Cellular Resilience and Fragility in Response to Environmental and Gene Expression Perturbations

Lars-Göran Ottosson

UNIVERSITY OF GOTHENBURG

AKADEMISK AVHANDLING

Akademisk avhandling för filosofie doktorsexamen i Naturvetenskap, inriktning biologi, som med tillstånd från Naturvetenskapliga fakulteten kommer att offentligt försvaras fredagen den 21 september 2012 kl. 10.00 i föreläsningssal K2320, Carl Kylberg, Institutionen för kemi och molekylärbiologi, Medicinaregatan 9, Göteborg.

Göteborg 2012

Cellular Resilience and Fragility in Response to Environmental and Gene Expression Perturbations

And Now for Something Completely Different

(Monty Python's Flying Circus)

Abstract

Cells are constantly subjected to perturbations. Whether these are extracellular or intracellular, they can be detrimental to cellular fitness. The cell has evolved elaborate systems and mechanisms that allow it to remain functional in the face of disturbances.

Cellular signal transduction can be summarised as the processes by which environmental stimuli is integrated with information on cellular status through the transmission of intracellular signals. This information is carried by specific proteins that operate jointly in signalling networks, or pathways. An important output of these pathways is to establish cellular responses to perturbations. To remain functional the signalling network must be robust to fluctuations in both environmental stimuli and levels of signalling components.

In this thesis it is investigated to what extent cellular fitness is affected by gene overexpression of signalling components. A high degree of fragility to increases in gene dosage was observed. This stands in stark contrast to overall system resilience to deletions of the same components. Fragile nodes were also dispersed over different classes of signalling components as well as throughout the signalling networks. The observed fragility patterns were further demonstrated to be largely independent of environmental and genotypic fluctuations suggesting fragility to be a product of local network architecture.

Cellular responses to the rare but toxic metalloid tellurite, in terms of gene-by-environment interactions, are also investigated. To genetically elucidate mechanisms of sensitivity and resistance to this compound a genome-wide collection of gene deletion mutants was screened in presence of tellurite. A metabolic pathway, the sulfate assimilation pathway, was found to be central to tellurite toxicity. Chemically related compounds were also shown to share a common toxicity mechanism.

Quantitative biology is central to this thesis and high-throughput high-resolution measurement regimes for microbial growth have been applied to all studies included herein. Phenomics is introduced and the different types of phenotyping strategies applied to studies in this thesis are elaborated on.

Keywords: *Saccharomyces cerevisiae*, phenomics, liquid microcultivation, tellurite, gene overexpression, gToW, cellular signalling, HOG, protein phosphatase

Abbreviations

ATP	Adenosine triphosphate
DNA	Deoxyribonucleic acid
DSP	Dual specificity phosphatase
GI	Genetic interaction
gToW	Genetic tug-of-war
HOG	High Osmolarity Glycerol
LPI	Logarithmic phenotypic index
LSC	Logarithmic strain coefficient
MAPK	Mitogen-activated protein kinase
NaCl	Sodium chloride
OD	Optical density
ORF	Open reading frame
PK	Protein kinase
PP	Protein phosphatase
PQ	Paraquat
PTM	Posttranslational modification
PTP	Protein tyrosine phosphatase
SAPK	Stress-activated protein kinase
SGA	Synthetic genetic array
TF	Transcription factor

List of papers

This thesis is based on the following papers, referred to in the text by their Roman numerals;

I. **Ottosson LG**, Logg K, Ibstedt S, Sunnerhagen P, Käll M, Blomberg A and Warringer J (2010). Sulfate assimilation mediates tellurite reduction and toxicity in *Saccharomyces cerevisiae*. *Eukaryotic Cell* Oct;9(10):1635-47

II. Krantz M, Ahmadpour D[*], **Ottosson LG**[*], Warringer J, Waltermann C, Nordlander B, Klipp E, Blomberg A, Hohmann S and Kitano H (2009). Robustness and fragility in the yeast high osmolarity glycerol (HOG) signal-transduction pathway. *Molecular Systems Biology* 5:281
[*] These authors contributed equally

III. Waltermann C, Warringer J, **Ottosson LG**, Ahmadpour D, Zackrisson M, Kitano H, Hohmann S, Klipp E, Krantz M and Blomberg A. Characterisation of the robustness landscape of yeast signal transduction.
Manuscript

IV. **Ottosson LG**, Kvarnström M, Kitano H, Hohmann S, Krantz M, Warringer J and Blomberg A. Overexpression of catalytic subunits of protein phosphatases in *S. cerevisiae* reveals new functional connections in the signalling network.
Manuscript

Contribution report

I. Planned and performed most experiments except for experiments based on microscopy and raman measurements. Took part in analysis of data and creation of figures. Read and commented the final draft of the manuscript.

II. Planned and performed characterisation of gToW strains through growth assays. Took part in analysis of data and creation of figures. Read and commented the final draft of the manuscript.

III. Took part in planning and characterization of gToW strains through growth assays. Took part in analysis of data and creation of figures. Read and commented the final draft of the manuscript.

IV. Main author. Planned and performed all experimental work except for image quantification. Took part in analysis of data. Created all figures.

Table of Contents

1 AIMS OF THIS THESIS

As a graduate student being part of a platform in quantitative biology at the Faculty of Science, University of Gothenburg, my research has primarily involved quantitatively measureable aspects of cellular resilience and fragility in response to environmental and/or internal perturbations. My focus has been on the study of system fragility towards gene overexpression of signalling components in the model eukaryotic organism *Saccharomyces cerevisiae* using a method for gene overexpression. Besides from this, I have also had an interest in studying cellular responses to a toxic metalloid by the determination of genetic factors attributed to those responses. Throughout this thesis I have made use of high-resolution high-throughput growth phenotyping methods as tools for scientific discovery. In short, the main aims of this thesis have been:

➢ To characterize the resilience and fragility landscape of the cell in response to gene overexpression of signalling components.

➢ To elucidate cellular mechanisms of sensitivity and resistance to the toxic metalloid tellurite.

1.1 Thesis outline

In the first part of this thesis the concept of functional genomics and systems biology will be introduced followed by a more detailed description of the field of phenomics. These approaches to scientific discovery have been used extensively throughout this thesis. Biological perturbations (genetic and environmental) have been central to this thesis and are described in the middle part. In the last part of this thesis cellular signalling is described and the effects on cellular fitness upon overexpression of signalling components elaborated on.

2 INTRODUCTION

The cell constitutes the smallest living entity, whether being part of a multicellular organism or leading a solitary life as a unicellular microorganism. It constantly monitors and interacts with its surrounding, integrating environmental stimuli with internal cellular status. In favourable environmental conditions microorganisms grow both in size (cell growth) and numbers through cell division (proliferation). On the level of a population, the ability of the organism to both survive and reproduce in a particular environment is called its fitness. Throughout this thesis, fitness will be considered equivalent to, in particular, growth rate on the level of the microbial cell population.

Microorganisms are surrounded by, and subjected to, an environment that rapidly can fluctuate in many of its parameters, e.g. osmolarity, exposure to toxic metals or oxidizing agents, pH and temperature, to name a few. Many of these small organisms are also immobile and lack the ability to escape a harsh environment, should the need arise. To cope with this, microorganisms have evolved the means to adequately respond and adapt to rapidly fluctuating environmental parameters. In **Paper I** cellular responses to exposure of the toxic metalloid tellurite is studied.

The typical response is mounted at the plasma membrane interface between the extracellular and intracellular milieu where specific receptor molecules sense environmental stimuli and triggers an intracellular response. These responses are transmitted throughout the cell by special signal transduction molecules, often arranged in signalling cascades, or pathways, forming a dense signalling network. The downstream output of these pathways in turn affects many cellular processes, e.g. cell cycle and metabolism that constantly remodel the cell machinery to sustain life in the face of changes in the external milieu.

Extracellular stimuli, as well as abundance of signal transduction components, fluctuate stochastically over time. To cope with this, cellular signalling networks have also evolved intrinsic robust properties that allow them to maintain functionality despite a certain level of environmental or mutational perturbations. In **Papers II-IV** effects on cellular fitness caused by gene overexpression of signal transduction components are studied.

The function of a large fraction of cellular components, e.g. genes and proteins, has been painstakingly elucidated over the last decades. However, roughly 25% of predicted genes in the eukaryotic model organism *S. cerevisiae* remain either uncharacterized or dubious. In large part this is due to a lack of scale and sensitivity of experimental procedures aimed at functional discovery. The picture is slowly changing today, however, as new techniques for large-scale high-throughput phenotyping of quantifiable cellular characteristics are emerging, allowing scientists to systematically study gene networks on a genome-wide scale in response to environmental or genetic perturbations.

Throughout this thesis high-throughput and/or high-resolution growth phenotyping strategies has been used to address the specific biological questions raised. In **Paper I** a high-throughput approach based on high-density arrays of yeast colonies on agar plates was used while a liquid microcultivation approach for generation of high-resolution growth profiles was better suited for the studies conducted in **Papers II-IV**.

A common denominator throughout this body of work has been the use of the model organism *Saccharomyces cerevisiae* as a representative of eukaryotic life.

2.1 *Saccharomyces cerevisiae*

S. cerevisiae or budding yeast to the scientist, brewer's yeast to the beer brewer and baker's yeast to the baker, this versatile unicellular fungus is known by many names. Despite its small size (only 5-10 micrometers across) and at first seemingly insignificant appearance, this eukaryotic organism has been a humble servant to humankind for thousands of years (Scherens and Goffeau 2004). This is in large part due to a quite unusual property; it prefers fermentation to respiration. In this metabolic process sugar is anaerobically converted to carbon dioxide and ethanol. This phenomenon has been utilized by man in the baking of bread and fermentation of alcoholic beverages. Through sequencing of ribosomal DNA, presence of *S. cerevisiae* has been detected in wine jars dating at least from 3150 B.C. (Cavalieri et al. 2003).

Yeasts exist in many different terrestrial environments, but have also been found in marine settings (Kutty and Philip 2008). Natural habitats for *S. cerevisiae* are the surfaces of fruit (Mortimer and Johnston 1986) or the bark of trees, mainly broad-leaved species such as oak (Sniegowski et al. 2002).

Besides from its household advantages, a number of additional features have made *S. cerevisiae* a tractable organism to study in molecular biology, so much so that this yeast has become a well-established model organism for the study of basic eukaryotic cell physiology (Goffeau 2000). Being a eukaryotic organism, *S. cerevisiae* shares many basic cellular features with higher multicellular organisms in terms of subcellular organelles, chromosome organisation, posttranslational modifications and signal transduction systems, to name a few. It is easily cultivated in the lab, grows rapidly with a short generation time (1.5 – 2 h in basal growth medium), exists in both a haploid and a diploid state and has a well-defined genetic system with a high degree of homologous recombination simplifying genetic manipulation in an unprecedented way (Sherman 2002). In addition, *S. cerevisiae* was the first eukaryote to have its genome completely sequenced in 1996 as the result of a worldwide collaboration (Goffeau et al. 1996; Williams 1996) and today several genome-wide mutant collections for the roughly 6000 genes exist. In **Paper I** a gene deletion mutant collection was used to search for genetic determinants of sensitivity and resistance to a particular metalloid derivative, tellurite. Systematic collections of mutants of this kind have been, and remain, pivotal in the struggle of scientists to describe the inner workings of the cell and how all the parts of the system work together in an orchestrated way to sustain life (Forsburg 2001).

Throughout this thesis, *S. cerevisiae* and yeast will be used interchangeably, unless stated otherwise.

2.2 Functional genomics and systems biology

The discovery of the structure of DNA in the 1950s (Watson and Crick 1953) and recombinant DNA techniques in the 1970s (Cohen et al. 1973) made it possible to genetically engineer the genomes (DNA sequence) of organisms. New and improved techniques for DNA sequencing, first manual (Maxam and Gilbert 1977; Sanger et al. 1977) and later made automatic, with the aid of sequencers (Smith et al. 1986), paved the way for whole genome sequencing of organisms such as *S. cerevisiae* (Goffeau et al. 1996), the bacterium *Escherichia coli* (Blattner et al. 1997), the fruit fly *Drosophila melanogaster* (Adams et al. 2000) and human (Lander et al. 2001; Venter et al. 2001). The bulk of sequence data gave rise to the fields of genomics and functional genomics that are concerned with the study of genome sequences and the functions of the genes

encoded within the genome. In **Papers II-IV** the available and annotated genome sequence of *S. cerevisiae* was used as source for construction of gene overexpression strains. Parallel advances in related fields such as transcriptomics (analysis of all RNA molecules) and proteomics (analysis of structure and function of all proteins) helped identify most of the components in the cell, although it should be stressed that new components are still being discovered.

While these approaches dominated research on molecular biology during the last half of the 20th century and were successful in establishing cellular "parts lists", and to some extent the function of individual components, they did not provide information on system properties, e.g. organisation, dynamic interactions and robustness of cellular networks (Strange 2005; Sauer et al. 2007). Over the last decade, it has become increasingly clear that the complexity of organisms is larger than the sum of their parts (Strange 2005). From this insight a new approach to scientific discovery emerged: systems biology.

Systems biology is aimed at investigating systems behaviour, or phenotypes, with regards to all components and their interactions both within and outside (environmental interactions) the system while in operation. This approach draws its energy from a multitude of complex high-throughput datasets on components and their interactions, and makes use of mathematical modelling to describe and simulate the dynamic properties of cellular networks, e.g. signalling networks (Gilbert et al. 2006). In **Paper II** this approach was used to incorporate phenotypic data on gene overexpression with an existing model of a signal transduction pathway, and a refinement to the model was proposed in which protein complex formation contributes to observed overexpression fragility (**Paper II, Fig. 2**). Over the last decade, systems biology has become a widely used approach in biological research and is now considered a norm in many instances (Kohl et al. 2010).

Throughout this thesis, functional genomics and systems biology have been core approaches in the study of cellular responses, mainly on a systems level represented by cellular fitness, towards gene perturbations and/or environmental variation.

2.3 Modularity in biology

The biology of the cell is incomprehensibly complex, with thousands of components working in an orchestrated fashion to sustain life. For the human mind to be able to penetrate this complexity and make sense out of a plethora of information, science has a long history of trying to reduce complexity through modularisation. This approach has also been fruitful as modularity indeed appears to be an integral part of the topology of cellular networks.

Many cellular functions are carried out by semi-autonomous "modules" (Hartwell et al. 1999). Signalling and metabolic pathways can be considered such functional modules. Besides from functional modules, also physical modules, e.g. the cell and its different organelles exist. This way of thinking has also been useful in designing new experiments to understand the cell machinery as it gives scientists a base from which to attack problems. Modularity is also considered one of the cornerstones of biological robustness analysis as these modules to some extent can encapsulate processes and functions, thereby isolating damaging perturbations from the rest of the system (Kitano 2004; Stelling et al. 2004). Although the notion of modularity is useful in research, it should be noted that no component or module is an isolated island in the cell. It is clear from genetic and biochemical data that components are interconnected to a high degree and often in intricate dynamic relationships. This can be exemplified by the yeast response to high osmolarity which activates components in a signal transduction pathway, the high osmolarity glycerol (HOG) pathway. Although this pathway can be modularised and defined to include a limited set of key components (in **Paper II** 29 genes are defined as core components of the HOG pathway) the complete picture of the cellular response to high osmolarity is more complex. In a genome-wide experiment performed on gene deletion mutants cultivated in presence of NaCl (known to induce osmostress), several hundreds of other genes were demonstrated to be of importance to cellular fitness in the response to hyperosmolarity (Warringer et al. 2003). Besides from cellular defence mechanisms, genes involved in cellular transport, cytoskeleton organization and metabolism contributed to responses to hyperosmolarity (Warringer et al. 2003).

The concept of modularity has also been of importance to the studies conducted in this thesis. In **Paper I** a search for genetic determinants of traits associated with sensitivity and resistance to the metalloid tellurite resulted in the finding that a metabolic

pathway, the sulfate assimilation pathway, is at the core of these traits. In **Paper III** the study was expanded to encompass other signalling pathways, each of which also can be considered functional modules assigned to different aspects of the cellular machinery. In **Paper IV** protein phosphatases were instead studied as a class of enzymes. Although they together cannot be considered a functional module, given their many roles in the cell, they nevertheless share a common catalytic mechanism around which they can be classified.

3 PHENOMICS

In the wake of the large sequencing projects of the late 90's and early 2000's scientists were able to identify most of the genes responsible for producing the biological complexity of organisms. Parallel innovations in DNA technology also made it possible to genetically manipulate these with relative ease. These scientific breakthroughs made it possible to approach the long-term goal in biology of understanding the relationship between the genotype, environmental variation and the phenotype of an organism. The phenotype emerges as the result of an interaction between the genotype and environmental factors:

Genotype + Environmental factors → Phenotype

A phenotype is best described as any observable trait or characteristic of an organism, where observable refers to being either visually distinguishable or detectable by means of any technical procedure in the laboratory. It is found on all levels of biological organisation, e.g. changes in gene expression profiles in response to stressful environmental factors (Gasch et al. 2000), protein expression (Ghaemmaghami et al. 2003), localisation (Huh et al. 2003) and phosphorylation (Ptacek et al. 2005), colony morphology (**Paper I**) (Granek and Magwene 2010) and the organisation of cell communities (Honigberg 2011), to name a few.

To be able to quantify the phenotypic effects of gene and environmental perturbations on a large-scale, phenomics emerged a little over a decade ago as an area of biology concerned with high-throughput quantitative measurements of phenotypes. Phenomics is interdisciplinary and combines biology, chemistry and physics with technological innovations in robotics, optics and image analysis. It is dependent on parallel development of automated, or semi-automated, high-throughput measurement systems for generation of high-resolution data required for scientific discovery. These systems are built on modules, each dedicated to a specific task, from the instrument performing the measurements to the computer and bioinformatics software analysing the results and providing the researcher with manageable datasets.

In our lab we use automated, or semi-automated, systems for phenotypic profiling of growth characteristics of microorganisms. The studies included in this thesis are mainly focused on microbial growth, and in particular growth rate, as a phenotypic representation of fitness effects. In **Paper I** and **IV** a system based on agar plates was used. In this approach yeast strains are automatically pinned as single colonies on an agar surface and growth and morphology phenotypes are measured after a defined time period. In **Papers II-IV** a microcultivation system was instead used to measure culture growth over time and automatically generate high-resolution growth curves.

3.1 Microbial growth

As microorganisms are introduced into an environment with ample nutrients, e.g. a batch culture with a fixed volume of growth medium in a controlled set of environmental conditions (**Papers II-IV**), the population grows according to a sigmoidal growth curve (Fig. 1B and C). The growth curve serves as an important empirical model describing population size over time and applies to most living matter. It consists of three distinct growth phases, the *lag* (or adaptation) phase during which the cells acclimatises to the environment before assuming growth, the *exponential* (or logarithmic) phase during which the population will double in size at a constant growth rate and the *stationary* phase at which growth rate slows down and levels out due to nutrient depletion and a build up of toxic metabolic by-products (Fig. 1C). The level of the stationary phase, in terms of population density increment, can also be called growth efficiency (or yield) and depends on the ability of the population to utilise available resources. During prolonged growth/cultivation, the curve also eventually enters a death phase during which cells die.

These growth phases can be considered variable as they are differently affected depending on the genetic and/or environmental perturbations being studied.
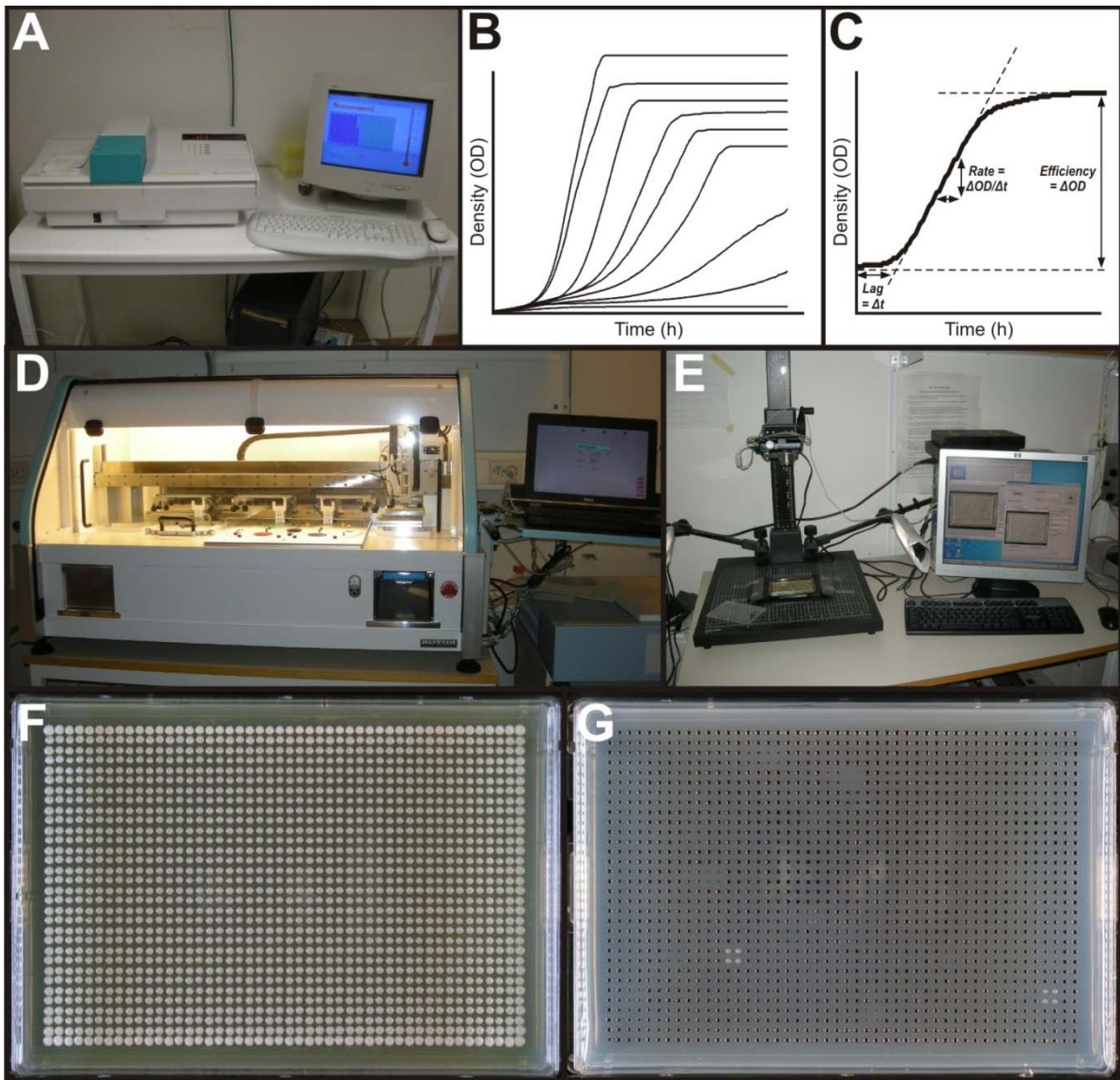
Figure 1. Strategies for microbial growth phenotyping. (A) Bioscreen C for liquid microcultivation. (B) Representative microbial growth curves as generated by frequent OD recordings by a Bioscreen C device. Each growth curve represents the growth of an individual yeast population. Presented are growth profiles with varying degree of effects on growth lag, growth rate and growth efficiency. (C) Extraction of the growth curve variables *lag*, *rate* and *efficiency*. For details on the growth variable extraction, see text and (Warringer and Blomberg 2003; Warringer et al. 2003). (D) RoToR HDA benchtop pinning robot for colony growth on agar plates. (E) Image capture rig for agar plates with colonies. Agar plates with yeast arrays are placed in a fixture and photographs taken using a digital camera. (F) Array with 1536 yeast colonies. (G) Array with 1536 yeast gene deletion mutants cultivated in presence of tellurite (**Paper I**). Most colonies have a dark colouration caused by intracellular accumulation of tellurium. Colonies with deviant colouration indicate mutants with reduced accumulation of tellurium.

3.2 Liquid microcultivation

Liquid microcultivation is a high-throughput high-resolution growth phenotyping strategy where microorganisms are cultivated in microscale, typically in microtiter plates (Blomberg 2011). Each well in the plate serves as a miniaturized reaction chamber filled with liquid growth medium. During cultivation, optical measurements are recorded for each well over time as an estimate of the increase in cell population density.

Measurements are made automatic by the use of robotic devices, which in many cases are combined incubators, shakers and readers that records optical density (OD) values from a number of parallel cultures that depends on the device (Blomberg 2011). In our lab we use a system named Bioscreen C (Oy Growth Curves Ab Ltd., Finland) that allows simultaneous measurements of 200 cultures (2x100 well plates) in one experimental run (Fig. 1A). This was the system used for the gene overexpression studies conducted in Papers **II-IV**.

Based on this system a procedure for automated microscale cultivation with subsequent data analysis and feature extraction was developed (Warringer and Blomberg 2003). In this setup, yeast strains to be tested are inoculated in a defined growth medium and cultivated under precise temperature control (typically 30°C for routine tests with yeast). OD values are recorded every 20 min for a time period of 72 h by default (these parameters can be user-defined) which results in high-resolution growth curves (Fig. 1B). From these curves, three variables are then automatically extracted; *growth lag* (response or adaptation time) in hours, *growth rate* (population doubling time during the exponential growth phase) in hours and *growth efficiency* (biomass gain as a result of available resources) as stationary phase OD increment (Fig. 1C). The frequent measurement regime together with automated data extraction makes this procedure sensitive, precise and reproducible.

It has been successfully applied in a genome-wide analysis of the, in many cases, marginal but functionally important growth phenotypes associated with gene deletions in *S. cerevisiae*, both in basal growth medium and as a result of environmental stress factors such as NaCl (Warringer et al. 2003). The procedure has also been applied in chemogenetic screens of mode-of-actions for bioactive compounds (Warringer et al. 2008), differences within and between species attributed to population genomics (Liti et

al. 2009) and in the study of essential genes through the use of conditional temperature-sensitive (ts) mutants (Li et al. 2011).

As this procedure generates large datasets, an in-house database named PROPHECY (PROfiling of PHEnotypic Characteristics in Yeast) was developed as a public resource for phenotypic data generated in the lab (Fernandez-Ricaud et al. 2005; Fernandez-Ricaud et al. 2007).

While precise details on the experimental setup can be found elsewhere (Warringer and Blomberg 2003; Warringer et al. 2003), a brief description of aspects important to this thesis will be given. Yeast strains to be tested are pre-cultivated in defined growth medium after which they are transferred to plates with growth medium depending on the experimental setup. Duplicate samples are typically placed on separate plates. As controls or references, wild-type strains are included in the experimental runs, the purpose of which are to be used for normalisation of strain behaviour by forming logarithmic ($\log_2$) strain coefficients (LSC) for each growth variable. In **Papers II-IV** a plasmid-based gene overexpression method was used. A strain carrying a plasmid without gene insert but otherwise identical sequence backbone was used as a control. For rate of growth, LSC is calculated as

$$LSC_{ij} = \frac{\sum_{r=1}^{2}\left[\left[\frac{1}{m}\sum_{k=1}^{m}\log_2\left(wt_{kj}^r\right)\right] - \log_2\left(x_{ij}^r\right)\right]}{2}$$

where wt_{kj} is the doubling time of the k:th measurement of the wt in environment j, x_{ij} is the doubling time of strain i in environment j, m is the number of included wt strains and r indicates the run. As these growth ratios are logarithmic, it implies that a tested strain that has a lower rate of growth compared to the wt has a negative LSC value. If the opposite holds true, the tested strain instead has a positive LSC value.

3.3 Colony growth on agar plates

An alternative strategy to liquid microcultivation is to cultivate microorganisms on solid nutrient surfaces, typically agar plates, and measure growth as an increase in colony area over time. With this strategy, microorganisms are pinned, or spotted, onto the agar surface in a grid/array format.

A multitude of technological innovations in recent years have made it possible to increase the density, or number of colonies, in the arrays. In our lab we use a system named RoToR HDA (Singer Instrument Co. Ltd., United Kingdom) for handling and pinning of high-density arrays of yeast (Fig. 1D), in the formats 96, 384, 786 and 1536 colonies per plate (Fig. 1F and **Paper I**). The base of 96 allows for an additional liquid step in which cells can be pinned from stock collections typically stored in ordinary 96-well microtiter plates.

Several genome-wide screens have been performed using this approach, e.g. synthetic genetic array (SGA) analysis (Tong et al. 2001; Tong et al. 2004; Costanzo et al. 2010), gene overexpression (Sopko et al. 2006) and tolerance to transition metals (Bleackley et al. 2011). While these studies have focused on colony growth as readout, other colony features such as morphology and colour can also be screened for.

In **Paper I** intracellular accumulation of the metalloid tellurium was used as a phenotypic readout. This accumulation results in a distinct cell darkening and visual inspection of gene deletion mutant colonies can be used to trace genetic determinants of this trait by quantifying colour deviation (Fig. 1G).

After a defined time-point (typically an end-point measurement), the arrays are captured visually using an imaging technique. For **Papers I and IV** an in-house image capture rig with a mounted digital camera was used (Fig. 1E). Colony features are then extracted either manually by visual inspection or automatically by using image analysis tools that perform colony segmentation and area quantification (Lamprecht et al. 2007; Memarian et al. 2007; Dittmar et al. 2010; Lawless et al. 2010). In **Paper IV** a tool was developed for automatic extraction of colony sizes. While available software tools mainly are aimed at quantification of colony size, a phenotype such as colour is more difficult to automatically quantify. In **Paper I** colonies were manually inspected and their areas and coloration quantified (Fig. 1G).

3.4 Liquid versus solid phenotyping - pros and cons

Although the different phenotyping strategies both measure microbial growth, there are pros and cons to them. A clear benefit with microcultivation is the frequent measurement regime and automated data extraction that makes this procedure sensitive, precise and easy to use. To obtain high-resolution data on growth effects caused by gene overexpression, this method was used for the studies conducted in **Papers II-IV**. In these studies reasonably small sets of strains were also studied with no set containing more than 200 strains, which corresponds to the maximum number of samples that can be measured simultaneously in one experimental run using the Bioscreen setup. Microcultivation therefore is not truly a high-throughput technology. Another issue with microcultivation is that cultivation in presence of certain compounds, e.g. tellurite and selenite (**Paper I**) that cause intracellular metalloid accumulation, can result in an increase in growth medium turbidity with an impact on the OD measurements (**Paper I, Fig. 1A**).

While not having the same degree of sensitivity and precision, cultivation of yeast colonies on solid growth media instead relies on scalability. In **Paper I**, an entire yeast gene deletion mutant collection was screened on 14 plates in the 1536 colony format, with each mutant being represented in quadruplicates. Compared with liquid microcultivation where a total of 200 samples can be measured in each experiment, a single agar plate has the potential of more than a 7-fold increase in sample size, truly making this a high-throughput strategy. Many scientists also argue that solid cultivation, to some extent, mimics the way many of the commonly studied yeast species exist in nature, living on different surfaces. An issue with cultivation on solid growth media in generating growth curves for yeast strains is that measurements of colony growth are typically not time-resolved. Instead an end-point measurement, or at best a few measurements at intervals of several hours, summarises the total growth behaviour of the colony that will be a composite of lag, rate and efficiency.

4 BIOLOGICAL PERTURBATIONS

A biological perturbation can be seen as any alteration or disturbance to the normal function of a biological system, whether it is a macroscopic ecosystem or a microscopic cell. These perturbations are caused by internal and/or external factors. For a cell, internal perturbations are frequently caused by mutations to the genetic material. Although these can be detrimental to cellular fitness, they are also essential as they create the genetic variation necessary for evolution. The ability to experimentally create specific mutations, and with relative ease for yeast, also gives scientists important tools in the study of gene functions. The cell should not be regarded as an isolated entity but rather part of a greater context in which it interacts with the environment, e.g. gene-by-environment interactions.

4.1 Gene perturbations

The genome contains the blueprint for life, in the form of thousands of genes. To gain an understanding of life requires an understanding of the functions of individual genes. A standard strategy among scientists in approaching this goal is to study the phenotypic consequences of gene perturbations, either under standard laboratory conditions or in combination with environmental perturbations.

In this context, a gene perturbation should be regarded as a manipulation of the DNA sequence for that gene that will result in either a gene product with altered function or a changed gene expression. Besides being of importance in gaining an understanding of gene function in organisms, mutations that affect gene expression are of a general interest as they contribute to the emergence and progression of many diseased states in humans. Gene deletions, rearrangements, translocations and point mutations have been associated to numerous human diseases, and gene overexpression or amplification has also been linked to the progression of pathological disorders, e.g. diabetes and cancer (Shastry 1995; Santarius et al. 2010). Overexpression of protein phosphatases that are involved in intracellular signal transduction have been demonstrated to occur in human cancers, e.g. breast and prostate cancer (Lee et al. 2000; Ray and Kiyokawa 2008).

To study genetic interactions different approaches have also been established to combine perturbations of different genes in a single cell. This strategy is particularly well suited for the elucidation of cellular networks of functionally related genes. As previously described, a phenotype emerges as the result of an interaction between the genotype and environmental factors. For that reason efforts are also being devoted to elucidate gene-by-environment interactions. For *S. cerevisiae* many genome-wide collections of mutants where each gene is perturbed in some way have been constructed during the last decade, and these are used routinely in labs worldwide in genome-wide functional genomics studies.

Gene perturbations that either abolish gene expression or cause gene overexpression have been exploited in this thesis. In **Paper I** a search for genes involved in sensitivity and resistance to the metalloid tellurite was performed by screening a genome-wide collection of gene deletion mutants, and in **Papers II-IV** a plasmid-based gene overexpression methodology was used to study the effects on cellular fitness caused by overexpression of signalling components.

4.1.1 Gene loss-of-function

Gene loss-of-function mutations can be of two types, full gene deletions (from start to stop of the coding region of a gene) that cause a complete loss of function mutation by abolished gene expression, or partial gene deletions that cause a partial loss of gene function. In **Papers II-IV** a defective metabolic gene allele, *leu2-d*, that has a large promoter truncation causing a weakened gene expression relative to wild-type *LEU2*, was used to drive an increase in plasmid copy number (Fig. 2A).

In an effort to study the consequences of gene deletions on a genome-wide scale, in terms of fitness effects, nearly every predicted gene was systematically deleted in *S. cerevisiae* (Winzeler et al. 1999; Giaever et al. 2002). The authors found that ~80% of genes are dispensable for growth under standard laboratory conditions. A later study, however, demonstrated a measurable growth phenotype for ~97% of the non-essential gene deletion mutants in at least one of several environmental conditions, stressing the high degree of interactions between the genotype and environmental factors (Hillenmeyer et al. 2008). The high fraction of dispensable genes can also to some extent be explained by functional redundancy as ~25% of genes in the yeast genome are considered duplicates, based on protein sequence alignments (Gu et al. 2003). It is now widely

believed that these duplications are the result of an ancient whole genome duplication event (Wolfe and Shields 1997; Kellis et al. 2004). It should be stressed, however, that many of the duplicated genes have differing functions and hence cannot compensate the function of each other.

Among the ~20% of genes essential to cell viability, genes involved in protein synthesis, mitochondrial function and respiration are overrepresented, demonstrating the need for unperturbed basic cellular processes (Giaever et al. 2002). It has also been demonstrated that the essential genes are overrepresented among subunits of multi-subunit protein complexes (Semple et al. 2008). This observation was further supported by the finding that a large fraction (>30%) of essential genes tend to associate with each other in large and essential protein complexes containing more protein constituents than non-essential complexes (Wang et al. 2009). A hypothesis that helps explain the strong loss-of-function phenotypes observed for essential genes is that decreases in the levels of certain subunits render the complexes stoichiometrically unstable and eventually destroy their function (Semple et al. 2008).

To study the consequences of gene loss-of-function mutations, several genome-wide mutant collections of yeast are now available to the scientific community. In a worldwide collaboration, a collection of gene deletion mutants in four different versions was constructed.

Two of these are haploid and constructed for both mating types (MATa and MATα), and two are diploid with one heterozygous and one homozygous collection (Winzeler et al. 1999; Giaever et al. 2002). A haploid gene deletion collection was used extensively in the studies conducted in this thesis (**Papers I-IV**). Gene loss-of-function mutants for many of the essential genes also exist, although these are more difficult to construct. This has been circumvented by the introduction of alleles with conditionally controllable gene expression, e.g. tetracycline (tet)-regulatable promoter swapping where expression is controlled by addition of the tetracycline derivative doxycycline to the growth medium (Mnaimneh et al. 2004). An alternative approach is the use of temperature-sensitive (ts) alleles that offer a finely tuned control over the corresponding protein's activity at different temperatures (Li et al. 2011).

4.1.2 Gene gain-of-function

Gene gain-of-function mutations either change the gene product in such a way that it gains new or abnormal functions, or enhance the expression of a gene (hypermorphic mutation). The latter can also be called gene overexpression and implies an increase in gene expression levels above normal. Gene overexpression is frequently caused by gene hyperactivation or amplification and is of a medical interest as it is prevalent in many human diseases as previously described.

In yeast, the use of gene overexpression as an alternative approach to gene deletion in understanding the function of genes has a long tradition (Burke et al. 1989). It can be accomplished experimentally by fusing genes (open reading frames) to endogenous and more powerful promoters (Rine 1991), although exogenous promoters also can be used as in the Tet-titratable promoter system (Mnaimneh et al. 2004). These promoters are either constitutive and drives the expression of the gene under most conditions, or inducible and drives the expression in response to a particular environmental change, e.g. various carbon sources. The promoter-gene fusions are either integrated into the genome, or inserted into a plasmid vector. In the latter case these vectors are often derived from the endogenous 2µ (micron) plasmid. These 2µ plasmids are maintained at 10-30 copies per cell (Rose and Broach 1990) and also propagate stably in cell populations (Ghosh et al. 2006). In that way they also provide additional influence over gene expression through their copy numbers (multicopy plasmid), effectively also increasing gene copy numbers. Derivatives of the 2µ plasmid were used in the overexpression studies performed in **Papers II-IV**.

A commonly used inducible promoter is p*GAL1* that induces gene expression more than a 1000-fold in presence of galactose in the growth medium, and represses expression when the carbon source is shifted to glucose (Johnston and Davis 1984). Examples of other inducible promoters used for gene overexpression includes the p*CUP1-1* that is induced in presence of copper ions (Etcheverry 1990) and the p*MET25* that is induced in absence of methionine (Mumberg et al. 1994).

Using p*GAL1*, a majority of genes in *S. cerevisiae* were systematically overexpressed (Sopko et al. 2006). It was demonstrated that ~15% of overexpressed genes caused fitness defects, and that genes involved in cell cycle, transport, transcriptional regulation and signal transduction, e.g. protein kinases and protein phosphatases, were overrepresented in this set (Sopko et al. 2006). Through this study a collection of mutants

each overexpressing an individual gene (covering >80% of the yeast genome) from plasmids was also constructed (Sopko et al. 2006).

These overexpression, or dosage sensitivity, phenotypes are more difficult to explain than loss-of-function phenotypes. They differ mechanistically to a large extent as the gene sets do not overlap significantly (Sopko et al. 2006; Yoshikawa et al. 2011; Kaluarachchi Duffy et al. 2012). While deletion of a duplicated gene can be functionally compensated by the remaining copy, it has interestingly been demonstrated that when one of the copies is overexpressed, a fitness defect has a significantly higher frequency of occurring (Yoshikawa et al. 2011). In **Paper IV**, many of the strongest fitness effects caused by gene overexpression of protein phosphatases were observed for duplicate genes (Kellis et al. 2004).

There are several proposed mechanistic reasons for gene overexpression phenotypes (Prelich 2012). Broadly, overexpression can either inhibit or activate proteins and a number of possible mechanisms explain the two. Inhibition can be caused by the reduction of steady-state levels of proteins, by the disruption of stoichiometrically balanced complexes (Sopko et al. 2006), by the sequestration of proteins by mutant enzymes or through functional inactivation of proteins. Activation on the other hand can be caused by the overexpression of inactive genes that activate otherwise silent pathways, by increasing levels of limiting proteins, by counteracting the activity of repressors or by increasing the specific activity of proteins, e.g. through posttranslational modifications (**Papers II and IV**). In a thorough study aimed at elucidating the causes of overexpression phenotypes, the authors tested a variety of different genomic and experimental features in their ability to predict the phenotypes (Vavouri et al. 2009). They demonstrate that intrinsic protein disorder and a tendency to participate in a large number of pairwise protein-protein interactions could be a likely cause. Disordered, or unstructured, regions of proteins can be seen as linear functional sites allowing for many recognition events, such as protein-protein associations and posttranslational modifications. These regions are more likely to take part in promiscuous interactions when proteins are overexpressed. This fact could, at least in part, explain why the dosage sensitive genes are enriched for transcription factors and signal transduction components (Sopko et al. 2006). This, however, did not seem to predict the overexpression phenotypes observed for protein phosphatases in **Paper IV**.

4.1.2.1 Genetic tug-of-war (gToW)

Genetic tug-of-war, or gToW, is a plasmid-based gene overexpression methodology (Fig. 2) (Moriya et al. 2006). Contrary to most gene overexpression strategies using artificial promoter-gene fusions to control gene expression, in gToW the DNA regulatory regions (promoter and terminator) for each gene are maintained to allow native regulation of gene expression (Fig. 2A).

Expression of the gene is linked to the expression of a defective metabolic gene, *leu2-d*, that has a functional but truncated promoter causing an expression of less than 5% compared to that of the wild-type *LEU2* gene (Erhart and Hollenberg 1983; Schneider and Guarente 1991). The low expression drives the plasmid to high copy number since the allele needs to be present in many copies to give leucine prototrophy and support unperturbed growth during cultivation in growth medium lacking leucine. As the plasmid copy number increases, the gene copy number increases proportionally and hence gene expression. If the gene overexpression causes fitness defects, a negative pressure on plasmid copy number will balance the positive pressure conferred by the *leu2-d* allele, causing a genetic tug-of-war (Fig. 2A). This compromise results in a growth rate that correlates to plasmid copy number and can be measured precisely and quantitatively using liquid high-resolution microcultivation (Fig. 2B) (Warringer and Blomberg 2003; Warringer et al. 2003). Plasmid copy number can also be determined. A more detailed view on the gToW plasmid can be found in **Paper II, Fig. S1A**. As a control in gToW experiments, plasmids without gene insert (empty vector or gToW-Ø) but otherwise identical in sequence are used. When cultivated in absence of leucine (-Leu), the growth curve profile is close to that observed in presence of leucine (+Leu) (**Paper II, Fig. 1B**).

To a large extent, the observed fitness defects will be a consequence of increased levels of gene product, as it has been demonstrated that protein abundance correlates to gene copy number (Springer et al. 2010). In a gToW study on the cell division cycle genes in *S. cerevisiae* it was also demonstrated that protein amount correlated to gene copy number (Moriya et al. 2006). In **Paper III** the average gToW plasmid copy number of a subset of strains was quantified and shown to correlate to growth rate also for this set of genes (**Paper III, Fig. S2**). It was further demonstrated in **Paper II** that protein expression levels increased in response to the gToW perturbation (**Paper II, Fig. S6B**) and also that part of the toxicity of certain genes overexpressed was caused by the function of the gene/protein as the toxicity could be repressed by deletion of genes located

upstream in the signal pathway (**Paper II, Fig. 1D and S6D**). A likely culprit in that sense is an increase in catalytic activity and it was demonstrated in **Paper IV** that overexpression phenotypes of protein phosphatases could be suppressed by addition of inhibitors of protein phosphatase catalytic activity (**Paper IV, Fig. 3**). It is also possible that other parameters such as levels of mRNA molecules can contribute, but this is unlikely a major cause of overexpression defects since it has been demonstrated that there is no correlation between dosage sensitivity and mRNA abundance (Vavouri et al. 2009).

The gToW methodology has been used to perform robustness analysis of cell division cycle (CDC) genes in both *S. cerevisiae* (Moriya et al. 2006; Kaizu et al. 2010) and the fission yeast *Schizosaccharomyces pombe* (Moriya et al. 2011).
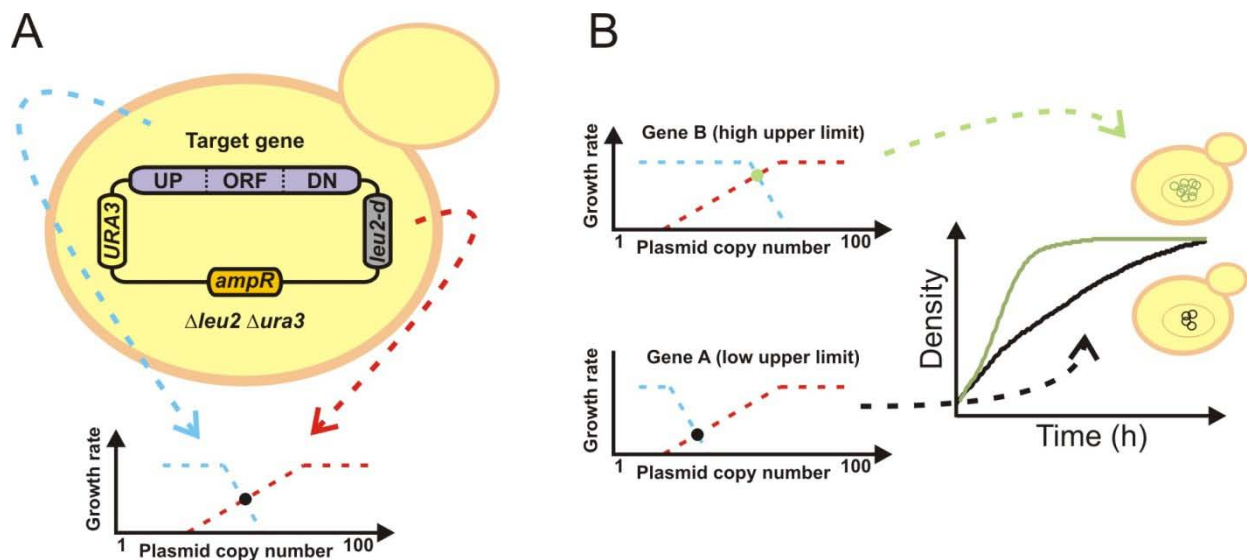
Figure 2. gToW methodology. (A) A target gene is inserted into a plasmid backbone together with its native regulatory regions both upstream (UP) and downstream (DN), including promoter and terminator sequences. Also on the plasmid is a defective metabolic allele, *leu2-d*, that must be present in many copies to support the cells need for leucine and so creates a bias towards increased plasmid copy numbers. The target gene copy number increases proportionally, causing overexpression. If the gene product causes fitness defects when overexpressed, this will create another bias towards decreased plasmid copy numbers. (B) The trade-off, or genetic tug-of-war, between these two selection biases, will cause a certain number of plasmids to be enriched in the cells. Gene B, with a high upper limit of gene copy numbers will allow cells to concentrate more plasmids and hence grow at a higher rate. The opposite scenario is observed for Gene A, with a low upper limit. This effect on cellular fitness can be measured quantitatively using microcultivation. Based on and inspired by (Moriya et al. 2006).

4.1.3 gToW versus GAL overexpression

Although promoter-gene fusions, e.g. p*GAL1*, are well established in the scientific community and provide valuable information on the functions of genes and proteins, a number of issues exist that can cloud interpretation of the results.

Inducible promoters ignore native gene regulation and can cause unrealistically high gene expression levels. Expression is also induced by the addition of chemicals, e.g. galactose, to the growth medium that might have an impact on global gene expression. In p*GAL1*-driven overexpression, the ORFs are also fused N-terminally to a glutathione *S*-transferase-polyhistidine (GST-HisX6) tag that allows purification of the target protein using glutathione-agarose beads (Zhu et al. 2001). The use of tags might also interfere with normal protein function, specifically if the protein function is dependent on an intact terminal domain (Caesar and Blomberg 2004).

In a study where we tried to use tandem affinity purification (TAP)-tagged gToW strains for protein measurements additional aggravating effects for many of the gToW phenotypes were observed (data not shown). In a later study it was demonstrated that the toxicity effect largely was due to increases of a promoter, the *EF-1α* that is frequently used in heterologous marker cassettes, e.g. HIS3MX, KanMX and NatMX, commonly used in mutant collections (Babazadeh et al. 2011). The use of tags on proteins can clearly have additional negative effects on cellular fitness in overexpression studies.

The major benefit of the gToW system, on the other hand, is the retained native gene regulatory sequences that allow a regulated gene expression that increases as the gene dosage increases. The gToW system is also free of protein tags.

Just as native gene regulation is a benefit of the gToW system it also opens up the possibility of feedback mechanisms and/or titration of transcription factors as contributing factors to observed overexpression phenotypes.

In **Paper III** a comparison between the effects of overexpression using gToW and p*GAL1*-driven overexpression demonstrated a reasonably strong correlation. A substantial amount of discrepancies, however, was also observed (**Paper III, Fig. S3**). The differences might, at least in part, be explained by varying absolute levels of overexpression in the different setups (**Paper II** and (Moriya et al. 2006)).

It should, however, be beneficial to retain as much as possible of the native gene regulatory mechanisms when studying the effects of an increased gene dosage.
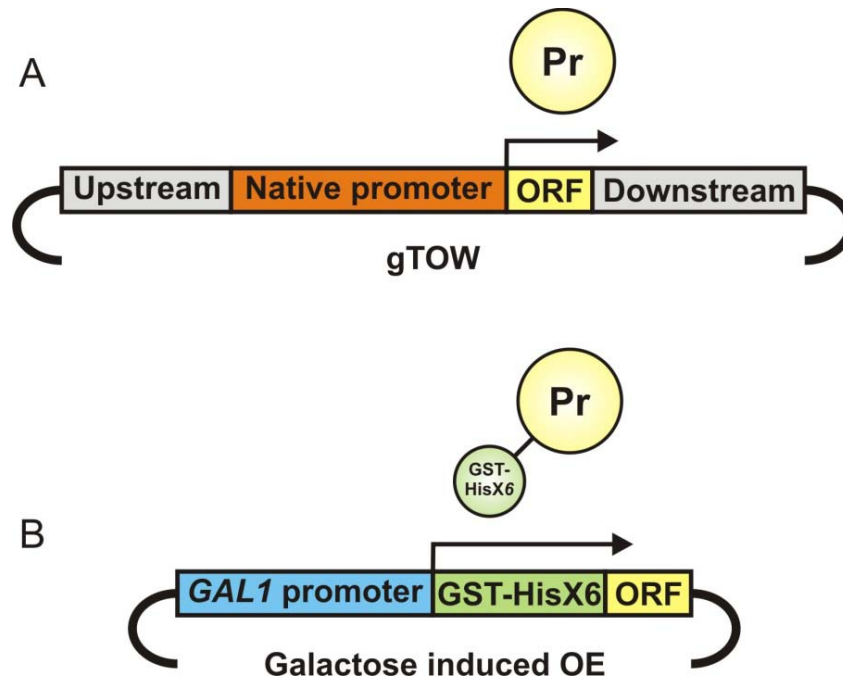
Figure 3. gToW gene overexpression compared with galactose inducible gene overexpression. (A) In the gToW system the ORF together with its native regulatory regions (promoter and downstream sequence including the terminator) is incorporated into a vector plasmid. Native gene regulation allows controlled gene expression that increases proportionally to plasmid copy number. A native untagged protein is expressed. (B) Inducible promoter-gene fusion systems exemplified by the commonly used p*GAL1*. Besides from being artificially fused to the *GAL1* promoter, the ORF is also N-terminally fused to a GST-HisX6 tag that allows protein purification but at the same time potentially can interfere with the function of the protein (Zhu et al. 2001).

4.2 Environmental perturbations

Microorganisms, e.g. yeasts, are found virtually everywhere on the planet, from deep below the surface of the ocean (Gadanho and Sampaio 2005) to high in the Arctic glaciers (Butinar et al. 2007) and have been isolated from locations in all major continents (Liti et al. 2009). As such yeasts are subjected to extremely varying environmental conditions that often fluctuate rapidly. To function properly, in terms of growth and survival, these organisms need to maintain highly regulated and controlled internal conditions despite external variation. Yeasts have evolved elaborate mechanisms that allow them to adapt and survive in response to numerous environmental challenges, e.g. high osmolarity and exposure to toxic chemicals, e.g. metals, to name a few. These mechanisms involves sensing of environmental cues and transmission of this information through signalling

pathways/networks to downstream effects, e.g. alterations of gene expression programs (Gasch et al. 2000).

4.2.1 High osmolarity

Given their lifestyle, yeasts are subjected to fluctuations in water activity. If a yeast cell is suddenly exposed to an increase in the concentration of a solute (high osmolarity), e.g. high concentration of sugars due to the breakage of a grape, a hyperosmotic shock (low water activity) follows. Through osmosis water molecules tend to travel from areas with high water activity to areas with low water activity to reduce the differences in solute concentration, e.g. from one side of the cell's plasma membrane to the other. In response to a hyperosmotic shock this causes water to flow out of the cell with an ensuing loss of turgor pressure (the force by which the plasma membrane pushes against the cell wall) and cell shrinkage. The opposite is observed if a cell is subjected to a hypo-osmotic shock (high water activity) upon which water flows into the cell causing an increase in turgor pressure and cell swelling (Hohmann 2002).

It is critical for a cell to maintain a proper intracellular water balance in terms of upheld structural features and favourable conditions for biochemical processes. To counteract conditions of high and low osmolarity, cells have evolved mechanisms and systems for osmoregulation that allow osmotic homeostasis to be maintained. In **Paper II** such a system, the high osmolarity glycerol (HOG) signal transduction pathway, is studied. This pathway is essential in the cellular response to increased external osmolarity and an important function is the production of the compatible osmolyte glycerol. A build-up of glycerol causes increased solute concentration with an ensuing flow of water back into the cell, effectively restoring cell shape and turgor pressure.

Conditions of high osmolarity can be induced experimentally by cultivating yeast cells in growth medium supplemented with a high concentration of a solute. In **Papers II-IV** NaCl, or table salt, was used to induce osmotic stress and a subsequent activation of the HOG pathway. A concentration of 0.85M NaCl was used that causes a 40-50% reduction in reference (wild-type) strain growth rate (Warringer et al. 2003).

4.2.2 Cellular responses to the metalloid tellurite (Paper I)

In nature, organisms are exposed to toxic metals and metalloids, among other chemical compounds, that often can reach high concentrations locally. These elements include arsenic, cadmium, mercury and copper, to name a few, and cellular responses to exposure of these have been well studied in yeast. These studies have identified many genes and gene products that are involved in these response mechanisms (Wysocki and Tamas 2010; Bleackley et al. 2011). Biological systems have evolved to deal with these elements through a variety of protective and detoxifying mechanisms.

Despite being one of the rarest elements in the Earth's crust, concentrations of the metalloid tellurium (Te) has increased in many locations due to expanded use in the electronics industry. The Te derivative and oxyanion tellurite is also highly toxic to organisms. In **Paper I** a study was conducted to trace the basis of sensitivity and resistance mechanisms in response to tellurite exposure.

Tellurite has been shown to be reduced to methylated forms or elemental Te in prokaryotic cells, the latter typically deposited as black aggregates. This accumulation results in a dark colouration of cell colonies on agar plates and is a hallmark of Te accumulation in bacteria. Black aggregates are accumulated also in cells of both *S. cerevisiae* and the fission yeast *Schizosaccharomyces pombe* (**Paper I, Fig. 1C and D**) and it was further demonstrated that these aggregates were caused by elemental Te (**Paper I, Fig. 1G**).

In an attempt to trace genetic determinants of Te-related traits in yeast, a gene-by-environment screen was performed using a genome-wide collection of nonessential gene deletion mutants (Fig. 1G) using our phenotyping strategy of colonies on solid agar plates (Fig. 1D and E). Colony colouration was used as a phenotypic readout of Te accumulation and colony size as a phenotypic readout of tellurite sensitivity/resistance. These phenotypes were quantified manually by visual inspection of colonies (Fig. 1G). It was demonstrated that reduction of tellurite and Te accumulation correlated with loss of cellular fitness and that there was a causal link between the two. A metabolic pathway, the sulfate assimilation pathway, was also found to be central to tellurite toxicity and reduction (**Paper I, Fig. 2B and 3A**). Mutants with altered tolerance to tellurite also displayed a similar deviation in tolerance to selenite suggesting that there is a shared toxicity mechanism between the two (**Paper I, Fig. 4**). It was further demonstrated that

tellurite reduction and toxicity was partially mediated via a mitochondrial respiratory mechanism that did not involve a substantial generation of oxidative stress.

Surprisingly, few protein kinases and protein phosphatases were picked up by the screen suggesting their dispensability for tellurite related traits in yeast.

5 CELLULAR SIGNALLING

Cellular signalling can be described as the process by which a cell senses extracellular and/or intracellular stimuli and transforms this into an intracellular signal that is both amplified and propagated throughout the cell to elicit proper responses. As many environmental fluctuations are stressful to the cell, an important cellular response is to adapt through changes is gene expression patterns and/or metabolism.

Cellular signalling typically involves the activation of specific signalling components in the cell, from sensing of stimulus by cell surface receptors to modification of specific signal transduction molecules within the cell, e.g. protein kinases and protein phosphatases. These signalling components operate in a coordinated fashion, referred to as signal transduction pathways or signalling networks. A generated signal is transmitted in these pathways through reversible posttranslational protein modifications, mainly protein phosphorylation. A typical downstream output is the interaction between the signalling network and the transcriptional network through associations with transcription factors (TFs) that guide transcription of genes (Wang et al. 2012).
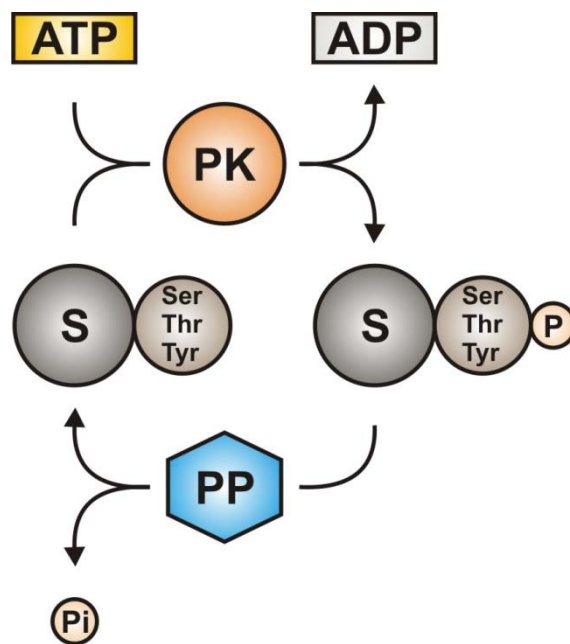
5.1 Phosphorylation

Posttranslational modifications (PTMs) of proteins are fundamental to life and present in both prokaryotic and eukaryotic cells. PTMs fulfil many physiological roles in the cell, and affect proteins in different ways, e.g. by changes to activity, conformation, protein interaction affinity and cellular localisation. PTMs are believed to be one of two major mechanisms (the other being alternative mRNA splicing) to extend the coding capacity of a limited set of genes in the genome of an organism, making the proteomes many orders of magnitude more complex than the corresponding genomes (Walsh et al. 2005).

A multitude of PTMs have been identified thus far, but among these protein phosphorylation remains one of the best characterised largely because of the relative ease of detecting protein phosphorylation both *in vivo* and *in vitro*. Protein phosphorylation is also important to cellular signalling and it has been demonstrated that more than 30% of the yeast proteome is affected by this modification at any given moment (Ficarro et al.

2002). This number might be substantially higher for a eukaryotic cell as it has been shown that a majority of proteins in a mammalian cell are phosphorylated on at least one amino acid residue (Olsen et al. 2006).

In this modification a phosphate group (PO_4^{3-}) is covalently attached to the side chains of amino acid residues in the protein, mainly serine (Ser), threonine (Thr) and tyrosine (Tyr) residues, by a class of enzymes called protein kinases. Another class of enzymes, protein phosphatases, counteract the activities of protein kinases by removing the phosphate groups through dephosphorylation. Phosphorylation is thus a reversible PTM (Fig. 4).

Figure 4. Phosphorylation cycle. A protein kinase (PK) covalently attaches a phosphate group (P) to a serine (Ser), threonine (Thr) or a tyrosine (Tyr) amino acid residue in a substrate (S) protein at the expense of an ATP molecule that is converted to an ADP molecule. A protein phosphatase (PP) will counteract the activity of the PK by dephosphorylating the protein substrate and releasing inorganic phosphate (Pi) into solution.

In yeast, as well as many other organisms, most phosphorylation events take place on Ser and Thr residues (82.3 and ~17.5%, respectively) while only 0.027% of Tyr residues being phosphorylated (Chi et al. 2007). Other amino acid residues, e.g. histidine, arginine, lysine, cysteine, aspartate and glutamate can also be phosphorylated, and it has become increasingly clear that these modifications also play important roles in various cellular processes (Ciesla et al. 2011). In **Paper II** it is demonstrated that overexpression

of protein kinases, using the gToW system, can cause phosphorylation of substrates (**Paper II, Fig. 1E**).

Phosphorylation and dephosphorylation have profound effects on the activity of proteins, toggling them between "on" and "off" states rapidly, often in response to environmental stimuli and are frequently observed in signal transduction.

5.2 Protein kinases (PKs)

Protein kinases (PKs) constitute a large and functionally diverse family of enzymes that comprises 1.5-2.5% of all eukaryotic genes (Manning et al. 2002). Although diverse in what biological processes they partake, protein kinases are highly conserved in their catalytic kinase domain (catalysing the phosphorylation event), and this domain is therefore typically used for phylogenetic classification of PKs (Hunter and Plowman 1997). Studies on domain similarity across species have demonstrated that ~80% of yeast kinases have homologues in higher organisms, e.g. worm, fly and human (Manning et al. 2002).

The catalytic kinase domain, located in a deep cleft between two lobes, completely accommodates an ATP molecule for covalent attachment to a protein substrate interacting with the kinase at the mouth of the cleft (Bossemeyer 1995). Substrate specificity is determined by the kinase domain, and in most cases is categorised in one of two classes; dual specificity Ser/Thr or Tyr kinases. A number of dual-specificity kinases are also able to phosphorylate Tyr residues in addition to Ser/Thr (Hanks et al. 1988).

The yeast genome contains 121 characterised or putative protein kinases (Zhu et al. 2000; Fiedler et al. 2009) of which most are Ser/Thr kinases (Hunter and Plowman 1997). Among these, yeast has seven dual specificity kinases (Hunter and Plowman 1997; Zhu et al. 2000), and one characterised histidine kinase, *SLN1*, which is an osmosensor in the high osmolarity glycerol (HOG) mitogen-activated protein (MAP) kinase pathway and part of a phosphorelay system with similarity to bacterial two-component regulatory systems (Ota and Varshavsky 1993). Although yeast has no Tyr kinase family members (Pincus et al. 2008), tyrosine phosphorylation plays a central role in yeast, e.g. MAP kinases are phosphorylated on conserved threonine and tyrosine residues in their activation loops in response to a wide range of often stressful environmental conditions (Gustin et al. 1998; Hohmann 2002). There are also tyrosine phosphorylations that are

believed to be of functional importance to other types of proteins. Roughly 13% of PKs are essential to viability under standard laboratory conditions (Winzeler et al. 1999).

In a study using proteome chip technology, over 1300 different proteins were found to be substrates for a majority of known protein kinases (Ptacek et al. 2005). Each of these kinases recognized up to 256 different substrate proteins with an average of 47, while 73% of substrates were recognized by fewer than three kinases indicating unique substrate recognition profiles. Transcription factors were found to be phosphorylated to a great extent further strengthening the importance of this modification in cellular responses based on alterations to gene expression.

Interestingly, many proteins contain multiple phosphorylation sites with different kinetics, suggesting their importance as platforms for integrating signals (Olsen et al. 2006). In the yeast pheromone signalling pathway, for instance, a fraction of phosphopeptides were at least 2-fold differentially regulated in response to mating pheromone (Gruhler et al. 2005). These results point towards a dynamic phosphoproteome as a missing link for an integrative view of cellular regulation.

5.3 Protein phosphatases (PPs)

As for PKs, it is the catalytic phosphatase domain that has been used to classify the protein phosphatases (PPs) into four major families or groups (Moorhead et al. 2009). A majority of Ser/Thr dephosphorylation is carried out by members of the phosphoprotein phosphatase (PPP) and metallo-dependent protein phosphatase (PPM) families. The latter is functionally Mn^{2+}/Mg^{2+} dependent, which has given rise to the classification (Moorhead et al. 2009). Although the overall sequence similarity between members of the PPP and PPM families differs, the catalytic centre appears to be similar in terms of three-dimensional structure and catalytic mechanism (Das et al. 1996). Tyr dephosphorylation on the other hand is carried out by a protein tyrosine phosphatase (PTP) family where all members are defined by the catalytic sequence motif $C(X)_5R$ (Kolmodin and Aqvist 2001). This family contains a major group of tyrosine specific phosphatases, the dual specificity phosphatases (DSPs) capable of dephosphorylating Ser/Thr and Tyr residues, the low molecular weight PTPs (LMPTPs) and the dual specificity Cdc25 phosphatases (Kolmodin and Aqvist 2001). A fourth group of phosphatases belongs to the haloacid

dehalogenase (HAD) superfamily where members are defined by the catalytic sequence motif DXDX(T/V).

Relative PKs, PPs constitute a smaller group of enzymes. In yeast 38 characterised or putative protein phosphatase catalytic subunits have been identified (Fiedler et al. 2009). Based on the fact that this small group of enzymes perform the protein dephosphorylations in the cell, it has for long been assumed that PPs are promiscuous by nature and less substrate specific compared to PKs. It has, however, become increasingly clear that the PPs constitute a group of highly regulated enzymes that achieve substrate specificity through complex formation with several different regulatory subunits (Virshup and Shenolikar 2009). This seems to be particularly evident for Ser/Thr PPs where the catalytic domain is believed to have functionally evolved by complex formation to numerous regulatory subunits (Moorhead et al. 2009). Yeast have at least 39 regulatory subunits (Fiedler et al. 2009), but it is possible that this number will increase as more functional knowledge is gained for still uncharacterised genes. Of the PPs roughly 11% are essential to viability under standard laboratory conditions. The number of PPs dispensable to growth is therefore similar to that observed for PKs (Winzeler et al. 1999).

As for PKs, PPs have important functions in the regulation of signal transduction pathways where their activity counteracts that of the PKs and not only reset the pathway after adaptive responses have occurred, following activation by environmental stimuli, but also reduce basal pathway activity (Macia et al. 2009). PPs also play an important role in regulating correct cellular localisation of proteins, e.g. MAPKs, through binding properties facilitated by specific conserved motifs containing docking sites (D-sites) (Martin et al. 2005).

5.3.1 gToW on PPs as a class of enzymes (Paper IV)

In large part due to the fact that PPs is a class of signalling components that has received less attention compared to PKs, in **Paper IV** the PP catalytic subunits in yeast were systematically overexpressed using the gToW method. A high prevalence of negative impacts on fitness was observed among PPs, with several of the most extreme effects for Ser/Thr PPs, e.g. *PPZ1/2* and *PPH21/22* (**Paper IV, Fig. 1A and D**). Using p*GAL1*-driven gene overexpression it has previously been demonstrated that Ser/Thr PPs are overrepresented among genes causing fitness defects when overexpressed (Sopko et al. 2006). Of the PPs 78% caused a significant growth defect upon overexpression. Although

growth defects were frequent, they ranged in severity and were also spread over different PP families. These strong overexpression phenotypes stand in stark contrast to an almost complete lack of effects conferred by deletions of the same genes (**Paper IV, Fig. 2**).

As earlier described it was demonstrated that gToW-PP overexpression phenotypes could be suppressed by PP inhibitors, suggesting that overexpression fitness defects, at least in part, stems from increased phosphatase catalytic activity. It was also observed that many of the strongest overexpression effects were with PPs that are considered duplicate gene pairs, e.g. *PPZ1/2*, *PPH21/22* and *PTC2/4* (Kellis et al. 2004), and also associated with protein complexes, e.g. *PPH21/22* in PP2A complex. It is therefore likely that another mechanism contributing to overexpression toxicity might stem from imbalances between PP catalytic subunits and other components in these complexes.

To test to what extent PP overexpression effects were dependent on environmental factors the gToW strains were exposed to stressors with broad effects on cellular functions. Interestingly, it was observed that the gToW-PP overexpression fragility profiles were largely independent of environmental variations since no condition-specific fragilities were observed for a majority of PPs (**Paper IV, Fig. 4**).

To test the dependence of PP overexpression on the organism's position in genotypic space, a genetic interaction screen was performed where PP overexpression was pairwise combined with individual gene deletions for a majority of the cellular signalling components. The synthetic genetic array (SGA) method was used by which gene mutations systematically can be combined in a cell (Tong and Boone 2006).

An interaction was considered to have occurred if it had an aggravating effect (negative) or buffering effect (positive) on the original overexpression phenotype. A striking observation was that the tested PPs had few, if any, significant genetic interactions and among these a majority were negative (**Paper IV, Fig. 6**). Together with the observation of environmental independence of overexpression phenotypes for PPs this suggests that the signalling network as such is rather robust against external or internal (genotypic) changes.

One way to interpret genetic interaction data mechanistically for PPs is that they act on the same substrate(s) as a corresponding PK. Thus, enhanced expression of a PP would generate a highly dephosphorylated substrate that could be detrimental for cellular fitness. A deletion of the corresponding PK would then aggravate the growth inhibition since more substrate would be non-phosphorylated by deletion of the PK. To follow this up for *PPH21/22* that had the highest number of significant negative interactions of all tested

gToW-PPs, a literature survey gave mechanistic support to some of the strongest interactions observed between *PPH21/22* and PKs. Interestingly there was also a high degree of overlap between PPH21/22 in their interaction profiles (**Paper IV, Fig. 8**).

Unfortunately no database on phosphatase-substrate relations yet exist which made the results more difficult to interpret. Several of these interactions showed clear links to well-established cellular features where PPs play central roles, e.g. cell cycle and cell polarity. For some of these, no earlier reports could be found suggesting that these interactions indicate new links to specific genes.

5.4 Mitogen-activated protein kinase (MAPK) pathways

Information in the cell is propagated through the cell by the activities of signalling components, e.g. PKs and PPs that are brought together in signalling networks, or pathways. Prominent among these signalling pathways are the mitogen-activated protein kinase (MAPK) pathways.

Eukaryotic organisms use MAPK pathways to sense fluctuations in the surrounding environment and respond accordingly. These pathways are well conserved through evolution and regulate many essential cellular processes, e.g. proliferation, differentiation, development and response to a wide range of stressful conditions (Qi and Elion 2005).

At the core of the MAPK pathway, a three-component kinase cascade is located, comprised of a MAP kinase kinase kinase (MAPKKK) that upon pathway activation phosphorylates a MAP kinase kinase (MAPKK) on a Ser and Thr residue. Active MAPKK dually phosphorylates a MAP kinase (MAPK) on conserved Thr and Tyr residues within the activation loop (TXY motif). Activated MAPK then phosphorylates targets on Ser and Thr residues, that in turn controls many cellular processes, e.g. gene expression (O'Rourke and Herskowitz 2004).

The protein kinase C (PKC) pathway senses, among other stimuli, cell wall perturbations and responds through measures to increase cell wall resilience (Levin 2011). The functions of the largely overlapping mating and pseudohyphal differentiation pathways (MAT/PHD) are to allow haploid cells to mate and both haploids and diploids to switch their mode of growth when faced with nutrient (haploids) and nitrogen (diploids) starvation. The mating pathway is best characterized in haploids where it, upon activation by pheromones secreted by cells of opposite mating type, triggers cell cycle

arrest and polar growth towards a partner cell (Bardwell 2004). The High Osmolarity Glycerol (HOG) pathway is essential to adaptation to increased extracellular osmolarity (Hohmann 2002).

5.4.1 High Osmolarity Glycerol (HOG) pathway

The High Osmolarity Glycerol (HOG) signal transduction pathway (Brewster et al. 1993) is one of the most extensively studied eukaryotic signal transduction pathways (Fig. 5 and **Paper II**). It is activated by high osmolarity and is also essential under this condition. At its core is a highly conserved MAPK module, upon which two independent upstream branches converge.

One of these is termed Sln1 branch after the membrane histidine kinase and presumed osmosensor Sln1p. Sln1p, the phosphotransfer protein Ypd1p and the response regulator Ssk1p constitutes a two-component phosphorelay module, a design more commonly found in prokaryotes. During a hyperosmotic shock, this module becomes inactive and dephosphorylated Ssk1p activates the MAPKKKs Ssk2p and Ssk22p (Saito and Tatebayashi 2004).

An input sensor for the other branch has been more difficult to elucidate, and a number of candidates exist. The protein initially believed to be the sensor is Sho1p, which has given name to the branch. Studies, however, indicate the two mucin-like proteins Msb2p and Hkr1p as the most upstream elements and thus potential osmosensors in this branch (O'Rourke and Herskowitz 2002; Tatebayashi et al. 2007). Upon pathway activation Sho1p, through its cytoplasmic domain, assembles a complex consisting of the MAPKK and scaffold Pbs2p that tethers the MAPKKK Ste11p and the MAPK Hog1p into close proximity. Transmission of signal also requires the membrane associated GTPase Cdc42p which forms a complex with Ste20p and the adaptor protein Ste50p.

When activated the MAPKKKs Ste11p, Ssk2p and Ssk22p are each able to phosphorylate and activate the MAPKK Pbs2p, which also function as a scaffold (Dard and Peter 2006). Pbs2p then phosphorylates and activates the MAPK Hog1p that helps establish a number of responses to the hyperosmotic shock.

Very rapidly it seems that the paramount role of Hog1p is to alter metabolic conditions to allow for an elevated glycerol production (Westfall et al. 2008). This was demonstrated experimentally by attaching Hog1p to the plasma membrane and thereby preventing it from shuttling to the nucleus where it partakes in another important

response, the transcription of osmoresponsive genes involved in many cellular processes such as production of the osmolyte glycerol (O'Rourke and Herskowitz 2004). Activated Hog1p also mediates S phase delay, likely to be of importance in allowing Hog1p induced gene expression before replication (Yaakov et al. 2009).
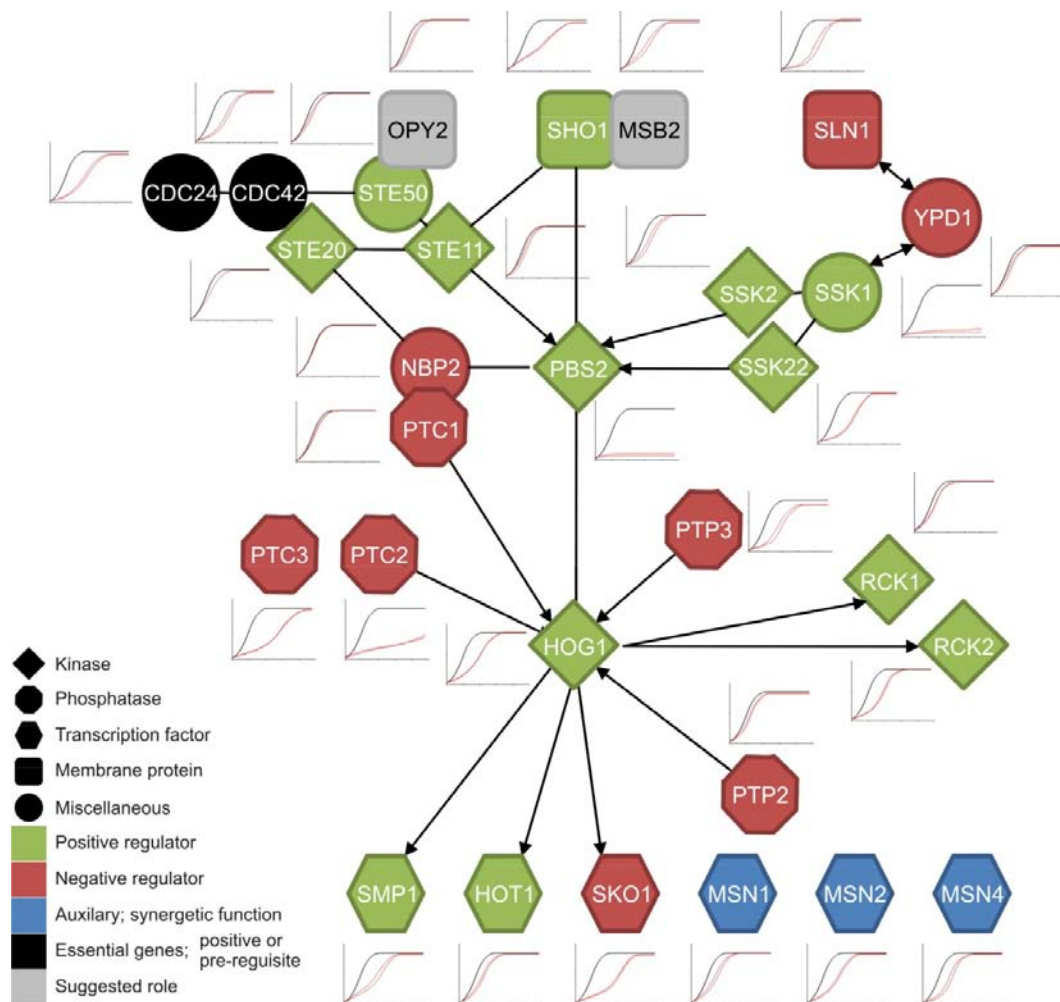
Figure 5. HOG pathway. Depicted are genes of fundamental importance to the cellular responses to high osmolarity in yeast. These genes were overexpressed using gToW and the effects measured quantitatively by microcultivation as demonstrated by the growth curves next to gene names. Reprinted from (Krantz et al. 2009) under the terms of the Creative Commons Attribution Licence.

To counteract pathway activity, a number of PPs monitor and regulate the activity of Hog1p. The Ser/Thr PPs Ptc1p, Ptc2p, Ptc3p as well as the Tyr PPs Ptp2p and Ptp3p (Wurgler-Murphy et al. 1997) function to down-regulate activity of the pathway by dephosphorylating Hog1p (Saito and Tatebayashi 2004; Martin et al. 2005). Having so many phosphatases performing essentially the same task might at first appear to be an unnecessary investment for the cell, but studies indicate that similar phosphatases can

have complementary roles. This is also the case for the PTCs where Ptc1p regulates basal activity and resets Hog1p activity after adaptation whereas Ptc2p and Ptc3p limit maximal Hog1p activity (Saito and Tatebayashi 2004; Martin et al. 2005). Ptc2p and Ptc3p also act outside the context of the HOG pathway by dephosphorylating the cyclin-dependent kinase Cdc28p (Cheng et al. 1999).

One of the substrates for activated Hog1p is the TF Sko1p that upon activation binds to, among many others, the promoter of *PTP3* and promotes its expression and thus demonstrates that negative feedback also plays an important role in control of pathway activity (Proft et al. 2005).

5.4.1.1 gToW on HOG pathway (Paper II)

To study the effects on fitness caused by overexpression of components in a single signal transduction pathway in yeast, 29 HOG-pathway related genes were systematically overexpressed using the gToW method in **Paper II**.

As for the PPs, a high prevalence of negative impacts on fitness from gene overexpression was observed as ~76% caused significant fitness defects (Fig. 5 and **Paper II, Fig. 1A and C**). Overexpression effects were also observed for all protein classes (PKs, PPs and TFs) involved in the pathway. These were also dispersed throughout the pathway seemingly independent of pathway topology. When comparing growth phenotypes to those caused by p*GAL1*-driven gene overexpression only 15% of the HOG pathway genes conferred detectable growth defects (Sopko et al. 2006) (**Paper II, Fig. S3A-C**). These discrepancies could, at least in part, be explained by basal protein expression levels (**Paper II, Fig. S3E**). *PTC2* is highly expressed and has a stronger gToW phenotype while *SSK2* is lowly expressed and has a stronger p*GAL1* phenotype (Ghaemmaghami et al. 2003; Sopko et al. 2006). The correlation between level of native expression and degree of response in the two systems were, however, rather weak over the entire set of genes.

The strongest overexpression effects were observed for the *PBS2* and *SSK1*. Their toxicity could be traced to depend on pathway hyperactivation as Hog1p was shown to be constitutively active through phosphorylation (**Paper II, Fig. 1E**). This finding was further supported by the fact that the toxicity of *PBS2* and *SSK1* could be partially suppressed by deletion of the downstream component *HOG1* (**Paper II, Fig. 1D**).

As for the PPs it was demonstrated that the HOG gToW overexpression toxicity profile was largely independent of environmental variation, even upon pathway activation by NaCl, and remained similar to that of unstressed conditions (**Paper II, Fig. 3A**). The gToW overexpression even caused stronger fitness defects under normal conditions than NaCl stress, indicating a general dampening effect of gToW phenotypes in stress conditions (**Paper II, Fig. 3B and C**).

5.4.2 gToW on additional MAPK and nutrient sensing pathways

To get a wider systems view on the effects on fitness in response to overexpression of cellular signalling components, in **Paper III** the gToW study on the HOG pathway was expanded to include all MAPK pathways in yeast and also the two most important nutrient sensing pathways, Target of Rapamycin (TOR) and Protein Kinase A (PKA). The latter two pathways are vital regulators of metabolism and growth in yeast (Zaman et al. 2008). The TOR pathway monitors nitrogen levels and coordinates uptake and production of amino acids while the PKA pathway monitors growth potential and is negatively by a wide range of stress conditions (Zaman et al. 2008). These five modules constitute a major part of the yeast signalling network with the core sections also being well conserved throughout Eukaryota.

A set of genes constituting either core components of the pathways (PKs, PPs and their regulatory subunits, downstream TFs) or components with suggested connections to the pathways were overexpressed. Also for this set of genes an extreme fragility was observed upon overexpression as ~58% of the overexpressed genes caused significant growth defects (**Paper III, Fig. 1C**). Similarly to that observed in the gToW studies on the HOG pathway and PPs, the high frequency of fitness defects caused by gene overexpression in general seems to stand in stark contrast to that observed for deletions of corresponding genes (**Paper III, Fig. 1F**). Overexpression phenotypes also tended to have a higher frequency of low to intermediate effects while deletions either caused no observable growth effects or inviability (essential genes). The signalling network thus seems much more resilient to decreases in gene expression, relative increases. A possible explanation might be that it is easier to compensate loss-of-function mutations via alternative pathways or mechanisms than to silence gain-of-function mutations.

As observed in the HOG study, fragile nodes were widely distributed also in the other pathways with no clear patterns as to biochemical function, spatial or topological

positions (**Paper III, Fig. 2**). It is therefore likely that patterns of fragility reflect component specific properties or local network architecture. A substantial fraction of the highly fragile nodes were found to be activators known to be stoichiometrically inhibited, e.g. Ste12p by Dig1p/Dig2p and Tpk1p/Tpk2p/Tpk3p by Bcy1p.

To test the impact of environmental variation on the fragility patterns caused by gToW overexpression also for the expanded signalling network, the constructs were probed in conditions known to activate or inhibit the pathways, NaCl (HOG pathway activation), heat (PKC pathway activation), caffeine (PKA pathway activation) and rapamycin (inhibition of TORC1 complex). Similarly to that observed for the HOG pathway and the PPs, tolerance to elevated levels of signalling components was largely independent of shifts in environmental conditions (**Paper III, Fig. 3**).

The observation of environmental independence of fragility patterns begged the question whether a similar independence also persisted with regards to fluctuations in genotypic space. To test that hypothesis, genetic manipulations were performed by removing components in vicinity to the signalling modules, e.g. transcription factors immediately downstream of the modules, followed by measurements of fitness effects upon overexpression of pathway components in these genetically manipulated strains (**Paper III, Fig. 4**). While a number of significant genetic interactions, in terms of suppression or aggravation of the original overexpression phenotype, were observed, the strength of most of the interactions was limited. It suggests that signalling pathway robustness to a large extent is independent of position in genotypic space, and also that crosstalk between signalling pathways seems to be overall limited.

6 CONCLUDING REMARKS

In conclusion,

I have in this thesis studied effects on cellular fitness in response to biological perturbations, both intracellular (genetic perturbations) and extracellular (environmental variation). Although a complex aspect of biology, cellular fitness has been considered equivalent to growth rate on the level of a cell population.

The first aim was to investigate the genetic basis of toxicity to the metalloid tellurite in yeast through gene-by-environment interactions (**Paper I**). By screening a genome-wide collection of gene deletion mutants and quantitatively measure fitness defects and accumulation of the reduced form of tellurite, elemental tellurium, genes belonging to the sulfate assimilation pathway were shown to be central to tellurite toxicity and its reduction to tellurium. It was also demonstrated that tellurite toxicity was partially mediated via a mitochondrial respiratory mechanism.

The second aim of this thesis was to characterize the fitness landscape in *S. cerevisiae* in response to gene overexpression of signalling components to a high resolution, with an emphasis on signal transduction pathways (**Papers II-III**) and protein phosphatases (**Paper IV**). This was accomplished using a gene overexpression method called genetic tug-of-war (gToW) that increases gene expression through an increase in gene dosage while retaining native gene regulation.

A high degree of negative impacts on cellular fitness was observed upon overexpression of signalling components. This observation stands in stark contrast to the overall system resilience to deletions of the same components. These fragile nodes were also dispersed over the different classes of signalling components analysed, e.g. protein kinases, protein phosphatases and transcription factors. They were also widely dispersed in signalling pathways with no clear patterns as to biochemical function, spatial or topological positions.

A likely culprit of the observed overexpression toxicity seems to stem from both an increased abundance of proteins in the cell, but also from an increased catalytic activity of the signalling components, e.g. phosphorylation and dephosphorylation. Surprisingly, fragility towards signal component overexpression was found to be largely independent of

environmental as well as genotypic fluctuations. This finding would suggest that fragile nodes are a product of local signalling network architecture and that crosstalk between signalling pathways is overall limited.

It has also been demonstrated that the use of gain-of-function alleles, besides from being an important tool in characterising the fitness landscape to a high-resolution upon overexpression, has the potential to unravel new interactions and links between components of cellular processes (**Paper IV**).

6.1 Future outlook for phenomics

Only a decade ago it was debated whether phenomics would be a concept for the future (Gerlai 2002). Ten years may at first seem a short time period, but on the level of technological innovation at lot has happened since 2002. We now have the technological know-how to create sophisticated devices to be used in phenotypic experiments, with increasing scale and resolution. It has become increasingly clear that phenomics is here to stay and also prove the next big challenge in modern biological research (Houle et al. 2010). Many scientists argue that phenomics thus naturally should pick up where genomics to some extent left of, as exemplified by a multitude of ongoing phenome projects in various organisms (Houle et al. 2010).

A task for future technological development for high-throughput phenotyping will be to maximise the number of phenotypes that can be measured simultaneously with as large sample sizes as possible without compromising reproducibility and precision (Houle et al. 2010). Especially the latter, precision, is of utmost importance since many phenotypes are only marginal, calling upon quantification of phenotypes with high resolution. That is particularly evident for the measurements of growth effects stemming from gene deletions that are frequently only marginal, even in presence of environmental perturbations. Throughout **Papers II-IV**, a liquid microcultivation approach was adopted with high precision and resolution. This strategy, however, is not suited for large-scale genome-wide phenotypic studies as it would be very time-consuming to conduct these experiments.

In **Paper I** a genome-wide study on the cellular responses to a metalloid compound was performed. To accomplish this, an alternative approach was used where gene deletion mutants were pinned onto the surface of agar plates in a grid format (array) and cultivated

for a fixed amount of time before images were taken using a digital camera. Growth phenotypes were then manually scored by visual inspection of the colonies. Although experiments of this kind are easy to perform with thousands of parallel samples being measured, the manual evaluation is subjective and could lead to errors in the judgement of individual phenotypes. Another problem with this strategy is a lack of resolution in terms of individual features of growth as the colony size after a defined period of time will be a composite of the growth variables; lag, rate and efficiency (Warringer et al. 2008).

To address this we are currently developing a workflow for high-throughput time-resolved phenotyping of microbial growth rate on solid media. In short, at the core of the system is an ordinary flatbed scanner in which four plates with yeast colony arrays are incubated in a fixture. The scanner is connected to a computer that controls its activity through in-house software to scan the plates at user-defined time intervals, typically 20 min to mimic the regime used throughout **Papers II-IV** for liquid microcultivation. Through this effort we hope to be able to generate time-resolved high-resolution growth curves by automatically extract colony area over time for thousands of parallel experiments. In that sense we attempt at establishing a phenotyping strategy that, to some extent, combines the benefits of liquid microcultivation and growth phenotyping on plates.

If successful, this approach will be beneficial to the field of phenomics as genome-wide gene perturbation mutant collections can be screened with relative ease and in a short period of time. It can also prove fruitful to the system-wide examination of fitness effects caused by gene perturbations.

7 ACKNOWLEDGEMENTS

Pursuing a PhD is not an easy task but given what I now know and have experienced during the past 6 years, I wouldn't want to have it any other way. Today, I get to write these very important words in the final paragraph of this thesis in large part because many people helped me, cared for me, supported me or in any other way walked next to me during my journey. These persons have meant all the difference to me, without whom none of this, or anything, would have been possible. In this section I want to acknowledge you.

For reasons of space, this list could never be complete, but for those not mentioned I want you to know that you are not forgotten! Where appropriate, the address will be given in Swedish.

First of all I would like to thank my supervisors **Anders** and **Stefan** for giving me the opportunity to pursue a career in science. Your input and guidance have been essential to the creation of this thesis. I also want to thank **Hiroaki Kitano** for giving me the opportunity to work in Tokyo and experience this very interesting and energetic city. I will definitely return to Japan one day!

Special thanks to **Jonas** and **Marcus** for all you help and assistance these years. I cannot accurately describe in words how much your help has meant to me.

Doryaneh: My "partner in crime" from the very beginning. We shared all of the initial struggles and successes and I wish you all the best for the future!

Markus: For nice discussions about science and for being so enthusiastic. I still owe you quite a number of photographs from various conferences. Soon there will be time for a photo session and a couple of cold ones!

Cissi: Our friendship goes back a long time, all the way back to the golden days of "Grindstugan" and student misbehaviour! =) Ahh, those were the days! How about an after work soon?

All members, past and present, of the Blomberg and Hohmann research groups. In particular I want to acknowledge the following persons; **Ellinor** and **Peter** for all your help in the lab, **Calle** and **Jimmy** for "Late Nights" with beer and discussions, **Maria** for being so friendly and cheerful, **Jonathan** (former cubicle mate), **Enikö** (present cubicle mate), **Martin**, **Magnus** and **Ulrika.**

8 REFERENCES

Adams, M. D., et al. (2000). "The genome sequence of Drosophila melanogaster." Science **287**(5461): 2185-95.

Babazadeh, R., et al. (2011). "The Ashbya gossypii EF-1alpha promoter of the ubiquitously used MX cassettes is toxic to Saccharomyces cerevisiae." FEBS Lett **585**(24): 3907-13.

Bardwell, L. (2004). "A walk-through of the yeast mating pheromone response pathway." Peptides **25**(9): 1465-76.

Blattner, F. R., et al. (1997). "The complete genome sequence of Escherichia coli K-12." Science **277**(5331): 1453-62.

Bleackley, M. R., et al. (2011). "High density array screening to identify the genetic requirements for transition metal tolerance in Saccharomyces cerevisiae." Metallomics **3**(2): 195-205.

Blomberg, A. (2011). "Measuring growth rate in high-throughput growth phenotyping." Curr Opin Biotechnol **22**(1): 94-102.

Bossemeyer, D. (1995). "Protein kinases--structure and function." FEBS Lett **369**(1): 57-61.

Brewster, J. L., et al. (1993). "An osmosensing signal transduction pathway in yeast." Science **259**(5102): 1760-3.

Burke, D., et al. (1989). "Dominant effects of tubulin overexpression in Saccharomyces cerevisiae." Mol Cell Biol **9**(3): 1049-59.

Butinar, L., et al. (2007). "Yeasts in high Arctic glaciers: the discovery of a new habitat for eukaryotic microorganisms." Antonie Van Leeuwenhoek **91**(3): 277-89.

Caesar, R. and A. Blomberg (2004). "The stress-induced Tfs1p requires NatB-mediated acetylation to inhibit carboxypeptidase Y and to regulate the protein kinase A pathway." J Biol Chem **279**(37): 38532-43.

Cavalieri, D., et al. (2003). "Evidence for S. cerevisiae fermentation in ancient wine." J Mol Evol **57 Suppl 1**: S226-32.

Cheng, A., et al. (1999). "Dephosphorylation of cyclin-dependent kinases by type 2C protein phosphatases." Genes Dev **13**(22): 2946-57.

Chi, A., et al. (2007). "Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry." Proc Natl Acad Sci U S A **104**(7): 2193-8.

Ciesla, J., et al. (2011). "Phosphorylation of basic amino acid residues in proteins: important but easily missed." Acta Biochim Pol **58**(2): 137-48.

Cohen, S. N., et al. (1973). "Construction of biologically functional bacterial plasmids in vitro." Proc Natl Acad Sci U S A **70**(11): 3240-4.

Costanzo, M., et al. (2010). "The genetic landscape of a cell." Science **327**(5964): 425-31.

Dard, N. and M. Peter (2006). "Scaffold proteins in MAP kinase signaling: more than simple passive activating platforms." Bioessays **28**(2): 146-56.

Das, A. K., et al. (1996). "Crystal structure of the protein serine/threonine phosphatase 2C at 2.0 A resolution." Embo J **15**(24): 6798-809.

Dittmar, J. C., et al. (2010). "ScreenMill: a freely available software suite for growth measurement, analysis and visualization of high-throughput screen data." BMC Bioinformatics **11**: 353.

Erhart, E. and C. P. Hollenberg (1983). "The presence of a defective LEU2 gene on 2 mu DNA recombinant plasmids of Saccharomyces cerevisiae is responsible for curing and high copy number." J Bacteriol **156**(2): 625-35.

Etcheverry, T. (1990). "Induced expression using yeast copper metallothionein promoter." Methods Enzymol **185**: 319-29.

Fernandez-Ricaud, L., et al. (2007). "PROPHECY--a yeast phenome database, update 2006." Nucleic Acids Res **35**(Database issue): D463-7.

Fernandez-Ricaud, L., et al. (2005). "PROPHECY--a database for high-resolution phenomics." Nucleic Acids Res **33**(Database issue): D369-73.

Ficarro, S. B., et al. (2002). "Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae." Nat Biotechnol **20**(3): 301-5.

Fiedler, D., et al. (2009). "Functional organization of the S. cerevisiae phosphorylation network." Cell **136**(5): 952-63.

Forsburg, S. L. (2001). "The art and design of genetic screens: yeast." Nat Rev Genet **2**(9): 659-68.

Gadanho, M. and J. P. Sampaio (2005). "Occurrence and diversity of yeasts in the mid-atlantic ridge hydrothermal fields near the Azores Archipelago." Microb Ecol **50**(3): 408-17.

Gasch, A. P., et al. (2000). "Genomic expression programs in the response of yeast cells to environmental changes." Mol Biol Cell **11**(12): 4241-57.

Gerlai, R. (2002). "Phenomics: fiction or the future?" Trends Neurosci **25**(10): 506-9.

Ghaemmaghami, S., et al. (2003). "Global analysis of protein expression in yeast." Nature **425**(6959): 737-41.

Ghosh, S. K., et al. (2006). "Mechanisms for chromosome and plasmid segregation." Annu Rev Biochem **75**: 211-41.

Giaever, G., et al. (2002). "Functional profiling of the Saccharomyces cerevisiae genome." Nature **418**(6896): 387-91.

Gilbert, D., et al. (2006). "Computational methodologies for modelling, analysis and simulation of signalling networks." Brief Bioinform **7**(4): 339-53.

Goffeau, A. (2000). "Four years of post-genomic life with 6,000 yeast genes." FEBS Lett **480**(1): 37-41.

Goffeau, A., et al. (1996). "Life with 6000 genes." Science **274**(5287): 546, 563-7.

Granek, J. A. and P. M. Magwene (2010). "Environmental and genetic determinants of colony morphology in yeast." PLoS Genet **6**(1): e1000823.

Gruhler, A., et al. (2005). "Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway." Mol Cell Proteomics **4**(3): 310-27.

Gu, Z., et al. (2003). "Role of duplicate genes in genetic robustness against null mutations." Nature **421**(6918): 63-6.

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

Hanks, S. K., et al. (1988). "The protein kinase family: conserved features and deduced phylogeny of the catalytic domains." Science **241**(4861): 42-52.

Hartwell, L. H., et al. (1999). "From molecular to modular cell biology." Nature **402**(6761 Suppl): C47-52.

Hillenmeyer, M. E., et al. (2008). "The chemical genomic portrait of yeast: uncovering a phenotype for all genes." Science **320**(5874): 362-5.

Hohmann, S. (2002). "Osmotic stress signaling and osmoadaptation in yeasts." Microbiol Mol Biol Rev **66**(2): 300-72.

Honigberg, S. M. (2011). "Cell signals, cell contacts, and the organization of yeast communities." Eukaryot Cell **10**(4): 466-73.

Houle, D., et al. (2010). "Phenomics: the next challenge." Nat Rev Genet **11**(12): 855-66.

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

Hunter, T. and G. D. Plowman (1997). "The protein kinases of budding yeast: six score and more." Trends Biochem Sci **22**(1): 18-22.

Johnston, M. and R. W. Davis (1984). "Sequences that regulate the divergent GAL1-GAL10 promoter in Saccharomyces cerevisiae." Mol Cell Biol **4**(8): 1440-8.

Kaizu, K., et al. (2010). "Fragilities caused by dosage imbalance in regulation of the budding yeast cell cycle." PLoS Genet **6**(4): e1000919.

Kaluarachchi Duffy, S., et al. (2012). "Exploring the yeast acetylome using functional genomics." Cell **149**(4): 936-48.

Kellis, M., et al. (2004). "Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae." Nature **428**(6983): 617-24.

Kitano, H. (2004). "Biological robustness." Nat Rev Genet **5**(11): 826-37.

Kohl, P., et al. (2010). "Systems biology: an approach." Clin Pharmacol Ther **88**(1): 25-33.

Kolmodin, K. and J. Aqvist (2001). "The catalytic mechanism of protein tyrosine phosphatases revisited." FEBS Lett **498**(2-3): 208-13.

Krantz, M., et al. (2009). "Robustness and fragility in the yeast high osmolarity glycerol (HOG) signal-transduction pathway." Mol Syst Biol **5**: 281.

Kutty, S. N. and R. Philip (2008). "Marine yeasts-a review." Yeast **25**(7): 465-83.

Lamprecht, M. R., et al. (2007). "CellProfiler: free, versatile software for automated biological image analysis." Biotechniques **42**(1): 71-5.

Lander, E. S., et al. (2001). "Initial sequencing and analysis of the human genome." Nature **409**(6822): 860-921.

Lawless, C., et al. (2010). "Colonyzer: automated quantification of micro-organism growth characteristics on solid agar." BMC Bioinformatics **11**: 287.

Lee, S. W., et al. (2000). "Overexpression of kinase-associated phosphatase (KAP) in breast and prostate cancer and inhibition of the transformed phenotype by antisense KAP expression." Mol Cell Biol **20**(5): 1723-32.

Levin, D. E. (2011). "Regulation of Cell Wall Biogenesis in Saccharomyces cerevisiae: The Cell Wall Integrity Signaling Pathway." Genetics **189**(4): 1145-75.

Li, Z., et al. (2011). "Systematic exploration of essential yeast gene function with temperature-sensitive mutants." Nat Biotechnol **29**(4): 361-7.

Liti, G., et al. (2009). "Population genomics of domestic and wild yeasts." Nature **458**(7236): 337-41.

Macia, J., et al. (2009). "Dynamic signaling in the Hog1 MAPK pathway relies on high basal signal transduction." Sci Signal **2**(63): ra13.

Manning, G., et al. (2002). "Evolution of protein kinase signaling from yeast to man." Trends Biochem Sci **27**(10): 514-20.

Martin, H., et al. (2005). "Protein phosphatases in MAPK signalling: we keep learning from yeast." Mol Microbiol **58**(1): 6-16.

Maxam, A. M. and W. Gilbert (1977). "A new method for sequencing DNA." Proc Natl Acad Sci U S A **74**(2): 560-4.

Memarian, N., et al. (2007). "Colony size measurement of the yeast gene deletion strains for functional genomics." BMC Bioinformatics **8**: 117.

Mnaimneh, S., et al. (2004). "Exploration of essential gene functions via titratable promoter alleles." Cell **118**(1): 31-44.

Moorhead, G. B., et al. (2009). "Evolution of protein phosphatases in plants and animals." Biochem J **417**(2): 401-9.

Moriya, H., et al. (2011). "Overexpression limits of fission yeast cell-cycle regulators in vivo and in silico." Mol Syst Biol **7**: 556.

Moriya, H., et al. (2006). "In vivo robustness analysis of cell division cycle genes in Saccharomyces cerevisiae." PLoS Genet **2**(7): e111.

Mortimer, R. K. and J. R. Johnston (1986). "Genealogy of principal strains of the yeast genetic stock center." Genetics **113**(1): 35-43.

Mumberg, D., et al. (1994). "Regulatable promoters of Saccharomyces cerevisiae: comparison of transcriptional activity and their use for heterologous expression." Nucleic Acids Res **22**(25): 5767-8.

O'Rourke, S. M. and I. Herskowitz (2002). "A third osmosensing branch in Saccharomyces cerevisiae requires the Msb2 protein and functions in parallel with the Sho1 branch." Mol Cell Biol **22**(13): 4739-49.

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**(2): 532-42.

Olsen, J. V., et al. (2006). "Global, in vivo, and site-specific phosphorylation dynamics in signaling networks." Cell **127**(3): 635-48.

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

Pincus, D., et al. (2008). "Evolution of the phospho-tyrosine signaling machinery in premetazoan lineages." Proc Natl Acad Sci U S A **105**(28): 9680-4.

Prelich, G. (2012). "Gene overexpression: uses, mechanisms, and interpretation." Genetics **190**(3): 841-54.

Proft, M., et al. (2005). "Genomewide identification of Sko1 target promoters reveals a regulatory network that operates in response to osmotic stress in Saccharomyces cerevisiae." Eukaryot Cell **4**(8): 1343-52.

Ptacek, J., et al. (2005). "Global analysis of protein phosphorylation in yeast." Nature **438**(7068): 679-84.

Qi, M. and E. A. Elion (2005). "MAP kinase pathways." J Cell Sci **118**(Pt 16): 3569-72.

Ray, D. and H. Kiyokawa (2008). "CDC25A phosphatase: a rate-limiting oncogene that determines genomic stability." Cancer Res **68**(5): 1251-3.

Rine, J. (1991). "Gene overexpression in studies of Saccharomyces cerevisiae." Methods Enzymol **194**: 239-51.

Rose, A. B. and J. R. Broach (1990). "Propagation and expression of cloned genes in yeast: 2-microns circle-based vectors." Methods Enzymol **185**: 234-79.

Saito, H. and K. Tatebayashi (2004). "Regulation of the osmoregulatory HOG MAPK cascade in yeast." J Biochem **136**(3): 267-72.

Sanger, F., et al. (1977). "DNA sequencing with chain-terminating inhibitors." Proc Natl Acad Sci U S A **74**(12): 5463-7.

Santarius, T., et al. (2010). "A census of amplified and overexpressed human cancer genes." Nat Rev Cancer **10**(1): 59-64.

Sauer, U., et al. (2007). "Genetics. Getting closer to the whole picture." Science **316**(5824): 550-1.

Scherens, B. and A. Goffeau (2004). "The uses of genome-wide yeast mutant collections." Genome Biol **5**(7): 229.

Schneider, J. C. and L. Guarente (1991). "Vectors for expression of cloned genes in yeast: regulation, overproduction, and underproduction." Methods Enzymol **194**: 373-88.

Semple, J. I., et al. (2008). "A simple principle concerning the robustness of protein complex activity to changes in gene expression." BMC Syst Biol **2**: 1.

Shastry, B. S. (1995). "Overexpression of genes in health and sickness. A bird's eye view." Comp Biochem Physiol B Biochem Mol Biol **112**(1): 1-13.

Sherman, F. (2002). "Getting started with yeast." Methods Enzymol **350**: 3-41.

Smith, L. M., et al. (1986). "Fluorescence detection in automated DNA sequence analysis." Nature **321**(6071): 674-9.

Sniegowski, P. D., et al. (2002). "Saccharomyces cerevisiae and Saccharomyces paradoxus coexist in a natural woodland site in North America and display different levels of reproductive isolation from European conspecifics." FEMS Yeast Res **1**(4): 299-306.

Sopko, R., et al. (2006). "Mapping pathways and phenotypes by systematic gene overexpression." Mol Cell **21**(3): 319-30.

Springer, M., et al. (2010). "A general lack of compensation for gene dosage in yeast." Mol Syst Biol **6**: 368.

Stelling, J., et al. (2004). "Robustness of cellular functions." Cell **118**(6): 675-85.

Strange, K. (2005). "The end of "naive reductionism": rise of systems biology or renaissance of physiology?" Am J Physiol Cell Physiol **288**(5): C968-74.

Tatebayashi, K., et al. (2007). "Transmembrane mucins Hkr1 and Msb2 are putative osmosensors in the SHO1 branch of yeast HOG pathway." Embo J **26**(15): 3521-33.

Tong, A. H. and C. Boone (2006). "Synthetic genetic array analysis in Saccharomyces cerevisiae." Methods Mol Biol **313**: 171-92.

Tong, A. H., et al. (2001). "Systematic genetic analysis with ordered arrays of yeast deletion mutants." Science **294**(5550): 2364-8.

Tong, A. H., et al. (2004). "Global mapping of the yeast genetic interaction network." Science **303**(5659): 808-13.

Vavouri, T., et al. (2009). "Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity." Cell **138**(1): 198-208.

Venter, J. C., et al. (2001). "The sequence of the human genome." Science **291**(5507): 1304-51.

Virshup, D. M. and S. Shenolikar (2009). "From promiscuity to precision: protein phosphatases get a makeover." Mol Cell **33**(5): 537-45.

Walsh, C. T., et al. (2005). "Protein posttranslational modifications: the chemistry of proteome diversifications." Angew Chem Int Ed Engl **44**(45): 7342-72.

Wang, H., et al. (2009). "A complex-based reconstruction of the Saccharomyces cerevisiae interactome." Mol Cell Proteomics **8**(6): 1361-81.

Wang, L., et al. (2012). "Integrating phosphorylation network with transcriptional network reveals novel functional relationships." PLoS One **7**(3): e33160.

Warringer, J., et al. (2008). "Chemogenetic fingerprinting by analysis of cellular growth dynamics." BMC Chem Biol **8**: 3.

Warringer, J. and A. Blomberg (2003). "Automated screening in environmental arrays allows analysis of quantitative phenotypic profiles in Saccharomyces cerevisiae." Yeast **20**(1): 53-67.

Warringer, J., et al. (2003). "High-resolution yeast phenomics resolves different physiological features in the saline response." Proc Natl Acad Sci U S A **100**(26): 15724-9.

Watson, J. D. and F. H. Crick (1953). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid." Nature **171**(4356): 737-8.

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

Williams, N. (1996). "Yeast genome sequence ferments new research." Science **272**(5261): 481.

Winzeler, E. A., et al. (1999). "Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis." Science **285**(5429): 901-6.

Wolfe, K. H. and D. C. Shields (1997). "Molecular evidence for an ancient duplication of the entire yeast genome." Nature **387**(6634): 708-13.

Wurgler-Murphy, S. M., et al. (1997). "Regulation of the Saccharomyces cerevisiae HOG1 mitogen-activated protein kinase by the PTP2 and PTP3 protein tyrosine phosphatases." Mol Cell Biol **17**(3): 1289-97.

Wysocki, R. and M. J. Tamas (2010). "How Saccharomyces cerevisiae copes with toxic metals and metalloids." FEMS Microbiol Rev **34**(6): 925-51.

Yaakov, G., et al. (2009). "The stress-activated protein kinase Hog1 mediates S phase delay in response to osmostress." Mol Biol Cell **20**(15): 3572-82.

Yoshikawa, K., et al. (2011). "Comprehensive phenotypic analysis of single-gene deletion and overexpression strains of Saccharomyces cerevisiae." Yeast **28**(5): 349-61.

Zaman, S., et al. (2008). "How Saccharomyces responds to nutrients." Annu Rev Genet **42**: 27-81.

Zhu, H., et al. (2001). "Global analysis of protein activities using proteome chips." Science **293**(5537): 2101-5.

Zhu, H., et al. (2000). "Analysis of yeast protein kinases using protein chips." Nat Genet **26**(3): 283-9.