 1998), in support of our findings. The stress-activated redistribution of Pol II from rESR to iESR genes is at least partly dependent on Hog1 (COOK and O'SHEA 2012; NADAL-RIBELLES *et al.* 2012), which we show phosphorylates the Rpb1-CTD *in vitro* (Fig 5D) and is required for its normal modification *in vivo* (Fig 8A). Thus, we propose that direct regulation of RNA Pol II, perhaps in part by the Hog1 kinase, plays a central role in coordinating these opposing transcriptional modules.

Establishing the correct balance between stress tolerance and growth rate is critical for surviving fluctuating environments in nature. But the enrichment for cancer-causing genes in the orthologous human subnetwork highlights the importance of this decision in disease biology, and it suggests that stress signaling in yeast may serve as a model for cancer signaling in humans. It is notable that orthologs of three key regulators in our network – Hog1, Cdc14, and CK2 – have all been implicated in regulating the mammalian tumor suppressor p53 (MEEK *et al.* 1990; BULAVIN *et al.* 1999; LI *et al.* 2000), which controls the growth/survival/apoptosis decision

in human cells and is mutated in many human cancers (CARVAJAL and MANFREDI 2013). These results underscore the importance of the growth/survival decision, and hint that the yeast subnetwork could be used to implicate as-yet unidentified human disease genes. An exciting area of future study will be to distinguish signaling dynamics and condition-specific versus common aspects of the signaling, with an eye toward their role in disease biology.

Materials and methods

Growth conditions

All strains were of the BY4741 background, primarily from the deletion collection (WINZELER *et al.* 1999) (Thermo Scientific, Waltham, MA), except for *cdc14-3* and its isogenic wild type (kindly provided by M. Weinreich (MILLER *et al.* 2009)). BY4741 *ckb1*□*ckb2*□ was kindly provided by C. Guthrie (BERGKESSEL *et al.* 2011). Knockout strains were verified by diagnostic PCR to ensure correct integration of the drug cassette and to confirm absence of the deleted gene. Unless otherwise noted, cells were grown to log phase in batch YPD cultures at 30°C for at least seven generations before addition of a final concentration of 0.7M NaCl, after which cells were grown for 30 min. *cdc14-3* and isogenic wild-type cells were grown at 25°C, shifted to the non-permissive temperature of 35°C for 90 minutes (or 120 min for experiments from Fig. 6G), and then treated with a final concentration of 0.7M NaCl at 35°C for an additional 30 min before sample collection. Relative physiological changes were compared to the timepoint collected immediately before addition of NaCl (*i.e.* 35°C for 90 min without NaCl).

Microarray analysis

Cell collection, RNA preparation, cDNA synthesis and labeling, array hybridization, and normalization were as previously described (BERRY and GASCH 2008; LEE *et al.* 2011), using cyanine dyes (Flownamics, Madison WI) and Superscript III (Life Technologies, Carlsbad, CA). Samples were hybridized to whole-genome tiled DNA microarrays (Roche Nimblegen, Madison,

WI), comparing cDNA from the salt-treated sample to cDNA generated from the unstressed culture. Dye orientation was performed on select samples to assess dye-specific biases; dye orientation for paired mutant-wild type samples was maintained for statistical analysis to avoid dye-specific effects. Comparison of unstressed strains was done as previously described (LEE *et al.* 2011) by retrieving and comparing single-channel data from mutant and wild type arrays. Array data are available through the NIH GEO accession # GSE60613.

Genes whose expression was altered in wild-type cells responding to NaCl were identified based on five biological replicates, using the Bioconductor package *limma* (SMYTH 2004) and Q-value (STOREY *et al.* 2005) to assess the false discovery rate (FDR) and taking q< 0.05 as significant. This analysis identified 5,056 genes with a significant change in expression in response to NaCl. Genes with a defect in NaCl-responsive expression in mutants shown in Fig. 1 were assessed in biological triplicate (for *hog1Δ , mck1Δ*, *pde2 Δ, msn2 Δ,* and *rim101Δ* strains) or duplicate (all other strains). Expression defects were identified using contrast matrices to wild-type expression in *limma* for triplicated samples, with q<= 0.025 taken as significant. For duplicated samples, expression defects were identified if both mutant replicates were outside the wild type mean + 2 standard deviations (95^{th} confidence level), based on 5 replicates of the wild-type samples. Identified targets (summarized in Table 1). Data for the *dot6 Δ tod6 Δ* mutant were taken from (LEE *et al.* 2011).

Validation experiments were performed on 10 deletion mutants responding to NaCl. For samples done in duplicate, significant expression changes were identified using *limma* with q<0.05 taken as significant. Expression defects from singleton experiments were identified based on a 1.5-fold difference in expression in that mutant versus the paired wild-type sample. Identified expression defects are summarized in Table 1. Expression in unstressed mutant cells was also assessed by comparing the mutant response to unstressed wild-type cells as described above. Unless

otherwise noted, we detected few expression differences in unstressed cells. For Figs. 6-8 where average expression values are plotted, data for some paired mutant-wild type experiments (namely *cdc14-3*, *dot6 Δtod6Δ*, and *ck2* mutants) were scaled such that their paired wild-type data matched the plotted wild-type data taken from other experiments, in order to accurately represent those mutant defects by accounting for day-to-day variation of paired samples.

Expression analysis for Fig 8 was done in strains generously provided by JL Corden (WEST and CORDEN 1995). Cells lacking endogenous *RPB1* carried a plasmid expressing *RPB1* with 14 wild-type CTD repeats (YSPTSPS) or a plasmid expressing chimeric *RPB1* genes: the Rpb1-CTD was composed of 5 repeats of S5A (YSPTAPS) followed by 7 wild-type-sequenced repeats in the so-called S5A mutant, or 8 S2A repeats (YAPTSPS) followed by 7 wild-type sequenced repeats in the S2A mutant. There was no difference in salt-responsive gene expression for control plasmids with 14 *versus* 21 wild-type repeats (not shown). Expression was measured as described above, before and at 30 min after treatment with 0.7M NaCl. There were few expression differences in the strains before stress (see Fig 8B).

Phospho-proteomic analysis

BY4741 was grown as described above, except that samples were taken before and at 5 min and 15 min after NaCl addition. Cells were lysed by three passages through the French press at 4°C in 3 mL of lysis buffer consisting of 50 mM Tris pH 8, 4M urea, 75 mM NaCl, 1 mM DTT, complete mini EDTA-free protease inhibitor (Roche Diagnostics, Indianapolis, IN), and phosSTOP phosphatase inhibitor (Roche Diagnostics). The lysate was centrifuged at 14000 r.p.m. for 10 min and the protein concentration determined by a bicinchoninic acid assay. Cysteine residues were reduced and alkylated by incubating lysate with 5 mM DTT for 45 min at 37°C followed by incubation in 15 mM IAA for 45 min at room temperature in the dark. After adding an additional aliquot of DTT to cap the alkylation reaction, the urea concentration was diluted to a final

concentration of 1M with 50 mM Tris and 1 mM $CaCl_2$. Proteins from each time point were digested overnight (37°C, pH 8) with trypsin (Promega, Madison, WI) at an enzyme:substrate ratio of 1:50. TFA was added to a final concentration of 0.5% to quench each digest and the resulting peptides were desalted via solid phase extraction on a 50-mg tC_{18} SepPak cartridge (Waters, Milford, MA) and the eluant lyophilized.

The desalted peptides from each time point were each labeled with a different tandem mass tag (TMT) isobaric label (Thermo-Pierce, Rockford, IL) according to the manufacturer's instructions. The differentially labeled TMT samples were pooled in equal volumes and dried down. Labeled peptides were fractionated by strong cation exchange (SCX) on a polysulfoethyl A column (9.4 mm x 200 mm; PolyLC) with mobile phases A: 5 mM KH_2PO_4 pH 2.65 and 30% acetonitrile; B: 5 mM KH_2PO_4 pH 2.65, 350 mM KCl, and 30% acetonitrile; C: 5 mM KH_2PO_4 pH 6.5 and 500 mM KCl; D: Water. The gradient was generated by a Surveyor LC quaternary pump (Thermo Scientific, Waltham, MA) at 3 mL/min flow rate. Peptides were eluted over the following gradient and detected via a PDA detector (Thermo Scientific): 0-2 min, 100% A; 2-5 min, 0-10% B; 5-41 min 10-100% B; 41-48 min 100% B; followed by washes with C and D prior to re-equilibration with mobile phase A. Fifteen fractions were collected, lyophilized, and desalted. A small portion, 5%, of each was retained for unmodified protein analysis and the remaining material used for phosphopeptide enrichment.

Each fraction was enriched for phosphopeptides using immobilized metal ion affinity chromatography (IMAC). Magnetic beads (Qiagen, Valencia, CA) were washed 3 times with water, incubated with 40 mM EDTA (pH 7.5) for 30 min, and washed with water again. The beads were then incubated with 100 mM $FeCl_3$ for 30 min and washed 4 times with 80% acetonitrile and 0.1% TFA. Peptides from each fraction were resuspended in 1 mL of 80% acetonitrile and 0.1% TFA and incubated with the beads for 30 min. Unbound peptides were removed from the beads by

washing 4 times with 80% acetonitrile and 0.1% TFA. Phosphopeptides were eluted using 1:1 acetonitrile:5% NH_4OH in water, immediately acidified with 4% formic acid, and lyophilized.

Phosphopeptide enriched and protein fractions were resuspended in 0.2% formic acid and analyzed by reverse phase liquid chromatography on a nanoAcquity LC (Waters) coupled to an ETD-enabled LTQ Orbitrap Velos (Thermo Scientific). Samples were first loaded onto a 10-cm, 75 µm i.d. pre-column packed with 5 µm C18 particles (Bruker-Michrom, Fremont, CA) in 98% A (0.2% formic acid in water), 2% B (0.2% formic acid in acetonitrile) and then separated across a 25-cm, 50 µm i.d. analytical column packed with 5 µm C18 particles (Bruker-Michrom) using the following gradient: 0-3 min, 2-5% B; 3-123 min, 5-35% B; 123-133 min, 35-70% B; 133-138 min, 70% B; 138-165 min, 2% B. Phosphopeptide and protein fractions were each analyzed in duplicate. Methods to acquire mass spectra started with one MS1 survey scan (R = 30000, 300-1500 Th) followed by data dependent MS2 fragmentation and analysis (R = 15000) of the ten most intense precursors. The exclusion duration was 60 seconds for -0.55 Th to +2.55 Th of the sampled precursor. Ions with an unassigned charge state or a single charge were excluded. The QuantMode instrument control method was employed to reduce reporter ion interference caused by co-isolation of multiple precursors (WENGER *et al.* 2011a).

Spectral reduction was performed using DTA Generator. Generated text files were searched for fully tryptic peptides with up to three missed cleavages against a UniProt target-decoy database populated with yeast plus isoforms (downloaded 29 July 2011) using the Open Mass Spectrometry Search Algorithm (OMSSA) (GEER *et al.* 2004). Carbamidomethylation of cysteine (+57.021464), TMT 6-plex on lysine (+229.162932), and TMT 6-plex on peptide N-terminus (+229.162932) were searched as fixed modifications for all samples. Phosphopeptide-enriched fractions were additionally searched for variable phosphorylation modifications. Search results were filtered to 1% FDR at the unique peptide level and identified peptides quantified within the

COMPASS software suite (WENGER *et al.* 2011b). Peptides were grouped into proteins according to previously reported rules and filtered to 1% FDR (NESVIZHSKII and AEBERSOLD 2005). Protein quantification was performed by summing all reporter ion intensities within each channel for each non-phosphorylated peptide mapping uniquely to that protein group.

Phosphorylation events were localized to specific residues using probabilistic methods (PHANSTIEL *et al.* 2011). Localized phosphorylated peptides were grouped together by identical modification sites and their reporter ion intensities were summed. For simplicity, phosphorylation isoforms are referred to as phospho-sites. The average of two technical replicates was taken per time point, and phospho-sites with at least two-fold change in recovery were taken as significant for downstream analysis.

Immunoprecipitation analysis

BY4741-*cka2Δ* and *cka1Δ* cells were transformed with empty vector or plasmids encoding GAL-inducible Cka2-GST or Cka1-GST (ZHU *et al.* 2001; SOPKO *et al.* 2006)(Thermo Scientific), respectively. Cells were grown in YP-2% galactose medium in log phase to 0.6-0.8 OD_{600}, subjected to osmotic stress (0.7N NaCl) for the indicated length of time, and lysed by bead-beating on ice. Cell lysates were incubated with glutathione-sepharose beads (GE Healthcare) at 4°C overnight in 1X PBS buffer with 1 mM DTT, 0.1% NP40, 10% glycerol and protease inhibitors (Millipore, Billerica, MA). Proteins were eluted with 1X SB buffer and resolved by SDS-PAGE and detected by immunoblotting. Antibodies used were goat polyclonal anti-Hog1 (Santa Cruz Biotech, Dallas, TX), rabbit polyclonal anti-phospho-p38 MAPK (Cell Signaling), mouse monoclonal anti-actin (Pierce Biotech), goat polyclonal anti-Cdc14 (Santa Cruz Biotech) and goat-polyclonal anti-GST (Abcam, Cambridge, MA). All blots shown in the manuscript are representatives of at least biological duplicates.

Microscopy

Harvested cells were fixed with 4% final concentration of formaldehyde for 15 min, and GFP was visualized on a Leica DM LB2 microscope with standard GFP filters. DNA was detected via cell staining with 1 µg/ml DAPI for 5 min. Viability of *cdc14-3* cells was measured with Live-Dead staining (Life Technologies), which showed that NaCl-treated *cdc14-3* maintained viability close to WT cells for over 30 min after treatment with 0.7M NaCl (not shown). Nuclear Hog1 was scored by visual inspection by comparing GFP signal to DAPI signal, in at least 75-100 cells per sample.

In vitro CTD phosphorylation

Cells expressing C-terminally TAP-tagged proteins (GHAEMMAGHAMI *et al.* 2003)(Thermo) were exposed to NaCl for the denoted times, snap frozen, and then cryo-lysed with a Retsch Mixer Mill MM 400 as described in (CHURCHMAN and WEISSMAN 2011). Ground yeast was added to TAP Buffer A, and TAP-tagged kinase was purified as described in (PUIG *et al.* 2001; LIU *et al.* 2004), with minor modifications. Kinases were eluted overnight at 4°C in 25 µL TAP Buffer A with 1mM DTT and 10 U AcTEV (Invitrogen).

Peptide substrate GST-CTD14 (fourteen repeats of YSPTSPS fused to GST), was purified essentially as described in (PATTURAJAN *et al.* 1998). Before elution, glutathione-sepharose beads were resuspended in 1 mL FastAP buffer (10 mM Tris-HCl pH 8.0, 5 mM $MgCl_2$, 100 mM KCl, 0.02% TritonX-100, 100 ug/mL BSA) and incubated with 100 U FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific) for 1 hour at 37°C to remove any phosphates placed by the bacteria. Beads were washed, and GST-CTD14 was eluted. Any remaining alkaline phosphatase was heat inactivated at 75°C for 5 min. The concentration of GST-CTD14 was determined via Bradford assays.

In vitro kinase assays were performed in at least biological duplicate using 5 µL of tandem affinity purified (TAP) kinase and 3 µM GST-CTD14 in 30 µL Buffer D as described in (ANSARI *et al.* 2005), with minor modifications. For Hog1 inhibition assays, the kinase was pre-incubated with the inhibitor 4-(4-fluorophenyl)-2-(4-methylsulfinylphenyl)-5-(4-pyridyl)imidazole (Cell Signaling Technology) for 10 min prior to the reaction. Reactions were performed at 30°C for two hours and resolved via SDS-PAGE and Western analysis using antibodies targeting CTD-Ser2P (Bethyl Laboratories), CTD-Ser5P (clone 3E8, gift from Dirk Eick), or the TAP tag (Thermo Scientific). Quantitation was performed using ImageJ. All images are representative of several biological replicates. The plot in Fig. 5D shows background-subtracted levels of Ser2P and Ser5P normalized to Hog1-TAP abundance in each lane, then referred to levels seen in unstressed cells to calculate fold-change in phosphorylation.

Analysis of novel predicted salt-response regulators

Fourteen predicted regulators not previously known to respond to NaCl were chosen for validation analysis. NaCl-responsive gene expression was measured in ten mutants, focusing on kinases and phosphatases not known to respond to NaCl and two RBPs (Scd6 and Arf3), as described above. Data for Rpd3, Bem1, Gal11, and Tpk2 were taken from previous studies probing the osmotic response (ALEJANDRO-OSORIO *et al.* 2009; GITTER *et al.* 2013), taking $q < 0.05$ from *limma*-q value analysis as significant. Mutants were considered to have a defect in NaCl-dependent expression if there were at least fifty affected genes. Overlap between measured and predicted target genes was based on the hypergeometric test, scoring the probability of getting the number of observations or more compared to random expectation from the 3,330 genes used for IP input. Genes affected in unstressed cells were also identified (see above) and compared to predicted genes in cases where the NaCl-measured targets did not significantly overlap with predicted targets.

We also assessed the connections predicted between the interrogated regulators and other nodes predicted to lie in their paths. From the nodes predicted to lie in each regulator's path (based on the consensus-paths network), we identified those with known downstream targets (*e.g.* TF and RPBs) or targets measured in this study. We then scored the enrichment of each predicted node's known targets within the measured targets of the interrogated regulator, taking $p<1e-6$ from the hypergeometric test as significant. Because the test lacks statistical power for large gene groups, we scored enrichment against the total list of measured targets as well as induced and repressed targets with defective or amplified expression changes considered separately. The results (Appendix A Table S2) indicate a lower bound of supported in-path nodes, since the hypergeometric test has lower statistical power for small gene groups (including known targets of several regulators), and targets of several in-path nodes were marginally enriched ($1e-5 < p < 0.01$) among measured targets of interrogated regulators but did not meet our stringent threshold. It is also possible that regulators that serve redundant roles are difficult to score with our assay, since single-gene knockouts may not identify all of the downstream targets.

Chromatin immunoprecipitation (ChIP)

ChIP was done similarly to as described in (TIETJEN *et al.* 2010), on cells before and 20 min after treatment with 0.7M NaCl. Rpb3 was immunoprecipitated using anti-Rpb3 antibody W0012 (Neoclone, Madison, WI) in strain Z26 carrying "wild-type" or "S5A" *RPB1* gene expressed on a CEN plasmid (WEST and CORDEN 1995), described above. Chromatin was sonicated on a Misonix 4000 machine (Qsonica, Newtown, CT), input and immunoprecipitated material was amplified using ligation-mediated PCR as previously described (TIETJEN *et al.* 2010) and hybridized to tiled Nimblegen arrays designed against the yeast genome (LEE *et al.* 2011). Data were normalized as in (TIETJEN *et al.* 2010), except without the baseline adjustment procedure. All two-color arrays from two biological replicates were quantile

normalized together before further analysis. This procedure did not change any of the trends reported in the manuscript but helped to adjust the baseline across biological replicates done on different days. ChIP-chip data area available in the NIH GEO database under accession # GSE60613

Ortholog analysis

To assess the relationship of the yeast consensus network ensemble to human diseases, we analyzed the orthologous set of human genes. We used the stringent RSD method of ortholog assignment (WALL *et al.* 2003), using a BLAST Evalue cutoff of 1e-5 and requiring fewer than 20% gapped positions in the global alignment. The method identified 2381 yeast-human orthologas; we focused on the 1619 of these genes that are reviewed in humans. We compared these genes to those annotated in the COSMIC v67 (FORBES *et al.* 2011) and OMIM (HAMOSH *et al.* 2005) databases. We also analyzed orthologous mouse proteins using the phenology.org database (WOODS *et al.* 2013).

NETWORK INFERENCE METHODS

Background network for IP method

To construct the background network, we identified a variety of binary interactions that are relevant to intracellular signaling and gene-expression regulation. The background network, gathered from numerous public databases, represents interactions between pairs of proteins (PTACEK *et al.* 2005; PU *et al.* 2009; FASOLO *et al.* 2011; SHARIFPOOR *et al.* 2011; HEAVNER *et al.* 2012), including kinase-substrate interactions, as well as protein-DNA interactions (GUELZIM *et al.* 2002; MACISAAC *et al.* 2006; EVERETT *et al.* 2009; NI *et al.* 2009; ABDULREHMAN *et al.* 2011; VENTERS *et al.* 2011; HUEBERT *et al.* 2012) and protein-RNA interactions (HOGAN *et al.* 2008; SCHERRER *et al.* 2010). After manual inspection of the background network neighborhoods of

the interrogated mutants, we added a set of 17 missing interactions between the mutants and nearby regulators based on known interactions in the literature.

While the types of biological interactions in the background network are rich and diverse, we use a simplified representation as input to the computational method (illustrated in Fig. 2 and 3A). The background network is represented as a graph, in which nodes represent genes and gene products, and edges represent interactions. A gene may be represented as two separate nodes in the background network: one representing the protein, and, for targets, one representing the DNA or mRNA. Each interaction may have a direction: for example, transcriptional regulatory interactions are directed, but most protein-protein interactions are not. The provenance of the background network and the interactions themselves are provided in Appendix A Information 1.2.3, Appendix A Table 2.

IP method input data and candidate paths

The primary goal of the IP approach is to provide explanations for the salt-specific transcriptomic changes measured for this article. We also use two additional sources of salt-specific experimental data. From these data, we generate directed, acyclic candidate paths that serve as input to the IP (Fig. 3B):

Source-target pairs and paths, source-source paths. From the transcriptomic data measured in each of the original signaling mutants, we identified the set of downstream genes with dysregulated salt-responsive expression. We then extracted what we refer to as *source-target pairs*, each consisting of a single *source* protein and a *target* gene that was dysregulated in the source mutant under salt stress. Next, for each source, we used the hypergeometric test to identify candidate transcription factors (TFs) and RNA-binding proteins (RBPs) whose known binding targets (promoters or transcripts, respectively) are significantly enriched with the genes

represented by the source's targets ($p < 0.05$). We also include TFs that are known to bind any number of targets under osmotic stress (NI *et al.* 2009; HUEBERT *et al.* 2012). Candidate source-target paths were enumerated to connect signaling mutants to their gene targets via candidate TFs/RBPs with up to three intermediate proteins between the source and TF/RBP (for a total of five interactions). Candidate path enumeration for each source was performed in an iterative deepening procedure, which was stopped at the path length at which at least 50% of candidate TFs/RBPs were reached.

Fitness-contribution hits and hit-source paths. Previously, we identified yeast mutants that conferred a defect in acquired stress resistance after salt pretreatment (BERRY *et al.* 2011). We refer to the gene products represented by these mutants as *fitness-contribution hits* because of the mutation's negative effect on yeast fitness under salt stress. *Candidate hit-source paths* and source-source paths were generated by finding short paths (including at most one intermediate protein) between these hits and the source proteins, and between pairs of source proteins. These paths are useful for interpreting the fitness-contribution hits in terms of connections to known regulators.

Phospho-proteomic hits. We use this name to refer to the proteins that showed differential phosphorylation under salt stress.

Receptor-source paths. Our method can take advantage of domain knowledge about the salt stress response in order to provide a scaffold for the inferred subnetwork. Here, we wanted to capture the most-upstream stress sensors that may otherwise be missed in connecting sources to their downstream targets. We identified well-known indirect relationships between two transmembrane receptors, Sln1 and Sho1, and one of the sources, Hog1 (SAITO and

TATEBAYASHI 2004). We enumerated candidate receptor-source paths (up to four intermediates) from Sln1 to Hog1 and Sho1 to Hog1, and provided them as input to the IP method.

Further details on the data and the generation of candidate paths are available in Appendix A Information 1.2.4. To measure the contribution of each input data set, we ran computational experiments in which each component was held aside. We also tested the effect of varying the length of the candidate paths. The results of these experiments are available in Appendix A Information Section 2.3 and 2.4.

IP notation and variables

The salt-specific signaling subnetwork is inferred by solving an integer linear program (IP, for short). We encode the relevance of each node, edge (physical interaction), and candidate path, and the direction of each edge, as binary variables. We characterize possible subnetworks using a set of linear constraints over those binary variables. Subnetwork inference is performed by choosing a union of relevant, directed paths that together satisfy our constraints and optimize a series of successively applied objective functions.

The values of some variables were determined by data provided as input to the inference process (for example, directions of directed edges), while others are inferred by solving the IP.

Notation (summarized in Appendix A *Table 3.)* The input to the method is represented as a graph of nodes $\mathcal{N}$, edges $\mathcal{E}$, and candidate paths $\mathcal{P}$. A node represents either a protein or a target gene/mRNA. Protein nodes may belong to one or more of the following subsets: sources $\mathcal{N}^S$, fitness-contribution hits $\mathcal{N}^F$, phospho-proteomic hits $\mathcal{N}^P$, and known membrane receptors $\mathcal{N}^R$. The set $\mathcal{N}^T$ describes targets, and for a given source node n, $\mathcal{N}^T(n)$ is the set of its targets.

The set of edges is $\mathcal{E} = (\mathcal{E}^D \cup \mathcal{E}^U)$, where $\mathcal{E}^D$ is the set of directed edges and $\mathcal{E}^U$ is the set of undirected edges. We denote an edge e between nodes n_i and n_j as $e = (n_i, n_j)$. $\mathcal{N}(e)$ refers to the nodes connected by a particular edge e, and $\mathcal{E}$(n) refers to the edges that touch a particular node n.

We consider four subsets of candidate paths $\mathcal{P}$: source-target paths between sources and their targets $\mathcal{P}^{ST}$, hit-source paths between fitness-contribution hits and sources $\mathcal{P}^{FS}$, source-source paths $\mathcal{P}^{SS}$, and receptor-source paths P^{RS} that connect known receptor proteins to sources. (Phosphoproteomic hits and additional fitness-contribution hits may appear in any of these paths.) To refer to the paths between a specific source s and target t, we use the notation $\mathcal{P}^{ST}(s,t)$. We use the same notation to refer to other kinds of paths with specific endpoints: $\mathcal{P}^{FS}(f,s), \mathcal{P}^{SS}(s_i, s_j), \mathcal{P}^{RS}(r,s)$.

Each path $\mathcal{P}$ specifies a direction for each of its undirected edges e, which is denoted as $dir(p,e)$. $\mathcal{E}(p)$ and $\mathcal{N}(p)$ refer to the edges and nodes in a particular path p.

Variables (summarized in Appendix A Table 4). The predicted relevance of a path $\mathcal{P}$, is represented with the variable σ_p which takes the value 1 if the path is included in the inferred subnetwork, and 0 if it is not. As many as two variables describe each edge. The predicted relevance of an edge e, is represented with the variable x_e, which takes the value 1 if the edge is in at least one relevant path. For undirected edges in the background network, the variable d_e represents the inferred direction of the edge. Each node n has one variable: y_n, representing whether or not the node is present in any relevant paths. Finally, for all pairs of sources (n_i, n_j), and also for all pairs consisting of one source and one fitness-contribution hit, the variable c_{ij}

represents whether or not the relevant subnetwork provides a directed path between the two nodes in the pair.

IP constraints

The following linear constraints define a subnetwork that, at minimum, provides consistently directed paths between source-target pairs and receptor-source pairs. Additional constraints are used to count up the number of connected fitness-contribution hit-source pairs and source-source pairs. These counts are optimized during the optimization procedure.

Provide at least one path between each source-target pair. Each source must be connected to each of its targets by at least one relevant path. The following constraint requires that, for each source s, for each of its targets t, at least one source-target path p in $\mathcal{P}^{ST}(s,t)$ from s to t must have $\sigma_p = 1$.

$$\sum_{\text{source-target paths } p \text{ in } \mathcal{P}^{ST}(s,t)} \sigma_p \geq 1 \qquad \text{for all sources } s \text{ in } \mathcal{N}^S, \text{ targets } t \text{ in } \mathcal{N}^T(s) \quad (1)$$

Provide at least one path between each receptor-source pair. We must provide at least one path showing the indirect relationship between an upstream receptor and a source. Similar to the previous constraint, this one requires that for each receptor r and each of its downstream sources s there must be at least one receptor-source path p in $\mathcal{P}^{RS}(r,s)$ for which $\sigma_p = 1$.

$$\sum_{\text{receptor-source paths } p \text{ in } \mathcal{P}^{RS}(r,s)} \sigma_p \geq 1 \qquad \text{for all receptors } r \text{ in } \mathcal{N}^R, \text{ sources } s \text{ in } \mathcal{N}^S(r) \quad (2)$$

Record whether or not there is a path between each fitness-contribution hit-source pair and source-source pair. Rather than require that each of these pairs is connected, we use the

optimization procedure to maximize the total count of connected pairs. We use the following constraints to count up the number of connected pairs.

If there is a path between a fitness-contribution hit f and a source s, set the variable $c_{fs} = 1$. Otherwise, set $c_{fs} = 0$:

$$\sum_{\substack{\text{hit-source paths} \\ p \text{ in } \left(P^{FS}(f,s) \cup P^{FS}(s,f)\right)}} \sigma_p - c_{fs} \geq 0 \qquad \text{for all fitness-contribution hits } f \text{ in } \mathcal{N}^F, \text{ sources } s \text{ in } N^S \tag{3}$$

$$c_{fs} - \sigma_p \geq 0 \qquad \text{for all fitness-contribution hits } f \text{ in } \mathcal{N}^F, \text{ sources } s \text{ in } \mathcal{N}^S, \text{ hit-source paths } p \text{ in } \mathcal{P}^{FS}(f,s) \cup \mathcal{P}^{FS}(s,f) \tag{4}$$

Similarly, if source s_i is connected to source s_j, we set $c_{ij} = 1$. Otherwise, set $c_{ij} = 0$.

$$\sum_{\substack{\text{source-source paths} \\ p \text{ in } \left(\mathcal{P}^{SS}(s_i,s_j) \cup \mathcal{P}^{SS}(s_j,s_i)\right)}} \sigma_p - c_{ij} \geq 0 \qquad \text{for all pairs of sources } (s_i, s_j) \text{ in } \mathcal{N}^S \times \mathcal{N}^S \tag{5}$$

$$c_{ij} - \sigma_p \geq 0 \qquad \text{for all pairs of sources } (s_i, s_j) \text{ in } \mathcal{N}^S \times \mathcal{N}^S, \text{ source-source paths } p \text{ in } \mathcal{P}^{SS}(s_i, s_j) \cup \mathcal{P}^{SS}(s_j, s_i) \tag{6}$$

All edges in a relevant path are relevant. For an edge e to be relevant (that is, have $x_e = 1$), there must be at least one relevant path that contains it (that is, a path $\mathcal{P}$ for which $\sigma_p = 1$). Similarly, a relevant path $\mathcal{P}$ must contain all relevant edges . The set $\mathcal{P}(e)$ refers to the paths that contain edge e.

$$\sum_{paths\ p\ in\ \mathcal{P}(e)} \sigma_p - x_e \geq 0 \qquad \text{for all edges } e \text{ in } \mathcal{E} \qquad (7)$$

$$x_e - \sigma_p \geq 0 \qquad \text{for all paths } p \text{ in } \mathcal{P}, \text{ edges } e \text{ in } \mathcal{E}(p) \qquad (8)$$

All nodes in a relevant edge are relevant. A node n is relevant if it is connected to a relevant edge e (where $x_e = 1$). Each node n for a relevant edge e must be relevant ($y_n = 1$).

$$\sum_{\text{edges } e \text{ in } \mathcal{E}(n)} x_e - y_n \geq 0 \qquad \text{for all nodes } n \text{ in } \mathcal{N} \qquad (9)$$

$$y_n - x_e \geq 0 \qquad \text{for all edges } e \text{ in } \mathcal{E}, \text{ nodes } n \text{ in } \mathcal{N}(e) \qquad (10)$$

All paths must be uniquely directed. For a relevant path p, all undirected edges e in that path (e in $\mathcal{E}(p) \cap \mathcal{E}^U$) must be uniquely oriented so that the path proceeds only in one direction. This required direction for each edge is determined when the candidate path is generated, and is given by $dir(p, e)$. (For source-target paths, the required direction allows the path to proceed from the source to the target.) The term including $I(\cdot)$, the indicator function, returns 1 if an edge's inferred direction corresponds to the direction that the path requires for it.

$$I\big(d_e = dir(p, e)\big) - \sigma_p \geq 0 \qquad \text{for all paths } p \text{ in } \mathcal{P}, \text{ undirected edges } e \text{ in } \mathcal{E}(p) \cap \mathcal{E}^U \qquad (11)$$

Solving the IP to find an ensemble of subnetworks

An optimal inferred subnetwork satisfies two goals: maximizing the inclusion of salt-response-relevant proteins that are supported by experimental evidence, and minimizing the number of additional nodes that are necessary for connecting each source to each target. To achieve this,

we apply four successive objective functions. To accompany the following description, a diagram of the process is depicted in Appendix A Fig. 8.

To model and solve the IP, we used the GAMS modeling system v. 23.9.3 and the ILOG CPLEX solver v. 12.4.0.1. Both are commercial packages for which an academic license available at a reduced cost.

Step 1: Maximize connections between hits and sources. This involves solving the IP to identify max_connections, the maximum number of connections possible between pairs of sources, and between pairs of fitness-contribution hits and sources. The purpose of this step is to reveal proximal connections between salt-responsive proteins, whether or not they occur between sources and targets. In this constraint, the set $(\mathcal{N}^S \times \mathcal{N}^S) \cup (\mathcal{N}^F \times \mathcal{N}^S)$ gives all source-source pairs and fitness-contribution-hit-source pairs, and the sum counts up the number of pairs that are connected by relevant paths.

$$\text{max\_connections} = \max \sum_{\text{hit-source pairs } (n_i,n_j) \text{ in } (\mathcal{N}^S\times\mathcal{N}^S)\cup(\mathcal{N}^F\times\mathcal{N}^S)} c_{ij} \tag{12}$$

After optimizing this criterion, we add a new constraint to the IP:

$$\sum_{\text{hit-source pairs } (n_i,n_j) \text{ in } (\mathcal{N}^S\times\mathcal{N}^S)\cup(\mathcal{N}^F\times\mathcal{N}^S)} c_{ij} = \text{max\_connections} \tag{13}$$

Step 2: Maximize inclusion of fitness and phospho hits. Next, we solve the IP to identify max_hits, the maximum number of fitness-contribution hits and phosphor-proteomic hits that

can be included in the relevant subnetwork. This step prioritizes the use of nodes with experimental evidence of being relevant to the salt stress response.

$$\text{max\_hits} = \max \sum_{\text{nodes } n \text{ in } (\mathcal{N}^F \cup \mathcal{N}^{\mathcal{P}})} y_n \tag{14}$$

After identifying the maximal number of hits that can be included in the subnetwork, we add a new constraint to the IP:

$$\sum_{\text{nodes } n \text{ in } (\mathcal{N}^F \cup \mathcal{N}^{\mathcal{P}})} y_n = \text{max\_hits} \tag{15}$$

Step 3: Minimize total nodes and find multiple solutions. Now we solve the IP with a new objective function, which minimizes the number of nodes required to satisfy all of the constraints. The resulting subnetwork will include only those nodes that are required to explain the experimental data.

$$\min \sum_{\text{nodes } n \text{ in } \mathcal{N}} y_n \tag{16}$$

At this point, we find an ensemble of solutions to the IP, where each solution identifies a minimum set of nodes (while still satisfying all other constraints). The CPLEX solver allows for the identification of multiple solutions. First, the CPLEX solver uses a branch-and-cut algorithm to find one optimal solution; this algorithm entails maintaining a tree of linear relaxations of the IP. Next, the solver proceeds down previously rejected branches of the tree to identify additional optimal solutions with different variable settings. For our experiments, we identified 10,000 solutions.

Step 4: Maximize the number of paths in each solution. After predicting the relevant nodes in the previous step, we would like to see all possible relevant connections between them, to aid in their interpretation. For each of the solutions identified in the previous step, we solve the IP again to maximize the number of relevant directed paths between the nodes included in the solution. This step does not change the node content of each solution, but instead reveals all possible directed paths that connect the node set chosen in the previous step.

For each solution:

First, we introduce constraints to fix each value of y_n to its value from the previous solution, $\widehat{y_n}$:

$$y_n = \widehat{y_n} \text{ for all nodes } n \tag{17}$$

Next, we maximize the number of relevant paths:

$$\max \sum_{\text{paths } p \text{ in } \mathcal{P}} \sigma_p \tag{18}$$

At this point, we assemble the solutions into an ensemble of inferred subnetworks. Using the ensemble, we assign a confidence value for a prediction based on the number of solutions in the ensemble that support the prediction. We performed several experiments to assess the effect of each component of our four-part objective function, as well as their ordering. These can be found in Appendix A Information Sections 2.3 and 2.5.

Precision-recall analysis

To assess the predictive accuracy of the ensemble (as shown in Fig. 4B), we curated a list of true positives and a list of likely negative proteins. True positives were defined as genes previously identified in the Hog network based on literature curation (DE NADAL and POSAS 2010; TIGER *et al.* 2012), genes with 'osmotic' or 'osmolarity' in their *Saccharomyces* Genome Database (SGD) (CHERRY *et al.* 2012) annotations, and genes with 'stress regulator' in their SGD annotations, if they were also linked to the osmotic response in at least one publication. In all, this identified 112 true positives. Likely negatives were taken as genes with no evidence for nuclear localization and whose GO compartment annotation was 'mitochondrion', 'mitochondrial envelope', 'peroxisome', 'vacuole', 'Golgi', and/or 'endoplasmic reticulum'. Proteins annotated in SGD as 'metabolic enzymes' were also added to this list of likely negatives. From this list we removed 32 well-known signaling proteins, many of which were already on the true positive list; in all, this left 1,865 likely negative proteins for the network assessment. Among these test cases, the background network contained 108 positives and 1512 likely negatives. In order to separate out the effect of the experimental hits on predictive accuracy, we omitted all hits from the test cases, leaving 70 true positives and 1416 likely negatives. For each test case (true positive or likely negative), we measured the inferred subnetwork ensemble's confidence that it is relevant to the salt response. This is calculated as the fraction of the 10,000 solutions in which the test case appears as a protein node in the subnetwork.

We compared our ensemble's precision-recall curve to two baselines, which we refer to as the *candidate* baseline (Fig. 4B, green) and *permuted* baseline (Fig. 4B, yellow). For the candidate baseline, we computed the precision and recall of the test cases using the complete set of protein nodes present in candidate paths. For the permuted baseline, we compared the inferred ensemble's accuracy to that of a set of 1,000 ensembles inferred using permuted experimental data. For each of 1,000 permutations, we randomly drew a set of sources, proteins with fitness

defects, and proteins with phospho-changes from the background network, equal in number and degree distribution to the true experimental data. To generate receptor-source pairs, we randomly drew two proteins from the background network and paired each with a randomly chosen source. To generate permuted source-target pairs, for each source, we randomly drew an equal number of targets from the entire background network. We inferred an ensemble of 1,000 solutions for each permutation, and measured the confidence of each test case as the average confidence over all 1,000 ensembles.

Ranking of putative ESR bifurcation points

We constructed the salt-relevant ESR consensus subnetwork shown in Fig. 7A and C as follows. First, we gathered three clusters of genes defined by (GASCH *et al.* 2000) based on expression profiles under multiple stress conditions: iESR (induced ESR) and two rESR (repressed ESR) subclusters, RiBi and RP. Using the protein-nucleic acid interactions from the background network, we identified potential transcriptional regulators of the three ESR gene clusters. These were TFs and RBPs whose targets were enriched for a cluster (determined by hypergeometric test, using a threshold of FDR=0.1, calculated by the Benjamini-Hochberg procedure). For iESR targets, we identified 25 total potential TFs/RBPs, of which 22 are TFs and three are RBPs. We found 16 TFs and 10 RBPs for the combined rESR clusters.

Next, we extracted the consensus source-target paths (having confidence ≥ 75%) that end in an interaction between an ESR-relevant TF/RBP and ESR-relevant target gene (of the same cluster). For each protein node in each ESR-relevant consensus path, we assigned a label based on the ESR cluster(s) represented by the downstream ESR-relevant TF/RBPs. These labels were used to perform the coloring in Fig. 7. Finally, we removed the targets that were not a member of any ESR cluster.

Using the ESR consensus paths, we identified candidate bifurcation points, defined as nodes that are upstream of both rESR and iESR targets (yellow and orange nodes in Fig. 7), according to how well their outgoing paths show a distinct division between the induced and repressed clusters. To rank the candidates, we defined a bifurcation score, $B(n)$, that is related to the concept of information gain ratio (QUINLAN 1986). $B(n)$ is calculated as follows. An illustration of this process is provided in Appendix A Fig. 8.

First, we define the *count* $C(T)$, which counts the number of bits required to represent the cluster membership of all of the targets in a set T. Considering the clusters $c \in$ {iESR, rESR}, let $T^c(n)$ be the set of targets downstream of n that belong to the ESR cluster c.

$$C\big(T(n)\big) = -\sum_{\text{clusters } c \text{ in } \{\text{iESR. rESR}\}} |T^c(n)| \log_2 \frac{|T^c(n)|}{|T(n)|} \tag{19}$$

An ideal bifurcation point would have a high $C(T(n))$ compared to the paths that emanate from it. To perform this comparison, we next calculate $C(\cdot)$ for each of the paths downstream from n. If the subnetwork were a tree, n's targets would simply be partitioned by n's children. However, since the paths leading out from n's children may converge on the same targets, we instead partition $T(n)$ into disjoint subsets of targets, each of which is reachable via a unique combination of n's children. We refer to n's outgoing partitions as $P_1(n)$ ... $P_m(n)$.

After having calculated $C\big(P_i(n)\big)$ for each partition, we then calculate the *information gain*, $I(n)$, which measures the number of bits that are saved by partitioning the targets downstream of n:

$$\mathrm{I(n)} = \mathrm{C\big(T(n)\big)} - \sum_{\mathrm{i=1}}^{\mathrm{m}} \mathrm{C\big(P_i(n)\big)} \tag{20}$$

Finally, to calculate the bifurcation score $B(n)$, we normalize $I(n)$ by the *split information* $S(n)$, which measures the number of bits required to describe the *partition* assignment of one of n's targets. $I(n)$ is strongly biased toward nodes whose outgoing partitions split each target each into its own partition. The normalized score $B(n)$ prioritizes nodes that have a small number of (relatively) cleanly-split outgoing paths and many downstream targets.

$$S(n) = -\sum_{i=0}^{m} \frac{|P_{i(n)}|}{|T(n)|} \log_2 \frac{|P_{i(n)}|}{|T(n)|} \tag{21}$$

$$B(n) = \frac{I(n)}{S(n)} \tag{22}$$

Acknowledgements

Thanks to S. Topper and A. Peterson for experimental and computational assistance, and M. Weinreich, C. Guthrie, F. Posas, A. Kumar, D. Eick, J.L. Corden for sharing strains and reagents, and M. Ferris for computational assistance and resources. Data are available in Appendix A and NIH GEO accession GSE60613. This work was funded by NIH R01 GM083989 to A.P.G., R01 GM080148 to J.J.C., NSF MCB0747197 to A.Z.A, NSF IIS-1218880 and NIH UL1 RR025011 to M.C., and NIH Training Grants in Genomic Sciences (T32HG002760 for ALM), Computation and Informatics in Biology and Medicine (5T15LM007359 for DC), and Chemistry-Biology Interface Training Program (T32 GM008505 for CMN). CMN was also supported by an NSF Graduate Research Fellowship. The authors declare no conflict of interest.

Author contributions

DC, MC, DBB and APG conceived of the project; DC, MC, and APG designed computational approach which was implemented by DC; AEM, MVL, JJC oversaw and conducted proteomic work; YHH, CMN, and AZA collaborated on CTD studies; YHH, DBB, MM, JH, JW conducted expression analysis; YHH performed all other molecular experiments, APG oversaw the project, and all authors contributed to writing the manuscript.

References

ABDULREHMAN, D., P. T. MONTEIRO, M. C. TEIXEIRA, N. P. MIRA, A. B. LOURENÇO *et al.*, 2011 YEASTRACT: providing a programmatic access to curated transcriptional regulatory associations in Saccharomyces cerevisiae through a web services interface. Nucleic Acids Res **39:** D136-D140.

ADROVER, M. A., Z. ZI, A. DUCH, J. SCHABER, A. GONZALEZ-NOVO *et al.*, 2011 Time-dependent quantitative multicomponent control of the G(1)-S network by the stress-activated protein kinase Hog1 upon osmostress. Sci Signal **4:** ra63.

ALEJANDRO-OSORIO, A. L., D. J. HUEBERT, D. T. PORCARO, M. E. SONNTAG, S. NILLASITHANUKROH *et al.*, 2009 The histone deacetylase Rpd3p is required for transient changes in genomic expression in response to stress. Genome Biol **10:** R57.

ALEPUZ, P. M., E. DE NADAL, M. ZAPATER, G. AMMERER and F. POSAS, 2003 Osmostress-induced transcription by Hot1 depends on a Hog1-mediated recruitment of the RNA Pol II. EMBO J **22:** 2433-2442.

ALEXANDER, M. R., M. TYERS, M. PERRET, B. M. CRAIG, K. S. FANG *et al.*, 2001 Regulation of cell cycle progression by Swe1p and Hog1p following hypertonic stress. Mol Biol Cell **12:** 53-62.

ANSARI, A. Z., A. OGIRALA and M. PTASHNE, 2005 Transcriptional activating regions target attached substrates to a cyclin-dependent kinase. Proc Natl Acad Sci U S A **102:** 2346-2349.

BELLI, G., E. GARI, M. ALDEA and E. HERRERO, 2001 Osmotic stress causes a G1 cell cycle delay and downregulation of Cln3/Cdc28 activity in Saccharomyces cerevisiae. Mol Microbiol **39:** 1022-1035.

BERGKESSEL, M., G. B. WHITWORTH and C. GUTHRIE, 2011 Diverse environmental stresses elicit distinct responses at the level of pre-mRNA processing in yeast. Rna **17:** 1461-1478.

BERRY, D. B., and A. P. GASCH, 2008 Stress-activated genomic expression changes serve a preparative role for impending stress in yeast. Mol Biol Cell **19:** 4580-4587.

BERRY, D. B., Q. GUAN, J. HOSE, S. HAROON, M. GEBBIA *et al.*, 2011 Multiple means to the same end: the genetic basis of acquired stress resistance in yeast. PLoS Genet **7:** e1002353.

BREITKREUTZ, A., H. CHOI, J. R. SHAROM, L. BOUCHER, V. NEDUVA *et al.*, 2010 A global protein kinase and phosphatase interaction network in yeast. Science **328:** 1043-1046.

BROACH, J. R., 2012 Nutritional control of growth and development in yeast. Genetics **192:** 73-105.

BULAVIN, D. V., S. SAITO, M. C. HOLLANDER, K. SAKAGUCHI, C. W. ANDERSON *et al.*, 1999 Phosphorylation of human p53 by p38 kinase coordinates N-terminal phosphorylation and apoptosis in response to UV radiation. EMBO J **18:** 6845-6854.

BURATOWSKI, S., 2003 The CTD code. Nat Struct Biol **10:** 679-680.

BUSNELLI, S., F. TRIPODI, R. NICASTRO, C. CIRULLI, G. TEDESCHI *et al.*, 2013 Snf1/AMPK promotes SBF and MBF-dependent transcription in budding yeast. Biochim Biophys Acta **1833:** 3254-3264.

CAPALDI, A. P., T. KAPLAN, Y. LIU, N. HABIB, A. REGEV *et al.*, 2008 Structure and function of a transcriptional network activated by the MAPK Hog1. Nat Genet **40:** 1300-1306.

CARVAJAL, L. A., and J. J. MANFREDI, 2013 Another fork in the road--life or death decisions by the tumour suppressor p53. EMBO Rep **14:** 414-421.

CAUSTON, H. C., B. REN, S. S. KOH, C. T. HARBISON, E. KANIN *et al.*, 2001 Remodeling of yeast genome expression in response to environmental changes. Mol Biol Cell **12:** 323-337.

CHERRY, J. M., E. L. HONG, C. AMUNDSEN, R. BALAKRISHNAN, G. BINKLEY *et al.*, 2012 Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res **40:** D700-705.

CHURCHMAN, L. S., and J. S. WEISSMAN, 2011 Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature **469:** 368-373.

CHYMKOWITCH, P., V. ELDHOLM, S. LORENZ, C. ZIMMERMANN, J. M. LINDVALL *et al.*, 2012 Cdc28 kinase activity regulates the basal transcription machinery at a subset of genes. Proc Natl Acad Sci U S A **109:** 10450-10455.

COOK, K. E., and E. K. O'SHEA, 2012 Hog1 controls global reallocation of RNA Pol II upon osmotic shock in Saccharomyces cerevisiae. G3 (Bethesda) **2:** 1129-1136.

DE AMICIS, F., F. GIORDANO, A. VIVACQUA, M. PELLEGRINO, M. L. PANNO *et al.*, 2011 Resveratrol, through NF-Y/p53/Sin3/HDAC1 complex phosphorylation, inhibits estrogen receptor alpha gene expression via p38MAPK/CK2 signaling in human breast cancer cells. FASEB J **25:** 3695-3707.

DE NADAL, E., and F. POSAS, 2010 Multilayered control of gene expression by stress-activated protein kinases. EMBO J **29:** 4-13.

ELLIOTT, B., and B. FUTCHER, 1993 Stress resistance of yeast cells is largely independent of cell cycle phase. Yeast **9:** 33-42.

EVERETT, L., A. VO and S. HANNENHALLI, 2009 PTM-Switchboard--a database of posttranslational modifications of transcription factors, the mediating enzymes and target genes. Nucleic Acids Res **37:** D66-71.

FASOLO, J., A. SBONER, M. G. F. SUN, H. YU, R. CHEN *et al.*, 2011 Diverse protein kinase interactions identified by protein microarrays reveal novel connections between cellular processes. Genes \& Development **25:** 767-778.

FORBES, S. A., N. BINDAL, S. BAMFORD, C. COLE, C. Y. KOK *et al.*, 2011 COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Res **39:** D945-950.

FRIEDMAN, N., 2004 Inferring cellular networks using probabilistic graphical models. Science **303:** 799-805.

GASCH, A. P., 2002 The Environmental Stress Response: a common yeast response to environmental stresses. In Yeast Stress Responses., pp. 11-70 in *Yeast Stress Responses*, edited by S. HOHMANN and P. MAGER. Springer-Verlag, Heidelberg.

GASCH, A. P., M. HUANG, S. METZNER, D. BOTSTEIN, S. J. ELLEDGE *et al.*, 2001 Genomic expression responses to DNA-damaging agents and the regulatory role of the yeast ATR homolog Mec1p. Mol Biol Cell **12:** 2987-3003.

GASCH, A. P., P. T. SPELLMAN, C. M. KAO, O. CARMEL-HAREL, M. B. EISEN *et al.*, 2000 Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell **11:** 4241-4257.

GAT-VIKS, I., and R. SHAMIR, 2007 Refinement and expansion of signaling pathways: the osmotic response network in yeast. Genome Res **17:** 358-367.

GAT-VIKS, I., A. TANAY, D. RAIJMAN and R. SHAMIR, 2006 A probabilistic methodology for integrating knowledge and experiments on biological networks. Journal of Computational Biology **13:** 165-181.

GEER, L. Y., S. P. MARKEY, J. A. KOWALAK, L. WAGNER, M. XU *et al.*, 2004 Open mass spectrometry search algorithm. J Proteome Res **3:** 958-964.

GHAEMMAGHAMI, S., W. K. HUH, K. BOWER, R. W. HOWSON, A. BELLE *et al.*, 2003 Global analysis of protein expression in yeast. Nature **425:** 737-741.

GITTER, A., M. CARMI, N. BARKAI and Z. BAR-JOSEPH, 2013 Linking the signaling cascades and dynamic regulatory networks controlling stress responses. Genome Res **23:** 365-376.

GITTER, A., J. KLEIN-SEETHARAMAN, A. GUPTA and Z. BAR-JOSEPH, 2011 Discovering pathways by orienting edges in protein interaction networks. Nucleic Acids Res **39:** e22.

GUELZIM, N., S. BOTTANI, P. BOURGINE and F. KÉPÈS, 2002 Topological and causal structure of the yeast transcriptional regulatory network. Nat Genet **31:** 60-63.

HALBEISEN, R. E., and A. P. GERBER, 2009 Stress-dependent coordination of transcriptome and translatome in yeast. PLoS Biol **7:** e1000105.

HAMOSH, A., A. F. SCOTT, J. S. AMBERGER, C. A. BOCCHINI and V. A. MCKUSICK, 2005 Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res **33:** D514-517.

HEAVNER, B. D., K. SMALLBONE, B. BARKER, P. MENDES and L. P. WALKER, 2012 Yeast 5 - an expanded reconstruction of the Saccharomyces cerevisiae metabolic network. BMC Syst Biol **6:** 55.

HENGARTNER, C. J., V. E. MYER, S. M. LIAO, C. J. WILSON, S. S. KOH *et al.*, 1998 Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. Mol Cell **2:** 43-53.

HILDESHEIM, J., J. M. SALVADOR, M. C. HOLLANDER and A. J. FORNACE, JR., 2005 Casein kinase 2- and protein kinase A-regulated adenomatous polyposis coli and beta-catenin cellular localization is dependent on p38 MAPK. J Biol Chem **280:** 17221-17226.

HIRASAWA, T., K. ASHITANI, K. YOSHIKAWA, K. NAGAHISA, C. FURUSAWA *et al.*, 2006 Comparison of transcriptional responses to osmotic stresses induced by NaCl and sorbitol additions in Saccharomyces cerevisiae using DNA microarray. J Biosci Bioeng **102:** 568-571.

HOGAN, D. J., D. P. RIORDAN, A. P. GERBER, D. HERSCHLAG and P. O. BROWN, 2008 Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. PLoS Biol **6:** e255.

HOHMANN, S., and P. MAGER (Editors), 2003 *Yeast Stress Responses*. Springer-Verlag, Heidelberg.

HONG, S. P., and M. CARLSON, 2007 Regulation of snf1 protein kinase in response to environmental stress. J Biol Chem **282:** 16838-16845.

HUANG, S. S., D. C. CLARKE, S. J. GOSLINE, A. LABADORF, C. R. CHOUINARD *et al.*, 2013 Linking proteomic and transcriptional data through the interactome and epigenome reveals a map of oncogene-induced signaling. PLoS Comput Biol **9:** e1002887.

HUANG, S. S., and E. FRAENKEL, 2009 Integrating proteomic, transcriptional, and interactome data reveals hidden components of signaling and regulatory networks. Sci Signal **2:** ra40.

HUANG, S. S., and E. FRAENKEL, 2012 Swimming upstream: identifying proteomic signals that drive transcriptional changes using the interactome and multiple "-omics" datasets. Methods Cell Biol **110:** 57-80.

HUEBERT, D. J., P. F. KUAN, S. KELES and A. P. GASCH, 2012 Dynamic changes in nucleosome occupancy are not predictive of gene expression dynamics but are linked to transcription and chromatin regulators. Mol Cell Biol **32:** 1645-1653.

IDEKER, T. E., V. THORSSON and R. M. KARP, 2000 Discovery of regulatory interactions through perturbation: inference and experimental design. Pac Symp Biocomput **5:** 305-316.

INGOLIA, N. T., S. GHAEMMAGHAMI, J. R. NEWMAN and J. S. WEISSMAN, 2009 Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science **324:** 218-223.

ISAEVA, A. R., and V. I. MITEV, 2011 CK2 is acting upstream of MEK3/6 as a part of the signal control of ERK1/2 and p38 MAPK during keratinocytes autocrine differentiation. Z Naturforsch C **66:** 83-86.

JONES, R. G., and C. B. THOMPSON, 2009 Tumor suppressors and cell metabolism: a recipe for cancer growth. Genes Dev **23:** 537-548.

LAN, A., I. Y. SMOLY, G. RAPAPORT, S. LINDQUIST, E. FRAENKEL *et al.*, 2011 ResponseNet: revealing signaling and regulatory networks linking genetic and transcriptomic screening data. Nucleic Acids Res **39:** W424-429.

LEE, M. V., S. E. TOPPER, S. L. HUBERL, J. HOSE, C. D. WENGER *et al.*, 2011 A Dynamic Model of Proteome Changes Reveals New Roles for Transcript Alteration in Yeast Mol Syst Biol.

LEVY, S. F., N. ZIV and M. L. SIEGAL, 2012 Bet hedging in yeast by heterogeneous, age-correlated expression of a stress protectant. PLoS Biol **10:** e1001325.

LI, L., M. LJUNGMAN and J. E. DIXON, 2000 The human Cdc14 phosphatases interact with and dephosphorylate the tumor suppressor protein p53. J Biol Chem **275:** 2410-2414.

LIANG, S., S. FUHRMAN and R. SOMOGYI, 1998 Reveal, a general reverse engineering algorithm for inference of genetic network architectures. Pac Symp Biocomput **3:** 18-29.

LIPSON, D., T. RAZ, A. KIEU, D. R. JONES, E. GILADI *et al.*, 2009 Quantification of the yeast transcriptome by single-molecule sequencing. Nat Biotechnol **27:** 652-658.

LIU, Y., C. KUNG, J. FISHBURN, A. Z. ANSARI, K. M. SHOKAT *et al.*, 2004 Two cyclin-dependent kinases promote RNA polymerase II transcription and formation of the scaffold complex. Mol Cell Biol **24:** 1721-1735.

LU, C., M. J. BRAUER and D. BOTSTEIN, 2009 Slow growth induces heat-shock resistance in normal and respiratory-deficient yeast. Mol Biol Cell **20:** 891-903.

MACISAAC, K., T. WANG, D. B. GORDON, D. GIFFORD, G. STORMO *et al.*, 2006 An improved map of conserved regulatory sites for Saccharomyces cerevisiae. BMC Bioinformatics **7:** 113+.

MARKOWETZ, F., J. BLOCH and R. SPANG, 2005 Non-transcriptional pathway features reconstructed from secondary effects of RNA interference. Bioinformatics **21:** 4026-4032.

MARLES, J. A., S. DAHESH, J. HAYNES, B. J. ANDREWS and A. R. DAVIDSON, 2004 Protein-protein interaction affinity plays a crucial role in controlling the Sho1p-mediated signal transduction pathway in yeast. Mol Cell **14:** 813-823.

MARTINEZ-MONTANES, F., A. PASCUAL-AHUIR and M. PROFT, 2010 Toward a genomic view of the gene expression program regulated by osmostress in yeast. OMICS **14:** 619-627.

MCCLEAN, M. N., A. MODY, J. R. BROACH and S. RAMANATHAN, 2007 Cross-talk and decision making in MAP kinase pathways. Nat Genet **39:** 409-414.

MEEK, D. W., S. SIMON, U. KIKKAWA and W. ECKHART, 1990 The p53 tumour suppressor protein is phosphorylated at serine 389 by casein kinase II. EMBO J **9:** 3253-3260.

MELAMED, D., L. PNUELI and Y. ARAVA, 2008 Yeast translational response to high salinity: global analysis reveals regulation at multiple levels. Rna **14:** 1337-1351.

MILLER, C., B. SCHWALB, K. MAIER, D. SCHULZ, S. DUMCKE *et al.*, 2011 Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. Mol Syst Biol **7:** 458.

MILLER, C. T., C. GABRIELSE, Y. C. CHEN and M. WEINREICH, 2009 Cdc7p-Dbf4p regulates mitotic exit by inhibiting Polo kinase. PLoS Genet **5:** e1000498.

MITCHELL, A., G. H. ROMANO, B. GROISMAN, A. YONA, E. DEKEL *et al.*, 2009 Adaptive prediction of environmental changes by microorganisms. Nature **460:** 220-224.

NADAL-RIBELLES, M., N. CONDE, O. FLORES, J. GONZALEZ-VALLINAS, E. EYRAS *et al.*, 2012 Hog1 bypasses stress-mediated down-regulation of transcription by RNA polymerase II redistribution and chromatin remodeling. Genome Biol **13:** R106.

NAGIEC, M. J., and H. G. DOHLMAN, 2012 Checkpoints in a yeast differentiation pathway coordinate signaling during hyperosmotic stress. PLoS Genet **8:** e1002437.

NESVIZHSKII, A. I., and R. AEBERSOLD, 2005 Interpretation of shotgun proteomic data: the protein inference problem. Mol Cell Proteomics **4:** 1419-1440.

NI, L., C. BRUCE, C. HART, J. LEIGH-BELL, D. GELPERIN *et al.*, 2009 Dynamic and complex transcription factor binding during an inducible response in yeast. Genes Dev **23:** 1351-1363.

NOVERSHTERN, N., A. REGEV and N. FRIEDMAN, 2011 Physical Module Networks: an integrative approach for reconstructing transcription regulation. Bioinformatics **27:** i177-i185.

O'ROURKE, S. M., and I. HERSKOWITZ, 1998 The Hog1 MAPK prevents cross talk between the HOG and pheromone response MAPK pathways in Saccharomyces cerevisiae. Genes Dev **12:** 2874-2886.

O'ROURKE, S. M., and I. HERSKOWITZ, 2004 Unique and redundant roles for HOG MAPK pathway components as revealed by whole-genome expression analysis. Mol Biol Cell **15:** 532-542.

OURFALI, O., T. SHLOMI, T. IDEKER, E. RUPPIN and R. SHARAN, 2007 SPINE: a framework for signaling-regulatory pathway inference from cause-effect experiments. Bioinformatics **23:** i359-366.

PATTERSON, J. C., E. S. KLIMENKO and J. THORNER, 2010 Single-cell analysis reveals that insulation maintains signaling specificity between two yeast MAPK pathways with common components. Sci Signal **3:** ra75.

PATTURAJAN, M., R. J. SCHULTE, B. M. SEFTON, R. BEREZNEY, M. VINCENT *et al.*, 1998 Growth-related changes in phosphorylation of yeast RNA polymerase II. J Biol Chem **273:** 4689-4694.

PESSINA, S., V. TSIARENTSYEVA, S. BUSNELLI, M. VANONI, L. ALBERGHINA *et al.*, 2010 Snf1/AMPK promotes S-phase entrance by controlling CLB5 transcription in budding yeast. Cell Cycle **9:** 2189-2200.

PHANSTIEL, D. H., J. BRUMBAUGH, C. D. WENGER, S. TIAN, M. D. PROBASCO *et al.*, 2011 Proteomic and phosphoproteomic comparison of human ES and iPS cells. Nat Methods **8:** 821-827.

PHATNANI, H. P., J. C. JONES and A. L. GREENLEAF, 2004 Expanding the functional repertoire of CTD kinase I and RNA polymerase II: novel phosphoCTD-associating proteins in the yeast proteome. Biochemistry **43:** 15702-15719.

PROFT, M., G. MAS, E. DE NADAL, A. VENDRELL, N. NORIEGA *et al.*, 2006 The stress-activated Hog1 kinase is a selective transcriptional elongation factor for genes responding to osmotic stress. Mol Cell **23:** 241-250.

PTACEK, J., G. DEVGAN, G. MICHAUD, H. ZHU, X. ZHU *et al.*, 2005 Global analysis of protein phosphorylation in yeast. Nature **438:** 679-684.

PU, S., J. WONG, B. TURNER, E. CHO and S. J. WODAK, 2009 Up-to-date catalogues of yeast protein complexes. Nucleic Acids Res **37:** 825-831.

PUIG, O., F. CASPARY, G. RIGAUT, B. RUTZ, E. BOUVERET *et al.*, 2001 The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods **24:** 218-229.

QUINLAN, J. R., 1986 Induction of decision trees. Machine Learning **1:** 81-106.

REISER, V., K. E. D'AQUINO, L. S. EE and A. AMON, 2006 The stress-activated mitogen-activated protein kinase signaling cascade promotes exit from mitosis. Mol Biol Cell **17:** 3136-3146.

SAITO, H., and K. TATEBAYASHI, 2004 Regulation of the osmoregulatory HOG MAPK cascade in yeast. J Biochem **136:** 267-272.

SAYED, M., S. O. KIM, B. S. SALH, O. G. ISSINGER and S. L. PELECH, 2000 Stress-induced activation of protein kinase CK2 by direct interaction with p38 mitogen-activated protein kinase. J Biol Chem **275:** 16569-16573.

SCHADT, E. E., J. LAMB, X. YANG, J. ZHU, S. EDWARDS *et al.*, 2005 An integrative genomics approach to infer causal associations between gene expression and disease. Nat Genet **37:** 710-717.

SCHERRER, T., N. MITTAL, S. C. JANGA and A. P. GERBER, 2010 A screen for RNA-binding proteins in yeast indicates dual functions for many enzymes. PLoS One **5:** e15499.

SCHWARTZ, M. A., and H. D. MADHANI, 2004 Principles of MAP kinase signaling specificity in Saccharomyces cerevisiae. Annu Rev Genet **38:** 725-748.

SHARIFPOOR, S., A. N. N. BA, J.-Y. YOUNG, D. VAN DYK, H. FRIESEN *et al.*, 2011 A quantitative literature-curated gold standard for kinase-substrate pairs. Genome Biol **12:** R39.

SHOCK, T. R., J. THOMPSON, J. R. YATES, 3RD and H. D. MADHANI, 2009 Hog1 mitogen-activated protein kinase (MAPK) interrupts signal transduction between the Kss1 MAPK and the Tec1 transcription factor to maintain pathway specificity. Eukaryot Cell **8:** 606-616.

SILVERBUSH, D., M. ELBERFELD and R. SHARAN, 2011 Optimally orienting physical networks. J Comput Biol **18:** 1437-1448.

SMETS, B., R. GHILLEBERT, P. DE SNIJDER, M. BINDA, E. SWINNEN *et al.*, 2010 Life in the midst of scarcity: adaptations to nutrient availability in Saccharomyces cerevisiae. Curr Genet **56:** 1-32.

SMYTH, G. K., 2004 Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Statistical applications in genetics and molecular biology **3:** Article 3.

SOPKO, R., D. HUANG, N. PRESTON, G. CHUA, B. PAPP *et al.*, 2006 Mapping pathways and phenotypes by systematic gene overexpression. Mol Cell **21:** 319-330.

SOUFI, B., C. D. KELSTRUP, G. STOEHR, F. FROHLICH, T. C. WALTHER *et al.*, 2009 Global analysis of the yeast osmotic stress response by quantitative proteomics. Mol Biosyst **5:** 1337-1346.

SPELLMAN, P. T., G. SHERLOCK, M. Q. ZHANG, V. R. IYER, K. ANDERS *et al.*, 1998 Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. Mol Biol Cell **9:** 3273-3297.

STARK, C., B.-J. J. BREITKREUTZ, T. REGULY, L. BOUCHER, A. BREITKREUTZ *et al.*, 2006 BioGRID: a general repository for interaction datasets. Nucleic Acids Res **34:** D535-539.

STOREY, J. D., W. XIAO, J. T. LEEK, R. G. TOMPKINS and R. W. DAVIS, 2005 Significance analysis of time course microarray experiments. Proc Natl Acad Sci U S A **102:** 12837-12842.

SUMNER, E. R., and S. V. AVERY, 2002 Phenotypic heterogeneity: differential stress resistance among individual cells of the yeast Saccharomyces cerevisiae. Microbiology **148:** 345-351.

SUTHRAM, S., A. BEYER, R. M. KARP, Y. ELDAR and T. IDEKER, 2008 eQED: an efficient method for interpreting eQTL associations using protein networks. Mol Syst Biol **4:** 162.

TIETJEN, J. R., D. W. ZHANG, J. B. RODRIGUEZ-MOLINA, B. E. WHITE, M. S. AKHTAR *et al.*, 2010 Chemical-genomic dissection of the CTD code. Nat Struct Mol Biol **17:** 1154-1161.

TIGER, C.-F., F. KRAUSE, G. CEDERSUND, R. PALMÈR, E. KLIPP *et al.*, 2012 A framework for mapping, visualisation and automatic model creation of signal-transduction networks. Mol Syst Biol **8:** 578.

TSVETANOVA, N. G., D. M. KLASS, J. SALZMAN and P. O. BROWN, 2010 Proteome-wide search reveals unexpected RNA-binding proteins in Saccharomyces cerevisiae. PLoS One **5**.

TU, Z., L. WANG, M. N. ARBEITMAN, T. CHEN and F. SUN, 2006 An integrative approach for causal gene identification and gene regulatory pathway inference. Bioinformatics **22:** e489-e496.

VAN WUYTSWINKEL, O., V. REISER, M. SIDERIUS, M. C. KELDERS, G. AMMERER *et al.*, 2000 Response of Saccharomyces cerevisiae to severe osmotic stress: evidence for a novel activation mechanism of the HOG MAP kinase pathway. Mol Microbiol **37:** 382-397.

VASKE, C. J., C. HOUSE, T. LUU, B. FRANK, C.-H. H. YEANG *et al.*, 2009 A factor graph nested effects model to identify networks from genetic perturbations. PLoS Comput Biol **5:** e1000274.

VENTERS, B. J., S. WACHI, T. N. MAVRICH, B. E. ANDERSEN, P. JENA *et al.*, 2011 A comprehensive genomic binding map of gene and chromatin regulatory proteins in Saccharomyces. Mol Cell **41:** 480-492.

WALL, D. P., H. B. FRASER and A. E. HIRSH, 2003 Detecting putative orthologs. Bioinformatics **19:** 1710-1711.

WALTERMANN, C., and E. KLIPP, 2010 Signal integration in budding yeast. Biochem Soc Trans **38:** 1257-1264.

WARNER, J. R., J. VILARDELL and J. H. SOHN, 2001 Economics of ribosome biosynthesis. Cold Spring Harb Symp Quant Biol **66:** 567-574.

WARRINGER, J., M. HULT, S. REGOT, F. POSAS and P. SUNNERHAGEN, 2010 The HOG pathway dictates the short-term translational response after hyperosmotic shock. Mol Biol Cell **21:** 3080-3092.

WENGER, C. D., M. V. LEE, A. S. HEBERT, G. C. MCALISTER, D. H. PHANSTIEL *et al.*, 2011a Gas-phase purification enables accurate, multiplexed proteome quantification with isobaric tagging. Nat Methods **8:** 933-935.

WENGER, C. D., D. H. PHANSTIEL, M. V. LEE, D. J. BAILEY and J. J. COON, 2011b COMPASS: a suite of pre- and post-search proteomics software tools for OMSSA. Proteomics **11:** 1064-1074.

WEST, M. L., and J. L. CORDEN, 1995 Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations. Genetics **140:** 1223-1233.

WESTFALL, P. J., J. C. PATTERSON, R. E. CHEN and J. THORNER, 2008 Stress resistance and signal fidelity independent of nuclear MAPK function. Proc Natl Acad Sci U S A **105:** 12212-12217.

WINZELER, E. A., D. D. SHOEMAKER, A. ASTROMOFF, H. LIANG, K. ANDERSON *et al.*, 1999 Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science **285:** 901-906.

WOODS, J. O., U. M. SINGH-BLOM, J. M. LAURENT, K. L. MCGARY and E. M. MARCOTTE, 2013 Prediction of gene-phenotype associations in humans, mice, and plants using phenologs. BMC Bioinformatics **14:** 203.

YE, T., K. ELBING and S. HOHMANN, 2008 The pathway by which the yeast protein kinase Snf1p controls acquisition of sodium tolerance is different from that mediating glucose regulation. Microbiology **154:** 2814-2826.

YEANG, C.-H., T. IDEKER and T. JAAKKOLA, 2004a Physical network models. Journal of Computational Biology **11:** 243-262.

YEANG, C. H., T. IDEKER and T. JAAKKOLA, 2004b Physical network models. J Comput Biol **11:** 243-262.

YEANG, C. H., H. C. MAK, S. MCCUINE, C. WORKMAN, T. JAAKKOLA *et al.*, 2005 Validation and refinement of gene-regulatory pathways on a network of physical interactions. Genome Biol **6:** R62.

YEGER-LOTEM, E., L. RIVA, L. J. SU, A. D. GITLER, A. G. CASHIKAR *et al.*, 2009 Bridging high-throughput genetic and transcriptional data reveals cellular responses to alpha-synuclein toxicity. Nat Genet **41:** 316-323.

YEUNG, K. Y., M. MEDVEDOVIC and R. E. BUMGARNER, 2004 From co-expression to co-regulation: how many microarray experiments do we need? Genome Biol **5:** R48.

YOSEF, N., L. UNGAR, E. ZALCKVAR, A. KIMCHI, M. KUPIEC *et al.*, 2009 Toward accurate reconstruction of functional protein networks. Mol Syst Biol **5:** 248.

YOU, C., H. OKANO, S. HUI, Z. ZHANG, M. KIM *et al.*, 2013 Coordination of bacterial proteome with metabolism by cyclic AMP signalling. Nature **500:** 301-306.

ZAKRZEWSKA, A., G. VAN EIKENHORST, J. E. BURGGRAAFF, D. J. VIS, H. HOEFSLOOT *et al.*, 2011 Genome-wide analysis of yeast stress survival and tolerance acquisition to analyze the central trade-off between growth rate and cellular robustness. Mol Biol Cell **22:** 4435-4446.

ZARRINPAR, A., R. P. BHATTACHARYYA, M. P. NITTLER and W. A. LIM, 2004 Sho1 and Pbs2 act as coscaffolds linking components in the yeast high osmolarity MAP kinase pathway. Mol Cell **14:** 825-832.

ZHANG, D. W., J. B. RODRIGUEZ-MOLINA, J. R. TIETJEN, C. M. NEMEC and A. Z. ANSARI, 2012 Emerging Views on the CTD Code. Genet Res Int **2012:** 347214.

ZHU, H., M. BILGIN, R. BANGHAM, D. HALL, A. CASAMAYOR *et al.*, 2001 Global analysis of protein activities using proteome chips. Science **293:** 2101-2105.

Chapter 3

The Environmental Stress Response is a true stress response, not a byproduct of cell cycle phase or slow growth rate

Introduction

Free-living microbes in nature must respond precisely and agilely to environmental changes that occur frequently in nature. Rapid changes to the external conditions can cause a direct impact on internal functions, threatening growth and survival. Therefore, microbial cells have evolved to respond rapidly to sudden environmental shifts by coordinating a complex physiological response. When the environment changes, budding yeast cells modulate the expression of a large fraction of the transcriptome to acclimate to the new conditions (GASCH *et al.* 2000; CAUSTON *et al.* 2001). While much of the transcriptome changes are specific to the conditions at hand, transcriptomic changes triggered by acute stress also include a common gene expression program called the environmental stress response (ESR) (GASCH *et al.* 2000). The ESR comprises ~300 induced transcripts (iESR) and ~600 reduced genes (rESR) whose anti-correlated expression is coordinated across a diverse range of stresses (GASCH 2002a). The iESR genes are broadly involved in stress defense, while the rESR transcripts encode ribosomal proteins (RPs) and ribosome biogenesis (RiBi) factors. In actively growing cells, rESR genes are highly expressed and translated to supply the needs of growing cells (WARNER 1999; GASCH 2002a). The function of the ESR was originally proposed to protect cells against stress. Since that time, the increased abundance of iESR transcripts was shown to be important for surviving subsequent stress exposure (BERRY and GASCH 2008; LU *et al.* 2009), while the role of rESR repression has been less clear. Importantly, ESR programs orthologous to the yeast response have been observed in bacteria (VAN DE GUCHTE *et al.* 2002; BATTESTI *et al.* 2011), other fungi (GASCH 2007; ROETZER *et al.* 2008; WOHLBACH *et al.* 2014), *Drosophila* (BROWN *et al.* 2014), and mammalian cells (HO and GASCH 2015).

While the ESR was originally proposed as a stress response, a number of studies have challenged this assertion. Several studies observed that ESR activation correlates with reduced

growth rates in nutrient-restricted chemostats (REGENBERG *et al.* 2006; CASTRILLO *et al.* 2007; BRAUER *et al.* 2008). In this culture system, the growth rate can be controlled by restricting the influx of fresh nutrients, such that cells with reduced access to nutrients grow at a reduced rate. As growth rate decreases across chemostat cultures, abundance of many iESR genes increases while levels of rESR genes drop in linear proportion to growth (REGENBERG *et al.* 2006; CASTRILLO *et al.* 2007; BRAUER *et al.* 2008). In particular, the correlation between growth and rESR transcript levels that encode ribosome components fits well with models in bacteria (MAALOE and KJELDGAARD 1966; NOMURA 1999). Together, such studies proposed that ESR activation is a secondary effect of the growth arrest triggered by stress, rather than a direct stress response (CASTRILLO *et al.* 2007; BRAUER *et al.* 2008), and furthermore that growth rate can be inferred simply based on ESR transcript abundance (BRAUER *et al.* 2008; AIROLDI *et al.* 2009). Another recent study by O'Duibhir *et al* proposed that ESR activation is instead a signature of prolonged G1 progression in slow-growing cells (O'DUIBHIR *et al.* 2014). These authors observed a common transcriptome signature that is highly similar to the ESR in slow growing gene-knockout mutants grown in batch cultures. Based on analysis of cell cycle distributions and previously published cell-cycle transcriptome data, these authors proposed that the ESR is not a stress response but simply an indirect reflection of changes in the cell cycle distribution (O'DUIBHIR *et al.* 2014). Since slower growing *S. cerevisiae* cells (including those in nutrient-restricted chemostats) spend a larger fraction of their cell cycle in G1 phase (JOHNSTON *et al.* 1980; SALDANHA *et al.* 2004; BRAUER *et al.* 2008), this model could also explain the relationship between ESR activation and growth rate in chemostat cultures. However, a major challenge with the above studies is that they could not decouple the effects of growth and cell-cycle progression from the stress of nutrient restriction and gene mutation that led to the growth effects in the first place. Thus, the function of the ESR – and the purposed of rESR repression in particular – remains ambiguous.

Understanding the true nature of the ESR is of fundamental importance in understanding environmental responses conserved across species, and in turn influences how transcriptomic data are interpreted. Here, we set out to address the responsiveness of the ESR by decoupling growth and cell cycle progression from the stress response. We find that the ESR cannot be explained by a linear relationship with growth rate, that the ESR does not correlate with cell cycle arrest point, and that the ESR is not associated with G1 phase. Instead, the ESR is actively regulated in response to stress even in growth- and cell-cycle arrested cells, in a stress dose-dependent manner. We propose that the ESR and growth rate are separate but inter-connected features of the cellular allostasis program, which fluctuates to accommodate the internal balance of translational capacity depending on the conditions at hand.

Results

We wanted to address the question of whether the ESR transcriptome program is an active response to stress or an indirect byproduct of growth or cell-cycle arrest. To do this, we decoupled growth and cell cycle progression from the stress response, by characterizing the stress-responsive transcriptome in arrested cells. Cells were arrested with a two-hour treatment of either nocodazole or mating pheromone that fully arrest cells at G2/M or G1 phases, respectively. Both treatments lead to cell-cycle arrest as well as growth reduction, as indicated by several features. Cell morphology changes in arrested cells showed that the majority of nocodazole treated cells (~85%) were dumbbell shaped as expected for G2/M arrest, while most alpha factor treated cells were unbudded (~98%) with a mating projection (Table 1). Arrested cells also showed a significant retardation in the rise of optical density (Fig 1A) and of biomass production (Fig 1B) over time, which we define here as ‘growth’. Since a significant portion of biomass is represented by protein accumulation, we further investigated bulk translation trends using polysome profiling. Arrested cells showed altered polysome profiles compared to actively growing asynchronous cells, with a reduced fraction of mRNAs in the

	T0	T30
Alpha factor		
budded %	2.05 ± 0.53	7.22 ± 2.64
Nocodazole		
dumbbell-shaped %	84.56 ± 5.63	86.56 ± 4.56

Table 1. Cell morphology changes in arrested cells.

Cells treated with either nocodazole or alpha factor were arrested at either at G2/M or G1 phases. Morphology change was monitored after two-hour treatment and also after applying 0.7N NaCl for 30 mins. The mean of two biological experiments ± SD is shown.

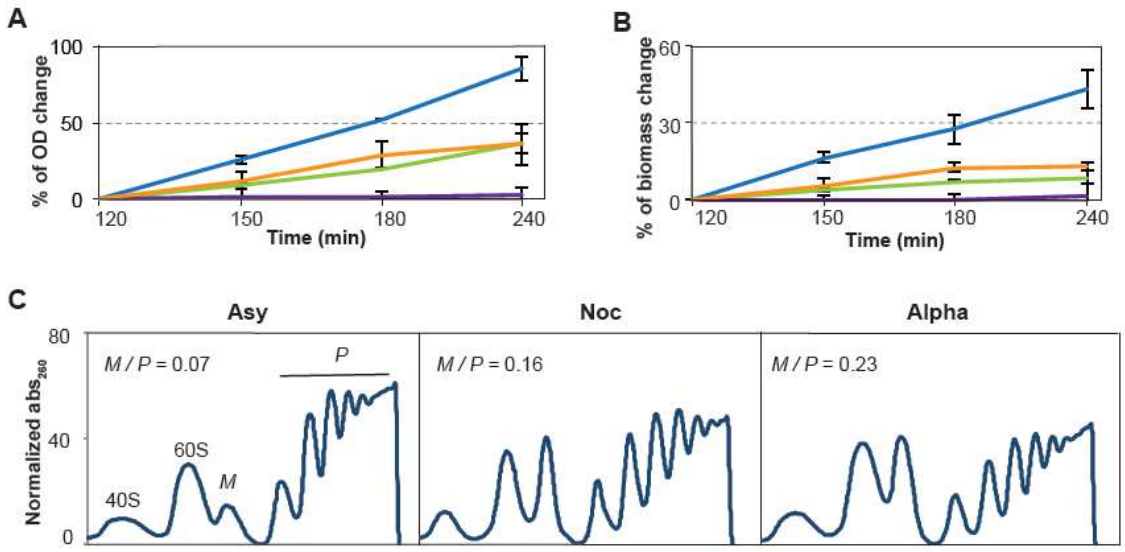


Fig 1. Characterization of nocodazole- and alpha-arrested cells.

A. Change in optical density (relative to starting point) starting two hours after treatment with nocodazole (orange) or alpha factor (green). Asynchronously growing cells (blue) and starved cells (purple) are also shown for reference.

B. Percent change in biomass (normalized to the starting point) of arrested or growing cells. Color codes are the same as in A.

C. Polysome profiles of asynchronously growing cells and cells treated for 2 h with nocodazole (Noc) or alpha factor (Alpha). Absorbance at 260 nm across fractions is shown relative to the starting baseline; 40S, 60S and 80S monosome (*M*) and polysome (*P*) peaks are indicated. The monosome/polysome (M/P) ratio was calculated based on the area under the curve.

actively translating polysome ('P') fractions and an increase of mRNAs in the monosome ('M') fraction that reflects stalled translation initiation (Fig 1C). In fact, arrested cells showed 2-3 fold higher M/P ratios, suggesting that they have reduced translation globally. In all, these results show that treatment with nocodazole or alpha factor fully arrests cells at G2/M or G1 phases, dramatically reduces biomass production, and significantly alters bulk translation compared to asynchronous cells.

ESR activation does not correlate with cell-cycle phase

To clarify the relation of ESR and cell cycle distribution, we compared the transcriptome profile of cells arrested at different phases to that in asynchronous cells, through RNA-seq analysis. O'Duibhi *et al* proposed that the ESR triggered by diverse stresses mainly reflects the resulting increase in G1 cells within the population, instead of a direct stress response (O'DUIBHIR *et al.* 2014). This model predicts that the ESR would be activated in G1 arrested cells but repressed in G2/M phase. However, this was not the result: both nocodazole- and pheromone-arrested cells displayed mild ESR activation. The magnitude of ESR activation was minor compared to that seen in response to stress (see below), with a 1.3 ~2.5 fold increase in iESR transcripts and 0.1 ~0.4 fold decrease in rESR messages compared to asynchronous cells, in response to nocodozole or pheromone, respectively. In the case of nocodazole-arrested cells, the magnitude of ESR activation was influenced by the growth temperature, since cells arrested at 25 degrees had lower ESR activation than cells arrested at 30 degrees. This indicates that the cells are responding to factors other than mere cell cycle phase (Fig 2). In response to pheromone treatment, the ESR was activated to a slightly higher level. This result may indicate that the ESR correlates with G1 phase or instead could be a response to the physiological and morphological changes triggered by alpha factor itself (BALTANÁS *et al.* 2013).

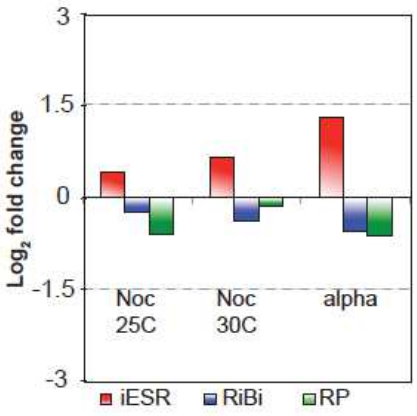


Fig 2. The ESR is not correlated with cell cycle phases nor growth rate.

Average $\log_2$ (fold-change) in transcript abundance for iESR, RP, and RiBi genes in nocodazole- and alpha factor-arrested cells compared to asynchronously growing cells.

To distinguish between these possibilities, we analyzed transcripts that fluctuate with the cell-cycle as reported by Spellman *et al* and Granovskaia *et al*, who monitored transcriptomic changes in synchronized, cycling cells (SPELLMAN *et al.* 1998; GRANOVSKAIA *et al.* 2010). Several lines of evidence show that the ESR is not associated with the cell cycle. First, less than 1% of ESR genes were classified as cycling, independent of arrest method, in either dataset (67 out of 853 ESR genes were identified by Spellman *et al* and 60 genes from Granovskaia *et al*). Second, ESR genes do not cycle in cells synchronized with mating pheromone (Fig 3) or other synchronization methods (SPELLMAN *et al.* 1998; GRANOVSKAIA *et al.* 2010). We furthermore found no correlation between ESR transcript abundance and transcript markers of different cell cycle phases. Thus, ESR activation is not correlated with cell cycle oscillation. One explanation for the G1 link reported by O'Duibhir *et al* is that synchronization by elutriation, the main dataset analyzed in that study, induces a strong stress response due to the centrifugation step (SPELLMAN *et al.* 1998). We also noticed a trace induction of the ESR at the beginning of the pheromone-synchronization experiments (see Fig 3), but this response vanished as cells progressed through the cell cycle. This indicates the ESR activation we observed with alpha factor treatment is likely responding to the pheromone itself, instead of the specific cell cycle phase.

ESR activation does not respond to growth reduction per se

Several studies found a linear correlation between ESR activation and growth rate in nutrient-restricted chemostats (REGENBERG *et al.* 2006; CASTRILLO *et al.* 2007; BRAUER *et al.* 2008). This model predicts that the complete growth arrest we report for nocodazole- or pheromone-treated cells should lead to maximal ESR activation compared to cells growing but with a reduced rate. To test this, we analyzed the data of Brauer *et al*, who monitored transcriptome changes across chemostat cultures with various growth rates (BRAUER *et al.* 2008), and fit our data into a linear model (Fig 4). Since arrested cells are not dividing, we set the dilution rate μ

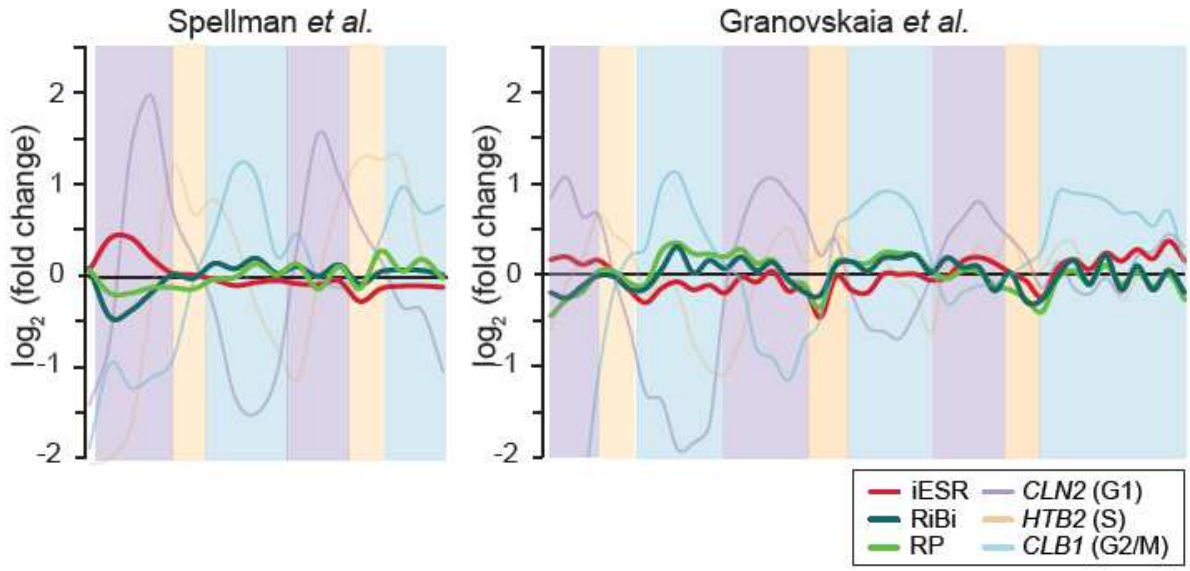


Fig 3. Average log2(fold-change) in transcript abundance for iESR, RP, and RiBi genes in alpha-factor synchronized cells from Spellman *et al* (left) and from Granovskaia *et al* (right). Expression of *CLN2* (which peaks in G1 phase), *HTB2* (S phase), and *CLB1* (G2/M phase) and boundaries of G1 (purple), S (beige), G2/M (blue) are shown for reference.

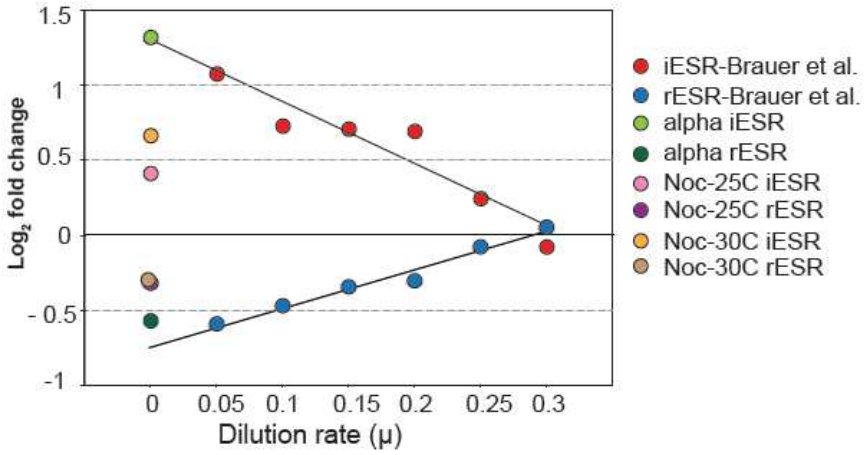


Fig 4. Magnitude of ESR expression does not fit the linear model.

The average ESR transcript abundance (relative to asynchronously growing cells) was plotted against the dilution rate (μ), as reported by Brauer *et al* in response to glucose limitation and as measured here. Arrested cells were set to a dilution rate of zero. The linear fit was set according to Brauer *et al* data.

(which is proportional to division rate) to zero. We found that the average abundance of ESR transcripts in arrested cells does not fit the linear expectation. With regard to the rESR genes in particular, transcript levels in both nocodazole- and pheromone-arrested cells (relative to the asynchronous baseline) were significantly higher than predicted, even though the cells are not growing or dividing (Fig 4). The mild induction of ESR genes in nocodazole-arrested cells falls well below the predicted expectation for iESR transcripts, although iESR transcripts were along the line in response to pheromone (Fig 4). Together, these results show that, at least in arrested cells, the ESR is not explained as simply a signature of reduced growth.

Arrested cells actively respond to stress stimuli

To further test if the ESR is a stress response or byproduct of halted growth, we challenged arrested cells with sodium chloride (NaCl) or heat-shock stress and monitored transcriptome changes. Both nocodazole- and pheromone-arrested cells responded to NaCl and heat shock, with dynamics that were comparable to asynchronous cells (Fig 5). At the peak response time, ESR activation was highly correlated in arrested versus asynchronous cells. In the case of nocodazole-arrested cells, the linear correlation with asynchronous cells was very high (R^2 = 0.9) in response to both stresses, with a slightly reduced slope (~ 0.8) (Fig 6A). A similar result was observed in alpha factor-arrested cells, albeit with slight weaker correlation (R^2 ~ 0.7) and lower slope (0.35 or 0.5 for the NaCl and heat responses, respectively) (Fig 6B). Furthermore, cells arrested with nocodazole and alpha factor both showed a dose-dependent transcriptome response to stress, with stronger ESR activation in response to a higher dose of salt or heat-shock stresses (Fig 7). Overall, these experiments demonstrate that arrested cells still actively respond to stress.

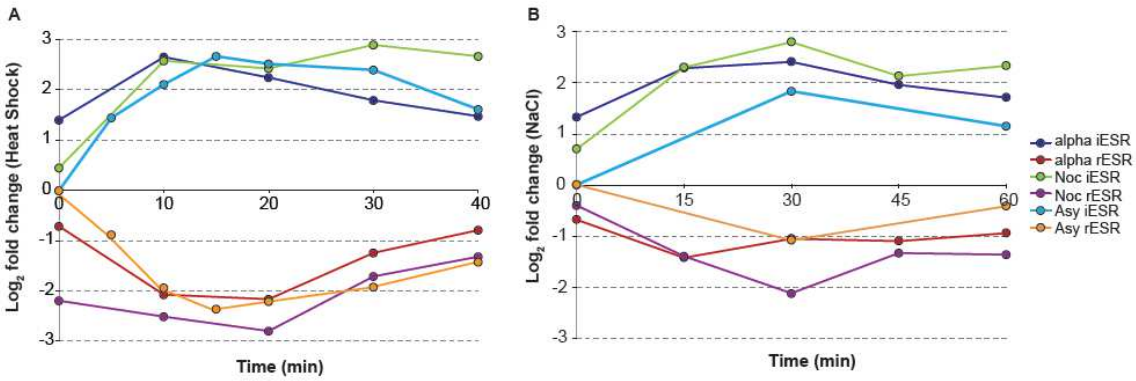


Fig 5. Dynamic changes in ESR transcript abundance upon stress treatment in growing and arrested cells. Average of log2 (fold change) in transcript abundance for iESR and rESR of asynchronous and arrested cells, normalized to the expression levels seen in synchronously growing cells. Data of asynchronously growing cells responding to heat shock (25°C to 37°C) or NaCl (0.7N) were taken from (GASCH *et al.* 2000; LEE *et al.* 2011).

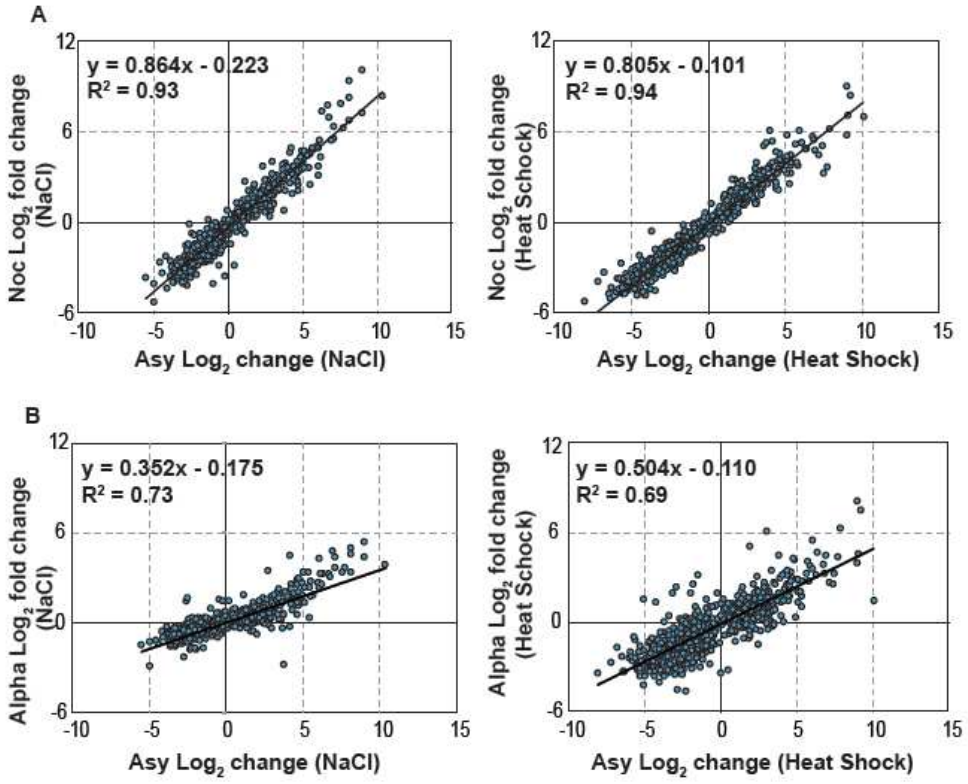


Fig 6. Correlation of ESR response between arrested and asynchronous cells upon stress.

A. Correlation of ESR between nocodazole-arrested and asynchronous cells upon salt (left) and heat shock (right), respectively. Linear relationship and R^2 are showed on top-left corner.

B. Correlation of ESR is alpha factor-arrested versus asynchronous cells upon salt and heat shock, respectively.

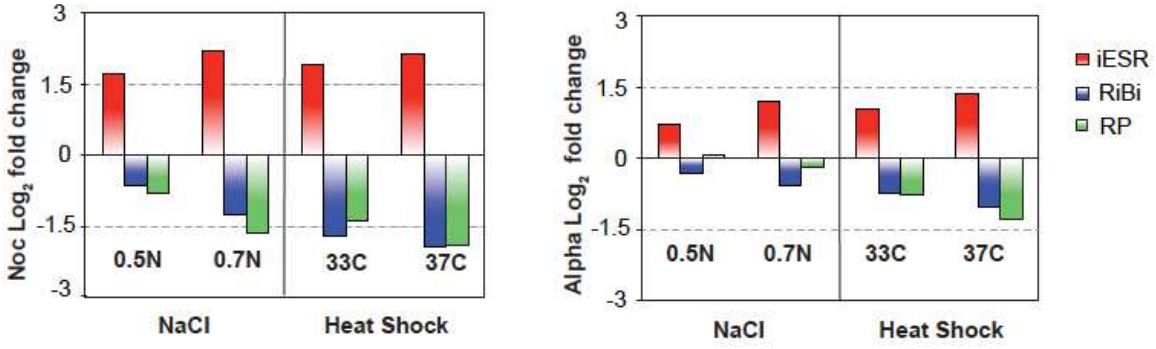


Fig 7. Arrested cells actively respond to stress in a dosage dependent manner

Nocodazole- and alpha factor–arrested cells (left and right panel, respectively) are challenged with different dosage of stimuli, slat and heat shock. Average of Log_2 fold change in iESR, RP and RiBi genes are shown. In heat shock experiments, cells grow at 25 degree before stress applied.

Absolute ESR transcript levels respond to stress similarly in arrested and growing cells.

Although the ESR was induced upon stress in arrested cells, the magnitude of the ESR gene expression changes before and after stress was generally smaller, particularly for pheromone-arrested cells, compared to asynchronous cells. One explanation for the smaller fold-change is that the ESR was already slightly activated upon nocodazole and alpha factor treatments, before the stress was applied (Fig 2). To account for this, we adjusted the absolute ESR transcript levels to the asynchronous baseline, by summing the response to arrest treatment and to stress. Remarkably, the absolutely iESR transcript levels after stress were largely similar to that seen in asynchronously growing cells across all conditions studied (Fig 8). This baseline adjustment also significantly improved the correlation and slope with the asynchronously growing cells (Fig 9). Interestingly, the baseline adjustment revealed that rESR transcripts remained at higher abundance (i.e. repressed less in response to stress) in arrested cells compared to asynchronously growing cells (Fig 8), especially in pheromone-arrested cells (see Discussion).

Discussion

A long-standing model in bacterial physiology is that cellular growth rate is dependent on the rate of production of new ribosomes (ECKER and SCHAECHTER 1963; KJELDGAARD and KURLAND 1963; MAALOE and KJELDGAARD 1966; NOMURA 1999). Since many of the rESR genes (including RP and RiBi genes) can be directly related to ribosome production, the link between rESR expression and growth rate fits this prevailing model. Indeed, reduced growth rate due to nutrient restriction (either in a chemostat or due to mutation in cellular transporters (SCOTT *et al.* 2014)) is correlated with reduced RP and RiBi proteins (HUI *et al.* 2015). However – when growth rate is restricted due to translational inhibition, the opposite trend emerges: slow growth

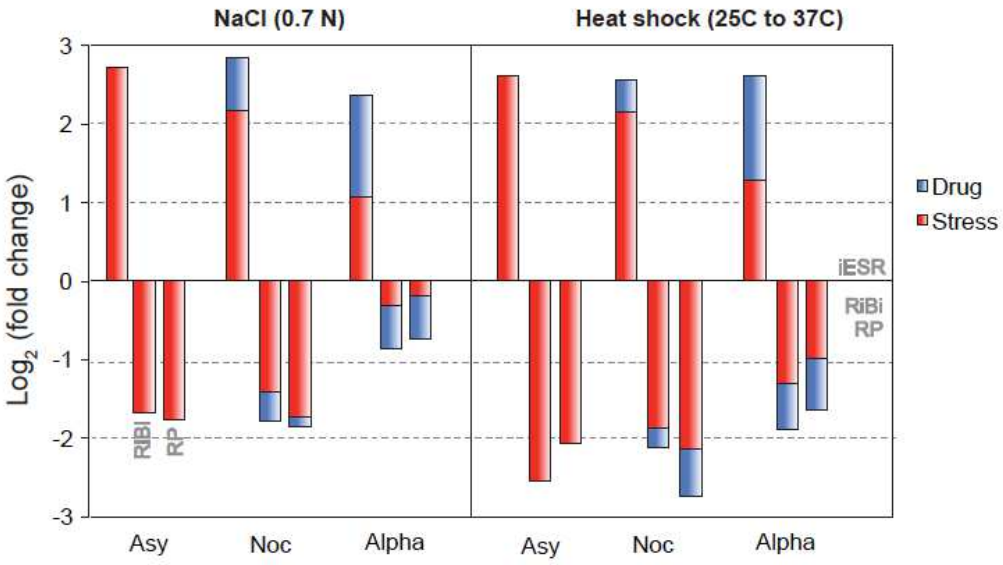


Fig 8. Total ESR transcript levels are similar across conditions.

Average of total transcript levels (relative to the asynchronous baseline) for iESR, RiBi and RP genes is shown as measured in asynchronously growing, nocodazole- or alpha factor arrested cells responding to the denoted stresses. The magnitude of ESR change due to arrest treatment is shown in blue and the magnitude of ESR change due to stress treatment is shown in red.

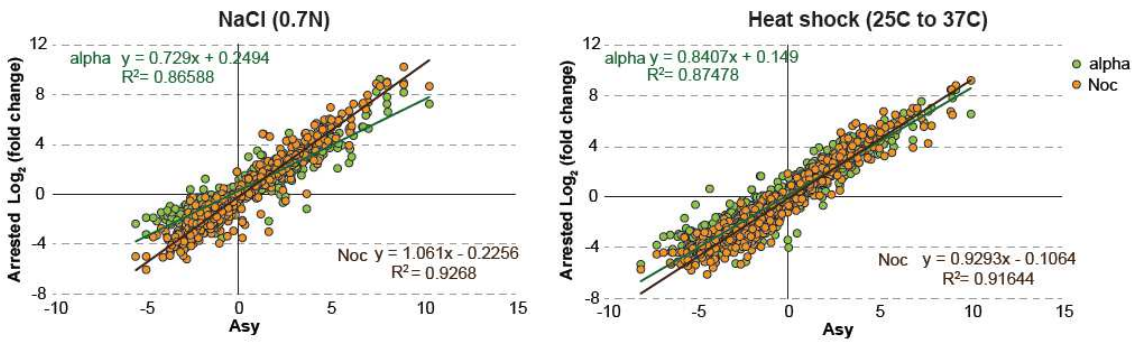


Fig 9. Correlation of total ESR expression between arrested cells and asynchronous cells upon stresses.

Total ESR expression is summed with the effect of arrest and stress response. Nocodazole (orange), alpha-arrested (green) and asynchronous cells were challenged with NaCl (left) and heat shock (right). Linear relation and R^2 are shown at corner: nocodazole (brown, lower right), alpha factor (green, upper left).

under these conditions is correlated with *higher* abundance of RPs and associated proteins, almost certainly because cells are trying to address the problem that actually causes growth reduction (HUI *et al.* 2015). Thus, the production of translational machinery does not respond exclusively to growth rate but instead is likely fine-tuned toward the specific challenges faced by the cell.

In this study, we show definitively that the yeast ESR cannot be simply explained as a byproduct of altered cell-cycle distribution or arrested growth. Arrest of growth and cell cycle progression did not trigger the strong ESR activation predicted by alternate models (Fig 2). Furthermore, ESR transcripts did not fluctuate with cell cycle phase in dividing cells and did not respond to the arrest point as proposed by O'Duibhir *et al* (Fig 3) (O'DUIBHIR *et al.* 2014). We argue that the mild activation of the ESR in response to arrest treatments used here is likely due to the physiological perturbation of the arrest method, rather than the arrest itself. Our results also showed that complete arrest of cell division did not activate the ESR according to expectation if ESR is linearly responsive to growth rate (Fig 4), implying ESR cannot be a solely signature of slow growth. One challenge with past chemostat studies is that it is difficult to distinguish the effect of slow growth from the external environmental disturbance, since the nutrient limitation that sets growth rate is likely a stress itself. Our results clearly demonstrates that growth rate should not be inferred based on the level of ESR activation, since the link between growth and ESR activation breaks down in arrested cells.

We also show that activation of the ESR is an active response to stress. Arrested cells show robust, dose-dependent ESR activation (albeit slightly weaker in some cases) in response to stress, even though they are not cycling or producing significant biomass (Fig 6 & 7). In fact, this is consistent with a study by Lu *et al*, who showed that despite growth reduction in a chemostat, the expression of over half of the stress responsive genes responded to both

external stress and chemostat-regulated growth rate, while the other half of responsive genes was not affected by growth rate and responded primarily to external stress (LU *et al.* 2009). Together, it is clear that the ESR is an active response to stress that can be completely decoupled from growth-rate changes.

Redirecting translational capacity: An alternate model for rESR repression

We have proposed an alternate model: that repression of rESR genes – including those encoding RPs and RiBi factors – is a stress-defense mechanism that serves to redirect transcriptional and translational capacity toward defense genes (LEE *et al.* 2011; CHASMAN *et al.* 2014). In actively growing cells, rESR transcripts are at highly transcribed and translated to ensure sufficient supply of ribosomes to new cells (WARNER 1999; GASCH 2002a). Nearly 90% of ribosomes are actively engaged in translation and largely directed to making rESR proteins themselves (WARNER 1999; ARAVA *et al.* 2003; VON DER HAAR 2008). It had been unclear how cells would have capacity to synthesize proteins from newly made iESR transcripts, which are generally well translated in response to stress (LEE *et al.* 2011), if the majority of ribosomes were engaged in rESR proteins production. We showed previously that the drop in rESR transcript levels upon salt stress is directly proportionate to the demand for ribosomes in order to translate newly made messages. Failure to repress rESR genes caused ribosomes to remain engaged with rESR transcripts after stress (LEE *et al.* 2011). More recent work from our lab suggests that competition for RNA polymerase may also play a role in the drop in rESR transcription in favor of the production of iESR and other stress transcripts (CHASMAN *et al.* 2014).

Consistent with this model, our results reported here show that, even when cells are arrested and barely growing, the rESR genes still become significantly repressed upon stress treatment.

This observation confirms that, under the conditions studied in this work, repression of the rESR genes is not directly linked to growth but instead has a specific role in the stress response. However, when accounting for the response to arrest method and stress treatment, the absolute levels of rESR transcripts were repressed less strongly in arrested cells compared to actively growing cells, particularly in response to pheromone arrest (Fig 8).

We propose that the magnitude of rESR repression is set by the relative demand for new translation compared to current translational capacity. Polysome profiles indicate that arrested cells, especially those treated with pheromone, are operating below their translational capacity, as indicated by the elevated ratio in M/P fractions (Fig 1). Absolute abundance of iESR transcripts is very similar across all treatments, whether cells are growing or not – but rESR transcripts show milder reduction in arrested cells. We propose that the reason for this is that cells are already functioning below their translational capacity, such that there is less need to redirect translation from rESR genes and transcripts. While further experiments are required to test the model, it would explain why rESR repression is so well anti-correlated with iESR production across a wide array of diverse environmental conditions.

Stress defense and growth control: interconnected responses in the allostasis program

Activation of the ESR is clearly a direct response to stress, but expression of the program (particularly the rESR genes) is also correlated with nutrient-influenced growth rate and cell cycle progression. How can we reconcile these results? We propose that ESR activation and growth/cell-cycle control are covariates in the cellular allostasis program. Allostasis is distinct from homeostasis in that different cellular states may be required to support the same cellular output – maximal fitness – depending on the external environmental conditions. Upon treatment with a particular stress, cells may require a distinct cellular program to maintain proper functions.

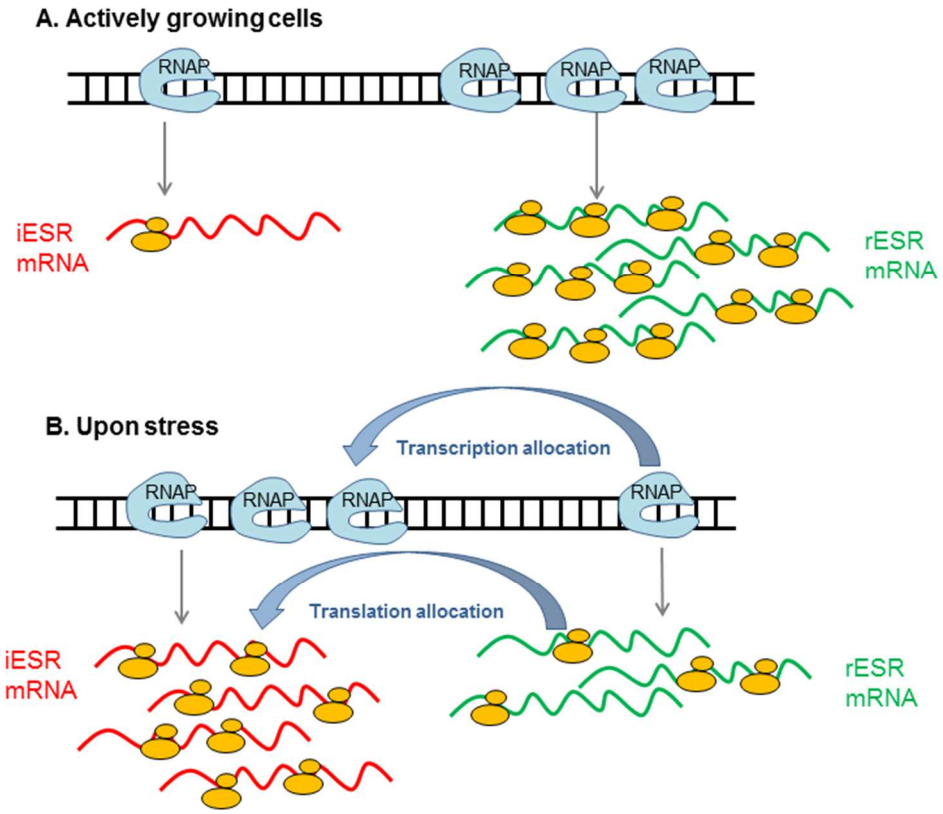


Fig 10. Model of cellular resources allocation upon stress.

A. In actively growing cells, rESR genes are highly transcribed and translated, occupying the majority of transcription and translation machinery.

B. Upon stress, cells reallocate transcription and translation machinery from rESR genes and transcripts to iESR genes and newly made transcripts.

We propose that, upon a sudden exposure to stress, cells induce expression of defense genes (while mediating other stress-dependent physiology changes to deal with the stress) (Fig 10). The induced expression of specific genes occurs at the expense of rESR transcription, in order to free up polymerase (CHASMAN *et al.* 2014). In order to translate the newly made mRNAs, cells reorganize their translational effort, reducing translation of rESR transcripts and rerouting machinery to iESR messages (LEE *et al.* 2011). To accommodate the sudden drop in new-ribosome synthesis, cells decrease growth rate and cell-cycle progression. Once cells have acclimated to the new conditions, growth often resumes but at a reduced rate. Thus, the reduction in growth rate frequently accompanies a stress response (either transiently in response to acute stress or perpetually in the presence of limited nutrients). This may be part of the cellular allostasis program that helps to prioritize and reallocate cellular transcription and translation capacity toward stress defense and acclimation, thus the need for cellular growth and biomass accumulation are suppressed and postponed. Resources reallocation in the cellular allostasis program thereby results in previously observed relation between the ESR and growth rate.

Materials and methods

Cell growth and RNA preparation

Strain BY4741 was grown > 7 generations to an optical density (OD_{600}) ~ 0.25 at 30°C in batch culture in YPD (1% yeast extract, 2% peptone and 2% dextrose) medium. Synchronized cells were treated with either 10 µg/ml Nocodazole (EMD Millipore, Bilerica, MA) or 40 µg/ml alpha factor (ZymoResearch, Irvine, CA) final concentrations for 2 hours before stress treatment. Asynchronous cells were grown > 7 generations to OD_{600} ~ 0.6 as described above. Samples were collected before synchronization, after synchronization and right before stress treatment, and after stress applied as indicated. The concentrations of nocodazole and alpha factor were

maintained during stress experiments despite media additions in some cases. Cells were harvested by spin at 3000 r.p.m. for 2 min followed by flash frozen in liquid nitrogen and maintained at −80°C until RNA extraction. Total RNA was extracted with hot phenol as previously described (GASCH 2002b), DNase-treated at 37°C for 30 min with TURBO DNase (Life Technologies), and then precipitated at −20°C in 2.5 M LiCl for 30 min. rRNA depletion of the DNase-treated total RNA was carried out with ScriptSeq Complete Kit H/M/R (Epicentre, Madison, WI), and samples were purified with a RNeasy MinElute Cleanup Kit (Qiagen).

RNA-Seq

cDNA library preparation of rRNA-depleted RNA was carried out with Index PCR Primers (Epicentre) and FailSafe PCR Enzyme Mix (Epicentre, Madison, WI) and purified with a Axy Prep MAG PCR Cleanup beads (Corning, Corning, NY). cDNA libraries were sequenced on Illumina's HiSeq 2500 System (UW-Madison DNA Sequencing Facility), generating single-end 100 bp reads. Sequencing reads were processed with Trimmomatic version 0.32 (BOLGER *et al.* 2014) and reads were mapped to the S288c reference 64 using Bowtie2 (LANGMEAD and SALZBERG 2012). HTseq version 0.6.1 (ANDERS *et al.* 2015) was used to sum read counts per gene, which were then normalized for gene size and the number of reads generated per library via reads per kilobase per million-mapped reads (RPKM).

Polysome profiling

Polysome profiles were collected as described in (ARAVA *et al.* 2003) using a 5-50% continuous sucrose gradients analyzed on a Teledyne Isco Fractionation unit.

References

AIROLDI, E. M., C. HUTTENHOWER, D. GRESHAM, C. LU, A. A. CAUDY *et al.*, 2009 Predicting cellular growth from gene expression signatures. PLoS Comput Biol **5:** e1000257.

ANDERS, S., P. T. PYL and W. HUBER, 2015 HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics **31:** 166-169.

ARAVA, Y., Y. WANG, J. D. STOREY, C. L. LIU, P. O. BROWN *et al.*, 2003 Genome-wide analysis of mRNA translation profiles in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A **100:** 3889-3894.

BALTANÁS, R., A. BUSH, A. COUTO, L. DURRIEU, S. HOHMANN *et al.*, 2013 Pheromone-induced morphogenesis improves osmoadaptation capacity by activating the HOG MAPK pathway. Sci Signal **6:** ra26.

BATTESTI, A., N. MAJDALANI and S. GOTTESMAN, 2011 The RpoS-mediated general stress response in Escherichia coli. Annu Rev Microbiol **65:** 189-213.

BERRY, D. B., and A. P. GASCH, 2008 Stress-activated genomic expression changes serve a preparative role for impending stress in yeast. Mol Biol Cell **19:** 4580-4587.

BOLGER, A. M., M. LOHSE and B. USADEL, 2014 Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30:** 2114-2120.

BRAUER, M. J., C. HUTTENHOWER, E. M. AIROLDI, R. ROSENSTEIN, J. C. MATESE *et al.*, 2008 Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast. Mol Biol Cell **19:** 352-367.

BROWN, J. B., N. BOLEY, R. EISMAN, G. E. MAY, M. H. STOIBER *et al.*, 2014 Diversity and dynamics of the Drosophila transcriptome. Nature **512:** 393-399.

CASTRILLO, J. I., L. A. ZEEF, D. C. HOYLE, N. ZHANG, A. HAYES *et al.*, 2007 Growth control of the eukaryote cell: a systems biology study in yeast. J Biol **6:** 4.

CAUSTON, H. C., B. REN, S. S. KOH, C. T. HARBISON, E. KANIN *et al.*, 2001 Remodeling of yeast genome expression in response to environmental changes. Mol Biol Cell **12:** 323-337.

CHASMAN, D., Y.-H. HO, D. B. BERRY, C. M. NEMEC, M. E. MACGILVRAY *et al.*, 2014 Pathway connectivity and signaling coordination in the yeast stress-activated signaling network. Mol Syst Biol **19:** 759.

ECKER, R. E., and M. SCHAECHTER, 1963 Ribosome content and the rate of growth of Salmonella Typhimurium. Biochim Biophys Acta **76:** 275-279.

GASCH, A. P., 2002a The Environmental Stress Response: a Common Yeast Response to Environmental Stresses. Springer-Verlag: Heidelberg.

GASCH, A. P., 2002b Yeast genomic expression studies using DNA microarrays. Methods Enzymol **350:** 393-414.

GASCH, A. P., 2007 Comparative genomics of the environmental stress response in ascomycete fungi. Yeast **24:** 961-976.

GASCH, A. P., P. T. SPELLMAN, C. M. KAO, O. CARMEL-HAREL, M. B. EISEN *et al.*, 2000 Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell **11:** 4241-4257.

GRANOVSKAIA, M. V., L. J. JENSEN, M. E. RITCHIE, J. TOEDLING, Y. NING *et al.*, 2010 High-resolution transcription atlas of the mitotic cell cycle in budding yeast. Genome Biol **11:** R24.

HO, Y. H., and A. P. GASCH, 2015 Exploiting the yeast stress-activated signaling network to inform on stress biology and disease signaling. Curr Genet.

HUI, S., J. M. SILVERMAN, S. S. CHEN, D. W. ERICKSON, M. BASAN *et al.*, 2015 Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. Mol Syst Biol **11:** 784.

JOHNSTON, G. C., R. A. SINGER, S. O. SHARROW and M. L. SLATER, 1980 Cell division in the yeast Saccharomyces cerevisiae growing at different rates. Microbiology **118:** 479-484.

KJELDGAARD, N. O., and C. G. KURLAND, 1963 The distribution of soluble and ribosomal RNA as a function of growth rate. J Mol Biol **6:** 341-348.

LANGMEAD, B., and S. L. SALZBERG, 2012 Fast gapped-read alignment with Bowtie 2. Nat Methods **9:** 357-359.

LEE, M. V., S. E. TOPPER, S. L. HUBLER, J. HOSE, C. D. WENGER *et al.*, 2011 A dynamic model of proteome changes reveals new roles for transcript alteration in yeast. Mol Syst Biol **19:** 514.

LU, C., M. J. BRAUER and D. BOTSTEIN, 2009 Slow growth induces heat-shock resistance in normal and respiratory-deficient yeast. Mol Biol Cell **20:** 891-903.

MAALOE, O., and N. O. KJELDGAARD, 1966 *Control of macromolecular synthesis*, New York.

NOMURA, M., 1999 Regulation of Ribosome Biosynthesis in Escherichia coli and Saccharomyces cerevisiae: Diversity and Common Principles. J Bacteriol **181:** 6857-6864.

O'DUIBHIR, E., P. LIJNZAAD, J. J. BENSCHOP, T. L. LENSTRA, D. VAN LEENEN *et al.*, 2014 Cell cycle population effects in perturbation studies. Mol Syst Biol **10:** 732.

REGENBERG, B., T. GROTKJAER, O. WINTHER, A. FAUSBØLL, M. AKESSON *et al.*, 2006 Growth-rate regulated genes have profound impact on interpretation of transcriptome profiling in Saccharomyces cerevisiae. Genome Biol **7:** R107.

ROETZER, A., C. GREGORI, A. M. JENNINGS, J. QUINTIN, D. FERRANDON *et al.*, 2008 Candida glabrata environmental stress response involves Saccharomyces cerevisiae Msn2/4 orthologous transcription factors. Mol Microbiol **69:** 603-620.

SALDANHA, A. J., M. J. BRAUER and D. BOTSTEIN, 2004 Nutritional homeostasis in batch and steady-state culture of yeast. Mol Biol Cell **15:** 4089-4104.

SCOTT, M., S. KLUMPP, E. M. MATEESCU and T. HWA, 2014 Emergence of robust growth laws from optimal regulation of ribosome synthesis. Mol Syst Biol **10:** 747.

SPELLMAN, P. T., G. SHERLOCK, M. Q. ZHANG, V. R. IYER, K. ANDERS *et al.*, 1998 Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. Mol Biol Cell **9:** 3273-3297.

VAN DE GUCHTE, M., P. SERROR, C. CHERVAUX, T. SMOKVINA, S. D. EHRLICH *et al.*, 2002 Stress responses in lactic acid bacteria. Antonie Van Leeuwenhoek **82:** 187-216.

VON DER HAAR, T., 2008 A quantitative estimation of the global translational activity in logarithmically growing yeast cells. BMC Syst Biol **2**.

WARNER, J. R., 1999 The economics of ribosome biosynthesis in yeast. Trends Biochem Sci **24:** 437-440.

WOHLBACH, D. J., N. ROVINSKIY, J. A. LEWIS, M. SARDI, W. S. SCHACKWITZ *et al.*, 2014 Comparative genomics of Saccharomyces cerevisiae natural isolates for bioenergy production. Genome Biol Evol **6:** 2557-2566.

Chapter 4

Future directions

Future directions

Proper stress responses are pivotal for cells to survive and adapt to new environments, particularly for free-living microbes that face constantly changing environments. In this thesis, I have discussed the integration and coordination of a signaling network upon stress in Chapter 2. The inferred signaling network revealed complex cross-talk among signaling pathways and identified new regulators that helped to formulate new hypotheses. We also proposed that repression of the rESR plays a role in allocating transcriptional capacity to iESR genes. In Chapter 3, I have shown that the ESR is truly a stress response, instead of a byproduct of growth arrest or increased G1 populations. We proposed that growth rate and ESR are covariates in the cellular allostatis program, instead of directly dependent on each other. We further proposed that repression of rESR genes upon stress is important to reallocate translational capacity to newly made mRNAs. Although much has already been studied in this field and discussed in this thesis, there are still many questions left to be answered about stress responses. Here I'd like to discuss several remaining questions that are directly related to our hypothesis and I am particularly interested in.

How is rESR repression related to translational capacity?

We proposed that rESR repression is important for translational reallocation in Chapter 3 and previous published work (LEE *et al.* 2011), since rESR repression would release ribosomes to translate newly made transcripts. I'd like to further test this model via measuring translation propensity change upon stresses in arrested cells, which have less rESR repression upon stress. Some pioneer studies showed cells reprogram translational profiles upon stress by mounting a global repression of the translation of pre-existing mRNAs, while subsets of newly transcribed transcripts are selectively translated (KUHN *et al.* 2001; SHENTON *et al.* 2006; MELAMED *et al.* 2008). One particular study addressed this question in osmotic stress (1M NaCl for 1 hour, which reaches the maximal translation inhibition) using microarray analysis of

monosomal and polysomal associated mRNAs pools (MELAMED *et al.* 2008). Consistent with others, they found targets of Msn2/Msn4 are translated actively (i.e. potentiated) in spite of global translation repression. Although these results provide a general idea that cells selectively translate stress-defense related transcripts, these experiments were carried out in actively growing cells where translation machinery is already engaged near capacity and devoted to cell growth. Therefore I'd like to further test our model by measuring changes in translation upon stresses in arrested cells, to test why they have less repression of rESR transcripts. In order to test our model, it will be interesting to see if the cellular status may affect translation profiles, i.e. in arrested cells.

To address this question, I proposed one particular experiment to stress asynchronously growing cells and arrested cells by collecting pre-stress and post-stress mRNAs fractionated by polysome association, and then subjecting samples to RNA-seq. This analysis will identify distributions of each mRNA across polysome fractions in different tested conditions. The correlation between abundance of transcripts and polysomal association will reveal if cells may selectively translate target transcripts to cope with stress – for example, mRNAs that are regulated at the translational level may have a shift in polysomal association in the absence of, or in excess of, the increase in mRNA levels.

The comparison among asynchronously growing and arrested cells will provide models on how cells reallocate translational capacity when cells are arrested and upon stress. We showed that arrested cells have a lower magnitude of rESR repression upon stress. One possibility is that this situation leads to fewer ribosomes being released and therefore there may be a delay or reduced polysome association, and thus protein production, from stress-induced transcripts. However, an alternative model is that arrested cells are already operating below their translational capacity. Arrested cells possess lower levels of polysomal-associated mRNAs and

presumably have more ribosomes available for potential translation. Therefore, upon stress, arrested cells may still have comparable translation and production of iESR proteins despite weaker rESR repression. Further experiments of testing levels of free/available ribosomes and measuring global proteome production in arrested cells will be needed to understand what factors affect how cells reallocate translational capacity.

Is translation reallocation important in stress defense?

As already demonstrated in bacteria, translational reallocation plays an important role in response to environment changes (SCOTT *et al.* 2010; SCOTT *et al.* 2014). I' d like to test this model in yeast upon stress. One potential experiment is to overexpress a foreign protein (for example, LacZ or GFP) in cells, to consume a large portion of their translational capacity, and challenge cells with stresses. Based on studies in bacteria, overexpressing proteins will consume significant translation capacity and leave a relative small portion of translational capacity for cells to grow and utilize for other physiological processes (SCOTT *et al.* 2010). Thus, these cells were shown to grow slower. The question is if they are more stress-sensitive. One experiment is to over-express a foreign protein in yeast and measure the cell's growth rate and stress sensitivity. If cells grow slower, as expected, and do possess higher stress sensitivity, it may reflect that they have lower or delayed production of stress-defense proteins, a model that can be tested with proteomics or targeted Western blots. If true, this result would be in conflict with previous observations that slower growing cells are more stress tolerant (LU *et al.* 2009), and thereby decouple the relation between growth rate and stress tolerance.

How are growth-promoting signaling pathways turned-off upon stress?

Our results in Chapter 3 suggest that ESR regulation and growth control are coordinated, but not necessarily directly regulating one another. How this occurs at the signaling level is not

clear. The TORC1 and PKA pathways are the two major nutrient-sensing and growth-promoting signaling pathways, and are also the two major pathways that regulate expression of rESR transcripts. The repression of these two pathways is generally observed upon diverse stresses; however the upstream regulators that repress these two pathways are still unclear. One possibility is that, instead of one single regulator/protein that dominates the regulation, there are multiple proteins as a regulatory network to regulate these two pathways. In fact, the complicated connectivity observed in our salt signaling network may support this model (CHASMAN *et al.* 2014), and this may indicate the growth control is sensitively adjusted to the demands. Another possibility is that, instead of regulators modulating the activity of these pathways, other secondary messengers (for example, cAMP) determine the activity of the pathways. Therefore, multiple factors along with other aspects of physiological status are involved in affecting the activity of these pathways. Actually, in our network bifurcation study (CHASMAN *et al.* 2014), several top regulators are found to be involved in cAMP regulation. Future experiments addressing this question will likely be informative. The result will also provide a general stress-defense model across various conditions.

Dynamic changes of signaling network upon stress

In Chapter 2, we have already provided an inferred network which reveals intensive cross-talk among different pathways, however, it lacks dynamic information – the network was generated based on the peak time points in the salt-responsive transcriptome and proteome. Upon stress, cells undergo several different phases to cope with stress and adaption: initial arrest stage (transcription, translation and cell-cycle arrest), stress-acclimation stage (ESR expression, translation reallocation and physiological reprograming), and restorative stage (cells adjust homeostasis to the new environment). The process may take up to hours dependent on the stimuli. Therefore, a dynamic model of a signaling network will reveal more details of how cells

sequentially modulate different physiological changes at each stage by coordinating these signaling pathways.

Taken together, proper stress response is required for cells to survive and adapt. Understanding the signaling networks activated during the stress response and deciphering the role of the ESR will provide fundamental concepts of how cells cope with stress in a systematic view. Also, these studies reveal a broader view of the relation between growth control and stress defense. And since many major signaling pathways are conserved in higher organisms, these observations may provide insights of stress response in higher organisms.

References

*CHASMAN, D., Y.-H. HO, D. B. BERRY, C. M. NEMEC, M. E. MACGILVRAY et al., 2014 Pathway connectivity and signaling coordination in the yeast stress-activated signaling network. Mol Syst Biol **19:** 759.*

*KUHN, K. M., J. L. DERISI, P. O. BROWN and P. SARNOW, 2001 Global and specific translational regulation in the genomic response of Saccharomyces cerevisiae to a rapid transfer from a fermentable to a nonfermentable carbon source. Mol Cell Biol **21:** 916-927.*

*LEE, M. V., S. E. TOPPER, S. L. HUBLER, J. HOSE, C. D. WENGER et al., 2011 A dynamic model of proteome changes reveals new roles for transcript alteration in yeast. Mol Syst Biol. **19:** 514.*

*LU, C., M. J. BRAUER and D. BOTSTEIN, 2009 Slow growth induces heat-shock resistance in normal and respiratory-deficient yeast. Mol Biol Cell **20:** 891-903.*

*MELAMED, D., L. PNUELI and Y. ARAVA, 2008 Yeast translational response to high salinity: global analysis reveals regulation at multiple levels. RNA **14:** 1337-1351.*

*SCOTT, M., C. W. GUNDERSON, E. M. MATEESCU, Z. ZHANG and T. HWA, 2010 Interdependence of cell growth and gene expression: origins and consequences. Science **330:** 1099-1102.*

*SCOTT, M., S. KLUMPP, E. M. MATEESCU and T. HWA, 2014 Emergence of robust growth laws from optimal regulation of ribosome synthesis. Mol Syst Biol **10:** 747.*

*SHENTON, D., J. B. SMIRNOVA, J. N. SELLEY, K. CARROLL, S. J. HUBBARD et al., 2006 Global translational responses to oxidative stress impact upon multiple levels of protein synthesis. J Biol Chem **281:** 29011-29021.*

Appendix A

Supplementary information of Chapter 2

Pathway connectivity and signaling coordination in the yeast stress-activated signaling network

Deborah Chasman, Yi-Hsuan Ho, David B Berry, Corey M Nemec, Matthew E MacGilvray, James Hose, Anna E Merrill, M Violet Lee, Jessica L Will, Joshua J Coon, Aseem Z Ansari, Mark Craven, Audrey P Gasch

This Work is published on *Molecular Systems Biology* (2014) 10: 759

Appendix A1 Computational methods

We developed an integer linear programming-based approach (IP, for short) to infer the path- ways through which the signaling regulator mutants perturb the regulation of their affected down- stream targets. Using a *background network* of directed and undirected intracellular interactions, the method infers a *subnetwork* that predicts the *paths* by which each interrogated *source* regulator is connected to its downstream *targets*, identified as dysregulated genes in the source mutant responding to NaCl treatment. Each path is a directed, linear chain of interactions between yeast gene products. The final interaction in each path represents the regulation of the target gene by a sequence-specific transcription factor or RNA-binding protein. The IP's objective function favors the inclusion of salt-responsive proteins, defined as those with differential phosphorylation upon NaCl treatment (phospho-proteomic hits) and those required for normal fitness during the NaCl response (fitness-contribution hits).

A1.1 Outline of the approach

In this section, we present a detailed description of the three steps of our inference method, illustrated in Figure 3.

Step 1: Gather the experimental data and a background network. Figure 3A shows the input to the method: data from experiments that interrogate the yeast salt stress response (as described in the manuscript), and a background network of intracellular interactions. The background net- work contains three types of nodes: proteins (shown as ellipses), TFs and RNA-binding proteins (diamonds), and genes/mRNAs (rectangles). The experimental data comes in the following forms:

- *Source-target pairs*, each consisting of one interrogated *source* regulator and one of its identified downstream gene/mRNA targets (sources are shown with bold outlines;

targets are shown as rectangles)

- *Receptor-source pairs*, where necessary, which represent directed relationships gathered from the literature. For example, to capture upstream stress sensors, we included knowledge about two transmembrane receptors that are upstream of Hog1.
- *Fitness-contribution hits*, genes that result in defective stress resistance acquisition when they are knocked out (shown in blue in Figure 3)
- *Phospho-proteomic hits*, proteins that are differentially phosphorylated under salt stress conditions (shown in yellow in Figure 3)

We refer to the fitness-contribution hits and phospho-proteomic hits together as *hits*. The sets are not mutually exclusive. The background network (described in Section S1.2.3) consists of a variety of types of bio- logical interactions, namely protein-protein and protein-nucleic acid interactions, gathered from public databases. It is represented as a graph in which the nodes represent either proteins or genes/mRNAs, and the edges represent interactions between them. If the direction of an inter- action is known (as in a kinase-substrate interaction), then the corresponding edge is directed. Otherwise, the edge is undirected.

Step 2: Generate candidate paths to explain experimental observations. The inferred sub-network must provide directed paths between each source and all of its targets, and so we first use the background network to enumerate all possible candidate paths. Each candidate source- target path must end with an interaction that represents regulation of the target mRNA: either the binding of a transcription factor (TF) to the target's gene, or the binding of an RNA-binding protein (RBP) to the target's mRNA. We also enumerate other kinds of candidate paths in order to capture connections among salt-responsive proteins that may be missed by cataloging only source-target paths. In Figure 3B, we show the three kinds of paths that we identify:

- *Source-target paths*, which connect interrogated sources to their targets
- *Receptor-source paths*, which offer paths between upstream stress sensors and downstream sources
- *Fitness-contribution hit-source* and *source-source paths*, which identify connections between potential regulators

The inferred subnetwork is a collection of directed paths that make a consistent assignment for the direction of each edge.

Step 3: Solve an IP to infer an ensemble of subnetworks. To perform inference, we construct an integer linear program (Figure 3C) whose constraints and objective functions describe a sub- network that conforms to the following desiderata:

- Each source is connected to of all of its targets by directed paths
- Any provided upstream relationships (receptor-source pairs) are explained by directed paths
- Directed paths reveal connections between fitness-contribution hits and sources that are proximal in the background network
- Each edge is assigned only one direction
- Proteins with phospho-proteomic or fitness evidence are favored for inclusion in the subnet- work
- The subnetwork includes a minimal number of nodes that are not supported by experimental evidence

Because both the experimental data and the background network are incomplete, there will be many possible inferred subnetworks that explain the experimental data equally well. To

quantify our confidence in the relevance of each protein and interaction, we infer an ensemble of optimal subnetworks. Confidence in a protein (interaction, path) is calculated as the fraction of the subnet- works in the ensemble that predict that the protein (interaction, path) is relevant.

Figure 3C shows two inferred relevant subnetworks for the given example. In each subnetwork, the predicted relevant edges (interactions) and nodes (proteins) are indicated with black outlines; rejected edges and nodes are in grey. A direction has been inferred for each formerly undirected interaction that is predicted to be relevant.

A1.2 Data supplied to subnetwork inference method

The inputs to the IP method are the data capturing the cell's response to salt stress (transcriptome changes, fitness-contributions, and phosphorylation responses), additional, known relationships between stress response factors curated from literature, and a background network consisting of publicly available yeast intracellular interactions.

A1.2.1 Experimental data

The primary goal of the approach is to provide explanations for the measured salt-specific transcriptomic changes. We also use two additional sources of salt-specific experimental data.

Source-target pairs. From the transcriptomic data measured in each of the original signaling mu- tants, we identified the set of downstream genes with dysregulated salt-responsive expression. We then extracted what we refer to as *source-target pairs*, each consisting of a single signaling protein (*source*) and a gene that was dysregulated in the mutant under salt stress (*target*). We divided the set of pairs into two categories, based on whether or not the source is a known transcription factor (TF). For non-TF sources, the effect on the targets is

assumed to be indirect, and mediated by interactions between other gene products. For sources that are TFs, it is assumed that they affect their targets directly.

Fitness-contribution hits. Previously, Berry *et al.* (2011) identified yeast mutants that conferred a defect in acquired stress resistance after salt pretreatment (BERRY *et al.* 2011). For each of these mutants, we ap- plied the label *fitness-contribution hit* to the gene product represented in our background network, because of the mutation's negative effect on yeast fitness under salt stress.

Phospho-proteomic hits. We applied the label *phospho-proteomic hit* to each protein that showed differential phosphorylation under salt stress).

A1.2.2 Indirect interactions from domain knowledge

Our method can take advantage of domain knowledge about the salt stress response in order to provide a scaffold for the inferred subnetwork. Here, we wanted to capture the upper-most stress sensors that may otherwise be missed in connecting sources to their downstream targets. We identified well-known indirect relationships between two transmembrane receptors, Sln1 and Sho1, and one of the sources, Hog1 (SAITO and TATEBAYASHI 2004). We provided this information to our method as *receptor-source* pairs: (Sln1, Hog1) and (Sho1, Hog1). During inference, the method selects directed paths from the two receptors to Hog1.

A1.2.3 Background network

In constructing our background network, we identified a variety of binary interactions that are

relevant to intracellular signaling and gene expression regulation. The background network represents protein-protein (including kinase-substrate), protein-DNA, and protein-RNA interactions. Physical protein-protein interactions (PPIs) were sourced primarily from the BioGRID database (STARK *et al.* 2006); these included both undirected PPIs and directed post-translational modifications, such as ubiquitination. To identify high-confidence interactions, we retrieved interactions that were supported by at least two separate experimental methods. In order to specifically capture signaling pathways, we supplemented the PPI network with additional high-confidence protein-protein interactions involving kinases and phosphatases (BERGKESSEL *et al.* 2011; FASOLO *et al.* 2011; SHARIFPOOR *et al.* 2011) . Directed kinase-substrate and phosphatase-substrate interactions were identified from BioGRID, data published by Ptacek *et al.*(PTACEK *et al.* 2005), and the KID database(SHARIFPOOR *et al.* 2011) . Additionally, we extracted binary interactions from metabolic pathway data (HEAVNER *et al.* 2012) represent interactions between enzymes that catalyze adjacent metabolic reactions, and information about which proteins are components of annotated protein complexes (PU *et al.* 2009; HEAVNER *et al.* 2012). To represent the direct regulation of the target genes/mRNAs, we gathered protein-DNA(GUELZIM *et al.* 2002; MACISAAC *et al.* 2006; EVERETT *et al.* 2009; ABDULREHMAN *et al.* 2011; VENTERS *et al.* 2011) and protein-RNA (HOGAN *et al.* 2008; SCHERRER *et al.* 2010; TSVETANOVA *et al.* 2010) interactions from several publications. We supplemented these with two sources of protein-DNA interactions gathered under stress conditions (NI *et al.* 2009; HUEBERT *et al.* 2012). After manual inspection of the background network neighborhoods of the interrogated mutants, we added a set of 17 missing interactions between the mutants and nearby regulators based on known interactions in the literature. While the types of biological interactions in the background network are rich and diverse, we use a simplified representation as input to the computational method. The background network is represented as a graph, in which nodes represent genes and gene products, and edges represent interactions. A gene may be represented as two separate nodes in the background

network: one representing the protein, and, for targets, one representing the DNA or mRNA. Each interaction may have a direction: for example, transcriptional regulatory interactions are directed, but most protein-protein interactions are not. (Protein-protein interactions representing post-translational modifications, such as phosphorylation, are directed.) We represented each protein complex as an additional node, and added a directed edge to the complex from each of its component proteins. Using protein-protein interactions, we also inferred undirected interactions between protein complexes and between a complex and a single protein. If more than 50% of the possible protein-protein interactions between two protein complexes (or between a complex and a protein) were present in the data, then we added an undirected edge between the two complexes (or complex and protein). Finally, we collapsed Dot6 and its paralog Tod6 into the same protein node in the background network, because they have been observed to function redundantly and because our measured targets were identified in a *dot6*Δ*tod6*Δ double-mutant strain. We accomplished this by replacing the two separate proteins with a new protein node, named Dot6/Tod6. In total, the background network consisted of 5,130 nodes representing proteins or protein complexes and 6,481 nodes representing DNA or RNA. There were 29,936 unique, interacting pairs of proteins (27% having a known direction) and 260,365 unique pairs of one protein and one nucleic acid sequence. Information about the provenance of the background network is provided in Appendix Table A2.

A1.2.4 Candidate paths

An inferred subnetwork is composed of a union of directed paths through the background network, each path providing an explanation for pairs identified from experimental data or domain knowledge. As a set-up step before proceeding with inference, we must enumerate candidate directed paths for each type of salt-relevant pair by performing searches through the

background network (Figure 3B). When we incorporate an undirected edge in a path, we record the direction of the edge that allows the path to move forward from the source to the target. Experiments varying the lengths of these paths are discussed in Section S2.4.

Candidate source-target paths. To account for the effect of a source mutant on its dysregulated targets, we must provide directed candidate paths that begin with the source and end with an interaction that represents the binding of a TF (or RBP) to the target gene (or mRNA). In this section we describe our process for enumerating these paths. For the five sources that are themselves transcription factors (Dot6/Tod6, Msn2, Rim101, Swc3, and Swc5), we assumed that they bind their targets directly. For these sources, each source-target path consisted of only one interaction. We enumerated longer paths to account for the sources that are not transcription factors. First, we identified which TFs and RBPs present in the background network could plausibly account for the effect of each source on its targets. We considered a relationship between a source and a TF (or RBP) if (a) the binding data for the TF/RBP was gathered under salt stress conditions (Ni *et al.* 2009) or (b) the set of genes bound by the TF/RBP were significantly overrepresented in the source's targets. Overrepresentation is decided by a p-value < 0.05 given by a hypergeometric test of the overlap between the two sets. Because we were only assembling a candidate list of TFs for each source, we did not correct for multiple testing at this step. After identifying candidate TFs and RBPs for each source, we enumerated possible directed paths between all source-target pairs through the background network, terminating with an interaction between one of the candidate TFs/RBPs and the target. Because the experimental data represents transcriptional changes that happen on a short time scale, the interactions in the path (except for the last) were limited to either interactions between proteins or between an RBP and the protein node for an mRNA that it binds. In other words, we did not allow connections from target genes/mRNAs to proteins,

which would require both transcription and subsequent translation of the protein encoded by the target. We applied a limit on the number of interactions in each path for the sake of tractability. For each source that was not a TF/RBP, we first enumerated all paths of up to three interactions. If by doing so we were unable to reach at least 50% of the source's candidate TFs/RBPs, we searched out until that goal percentage was reached, allowing up to five interactions total. By this process we were able to reach on average 78% of each source's targets and 75% of each source's candidate TFs/RBPs. Appendix Table A5 shows the coverage of each source's targets and candidate TFs/RBPs by the candidate paths.

Candidate fitness-contribution hit-source and source-source paths. In addition to explaining the effects of the sources on their targets, we were also interested in identifying other connections among salt-responsive proteins that may not lie along source-target paths: among fitness- contribution hits and sources, and among the sources themselves. We enumerated short candidate paths (up to two interactions) between each fitness-contribution hit and each source, and between each pair of sources. None of the interactions in these paths could represent the binding of a TF to a gene. For each pair, we searched for paths that proceed in both directions.

Candidate paths to explain indirect relationships from domain knowledge (receptor-source paths). In this step, we enumerated directed paths of up to five interactions that could connect the two upstream receptors, Sho1 and Sln1, to Hog1. As we did for the hit-source and source-source paths, we only allowed protein-protein and RBP-protein interactions in these paths.

A2 Computational analysis

We used the ensemble of subnetworks inferred by the IP method to make predictions about different aspects of the salt-specific subnetwork. We inspected the ensemble from a different view for each type of prediction. By applying a threshold on confidence values represented by the ensemble, we defined three high-confidence, or consensus, views:

- **Consensus nodes**, defined as the set of nodes with at least 75% confidence; they were used to make predictions about the identity of key players in the salt response
- **Consensus edges**, defined as the set of edges with at least 75% confidence; they were used to make predictions about direct interactions between consensus nodes
- **Consensus paths**, defined as the set of paths with at least 75% confidence; they were used to make predictions about indirect relationships between proteins and genes/mRNAs (such as, which proteins mediate the signal between a given source and its targets)

In this section we describe the methodology used for each computational result presented in the manuscript, as well as a supplementary analysis of the contributions of multiple components of the IP method. A list of the 15 highest-degree consensus nodes is provided in Appendix Table A6.

A2.1 Enrichment of consensus subnetwork with known relevant proteins

Because our curated list of true positives is known to be incomplete, we additionally evaluated our consensus protein node set (excluding target genes/mRNAs) against several external sources of likely-relevant genes and interactions. We tested the enrichment of consensus proteins with two sets of genes that are related to cell signaling and stress: (a) kinases and phosphatases, and (b) proteins retrieved from SGD's YeastMine (BALAKRISHNAN *et*

al. 2012; CHERRY *et al.* 2012) with the queries 'stress regulation' or 'nutrient regulation'. We also tested for enrichment with essential genes from SGD (CHERRY *et al.* 2012), and for genetic interactions from BioGRID (STARK *et al.* 2006). Finally, we also tested the consensus subnetwork for overrepresentation of the true positives and underrepresentation of the likely negatives described in Methods and Figure 4B. Using the hypergeometric test, enrichment was tested both relative to the candidate network (the network given by the set of candidate paths) and to the background network. To separate out the effect of experimental hits on the enrichment score, we excluded them from the consensus nodes and the external gene sets. After filtering, 160 consensus nodes, 736 candidate nodes, and 4703 background network nodes remained. With experimental hits removed, the external gene sets consisted of 147 kinases and phosphatases, 237 general stress proteins, 910 essential genes, 70 true positives, and 1416 likely negatives. Additionally, we tested for enrichment of genetic interactions based on the assumption that genetic interactions are more likely to occur between functionally related genes. For this test, we did not omit experimental hits. We extracted 141,507 genetic interactions reported in BioGRID (STARK *et al.* 2006) (downloaded February 2013) and assumed a total of 16 million possible interactions among 5,700 yeast genes. Consensus subnetwork enrichment for genetic interactions was calculated using the number of pairs of consensus nodes that have a reported genetic interaction. We also tested the consensus subnetwork against the permuted subnetworks described in Methods and Figure 4B. For each ensemble inferred from permuted data, we defined a consensus node set using a 75% confidence threshold, and, for each external gene/interaction set, measured the proportion of each consensus node set that was contained in the external set. For each external gene set, we calculated the *p*-value as the fraction of the 1,000 permuted node sets that had an equal or higher proportion than the 'true' consensus node set. (For the likely negatives, we counted the number of permuted consensus node sets with equal or lower representation.) We did not filter out hits for this experiment, as the hit sets were different for each permutation. In Appendix

Table A7, we show the proportion of the consensus subnetwork, candidate network, background network, and permuted subnetworks that is represented in each external gene set, and the *p*-values for comparisons to the consensus subnetwork. These results are summarized in the main manuscript.

A2.2 Stability analysis

As it is practically infeasible to identify all optimal solutions to the IP, we specify the number of sub- networks in the ensemble as an input to our method. The question arises of how the ensemble's predictions change as more solutions are identified. In this analysis, we compared the predictions made by the ensemble of 10,000 solutions to two smaller ensembles of 100 and 1,000 solutions each. As shown in Appendix Figure A10A, the precision-recall curves for each ensemble were approximately the same, although the recall of low-confidence predictions increased slightly as more solutions were gathered. The representation of general stress proteins, kinases and phosphatases, and genetic interactions was also approximately the same across the different ensemble sizes. Additionally, we compared the different sizes of ensembles based on the Jaccard similarity between their predictions. We calculate the similarity between two ensembles E (10,000 solutions) and E' (100 or 1,000 solutions) as follows. Using the variable y_n (node relevance) as an example, $p^E(y_n = 1)$ is E's confidence that node n is relevant. $\mathcal{N}$ refers to the set of all nodes (proteins).

$$similarity(E, E') = \frac{\sum_{n \in \mathcal{N}} |p^E(y_n = 1) - p^{E'}(y_n = 1)|}{\sum_{n \in \mathcal{N}} p^E(y_n = 1) + p^{E'}(y_n = 1) - |p^E(y_n = 1) - p^{E'}(y_n = 1)|} \quad (1)$$

We compared the 10,000-solution ensemble to the two smaller ensembles based on four variable types: node relevance, edge relevance, edge direction, and path relevance. Sources, targets, and upstream receptors were omitted from the node stability calculations because their relevance variables were fixed. When calculating the similarity of edge directions, only edges that were undirected in the background network were counted. For all comparisons, the similarity between the 10,000-solution ensemble and the smaller ensemble was very high (≥ 90%). In Appendix Figure A10 B, we provide a visualization of the results of each comparison, showing the number of elements that were shared or unique to each ensemble. All together, our stability analysis results demonstrate that the distribution of solutions captured by CPLEX is not changed significantly by the collection of more solutions.

A2.3 Lesion testing

Our approach combines multiple sources of experimental data using a multi-step optimization pro- cedure. It is useful to investigate the contribution of each ingredient to the predictions made by the inferred subnetwork ensemble. To do so, we ran lesioned versions of the IP, in each of which one component of the optimization procedure or one experimental data source was held aside, and inferred an ensemble of solutions, which we refer to as a *lesioned ensemble*. We evaluated the lesioned ensembles using the same computational evaluation metrics that we applied to the ensemble inferred by the complete IP (the *complete ensemble*): namely, accuracy in predicting held-aside test cases, and enrichment of relevant, externally-derived gene and interaction sets. Experimental hits were omitted from the test cases and the relevant, external gene sets. As an additional evaluation, we compared the lesioned ensembles to the complete ensemble based on the similarity of their predictions, as in the stability analysis presented in Section 2.2. Having observed that predictions were highly stable regardless of the size of the ensemble, we performed the lesion experiments using an ensemble size of 1,000. Correspondingly, we compared the lesioned ensembles to the complete ensemble

with 1,000 solutions. Appendix Figure A11A-F show comparative precision-recall curves for each lesioned ensemble. In Appendix Table A8, we show the results of the enrichment analyses, including the proportion of each lesioned consensus subnetwork that is contained in each external gene and interaction set. For each relevant set, we compared its representation in the lesioned consensus to the complete consensus using a two-tailed *z*-test of proportions. We discuss the contribution of each lesioned component:

Maximizing the number of connections between hits and sources (Appendix Figure A11A). Removing this step alone slightly increased the predictive accuracy of the inferred ensemble, but did not significantly change the proportions of the consensus nodes that were represented by any relevant gene sets. We suggest three possible explanations for the change in precision:

(a) the experimental hit set is noisy or includes proteins that are required for stress survival but not for signaling *per se*, (b) the set of test cases is relatively small and based on focused labora- tory experimentation, and is thus likely to be incomplete, and (c) removing this objective function component left the resulting IP less constrained, allowing greater variation between the different solutions in the ensemble and more stratification between high-confidence predictions. There was also a small but statistically significant change in the proportion of genetic interactions. Despite the small change in precision that it confers, including this step assists in the interpretation of the experimental data, as it reveals proximal connections between fitness-contribution hits and sources that are not captured by source-target paths.

Maximizing the number of experimental hits (Appendix Figure A11B). Even though the hit nodes themselves were not considered in the accuracy analysis, maximizing their inclusion results in more accurate choices among the other nodes: the lesioned IP had a slightly lower PR curve compared to the complete IP. The representation of relevant gene sets did not change significantly; however, the lesioned consensus had a higher proportion of genetic interactions than the complete consensus node set.

The combined contribution of maximizing connections and hits (Appendix Figure A11C). These two objective components are partially redundant, as maximizing the number of connections indirectly also maximizes the inclusion of hits. Therefore, we assessed the effect of removing both components. This resulted in a slight increase in precision. Additionally, while the lesioned consensus node set did not differ in the representation of relevant gene sets, it did contain a significantly higher proportion of genetically interacting pairs of nodes. The next lesion we performed attempted to tease apart the contributions of the two experimental hit sets.

Maximizing the inclusion of either of the two hit sets (Appendix Figure A11D). To assess the hit sets separately, we performed two lesions, separately holding aside the fitness-contribution hits and the phospho-proteomic hits. The inclusion of phospho-proteomic hits in the IP appears to be useful for identifying higher precision subnetworks, demonstrating that some of the information contained in this hit set is congruent with current knowledge about the salt stress response. Conversely, holding aside the entire set of fitness-contribution hits (modulo the sources that are fitness-contribution hits) resulted in a moderate increase in precision, suggesting that this set of hits is noisy or partially disjoint to existing knowledge, and was responsible for the increased precision observed when both types of hits are held aside. Neither lesioned consensus node set shows a significant change in the representation

of relevant node sets, though the lesioning of phospho- proteomic hits results in a statistically significant increase in genetic interactions. It is notable that most genetic interactions have been measured under standard growth conditions; it is likely that NaCl-specific genetic interactions are not represented in the test set. Even though the predictive accuracy dropped slightly when fitness-contribution hits were maximized, our results may actually represent novel discoveries, as the set of true positives is heavily biased toward HOG pathways proteins. Regardless, we are willing to tolerate a slight drop in ac- curacy in exchange for increased interpretability. We believe that including the fitness-contribution hits in the IP is still useful; after all, part of the purpose of this method is to aid in the interpretation of the experimental hits. The resulting subnetwork reveals connections between the fitness-contributions and other components of the salt stress response, which may inspire further experimentation.

Minimizing the total number of nodes (Appendix Figure A11E). Removing this step resulted in decreased precision and enrichment across the board, demonstrating that minimizing nodes is useful for reducing false positives.

Connecting receptor-source pairs (Appendix Figure A11F). Including the receptor-source pairs appeared to significantly increase recall, but did not appear to have a great effect on precision or enrichment with relevant gene sets or genetic interactions. We believe that including these pairs makes the inferred subnetwork more interpretable by showing connections to well-understood interactions in the stress response.

Similar to our evaluation of the stability of the complete IP across different ensemble sizes,

we measured the similarity between the complete and lesioned ensembles. A visualization of the differences and intersections between the ensembles' predictions is shown in Appendix Figure A12. Except when nodes were not minimized, the lesioned ensembles contained fewer nodes, edges, and paths, and were mostly subsets of the complete ensemble. (In the figure, we observe that the size of the left-hand portion of each bar generally dominates the right-hand portion.) When nodes were not minimized, the reverse is observed. Path relevance and edge direction showed the most variablility, demonstrated by the longer right-hand ends to those bars.

A2.4 Testing variations on candidate path length

To supplement the lesion tests, we measured the effect of increasing and decreasing the maximum length of the candidate source-target and hit-source paths by one interaction. For each variation, we inferred an ensemble of 1,000 subnetworks, and compared the results to the original consensus ensemble on the basis of precision-recall and stability analyses (Appendix Figure A13), as well as enrichment analysis on the nodes with ≥ 75% confidence (Appendix Table A9). Because the candidate path sets are different between the original, complete IP and the new variations, we do not measure the stability of paths for these comparisons, and only compare node and edge relevance and edge direction.

Source-target paths (Appendix Figure A13 A-B). In the original IP, we enumerated paths of up to five interactions, stopping search early at the depth at which at least 50% of candidate TFs/RBPs for a given source were reached. We tested the effect of this early-stopping option by inferring subnetworks using candidate paths enumerated at three different lengths: three, four, or five interactions for all sources. At a path length of five, we were unable to complete the

IP solution portion of the method due to the large number of paths generated by two sources, Hog1 and Mck1. As a compromise, we truncated the paths for those two sources to four, but searched for paths of length five for all other sources. In general, the precision-recall curves and most enrichment scores do not appear to be very sensitive to the path length in the range tested; however, we see some patterns. The ensemble inferred from candidate paths of length four appears to be nearly identical to the original IP under all measures, which is unsurprising, as search stopped at that length for most sources in the original setting. Stopping all search at length three results in a slight increase in precision, a slight decrease in recall, and a significant increase in genetic interactions. Stopping all search at length five results in a slight decrease in precision and a significant decrease in genetic interactions. As shown by the relative lengths of the ends of the bars in the stability analysis figure (Appendix Figure A13B), increasing candidate path length also results in increased size of inferred subnetworks. The node content of the inferred ensembles is fairly similar, with more variation shown in edge relevance and direction.

Hit-source and source-source paths (Appendix Figure A13 C-D). In the original IP, we sought candidate paths consisting of up to two edges between the fitness-contribution hits and the tested sources. To assess the sensitivity of the method to this path length, we tested two alternatives: allowing only direct interactions (no intermediate nodes), and allowing longer paths of up to three edges (two intermediate nodes). The shorter path length confers increased precision in the precision-recall curve; however, the 75%-confidence nodes from this subnetwork do not have statistically significantly different pro- portions of the gene sets tested for enrichment. There is a small but statistically significant in- crease in genetic interactions in the ensemble with shorter paths. This subnetwork includes 49 fitness-contribution hits, in contrast to the 106 included in the original consensus subnetwork, and is somewhat smaller. The longer path length results in a large decrease in precision, weakly significant decreases in

enrichment of true positives and genetic interactions, and a corresponding weakly significant increase in essential gene enrichment and likely negatives. This subnetwork includes 151 fitness- contribution hits and is much larger overall. Considering these results, it appears that limiting the length of candidate hit-source and source- source paths is beneficial: longer paths result in larger subnetworks with decreased precision and decreased enrichment with other relevant gene and interaction sets. The preferred stringency of this limit may depend on how the inferred subnetwork will be used to guide further experiments. The shortest length path provides an increase in accuracy at the expense of not being able to provide connections to as many fitness-contribution hits.

A2.5 Testing the effect of reordering the objective function components

Our four objective functions can be summarized as representing two concepts: explain a *maximal* amount of the experimental data using a *minimal* subnetwork. To assess the sensitivity of the subnetwork inference method to the ordering of these concepts, we performed an experiment in which we reversed the order. The altered sequence of objective functions proceeds as follows, and can be compared to the sequence shown in Appendix Figure A8:

1. Minimize nodes (one solution)
2. Maximize connections (one solution)
3. Maximize fitness-contribution and phospho-proteomic hits (one thousand solutions)
4. For each previous solution, maximize paths (one solution each)

Using this ordering of the objective functions, we inferred an ensemble of 1,000 subnetworks, and compared the results to the original consensus ensemble using precision-recall and

stability analysis (Appendix Figure A14), and enrichment analysis (Appendix Table A10). Compared to the original IP, the IP with reversed objective function components (which we name "min-then-max") achieves lower recall of the known regulators, no significantly different proportions of relevant gene sets, and a significant increase in genetic interaction enrichment. The resulting subnetworks are also overall smaller, as shown by the relative lengths of the colored portions of the bars in the stability analysis chart in Appendix Figure A14B.

A3 Analysis of newly implicated regulators

The array results from the validation analysis provide new insights into the newly identified NaCl-responsive regulators. Several mutants had defects in ESR regulation. Measured targets of the Pho85 kinase, the Arf3 RBP, and CK2 subunit Cka2 were enriched for genes induced in the iESR. Cells lacking the cyclin-dependent Pho85 kinase also had a defect in repressing RP genes but displayed amplified induction of genes involved in respiration. Pho85 is thought of as a nutrient and cell cycle regulator but was also previously implicated in suppressing the ESR(MOFFAT *et al.* 2000; CARROLL *et al.* 2001; CARROLL and O'SHEA 2002). The *pho85*Δ mutant indeed had an apparent derepression of the ESR before stress (however this is hard to distinguish from ESR activation due to inherent stress in the sickly mutant cells). Cells lacking the Yak1 kinase had a defect in repressing RP genes in response to NaCl, likely due to the known Yak1-dependent regulation of RP regulators Ifh1 and/or Crf1(MARTIN *et al.* 2004; KIM *et al.* 2011). Surprisingly, cells lacking Bck1 — a key MAPKKK in the cell wall integrity pathway — also displayed a defect in repressing RP genes upon NaCl treatment. The subnetwork inference predicts that Bck1 lies upstream of Yak1, based on the observation that Yak1 is an *in vitro* substrate of Bck1(PTACEK *et al.* 2005). As several studies have uncovered genetic interactions between these kinases (HERMANSYAH *et al.* 2010; CARDONA *et al.* 2012; LAVIÑA *et al.* 2014), our results suggest that Bck1 is a previously unrealized regulator of Yak1 in the

NaCl response. Another example is the CK2 kinase complex, which regulates the iESR repressor Nrg1 and is known to affect stress-specific splicing and abundance of ribosome-related transcripts(BERKEY and CARLSON 2006; RUDRA *et al.* 2007; BERGKESSEL *et al.* 2011). We found that cells lacking either catalytic subunit (*CKA1* or *CKA2*) or both regulatory subunits (*CKB1/CKB2*) displayed a defective NaCl response, but that the defect differed for each mutant: *cka2*Δ cells displayed a defect in iESR gene induction, whereas the *cka1*Δ and *ckb1*Δ*ckb2*Δ mutants instead produced amplified repression of rESR genes. For both Cdc14 and CK2 subunits, the gene targets predicted by the subnetwork significantly overlapped the measured targets, and 70-80% of scorable nodes predicted in their paths were supported by our analysis (see Appendix Table A1). We also found defective NaCl-provoked expression in two mutants not previously linked to stress. Kin2 is a poorly characterized serine/threonine kinase in yeast that is linked to exocytosis (ELBERT *et al.* 2005; GASSER *et al.* 2007). Genes that displayed defective expression either basally or upon NaCl treatment were enriched for nutrient transporters, nutrient responsive regulators, and genes whose deletion produces morphology defects. Genes with amplified NaCl-responsive induction in the *kin2*Δ mutant were also enriched for targets of the meiosis regulator and transcriptional repressor, Sum1. These results are consistent with nutrient response and effects at the cellular periphery, where Kin2 is localized (TIBBETTS *et al.* 1994). Little is known about the kinase Nnk1, except from large-scale studies that show an interaction with Tor1 (BREITKREUTZ *et al.* 2010). Most of the affected genes displayed amplified expression changes compared to wild type cells. There was little functional enrichment among the genes, aside of mobile-intron containing mitochondrial genes and genes regulated by TATA-dependent transcription.

A4 Cdc14 is a central regulator in the NaCl response

Cdc14 phosphatase is a key regulator of mitotic exit and acts in part as a Cdc28/Cdk1 antagonist, by reversing phosphorylation of Cdc28 targets (VISINTIN *et al.* 1998). Cdc14 is

sequestered in the nucleolus during most of the cell cycle but is released and distributed throughout the cell at anaphase (VISINTIN *et al.* 1998; DE ALMEIDA *et al.* 1999). In addition to triggering mitotic exist, Cdc14 also helps to set up subsequent cell-cycle phases, for example by priming Swi6 for nuclear import in G1 (GEYMONAT *et al.* 2004) and licensing origins of replication for proper replication timing in S-phase (BLOOM and CROSS 2007; DULEV *et al.* 2009; ZHAI *et al.* 2010). Here we show that Cdc14 is critical for mounting a normal NaCl response. Cells harboring the temperature sensitive *cdc14-3* allele had a major defect in several aspects of the NaCl response that could not be explained by cell-cycle defects. For these experiments, the mutant and its isogenic wild type were grown to log phase at 25˚C then shifted to the nonpermissive 35˚C for 90 minutes, leading to near-complete M-phase arrest of *cdc14-3* cells. At this time cells were ex- posed to 0.7M NaCl and followed for up to 2 hours. Notably, wild-type cells had a near normal NaCl response across the transcriptome under these conditions, with slightly smaller fold-changes compared to cells growing constitutively at 30˚C. Thus the prior heat treatment had little effect on the wild-type response (as long as cells had 90 min to acclimate to the higher temperature, not shown). Mutant *cdc14-3* cells showed normal viability after heat shock (as scored by live-dead staining), but were more sensitive to short-term NaCl than wild-type cells (80% viability compared to 92% viability for wild-type). To ensure that the apparent defect in NaCl response was not obscured by unique *cdc14-3* gene expression differences due to heat pretreatment, we compared expression in heat-treated *cdc14-3* and wild-type cells after the 90 min heat treatment but before addition of salt. Heat inactivation of *cdc14-3* caused only subtle differences in gene expression for most affected genes (Appendix Figure A6A). The exception was cell cycle-regulated genes, whose expression reflected the M- phase arrest of the mutant (which produces reduced abundance of transcripts that peak in other cell-cycle phases, Figure 6F). The *cdc14-3* mutant had a slight induction of many stress-induced genes after heat inactivation but before NaCl addition. However, this cannot explain the defect

in NaCl response of the mutant. First, many genes showed defective NaCl-responsive induction in the *cdc14-3* mutant despite no difference in pre-NaCl expression levels. This was also true for most NaCl-repressed genes, which showed little significant difference before NaCl addition in the *cdc14-3* mutant. Second, there were many genes for which pre-NaCl levels were slightly elevated in the inactivated *cdc14-3* mutant yet the fold-change in expression upon NaCl treatment happened at wild-type levels. Thus, expression differences in the *cdc14-3* mutant cannot be simply explained by highly aberrant transcript differences before NaCl addition. In further support of a specific role for Cdc14 in the NaCl response, we found *cdc14-3* mutant cells growing at the non-permissive temperature had a partial defect in Hog1 localization (see main text). This was not due to M-phase arrest, since wild-type cells in G2/M phase showed robust Hog1 nuclear localization (Appendix Figure A6B); furthermore, we found no defect in nuclear Hog1 accumulation in wild type cells arrested with nocodozole (not shown). These results strongly suggest that the role of Cdc14 in the NaCl response is due to specific activation of the phosphatase, rather than an indirect effect of cell-cycle arrest. The NaCl-enhanced interaction between Cdc14 and Cka2 supports this conclusion. Because the mitotic function of Cdc14 is regulated via localization changes, we looked for nucleolar release of Cdc14 upon salt treatment. There was no gross relocalization of Cdc14-GFP after NaCl treatment (Appendix Figure A6C); however, we cannot exclude a partial release of the phosphatase. Interestingly, a prior study found that nucleolar release of Cdc14 at anaphase is in part triggered by the Hog1 MAPKK Pbs2 (REISER *et al.* 2006); this observation bolsters our subnetwork's prediction that Pbs2 regulates Cdc14 activity upon NaCl treatment as well. Notably, we detected no interaction between immunoprecipitated Hog1 and Cdc14, either with or without prior cellular NaCl treatment (not shown). Cdc14 is a known CTD phosphatase in yeast (CLEMENTE-BLANCO *et al.* 2011) and humans (GUILLAMOT *et al.* 2011), prompting us to investigate whether it was required for the immediate drop in Ser2 and Ser5 phosphorylation upon NaCl treatment (see Figures 5E and Appendix Figure A5). However, we found no obvious defect in the *cdc14-3*

mutant responding to salt, relative to Pol II subunit Rpb3 (Appendix Figure A6D). We noticed in several replicates that Rbp3 (normalized to a second internal control, Pab1) showed decreasing levels over time in the *cdc14-3* strain responding to NaCl at the non-permissive temperature, but not at 25˚C (not shown). However, this was not always the case, raising question to its significance. It is intriguing, however, that the subnetwork connects Cdc14 to Rpb3 through the ubiquitin ligase Rsp5, which is known to control stability of polymerase subunits in response to other stresses (BEAUDENON *et al.* 1999). Future work will be required to dissect this interplay.

References

ABDULREHMAN, D., P. T. MONTEIRO, M. C. TEIXEIRA, N. P. MIRA, A. B. LOURENÇO *et al.*, 2011 YEASTRACT: providing a programmatic access to curated transcriptional regulatory associations in Saccharomyces cerevisiae through a web services interface. Nucleic Acids Res **39:** D136-140.

ALEJANDRO-OSORIO, A. L., D. J. HUEBERT, D. T. PORCARO, M. E. SONNTAG, S. NILLASITHANUKROH *et al.*, 2009 The histone deacetylase Rpd3p is required for transient changes in genomic expression in response to stress. Genome Biol **10:** R57.

BALAKRISHNAN, R., J. PARK, K. KARRA, B. C. HITZ, G. BINKLEY *et al.*, 2012 YeastMine--an integrated data warehouse for Saccharomyces cerevisiae data as a multipurpose tool-kit. Database (Oxford) **2012:** bar062.

BEAUDENON, S. L., M. R. HUACANI, G. WANG, D. P. MCDONNELL and J. M. HUIBREGTSE, 1999 Rsp5 ubiquitin-protein ligase mediates DNA damage-induced degradation of the large subunit of RNA polymerase II in Saccharomyces cerevisiae. Mol Cell Biol **19:** 6972-6979.

BERGKESSEL, M., G. B. WHITWORTH and C. GUTHRIE, 2011 Diverse environmental stresses elicit distinct responses at the level of pre-mRNA processing in yeast. RNA **17:** 1461-1478.

BERKEY, C. D., and M. CARLSON, 2006 A specific catalytic subunit isoform of protein kinase CK2 is required for phosphorylation of the repressor Nrg1 in Saccharomyces cerevisiae. Curr Genet **50:** 1-10.

BERRY, D. B., Q. GUAN, J. HOSE, S. HAROON, M. GEBBIA *et al.*, 2011 Multiple means to the same end: the genetic basis of acquired stress resistance in yeast. PLoS Genet **7:** e1002353.

BLOOM, J., and F. R. CROSS, 2007 Novel role for Cdc14 sequestration: Cdc14 dephosphorylates factors that promote DNA replication. Mol Cell Biol **27:** 842-853.

BODENMILLER, B., L. N. MUELLER, M. MUELLER, B. DOMON and R. AEBERSOLD, 2007 Reproducible isolation of distinct, overlapping segments of the phosphoproteome. Nat Methods **4:** 231-237.

BREITKREUTZ, A., H. CHOI, J. R. SHAROM, L. BOUCHER, V. NEDUVA *et al.*, 2010 A global protein kinase and phosphatase interaction network in yeast. Science **328:** 1043-1046.

CARDONA, F., M. L. DEL OLMO and A. ARANDA, 2012 Phylogenetic origin and transcriptional regulation at the post-diauxic phase of SPI1, in Saccharomyces cerevisiae. Cell Mol Biol Lett **17:** 393-407.

CARROLL, A. S., A. C. BISHOP, J. L. DERISI, K. M. SHOKAT and E. K. O'SHEA, 2001 Chemical inhibition of the Pho85 cyclin-dependent kinase reveals a role in the environmental stress response. Proc Natl Acad Sci U S A **98:** 12578-12583.

CARROLL, A. S., and E. K. O'SHEA, 2002 Pho85 and signaling environmental conditions. Trends Biochem Sci **27:** 87-93.

CHERRY, J. M., E. L. HONG, C. AMUNDSEN, R. BALAKRISHNAN, G. BINKLEY *et al.*, 2012 Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res **40**.

CLEMENTE-BLANCO, A., N. SEN, M. MAYAN-SANTOS, M. P. SACRISTÁN, B. GRAHAM *et al.*, 2011 Cdc14 phosphatase promotes segregation of telomeres through repression of RNA polymerase II transcription. Nat Cell Biol **13:** 1450-1456.

DE ALMEIDA, A., I. RACCURT, S. PEYROL and M. CHARBONNEAU, 1999 The Saccharomyces cerevisiae Cdc14 phosphatase is implicated in the structural organization of the nucleolus. Biol Cell **91:** 649-663.

DULEV, S., C. DE RENTY, R. MEHTA, I. MINKOV, E. SCHWOB *et al.*, 2009 Essential global role of CDC14 in DNA synthesis revealed by chromosome underreplication unrecognized by checkpoints in cdc14 mutants. Proc Natl Acad Sci U S A **106:** 14466-14471.

ELBERT, M., G. ROSSI and P. BRENNWALD, 2005 The yeast par-1 homologs kin1 and kin2 show genetic and physical interactions with components of the exocytic machinery. Mol Biol Cell **16:** 532-549.

EVERETT, L., A. VO and S. HANNENHALLI, 2009 PTM-Switchboard--a database of posttranslational modifications of transcription factors, the mediating enzymes and target genes. Nucleic Acids Res **37:** D66-71.

FASOLO, J., A. SBONER, M. G. SUN, H. YU, R. CHEN *et al.*, 2011 Diverse protein kinase interactions identified by protein microarrays reveal novel connections between cellular processes. Genes Dev **25:** 767-778.

GASSER, B., M. SAUER, M. MAURER, G. STADLMAYR and D. MATTANOVICH, 2007 Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in yeasts. Appl Environ Microbiol **73:** 6499-6507.

GEYMONAT, M., A. SPANOS, G. P. WELLS, S. J. SMERDON and S. G. SEDGWICK, 2004 Clb6/Cdc28 and Cdc14 regulate phosphorylation status and cellular localization of Swi6. Mol Cell Biol **24:** 2277-2285.

GITTER, A., M. CARMI, N. BARKAI and Z. BAR-JOSEPH, 2013 Linking the signaling cascades and dynamic regulatory networks controlling stress responses. Genome Res **23:** 365-376.

GUELZIM, N., S. BOTTANI, P. BOURGINE and F. KÉPÈS, 2002 Topological and causal structure of the yeast transcriptional regulatory network. Nat Genet **31:** 60-63.

GUILLAMOT, M., E. MANCHADO, M. CHIESA, G. GÓMEZ-LÓPEZ, D. G. PISANO *et al.*, 2011 Cdc14b regulates mammalian RNA polymerase II and represses cell cycle transcription. Sci Rep **1:** 189.

HEAVNER, B. D., K. SMALLBONE, B. BARKER, P. MENDES and L. P. WALKER, 2012 Yeast 5 - an expanded reconstruction of the Saccharomyces cerevisiae metabolic network. MNC Syst Biol **6:** 55.

HERMANSYAH, W. A. LAVIÑA, M. SUGIYAMA, Y. KANEKO and S. HARASHIMA, 2010 Identification of protein kinase disruptions as suppressors of the calcium sensitivity of S. cerevisiae Deltaptp2 Deltamsg5 protein phosphatase double disruptant. Arch Microbiol **192:** 157-165.

HOGAN, D. J., D. P. RIORDAN, A. P. GERBER, D. HERSCHLAG and P. O. BROWN, 2008 Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. PLoS Biol **6:** e255.

HUEBERT, D. J., P. F. KUAN, S. KELEŞ and A. P. GASCH, 2012 Dynamic changes in nucleosome occupancy are not predictive of gene expression dynamics but are linked to transcription and chromatin regulators. Mol Cell Biol **32:** 1645-1653.

KIM, E. M., J. KIM, Y. G. KIM, P. LEE, D. S. SHIN *et al.*, 2011 Development of high-throughput phosphorylation profiling method for identification of Ser/Thr kinase specificity. J Pept Sci **17:** 392-397.

LAVIÑA, W. A., H. SHAHSAVARANI, A. SAIDI, M. SUGIYAMA, Y. KANEKO *et al.*, 2014 Suppression mechanism of the calcium sensitivity in Saccharomyces cerevisiae ptp2Δmsg5Δ double disruptant involves a novel HOG-independent function of Ssk2, transcription factor Msn2 and the protein kinase A component Bcy1. J Biosci Bioeng **117:** 135-141.

MACISAAC, K. D., T. WANG, D. B. GORDON, D. K. GIFFORD, G. D. STORMO *et al.*, 2006 An improved map of conserved regulatory sites for Saccharomyces cerevisiae. BMC Bioinformatics **7:** 113.

MARTIN, D. E., A. SOULARD and M. N. HALL, 2004 TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1. Cell **119:** 969-979.

MOFFAT, J., D. HUANG and B. ANDREWS, 2000 Functions of Pho85 cyclin-dependent kinases in budding yeast. Prog Cell Cycle Res **4:** 97-106.

NI, L., C. BRUCE, C. HART, J. LEIGH-BELL, D. GELPERIN *et al.*, 2009 Dynamic and complex transcription factor binding during an inducible response in yeast. Genes Dev **23:** 1351-1363.

PTACEK, J., G. DEVGAN, G. MICHAUD, H. ZHU, X. ZHU *et al.*, 2005 Global analysis of protein phosphorylation in yeast. Nature **438:** 679-684.

PU, S., J. WONG, B. TURNER, E. CHO and S. J. WODAK, 2009 Up-to-date catalogues of yeast protein complexes. Nucleic Acids Res **37:** 825-831.

REISER, V., K. E. D'AQUINO, L. S. EE and A. AMON, 2006 The stress-activated mitogen-activated protein kinase signaling cascade promotes exit from mitosis. Mol Biol Cell **17:** 3136-3146.

RUDRA, D., J. MALLICK, Y. ZHAO and J. R. WARNER, 2007 Potential interface between ribosomal protein production and pre-rRNA processing. Mol Cell Biol **27:** 4815-4824.

SAITO, H., and K. TATEBAYASHI, 2004 Regulation of the osmoregulatory HOG MAPK cascade in yeast. J Biochem **136:** 267-272.

SCHERRER, T., N. MITTAL, S. C. JANGA and A. P. GERBER, 2010 A screen for RNA-binding proteins in yeast indicates dual functions for many enzymes. PLoS One **5:** e15499.

SHARIFPOOR, S., A. N. NGUYEN BA, J. Y. YOUN, D. VAN DYK, H. FRIESEN *et al.*, 2011 A quantitative literature-curated gold standard for kinase-substrate pairs. Genome Biol **12:** R39.

SOUFI, B., C. D. KELSTRUP, G. STOEHR, F. FRÖHLICH, T. C. WALTHER *et al.*, 2009 Global analysis of the yeast osmotic stress response by quantitative proteomics. Mol Biosyst **5:** 1337-1346.

STARK, C., B. J. BREITKREUTZ, T. REGULY, L. BOUCHER, A. BREITKREUTZ *et al.*, 2006 BioGRID: a general repository for interaction datasets. Nucleic Acids Res **34:** D535-539.

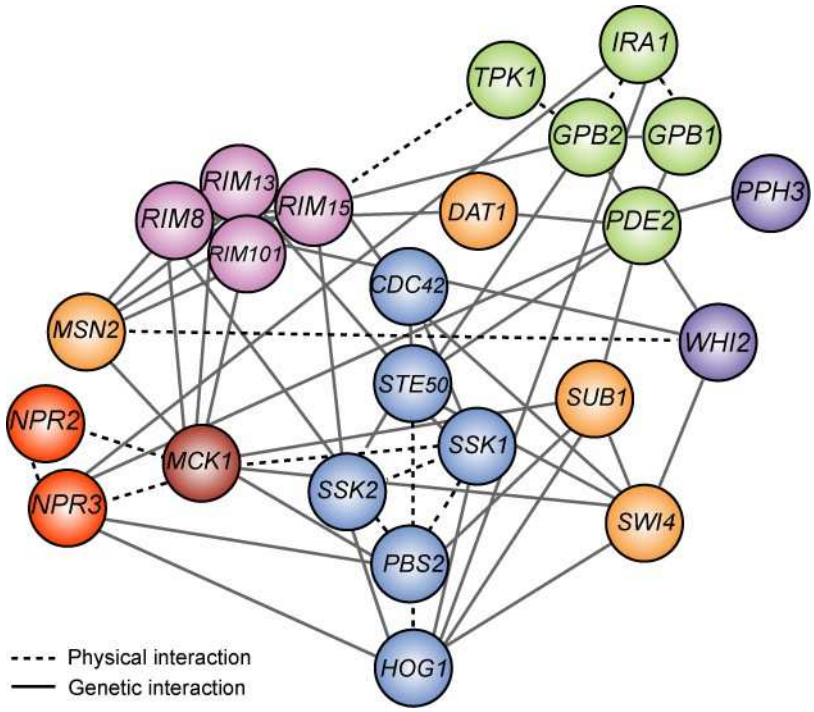
TIBBETTS, M., M. DONOVAN, S. ROE, A. M. STILTNER and C. I. HAMMOND, 1994 KIN1 and KIN2 protein kinases localize to the cytoplasmic face of the yeast plasma membrane. Exp Cell Res **213:** 93-99.

TSVETANOVA, N. G., D. M. KLASS, J. SALZMAN and P. O. BROWN, 2010 Proteome-wide search reveals unexpected RNA-binding proteins in Saccharomyces cerevisiae. PLoS One **5:** e12671.

VENTERS, B. J., S. WACHI, T. N. MAVRICH, B. E. ANDERSEN, P. JENA *et al.*, 2011 A comprehensive genomic binding map of gene and chromatin regulatory proteins in Saccharomyces. Mol Cell **41:** 480-492.

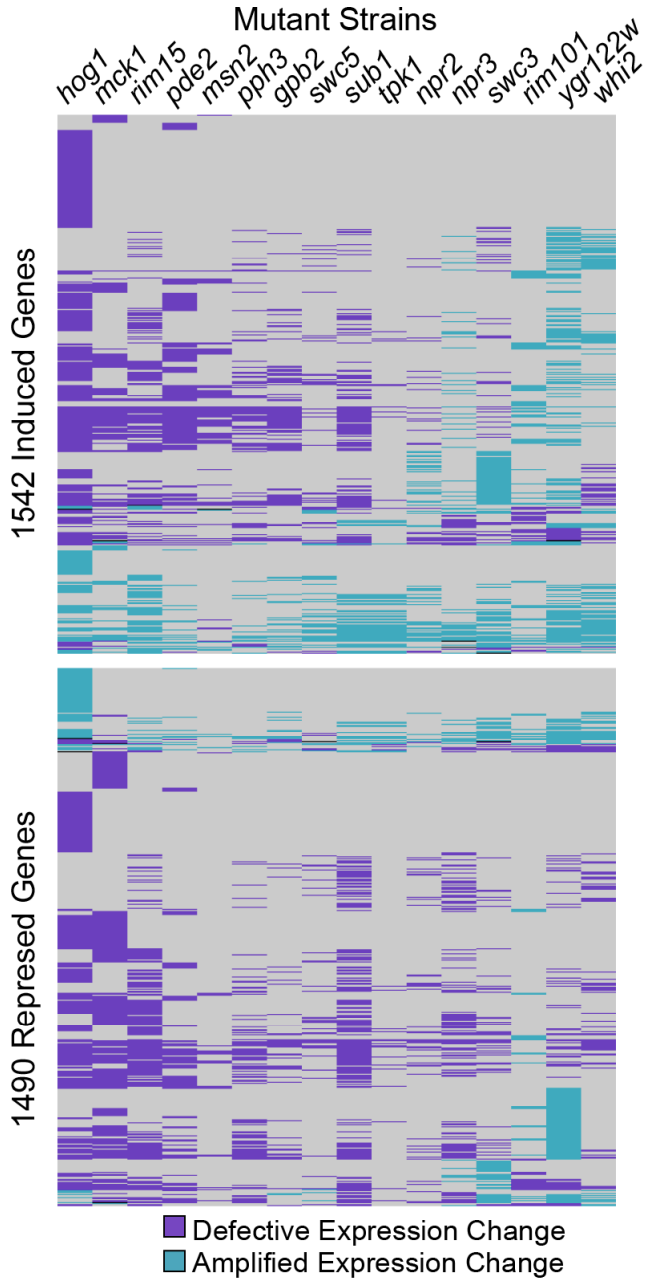
VISINTIN, R., K. CRAIG, E. S. HWANG, S. PRINZ, M. TYERS *et al.*, 1998 The phosphatase Cdc14 triggers mitotic exit by reversal of Cdk-dependent phosphorylation. Mol Cell **2:** 709-718.

ZHAI, Y., P. Y. YUNG, L. HUO and C. LIANG, 2010 Cdc14p resets the competency of replication licensing by dephosphorylating multiple initiation proteins during mitotic exit in budding yeast. J Cell Sci **123:** 3933-3943.



Appendix Figure A1. Signaling proteins required for acquisition of H_2O_2 resistance. Genes required for acquisition of severe H_2O_2 resistance after 60 min treatment with 0.7M NaCl were identified in (Berry *et al.*, 2011). Shown are known signaling proteins from the selection and their physical or genetic interactions reported in the BioGRID database (Stark *et al.*, 2006). Nodes are colored according to known players in the Hog (blue), Rim (pink), Tor (dark orange), PKA (green), and Mck1 (brick) pathways. Transcription factors are colored orange, and phosphatases are colored dark purple.

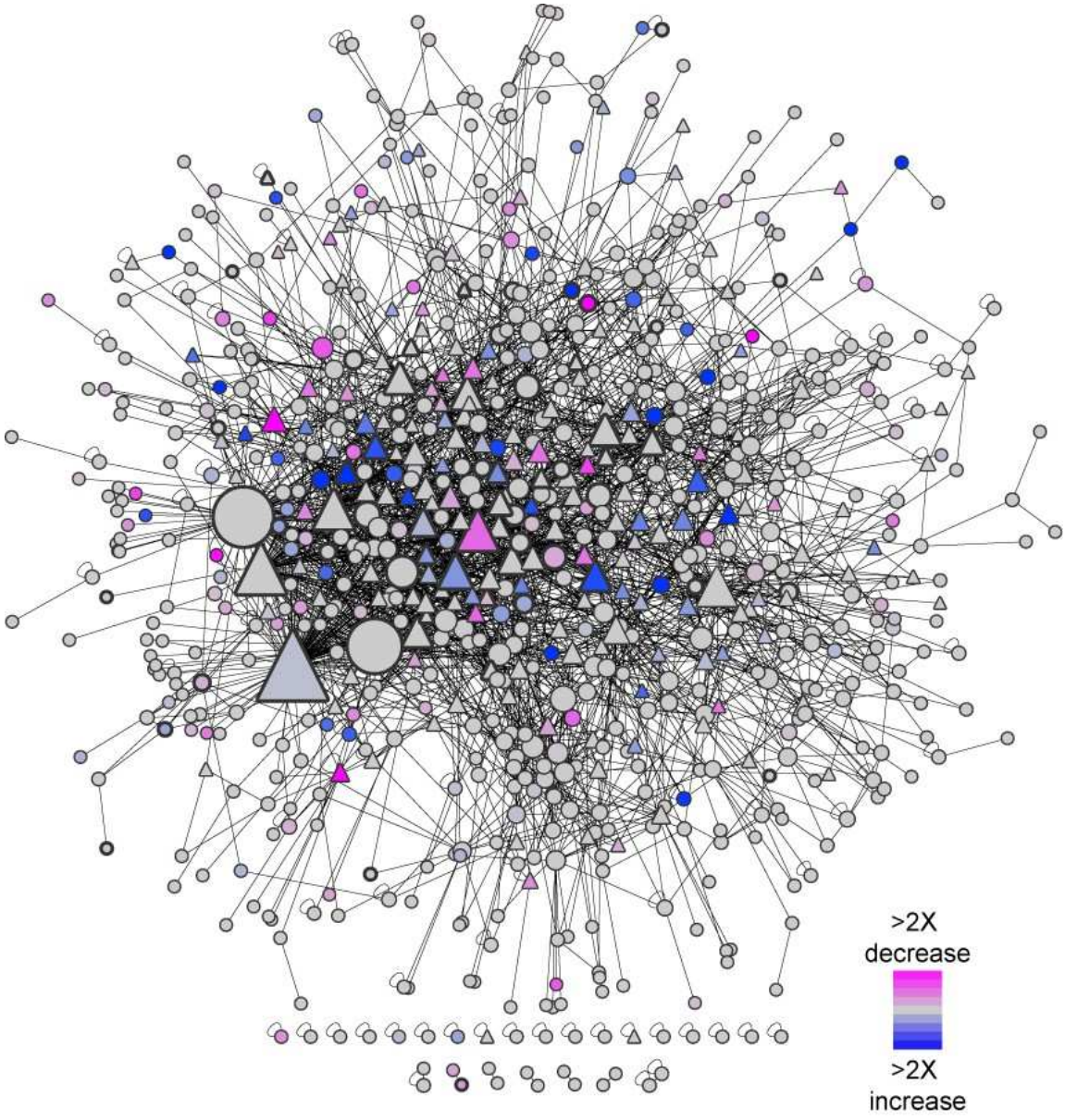
Appendix Figure A2. Graphical representation of affected genes. Genes induced (top) or repressed (bottom) in wild type cells responding to NaCl were identified (FDR < 0.05). Genes (rows) affected in each designated mutant (columns) were identified as described in Methods. Genes with a significant defect in expression compared to wild type (i.e. a smaller fold-change) are colored in purple, and genes with a significant amplification compared to wild type (i.e. larger fold-change) are colored in cyan. Of the 5056 genes with significant NaCl-responsive changes in wild-type cells (FDR < 0.05), 3,300 genes shown here displayed a significant defect in at least one mutant. A third of the affected genes were dependent on two or more regulators. There was also significant overlap in the targets of several regulators (see also Figure 1). For example, many genes regulated by the NaCl-activated Hog1 kinase were also affected by the phosphodiesterase Pde2, which inhibits PKA signaling, and the nutrient-responsive Rim15 and Mck1 kinases not previously linked to NaCl stress.

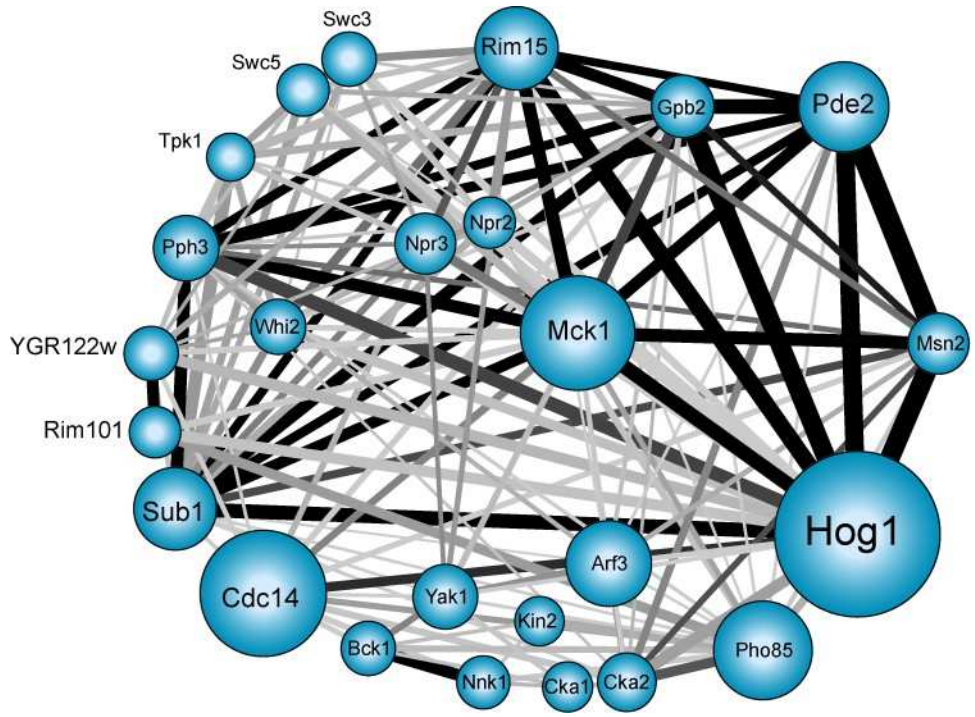

Mutant Strains
hog1
mck1
rim15
pde2
msn2
pph3
gpb2
swc5
sub1
tpk1
npr2
npr3
swc3
rim101
ygr122w
whi2
1542 Induced Genes
1490 Represed Genes
Defective Expression Change
Amplified Expression Change

Appendix Figure A3. Phospho-proteome network.

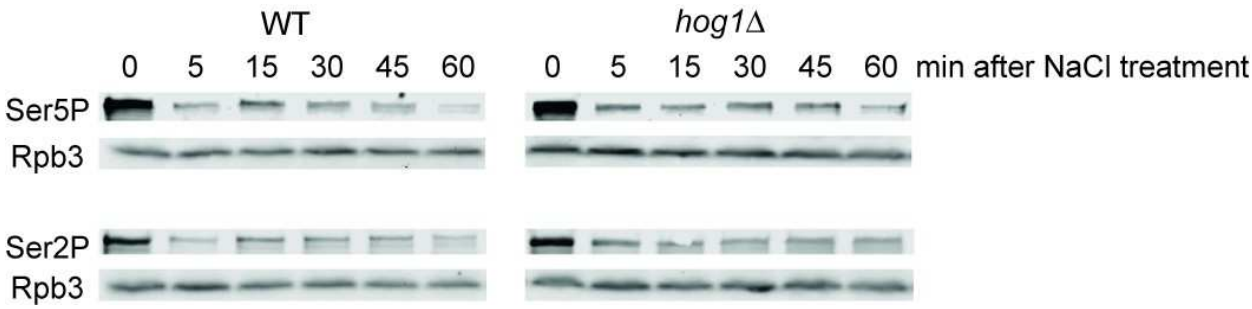
Changes in the phospho- proteome were identified by quantifying phospho-peptides before and at 5 and 15 min after NaCl treatment, as described in Methods. We identified 1937 phospho-peptides mapping to 973 different proteins. Roughly 600 sites displayed a two-fold change in expression at one or both time points, roughly split between proteins with increased and decreased phosphorylation. We compared our dataset to a prior phospho-proteome analysis that identified 169 proteins with a ≥ two-fold change in phosphorylation upon treatment with 0.4M NaCl for 5 min and 20 min (SOUFI *et al.* 2009). Not surprisingly given the sampling challenges of phosphoproteomics and the different experimental setups, the overlap between the datasets was low (only 12 changing phospho-proteins identified by both studies). This discrepancy could be caused from the difference of culturing conditions and stress dose (0.4M in that study versus 0.7 M used here), as well as the variable peptide sampling in different mass spectrophotometer runs, a known challenge in phosphoproteomics (*e.g.*, (BODENMILLER *et al.* 2007)). We used our phospho-proteomic data to build a network of the phospho-proteome, based on previously measured protein-protein interactions. The figure shows physical interactions between phospho-proteins, where triangles represent proteins included in the inferred consensus subnet- work, and circles indicate phospho-proteins that were not included in the consensus signaling sub-network. Node size is proportionate to degree in the phospho-proteome network, and node color is proportionate to the average fold-change in phosphorylation across both timepoints according to the color key. Nodes with more than one changing phospho-site are colored according to the site with the largest fold-change. Nodes with dark outlines indicate kinases. The network reveals high connectivity among phospho-proteins, with many interactions corresponding to protein complexes. Two hundred (21%) of the 973 phospho-proteins were included in the final signaling network, demonstrating that these networks are partially overlapping at best. Proteins participating in the phospho-proteomic network were enriched for specific functional categories,

including signal transduction, cell-cycle progression, cytokinesis/bud-site selection, and actin organization (Bonferroni-corrected $p < 0.05$), suggesting downstream effects on multiple physiological responses. Among the hubs unique to the protein network were palmitoylated plasma membrane-bound casein kinase 1 isoform Yck2 and a poorly characterized kinase, Ptk1, linked to spermine uptake. Proteins interacting with Yck2 were enriched for those involved in cell morphogenesis and linked to the cell periphery, whereas proteins linked to Ptk1 were enriched for kinases and proteins in the actin cytoskeleton. Other hubs in the phospho-protein network were Yck1, whose interacting proteins are involved in RNA transport, cell morphogenesis, and localized to the bud neck, and Atg1, a kinase regulating autophagy and vesicular trafficking and whose interacting partners were enriched for RBPs and proteins localized to the nucleolus.


>2X
decrease
>2X
increase



Appendix Figure A4. Overlap in targets of all interrogated regulators. As shown in Figure 1, but including regulators profiled through validation studies. Edge thickness represents the fraction of the smaller node's targets that overlap between two nodes. Edge color is proportional to significance of the overlap, where black represents a -log(*p*-value) of 5 or greater. The figure represents overlap only in targets with defective NaCl-dependent expression changes (i.e. smaller change in transcript abundance compared to wild type).



Appendix Figure A5. Defective *in vivo* CTD phosphorylation in *hog1*Δ cells responding to NaCl. This figure shows the blots that were quantified for Figure 5D and represents reproducible biological duplicate experiments.

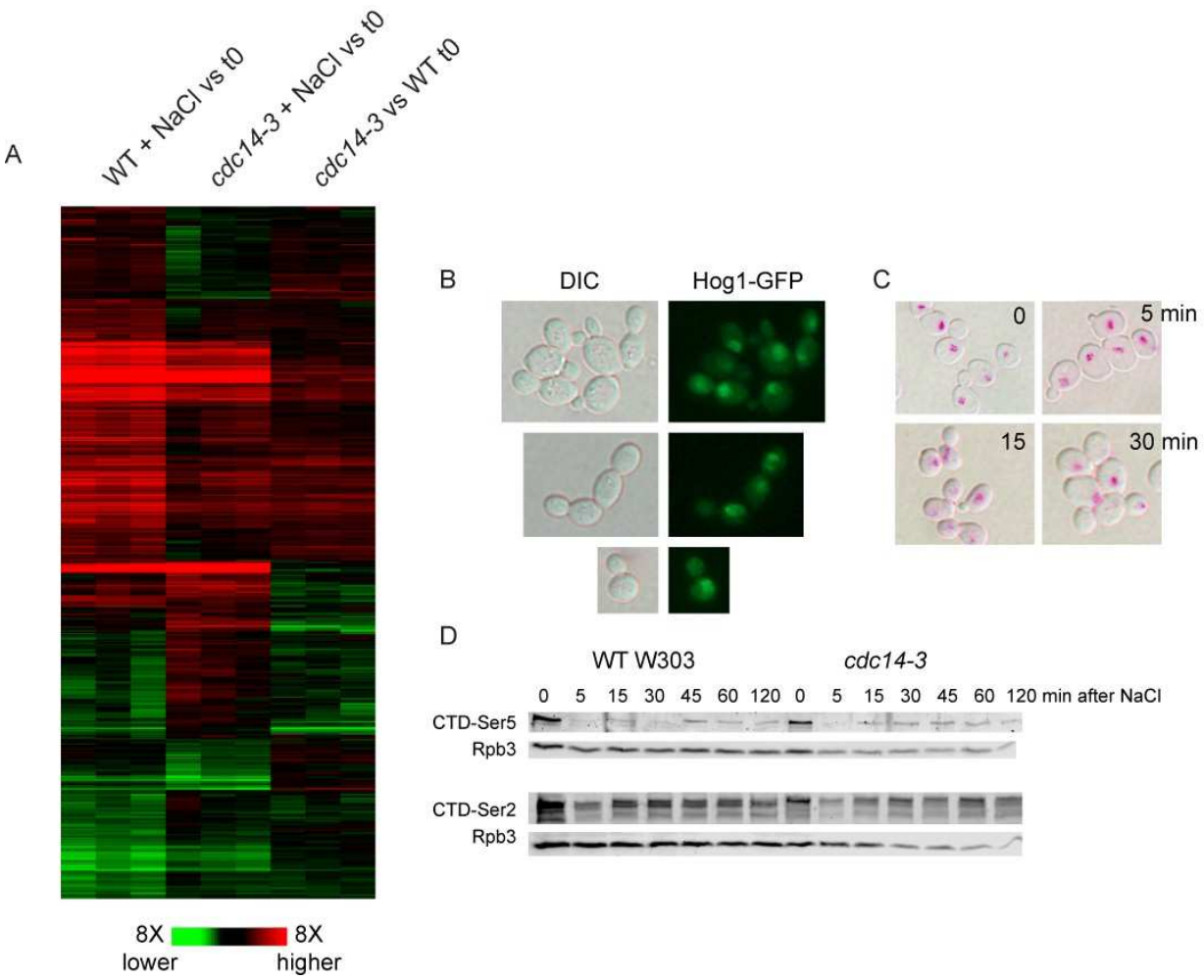
Appendix Figure A 6. Cdc14 is required for normal NaCl response.

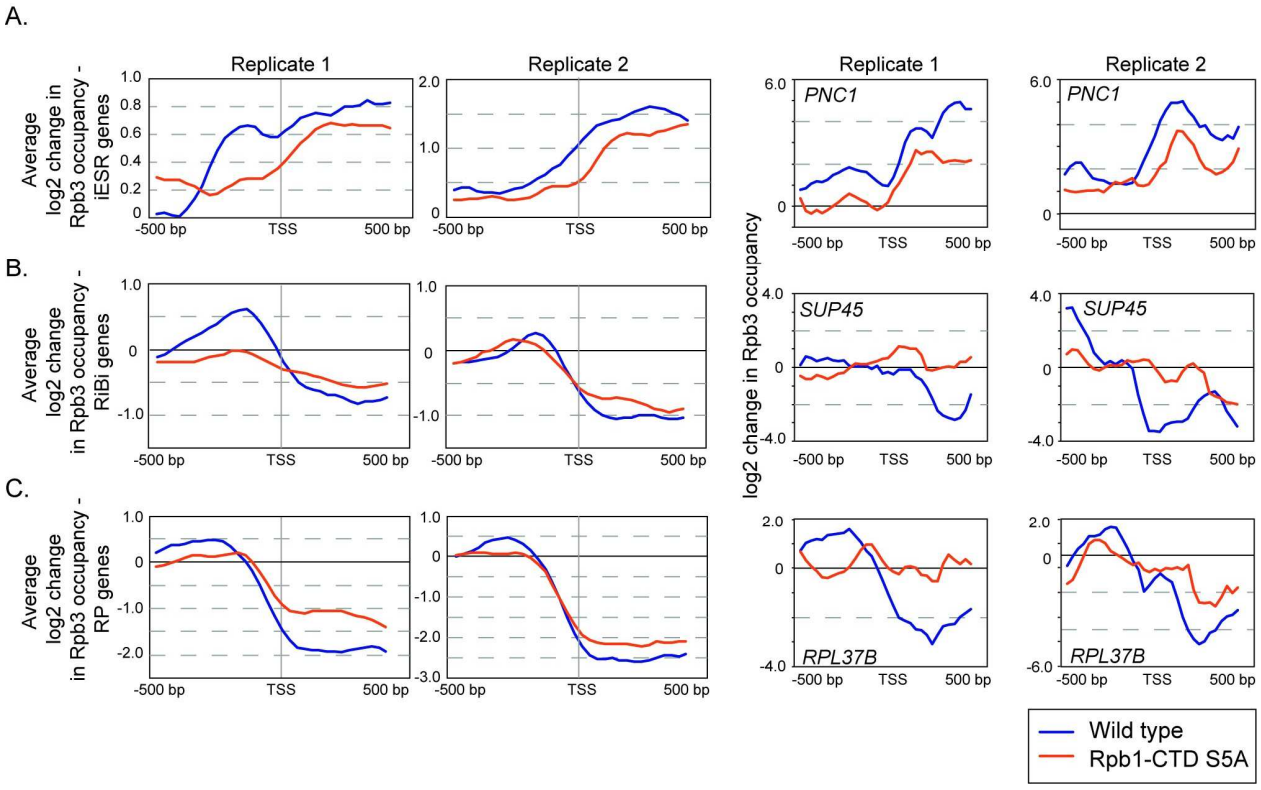
A. **A** Expression of 940 genes (excluding cell-cycle regulated genes (Spellman *et al.*, 1998)) scored as differentially expressed in the *cdc14-3* mutant responding to NaCl ($q < 0.05$). Each experiment was done in biological triplicate.

B. **B** Examples of dumbbell-shaped wild type W303 cells with nuclear Hog1-GFP at 10 min after NaCl treatment.

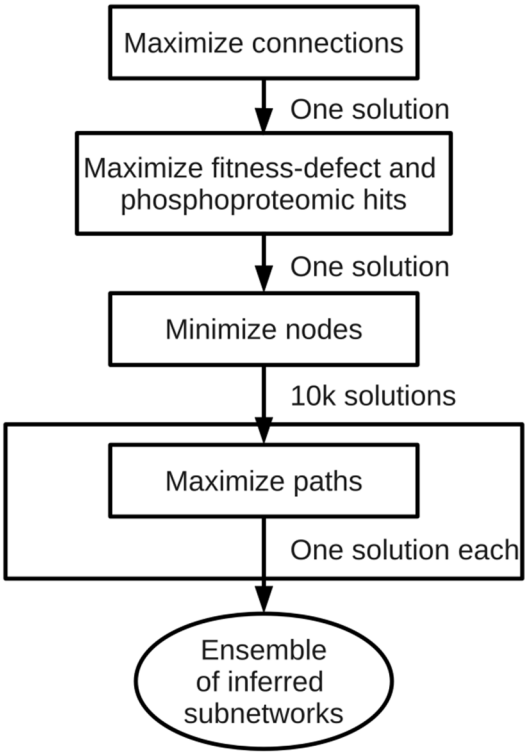
C. **C** Representative pseudo-colored overlays of Cdc14-GFP cells responding to 0.7M NaCl at denoted times before (0 min) or after treatment. GFP signal is pseudocolored as magenta.

D. **D** Western blots of bulk CTD Ser5-P and Rpb3 loading control (top) and Ser2-P and Rpb3 loading control (bottom) in wild-type and *cdc14-3* cells grown at 35˚C for 90 min ("0 min" NaCl treatment) and then exposed to 0.7M NaCl at 35˚C for 120 min after salt treatment.


A
WT + NaCl vs t0
cdc14-3 + NaCl vs t0
cdc14-3 vs WT t0
8X lower
8X higher
B
DIC
Hog1-GFP
C
0
5 min
15
30 min
D
WT W303
cdc14-3
0 5 15 30 45 60 120 0 5 15 30 45 60 120 min after NaCl
CTD-Ser5
Rpb3
CTD-Ser2
Rpb3

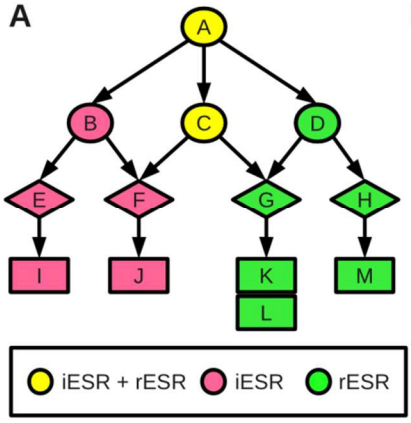


Appendix Figure A7. Change in Rpb3 occupancy at ESR-gene promoters. The log2(fold-change) in Rpb3 occupancy is shown from -500bp to +500bp around the transcription start site (TSS) of ESR genes, for A) iESR genes, B) RiBi genes, and C) RP genes. Wild-type profiles are shown in blue and profiles from the S5A mutant are shown in orange. Data from two biological replicates is shown, averaged across all genes in the respective groups (left two panels) and for a representative gene in each group (right two panels), including iESR gene *PNC1*, RiBi gene *SUP45*, and RP gene *RPL37B*.



Appendix Figure A8. Diagram of the procedure for optimization in the IP.

A



B

Target set	*C(T)*
T(A)	4.86
T(C)	2.76
T(B only)	0.0
T(B and C)	0.0
T(C and D)	0.0
T(D only)	0.0
T(F)	0.0
T(G)	0.0

C

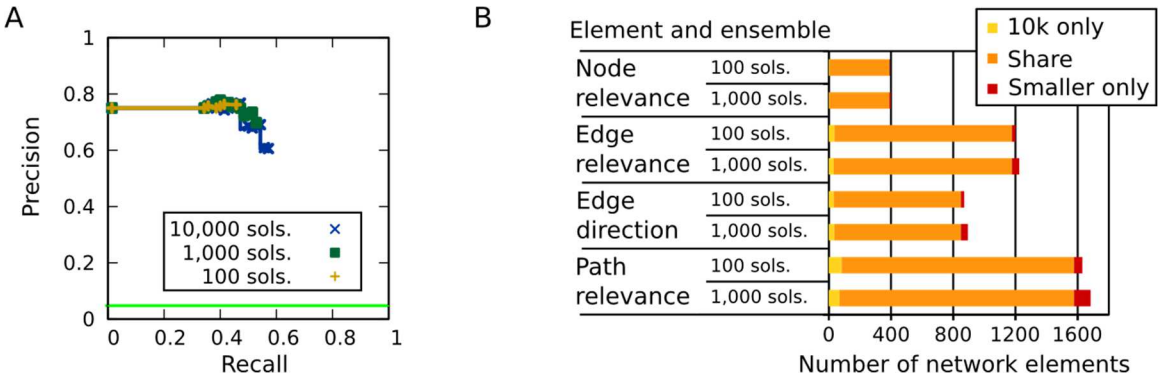
n	*I(n)*	*S(n)*	*B(n)*
A	4.86	1.92	2.53
C	2.76	0.92	3.00

Appendix Figure A9. Example of ESR bifurcation score calculation.

A. Example ESR network with nodes labeled according to which ESR targets are downstream. Nodes *a* and *c* are clearly candidate bifurcation points: they are upstream of both clusters, and we can see that the paths leading out from them divide the two clusters. However, choosing and ranking possible bifurcation points in the ESR consensus subnetwork is not easy to do by visual inspection.

B. Calculation of *C(T)*, count, for each candidate bifurcation point (*a* and *c*) and their outgoing partitions. At a given node, the outgoing partitions are defined by combinations of the node's children. Node *a*'s children define four outgoing partitions: ([*b* only], [*b* and *c*], [*c* and *d*], [*d* only]), and *c*'s children define two outgoing partitions ([*f*], [*g*]). This table shows all of the values of *C( )* that must be calculated for the example network: that of the two candidate bifurcation points, *a* and *c*, and each of their outgoing partitions.

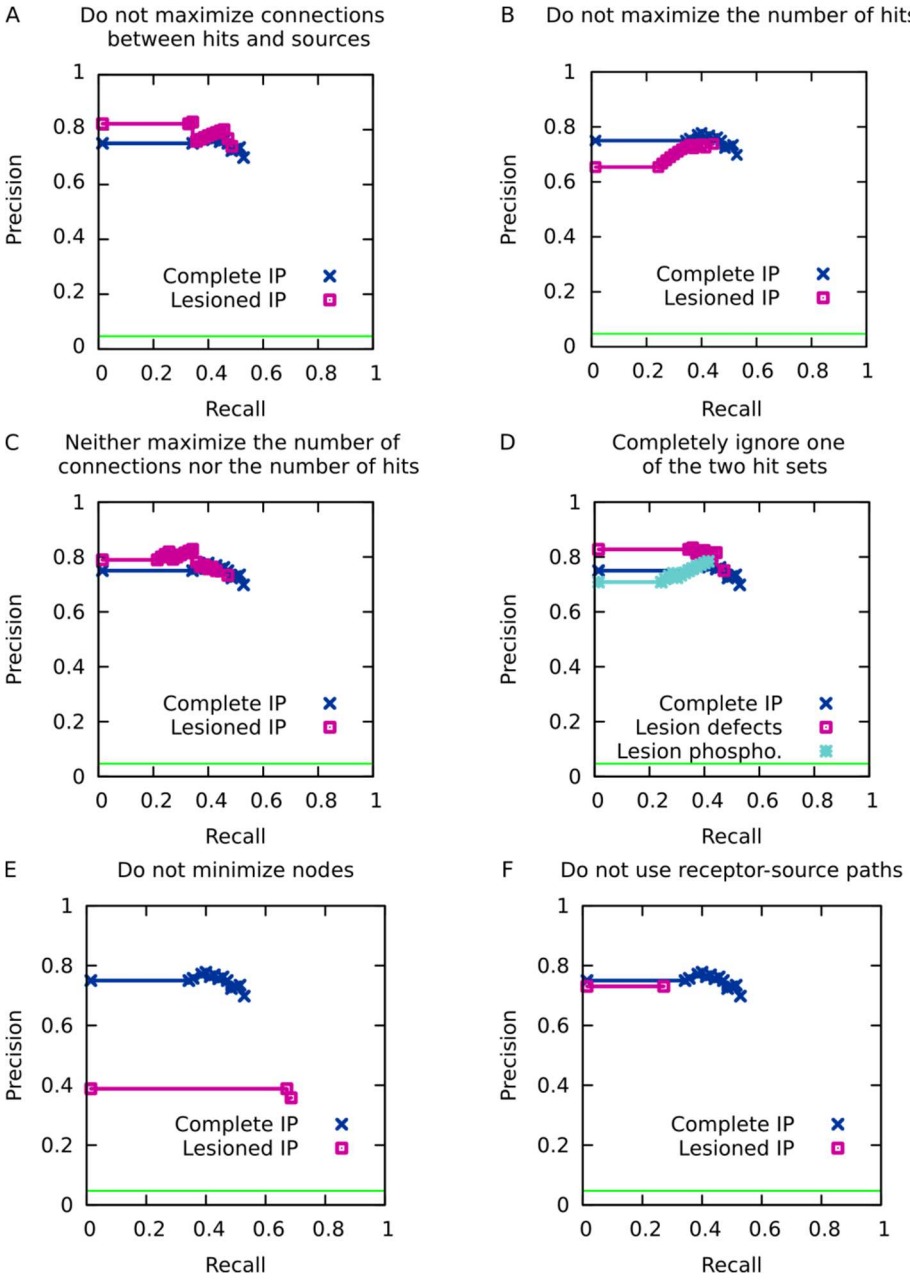
C. Final scores for the candidate bifurcation points. Note that while node *a* has superior information gain */(a)* compared to *c*, its bifurcation score *B(a)* is penalized by its high split information. Node *c* is therefore the higher-ranked bifurcation point.



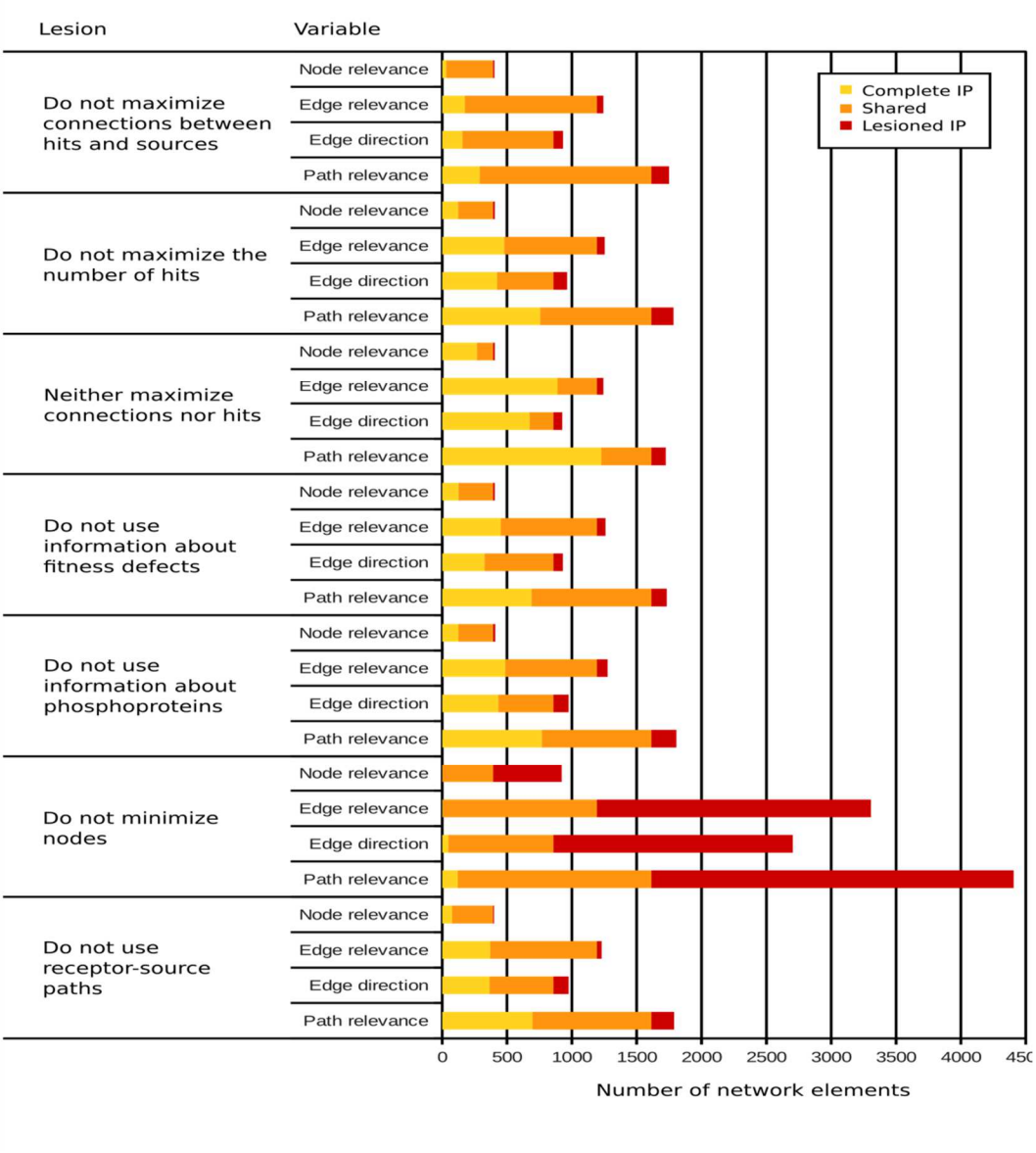
Appendix Figure A10. IP stability analysis results.

A. Precision-recall curves for inferred ensembles of three sizes. The horizontal green line in each panel indicates the precision achieved by predicting that all test cases are positive examples.

B. Similarity of confidence values for different types of network elements (nodes, edges, paths) between different sizes of ensembles. Each bar represents the total number of elements in the union of the complete, 10,000-solution ensemble with a smaller ensemble. The central portion of the bar, "Shared", represents the number of elements that are in the intersection of both ensembles. The end caps, labeled "10k only" and "Smaller only", represent the number of elements unique to each ensemble. The longer the center is relative to the ends, the more similar are the two ensembles in their predictions.



Appendix Figure A11. Precision-recall curves for the lesioned IP. The horizontal green line in each panel indicates the precision achieved by predicting that all test cases are positive examples. The lesioned components are: A maximization of connections, B maximization of hits, C maximization of both connections and hits, D an entire set of experimental node labels (fitness-contributions or pospho-proteins), E minimization of nodes, F inclusion of receptor-source pairs curated from literature.



Appendix Figure A12. Similarity of ensembles inferred by the lesioned IPs. Comparison of the complete ensemble to the lesioned ensembles on the basis of the confidence values given to each type of network element (nodes, edges, paths). Each bar represents the total number of elements in the union of the complete ensemble with a lesioned ensemble. The central portion of the bar, "Shared", represents the number of elements that are in the intersection of both ensembles. The end caps, labeled "Complete IP" and "Lesioned IP", represent the number of elements unique to each ensemble. The longer the center is relative to the ends, the more similar are the two ensembles in their predictions.

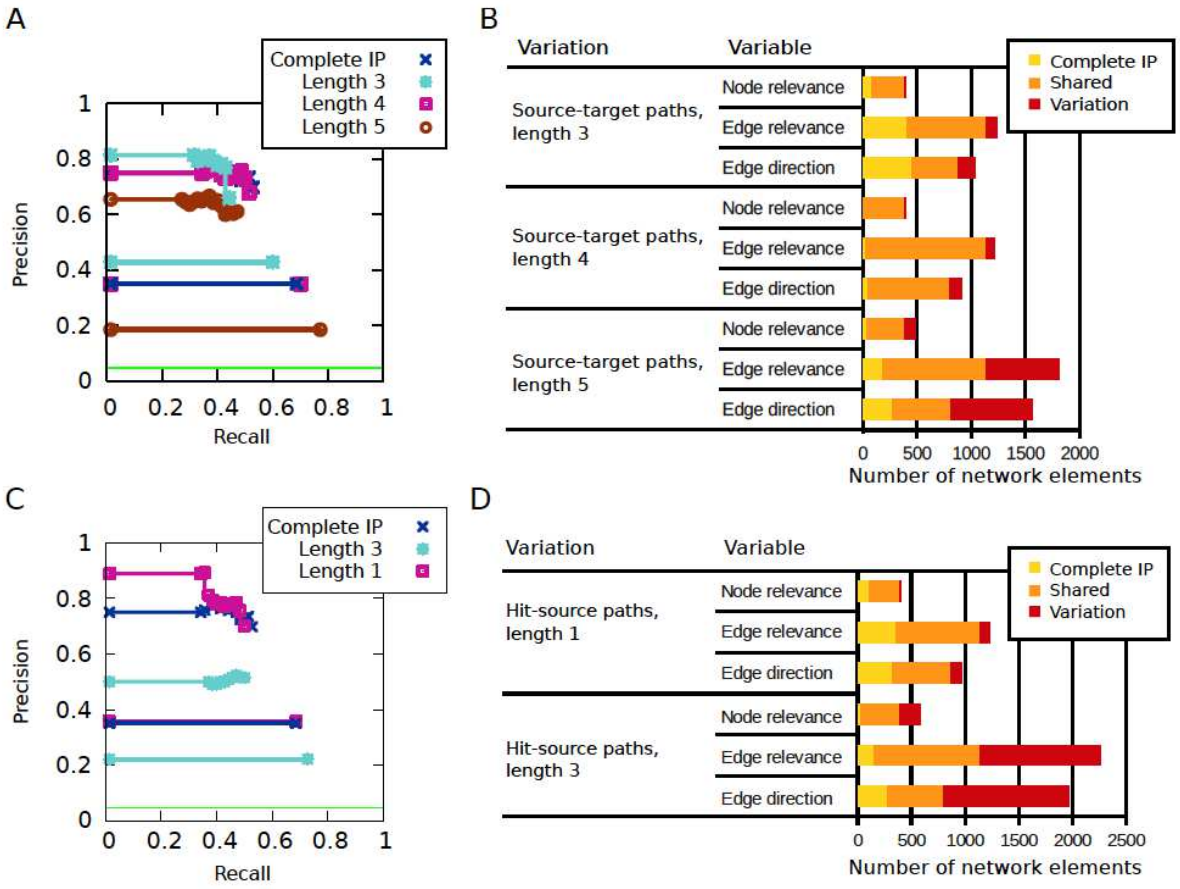
Appendix Figure A13. Precision-recall curves and stability results derived from changing the length of candidate paths. For each variation, we show two PR curves in the same color. The top curve represents the inferred ensemble, while the bottom curve represents the performance of candidate paths provided as input to the inference m. The horizontal green line in PR plot indicates the precision achieved by predicting that all test cases are positive examples. In the stability analysis panel, each bar represents the total number of elements in the union of the complete ensemble with a lesioned ensemble. The central portion of the bar, "Shared", represents the number of elements that are in the intersection of both ensembles. The end caps, labeled "Complete IP" and "Variation", represent the number of elements unique to each ensemble. The longer the center is relative to the ends, the more similar are the two ensembles in their predictions. In contrast to previous stability analyses, we do not compare the ensembles based on path relevance, as the input sets of paths are different.

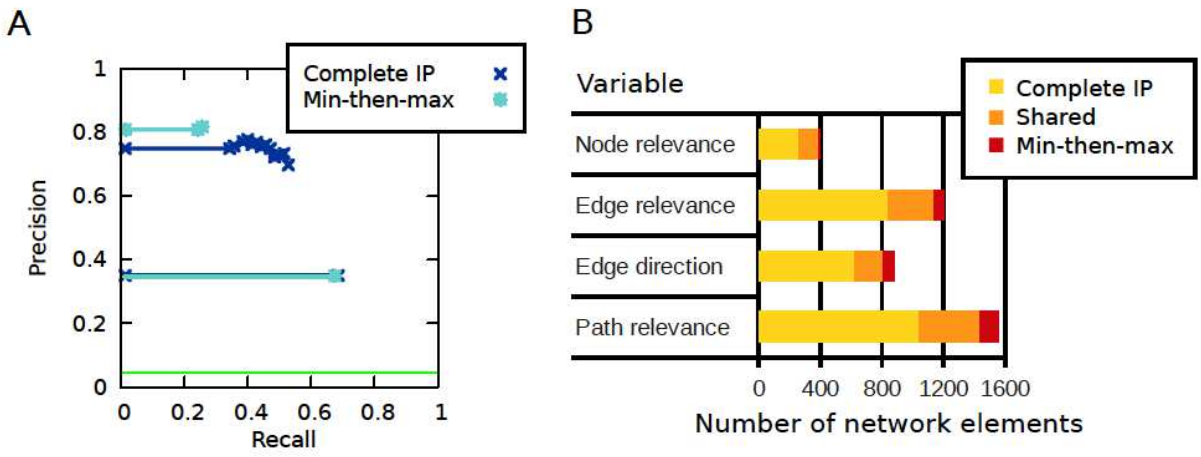
A Precision-recall analysis for varying source-target path length

B Stability analysis for varying source-target path length

C Precision-recall analysis for varying hit-source path length

D Stability analysis for varying hit-source path length


A
Complete IP
Length 3
Length 4
Length 5
Precision
Recall
B
Variation
Variable
Source-target paths, length 3
Source-target paths, length 4
Source-target paths, length 5
Node relevance
Edge relevance
Edge direction
Complete IP
Shared
Variation
Number of network elements
C
Complete IP
Length 3
Length 1
Precision
Recall
D
Variation
Variable
Hit-source paths, length 1
Hit-source paths, length 3
Node relevance
Edge relevance
Edge direction
Complete IP
Shared
Variation
Number of network elements



Appendix Figure A14. Precision-recall curves and stability results derived from reversing the order in which the objective function components are solved.

A. A Precision-recall analysis. For each ensemble, two PR curves are shown, one for each variation on the objective function procedure. The top curve represents the performance of the inferred sub- network, while the bottom curve represents the performance of input candidate paths. The curves for candidate paths are identical, as the input was the same for both versions. The horizontal green line in each panel indicates the precision achieved by predicting that all test cases are positive examples.

B. B Stability analysis. Each bar represents the total number of elements in the union of the complete ensemble with a lesioned ensemble. The central portion of the bar, "Shared", represents the number of elements that are in the intersection of both ensembles. The end caps, labeled "Complete IP" and "Min-then-max", represent the number of elements unique to each ensemble. The longer the center is relative to the ends, the more similar are the two ensembles in their predictions.

Appendix Table A1: Validation of predicted regulators.

* = data from (GITTER *et al.* 2013). ^T0 = unstressed mutant. + = data from (ALEJANDRO-OSORIO *et al.* 2009).

NA = no targets predicted, by nature of the node inclusion.

Predicted regulators	Defective NaCl transcriptome response	Significant overlap in measured vs. predicted targets (*p*-value)	Number of in-path validated nodes	Number of scorable in-path nodes	Percent validated in-path nodes
Arf3	Yes	Yes (*p* =6e-5)	2	2	100%
Bck1	Yes	NA	2	4	50%
Bem1*	Yes	NA	1	4	25%
Cdc14	Yes	Yes (2e-14)	14	18	78%
Cka2	Yes	Yes (3e-3)	7	10	70%
Ckb1/2	Yes	No (1); overlap with Cka2 targets (*p* =3e-12)	8	21	38%
Gal11*	Yes	Yes (*p* =2e-4)	5	10	50%
Kin2^	Yes	No (*p* =0.57); marginal overlap with T0-affected (*p* = 0.005)	3	17	18%
Nkk1	Yes	No (*p* =1)			
Pho85	Yes	Yes (*p* =5e-8)	14	21	67%
Rpd3+	Yes	NA			
Scd6	No				
Tpk2*	Yes	Yes (*p* =5e-20)	10	24	42%
Yak1	Yes	Yes (*p* =1e-10)	5	10	50%

Appendix Table A2: Provenance of interactions in the background network. Notes about database extraction procedures are given in italicized text.

Interaction	Source	Directed	Count
	Protein-protein interactions		
General protein-protein	Stark *et al.* (2006) (BioGRID)	Both	17,306
	Interactions supported by at least two categories of experimental methods; treated phosphorylation edges separately; downloaded 3/2011		
PPIs involving kinases and phosphatases	Breitkreutz *et al.* (2010) (YeastKinome.org)	No	989
	Fasolo *et al.* (2011)	No	1,028
	Sharifpoor *et al.* (2011) (Yeast KID)	No	138
	Interactions without evidence of direct phosphorylation and annotated with p<0.01; downloaded 9/2012		
	Sharifpoor *et al.* (2011)	Yes	414
Kinase-substrate interactions	*Annotated with p < 0.01 and one of the following evidence codes: "(LTP in vitro kinase assays OR /n vitro phosphorylation site mapping (Mass Spec, Phospho-specific antibodies by Western, in vitro site-directed mutagenesis) OR In vivo site-directed mutagenesis in substrate showing same biological consequence as the kinase delete OR Reduction in phospho-peptide in vivo by mass-spec OR in vivo phosphorylation site mapping using phospho-specific antibodies (Western blot) or by phospho-peptide mapping)"; downloaded 9/2012*		
	Ptacek *et al.* (2005); Stark *et al.* (2006)	Yes	5,315
	All interactions from Ptacek et al. *(2005), plus low-throughput phosphorylation/dephosphorylation interactions from BioGR/D, March 2011*		
Manually curated	This work	Both	17
	Hand-constructed after inspection of neighborhoods of interrogated sources		

Continued on next page

Interaction	Source	Directed	Count
	Protein-nucleic acid interactions		
Protein-DNA	Abdulrehman *et al.* (2011); Everett *et al.* (2009); Guelzim *et al.* (2002); MacIsaac *et al.* (2006); Venters *et al.* (2011)	Yes	259,565
Osmotic-stress specific, protein-DNA	Huebert *et al.* (2012)	Yes	1,225
Salt-stress specific protein-DNA	Ni *et al.* (2009)	Yes	2,144
Protein-RNA	Hogan *et al.* (2008); Scherrer *et al.* (2010); Tsvetanova*et al.* (2010)	Yes	17,868
	Other interaction types		
Between metabolic enzymes	Heavner *et al.* (2012)	Both	1,153
	Created binary interactions from enzymes reported to catalyze adjacent reactions; reported as "directed" if the reaction was annotated as "irreversible"		
Complex membership	Heavner *et al.* (2012); Pu *et al.* (2009)	Yes	2,183
	Directed interactions from protein to complex		
Inferred complex-complex interactions	Heavner *et al.* (2012); Pu *et al.* (2009), Stark *et al.* (2006)	No	22
	Interaction between two complexes inferred if >50% possible protein pairs have interactions in BioGR/D		
Inferred complex-protein interactions	Heavner *et al.* (2012); Pu *et al.* (2009), Stark *et al.* (2006)	No	1,128
	Interaction between complex and protein inferred if >50% possible protein pairs have interactions in BioGRID		

Appendix Table A3: Sets of network elements that are provided as input to the method.

Network elements	Set	Description
Nodes	N	All nodes
	N^S	Source nodes
	N^R	Receptor nodes
	N^F	Fitness-contribution hits
	N^P	Phospho-proteomic hits
	N^T	Target nodes
	$N^T(s)$	Dysregulated targets of source s
	$N^S(r)$	Sources that are downstream of receptor r
	$N(e)$	Nodes in edge e
Edges	E	All edges
	E^D	Directed edges
	E^U	Undirected edges
	$E(n)$	Edges that touch node n
Paths	P	All paths
	P^{ST}	Source-target paths
	$P^{ST}(s, t)$	Source-target paths between source s and target t
	P^{RS}	Receptor-source paths
	$P^{RS}(r, s)$	Receptor-source paths between receptor r and source s
	P^{FS}	fitness-contribution hit-source paths
	$P^{FS}(f, s)$	fitness-contribution hit-source paths between hit f and source s
	P^{SS}	Source-source paths
	$P^{SS}(s_i, s_j)$	Paths between source s_i and source s_j

Appendix Table A4: Integer program variables. Binary variables represent the status of nodes, edges, paths, and pairs in the network.

Network elements	Variable	Interpretation	Values
Paths p	σ_p	Relevant	no=0, yes=1
Edges e	x_e	Relevant	no=0, yes=1
	d_e	Direction	back=0, forward=1
Nodes n	y_n	Relevant	no=0, yes=1
Pairs (n_i, n_j)	c_{ij}	Connected	no=0, yes=1

Appendix Table A5: Coverage of each source's targets and candidate TF/RBPs by the candidate paths. 'Prop.' columns give the proportion of targets covered by candidate paths. Sources marked with the message 'Self' are sources that are themselves TFs or RBPs; for these sources, all targets were covered by the addition of inferred regulatory interactions.

	Targets			Candidate TF/RBPs		
Source	Covered	Total	Prop.	Covered	Total	Prop.
Gpb2	174	234	0.74	12	17	0.71
Hog1	1843	1912	0.96	41	44	0.93
Mck1	861	886	0.97	48	52	0.92
Npr2	71	131	0.54	9	12	0.75
Npr3	218	267	0.82	16	28	0.57
Pde2	526	573	0.92	26	41	0.63
Pph3	209	247	0.85	12	22	0.55
Rim15	513	530	0.97	36	41	0.88
Sub1	422	515	0.82	15	19	0.79
Tpk1	53	124	0.43	6	10	0.6
Whi2	167	296	0.56	8	9	0.89
Dot6/Tod6	–	259	–	–	Self	–
Msn2	–	209	–	–	Self	–
Rim101	–	302	–	–	Self	–
Swc3	–	365	–	–	Self	–
Swc5	–	138	–	–	Self	–

Appendix Table A6: Top 15 consensus nodes ranked by degree.

Name	ORF	Degree	Annotation
Tpk1	YJL164C	69	Source, kinase
Mck1	YNL307C	44	Source, kinase
Hog1	YLR113W	42	Source, kinase
Pho85	YPL031C	31	Kinase
Cdc28	YBR160W	28	Kinase
Rpb1	YDL140C	27	RNA Pol II core subunit
Kin2	YLR096W	24	Kinase
Tpk2	YPL203W	24	Source, kinase
Cdc14	YFR028C	23	Kinase
Rad53	YPL153C	23	Kinase
Ubi4	YLL039C	23	Ubiquitin
Smt3	YDR510W	21	SUMO
Bmh1	YER177W	20	14-3-3 scaffold
Hrr25	YPL204W	20	CTD-binding kinase
Tpk3	YKL166C	19	Kinase

Appendix Table A7: Enrichment analysis results. The 'Prop.' columns give the proportion of each subnetwork that is in the relevant gene or interaction set; for the 'Permutations' row, this value is the average over all 1,000 permutations. The '*p*-value' columns are calculated by a comparison to the consensus subnetwork (derived from the 10,000-solution ensemble). Asterisks (*) indicate a result that is significant at $p < 0.05$. For comparisons to the candidate and background networks, *p*-values are calculated using the hypergeometric test. For comparisons to the permuted ensembles, *p*-values are calculated as the fraction of permutations having an equal or greater proportion of relevant genes (or lower propoerion of likely negatives) compared to the consensus subnetwork.

	Relevant gene sets											
	True positives		Likely negatives		Kinases and phosphatase		General stress proteins		Essential genes		Genetic interaction	
Subnetwork	Prop.	*p*-value	Prop.	*p*-value	Prop.	*p*-value	Prop.	*p*-value	Prop.	*p*-value	Prop.	*p*-value
Consensus	0.156		0.050		0.200		0.269		0.294		0.051	
Candidate network	0.065	2e-6 *	0.121	2e-4 *	0.124	0.001 *	0.145	2e-6 *	0.319	0.810	0.033	≈ *0**
Background network	0.015	5e-20 *	0.302	6e-17 *	0.032	8e-18 *	0.050	5e-21 *	0.194	0.001 *	0.008	≈ *0**
Permutations	0.072	0.002 *	0.133	0.007 *	0.183	0.309	0.185	0.007 *	0.348	0.899	0.035	0.003 *

Appendix Table A8: Enrichment analysis of lesioned IPs. 'Prop' columns show the proportion of each lesioned or complete consensus subnetwork that is composed of the relevant gene set being tested. *P*-values are calculated by a two-tailed *z*-test of proportions comparing the complete consensus subnetwork (derived from the 1,000 solution ensemble) to the consensus subnetwork inferred using a different path length (also derived from 1,000-solution ensembles). Asterisks (*) indicate significance at $p < 0.05$. For significant results, bolded proportions and *p*-values indicate comparisons in which the original, complete IP has a higher proportion of relevant genes/interactions (or lower proportion of likely negatives). Italicized proportions and *p*-values indicate the opposite.

	Relevant gene sets											
	True positives		Likely negatives		Kinases and phosphatase		General stress proteins		Essential genes		Genetic interaction	
IP Version	Prop.	*p*-val.	Prop.	*p*-val.	Prop.	*p*-val.	Prop.	*p*-val.	Prop.	*p*-val.	Prop.	*p*-val.
Complete IP (1k sols.)	0.168		0.054		0.197		0.257		0.293		0.052	
Do not maximize connections	0.179	0.801	0.057	0.901	0.200	0.958	0.250	0.881	0.279	0.775	**0.047**	**≈ 0** *
Do not maximize	0.131	0.380	0.066	0.664	0.182	0.738	0.255	0.968	0.263	0.554	*0.061*	*≈0**
Neither maximize connections nor hits	0.183	0.757	0.043	0.700	0.172	0.613	0.237	0.709	0.247	0.426	*0.071*	*≈0**
Do not use fitness-defects	0.202	0.454	0.047	0.774	0.202	0.933	0.256	0.974	0.310	0.757	0.050	0.242
Do not use phospho-hits	0.133	0.409	0.052	0.937	0.193	0.913	0.252	0.911	0.259	0.510	*0.060*	*≈0**
Do not minimize	**0.067**	**0.000 ***	**0.113**	**0.024 ***	**0.121**	**0.009 ***	**0.147**	**0.001 ***	0.315	0.591	**0.033**	**≈0***
Do not use receptor-source paths	0.133	0.394	0.049	0.845	0.182	0.724	0.273	0.762	0.273	0.687	0.052	0.553

Appendix Table A9: Enrichment analysis of subnetworks inferred from varied candidate path lengths. 'Prop' columns show the proportion of each consensus subnetwork that is composed of the relevant gene set being tested. *P* -values are calculated by a two-tailed *z*-test of proportions comparing the complete consensus subnetwork (derived from the 1,000 solution ensemble) to the consensus subnetwork inferred using a different path length (also derived from 1,000-solution ensembles). Asterisks (*) indicate significance at $p < 0.05$. For significant results, bolded proportions and *p*-values indicate comparisons in which the original, complete IP has a higher proportion of relevant genes/interactions (or lower proportion of likely negatives). Italicized proportions and p-values indicate the opposite.

	Relevant gene sets											
	True positives		Likely negatives		Kinases and		General stress proteins		Essential genes		Genetic interactions	
IP Version	Prop.	*p*-val.	Prop.	*p*-val.	Prop.	*p*-val.	Prop.	*p*-val.	Prop.	*p*-val.	Prop.	*p*-val.
Complete IP (1k sols.)	0.168		0.054		0.197		0.257		0.293		0.052	
Source-target paths, length 3	0.176	0.857	0.046	0.751	0.221	0.616	0.260	0.968	0.328	0.518	*0.057*	*≈0* *
Source-target paths, length 4	0.168	1.000	0.048	0.803	0.204	0.891	0.258	1.000	0.287	0.904	0.052	0.504
Source-target paths,	0.124	0.260	0.071	0.517	0.195	0.960	0.225	0.484	0.343	0.327	**0.044**	**≈0***
Source-hit paths, length 1	0.191	0.604	0.038	0.524	0.214	0.731	0.260	0.970	0.282	0.836	*0.061*	*≈0* *
Source-hit paths, length 3	**0.100**	**0.039** *	0.107	0.055	0.184	0.724	0.207	0.223	*0.387*	*0.048* *	**0.041**	**≈0***

Appendix Table A10: Enrichment analysis of reordered IPs. 'Prop' columns show the proportion of each consensus subnetwork that is composed of the relevant gene set being tested. *P* -values are calculated from a two-tailed *z*-test of proportions comparing the complete consensus subnetwork (derived from 1,000 solutions) to the consensus subnetwork inferred using a different ordering of the objective function procedure (also derived from 1,000 solutions). Asterisks (*) indicate significance at $p < 0.05$. For the significant result, the italicized proportion (and *p*-value) indicates a comparison in which the original, complete IP has a lower proportion of relevant genetic interactions.

	Relevant gene sets											
	True positives		Likely negatives		Kinases and phosphatase		General stress proteins		Essential genes		Genetic interaction	
IP Version	Prop.	*p*-value	Prop.	*p*-value	Prop.	*p*-value	Prop.	*p*-value	Prop.	*p*-value	Prop.	*p*-value
Complete IP (1k sols.)	0.168		0.054		0.197		0.257		0.293		0.052	
Min-then-max	0.181	0.787	0.043	0.686	0.170	0.586	0.255	0.969	0.245	0.398	*0.076*	*≈ 0 **

Appendix B

Exploiting the yeast stress-activated signaling network to inform on stress biology and disease signaling

Yi-Hsuan Ho & Audrey P. Gasch

This Work is published on *Current Biology* (2015) May 10. [Epub ahead of print]

I contributed to write up the yeast section.

Abstract

Healthy cells utilize intricate systems to monitor their environment and mount robust responses in the event of cellular stress. Whether stress arises from external insults or defects due to mutation and disease, cells must be able to respond precisely to mount the appropriate defenses. Multi-faceted stress responses are generally coupled with arrest of growth and cell-cycle progression, which both limits the transmission of damaged materials and serves to reallocate limited cellular resources toward defense. Therefore, stress defense *versus* rapid growth represent competing interests in the cell. How eukaryotic cells set the balance between defense *versus* proliferation, and in particular knowledge of the regulatory networks that control this decision, are poorly understood. In this *Perspective*, we expand upon our recent work inferring the stress-activated signaling network in budding yeast, which captures pathways controlling stress defense and regulators of growth and cell-cycle progression. We highlight similarities between the yeast and mammalian stress responses and explore how stress-activated signaling networks in yeast can inform on signaling defects in human cancers.

Introduction

All cells must allocate resources to balance conflicting physiological demands. From free-living microbes to multicellular animals, actively growing cells funnel resources to biosynthesis needed for division. A key consumer of cellular resources is translation, since production of ribosomes and translation itself require substantial amounts of energy (WARNER 1999; THOMAS 2000). The drive for proliferation often comes at the cost of stress tolerance. For example, rapidly growing yeast cells are the most sensitive to environmental insults (ZAKRZEWSKA *et al.* 2011), whereas slow-growing clones are extremely resistant to adversity (ELLIOTT and FUTCHER 1993; LU *et al.* 2009; LEVY *et al.* 2012). The same relationship exists in higher organisms, most notably in rapidly dividing cancer cells that are often the most susceptible to chemotherapy drugs and treatments (JONES and THOMPSON 2009). Understanding how cells allocate cellular

resources toward proliferation *versus* stress defense is therefore critically important for understanding stress defense and, in turn, human disease.

The balance between cellular growth and stress defense can be modulated in response to stressful situations, and thus the mobilization of defense strategies is often coordinated with arrest of growth and cell-cycle progression (LOPEZ-MAURY *et al.* 2008). At the same time, defense responses include orchestrated changes to transcription, translation, and post-translational protein modifications that mediate changes in protein abundance, localization, and function. Although much is known about the signaling pathways that control individual physiological responses to stress, how signaling is integrated into a single cellular network that coordinates a multi-faceted response is only beginning to emerge. To address this question, we recently developed an experimental and computational pipeline in budding yeast *Saccharomyces cerevisiae*. The approach integrates stress-activated transcriptome alterations, phospho-proteomic changes, and gene-fitness contributions with large-scale protein interaction data, to implicate a single network coordinating the response to, in this case, salt stress (CHASMAN *et al.* 2014). The resulting network identified known and novel salt-regulated signaling proteins, uncovered previously unrecognized cross-connections between signaling pathways, and implicated important decision points in the growth-*versus*-defense decision. Importantly, the network is enriched with statistical significance for genes whose human orthologs cause cancer when mutated in somatic tissues.

In this *Perspective*, we first provide a broad overview of signaling pathways that control proliferation *versus* stress defense in yeast and mammalian cells (Fig 1). We then highlight

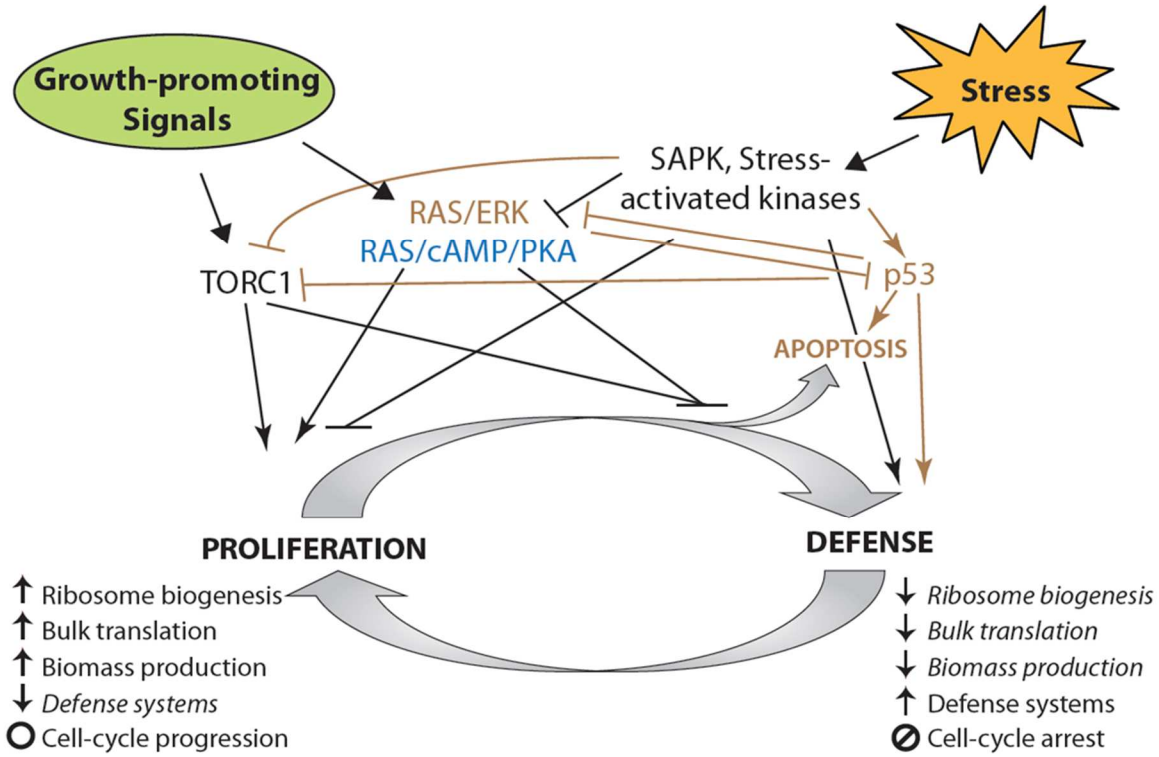


Figure 1: Proliferation and stress defense compete for limited resources in the cell. A simplified view of the signaling processes that occur in yeast and/or mammalian cells responding to growth promoting signals versus cellular stress signals. Signaling molecules specific to yeast are shown in blue and those unique to mammalian cells are shown in brown. See text for details.

opportunities for using the stress-activated signaling network in yeast to understand the growth-versus-defense decision points in eukaryotic cells.

The Environmental Stress Response: a common yeast response to diverse stresses.

Actively growing microbial cells maintain high translational capacity to support division, in part by promoting ribosome biogenesis. Supporting high translational capacity therefore consumes a significant amount of the cell's resources. Under optimal conditions, cells produce ~2000 ribosomes per minute (WARNER 1999). The high demand for ribosomes requires active transcription of the rDNA locus; nearly 60% of cellular transcription is devoted to rRNA production by RNA polymerase (Pol) I. Furthermore, over half the activity of Pol II transcription occurs at genes encoding ribosomal proteins (RPs), while Pol III is dedicated to transcribing 5S rRNA and tRNAs (WARNER 1999; RUDRA and WARNER 2004). A significant fraction of Pol II is also dedicated to transcribing genes encoding ribosome assembly factors (commonly called the RiBi regulon) (JORGENSEN *et al.* 2004). In addition to consuming transcriptional capacity, transcripts emerging from RP and RiBi genes are highly translated and thus monopolize a large fraction of the cell's ribosomes (90% of which are actively involved in translation during rapid growth) (ARAVA *et al.* 2003; VON DER HAAR 2008). Thus, significant resources support ribosome biogenesis and translational capacity in cells actively growing under optimal conditions.

Upon a sudden shift to stressful conditions, cells must rapidly adjust resource allocation to mount cellular defense strategies. Much has been gleaned from studying transcriptome changes in response to stress. Stressed yeast cells activate condition-specific transcript changes at specialized genes that specifically address the particular stress condition. Concurrently, cells activate the environmental stress response (ESR) (GASCH *et al.* 2000; CAUSTON *et al.* 2001), a common response to diverse types of stress. The ESR includes ~300 induced ESR (iESR) genes that are broadly involved in stress defense and ~600 repressed

(rESR) genes that encode RPs and RiBi factors (GASCH *et al.* 2000; GASCH 2002). The expression patterns of the iESR and rESR genes are strikingly anti-correlated across diverse types of stress, indicating that they are likely regulated by the same upstream signaling systems (GASCH 2002).

The ESR is postulated to relate to stress defense, given the functions of the encoded proteins. But surprisingly, activation of the ESR – and of stress-activated transcript changes in general – is not required to survive the initial stress treatment, at least in yeast (BERRY and GASCH 2008; WESTFALL *et al.* 2008; MITCHELL *et al.* 2009; BERRY *et al.* 2011). Instead, transcript changes are critical for surviving subsequent stressful insults, through a phenomenon known as acquired stress resistance (BERRY and GASCH 2008). Induction of iESR transcripts correlates with increased abundance of the encoded proteins, which most likely serve to defend against stress (BERRY and GASCH 2008; LEE *et al.* 2011; VOGEL *et al.* 2011). However, the role of the rESR transcripts has remained somewhat elusive. Previous observations found that the abundance of RP and RiBi transcripts correlates with growth rate upon nutrient restriction (JORGENSEN *et al.* 2002; JORGENSEN *et al.* 2004; REGENBERG *et al.* 2006; CASTRILLO *et al.* 2007; BRAUER *et al.* 2008), leading to the suggestion that rESR gene repression underlies growth and translational arrest during stress. However, we later showed that stress-dependent repression of RP and RiBi genes is not required for translational arrest or growth reduction, at least during salt stress (LEE *et al.* 2011). Instead, we proposed that rESR transcript reduction serves to release engaged ribosomes, thereby redirecting limited translational capacity to newly made mRNAs (LEE *et al.* 2011). More recent work from our lab also implies that reduced transcription of rESR genes is required to reallocate RNA Pol II to iESR and other defense genes (CHASMAN *et al.* 2014). Thus, the anti-correlation between rESR and iESR gene expression is at least partly explained by competition of these gene modules for limited resources.

Regulation of the growth-versus-defense decision in yeast.

Two central signaling pathways regulating ribosome biogenesis are the target of rapamycin (TOR) and the RAS/cAMP/Protein Kinase A (PKA) signaling pathways, which respond to nitrogen and carbon availability, respectively (NEUMAN-SILBERBERG *et al.* 1995; WANG *et al.* 2004; XIAO and GROVE 2009; BROACH 2012). Both pathways are largely conserved from yeast to mammals. The TOR complex 1 (TORC1) directly binds the rDNA locus under optimal conditions (LI *et al.* 2006) and promotes rDNA transcription, by promoting activity of the rDNA transcription factor Rrn3 and, indirectly, by inhibiting the Pol III repressor Maf1 (CLAYPOOL *et al.* 2004; LEE *et al.* 2009; PHILIPPI *et al.* 2010). Maf1 is also directly regulated by PKA, which suppresses the inhibitory activity of Maf1 during rapid growth (MOIR *et al.* 2006; WILLIS and MOIR 2007). Both TORC1 and PKA pathways have been implicated in regulating the RP genes and RiBi regulon, by modulating transcriptional activators (Fhl1/Crf1/Ifh1, Rap1, Sfp1) and repressors (Dot6/Tod6) (KLEIN and STRUHL 1994; JORGENSEN *et al.* 2004; MARION *et al.* 2004; MARTIN *et al.* 2004; LIPPMAN and BROACH 2009; HUBER *et al.* 2011). The TOR and RAS/PKA pathways are clearly interconnected, even though the precise relationships in regulating ribosomal biogenesis are still controversial. Although some studies have suggested PKA as a downstream effector of the TOR pathway, possibly via the TORC1 downstream target Sch9 (MARTIN *et al.* 2004; SCHMELZLE *et al.* 2004; SOULARD *et al.* 2010), other evidence has indicated that PKA and TOR are two parallel pathways activating RP gene expression (ZURITA-MARTINEZ and CARDENAS 2005; RAMACHANDRAN and HERMAN 2011).

Along with promoting biogenesis of ribosomes and translational machinery, PKA and TOR act in parallel to antagonize ESR activation, in part by suppressing Msn2 and Msn4, the so-called 'general stress' transcription factors (GÖRNER *et al.* 1998; SMITH *et al.* 1998; BECK and HALL 1999). Thus, under stressful conditions TOR and PKA pathways must be suppressed – although the mechanisms remain poorly understood – while stress-activated signaling pathways

are mobilized. For example, the MAP kinase Hog1 (orthologous to human kinase p38) is activated by osmotic shock and related stresses to coordinate myriad responses, including induction and repression of ~2000 transcripts (including iESR and rESR genes) (O'ROURKE and HERSKOWITZ 2004; CHASMAN *et al.* 2014), reallocation of RNA Pol II distribution (COOK and O'SHEA 2012; NADAL-RIBELLES *et al.* 2012; CHASMAN *et al.* 2014), alterations in translational capacity (TEIGE *et al.* 2001; UESONO and TOH-E 2002; NAGIEC and DOHLMAN 2012), altered metabolism including osmolyte production (ALBERTYN *et al.* 1994; PROFT and STRUHL 2004), and transient cell-cycle arrest (BELLÍ *et al.* 2001; ESCOTÉ *et al.* 2004; CLOTET *et al.* 2006; YAAKOV *et al.* 2009). Alternate signaling pathways are specifically activated by other stresses (such as PKC in the case of cell wall stress (HEINISCH *et al.* 1999), ATM/ATR in response to DNA damage (ABRAHAM 2001), and the AMPK ortholog Snf1 during starvation (CONRAD *et al.* 2014)). How activation of these stress-regulated pathways is coordinated with suppression of the TOR and PKA pathways remains unclear. Nonetheless, coordination and integration of signaling pathways is likely to be key for precise allocation of limited resources and adaptation to a new environment.

Stress responsive signaling in mammalian cells.

There are many similarities in the molecular responses to stress in yeast and mammalian systems, despite the added complexities of multi-celled organisms. Studying cellular stress responses in the organismal context is particularly challenging, and thus much of the knowledge comes from culturing immortalized lines or primary cells responding to external stresses. A variety of environmental stresses have been studied extensively in mammalian systems, with the greatest focus on stresses linked to infection and disease, including DNA damaging agents, oxidants, elevated temperature, chemotherapy and other drugs, and hypoxic conditions that mimic the inter-tumor environment. Like yeast, cells must decide between maintaining proliferation or mediating cell-cycle arrest and stress defense. In addition, cells in a multicellular

context have a third option in apoptosis, to clear cells that simply cannot recover from the insult. Thus, during adversity mammalian cells face a multi-pronged decision point directing cells toward proliferation, defense, or death. Although there are many specific aspects of stress responses that are critically influenced by tissue type, developmental phase, and stress identity, several themes can be generalized from research in this area.

As in yeast, high ribosome biogenesis and translational activity are maintained in actively growing mammalian cells, promoted via signaling through mTORC1 and RAS to the ERK MAP kinase (HANNAN *et al.* 2011). mTORC1 supports ribosome biogenesis through rRNA production, by stimulating transcription of rDNA genes as well as genes encoding rRNA processing factors (MAYER *et al.* 2004; MAYER and GRUMMT 2006; KANTIDAKIS *et al.* 2010; CHAUVIN *et al.* 2014). mTOR signaling also promotes ribosome production by favoring translation of so-called 5-TOP transcripts that encode RPs and translation factors (MEYUHAS and DREAZEN 2009). mTOR broadly enables cap-dependent translation by suppressing the eIF4e inhibitory binding protein 4E-BP1 (GINGRAS *et al.* 1998) and by activating the S6K kinase (orthologous to yeast Sch9), which further stimulates translation initiation and elongation (MAGNUSON *et al.* 2012; ROUX and TOPISIROVIC 2012). Mitogen-activated RAS/ERK signaling also provides a dual boost to translational activity: ERK signaling stabilizes the growth-enhancing transcription factor Myc, which induces Pol I, II, and III-dependent transcription of ribosome components, and phosphorylates rDNA transcription factors Tif1-A (orthologous to Rrn3 in yeast) and UBF to stimulate rDNA transcription (HANNAN *et al.* 2011; KUSNADI *et al.* 2015). At the same time, ERK phosphorylation activates downstream kinases that promote cap-dependent translation initiation and elongation (ROUX and BLENIS 2004; ROUX and TOPISIROVIC 2012). As in yeast, there are multiple points of cross-talk between the RAS and mTORC pathways that are thought to produce precisely tuned growth behavior (MENDOZA *et al.* 2011).

In response to stressful situations, many cell types mount a response that shares hallmarks with the yeast ESR. Murray *et al.* was one of the first to compare and contrast stress-activated transcriptome changes in primary and immortalized human cells: at least under the conditions and time frames studied, there were relatively few commonly induced or repressed genes upon diverse stress treatments (that included heat, oxidants, ER stress, and crowding) (MURRAY *et al.* 2004). However, more recent studies have identified common responses to different stresses within a given cell type, albeit with smaller magnitude transcript changes than seen in yeast. Nayak *et al.* leveraged statistical power in a large study of B-cell responses to ER stress and ionizing radiation, finding substantial overlap in response to the two stresses (NAYAK *et al.* 2014). Several studies interrogating p53 activity (see below) identified a common response that persists across several cell lines and conditions, with induced genes related to stress defense and regulation of cell cycle or apoptosis and repressed genes linked to rDNA transcription, ribosome biogenesis, translation and cell cycle/apoptotic factors that work antagonistically to induced genes (CAIRNS and WHITE 1998; BUDDE and GRUMMT 1999; ZHAI and COMAI 2000; WEI *et al.* 2006; MENENDEZ *et al.* 2009; NIKULENKOV *et al.* 2012; SCHLERETH *et al.* 2013). The coordinated induction of defense genes with repression of rDNA/protein synthesis genes is reminiscent of the coordinated expression changes of the induced and repressed genes modules of the yeast ESR.

Beyond the level of gene expression, other aspects of the stress response are conserved from yeast to humans. Upon stressful insults, transcription of rDNA is generally sharply decreased and coupled to an overall drop in cap-dependent translation (SPRIGGS *et al.* 2010; LIU and QIAN 2014; KUSNADI *et al.* 2015). These responses are mediated by abrogated signaling through the mTOR/S6K and RAS/ERK pathways and via activation of the stress-activated protein kinases (SAPKs) JNK and p38. p38 isoforms, including the broadly expressed p38α and β along with more tissue-specific γ and δ isoforms, vary in their responsiveness according to tissue type and

developmental stage, but they generally respond to diverse types of cellular stress, cytokines and cell-cell contact but also mitogens and cellular development (ZARUBIN and HAN 2005; ASHWELL 2006). The broad responsiveness of p38 isoforms distinguishes them from the yeast ortholog, Hog1, which responds primarily to osmotic and related stresses (SAITO and POSAS 2012). Both JNK and p38 are activated by a number of upstream kinases, some that have overlapping affinity for both SAPKs and others with specificity for only one of the kinases (ROUX and BLENIS 2004). Both SAPKs in turn activate a slew of downstream kinases and a host of transcription factors that mediate many aspects of the stress response.

Among the most famous of the SAPK targets is the transcription factor p53, which lies at the crux of the growth/defense/apoptosis decision (BECKERMAN and PRIVES 2010). Recent transcriptomic experiments have identified at least 280 common p53 targets activated by diverse stresses and in multiple tissues, with perhaps hundreds of additional targets identified by studies looking at different tissue and stress types (MIRZA *et al.* 2003; WEI *et al.* 2006; MENENDEZ *et al.* 2009; HUARTE *et al.* 2010; NIKULENKOV *et al.* 2012; GAMBINO *et al.* 2013; SCHLERETH *et al.* 2013; LEVEILLE *et al.* 2015). Upon activation and translocation from the cytosol to nucleus, p53 can function both as an inducer and as a repressor (working directly and indirectly via its transcriptional targets); the genes regulated by p53 can vary in a given tissue depending on severity of the stress (see below)(VOUSDEN and LU 2002; MENENDEZ *et al.* 2009; LEVAV-COHEN *et al.* 2014). Commonly induced genes include those linked to stress defense, regulators of cell-cycle arrest, and pro- or anti-apoptotic factors (WEI *et al.* 2006; MENENDEZ *et al.* 2009; NIKULENKOV *et al.* 2012; SCHLERETH *et al.* 2013), while repressed genes include genes that promote cell-cycle progression and proliferation and genes inked to ribosome biogenesis and translation. In fact, p53 can repress activity of all three RNA polymerases, and in the case of Pol I and Pol III it does so by interfering with proper assembly of general transcriptional machinery (CAIRNS and WHITE 1998; BUDDE and GRUMMT 1999; ZHAI and COMAI 2000;

BECKERMAN and PRIVES 2010). Interestingly, the severity and duration of p53 activation are thought to influence whether cells arrest the cell cycle or sacrifice themselves through apoptosis (VOUSDEN and LU 2002). p53 binding elements upstream of pro-arrest and anti-apoptosis genes have higher affinity for p53 and lower dependence on cooperative binding, while promoters of pro-apoptosis factors are more likely to harbor degenerate p53 elements that require cooperative tetrameric association (SCHLERETH *et al.* 2013). This has led to the model that weak p53 activation (producing transient or incomplete nuclear localization) may promote cell survival while strong or prolonged activation directs cells toward death (BECKERMAN and PRIVES 2010).

Signaling crosstalk and the role of the growth-defense-death decision in human cancers. The correct balance between growth and stress defense or apoptosis is fundamental for proper organismal function. Improper balance leading to unchecked growth is thought to be a critical driver in diseases such as cancer (JONES and THOMPSON 2009). Positive regulators promoting RAS/ERK and mTORC signaling harbor gain-of-function mutations in many human cancers (SHAW and CANTLEY 2006; FERNÁNDEZ-MEDARDE and SANTOS 2011), underscoring the importance of these pathways in promoting growth. Furthermore, many cancerous cells show aberrantly high ribosome production and altered translation regulation (consistent with the oncogene status of eIF4e (MAMANE *et al.* 2004)). Thus, elevated flux toward growth and cell division is a driving force in cancer emergence.

But mutations in stress-activated regulators can also contribute to cancer. For example, 5-10% of diverse cancers harbor mutations in the SAPK-activating kinase MEK4, which is also associated with poor prognosis in several cancer types (TAYLOR *et al.* 2008), while a striking 50% of human cancers have inactivating alleles of p53 (VOUSDEN and LU 2002). Several stress-activated signaling pathways play dual roles in suppressing cancer, by triggering

apoptosis and by inhibiting signaling through the growth-promoting pathways. Thus, their mutation both prevents cells death and results in unchecked proliferation signaling. For example, p53 suppresses tumorogenesis not only by activating arrest or apoptosis, but also by suppressing mTORC and ERK via up-regulating expression of their inhibitors, at least in certain tissues (MATTHEW *et al.* 2009; FENG and LEVINE 2010; HASTY *et al.* 2013; DROSTEN *et al.* 2014; AKENO *et al.* 2015). Another example of stress-dependent pathways inhibiting proliferation signaling is seen upon nutrient limitation in mouse embryonic fibroblasts, when activated AMPK and p38β suppress mTORC activity through at least three independent routes (ZHENG *et al.* 2011). Thus, it is clear that stress-activated pathways work at several levels to shift the balance away from proliferation and toward a productive stress response.

Stress-activated networks in yeast: a model for understanding design principles.

The design principles of stress-activated signal integration remain poorly understood in both yeast and higher mammals. One strategy to investigate signaling organization is through systems-biology approaches to infer signaling networks. Several recent studies have conducted computational inference of the stress-activated signaling networks, most commonly based on transcriptome data (FRIEDMAN 2004; SCHADT *et al.* 2005; GAT-VIKS and SHAMIR 2007; GITTER *et al.* 2013; WU *et al.* 2013). In our recent work, we integrated disparate high-throughput yeast datasets using an integer linear programming approach to infer the salt-activated signaling network (CHASMAN *et al.* 2014). The resulting network of ~400 proteins captured known and novel salt-activated pathways as well as key regulators in the growth-promoting TORC1, RAS, and cAMP/PKA pathways, which are suppressed upon salt treatment. It also uncovered previously unrecognized cross-connections between what are generally studied as discrete pathways. We defined pathways based on the literature and scored the number of cross-pathways connections between them. Among those with the greatest connections to other pathways were the TORC1, RAS, and cAMP/PKA pathways. The consequences of this

inter-pathway connectivity remain to be tested, but we hypothesize that it reflects an intricate level of control exerted by stress-activated pathways on growth-promoting signaling.

Many of the regulatory connections in the yeast salt-responsive signaling network are orthologous to known signaling connections in mammals. But even more striking is the link to disease: the salt-activated signaling network we inferred is significantly enriched (p=8e-4) for proteins whose human orthologous cause cancer when mutated in somatic tissues, according to the latest release of the COSMIC database (FORBES *et al.* 2015). Using a stringent method of identifying orthologs (DELUCA *et al.* 2006), we found 11 of 49 orthologs from the COSMIC database in the salt-activated network: nearly half were chromatin regulators or transcription factors (including the ortholog of Sfp1 which regulates RP genes), two were involved in RNA metabolism, and one was the Hog1 activator Pbs2 (which was assigned orthologous to the ERK-activating MKK2 kinase). Remarkably, the network is also enriched nearly 3.5-fold for orthologs of p53 interacting proteins captured in the Biogrid database (p =1e-5, (CHATR-ARYAMONTRI *et al.* 2015)) – even though budding yeast lacks a p53 ortholog. The group of yeast genes orthologous to p53 interactors is enriched for kinases (p = 1.5e-7) and cell-cycle regulators (p = 4e-5). These results suggest that the salt-activated signaling network in yeast shares key features with cancer-activated signaling in humans, suggesting that the networks represent modern-day renditions of an ancient signaling system.

Much remains to be dissected about how diverse signals are integrated into a single signaling network, and how cells set the balance between growth versus stress defense. It is in this light that yeast research can contribute fundamental insights. Yeast provides an excellent test bed for systems-biology approaches to this question, where detailed follow-up studies can test predictions from network science. Despite the differences in complexity between yeast and mammalian systems, we believe that continued exploration of stress-activated yeast signaling

networks in the context of disease-causing orthologs could provide a new perspective on such signaling decisions. Understanding how these fundamental decision are made and disseminated in cells will expand our understanding of stress biology and foster our eventually ability to modulate it.

Acknowledgements

We apologize to the authors of many important research studies that we were unable to cite due to space constraints. We thank M. MacGilvray for useful comments on the manuscript. This work was supported by NIH R01 GM083989 to A.P.G.

References

ABRAHAM, R. T., 2001 Cell cycle checkpoint signaling through the ATM and ATR kinases. Genes Dev. **15:** 2177-2196.

AKENO, N., A. L. MILLER, X. MA and K. A. WIKENHEISER-BROKAMP, 2015 p53 suppresses carcinoma progression by inhibiting mTOR pathway activation. Oncogene **34:** 589-599.

ALBERTYN, J., S. HOHMANN, J. M. THEVELEIN and B. A. PRIOR, 1994 GPD1, which encodes glycerol-3-phosphate dehydrogenase, is essential for growth under osmotic stress in Saccharomyces cerevisiae, and its expression is regulated by the high-osmolarity glycerol response pathway. Mol Cell Biol. **14:** 4135-4144.

ARAVA, Y., Y. WANG, J. D. STOREY, C. L. LIU, P. O. BROWN *et al.*, 2003 Genome-wide analysis of mRNA translation profiles in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A. **100:** 3889-3894.

ASHWELL, J. D., 2006 The many paths to p38 mitogen-activated protein kinase activation in the immune system. Nat Rev Immunol. **6:** 532-540.

BECK, T., and M. N. HALL, 1999 The TOR signalling pathway controls nuclear localization of nutrient-regulated transcription factors. Nature **402:** 689-692.

BECKERMAN, R., and C. PRIVES, 2010 Transcriptional regulation by p53. Cold Spring Harb Perspect Biol **2:** a000935.

BELLÍ, G., E. GARÍ, M. ALDEA and E. HERRERO, 2001 Osmotic stress causes a G1 cell cycle delay and downregulation of Cln3/Cdc28 activity in Saccharomyces cerevisiae. Mol Microbiol. **39:** 1022-1035.

BERRY, D. B., and A. P. GASCH, 2008 Stress-activated genomic expression changes serve a preparative role for impending stress in yeast. Mol Biol Cell. **19:** 4580-4587.

BERRY, D. B., Q. GUAN, J. HOSE, S. HAROON, M. GEBBIA *et al.*, 2011 Multiple means to the same end: the genetic basis of acquired stress resistance in yeast. PLoS Genet. **7:** e1002353.

BRAUER, M. J., C. HUTTENHOWER, E. M. AIROLDI, R. ROSENSTEIN, J. C. MATESE *et al.*, 2008 Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast. Mol Biol Cell **19:** 352-367.

BROACH, J. R., 2012 Nutritional control of growth and development in yeast. Genetics **192:** 73-105.

BUDDE, A., and I. GRUMMT, 1999 p53 represses ribosomal gene transcription. Oncogene **18:** 1119-1124.

CAIRNS, C. A., and R. J. WHITE, 1998 p53 is a general repressor of RNA polymerase III transcription. EMBO J **17:** 3112-3123.

CASTRILLO, J. I., L. A. ZEEF, D. C. HOYLE, N. ZHANG, A. HAYES *et al.*, 2007 Growth control of the eukaryote cell: a systems biology study in yeast. J Biol. **6:** 4.

CAUSTON, H. C., B. REN, S. S. KOH, C. T. HARBISON, E. KANIN *et al.*, 2001 Remodeling of yeast genome expression in response to environmental changes. Mol Biol Cell. **12:** 323-337.

CHASMAN, D., Y.-H. HO, D. B. BERRY, C. M. NEMEC, M. E. MACGILVRAY *et al.*, 2014 Pathway connectivity and signaling coordination in the yeast stress-activated signaling network. Mol Syst Biol. **19:** 759.

CHATR-ARYAMONTRI, A., B. J. BREITKREUTZ, R. OUGHTRED, L. BOUCHER, S. HEINICKE *et al.*, 2015 The BioGRID interaction database: 2015 update. Nucleic Acids Res **43:** D470-478.

CHAUVIN, C., V. KOKA, A. NOUSCHI, V. MIEULET, C. HOAREAU-AVEILLA *et al.*, 2014 Ribosomal protein S6 kinase activity controls the ribosome biogenesis transcriptional program. Oncogene **33:** 474-483.

CLAYPOOL, J. A., S. L. FRENCH, K. JOHZUKA, K. ELIASON, L. VU *et al.*, 2004 Tor pathway regulates Rrn3p-dependent recruitment of yeast RNA polymerase I to the promoter but does not participate in alteration of the number of active genes. Mol Biol Cell **15:** 946-956.

CLOTET, J., X. ESCOTE, M. A. ADROVER, G. YAAKOV, E. GARI *et al.*, 2006 Phosphorylation of Hsl1 by Hog1 leads to a G2 arrest essential for cell survival at high osmolarity. EMBO J **25:** 2338-2346.

CONRAD, M., J. SCHOTHORST, H. N. KANKIPATI, G. VAN ZEEBROECK, M. RUBIO-TEXEIRA *et al.*, 2014 Nutrient sensing and signaling in the yeast Saccharomyces cerevisiae. FEMS Microbiol Rev. **38:** 254-299.

COOK, K. E., and E. K. O'SHEA, 2012 Hog1 controls global reallocation of RNA Pol II upon osmotic shock in Saccharomyces cerevisiae. G3 (Bethesda) **2:** 1129-1136.

DELUCA, T. F., I. H. WU, J. PU, T. MONAGHAN, L. PESHKIN *et al.*, 2006 Roundup: a multi-genome repository of orthologs and evolutionary distances. Bioinformatics **22:** 2044-2046.

DROSTEN, M., E. Y. SUM, C. G. LECHUGA, L. SIMÓN-CARRASCO, H. K. JACOB *et al.*, 2014 Loss of p53 induces cell proliferation via Ras-independent activation of the Raf/Mek/Erk signaling pathway. Proc Natl Acad Sci U S A. **111:** 15155-15160.

ELLIOTT, B., and B. FUTCHER, 1993 Stress resistance of yeast cells is largely independent of cell cycle phase. Yeast **9:** 33-42.

ESCOTÉ, X., M. ZAPATER, J. CLOTET and F. POSAS, 2004 Hog1 mediates cell-cycle arrest in G1 phase by the dual targeting of Sic1. Nat Cell Biol. **6:** 997-1002.

FENG, Z., and A. J. LEVINE, 2010 The regulation of energy metabolism and the IGF-1/mTOR pathways by the p53 protein. Trends Cell Biol. **20:** 427-434.

FERNÁNDEZ-MEDARDE, A., and E. SANTOS, 2011 Ras in cancer and developmental diseases. Genes Cancer **2:** 344-358.

FORBES, S. A., D. BEARE, P. GUNASEKARAN, K. LEUNG, N. BINDAL *et al.*, 2015 COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res **43:** D805-811.

FRIEDMAN, N., 2004 Inferring cellular networks using probabilistic graphical models. Science **303:** 799-805.

GAMBINO, V., G. DE MICHELE, O. VENEZIA, P. MIGLIACCIO, V. DALL'OLIO *et al.*, 2013 Oxidative stress activates a specific p53 transcriptional response that regulates cellular senescence and aging. Aging Cell **12:** 435-445.

GASCH, A. P., 2002 The Environmental Stress Response: a Common Yeast Response to Environmental Stresses. Springer-Verlag: Heidelberg.

GASCH, A. P., P. T. SPELLMAN, C. M. KAO, O. CARMEL-HAREL, M. B. EISEN *et al.*, 2000 Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell. **11:** 4241-4257.

GAT-VIKS, I., and R. SHAMIR, 2007 Refinement and expansion of signaling pathways: the osmotic response network in yeast. Genome Res **17:** 358-367.

GINGRAS, A. C., S. G. KENNEDY, M. A. O'LEARY, N. SONENBERG and N. HAY, 1998 4E-BP1, a repressor of mRNA translation, is phosphorylated and inactivated by the Akt(PKB) signaling pathway. Genes Dev. **12:** 502-513.

GITTER, A., M. CARMI, N. BARKAI and Z. BAR-JOSEPH, 2013 Linking the signaling cascades and dynamic regulatory networks controlling stress responses. Genome Res **23:** 265-276.

GÖRNER, W., E. DURCHSCHLAG, M. T. MARTINEZ-PASTOR, F. ESTRUCH, G. AMMERER *et al.*, 1998 Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity. Genes Dev. **12:** 586-597.

HANNAN, K. M., E. SANIJ, N. HEIN, R. D. HANNAN and R. B. PEARSON, 2011 Signaling to the ribosome in cancer--It is more than just mTORC1. IUBMB Life. **63:** 79-85.

HASTY, P., Z. D. SHARP, T. J. CURIEL and J. CAMPISI, 2013 mTORC1 and p53: clash of the gods? Cell Cycle **12:** 20-25.

HEINISCH, J. J., A. LORBERG, H. P. SCHMITZ and J. J. JACOBY, 1999 The protein kinase C-mediated MAP kinase pathway involved in the maintenance of cellular integrity in Saccharomyces cerevisiae. Mol Microbiol. **32:** 671-680.

HUARTE, M., M. GUTTMAN, D. FELDSER, M. GARBER, M. J. KOZIOL *et al.*, 2010 A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell **142:** 409-419.

HUBER, A., S. L. FRENCH, H. TEKOTTE, S. YERLIKAYA, M. STAHL *et al.*, 2011 Sch9 regulates ribosome biogenesis via Stb3, Dot6 and Tod6 and the histone deacetylase complex RPD3L. EMBO J. **30:** 3052-3064.

JONES, R. G., and C. B. THOMPSON, 2009 Tumor suppressors and cell metabolism: a recipe for cancer growth. Genes Dev. **23:** 537-548.

JORGENSEN, P., J. L. NISHIKAWA, B. J. BREITKREUTZ and M. TYERS, 2002 Systematic identification of pathways that couple cell growth and division in yeast. Science **297:** 395-400.

JORGENSEN, P., I. RUPES, J. R. SHAROM, L. SCHNEPER, J. R. BROACH *et al.*, 2004 A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. Genes Dev. **18:** 2491-2505.

KANTIDAKIS, T., B. A. RAMSBOTTOM, J. L. BIRCH, S. N. DOWDING and R. J. WHITE, 2010 mTOR associates with TFIIIC, is found at tRNA and 5S rRNA genes, and targets their repressor Maf1. Proc Natl Acad Sci U S A. **107:** 11823-11828.

KLEIN, C., and K. STRUHL, 1994 Protein kinase A mediates growth-regulated expression of yeast ribosomal protein genes by modulating RAP1 transcriptional activity. Mol Cell Biol **14:** 1920-1928.

KUSNADI, E. P., K. M. HANNAN, R. J. HICKS, R. D. HANNAN, R. B. PEARSON *et al.*, 2015 Regulation of rDNA transcription in response to growth factors, nutrients and energy. Gene **556:** 27-34.

LEE, J., R. D. MOIR and I. M. WILLIS, 2009 Regulation of RNA polymerase III transcription involves SCH9-dependent and SCH9-independent branches of the target of rapamycin (TOR) pathway. J Biol Chem. **8:** 12604-12608.

LEE, M. V., S. E. TOPPER, S. L. HUBLER, J. HOSE, C. D. WENGER *et al.*, 2011 A dynamic model of proteome changes reveals new roles for transcript alteration in yeast. Mol Syst Biol. **19:** 514.

LEVAV-COHEN, Y., Z. GOLDBERG, K. H. TAN, O. ALSHEICH-BARTOK, V. ZUCKERMAN *et al.*, 2014 The p53-Mdm2 loop: a critical juncture of stress response. Subcell Biochem. **85:** 161-186.

LEVEILLE, N., C. A. MELO, K. ROOIJERS, A. DIAZ-LAGARES, S. A. MELO *et al.*, 2015 Genome-wide profiling of p53-regulated enhancer RNAs uncovers a subset of enhancers controlled by a lncRNA. Nat Commun **6:** 6520.

LEVY, S. F., N. ZIV and M. L. SIEGAL, 2012 Bet hedging in yeast by heterogeneous, age-correlated expression of a stress protectant. PLoS Biol **10:** e1001325.

LI, H., C. K. TSANG, M. WATKINS, P. G. BERTRAM and X. F. ZHENG, 2006 Nutrient regulates Tor1 nuclear localization and association with rDNA promoter. Nature **442:** 1058-1061.

LIPPMAN, S. I., and J. R. BROACH, 2009 Protein kinase A and TORC1 activate genes for ribosomal biogenesis by inactivating repressors encoded by Dot6 and its homolog Tod6. Proc Natl Acad Sci U S A. **106:** 19928-19933.

LIU, B., and S. B. QIAN, 2014 Translational reprogramming in cellular stress response. Wiley Interdiscip Rev RNA. **5:** 301-315.

LOPEZ-MAURY, L., S. MARGUERAT and J. BAHLER, 2008 Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. Nat Rev Genet **9:** 583-593.

LU, C., M. J. BRAUER and D. BOTSTEIN, 2009 Slow growth induces heat-shock resistance in normal and respiratory-deficient yeast. Mol Biol Cell **20:** 891-903.

MAGNUSON, B., B. EKIM and D. C. FINGAR, 2012 Regulation and function of ribosomal protein S6 kinase (S6K) within mTOR signalling networks. Biochem J. **441:** 1-21.

MAMANE, Y., E. PETROULAKIS, L. RONG, K. YOSHIDA, L. W. LER *et al.*, 2004 eIF4E--from translation to transformation. Oncogene **23:** 3172-3179.

MARION, R. M., A. REGEV, E. SEGAL, Y. BARASH, D. KOLLER *et al.*, 2004 Sfp1 is a stress- and nutrient-sensitive regulator of ribosomal protein gene expression. Proc Natl Acad Sci U S A. **101:** 14315-14322.

MARTIN, D. E., A. SOULARD and M. N. HALL, 2004 TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1. Cell **119:** 969-979.

MATTHEW, E. M., L. S. HART, A. ASTRINIDIS, A. NAVARAJ, N. G. DOLLOFF *et al.*, 2009 The p53 target Plk2 interacts with TSC proteins impacting mTOR signaling, tumor growth and chemosensitivity under hypoxic conditions. Cell Cycle **8:** 4168-4175.

MAYER, C., and I. GRUMMT, 2006 Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases. Oncogene **25:** 6384-6391.

MAYER, C., J. ZHAO, X. YUAN and I. GRUMMT, 2004 mTOR-dependent activation of the transcription factor TIF-IA links rRNA synthesis to nutrient availability. Genes Dev. **18:** 423-434.

MENDOZA, M. C., E. E. ER and J. BLENIS, 2011 The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation. Trends Biochem Sci **36:** 320-328.

MENENDEZ, D., A. INGA and M. A. RESNICK, 2009 The expanding universe of p53 targets. Nat Rev Cancer. **9:** 724-737.

MEYUHAS, O., and A. DREAZEN, 2009 Ribosomal protein S6 kinase from TOP mRNAs to cell size. Prog Mol Biol Transl Sci. **90:** 109-153.

MIRZA, A., Q. WU, L. WANG, T. MCCLANAHAN, W. R. BISHOP *et al.*, 2003 Global transcriptional program of p53 target genes during the process of apoptosis and cell cycle progression. Oncogene **22:** 3645-3654.

MITCHELL, A., G. H. ROMANO, B. GROISMAN, A. YONA, E. DEKEL *et al.*, 2009 Adaptive prediction of environmental changes by microorganisms. Nature **460:** 220-224.

MOIR, R. D., J. LEE, R. A. HAEUSLER, N. DESAI, D. R. ENGELKE *et al.*, 2006 Protein kinase A regulates RNA polymerase III transcription through the nuclear localization of Maf1. Proc Natl Acad Sci U S A. **103:** 15044-15049.

MURRAY, J. I., M. L. WHITFIELD, N. D. TRINKLEIN, R. M. MYERS, P. O. BROWN *et al.*, 2004 Diverse and specific gene expression responses to stresses in cultured human cells. Mol Biol Cell **15:** 2361-2374.

NADAL-RIBELLES, M., N. CONDE, O. FLORES, J. GONZÁLEZ-VALLINAS, E. EYRAS *et al.*, 2012 Hog1 bypasses stress-mediated down-regulation of transcription by RNA polymerase II redistribution and chromatin remodeling. Genome Biol. **13:** R106.

NAGIEC, M. J., and H. G. DOHLMAN, 2012 Checkpoints in a yeast differentiation pathway coordinate signaling during hyperosmotic stress. PLoS Genet **8:** e1002437.

NAYAK, R. R., W. E. BERNAL, J. W. LEE, M. J. KEARNS and V. G. CHEUNG, 2014 Stress-induced changes in gene interactions in human cells. Nucleic Acids Res. **42:** 1757-1771.

NEUMAN-SILBERBERG, F. S., S. BHATTACHARYA and J. R. BROACH, 1995 Nutrient availability and the RAS/cyclic AMP pathway both induce expression of ribosomal protein genes in Saccharomyces cerevisiae but by different mechanisms. Mol Biol Cell **15:** 3187-3196.

NIKULENKOV, F., C. SPINNLER, H. LI, C. TONELLI, Y. SHI *et al.*, 2012 Insights into p53 transcriptional function via genome-wide chromatin occupancy and gene expression analysis. Cell Death Differ. **19:** 1992-2002.

O'ROURKE, S. M., and I. HERSKOWITZ, 2004 Unique and redundant roles for HOG MAPK pathway components as revealed by whole-genome expression analysis. Mol Biol Cell **15:** 532-542.

PHILIPPI, A., R. STEINBAUER, A. REITER, S. FATH, I. LEGER-SILVESTRE *et al.*, 2010 TOR-dependent reduction in the expression level of Rrn3p lowers the activity of the yeast RNA Pol I machinery, but does not account for the strong inhibition of rRNA production. Nucleic Acids Res. **38:** 5315-5326.

PROFT, M., and K. STRUHL, 2004 MAP kinase-mediated stress relief that precedes and regulates the timing of transcriptional induction. Cell **118:** 351-361.

RAMACHANDRAN, V., and P. K. HERMAN, 2011 Antagonistic Interactions Between the cAMP-Dependent Protein Kinase and Tor Signaling Pathways Modulate Cell Growth in Saccharomyces cerevisiae. Genetics **187:** 441-454.

REGENBERG, B., T. GROTKJAER, O. WINTHER, A. FAUSBØLL, M. AKESSON *et al.*, 2006 Growth-rate regulated genes have profound impact on interpretation of transcriptome profiling in Saccharomyces cerevisiae. Genome Biol. **7:** R107.

ROUX, P. P., and J. BLENIS, 2004 ERK and p38 MAPK-activated protein kinases: a family of protein kinases with diverse biological functions. Microbiol Mol Biol Rev **68:** 320-344.

ROUX, P. P., and I. TOPISIROVIC, 2012 Regulation of mRNA translation by signaling pathways. Cold Spring Harb Perspect Biol. **4:** a012252.

RUDRA, D., and J. R. WARNER, 2004 What better measure than ribosome synthesis? Genes Dev. **18:** 2431-2436.

SAITO, H., and F. POSAS, 2012 Response to hyperosmotic stress. Genetics **192:** 289-318.

SCHADT, E. E., J. LAMB, X. YANG, J. ZHU, S. EDWARDS *et al.*, 2005 An integrative genomics approach to infer causal associations between gene expression and disease. Nat Genet **37:** 710-717.

SCHLERETH, K., C. HEYL, A. M. KRAMPITZ, M. MERNBERGER, F. FINKERNAGEL *et al.*, 2013 Characterization of the p53 cistrome--DNA binding cooperativity dissects p53's tumor suppressor functions. PLoS Genet. **9:** e1003726.

SCHMELZLE, T., T. BECK, D. E. MARTIN and M. N. HALL, 2004 Activation of the RAS/cyclic AMP pathway suppresses a TOR deficiency in yeast. Mol Biol Cell **24:** 338-351.

SHAW, R. J., and L. C. CANTLEY, 2006 Ras, PI(3)K and mTOR signalling controls tumour cell growth. Nature **441:** 424-430.

SMITH, A., M. P. WARD and S. GARRETT, 1998 Yeast PKA represses Msn2p/Msn4p-dependent gene expression to regulate growth, stress response and glycogen accumulation. EMBO J. **17:** 3556-3564.

SOULARD, A., A. CREMONESI, S. MOES, F. SCHÜTZ, P. JENÖ *et al.*, 2010 The rapamycin-sensitive phosphoproteome reveals that TOR controls PKA toward some but not all substrates. Mol Biol Cell **21:** 3475-3486.

SPRIGGS, K. A., M. BUSHELL and A. E. WILLIS, 2010 Translational regulation of gene expression during conditions of cell stress. Mol Cell. **40:** 228-237.

TAYLOR, J. L., R. Z. SZMULEWITZ, T. LOTAN, J. HICKSON, D. V. GRIEND *et al.*, 2008 New paradigms for the function of JNKK1/MKK4 in controlling growth of disseminated cancer cells. Cancer Lett **272:** 12-22.

TEIGE, M., E. SCHEIKL, V. REISER, H. RUIS and G. AMMERER, 2001 Rck2, a member of the calmodulin-protein kinase family, links protein synthesis to high osmolarity MAP kinase signaling in budding yeast. Proc Natl Acad Sci U S A **98:** 5625-5630.

THOMAS, G., 2000 An encore for ribosome biogenesis in the control of cell proliferation. Nat Cell Biol **2:** E71-72.

UESONO, Y., and A. TOH-E, 2002 Transient inhibition of translation initiation by osmotic stress. J Biol Chem. **277:** 13848-13855.

VOGEL, C., G. M. SILVA and E. M. MARCOTTE, 2011 Protein expression regulation under oxidative stress. Mol Cell Proteomics. **10:** M111.009217.

VON DER HAAR, T., 2008 A quantitative estimation of the global translational activity in logarithmically growing yeast cells. BMC Syst Biol. **2**.

VOUSDEN, K. H., and X. LU, 2002 Live or let die: the cell's response to p53. Nat Rev Cancer **2:** 594-604.

WANG, Y., M. PIERCE, L. SCHNEPER, C. G. GÜLDAL, X. ZHANG *et al.*, 2004 Ras and Gpa2 mediate one branch of a redundant glucose signaling pathway in yeast. PLoS Biol. **2:** E128.

WARNER, J. R., 1999 The economics of ribosome biosynthesis in yeast. Trends Biochem Sci **24:** 437-440.

WEI, C. L., Q. WU, V. B. VEGA, K. P. CHIU, P. NG *et al.*, 2006 A global map of p53 transcription-factor binding sites in the human genome. Cell **124:** 207-219.

WESTFALL, P. J., J. C. PATTERSON, R. E. CHEN and J. THORNER, 2008 Stress resistance and signal fidelity independent of nuclear MAPK function. Proc Natl Acad Sci U S A. **105:** 12212-12217.

WILLIS, I. M., and R. D. MOIR, 2007 Integration of nutritional and stress signaling pathways by Maf1. Trends Biochem Sci. **32:** 51-53.

WU, C., N. YOSEF, T. THALHAMER, C. ZHU, S. XIAO *et al.*, 2013 Induction of pathogenic TH17 cells by inducible salt-sensing kinase SGK1. Nature **496:** 513-517.

XIAO, L., and A. GROVE, 2009 Coordination of Ribosomal Protein and Ribosomal RNA Gene Expression in Response to TOR Signaling. Curr Genomics **10:** 198-205.

YAAKOV, G., A. DUCH, M. GARCÍA-RUBIO, J. CLOTET, J. JIMENEZ *et al.*, 2009 The stress-activated protein kinase Hog1 mediates S phase delay in response to osmostress. Mol Biol Cell **20:** 3572-3582.

ZAKRZEWSKA, A., G. VAN EIKENHORST, J. E. BURGGRAAFF, D. J. VIS, H. HOEFSLOOT *et al.*, 2011 Genome-wide analysis of yeast stress survival and tolerance acquisition to analyze the central trade-off between growth rate and cellular robustness. Mol Biol Cell **22:** 4435-4446.

ZARUBIN, T., and J. HAN, 2005 Activation and signaling of the p38 MAP kinase pathway. Cell Res. **15:** 11-18.

ZHAI, W., and L. COMAI, 2000 Repression of RNA polymerase I transcription by the tumor suppressor p53. Mol Cell Biol **20:** 5930-5938.

ZHENG, M., Y. H. WANG, X. N. WU, S. Q. WU, B. J. LU *et al.*, 2011 Inactivation of Rheb by PRAK-mediated phosphorylation is essential for energy-depletion-induced suppression of mTORC1. Nat Cell Biol **13:** 263-272.

ZURITA-MARTINEZ, S. A., and M. E. CARDENAS, 2005 Tor and Cyclic AMP-Protein Kinase A: Two Parallel Pathways Regulating Expression of Genes Required for Cell Growth. Eukaryot Cell **4:** 63-71.