Title
Quantitative analysis of nutrient-responsive signal transduction pathways in Saccharomyces cerevisiae

Author
Raser, Jonathan M

Publication Date
2005

Peer reviewed|Thesis/dissertation

Quantitative analysis of nutrient-responsive signal transduction pathways in *Saccharomyces cerevisiae*

by

Jonathan M. Raser

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

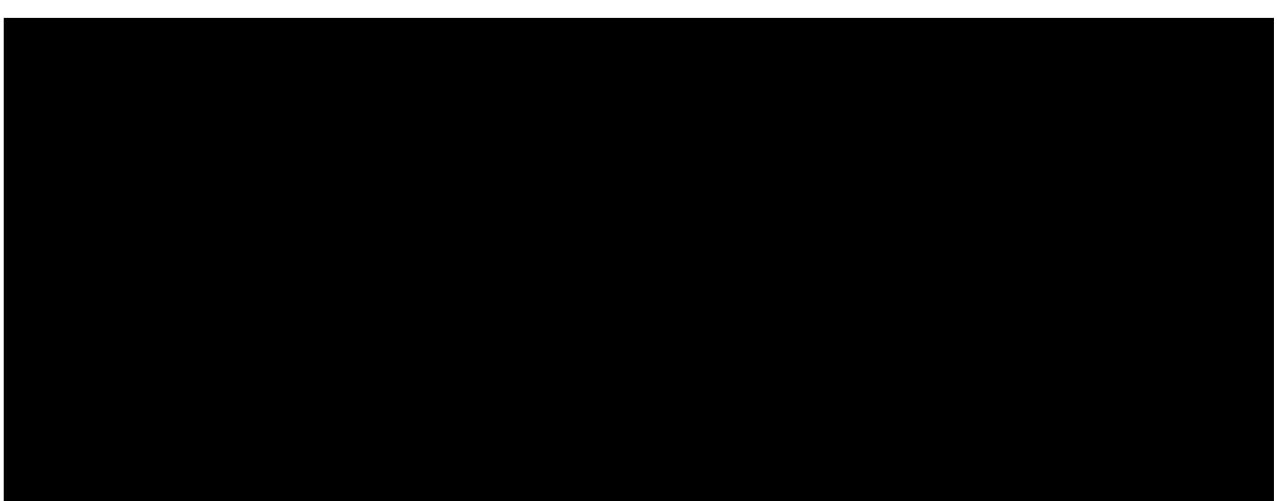
Biochemistry and Molecular Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO



Date University Librarian

Degree Conferred:.....

This thesis is dedicated to my family, present and future.

ACKNOWLEDGMENTS

I thank my advisor, Erin O'Shea, for all that she has taught me and for all the opportunities she has provided me. I am grateful to those past advisors who have supported my scientific education, including Mark Solomon, Mahboubeh Eghbali-Webb, and especially Jim and Anita Hopper, who continue to encourage my scientific work. I am thankful to have worked with and learned so much from the wonderful members of the O'Shea, Solomon, Panning, and Bargmann laboratories, including but not limited to Amy Bany, Deb Enke, Karen Ross, Beth Egan, Janet Burton, and Jen Holmes in the Solomon lab; and Dennis Wykoff, Liz Haswell, Arash Komeili, Marc Lenburg, Adam Carroll, Doug Jeffery, Noah Dephoure, Meg Byrne, Narendra Maheshri, Archana Belle, Rosa Marion, Andrew Capaldi, Melissa Thomas, Nicole Miller, Rusty Howson, Walter Lau, Susan Shih, Dave Steger, Reed Kelso, Eileen Wu, Brian Margolin, James Falvo, Felix Lam, Sid Huang, Won-ki Huh, and Michael Springer of Team O'Shea.

I thank Hiten Madhani and Jonathan Weissman for their contributions as members of my thesis committee; Jana and Catherine of the MSTP for their continual support of my broader curriculum; and the UCSF Medical Scientist Training Program, the American Diabetes Association, and the Burroughs Wellcome/UCSF Interfaces of Science Program for financial support.

I thank also my friends and family, especially my parents and brother and sister, who have cheerfully endured my absence, distraction and occasional, brief periods of despair. In particular, I must thank Dennis Wykoff for the many hours of scientific conversations mixed with humor, and humorous conversations mixed with a little science, without which I might never have completed this phase of my education.

License to Publish

The following Grant of License ("License") must be signed and returned to the American Association for the Advancement of Science ("AAAS") before a manuscript can be accepted for publication. If the contribution to SCIENCE, described below, is a joint work, AAAS requires that each joint author sign this License or elect one author to sign it as their duly authorized agent. If the copyright in the contribution is owned by the author's employer, the employer or an authorized representative must sign this form. In the event that AAAS decides not to publish the Work in SCIENCE, this License shall be null and void.

1. In consideration of publication in SCIENCE of the work currently titled Noise in gene expression: origins, control, and consequences (the "Work") and authored by Jonathan M. Raser ("Author"), the sole and exclusive, irrevocable right is hereby granted to AAAS to publish, reproduce, distribute, transmit, display, store, translate, create derivative works from and otherwise use the Work in any form, manner, format, or medium, whether now known or hereafter developed, throughout the world and in any language, for the entire duration of any such right and any renewal or extension thereof and to permit/sublicense others to do any or all of the foregoing as well.

2. Ownership of the copyright shall remain with the Author, subject to the rights granted to AAAS in paragraph 1. AAAS shall be the holder of title for purpose of registration.

3. Author also retains the non-exclusive right to use the Work in the following ways without further permission but only after publication of the Work in SCIENCE and subject to the requirement that credit be given to its first publication in the appropriate issue of SCIENCE: 1) Reprint the Work in print collections of Author's own writings; 2) Reprint the Work in print format for inclusion in a thesis or dissertation that the author writes; 3) Present the Work orally; 4) Reproduce the Work for use in courses the Author is teaching (if the author is employed by an academic institution, that institution may also reproduce the Work for course teaching); 5) Distribute photocopies of the Work to colleagues for non-commercial purposes only (providing that recipients are informed that they may not further distribute or copy the Work); 6) Post a copy of the accepted version of the Work (the accepted version of the paper before Science's copy editing and production) on the author's personal website, provided a hyperlink to the Work on the SCIENCE web site is included; 7) Reuse figures and tables created by the Author in future works the Author writes; 8) If the Work is prepared as a work made for hire, the author's employer may make photocopies for internal use only; 9) Author may use or authorize use of Supporting Online Material associated with the Work for any purpose and in any format.

4. In addition to the rights retained by the author in 3 above, For Works that are both (A) created under a grant from NIH and (B) accepted for publication in Science on or after 2 May 2005: The author may cause the posting, no sooner than 6 months after final publication of the Work in Science, of the "Accepted Version" of the Work in PMC. The "Accepted Version" shall be defined as the version of the paper accepted for publication in Science including changes resulting from peer review but prior to Science's copy editing and production.

5. In order that we may preserve the commercial value of our publication, Author agrees to obtain authorization from AAAS in advance for any uses not expressly described in paragraphs 3 or 4.

6. Author warrants and represents that the Work is original, that all the facts contained therein are true and accurate, and that the Work has not been published elsewhere, and does not infringe upon any copyright, proprietary, or personal right of any third party. If the Work contains any material that is owned or controlled by a third party, Author certifies that he/she has obtained permission for its use (please attach evidence of this) and that the material is clearly acknowledged within the text.

By signing this agreement, Author warrants that he/she has the full power to enter into this agreement. This agreement shall remain in effect throughout the term of copyright in the Work and may not be revoked without the express written consent of both parties.

This Agreement shall be governed and construed, and any dispute arising hereunder resolved, in accordance with the laws of the District of Columbia, United States of America, without resort to the conflicts of laws principles thereof.

Jonathan M. Raser

Author's Name (please print)

Author's Signature

For Authors employed by the United States Government [This section to be signed only if ALL authors are U.S. Government employees]: I attest that the above article was written as part of the official duties of the authors as employees of the U.S. Government and therefore the article is in the public domain.

Author's Name (please print)

Author's Signature

For works created under U.S. Government Contract: Please sign the top portion of this agreement. The AAAS recognizes the U.S. Government's non exclusive rights to use the Work for non commercial, governmental purposes where such rights are established in the grant or contract.
Works subject to Crown Copyright: If ALL authors are employees of the Australian, Canadian, and/or UK Governments, please contact AAAS for an alternate form.

License to Publish

The following Grant of License ("License") must be signed and returned to the American Association for the Advancement of Science ("AAAS") before a manuscript can be accepted for publication. If the contribution to SCIENCE, described below, is a joint work, AAAS requires that each joint author sign this License or elect one author to sign it as their duly authorized agent. If the copyright in the contribution is owned by the author's employer, the employer or an authorized representative must sign this form. In the event that AAAS decides not to publish the Work in SCIENCE, this License shall be null and void.

1. In consideration of publication in SCIENCE of the work currently titled Control of stochasticity in eukaryotic gene expression (the "Work") and authored by Jonathan M. Raser ("Author"), the sole and exclusive, irrevocable right is hereby granted to AAAS to publish, reproduce, distribute, transmit, display, store, translate, create derivative works from and otherwise use the Work in any form, manner, format, or medium, whether now known or hereafter developed, throughout the world and in any language, for the entire duration of any such right and any renewal or extension thereof and to permit/sublicense others to do any or all of the foregoing as well.

2. Ownership of the copyright shall remain with the Author, subject to the rights granted to AAAS in paragraph 1. AAAS shall be the holder of title for purpose of registration.

3. Author also retains the non-exclusive right to use the Work in the following ways without further permission but only after publication of the Work in SCIENCE and subject to the requirement that credit be given to its first publication in the appropriate issue of SCIENCE: 1) Reprint the Work in print collections of Author's own writings; 2) Reprint the Work in print format for inclusion in a thesis or dissertation that the author writes; 3) Present the Work orally; 4) Reproduce the Work for use in courses the Author is teaching (if the author is employed by an academic institution, that institution may also reproduce the Work for course teaching); 5) Distribute photocopies of the Work to colleagues for non-commercial purposes only (providing that recipients are informed that they may not further distribute or copy the Work); 6) Post a copy of the accepted version of the Work (the accepted version of the paper before Science's copy editing and production) on the author's personal website, provided a hyperlink to the Work on the SCIENCE web site is included; 7) Reuse figures and tables created by the Author in future works the Author writes; 8) If the Work is prepared as a work made for hire, the author's employer may make photocopies for internal use only; 9) Author may use or authorize use of Supporting Online Material associated with the Work for any purpose and in any format.

4. In order that we may preserve the commercial value of our publication, Author agrees to obtain authorization from AAAS in advance for any uses not expressly described in paragraph 3.

5. Author warrants and represents that the Work is original, that all the facts contained therein are true and accurate, and that the Work has not been published elsewhere and does not infringe upon any copyright, proprietary, or personal right of any third party. If the Work contains any material that is owned or controlled by a third party, Author certifies that he/she has obtained permission for its use (please attach evidence of this) and that the material is clearly acknowledged within the text.

By signing this agreement, Author warrants that he/she has the full power to enter into this agreement. This agreement shall remain in effect throughout the term of copyright in the Work and may not be revoked without the express written consent of both parties.

This Agreement shall be governed and construed, and any dispute arising hereunder resolved, in accordance with the laws of the District of Columbia, United States of America, without resort to the conflicts of laws principles thereof.

Jonathan M. Raser

Author's Name (please print)

Author's Signature

For Authors employed by the United States Government [This section to be signed only if ALL authors are U.S. Government employees]: I attest that the above article was written as part of the official duties of the authors as employees of the U.S. Government and therefore the article is in the public domain.

Author's Name (please print)

Author's Signature

For works created under U.S. Government Contract: Please sign the top portion of this agreement. The AAAS recognizes the U.S. Government's non-exclusive rights to use the Work for non-commercial, governmental purposes where such rights are established in the grant or contract.

Works subject to Crown Copyright: If ALL authors are employees of the Australian, Canadian, and/or UK Governments, please contact AAAS for an alternate form.

QUANTITATIVE ANALYSIS OF NUTRIENT-RESPONSIVE SIGNAL TRANSDUCTION PATHWAYS IN *SACCHAROMYCES CEREVISIAE*

Jonathan M. Raser

Erin K. O'Shea

Organisms and their component cells exist in a dynamic and unpredictable environment. Many of these environmental changes are potentially harmful or lethal. It is therefore unsurprising that cells are capable of adaptation to many of these potentially harmful environmental conditions. This adaptation requires both the ability to gather information about the environment, and the ability to convert the information gathered into the appropriate cellular response. Cells have developed specific signal transduction pathways that respond to a specific environmental stimulus and effect a specific, corresponding cellular response. Such signal transduction pathways are quantitative, in that cells do not simply display binary information processing but are capable of more subtle interpretation of different magnitudes of a particular stimulus. In particular, many signaling pathways enable adaptation to changing levels of environmental nutrients by control of cellular gene expression. We employed the single-cell eukaryote *Saccharomyces cerevisiae* as a model to study several facets of nutrient-responsive signal transduction in a quantitative manner. In the following text, we present our study of the sources of heterogeneity in gene expression in a population of genetically identical cells. In addition, we review recent advances in understanding of heterogeneity or noise in gene expression. Also, we present work characterizing the zinc-responsive signaling

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION:

Nutrient-responsive signal transduction pathways

The cellular response to a dynamic environment

Individual cells exhibit a remarkable degree of control of the conditions that exist within the membranes that demarcate the borders between the intracellular and extracellular spaces. Cells exert impressive control over the bulk chemical composition, the spatial organization, and the physical properties, such as temperature and pressure, of the intracellular space. This control is maintained despite a changing, uncertain extracellular environment. In many cases, the intracellular aspect remains unchanged despite significant alterations in the extracellular environment. This tendency to minimize intracellular changes can be labeled "homeostasis." It can be broadly stated that cells expend energy in order to maintain an intracellular environment that is favorable for continued growth and reproduction. Often, the maintenance of ideal intracellular conditions, or cellular homeostasis, in the face of dramatic extracellular environmental changes requires a significant alteration in the composition or activity of cellular constituents. The term "adaptation" is used to refer to the beneficial cellular response to environmental changes that allows the cell to continue to grow and reproduce.

In some circumstances, especially for multicellular organisms, the response to a particular environmental change is a dramatic change in intracellular or intraorganismal conditions. For example, the sporulation program in *Bacillus subtilis* (Piggot and Hilbert, 2004) or the dauer transition of *Caenorhabditis elegans* (Vanfleteren and Braeckman, 1999) represent a non-homeostatic adaptation to changing environmental conditions, in which the cell or organism enters a state of reduced metabolic activity and reproductive potential. This qualitatively distinct state is adaptive or beneficial on a longer timescale, as it allows the organism to survive extremely harsh and otherwise lethal environmental

conditions until the environment returns to conditions more favorable for growth and reproduction. However, many environmental changes result in less dramatic adaptive cellular responses that enable the cell to maintain a homeostatic intracellular environment, i.e. to continue to grow and divide in the new extracellular environment.

Signal transduction pathway components

A "signal transduction pathway" is the set of cellular components responsible for sensing changes in the extracellular environment and producing the appropriate response to the specific changes sensed. Signal transduction pathways consist of three categories of components. There must exist one or more sensors, whose activity changes in response to some extracellular stimulus or stimuli. Sensors are generally proteins whose conformation, enzymatic activity or binding surfaces are altered in response to the signal. Components that act as transducers to receive the signal from the sensor and pass the signal on are not mandatory but are usually part of a signal transduction pathway. Transducers can be either small molecules such as cAMP, inositol polyphosphates, or Ca^{++} ions (Berridge, 2005), or can be protein components. Finally, effector components are necessary to produce the change in the intracellular composition or activity. Common examples of effectors include transcription factors, whose activity results in a change in abundance of numerous cellular components, and protein kinases, whose enzymatic activity results in phosphorylation and subsequent changes in abundance, activity, or localization of a number of cellular components. This simple categorization of signal transduction components is complicated by the possibility of multifunctional roles – a particular protein could act both as a transducer, receiving and transmitting the signal onward, and as an effector, directly modifying a

cellular component necessary for adaptation.

Signal transduction pathway characteristics

It is possible to characterize the behavior of signal transduction pathways in a number of different dimensions. One problem faced by the cell in producing an effective adaptive response to environmental changes is that of specificity. In addition to sensing changes in extracellular physical properties such as temperature, pressure, viscosity, osmolarity, and radiation of numerous wavelengths, the cell must be able to sense changes in concentration of a large and diverse set of chemical compounds. Changes in some extracellular molecule concentrations require very specific responses for cellular adaptation; other responses can be protective against a wide range of extracellular changes. Therefore, the specificity of the sensor activity and the specificity of the effector are both of significance.

A second problem facing the cell in developing an effective signal transduction pathway is the spatial constraints produced by the subcellular compartmentalization of eukaryotic cells. Information about changes in the extracellular environment must be physically transmitted within the cell in order to produce changes in the intracellular environment. This spatial transmission of information can occur by as simple a mechanism as membrane-associated protein that acts as both extracellular sensor and intracellular effector. Often, especially when the effector is a transcription factor, the information sensed must pass from the extracellular environment across the plasma membrane, through the cytoplasm and into the nucleus. This more elaborate spatial transmission can be accomplished by a diffusible sensor/effector, or by a series of transducers that move the information from one subcellular compartment to the next.

A third property of all signal transduction pathways relates to the temporal nature of the signal compared to the response. While an virtually instantaneous response would seem ideally adaptive in many circumstances, there is often an energetic cost to adaptation. It seems likely that cells would balance the need to respond immediately to environment changes with the energetic cost of such a response. Time-averaging or temporal filtering of the extracellular signal by the signal transduction pathway allows for such a balance. Alternatively, some signal transduction pathways may respond only to pulsatile signals.

A fourth characteristic of a signal transduction pathway is the quantitative relationship between the signal and the response. This property is often measured at steady-state conditions over a range of levels or concentrations of the signal in question, and may display very different relationships dependent on the definition of "signal" and "response." One term often used in characterizing the input-output or signal-response relationship is the sensitivity, which refers to the magnitude of the change of the response relative to the fold-change in the signal (Koshland et al., 1982). A monotonic input-output response is often characterized as Michaelian when the relationship of change in signal concentration to change in response is consistent with the relationship between the substrate concentration and the reaction velocity in a enzymatic system that behaves according to the Michaelis-Menten approximation. Input-output relationships where the output changes more rapidly than the Michaelian case are denoted as "ultrasensitive," and those where the output changes less rapidly for a given change in the signal than the Michaelian case are denoted "subsensitive." An apparent Hill coefficient and an apparent Michaelis-Menten constant or K_m can often be calculated for such input-output relationships. However, signal transduction pathways

can display much more complex input-output relationships, such as non-monotonic or multiphasic relationships.

Additional concerns in the characterization of a signal transduction pathway include whether the response is history-dependent, or displays the ability to "remember" past conditions. Signal transduction pathways can display hysteresis, wherein the steady-state output at a particular level of signal depends on the initial conditions of the pathway, or multistability, where the output is quantized into a few discrete values, dependent on the current and past conditions. Another variable in signal transduction is the fidelity and homogeneity of the response in a population of cells, either identical or displaying typical diversity of cell-cycle, shape, size, and age. Finally, the combinatorial aspects of signal transduction are significant as well. The signal transduction pathway sensors or transducers may integrate information from multiple stimuli, or the action of the specific effector may depend on the activity of other distinct signaling effectors.

Feedback and signal transduction

Theoretically, signal transduction pathways can be categorized into two distinct types, closed loop and open loop systems. Open loop systems involve the transfer of information about the extracellular environmental signal from the sensor to the effector, but the action of the effector does not result in any change in either the signal or the flow of information through the signal transduction pathway. Most biological signal transduction pathways are closed loop systems, in that the action of the effect feeds back on either the signal or the flow of information from the sensor to the effector. Closed loop systems often incorporate negative feedback, in which the action of the effector tends to negate the signal or reduce the magnitude of signaling. Negative feedback is

often found in systems responsible for homeostasis, in which it is desirable that deviations in a particular signal cause a response that tends to correct such deviations. Positive feedback, or activity of the effector that intensifies signaling through the pathway, can be present as well, and can confer properties such as increased sensitivity, hysteresis, multistability, and a faster approach to steady-state values to the signal transduction pathway.

Nutrient-sensing signal transduction pathways

Numerous signal transduction pathways sense and respond to changes in the concentration of nutrients, or environmental chemicals involved in cellular catabolism. Changes in extracellular nutrient concentration can trigger dramatic adaptive changes, such as the metabolic and morphological changes associated with sporulation in the budding yeast *Saccharomyces cerevisiae* (Schneper et al., 2004). Cells are capable, however, of continual growth and division in extracellular environments containing a wide range of concentrations of a large number of nutrients. Signal transduction pathways are thought to be required for full adaptation, defined as cellular growth and division at the maximal rate, over most extracellular nutrient concentration ranges.

Signal transduction pathways that sense and respond to extracellular nutrient concentration can be categorized by whether the nutrient in question is absolutely essential for cellular growth, is a preferred nutrient, or represents an alternative nutrient. In budding yeast, carbon and nitrogen sources are organized by hierarchical preference, in that certain carbon or nitrogen sources are used preferentially (Schneper et al., 2004). Only when the preferred source is not present will an alternative source be utilized; often there are preferences displayed among the alternative sources as well. For example,

the fermentable hexose glucose is a preferred carbon source for *S. cerevisiae*. Galactose, another fermentable hexose, will be utilized only in the absence of glucose, and therefore represents an alternative carbon source. Galactose is preferred over the nonfermentable trioses glycerol and lactate, which are utilized only in the absence of fermentable carbon sources (Schneper et al., 2004). In the case of hierarchical nutrient preference, homeostasis reflects the maintenance of carbon and nitrogen metabolism, and is a property of the integration of a number of specific signal transduction pathways (Schneper et al., 2004).

Essential nutrients such as the trace metals, copper, iron, and zinc, are usually not available in multiple molecular formulations, and therefore do not display such hierarchical preference. Essential nutrient signal transduction pathways are thought to result in homeostasis by minimizing fluctuations in the intracellular concentration of a particular nutrient. This preservation of a constant concentration is especially important if the nutrient is toxic in excess in addition to being essential for growth.

Extracellular and intracellular nutrient sensors

While all nutrient sensors respond ultimately to changes in the extracellular nutrient concentration, some signal transduction pathways contain sensors that directly measure the extracellular nutrient concentration. Typically, these extracellular sensors are plasma membrane proteins that may or may not be capable of transporting the nutrient in question. For example, the glucose transporter gene homologs *SNF3* and *RGT2* encode for plasma membrane proteins that do not display any glucose transport activity but instead are necessary to signal the presence of extracellular glucose to the glucose-activated transcription factor Rgt1 (Johnston and Kim, 2005). The seven-

transmembrane domain protein Gpr1 also is thought to bind and respond to extracellular glucose (Lemaire et al., 2004). Similarly, the amino acid permease homolog Ssy1 appears to signal extracellular amino acid concentration in the absence of transport capability (Poulsen et al., 2005). In general, extracellular sensors are advantageous in that they directly sample the extracellular environment, but are not reliable sensors of intracellular nutrient status.

Other nutrient sensors are proposed to sense directly the intracellular concentration of nutrient. The processes of growth and cell division inevitably reduce the intracellular nutrient concentration, and intracellular nutrients are replenished from the extracellular environment. Most nutrients are small hydrophilic molecules that do not diffuse freely across the plasma membrane, and therefore rely on protein plasma membrane transporters to move from the extracellular to the intracellular compartment. The rate of transport can depend on the concentration of plasma membrane transporter, the extracellular concentration of nutrient, and the intracellular concentration of nutrient. Therefore, intracellular nutrient sensors may not necessarily respond to changes in the extracellular nutrient concentration in a linear manner. Additionally, cells often store nutrients in the form of polymers and in extracytoplasmic subcellular compartments. The flux of nutrients from these storage forms into the cytoplasmic pool has the potential to affect the kinetic and steady-state characteristics of the response of an intracellular sensor to changes in extracellular nutrient concentration. Intracellular sensors are advantageous in that they directly report the nutrient status of the cytoplasm, where most nutrients are utilized. However, intracellular sensors do not report on environmental conditions directly, which may represent a disadvantage of this approach.

Gene expression in response to nutrient signaling

Frequently, the outcome of a nutrient-responsive signaling is a change in the abundance of a set of gene products. This change in abundance is often accomplished through alteration of the rate of transcription of the corresponding genes. Therefore, in many nutrient-responsive signal transduction pathways, the effector acts as a transcription factor. The transcription factor usually affects a specific subset of the genome by means of sequence-specific binding and transcriptional activation or repression.

Across all nutrient-responsive signal transduction pathways that modulate gene expression, it is possible to identify common categories of gene products whose abundance is modulated. These gene products include plasma membrane proteins that transport the nutrient, intracellular enzymes that are responsible for nutrient metabolism, and extracellular enzymes that hydrolyze the nutrient or nutrient precursors. Also, if the nutrient is stored as a polymer or in a subcellular compartment, the genes whose products mobilize or release the nutrient are often transcriptionally modulated. All of these different types of gene products can affect the extracellular or intracellular levels of nutrient, and thus regulation of these genes may constitute significant feedback, either positive or negative. Finally, the genes whose products that act as signal transduction pathway components are often themselves regulated by that particular signal transduction pathway, a direct type of feedback.

Endocytosis in response to nutrient signaling

Another possible effect of a nutrient signal transduction pathway is the

posttranslational modulation of the abundance of a plasma membrane transporter by nutrient availability. This posttranslational regulation is accomplished by a change in the rate of endocytosis or degradation of the plasma membrane protein. In the cases where a signal transduction pathway controls gene expression in response to a nutrient, the accompanying regulation of endocytosis is mediated by a distinct signal transduction pathway. For most cases, it remains unclear whether the signal that triggers endocytosis is dependent on intracellular nutrient, extracellular nutrient, or nutrient flux across the plasma membrane through the regulated transporter. It has been proposed that the signal for endocytosis involves direct binding of the nutrient ligand to the transporter (Dupre et al., 2004; Felice et al., 2005). In the case of the uracil permease, it has been demonstrated that in addition to endocytosis from the plasma membrane, the transporter moves directly from the Golgi to the vacuole in the presence of nutrient (Blondel et al., 2004). In this case, then, if direct ligand-binding regulates the localization of the transporter, some form of intracellular nutrient must be responsible. However, it is still possible that the nutrient within the endoplasmic reticulum or Golgi compartments, which is topologically equivalent to the extracellular space, is signaling this effect.

Integrated models of nutrient signal transduction

Current goals in the field of nutrient signal transduction include the accurate measurement of the characteristics of nutrient-sensing signal transduction pathways, the demonstration of the biochemical mechanisms by which pathway components interaction, and the discovery of the qualitative result and magnitude of pathway effector activity. When these aims have been accomplished, it becomes possible to create

mathematical models that integrate the discrete steps of signaling and predict the behavior of the pathway across a range of nutrient or signal concentrations. Because the pathways often contain multiple feedback loops, both positive and negative in nature, pathway behavior is often multidimensional and nonintuitive. Thus, basic understanding of pathway behavior will require these types of integrated models. In addition, these models will provide insight on how nutrient-sensing signal transduction pathways develop and evolve.

PREFACE TO CHAPTERS ONE AND TWO

The cellular response to changes in the extracellular environmental is not necessarily identical for every cell within a population, even if the population is genetically identical. Such heterogeneity in response may represent a fundamental barrier to fidelity in signal transduction, or may indicate the utility of a strategy of diversification in environmental response. Regardless of the consequence of variability in the cellular response to the environment, little work has been done to identify and characterize the sources of such variability in eukaryotic cells. In the following two chapters, we examine the sources of variability or noise in gene expression in the eukaryote *Saccharomyces cerevisiae*, and we review recent experimental work relevant to the sources, control, and consequences of noise in gene expression.

CHAPTER ONE:

Experimental investigations into noise in eukaryotic gene expression

Introduction

The stochastic, or random and probabilistic, nature of chemical reactions has the potential to create variation in an identical population of cells (McAdams and Arkin, 1999). Stochasticity can create detectable variation when small numbers of molecules and infrequent reactions are characteristic of a cellular process. These requirements are potentially satisfied by the ordered series of intermolecular reactions that defines the process of gene expression; it is particularly notable that genes are present on the order of two molecules per cell (Schroedinger, 1944). Stochasticity in gene expression is of interest both as a mechanism to generate population variation when phenotypic diversity would be advantageous, and as a theoretical obstacle when fidelity in cellular behavior is required (Rao et al., 2002).

Measurements of stochasticity have been made by single-cell assays of gene expression of fluorescent protein reporters and calculation of population variability (Blake et al., 2003; Elowitz et al., 2002; Ozbudak et al., 2002). However, stochasticity in the reactions of gene expression is not the only potential source of population variability (Paulsson, 2004). Heterogeneity or noise in factors extrinsic to the reactions of gene expression can also create variability. A technique that separates these two different noise sources by use of two independent reporters of gene expression has been developed (Elowitz et al., 2002; Swain et al., 2002). In order to attribute variability to stochasticity in gene expression, measurements must be made either in a system that displays minimal noise from extrinsic factors, or utilizing the dual-reporter technique (Paulsson, 2004).

Prokaryotic noise measurements, using either two reporters in *Escherichia coli* or a single reporter in *Bacillus subtilis*, have produced profiles consistent with stochasticity in the production of mRNA molecules (Elowitz et al., 2002; Ozbudak et al., 2002). In

eukaryotes, the activity of transcriptional enhancers has been postulated to display stochasticity (Fiering et al., 2000). No measurements have directly assayed stochasticity in gene expression in an individual eukaryotic cell (Paulsson, 2004), though noise measurements conducted in the eukaryote *Saccharomyces cerevisiae* with a single reporter have been interpreted in the context of stochastic promoter activation (Blake et al., 2003).

Results

Implementation of the dual reporter methodology in *S. cerevisiae*

In order to measure explicitly the contribution of stochasticity to total variability in gene expression in *S. cerevisiae*, we implemented the dual-reporter technique (Elowitz et al., 2002; Swain et al., 2002) to separate noise intrinsic to the processes of gene expression from other extrinsic sources of variability. We constructed yeast strains expressing both cyan and yellow fluorescent proteins (CFP and YFP) from identical promoters, integrated at the same locus on homologous chromosomes in diploid yeast. Epifluorescence microscopy was used to measure independent CFP and YFP expression levels for individual cells in a population (Fig. 1.1).

Fluorescent reporter equivalence and independence

The separation of noise into extrinsic and intrinsic components by a dualreporter technique requires that the reporters be both equivalent and independent (Paulsson, 2004; Swain et al., 2002). To be equivalent, reporters must have similar kinetics for all the reactions of gene expression. Independence requires that the presence of one reporter gene or protein product does not affect expression or measurement of the other

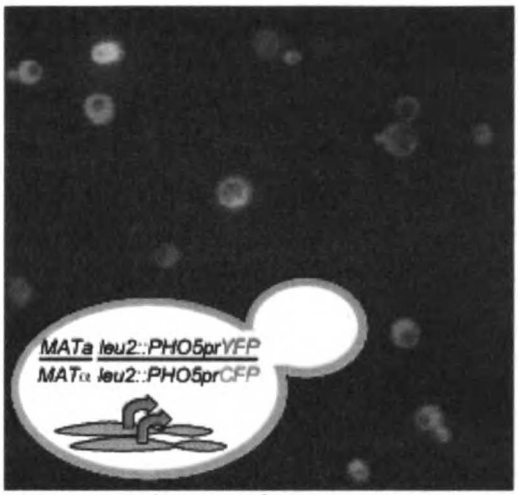


Figure 1.1. Schematic of measurement technique. False-colored overlay of YFP (red) and CFP (green) fluorescence micrographs from a diploid yeast strain expressing YFP and CFP from identical promoters at homologous loci, as diagrammed in the inset.

reporter gene. *A priori*, cyan fluorescent protein and yellow fluorescent protein should satisfy the requirement of equivalence, as they represent allelic variants of the green fluorescent protein, and are therefore nearly identical in sequence. Also, the yeast-optimized variants used in this study were constructed to be as similar as possible in terms of stability and thermosensitivity (see materials and methods). In order to be independent, the reporter genes, mRNAs and proteins must not selectively saturate factors required for gene expression (Paulsson, 2004). We expect that the promoter-specific transcriptional activators and general transcriptional and translational machinery are not limiting; we also assume that the fluorescent protein products are not capable of feedback regulation of their own transcription or translation. Also, we assume that the two identical promoters do not display coordinate regulation *in trans*; such regulation would cause underestimation of the stochastic contribution to population variability by our measurement technique.

To test these assumptions of equivalence and independence, we constructed yeast strains containing each single reporter under the control of the *PHO5* promoter, and measured the population mean and noise for four time points from an induction time course in response to phosphate starvation. The strain expressing only CFP displayed CFP fluorescence distributions indistinguishable from those of the CFP/YFP strain at

each time point (P = 0.71, 0.54, 0.78, 0.97 by Kolmogorov-Smirnov tests). The strain expressing only YFP displayed YFP fluorescence distributions indistinguishable from those of the CFP/YFP strain at each time point (P = 0.27, 0.56, 0.43, 0.44 by Kolmogorov-Smirnov tests). In the strain expressing both reporters, the CFP and YFP fluorescence distributions were indistinguishable for each time point (P = 0.74, 0.82, 1.00, 0.99 by Kolmogorov-Smirnov tests). In addition, we used the single-fluorophore measurements to confirm that YFP and CFP fluorescence are fully separable in our measurement system; the contribution of YFP fluorescence to measured CFP fluorescence is <0.1% and the contribution of CFP fluorescence to measured YFP fluorescence is <0.2%.

Definitions of intrinsic, extrinsic, and total noise and noise strength

Measurements of CFP and YFP values for single cells allow differentiation of two types of noise. Intrinsic noise represents the average difference between CFP and YFP per cell expressed as a percentage of the population mean. Intrinsic noise is attributable to stochastic events during gene expression, spanning promoter-activator binding events through protein degradation (Swain et al., 2002). Extrinsic noise can be due to any heterogeneous factor that affects gene expression, such as cell size, cell cycle, or translational capacity, or to stochastic events in upstream signal transduction (Swain et al., 2002). For each cell population, we calculate the extrinsic, intrinsic, and total noise in terms of the Fano factor or noise strength. We have chosen noise strength as the metric for presentation of noise data for two reasons: first, the extrinsic and intrinsic noise strengths are linearly additive, and secondly, for a single stochastic process the noise strength is independent of population mean, allowing direct comparison of populations with different mean levels of expression. We rely on this approximation to

a univariate stochastic process in order to compare our measurements across different population means.

In the following description, we adhere to the basic definitions of the noise η as $\sigma_p/\langle p\rangle$, or the standard deviation of the protein number per cell divided by the mean protein number. Normalized variance η^2 is defined as $\sigma_p^2/\langle p\rangle^2$ and the Fano factor or noise strength ν is $\sigma_p^2/\langle p\rangle$. Elowitz et al. and Swain et al. defined the total, intrinsic, and extrinsic normalized variance in their dual independent reporter system in the following manner (Elowitz et al., 2002; Swain et al., 2002):

$$\begin{aligned}\eta^2_{\text{total}} &= \eta^2_{\text{intrinsic}} + \eta^2_{\text{extrinsic}} \\ \eta^2_{\text{intrinsic}} &= \frac{\langle (c-y)^2\rangle}{2\langle c\rangle\langle y\rangle} \\ \eta^2_{\text{extrinsic}} &= \frac{\langle cy\rangle - \langle c\rangle\langle y\rangle}{\langle c\rangle\langle y\rangle}\end{aligned} \qquad [1]$$

where c and y represent CFP and YFP fluorescence per cell, respectively.

The corresponding noise strength ν can be calculated by multiplication of the normalized variance η^2 by the mean:

$$\begin{aligned}\nu_{\text{total}} &= \eta^2_{\text{total}}\sqrt{\langle c\rangle\langle y\rangle} \\ \nu_{\text{intrinsic}} &= \frac{\langle (c-y)^2\rangle}{2\sqrt{\langle c\rangle\langle y\rangle}} \\ \nu_{\text{extrinsic}} &= \frac{\langle cy\rangle - \langle c\rangle\langle y\rangle}{\sqrt{\langle c\rangle\langle y\rangle}}\end{aligned} \qquad [2]$$

As the intrinsic and extrinsic components of the normalized variance are directly additive, the intrinsic and extrinsic noise strength are additive as well.

Extrinsic noise of *PHO5* predominates during an induction time-course

We induced the expression of CFP and YFP from the budding yeast *PHO5* promoter and quantified the CFP and YFP fluorescence of single cells in a population over time (Fig. 1.2A). At every time point the total noise of gene expression from the *PHO5* promoter is dominated by the contribution from extrinsic factors (Fig. 1.2B); the intrinsic noise strength represents only between 2% and 20% of the total noise strength for these measurements.

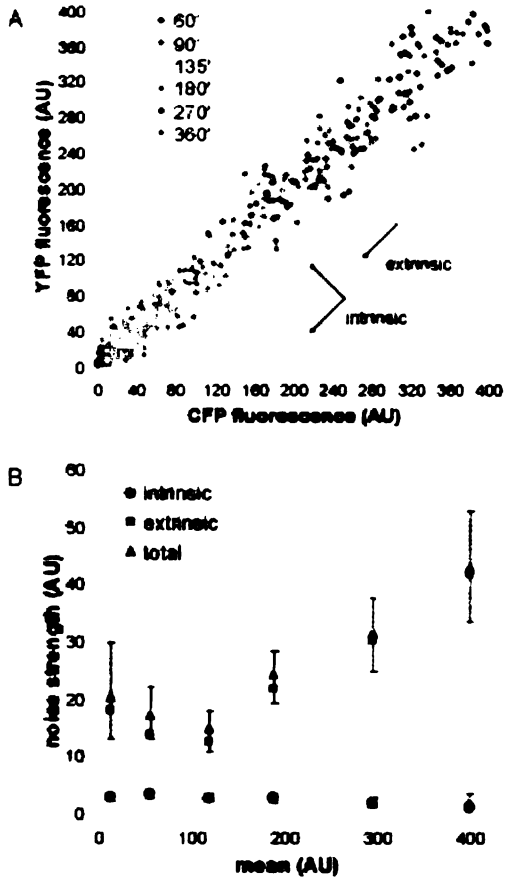


Figure 1.2. Separation of intrinsic and extrinsic noise for the *PHO5* promoter.. **(A)** Scatter plots showing CFP and YFP values for each cell (•) during a time-course of *PHO5* induction by phosphate starvation. Populations from different time points are indicated with different colors. Extrinsic noise is manifested as scatter along the diagonal and intrinsic noise as scatter perpendicular to the diagonal. Hereafter, AU indicates arbitrary units of fluorescence. **(B)** Total, extrinsic, and intrinsic noise strength as a function of population mean for (A). The solid line represents expectations for a single stochastic process, and error bars represent bootstrap values.

Extrinsic noise predominates during steady-state expression of the *GAL1* promoter

We confirmed the general observation that extrinsic noise contributes substantially more to total variation than intrinsic noise for a second yeast promoter, *GAL1*. For *GAL1*, we measured the intrinsic, extrinsic, and total noise at a number of different levels of steady-state induction (Fig. 1.3). Differences in steady-state induction were achieved by growth in media containing different levels of galactose. The intrinsic noise strength represents less than 3% of the total noise

strength at all levels of induction. For *GAL1*, very little of the total measured noise is due to intrinsic noise factors.

A previous study measured noise strength in gene expression in *S. cerevisiae* with a single fluorescent reporter (Blake et al., 2003). This study examined the noise of a modified version of the *GAL1* promoter. Our observations for *GAL1* differ from the conclusions of this study which reported a profile of noise strength as a function of

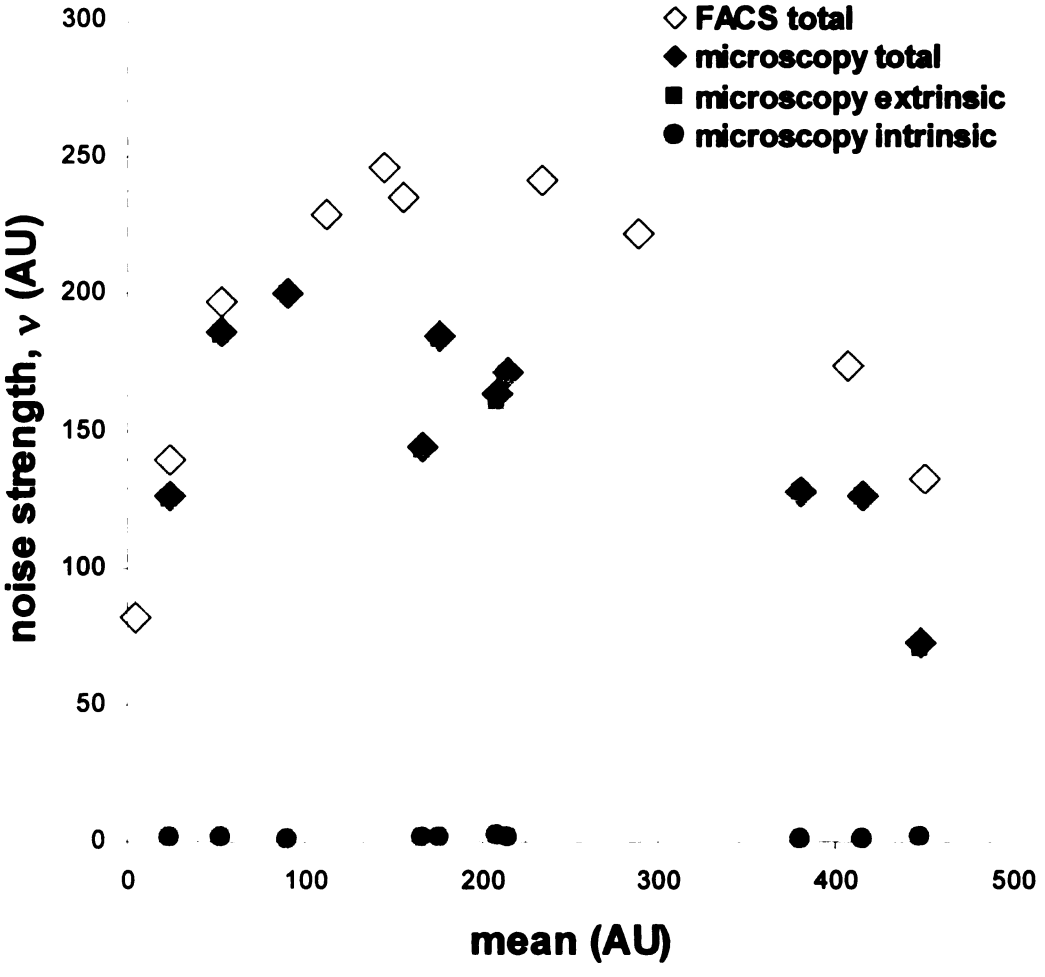


Figure 1.3. Intrinsic, extrinsic, and total noise measurements for the yeast *GAL1* promoter. Noise strength of the *GAL1* promoter after eight hours of exposure to various levels of galactose. Noise strength was measured by flow cytometric quantification of YFP (◇), or dual-reporter microscopy and calculate of total (◆), intrinsic (•), and extrinsic (■) noise strengths. The flow cytometric measurements were scaled to microscopy arbitrary fluorescent units by maximal induction population mean.

transcription rate qualitatively similar to that which we observe for the extrinsic and total noise strength (Fig. 1.3). This discrepancy may result from differences in strains or experimental details, or may reflect the fact that the technique used previously was unable to discriminate between intrinsic and extrinsic noise sources. We conclude from the data concerning both *PHO5* and *GAL1* that the stochasticity of gene expression is not necessarily reflected by measurements of total noise employing single reporter techniques.

A substantial portion of extrinsic noise is not gene-specific

Heterogeneity in a number of factors that affect gene expression may underlie extrinsic noise, including heterogeneity in cell size and shape, cell cycle stage, or gene-specific signaling. We tested if extrinsic noise factors can be eliminated by isolation of subpopulations of cells that are homogeneous in size and shape by flow cytometry. We found that while diminished by this process, extrinsic noise predominates relative to intrinsic noise in these subpopulations (Fig. 1.4A). Similarly, neither correction for individual cellular volume nor segregation by cell cycle stage (6) resulted in more than a ~25% decrease in extrinsic noise (Fig. 1.4B). To distinguish extrinsic noise that is gene-specific from a global, nonspecific source, we examined the correlation between CFP and YFP fluorescence in a strain expressing CFP from one promoter and YFP from a second, distinctly regulated promoter. At steady-state levels, expression from the *PHO84* and *GAL1* promoters is correlated with a R^2 value of 0.88 (Fig. 1.5A); additionally, expression from the *PHO84* and *ADH1* promoters is correlated (R^2 = 0.93) (Fig. 1.5B). Therefore, the majority of extrinsic noise in these cases is not

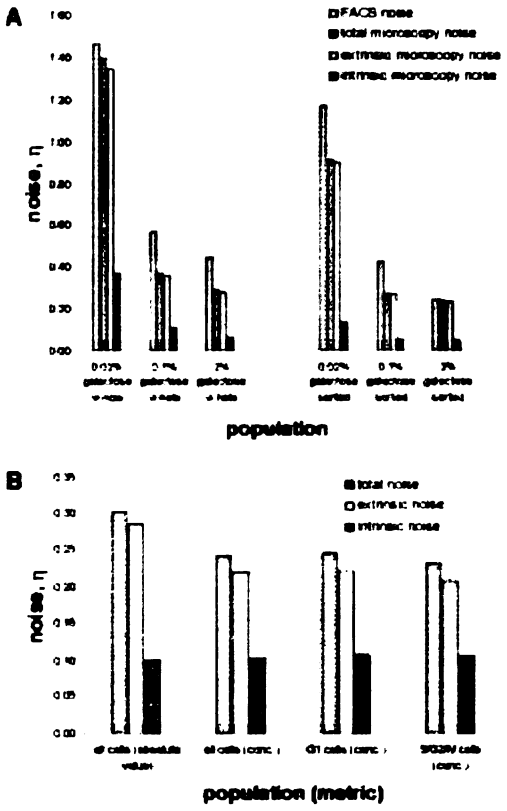


Figure 1.4. Extrinsic noise reduction by cell shape, size, and cell cycle stage. (A) Total, intrinsic, and extrinsic noise of *GAL1*-expressing populations before ("whole") and after sorting ("sorted") by flow cytometry cell size criteria (forward-scatter and side-scatter gating). The *GAL1* promoter was induced by the indicated concentrations of galactose for one hour prior to measurement. (B) Total, intrinsic, and extrinsic noise of *PHO5*-expressing populations before ("absolute value") and after correction for variation in cellular volume ("conc."), and in populations stratified by cell cycle stage after correction for cellular volume.

promoter-specific and will cause gene products in a cell to be maintained in constant relative concentration.

Transient extrinsic noise of *PHO5*

Extrinsic noise decreases substantially during an induction time-course of *PHO5* expression (Fig. 1.6). To examine if some component of this transient noise is dependent on signal transduction, we measured the extrinsic noise during *PHO5* induction by two different stimuli. We induced *PHO5* both by phosphate starvation and by direct chemical inhibition of an inhibitor-sensitive allele of the cyclin-dependent kinase *PHO85*. Inhibition of the inhibitor-sensitive Pho85 results in activation of the transcription factor Pho4 and therefore expression of the *PHO5* gene. The extrinsic noise of *PHO5* gene expression is substantially reduced in the inhibitor time-course relative to the phosphate-starved induction (Fig. 1.6). This relative difference suggests that some portion of extrinsic noise is dependent on the specific characteristics of the nutrient-responsive signal transduction pathway that controls *PHO5* expression.

The intrinsic noise of *PHO5* is not due to measurement error

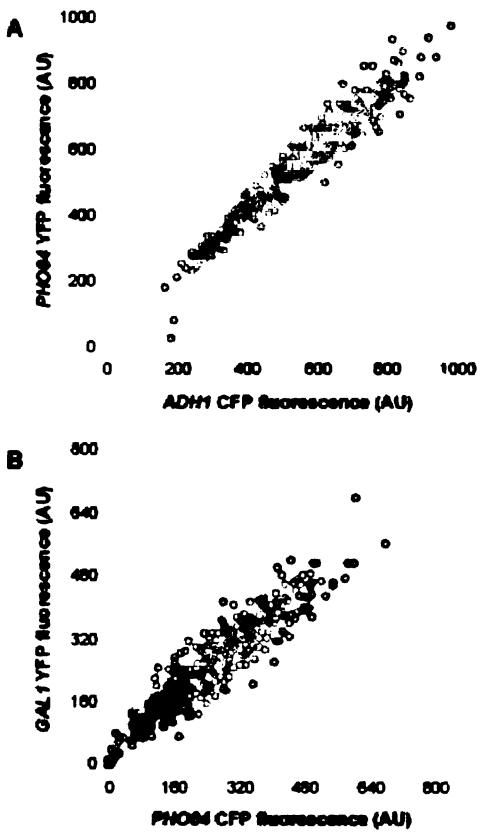


Figure 1.5. Correlated expression from two different promoters. (A) Scatter plot of cells expressing YFP from the *PHO84* promoter and CFP from the *ADH1* promoter. $R^2 = 0.93$; *PHO84* noise is 30% and *ADH1* noise is 30%. (B) Scatter plot of cells expressing YFP from the *GAL1* promoter and CFP from the *PHO84* promoter. $R^2 = 0.88$; *PHO84* noise is 58% and *GAL1* noise is 53%. For each scatter plot, cells from three separate measurements are shown.

We note that the intrinsic noise, while less substantial than extrinsic factors, is nonzero and substantially larger than the error of measurement of our system. A component of the intrinsic noise or intrinsic noise strength will be due to systematic errors of measurement in our system. To estimate the magnitude of the intrinsic error of measurement, we constructed a yeast strain expressing green fluorescent protein (GFP) from the *MET3* promoter, induced expression of GFP by methionine starvation, and measured the GFP fluorescence per cell using the YFP excitation/emission filter set, and using the CFP excitation/YFP emission filters. We then calculated intrinsic noise and intrinsic noise strength for these cell populations as if the YFP and CFP/YFP measurements were independent fluorophores. The intrinsic noise strength was <0.25 AU in all cases, and was mainly attributable to yeast autofluorescence in the CFP excitation/YFP emission measurement. During the actual measurements of strains containing CFP and YFP, autofluorescence in the CFP excitation/CFP emission filter set was significantly smaller than the autofluorescence of this control (data not shown), so we suspect this intrinsic error of measurement is an overestimate of the error in our measurement system. For *PHO5*, the intrinsic noise strength measurements were over

>0.75 AU for all measurements (Fig. 1.2B), representing at least three times the maximal intrinsic noise strength due to variation in the measurement process.

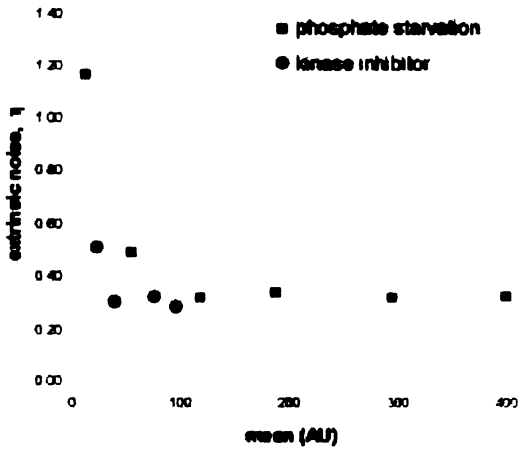


Figure 1.6. Transient extrinsic noise is signal-dependent. Extrinsic noise as a function of population mean was measured during an induction time-course of *PHO5*, induced by either phosphate starvation (green squares) or chemical inhibitor of the upstream kinase Pho85 (blue circles).

Measured *PHO5* intrinsic noise is consistent with a stochastic process

For the *PHO5* promoter, the intrinsic noise decreases from 45% to 5% over the induction time course shown in Fig. 1.2B (Fig. 1.7A). In contrast to the extrinsic or total noise, the intrinsic noise decreases in proportion to the inverse square-root of the mean of gene expression, consistent with a single underlying stochastic process. This relationship is equivalent to both a linear relationship between the inverse mean and the normalized variance (Fig. 1.7B) and the constant relationship seen between population mean and the intrinsic noise strength (Fig. 1.2B).

The intrinsic noise measured in our experiments could result from ongoing stochastic events in gene expression or from a stochastic delay between activation of the two alleles followed by deterministic gene expression. The delay model predicts that the maximal and average difference between CFP and YFP per cell in a population of cells during induction will decrease over time and reach zero at steady-state levels of induction. We measured the intrinsic noise of the constitutively active promoter *ADH1* at steady-state and find that the intrinsic noise strength is >4-fold over the limit of

detection. Additionally, during induction time courses of the *PHO5* promoter, the maximal and average difference between CFP and YFP per cell increases after >90% of the cells have induced both alleles for 7 of 7 time courses examined. For example, for the six time points of the induction time course shown in Fig. 1.2A, the average differences between CFP and YFP per cell are 5, 16, 18, 23, 21, and 21 AU for the 60', 90', 135', 180', 270' and 360' populations, and the maximal differences between CFP and YFP for any one cell are 27, 45, 75, 84, 73, and 98 AU, respectively. 100% of the cells have induced both CFP and YFP expression by the second time point (90') in this series.

Both the observation of intrinsic noise at steady-state for *ADH1*, and the increase in the difference between CFP and YFP after both alleles have induced *PHO5*, support a model in which intrinsic noise is not solely a consequence of an initial difference in the timing of induction of each allele followed by deterministic gene expression. We conclude that the intrinsic noise in our measurement system is attributable to stochasticity in gene expression, and that stochastic events are not limited to the initial

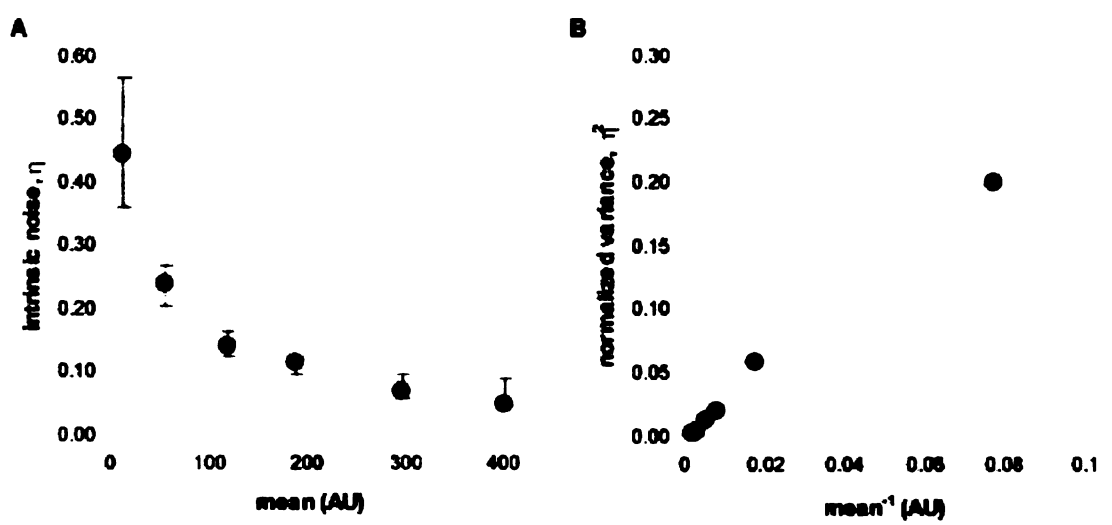


Figure 1.7. Intrinsic noise is consistent with a stochastic process. (A) Intrinsic noise as a function of population mean for populations from the *PHO5* induction time-course shown in Fig. 1.2A. Error bars are calculated by the bootstrap method. (B) Intrinsic normalized variance as a function of the inverse mean for the same data. For both plots, the solid lines represent expectations for a single underlying stochastic process.

induction period but occur during active gene expression.

Intrinsic noise of the *GAL1*, *PHO84*, and *PHO5* gene promoters

To characterize stochasticity in eukaryotic gene expression further, we measured the intrinsic noise of native yeast promoters of the *GAL1*, *PHO84*, and *PHO5* genes. *GAL1* and *PHO5* represent well-studied regulated eukaryotic promoters, and *PHO84* and *PHO5* share a common upstream signal transduction pathway for activation (Johnston, 1987; Lenburg and O'Shea, 1996; Oshima, 1997). We induced expression of each promoter at different rates by application of the appropriate external stimulus, measured the population mean and intrinsic noise strength at several time points, and calculated a rate of expression and average intrinsic noise strength over time for each level of stimulus (Fig. 1.8). We found that of the three promoters examined, both *GAL1* and *PHO84* display a relatively low level of intrinsic noise strength that does not vary significantly with changes in the rate of gene expression (Fig. 1.8A-B). The average intrinsic noise strength for each promoter was ~4-fold above the detection limit but ~3-fold lower than that of *PHO5* during maximal expression by phosphate starvation. Our results for *GAL1* conflict with a previous study (Blake et al., 2003) which reported a profile of noise strength as a function of transcription rate qualitatively similar to that which we observe for the extrinsic and total noise strength (Fig. 1.3). This discrepancy may result from differences in strains or experimental details, or may reflect the fact that the technique used previously was unable to discriminate between intrinsic and extrinsic noise sources. Therefore, previous claims regarding stochastic fluctuations in eukaryotic gene expression (Blake et al., 2003) obtained in studies with a single color reporter should be

re-evaluated, as these measurements may be dominated by contributions from noise extrinsic to the process of gene expression.

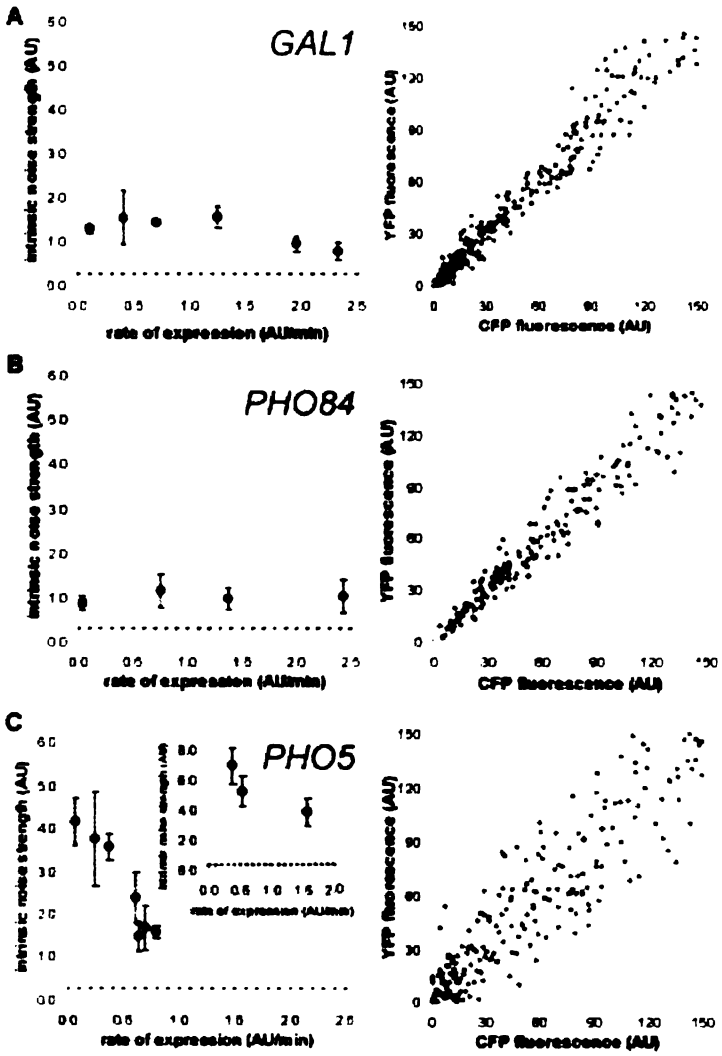


Figure 1.8. Intrinsic noise strength as a function of rate of expression. (A) Intrinsic noise strength as a function of rate of expression for *GAL1* and a scatter plot of *GAL1*-expressing cells at maximal induction (right). To produce different rates of expression, cells were induced with different galactose concentrations (left). (B) Intrinsic noise strength as a function of rate of expression for *PHO84* and a scatter plot of *PHO84*-expressing cells at maximal induction (right). Cells were induced with different phosphate concentrations. (C) Intrinsic noise strength as a function of rate of expression for *PHO5* and scatter plot of *PHO5*-expressing cells at the maximal level of induction by phosphate starvation (right). Cells were induced with various levels of chemical inhibition of an upstream kinase (left) or various organic phosphate concentrations (inset). The dashed line indicates the intrinsic error of measurement, and error bars represent standard deviations. Scatter plots contain cells from multiple time points.

In contrast to the other promoters, *PHO5* has a substantially larger intrinsic noise strength that decreases with increasing rate of gene expression (Fig. 1.8C); at a low expression rate, *PHO5* displays more than twice the intrinsic noise strength than at maximal expression. This decrease in intrinsic noise strength relative to the rate of gene expression is not sensitive to the exact stimulus used to induce *PHO5* expression to different levels (*cf.* Fig. 1.8C with inset). We conclude that noise intrinsic to gene expression is promoter-specific, and does not depend absolutely on rate of

expression, induction stimulus, or identity of sequence-specific transcriptional activator.

An extended, eukaryotic model for stochastic gene expression

The signaling pathways that modulate expression from these three promoters change the rate of mRNA production in order to effect changes in rate of gene expression (Johnston, 1987; Lenburg and O'Shea, 1996; Oshima, 1997). Previous work in a prokaryotic system studying a lacI-repressible promoter demonstrated that noise strength did not change substantially with changing rates of mRNA production (Ozbudak et al., 2002), similar to our observations of *GAL1* and *PHO84*. The authors proposed a basic model of noise generation in gene expression in which gene regulation is accomplished by direct modulation of transcriptional rate (Ozbudak et al., 2002; Thattai and van Oudenaarden, 2001). This model does not explicitly account for promoter regulation by conversion from an inactive state to an active state that is competent for transcription. Furthermore, this model does not predict the noise profile of the *PHO5* promoter, which is known to be regulated by a discrete promoter transition step that is upstream of and independent of transcription (Fascher et al., 1993). Because the paradigmatic mechanisms of transcriptional regulation in both eukaryotic and prokaryotic systems involve the regulated binding of factors to the promoter DNA, we constructed a model of stochastic gene expression that incorporates two distinct promoter states: an inactive state not permissive for transcription and an active state that is competent for transcription (Fig. 1.9A). This general model has been proposed previously with various elaborations (Blake et al., 2003; Ko, 1991; Paulsson, 2004; Peccoud and Ycart, 1995; Swain et al., 2002). In our model we assume that multiple transcripts can be generated by a single promoter activation event, and therefore that the

promoter activation event does not represent the binding of RNA polymerase.

Calculation of steady-state noise in terms of stochastic kinetic constants

We used the master equation to calculate a general solution of steady-state noise in terms of the stochastic kinetic constants of the model. Construction and steady-state solution of the moment generating function for the master equation for the model shown in figure 1.9A was undertaken as previously described (Thattai and van Oudenaarden, 2001). The resulting steady-state noise equation, expressed in the form of normalized variance, must conform to the general equation described previously (Paulsson, 2004):

$$\eta_{\text{int}}^2 = \underbrace{\frac{\sigma_p^2}{\langle p \rangle^2}}_{\text{protein}} + \underbrace{\frac{\sigma_m^2}{\langle m \rangle^2} \cdot t_{m \to p}}_{\text{mRNA}} + \underbrace{\frac{\sigma_a^2}{\langle a \rangle^2} \cdot t_{a \to p}}_{\text{active gene}} \qquad [3]$$

where the t represents the time-averaging contributions, p refers to protein, m to mRNA, a to the active gene, and the normalized variance terms refer to each molecular species in isolation. These terms are part of the intrinsic noise definition of the dual-reporter system; any noise terms due to signaling upstream of the promoter or constant noise contributions due to population heterogeneity are part of the extrinsic noise definition and therefore do not appear in this expression.

We note that the binomial distribution applies to the active DNA species, and set the gene number to one. We incorporate the results from the calculation of variance from differentiation of the master equation moment generating function and rewrite the normalized variance expression as noise strength by multiplication by the protein mean.

$$\nu_{\text{int}} = 1 + \mathrm{E}_{m \to p} \cdot \frac{\gamma_{\mathrm{m}}}{\gamma_{\mathrm{m}} + \gamma_{\mathrm{p}}} + \mathrm{E}_{m \to p}\mathrm{E}_{a \to m} \cdot \frac{(1 - \langle a \rangle)(\mathrm{k}_{\mathrm{a}} + \gamma_{\mathrm{a}} + \gamma_{\mathrm{m}} + \gamma_{\mathrm{p}})\gamma_{\mathrm{a}}\gamma_{\mathrm{m}}}{(\mathrm{k}_{\mathrm{a}} + \gamma_{\mathrm{a}} + \gamma_{\mathrm{m}})(\gamma_{\mathrm{m}} + \gamma_{\mathrm{p}})(\mathrm{k}_{\mathrm{a}} + \gamma_{\mathrm{a}} + \gamma_{\mathrm{p}})} \qquad [4]$$

Note that the noise strength can be thought of in integer units of molecules of protein. In our experimental setup, we measure noise strength in terms of arbitrary fluorescent units (AU); 1 AU corresponds to some unknown number of protein molecules. In Eq. 4, we have defined the average number of proteins produced per mRNA as the translational efficiency $E_{m\to p}$ (equal to k_p/γ_m), and the average number of mRNA produced per promoter activation event as the transcriptional efficiency $E_{a\to m}$ (equal to k_m/γ_a). This formulation eliminates the ratio between the means of various species, leaving behind the translational and transcriptional efficiency and obscuring the origin of each noise term. However, we can label each term relative to the original normalized variance terms in order to understand the source of the noise in protein levels:

$$\nu_{\text{int}} = \underbrace{1}_{\text{translation}} + \underbrace{E_{m\to p}\cdot\frac{\gamma_m}{\gamma_m+\gamma_p}}_{\text{transcription}} + \underbrace{E_{m\to p}E_{a\to m}\cdot\frac{(1-\langle a\rangle)(k_a+\gamma_a+\gamma_m+\gamma_p)\gamma_a\gamma_m}{(k_a+\gamma_a+\gamma_m)(\gamma_m+\gamma_p)(k_a+\gamma_a+\gamma_p)}}_{\text{gene activation}} \quad [5]$$

For example, the first term derives from translation. Our model contains the assumption that protein production from mRNA displays Poissonian statistics. If protein production in our model involved competition between the translation and degradation machinery, this first term would be larger than one, because translational events would occur with a geometric rather than Poissonian distribution, and the ratio of variance to mean of the geometric distribution is larger than one (McAdams and Arkin, 1997). Alternatively, if translation were not a stochastic reaction, the first term would become zero. In such a case, the amount of noise strength would still change as translational efficiency changes. This can be understood intuitively by considering that when we multiplied all three terms of the normalized variance equation by protein mean to

produce noise strength, we introduced a dependence on protein mean to the gene activation and transcription components of the noise strength.

Three different cases of gene expression

The noise strength equation can be simplified by assuming that protein degradation is infrequent compared to all other reactions in the model, and calculating the steady-state population mean of the active promoter in terms of the kinetic constants of promoter activation and deactivation:

$$\nu_{\text{int}} \approx \underbrace{1}_{\text{translation}} + \underbrace{E_{m\to p}}_{\text{transcription}} + \underbrace{E_{m\to p}\cdot\frac{k_m\gamma_a}{(k_a+\gamma_a)^2}}_{\text{gene activation}} \qquad [6]$$

This expression allows intuition of three different cases, or sets of behavior regarding noise that may apply to a particular promoter.

Case I

When both k_a and γ_a are similar in magnitude and much smaller than k_m, the third term is large relative to the other terms and is highly dependent on the k_a, and case I results. The noise strength in this case will decrease with an increasing rate of promoter activation, but will increase with increasing rate of transcription or translational efficiency (Fig. 1.9D). Intuitively, the effect of increasing the rate of promoter activation can be thought of as removing available substrate for the promoter activation step, which makes that promoter activation step behave in a less-than-Poissonian manner and causes the noise strength to decrease.

Case II

When k_a is much smaller than γ_a, and γ_a is approximately equal to or less than k_m, then the third term can be substantial relative to the other terms. The third term depends only on $E_{a\to m}$ and not on k_a, and case II results:

$$\nu_{int} \approx \underbrace{1}_{\text{translation}} + \underbrace{E_{m\to p}}_{\text{transcription}} + \underbrace{E_{m\to p}\cdot\frac{k_m}{\gamma_a}}_{\text{gene activation}} \quad [7]$$

The noise strength is not sensitive to changes in promoter activation but does scale with both the efficiency of transcription $E_{a\to m}$ and the efficiency of translation $E_{m\to p}$ (Fig. 1.9E).

Case III

When k_a and γ_a are both much larger than k_m, the third term, or contribution from the noise of gene activation, is very small and case III results:

$$\nu_{int} \approx \underbrace{1}_{\text{translation}} + \underbrace{E_{m\to p}}_{\text{transcription}} \quad [8]$$

The noise strength is not sensitive to changes in promoter activation or transcription, only to the efficiency of translation $E_{m\to p}$ (Fig. 1.9F).

A previously proposed model for noise generation in prokaryotes (Ozbudak et al., 2002; Thattai and van Oudenaarden, 2001) suggested that the stochastic contribution of transcription to noise strength is modulated by the translational efficiency. In such a model, the actual stochasticity of translation contributes minimally to the variability in protein levels, but the measured noise strength scales with translational efficiency. This scaling factor led the authors to propose that, given a set of genes with equivalent stochasticity in transcription, those genes that are relatively inefficiently translated will be less phenotypically variable. Our model encompasses this possibility, and suggests

additional situations where the stochasticity of transcription contributes little to population variability relative to the stochasticity of promoter activation.

We were interested in the biological relevance of these three different conditions, representing three distinct kinetic models of promoter transcriptional activation (Fig. 1.9B). Case I occurs when the activation step is infrequent relative to transcription and the active promoter state is stable ($k_a, \gamma_a \ll k_m$). This could correspond to a promoter that is activated by a slow chromatin-remodeling step where positioned nucleosomes are modified and removed from the promoter, and that is slowly inactivated by replacement of the positioned nucleosomes. In case II, the activation step is infrequent relative to transcription and the active promoter state is unstable ($k_a \ll \gamma_a, k_m$), possibly corresponding to a relatively infrequent but rapidly reversible activation step such as nucleosomal sliding, or acetylation and deacetylation reactions at a nucleosome, or prokaryotic promoter DNA looping. Finally, case III occurs when the activation step is frequent relative to transcription and the activated promoter is highly unstable ($k_m \ll k_a, \gamma_a$). The third case in its limit corresponds to the previously proposed prokaryotic model (Thattai and van Oudenaarden, 2001), and could represent rapid activator binding-dissociation reactions where transcription occurs only for a fraction of the binding events. All three conditions may occur in both eukaryotic and prokaryotic systems, dependent on the relative kinetic parameters of promoter activation, inactivation, and transcription.

Stochastic simulation confirms steady-state noise equation predictions

In addition to calculating steady-state noise strength from the general noise strength equation, we performed stochastic simulations in order to approximate our experimental measurements where each intrinsic noise strength measurement is the average of multiple time points during an induction time course. We examined how varying the rate of promoter activation (k_a) in order to change the steady-state mean of gene expression affects the intrinsic noise strength for the three different cases detailed above (Fig. 1.9C). We also compared how varying the rate of promoter activation (k_a), transcription (k_m), or translation (k_p) affects the intrinsic noise strength for each case (Fig. 1.9D-F). We find that the time course simulations and the steady-state noise equation values are in agreement, suggesting that our observations can be interpreted meaningfully in the context of this model for noise generation.

Mutations in the Pho4-binding sites of *PHO5* result in increased intrinsic noise

The noise strength profile of *PHO5* (Fig. 1.8C) is strikingly similar to that of the simulations and predicted values for case I as the promoter activation rate is changed (Fig. 1.9C). Therefore, we wished to test the model that noise generation at *PHO5* is dependent on the rate of a slow upstream promoter transition. In the inactive state the *PHO5* promoter displays positioned nucleosomes (Almer and Horz, 1986); upon binding of the Pho4 transcription factor to upstream activating sequences (UAS1 and UAS2) in the promoter, nucleosome positioning is disrupted and the nucleosomes are removed from the promoter region (Boeger et al., 2003; Fascher et al., 1990). Multiple chromatin-remodeling complexes, including SWI/SNF, INO80, and SAGA,

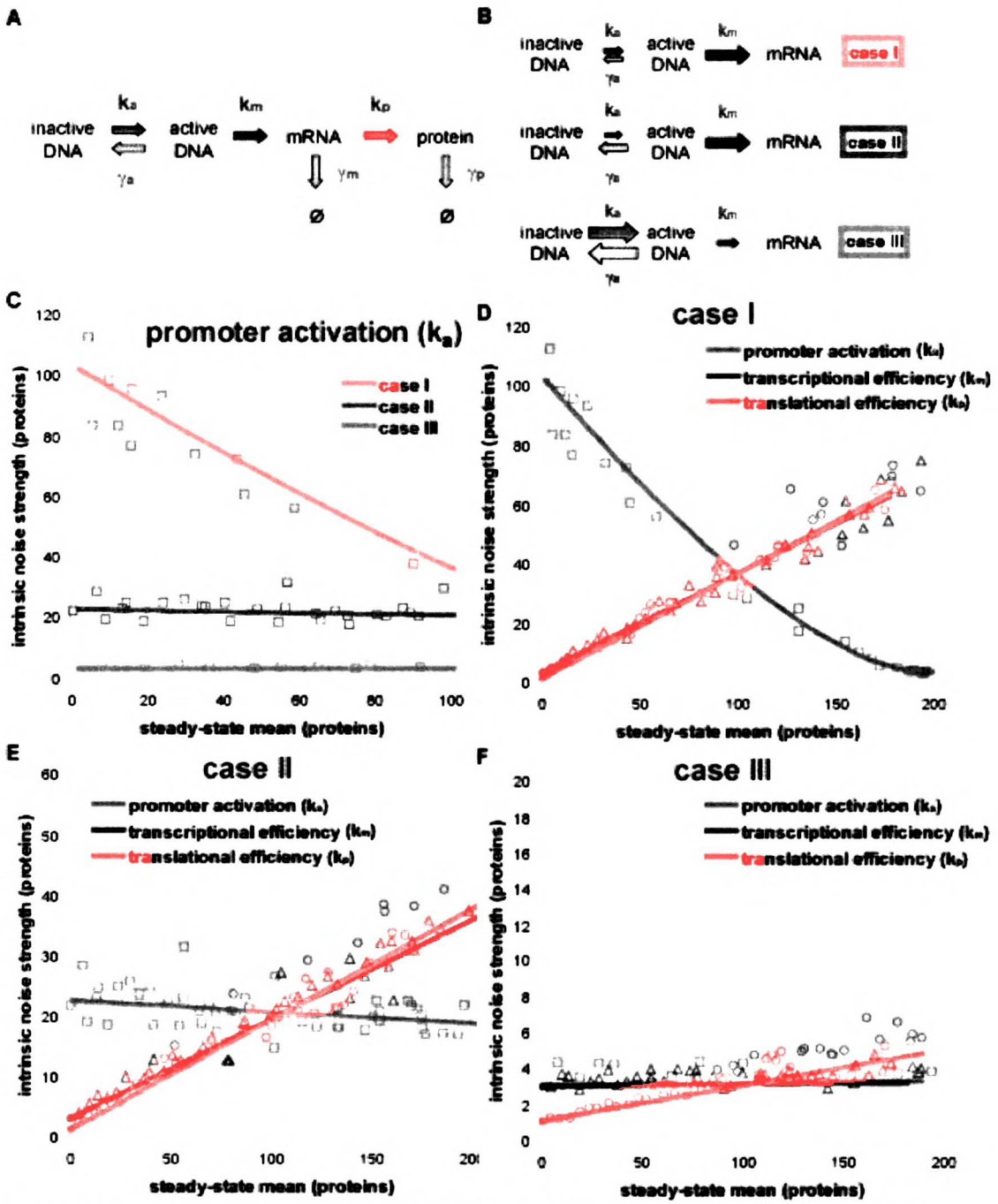


Figure 1.9. General stochastic model of gene activation and expression. (A) Schematic of reactions with stochastic rate constants of production (k) and degradation (γ). Ø indicates the null product of degradation. (B) Three cases, where the relative size of the arrows indicates the relative magnitude of the constants within each case. (C) The effect on intrinsic noise strength of changing the promoter activation rate to change the steady-state mean of expression for case I (orange □), case II (violet □), and case III (cyan □). (D-F) The effect on intrinsic noise strength of changing promoter activation (□), transcriptional efficiency (○), and translational efficiency (△) to change the steady-state mean of expression for case I (D), case II (E), and case III (F). For (C-F), the solid lines display the predicted values, and the open symbols are averages from stochastic simulations (6).

participate in this promoter transition (Barbaric et al., 2003; Barbaric et al., 2001; Ebbert et al., 1999; Steger et al., 2003), which can occur independent of subsequent transcription (Fascher et al., 1993). Individual promoter mutants in these UAS sites are able to activate gene expression weakly when the signaling pathway is maximally activated, and display a significant defect in the disruption of positioned nucleosomes during activation (Fascher et al., 1993). As predicted by our noise model, the two *PHO5* UAS mutant promoters have a lower rate of gene expression and increased intrinsic noise strength compared to the wild-type *PHO5* promoter (Fig. 1.10A); the intrinsic noise of the UASml promoter is 77% at a mean of 26 AU, compared to 43% intrinsic noise at a mean of 29 AU for the wild-type promoter. This observation is consistent with a model in which chromatin-remodeling is the stochastic promoter transition whose rate determines the noise of the *PHO5* promoter.

Strains lacking chromatin-remodeling activities display increased intrinsic noise

Though the UAS mutants do affect chromatin-remodeling, the transcription factor Pho4 that binds to these promoter sites is known to affect subsequent steps in promoter activation and transcription (Magbanua et al., 1997; Wu and Hampsey, 1999). Therefore, to test further the hypothesis that chromatin-remodeling is the stochastic promoter activation step in question, we examined the noise strength of the maximally induced *PHO5* promoter in yeast strains lacking single components of each of three chromatin-remodeling complexes known to act at the *PHO5* promoter (Fig. 1.10B). The deletion of components of SWI/SNF (*snf6Δ*), INO80 (*arp8Δ*), or SAGA (*gcn5Δ*) each results in a decreased rate of gene expression and an increased intrinsic noise strength,

consistent with the predictions of the noise model. The intrinsic noise of the *snf6Δ* strain is 73% at a mean of 35 AU. We note that there are substantial differences in the magnitude of increase of noise strength among the mutants, which may reflect different roles in the promoter transition process.

Mutations in the PHO5 TATA box lower intrinsic noise

Finally, we wished to confirm a qualitatively different prediction of the noise generation model for *PHO5*. The model predicts that the noise strength should scale directly with the efficiency of a transcription step downstream of promoter activation (Fig. 1.9D). The TATA element of the *PHO5* promoter is absolutely required for transcription but dispensable for efficient chromatin-remodeling (Fascher et al., 1993). We measured the noise strength at maximal induction of a series of *PHO5* promoter variants with various TATA box sequences (Fig. 1.10C). As predicted, the mutant TATA box promoters display decreasing noise strength with a decreasing rate of gene expression; the intrinsic noise of the TATA-C2 promoter is 15% at a mean of 25 AU. These observations provide further evidence that our model, and specifically case I, represents a useful framework for rationalization of noise generation at the *PHO5* promoter.

Differentiating between case II and case III for *GAL1*

Both the *PHO84* and *GAL1* promoters display profiles that are consistent with either case II or case III of the model. If the *GAL1* promoter is an example of case III, the intrinsic noise strength has only transcriptional and translational contributions. The

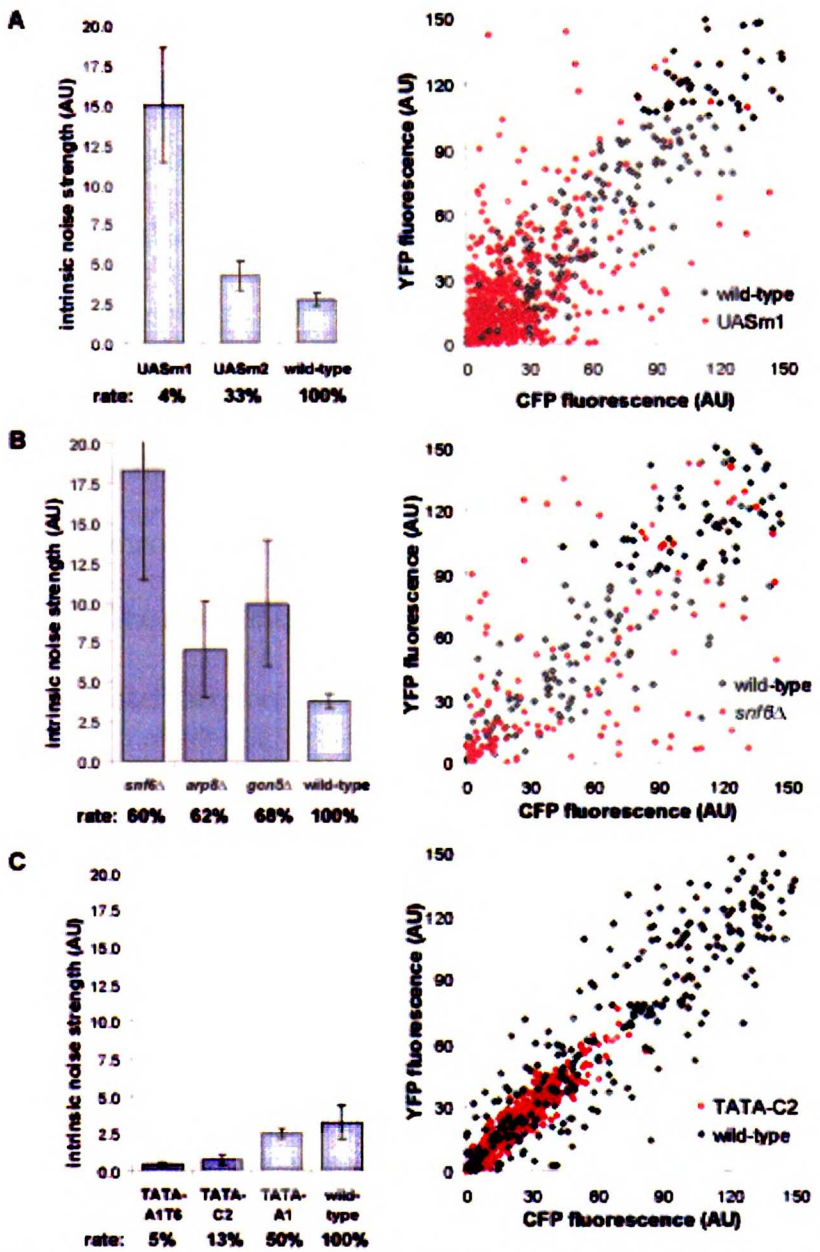
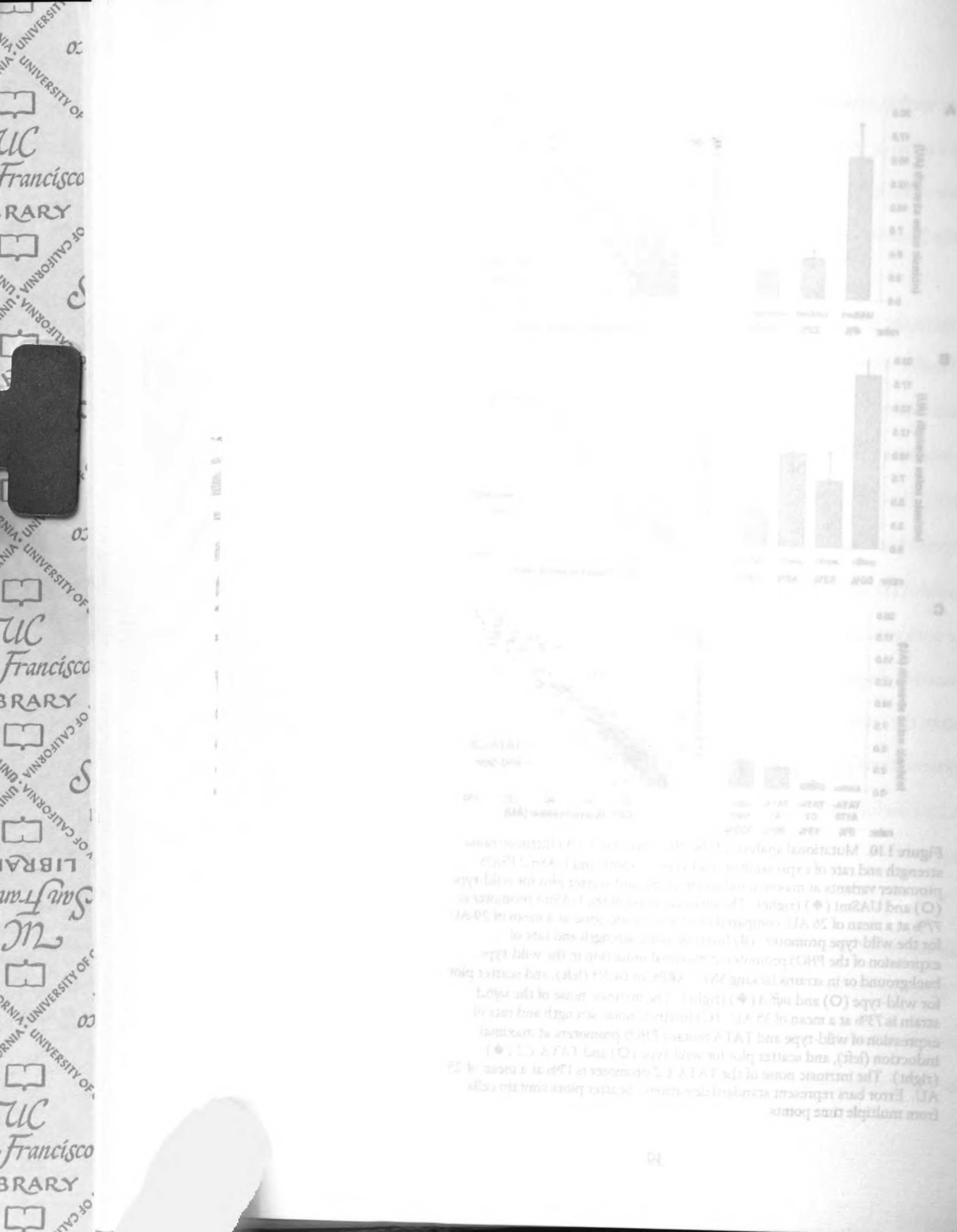


Figure 1.10. Mutational analysis of the *PHO5* promoter. (A) Intrinsic noise strength and rate of expression of wild-type, UASml, and UASm2 *PHO5* promoter variants at maximal induction (left), and scatter plot for wild-type (O) and UASml (◆) (right). The intrinsic noise of the UASml promoter is 77% at a mean of 26 AU, compared to 43% intrinsic noise at a mean of 29 AU for the wild-type promoter. (B) Intrinsic noise strength and rate of expression of the *PHO5* promoter at maximal induction in the wild-type background or in strains lacking *SNF6*, *ARP8*, or *GCN5* (left), and scatter plot for wild-type (O) and *snf6Δ* (◆) (right). The intrinsic noise of the *snf6Δ* strain is 73% at a mean of 35 AU. (C) Intrinsic noise strength and rate of expression of wild-type and TATA mutant *PHO5* promoters at maximal induction (left), and scatter plot for wild-type (O) and TATA-C2 (◆) (right). The intrinsic noise of the TATA-C2 promoter is 15% at a mean of 25 AU. Error bars represent standard deviations. Scatter plots contain cells from multiple time points.



data from the TATA box variants of *PHO5* allow calculation of the maximal noise contribution from transcription and translation combined; this noise strength value is found at the limit as the k_m or transcriptional efficiency approaches zero (see the noise strength equation above). This value for *PHO5* is less than 0.25 AU, substantially smaller than the measured noise strength of the *GAL1* promoter of ~1 AU. If *GAL1* represents case III then *GAL1* must display a much higher noise strength from transcription than *PHO5*. The magnitude of the intrinsic noise strength contribution from transcription and translation is in theory dependent only on the translational efficiency of the reporter mRNA (Eq. 8). The *PHO5* and *GAL1* reporter mRNAs differ only in the 5' untranslated (5'UTR) region, and therefore in order for *GAL1* to fit the profile of case III, the 5'UTR of the *GAL1* reporter mRNA must confer a substantially higher translational efficiency than the 5'UTR of the *PHO5* reporter mRNA. Because the start sites of transcription have been identified for both *PHO5* and *GAL1* (Giardina and Lis, 1993; Svaren and Horz, 1997), we were able to replace the entire 5'UTR regions of the two promoters in order to test if the 5'UTR from *GAL1* conferred a substantial increase in both the rate of transcription and stochastic noise strength on the *PHO5* upstream promoter sequence, and vice versa. We found that the promoters retain similar rates of induction and noise strength independent of 5'UTR identity (Fig. 1.11). Therefore we can hypothesize that the noise strength of *GAL1* does contain a substantial contribution from a promoter activation step prior to transcription, and corresponds to case II rather than case III. This assertion is consistent with the known role of chromatin-remodeling complexes in the activation of the *GAL1* promoter (Larschan and Winston, 2001; Santisteban et al., 2000).

Discussion

Our results support an extended general model for noise generation applicable to both eukaryotic and prokaryotic promoters, in which the kinetic constants of promoter activation, deactivation, and transcription determine the contribution of stochastic promoter activation to variability in gene expression. In such a model, two promoters can produce the same mean mRNA population with very different noise characteristics: a promoter that undergoes frequent activation steps followed by inefficient transcription will produce a cellular population with little variability, while a promoter that undergoes infrequent activation steps followed by efficient transcription can display very large differences from cell to cell. We have identified simple sequence mutations in the *PHO5* promoter that exemplify these two extremes: the *PHO5-UASm1* and the *PHO5-TATA-A1T6* promoters have approximately equal rates of gene expression on a population level, but strikingly divergent levels of stochasticity, a more than 30-fold change in intrinsic noise strength. Because the intrinsic noise is altered by small changes in promoter sequence and independent of the absolute rate of gene expression, these mutants suggest that stochasticity is an evolvable characteristic of each eukaryotic gene, determined by both *cis-* and *trans*-acting factors.

A previously proposed model for noise generation in prokaryotes (Ozbudak et al., 2002; Thattai and van Oudenaarden, 2001) suggested that the stochastic contribution of transcription to noise strength is modulated by the translational efficiency. In such a model, the actual stochasticity of translation contributes minimally to the variability in protein levels, but the measured noise strength scales with translational efficiency. This scaling factor led the authors to propose that, given a set of genes with equivalent stochasticity in transcription, those genes that are relatively inefficiently translated will

be less phenotypically variable. Our model encompasses this possibility, and suggests additional situations where the stochasticity of transcription contributes little to population variability relative to the stochasticity of promoter activation. In these cases, the intrinsic noise strength of gene expression will scale with the efficiency of the transcriptional, as well as translational, step. This allows direct selection of the noise in mRNA levels without modulation of translational efficiency.

Both our noise measurements and our model contradict the previous assertion that noise generation in gene expression is fundamentally different between prokaryotes and eukaryotes (Blake et al., 2003). Rather, we assert that the diverse mechanisms of gene regulation in prokaryotic and eukaryotic systems fall into three general categories of noise profiles, dependent upon the relative frequency of kinetic steps in promoter activation, deactivation, and transcription. Eukaryotic regulatory mechanisms such as chromatin remodeling may provide more flexibility in determining the relative rates of these reactions, but do not necessarily represent novel noise control abilities when compared to prokaryotic mechanisms.

Stochasticity in gene expression has been postulated to create substantial phenotypic variation in many biological cases. For example, stochastic fluctuations in the expression of regulatory proteins have been suggested to affect the lysis-lysogeny switch in populations of lambda phage-infected *E. coli* (Arkin et al., 1998); to create functional diversity in populations of mammalian olfactory neuronal cells (Serizawa et al., 2003); and to promote tumor formation in response to haploinsufficiency of a transcription factor (Magee et al., 2003). The importance of stochasticity in gene expression in these and other cases has yet to be tested directly.

The relative contribution of stochasticity in gene expression to the total noise is small in nearly all measurements made in budding yeast. However, we believe that measurement techniques based on either microscopy or flow cytometry may overestimate the physiologic extrinsic noise relative to the intrinsic noise (supplemental text online). Importantly, even if extrinsic noise dominates the contribution to physiologic variability in the amount of gene expression, the stochasticity intrinsic to gene expression is the only type of noise capable of generating substantial differences between two different alleles in a heterozygous diploid organism. If the gene products of the two alleles are functionally distinct, then intrinsic noise can create diversity due to this difference in gene product activity. Our results and the proposed model suggest that stochasticity in gene expression is not necessarily an obstacle to invariant cellular behavior, but that in a diploid population, with multiple heterozygous genetic loci, stochastic events in gene expression may constitute a fundamental source of advantageous population diversity.

Materials and Methods

YEAST STRAINS AND MEDIA

The coding sequence for yeast-optimized enhanced cyan fluorescent protein (yECFP) was constructed by incorporation of the following mutations into the yEGFP1 sequence (Cormack et al., 1997): Y66W, N146I, M153T, V163A, S175G (Crameri et al., 1996; Heim and Tsien, 1996; Siemering et al., 1996). Yeast-optimized Venus yellow fluorescent protein (yVYFP) was constructed by incorporation of the following mutations into the yEGFP3 sequence (Cormack et al., 1997): F46L, F64L, V68L, Q69M, M153T, V163A, S175G, T203Y (Griesbeck et al., 2001; Nagai et al., 2002). The excitation

and emission spectra were verified by fluorimetry of bacterial lysates expressing the yECFP and yVYFP proteins. A yeast chromosomal integration vector, pJRL2, that replaces the *LEU2* chromosomal locus by homologous recombination, was constructed by insertion of the 200 bp upstream of the *LEU2* start codon, an Asc I restriction enzyme site, the 200 bp downstream of the *LEU2* stop codon and a 5.5 kb *hisG::URA3::kanR::hisG* cassette (Alani et al., 1987) into pBluescript (Stratagene).

All constructs for dual-reporter noise measurements contained 1000 bp of promoter sequence, from the nucleotide immediately upstream of the relevant start codon, followed by six base-pairs of idealized Kozak sequence AACAAA for optimal translational initiation (Miyasaka, 1999), the coding sequence of the appropriate fluorophore (yECFP or yVYFP) and the 500 bp 3' to the stop codon of *ACT1* in the pJRL2 integrating vector. Constructs expressing YFP were integrated into W303-derivative EY1555 (*MAT***a** *trp1 HIS3 ADE2*), constructs expressing CFP were integrated into W303-derivative EY1556 (*MAT*α *TRP1 his3 ADE2*), and diploids were obtained by mating and double selection.

PHO5 promoter variants were constructed by PCR using primers containing the desired mutations. UASm1 and UASm2 mutants have been previously described (Steger et al., 2003). The TATA variant nomenclature indicates the mutation introduced into the wild-type PHO5 TATA box, e.g. TATA-A1 converts TATATAAG to AATATAAG. *PHO5-GAL1* and *GAL1-PHO5* fusions were constructed by PCR using primers containing the desired junction sequences (for *GAL1-PHO5*, the *PHO5* promoter sequence from -83 to the -1 position relative to the *PHO5* ATG was fused to the upstream portion of the *GAL1* promoter after the -129 position relative to the *GAL1* ATG; for *PHO5-GAL1*, the *GAL1*

promoter sequence from -128 to the -1 position relative to the *GAL1* ATG was fused to the upstream portion of the *PHO5* promoter after the -84 position relative to the *PHO5* ATG; both fusions preserve the native promoter TATA box). The yeast deletion strains were made as described previously (Steger et al., 2003) from strains containing integrated *PHO5prCFP* and *PHO5prYFP*. The chemical inhibitor-sensitive (*PHO85*^{F82G}) strains containing *PHO5prCFP* and *PHO5prYFP* were constructed as described previously (Steger et al., 2003).

Unless otherwise indicated, yeast strains were grown at 30°C in synthetic complete medium containing glucose with addition of 0.1 mg/mL adenine and 0.1 mg/mL tryptophan to suppress autofluorescence. The *PHO84* promoter was induced with synthetic phosphate-free medium containing levels of inorganic

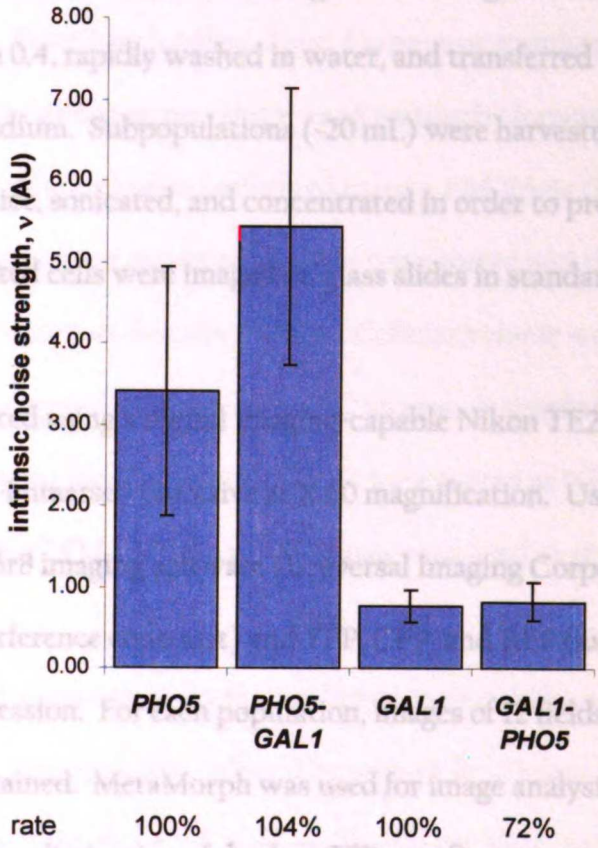


Figure 1.11. Exchange of 5' untranslated regions of *GAL1* and *PHO5* promoters. Intrinsic noise strength was measured for the *PHO5* and *GAL1* promoters and for fusion promoters containing the *PHO5* upstream sequences and *GAL1* transcription start site region (*PHO5-GAL1*), or the *GAL1* upstream sequence followed by the *PHO5* transcription start site region (*GAL1-PHO5*). The rates of gene expression as a percentage of the rates of the wild-type *PHO5* and *GAL1* promoters are shown below the graph.

phosphate ranging from 0 to 500 micromolar (Springer et al., 2003). The *GAL1* promoter was induced with medium containing 2% raffinose and levels of galactose from 0 to 2%. The *PHO5* promoter was induced either in medium containing levels from 0 to 10 micromolar of a chemical inhibitor (1-NaPP1) of a specific allele of *PHO85* (Carroll et al., 2001); or in synthetic phosphate-free medium containing levels of phytic acid from 0 to 600 micromolar as an organic phosphate source.

NOISE MEASUREMENTS

For noise measurement time-courses, cells were grown overnight in 50 mL batch cultures to an OD_{600} of less than 0.4, rapidly washed in water, and transferred to 200 mL of the appropriate induction medium. Subpopulations (~20 mL) were harvested at different time points, placed on ice, sonicated, and concentrated in order to prepare for microscopy. The chilled, sonicated cells were imaged on glass slides in standard medium.

Microscopy was conducted using a digital imaging-capable Nikon TE200/300 inverted microscope with an oil-immersed objective at X 60 magnification. Using a script in MetaMorph version 4.6r8 imaging software (Universal Imaging Corporation), Nomarski/DIC (differential interference constrast) and YFP, CFP, and RFP fluorescence images were taken in rapid succession. For each population, images of 12 fields containing 80-150 cells were obtained. MetaMorph was used for image analysis, which included background subtraction, elimination of dead or CFP-autofluorescent cells by RFP channel fluorescence (~5% of cells), and quantitation of CFP and YFP values for each individual cell. The CFP population mean value was scaled to the YFP population mean for each independent population of cells, and a maximal induction control was performed in parallel for each experiment in order to account for daily variation in

fluorescence intensity measurements (though such variation was minimal). For low level inductions of the *GAL1* promoter, bimodal populations were scaled individually. Less than 0.1% of all cells clearly expressed only one of the two fluorophores, and were eliminated from analysis. Rates of expression were calculated from linear fits of the increase in mean population fluorescence over time. All average noise strength values presented are the average of values from a single induction time course. All experiments were repeated at least three times and resulted in the same trends each time. Bootstrap values (90%) for Fig. 1C were calculated by random selection of populations of the same size as the original populations. Error bars for Figs. 2 and 4 represent standard deviations. Flow cytometry fluorescence measurements and sorting by forward scatter and side scatter (‹5% of total population) were performed using a FACSAria (BD Biosciences). For flow cytometry sorting experiments, cells were sorted and then processed for fluorescence microscopy as described above. Cellular volume was estimated by measurement of the area of each cell from DIC images. A simple budding index, in which cells that were unbudded or had small buds were scored as "G1" and cells with large buds were scored as "S/G2/M" based on DIC images, was used to stratify by cell cycle stage.

For the measurement of correlation of expression between two different promoters, expression from the promoters was induced by growth for 6 hours in synthetic phosphate-free medium containing 2% dextrose for the *PHO84/ADH1* measurements, and by synthetic phosphate-free medium containing 2% galactose for the *PHO84/GAL1* measurements.

STOCHASTIC SIMULATIONS

Efficient exact stochastic simulations were performed according to the Next Reaction Method (Gibson and Bruck, 2000; Gillespie, 1977) and implemented in MatLab (The MathWorks). The base constants chosen for the simulations were as follows in the form $\{k_a\ k_m\ k_p\ \gamma_a\ \gamma_m\ \gamma_p\}$: case I, {0.1 10 10 0.1 5 0.1}; case II, {0.53 100 10 10 5 0.1}; case III, {53 100 10 1000 5 0.1}. The forward rates of promoter activation, transcription, and translation were varied individually as necessary to produce a range of steady-state protein levels. Induction time course simulations were analyzed at t = 100, 200, 300 for cases I and II, and t = 10, 20, 30 for case III with independent populations of 100 simulated cells with a single gene. No extrinsic noise factors were included in the simulation, and therefore the intrinsic noise strength was calculated from the total noise strength.

CHAPTER TWO:

Origins, control, and consequences of noise in gene expression

Introduction

Every individual is unique and special. Not just a truism for self-help books and motivational speakers, this statement applies to any population of living organisms or cells (Fig. 2.1). Much of population variability is due to genetic differences. Still, genetically identical populations display variability in cellular phenotype due to differences in environment and history. For example, development and differentiation in multicellular organisms results in clonal populations of cells with amazingly diverse functions. However, even cells or organisms with the same genes, in the same environment, with the same history, display differences in form and behavior both subtle and dramatic. Recently, investigations have focused on the possibility that such variability is inevitable in biological systems because of the random nature of molecular interactions (Rao et al., 2002). In this view, the cell is best described as a hodgepodge of molecules undergoing stochastic chemical reactions. Each reaction is governed by a stochastic rate that describes the probability of molecules reacting in a certain time period (Gillespie, 1977; van Kampen, 1992). When large numbers of molecules are present, the reaction may proceed in a predictable manner. However, when only a few molecules of a specific type exist in a cell, stochastic effects can become much more prominent. This prominence is due to the relative increase in the fluctuation size, which is generally proportional to the square root of the average number of molecules (van Kampen, 1992). This stochasticity could create variability in cellular phenotype in an otherwise identical population. Because phenotypic differences could be either advantageous or deleterious for a clonal population, it is likely that cellular processes either harness or suppress the consequences of random molecular fluctuations to maximize fitness.

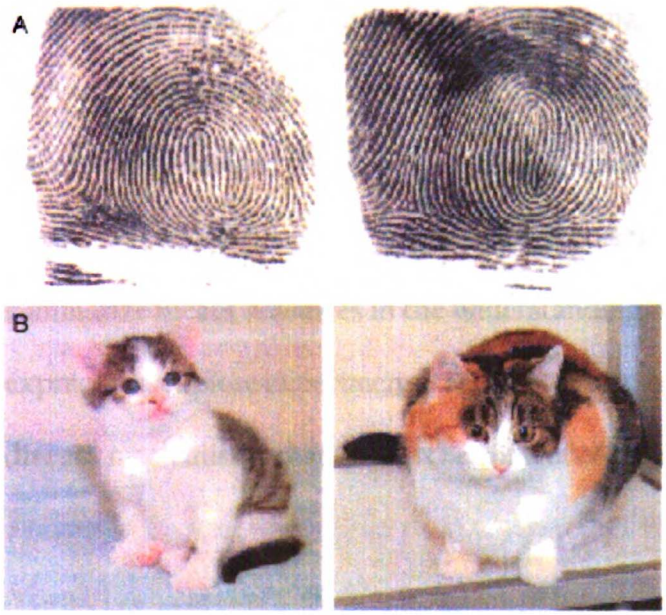


Figure 2.1. Examples of possible stochastic influences on phenotype. (A) The fingerprints of identical twins are readily distinguished on close examination Reprinted from Pattern Recognition 35, A. K. Jain et al., On the similarity of identical twin fingerprints, 2653-63, (2002), with permission from Elsevier.. (B) Cc, the first cloned cat (left) and Rainbow, Cc's genetic mother (right) display different coat patterns and personalities T. Shin et al., Nature 415, 859 (2002). Photo credit of College of Veterinary Medicine and Biomedical Sciences, Texas A&M University..

Gene expression, defined as the set of reactions that control the abundance of gene products, influences most aspects of cellular behavior. Therefore, variation in gene expression is often invoked to explain phenotypic differences in a population of cells. The abundance of gene products is theoretically sensitive to infrequent events involving small numbers of reactants, because DNA, RNA, and proteins can be present and active at few copies per cell (Holland, 2002). However, stochastic molecular events during gene expression are not the only potential source of variability in gene expression, nor must they inevitably result in substantial, detectable variation in gene product level. Broadly, four different sources of variation in gene expression must be considered (Fig. 2.2): (A) the stochasticity or randomness inherent in biochemical processes that are dependent on infrequent molecular events involving small numbers of molecules; (B) variation in gene expression due to differences in the internal states of a population of cells, either from predictable processes such as cell cycle progression, or from a random process such as partitioning of mitochondria during cell division; (C) subtle environmental differences, such as morphogen gradients in multicellular development, or local microenvironments in a

colony of unicellular organisms; and (D) ongoing genetic mutation, either random or directed, as in somatic hypermutation in human B cells. We use the term "noise" in gene expression to refer to the measured level of variation in gene expression among cells, regardless of source, within a supposedly identical population. In this review, we summarize recent advances in the understanding of the sources of noise in gene expression, explore experimental and theoretical consequences of such noise, and discuss mechanisms that enable the control of noise.

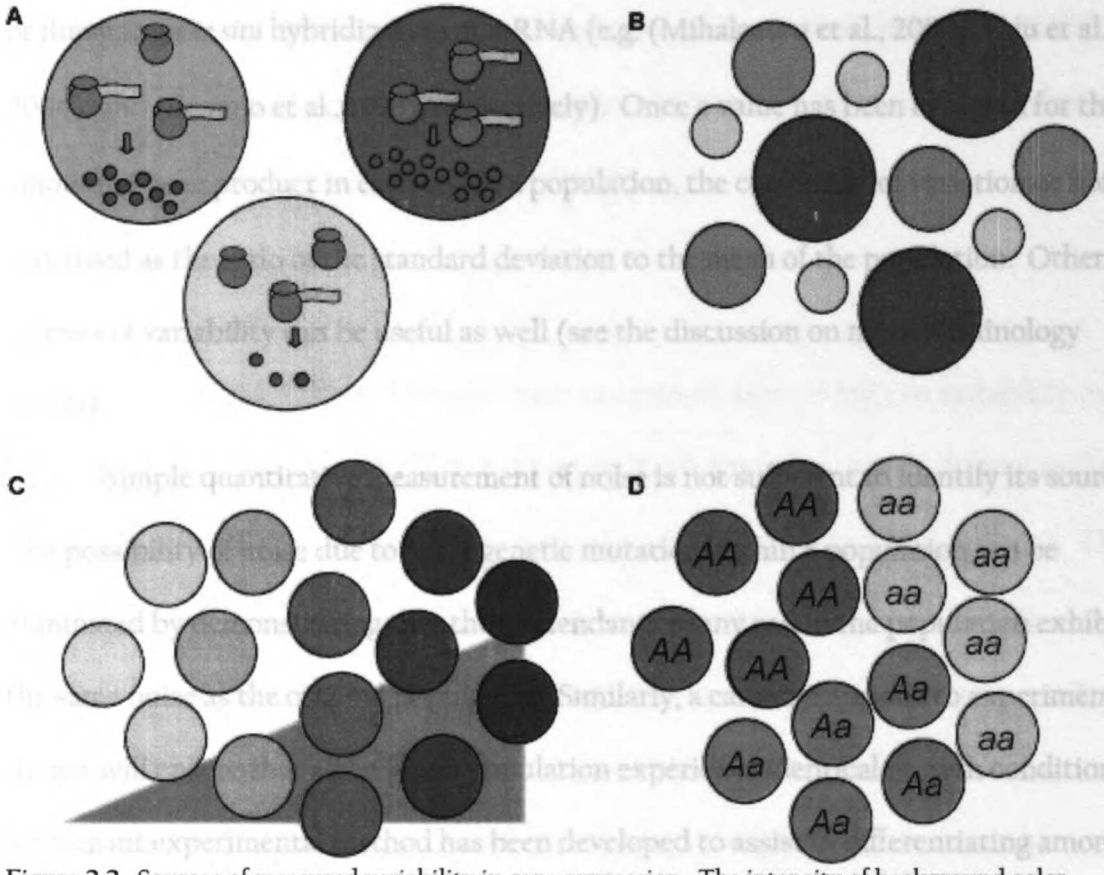


Figure 2.2. Sources of measured variability in gene expression. The intensity of background color represents gene product abundance within each cell. (A) Stochastic events in gene expression due to small numbers of molecules; in this example, fluctuations in the abundance of a rare mRNA create differences in protein levels among cells. (B) Heterogeneity or extrinsic noise in gene expression; the larger cells express more gene product. (C) Local microenvironments, such as a morphogen gradient, result in differences in a population. (D) Rapid genetic mutation in a population results in variability in gene expression.

Measurement techniques

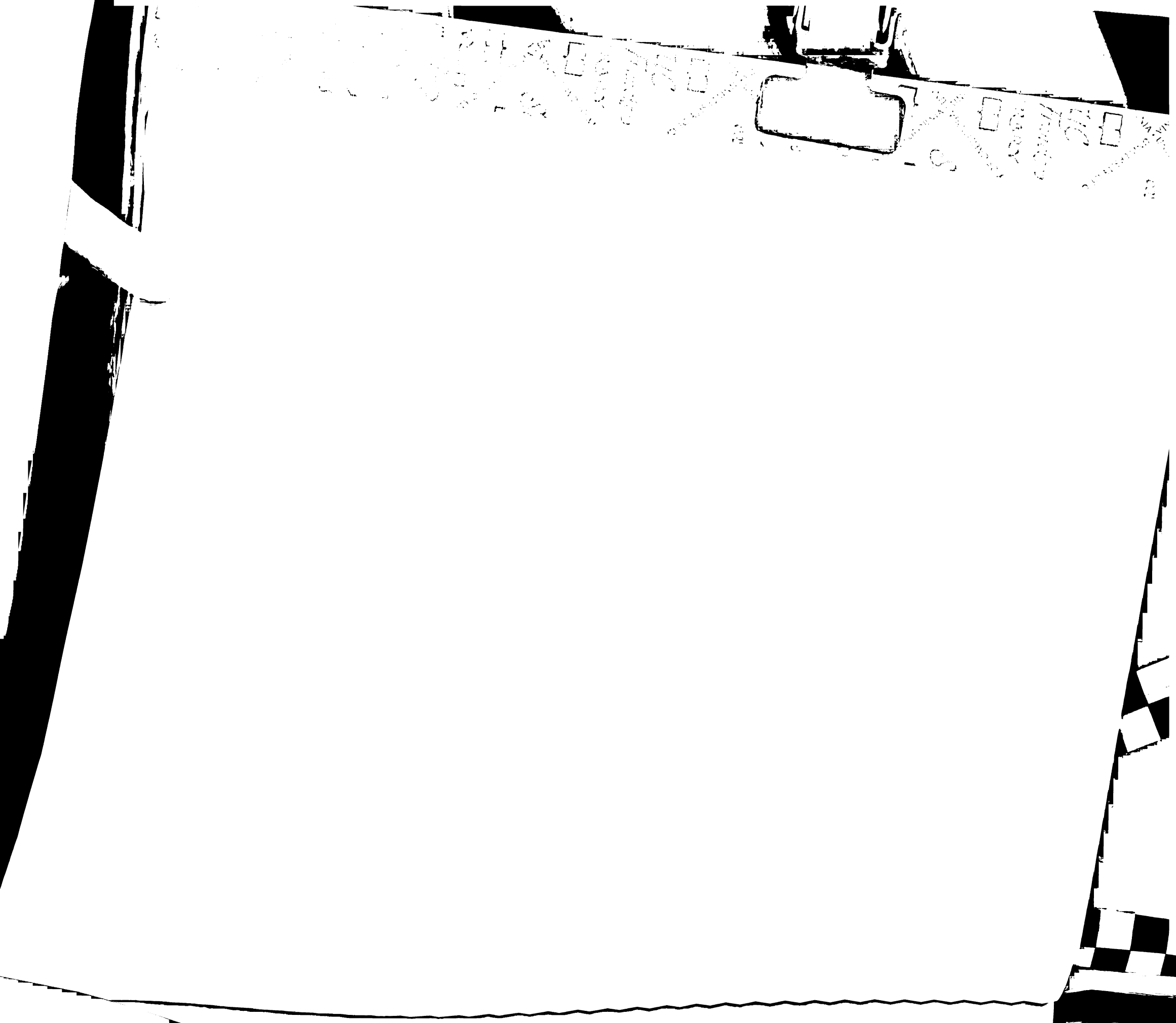
Techniques that precisely quantify the abundance of gene products in single cells are essential for the measurement of noise in gene expression. Useful assays both preserve a linear relationship between the amount of gene product and the measured signal, and have little error in measurement. Recent investigations have employed green fluorescent protein (GFP) variants, which allow the precise quantification of protein levels in living cells by flow cytometry or fluorescence microscopy (Chalfie et al., 1994). Additionally, gene product levels in single cells may be measured by bioluminescent reporter proteins, reverse transcription of mRNA followed by amplification (RT-PCR), or fluorescent *in situ* hybridization of mRNA (e.g. (Mihalcescu et al., 2004), (Liu et al., 2004), and (Femino et al., 1998), respectively). Once a value has been obtained for the amount of gene product in each cell in a population, the coefficient of variation or noise η is defined as the ratio of the standard deviation to the mean of the population. Other metrics of variability can be useful as well (see the discussion on noise terminology below).

Simple quantitative measurement of noise is not sufficient to identify its source. The possibility of noise due to rapid genetic mutation within a population can be eliminated by demonstrating that the descendants of any cell in the population exhibit the same noise as the original population. Similarly, a careful approach to experimental design will ensure that all cells in a population experience identical growth conditions. An elegant experimental method has been developed to assist in differentiating among the remaining sources of noise (Swain et al., 2002). This method involves quantifying expression of two equivalent, independent gene reporters placed in the same cell, which then allows noise sources to be partitioned into two categories: intrinsic, meaning noise

sources that create differences between the two reporters within the same cell (Fig. 2.3A); and extrinsic, referring to sources that impact the two reporters equally in any given cell, but create differences between two cells (Fig. 2.3B). Stochastic events during the process of gene expression, from the level of promoter-binding to mRNA translation to protein degradation, will manifest as intrinsic noise. Differences between cells, either in local environment, or in the concentration or activity of any factor that affects gene expression, will result in extrinsic noise. Extrinsic noise should be further subdivided into two categories (Pedraza and van Oudenaarden, 2005; Raser and O'Shea, 2004): global noise, or fluctuations in the rates of the basic reactions of gene expression that affect all genes (Fig. 2.3C); and gene- or pathway-specific extrinsic noise (Fig. 2.3D), such as fluctuations in the abundance of a particular transcription factor or stochastic events in a specific signal transduction pathway.

Noise terminology

The term **noise** is used in two ways: as a generic term to refer to variability in a population, and more specifically to refer to the coefficient of variation, defined as the standard deviation of a population (σ) divided by the mean of the population (μ). The coefficient of variation (σ/μ) provides a straight-forward estimate of the overall population variability. For example, a normally distributed population with a coefficient of variation or noise of 30% will contain ~17 individuals per 100 whose expression level is less than 70% of the mean. The **normalized variance** (σ^2/μ^2), or the square of the noise, is often used for both theoretical predictions and experimental measurements, as normalized variance from independent sources is additive. When more than one source



contributes to the total noise, normalized variance is a valuable metric that allows understanding of the relative contributions of different noise sources (Paulsson, 2004; Pedraza and van Oudenaarden, 2005). The **noise strength** (σ^2/μ), defined as the variance of the population divided by the mean of the population, has been used to compare experimental data to predictions derived from stochastic models (Blake et al., 2003; Ozbudak et al., 2002; Raser and O'Shea, 2004; Thattai and van Oudenaarden, 2001). Unless the measured signal is calibrated in terms of molecules, the noise strength is reported in arbitrary units that are meaningful only for comparative purposes.

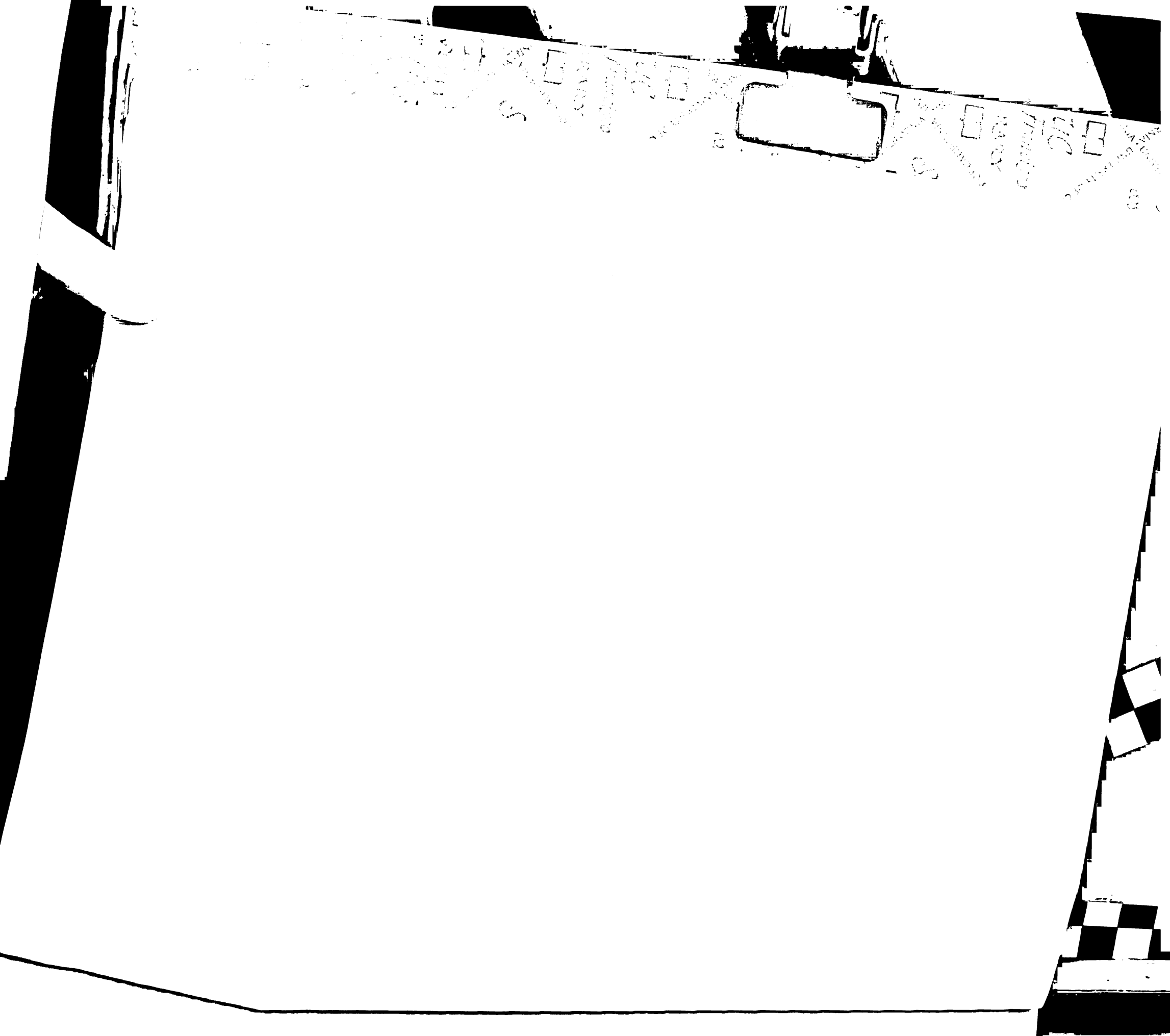
The **intrinsic noise** is defined as noise that creates differences in two identical reporters of gene expression contained within the same cell (Figs. 2.2A and 2.3A) (Swain et al., 2002). Intrinsic noise may be attributed to any stochastic event during gene expression, from the level of promoter-binding events, to mRNA splicing, to translation, to partitioning of proteins during cell division. Intrinsic noise may also arise from differences in reporter copy number or gene replication timing.

Extrinsic noise affects two identical reporters of gene expression within a given cell equally, but generates differences in reporter expression among cells (Figs. 2.2B and 2.3B) (Swain et al., 2002). Extrinsic noise may be attributed to any cellular heterogeneity that affects the processes of gene expression; this heterogeneity may be stochastic in nature, or may be due to a predictable process such as cell cycle progression. Extrinsic noise may also result from stochastic events in gene expression that occur prior to promoter binding, such as stochastic autophosphorylation of a receptor tyrosine kinase that results in the subsequent activation of a transcription factor. Extrinsic noise can be divided into two types: **global extrinsic noise**, which is extrinsic noise that affects all genes in a cell equally (Fig. 2.3C), and **gene- or pathway-**

specific extrinsic noise, which affects a subset of genes in the cells (Fig. 2.3D). Global extrinsic noise may be due to fluctuations in factors that affect the expression of all genes, such as nucleotide concentration or (for proteins) cellular ribosome number. Gene-specific extrinsic noise may be due to fluctuations in the specific regulators of a gene product, such as a specific transcription factor.

Two parameters of noise have been measured experimentally: the **magnitude** and the **autocorrelation time** of the noise (Rosenfeld et al., 2005). The autocorrelation time refers to the average amount of time over which a single cell in a noisy population retains a similar level of gene expression. While the magnitude of noise can be measured from a single observation, calculation of an autocorrelation time requires monitoring the same population of cells over time. The autocorrelation time is pertinent both to the source of the noise and the biological consequence. For example, an autocorrelation time of ~5 minutes is not consistent with the noise of partitioning during cellular division if the cell-cycle length is 24 hours. The effect of noise on cellular fitness must be considered in the context of how long the noise persists relative to how quickly selective pressure acts in a population. Populations can have different noise magnitudes and noise autocorrelation times (Fig. 2.3E-G). A more complete description of noise and its consequences would include noise magnitude, duration, and frequency spectrum (Samoilov et al., 2005).

Finally, it should be emphasized that in the traditional physical terminology, noise is equated with stochastic factors; however, there may exist predictable factors that result in apparent extrinsic noise in biology. For example, in a population of cells growing asynchronously, differences in cell cycle position may generate differences in gene copy number that directly affect gene expression. If the cell cycle length is



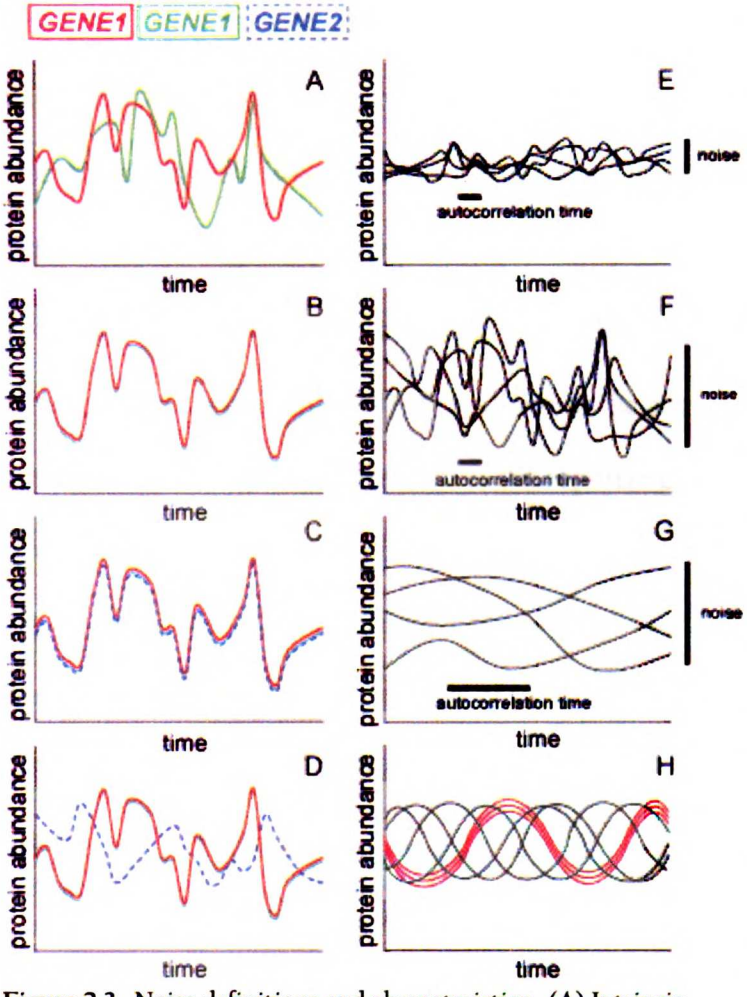
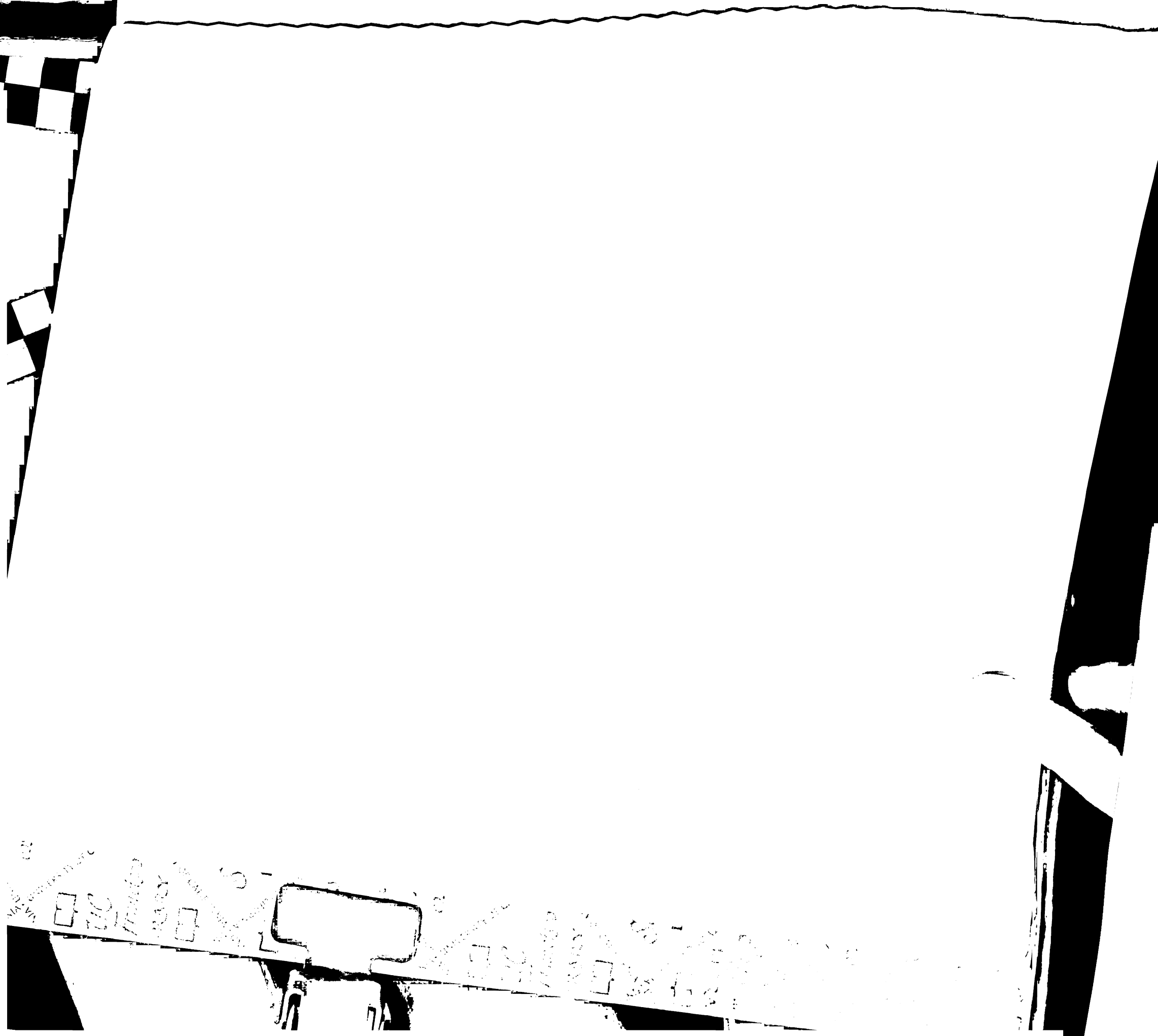


Figure 2.3. Noise definitions and characteristics. (A) Intrinsic noise will result in differences between two reporters of the same gene in a single cell. (B) Extrinsic noise will affect two reporters of the same gene equally in a single cell, but will cause differences from cell to cell or in a single cell over time. (C) Global noise affects two distinct genes equally but results in differences from to cell to cell or in a single cell over time. (D) Gene- or pathway-specific extrinsic noise affects two reporters of the same gene equally but causes differences from a reporter of a second distinct gene in a single cell. (E-H) Noise in a population; each line represents a different cell. (E) Noise of low magnitude and short autocorrelation time. (F) Noise of high magnitude and short autocorrelation time. (G) Noise of high magnitude and long autocorrelation time. (H) "Deterministic" noise: in a synchronized population of cells, cell cycle progression results in predictable changes in protein abundance over time (red lines); when the cells grow asynchronously, the population displays variability (black lines).

consistent within this population, each cell will display a predictable pattern of gene expression; the population will display noise in gene expression that is in fact "deterministic" at the level of a single cell (Fig. 2.3H). In other words, knowledge of the extrinsic factor that affects the efficiency of gene expression allows the elimination of extrinsic noise.

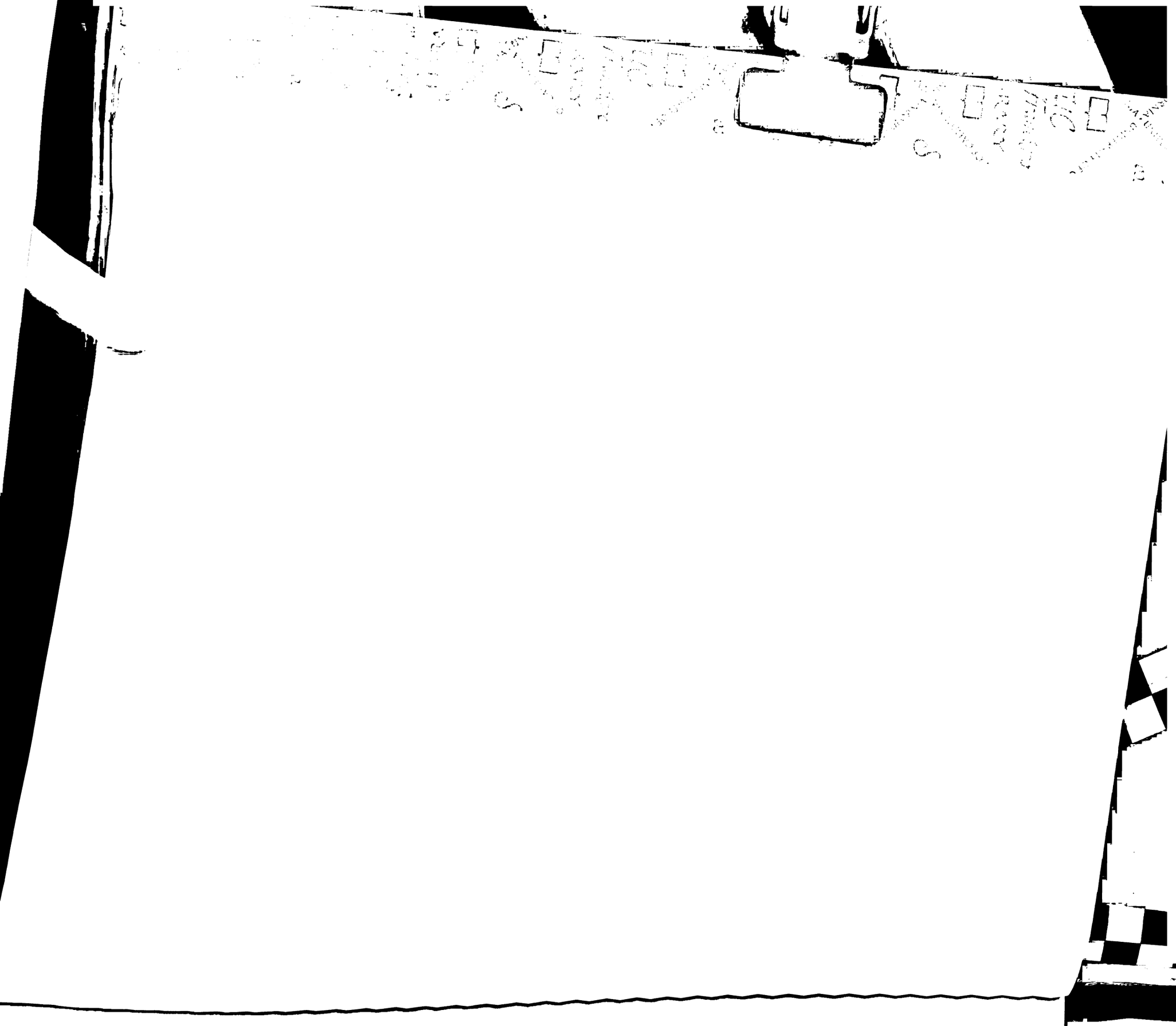
Experimental investigation of noise sources

The division of noise into 'extrinsic' and 'intrinsic' categories has proven practical experimentally. Elowitz *et al.* pioneered the two-reporter method in studies of noise in gene expression in *Escherichia coli* which quantified levels of cyan and yellow fluorescent proteins expressed from identical promoters on the same prokaryotic chromosome (Elowitz et al., 2002).



These studies demonstrated that the stochastic nature of gene expression gives rise to noise in protein levels in a clonal population of *E. coli*, and that the relative contributions of extrinsic and intrinsic components to the total noise vary with expression level. Ozbudak *et al.* used a single reporter to quantify noise in gene expression in *Bacillus subtilis* (Ozbudak et al., 2002). In this system, the extrinsic noise fortuitously appears to be negligible (Paulsson, 2004), thus allowing the attribution of single-reporter noise to stochastic events in the expression of the reporter gene. By comparing the noise observed in reporters with altered efficiency of transcription and translation, they concluded that prokaryotic transcription is the dominant source of noise in protein levels, as predicted by basic models of stochastic gene expression (Kepler and Elston, 2001; Kierzek et al., 2001; McAdams and Arkin, 1997; Thattai and van Oudenaarden, 2001).

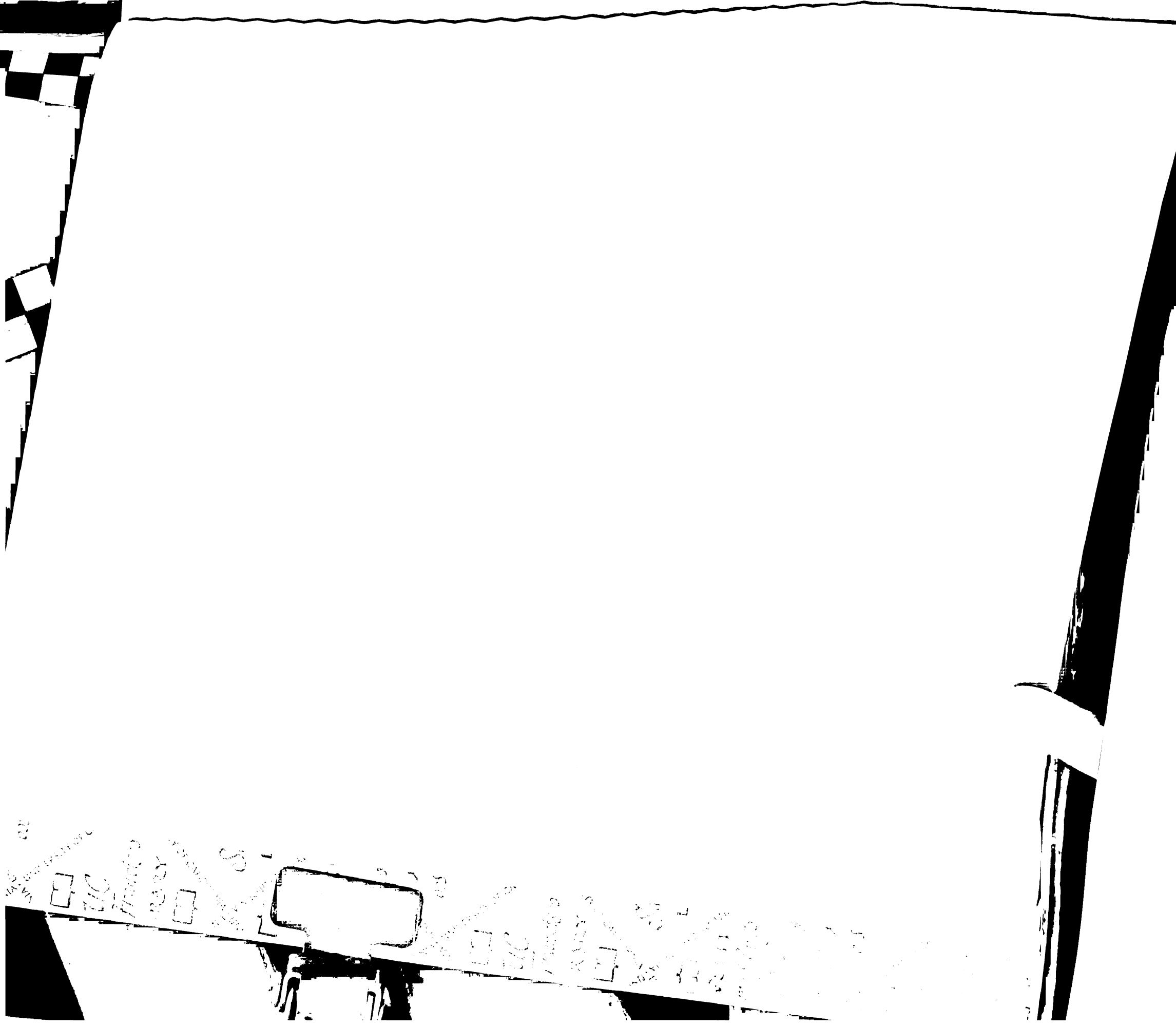
We have employed the two-reporter system to measure gene expression noise in cells of the diploid eukaryote *Saccharomyces cerevisiae* (Raser and O'Shea, 2004). These studies revealed that intrinsic noise in reporter protein levels is detectable, gene-specific and not determined exclusively by expression level or transcription factor identity. For one gene, intrinsic noise results from slow interconversion between inactive and active promoter states due to stochastic chromatin-remodeling events. However, extrinsic noise is the predominant form of noise for all gene promoters measured in these experiments. Surprisingly, simultaneous measurement of two independent, unrelated gene promoters indicated that much of this extrinsic noise is global in nature, presumably due to fluctuations in some factor that affects expression of all genes, and not due to fluctuations in extrinsic factors that affect a particular gene. Blake *et al.*



quantified noise in *S. cerevisiae* using a single-reporter method (Blake et al., 2003). Their results for the *GAL1* gene are consistent with the extrinsic noise profile of *GAL1* measured by the two-reporter method, suggesting that this noise is not the result of stochastic chromatin-remodeling or transcription. Rather, most of the single-reporter noise is likely due to extrinsic factors such as global noise or noise in *GAL* signaling.

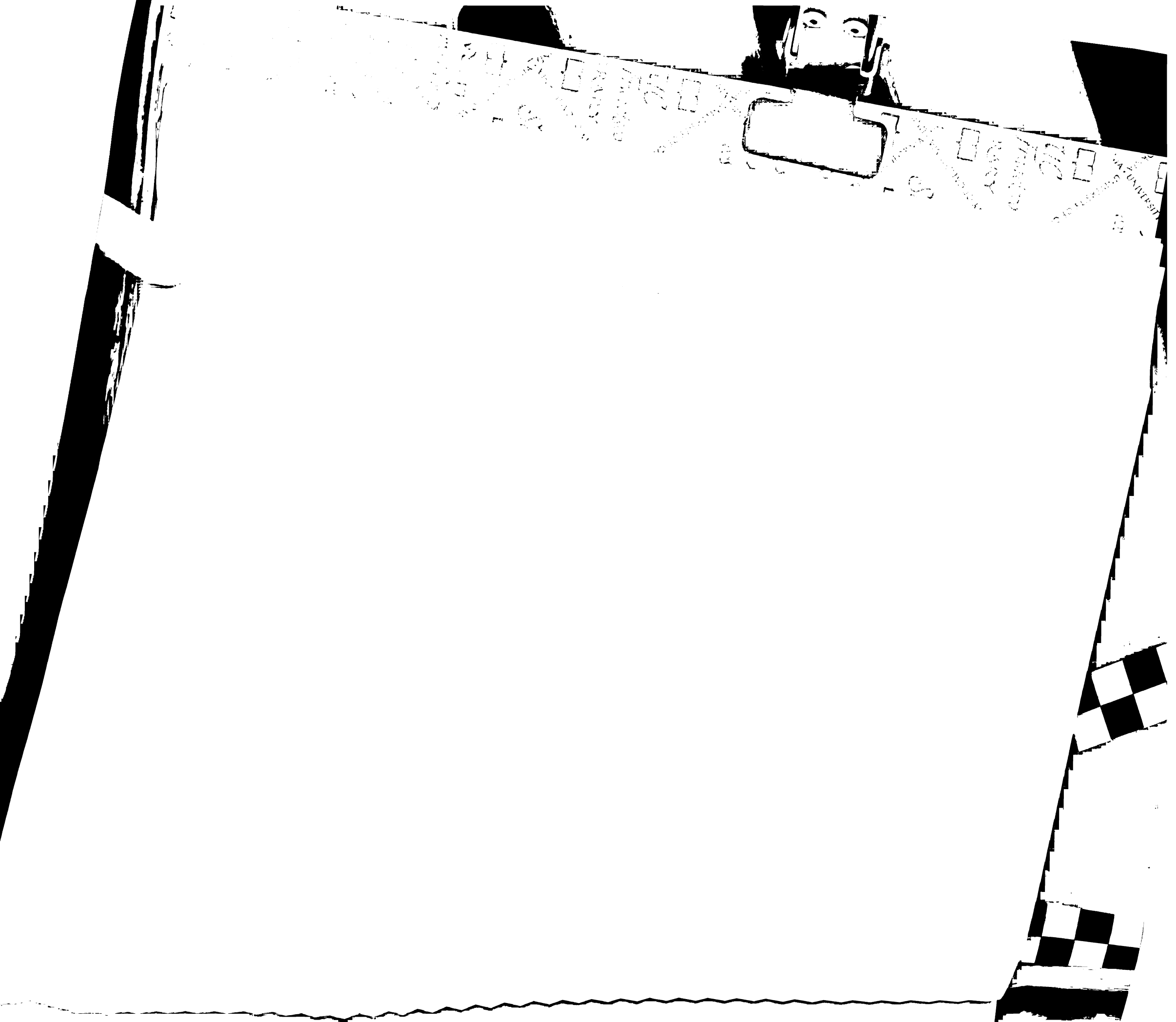
Recent measurements of gene expression in single *E. coli* cells over long periods of time have provided insights into the relative amplitude and time-scales of intrinsic and extrinsic noise (Rosenfeld et al., 2005). In these measurements, extrinsic noise is the primary source of variability in gene expression, similar to the observation in budding yeast. The authors calculated autocorrelation times for noise, or the time-scale over which the protein production rate fluctuates in any given cell (*cf.* Fig. 2.3E-G). The autocorrelation time for intrinsic noise is equal to or less than 10 minutes, consistent with the hypothesis that rapid fluctuations in mRNA numbers are the source of intrinsic noise. The autocorrelation time for global noise factors in protein production rate is ~40 minutes, similar to the observed cell cycle length, suggesting that whatever factors result in global noise persist on average for about one cell cycle. Protein production rates were normalized to account for DNA replication, so this unknown global factor is not simply variability in gene copy number.

Experiments in two metazoan organisms have produced results consistent with ongoing stochastic gene expression. In *Drosophila melanogaster*, Ahmad and Henikoff studied the behavior of a GFP reporter subject to position effect variegation (Ahmad and Henikoff, 2001). The reporter gene displays an expression pattern suggestive of repeated rounds of stochastic activation and inactivation of gene expression, resulting in patches



of cells expressing the reporter. The frequencies of activation and inactivation are sensitive to the level of the transcription factor; more transcription factor results in more frequent activation and less frequent inactivation. The authors propose that fluctuations in the chromatin state of the reporter gene uncover transcription factor binding sites, and that recruitment of chromatin-modulating activities subsequent to transcription factor binding can slow the rate of heterochromatin reformation. In another eukaryotic system, two groups reported that each allele of the gene encoding the cytokine IL-4 is expressed in a probabilistic manner in response to signaling through the T cell receptor (TCR) in mouse T helper 2 lymphocytes (Bix and Locksley, 1998; Riviere et al., 1998). The probability of expression for each allele is independent of its parental origin and increased with the strength of TCR stimulation, leading to biallelic expression at higher levels of stimulation. Furthermore, the pattern of allelic expression is different over multiple rounds of activation for some clonal populations, suggesting that the stochastic gene activation observed is a reversible process. Other mammalian genes may show similar patterns of both monoallelic and biallelic expression.

These reports of ongoing stochastic gene expression, in which the probability of expression of each allele is independent, contrast with the many examples of fixed monoallelic expression, either random or parentally imprinted (Rakyan et al., 2001). Random monoallelic expression relies on stochastic choice followed by a mechanism to stabilize that choice and involves epigenetic markers such as heterochromatin or DNA methylation. However, a stochastic choice does not necessarily rely on stochastic events in gene expression. For instance, prior to the random inactivation of one allele of the B cell receptor, the two allelic loci display different replication timing. The earlier replicating allele correlates with the subsequently active allele (Mostoslavsky et al.,

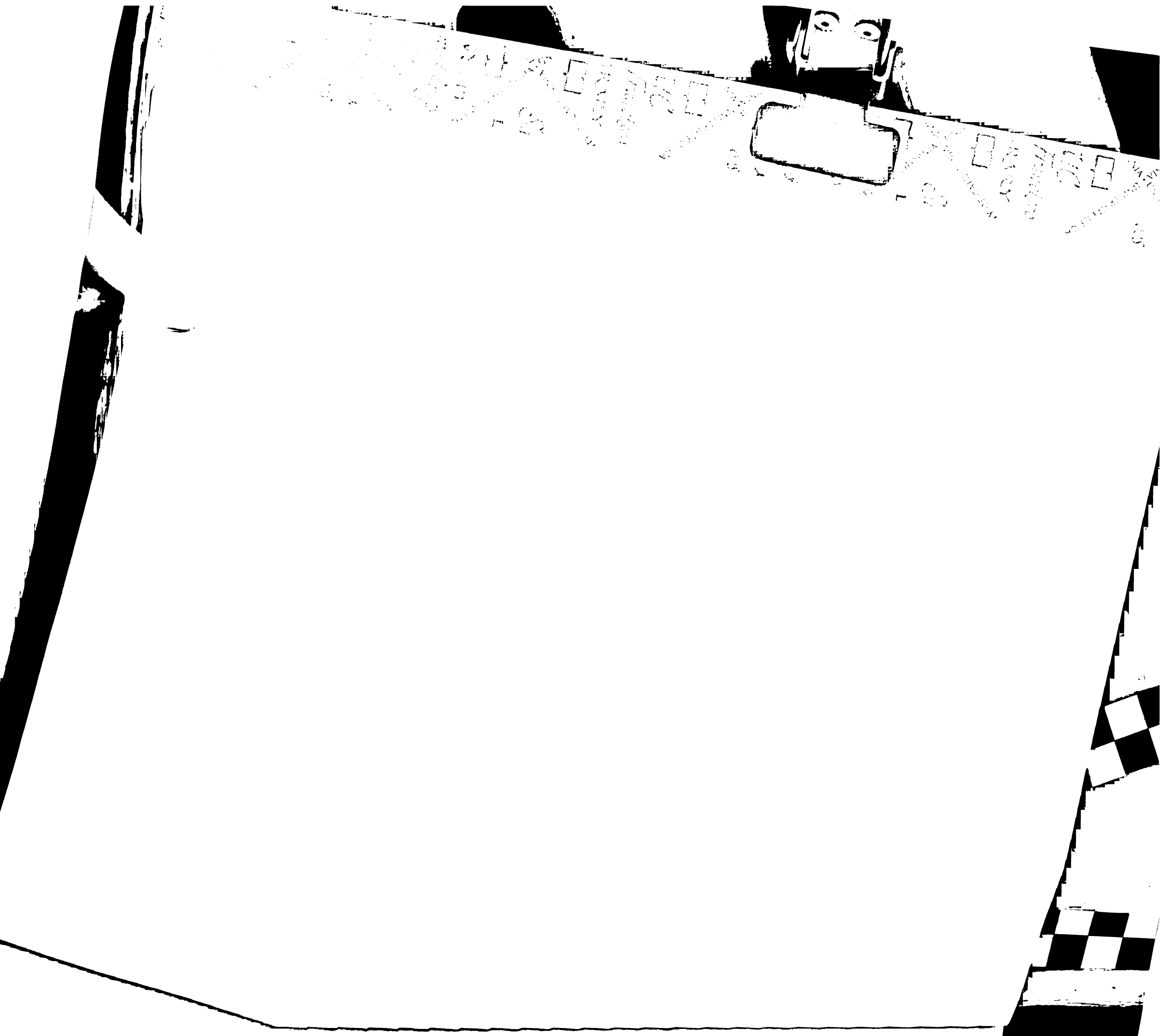


2001). In a survey of allelic imbalances in gene expression in heterozygous human cells, 18% of the more than 120 genes assayed displayed consistent biases in expression patterns toward one allele (Pastinen et al., 2004). It remains to be investigated what portion of such imbalanced expression was due to slow, reversible, stochastic fluctuations in gene expression and what portion was due to epigenetic factors or polymorphism in regulatory sequences.

Consequences of noise in gene expression

The consequences of noise in gene expression in a population of cells have been postulated to be beneficial, neutral or deleterious. Both the magnitude and the frequency of the noise impact the consequence. Small changes in protein abundance may have dramatic effects on fitness if they persist long enough, whereas large fluctuations in abundance may not have any effect if they occur too frequently to affect a cellular process (Rosenfeld et al., 2005; Samoilov et al., 2005). The observation that the time-scale for intrinsic noise fluctuations is much shorter than that for extrinsic noise suggests that extrinsic noise may impact cellular phenotypes more strongly than intrinsic noise, at least in *E. coli* (Rosenfeld et al., 2005).

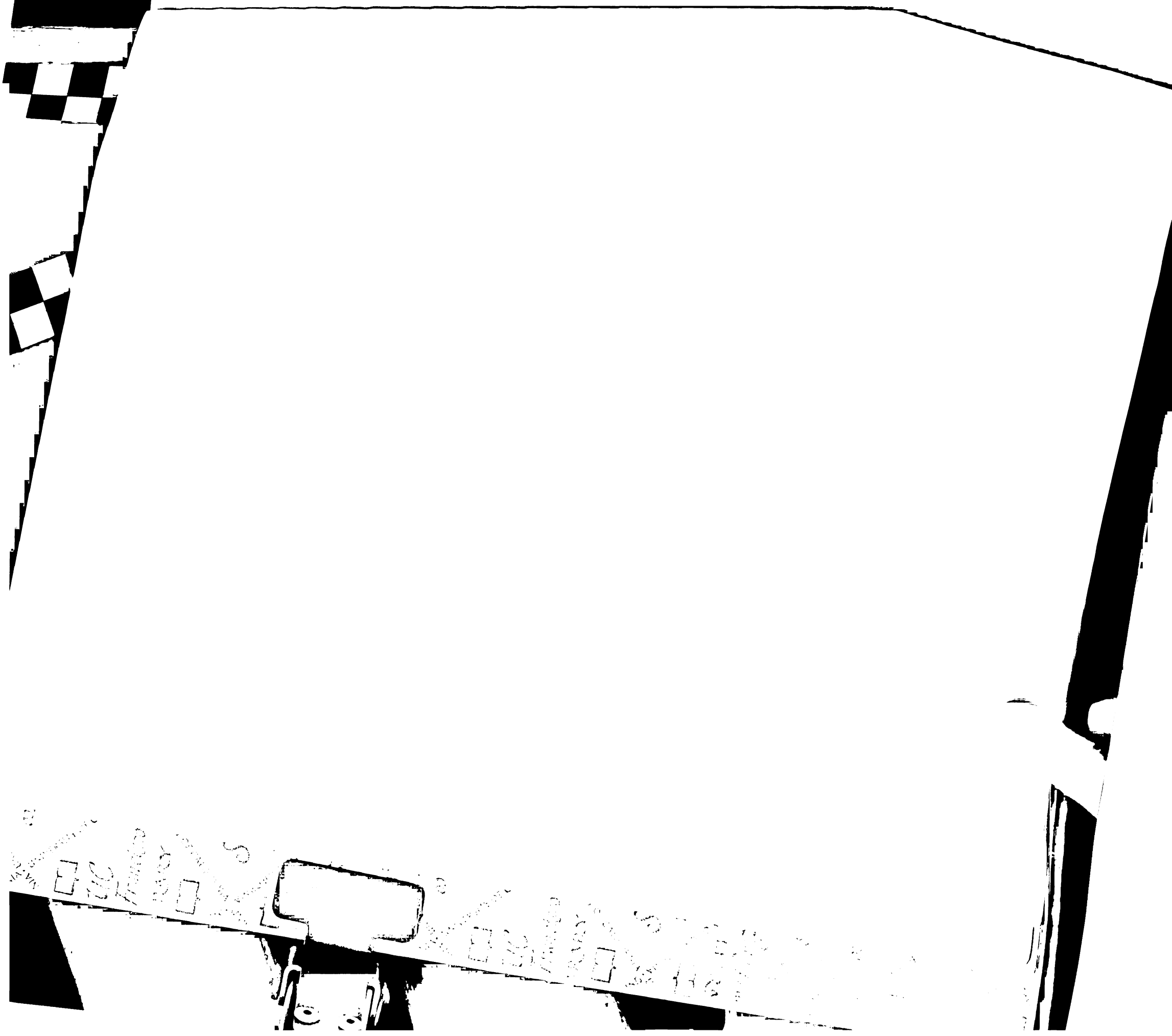
All types of noise by definition can create fluctuations in the abundance of a gene product. Small differences in protein abundance may confer a fitness advantage or disadvantage (Fig. 2.4A). For example, small increases in expression from a heat-shock-regulated promoter in a clonal population of *S. cerevisiae* correlate well with cell survival at extreme high temperatures (Attfield et al., 2001). Intrinsic noise can produce fluctuations in the relative expression of two alleles of the same gene in a heterozygote, potentially resulting in cells that express no allele, either individual allele, or both alleles.



If the two alleles are functionally divergent, the population of cells could acquire heterogeneity (Fig. 2.4B). Such fluctuations may contribute to the still-debated phenomenon of hybrid vigor (Hansson and Westerberg, 2002). Alternatively, intrinsic noise in the case of haploinsufficiency may result in increased levels of noise, or complete loss of function in a subset of cells. Such a mechanism has been proposed in the case of the human tumor suppressor gene *NF1* (Kemkemer et al., 2002) and prostate neoplasia formation in the mouse (Magee et al., 2003).

A brief period of intrinsic noise followed by feedback may allow for stochastic choice in stable monoallelic expression. This model has been proposed to explain the process of odorant receptor choice in olfactory neurons (Chess et al., 1994; Serizawa et al., 2004). Each murine olfactory neuron expresses a single allele of one odorant receptor gene, out of a choice of ~1500 odorant receptor genes. A functional odorant receptor is required to prevent expression of other odorant receptors, suggesting that receptor choice occurs through stochastic activation of a single promoter followed by inhibitory signaling to the inactive odorant receptor promoters. The resulting heterogeneous population of olfactory neurons enables sensitive differentiation of odorant molecules – stochastic gene expression helps create a functional sense of smell.

Many reports have been made of differentiation of a population of unicellular organisms into two distinct states of gene expression, for example, for the lysis-lysogeny decision in lambda phage-infected *E. coli* (Arkin et al., 1998; Isaacs et al., 2003), the *GAL1* gene in *S. cerevisiae* (Biggar and Crabtree, 2001) and the lac operon in *E. coli* (Ozbudak et al., 2004). The stochastic factor that generates the two states of expression has not been experimentally confirmed in any such case, but noise in gene expression remains a plausible culprit (Arkin et al., 1998; Hasty et al., 2000; Kepler and Elston, 2001; Thattai



and van Oudenaarden, 2001). Noise in gene expression in the context of positive feedback may be sufficient to create switching between the two stable states, as may have occurred in an engineered positive feedback loop in S. *cerevisiae* (Becskei et al., 2001). The use of stochasticity to populate multiple steady-states may play an important role in differentiation in multicellular organisms (Fig. 2.4C), or in survival in fluctuating environments for unicellular organisms, e.g. (Meyers and Bull, 2002; Thattai and van Oudenaarden, 2004; Wolf et al., 2005).

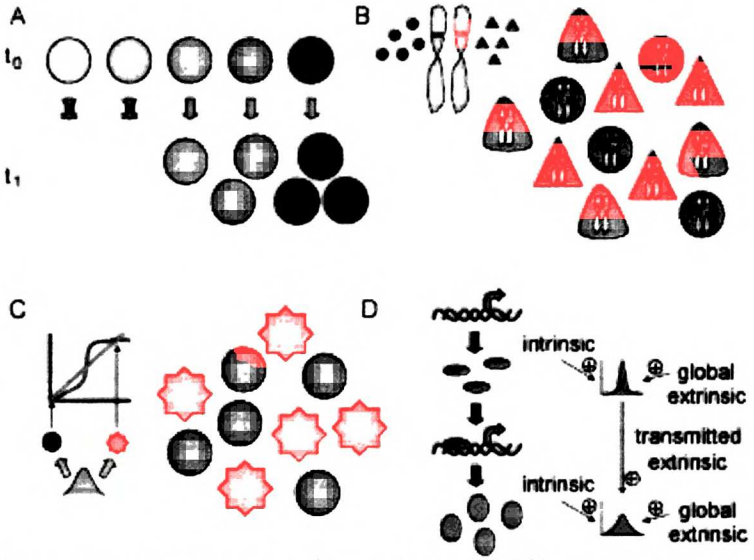


Figure 2.4. Consequences of noise. (A) Small differences in gene product abundance affect reproductive fitness. (B) In a heterozygous diploid population, cells display the phenotypes associated with each homozygote as well as the heterozygote. (C) Noise allows simultaneous achievement of multiple steady-state phenotypes in a population. (D) Noise can be transmitted from one gene, in this case a transcription factor, to a downstream target. The intrinsic and global extrinsic noise of the transcription factor can cause extrinsic noise in the downstream gene.

Genes are organized into regulatory circuits where the expression of one gene can influence the expression of another. A consequence of this organization is that noise in the expression of one gene may propagate to affect noise in the expression of a downstream gene (Fig. 2.4D). Recent work in *E. coli* has demonstrated that a synthetic cascade of three transcription factors produces more noise in output than a linear cascade of two transcription factors, or than one transcription factor alone (Hooshangi et al., 2005). Such transmission of noise has been analyzed further in other recent work, also in *E. coli*, where the authors examined the various sources of noise in a synthetic

transcriptional cascade (Pedraza and van Oudenaarden, 2005). Intrinsic noise in the expression of a transcription factor causes extrinsic noise in a downstream target gene. Additionally, global noise affecting expression of the transcription factor propagates to the downstream target. Because the factor acts to repress transcription, global fluctuations in the repressor counteract the effects of global fluctuations in the expression of the downstream gene. These results predict that global fluctuations in a transcriptional activator will exacerbate the noise in the target gene.

Stochastic events in gene expression allow for several counterintuitive behaviors that are not displayed by equivalent deterministic systems. Cellular switches are theoretically susceptible to noise in gene expression; in a certain regime, a deterministic switch, or predicted area of bistability, can be converted to a homogeneous graded response by stochastic factors (Kepler and Elston, 2001; Thattai and van Oudenaarden, 2001). Stochasticity in nonlinear systems provides higher sensitivity and results in decreased variability in gene induction than the deterministic case (Paulsson et al., 2000). Additionally, stochasticity can destabilize an oscillator, or provide for regular oscillations in a regime in which a deterministic system would not oscillate (Vilar et al., 2002). These predictions await experimental validation.

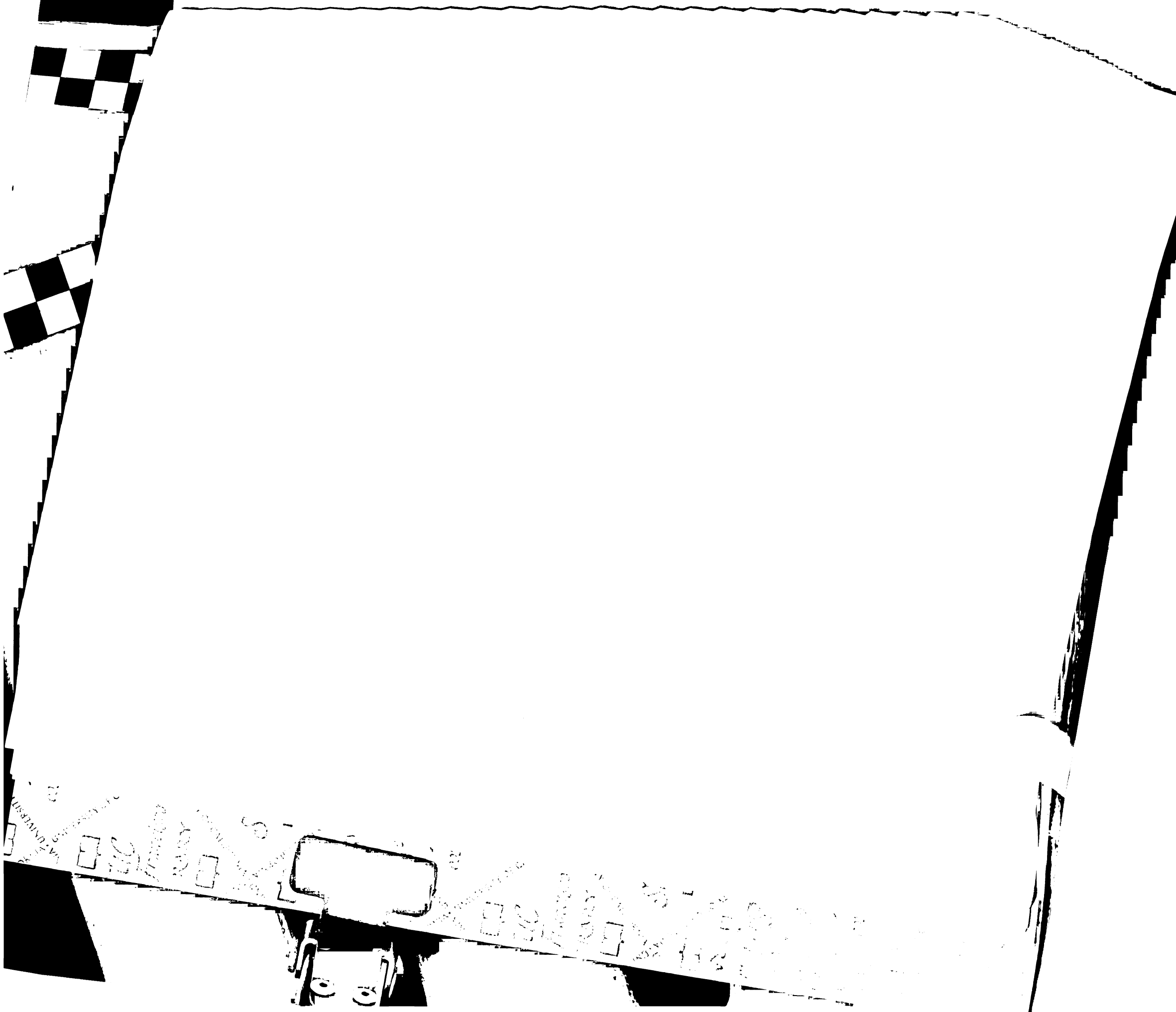
Control of noise in gene expression

Considering the consequences, it would be surprising if the control of noise in gene expression were not under evolutionary pressure. Several models suggest how such control could generate or suppress the intrinsic noise of gene expression. A theoretical model (Kepler and Elston, 2001; Kierzek et al., 2001; McAdams and Arkin, 1997; Thattai and van Oudenaarden, 2001), consistent with experimental evidence from

B. subtilis (Ozbudak et al., 2002), suggests that frequent transcription followed by inefficient translation results in lower intrinsic noise in protein levels than infrequent transcription followed by efficient translation (Fig. 2.5A). Similarly, when promoter fluctuations contribute to intrinsic noise as in *S. cerevisiae*, frequent promoter activation events followed by inefficient transcription will result in less noise in mRNA levels than infrequent promoter fluctuations followed by efficient transcription (Fig. 2.5B) (Blake et al., 2003; Ko, 1991; Raser and O'Shea, 2004). In these models, the control of noise comes at the energetic cost of producing few proteins from numerous mRNA, or the cost of repeated rounds of promoter remodeling resulting in a few mRNA. It has been noted that key regulatory proteins in *E. coli* display low translation rates, which could lower noise in protein levels (Ozbudak et al., 2002). Similarly, yeast genes that are essential or encode proteins involved in multisubunit complexes tend to have higher rates of transcription and lower rates of translation (Fraser et al., 2004).

The control of gene copy number represents a second way to lower the intrinsic noise in gene expression (Fig. 2.5C), which is predicted to scale with the inverse square root of the gene copy number. Noise control can therefore be added to the reasons cited for the widespread presence of polyploidy, especially in plants (Finnegan, 2002). Also, noise control by increased copy number provides an evolutionary rationalization for exact gene duplications and the maintenance of identical copies of the same gene. For example, two copies of many ribosomal protein genes have been maintained in *S. cerevisiae* following an ancient genome duplication event (Kellis et al., 2004).

Another possibility for control of noise lies in the presence of negative feedback linking the output of gene expression back to the control of gene expression. Becksei



and Serrano demonstrated the reduction of noise by means of negative feedback in the simplest such model in which a transcription factor negatively regulates its own synthesis (Fig. 2.5D) (Becskei and Serrano, 2000). More subtle forms of feedback may exist in which the rates of earlier gene expression steps are impacted by subsequent events. For example, the association of histone methylase activity with the elongating RNA polymerase complex during transcription results in the stable methylation of histones, which may affect the subsequent activation of the promoter (Ng et al., 2003). The consequences for noise of these types of feedback control of gene expression remain to be investigated experimentally; feedback violates the assumptions of the two-reporter methodology (Elowitz et al., 2002; Swain et al., 2002) and will require a modified analysis.

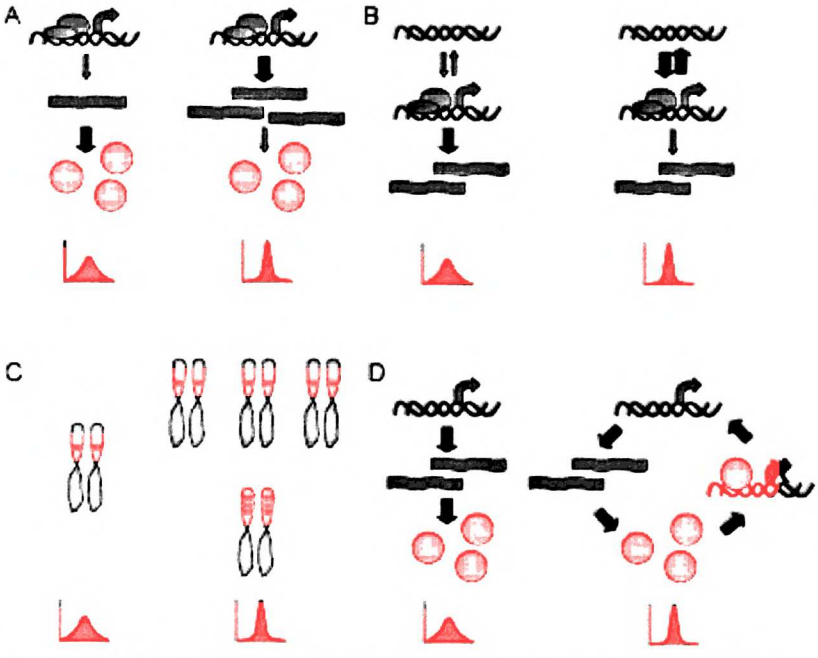
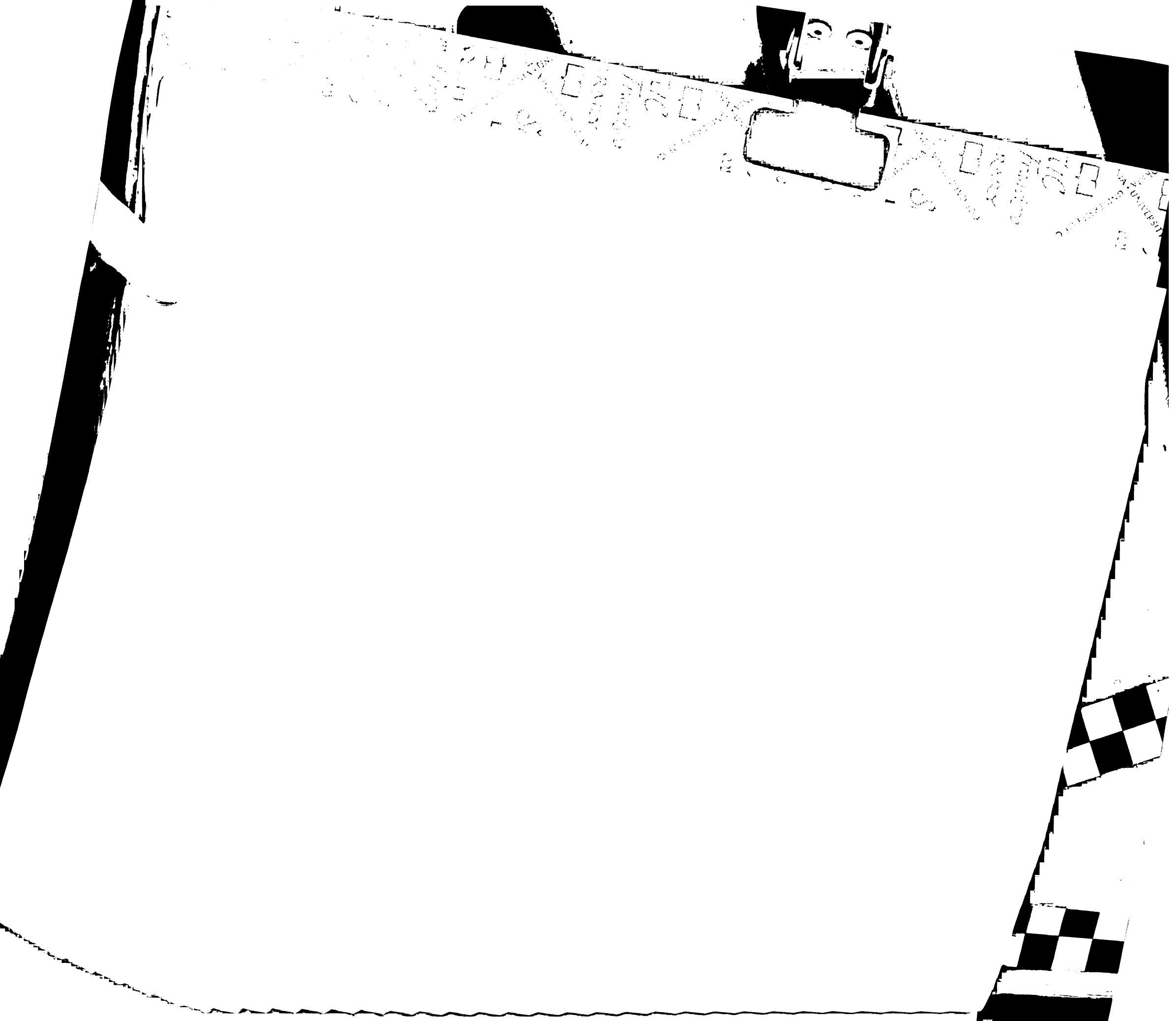


Figure 2.5. Control of noise. (A) Infrequent transcription followed by efficient translation results in high intrinsic noise in protein levels (left); frequent transcription and inefficient translation results in low intrinsic noise (right). (B) Infrequent promoter transitions between inactive and active states followed by efficient transcription result in high intrinsic noise in mRNA levels (left); frequent promoter transitions followed by inefficient transcription result in low intrinsic noise (right). (C) Increases in gene copy number, either through polyploidy (top right) or through gene duplication (bottom right), result in decreased intrinsic noise relative to a single gene copy (left). (D) Negative feedback, such as a transcription factor repressing its own transcription (right), results in decreased noise relative to a linear pathway (left).

Noise in the processes of gene expression may fundamentally limit the accuracy of certain cellular processes, such as the circadian oscillator. For example, transmission of noise may have resulted in the loss of coordination among cells in a synthetic



oscillator based on three transcription factors in *E. coli* (Elowitz and Leibler, 2000). Mihalcescu *et al.* demonstrated in the unicellular cyanobacterium *Synechococcus elongatus* that circadian oscillations persist with a stable oscillatory period for weeks without extracellular entraining cues or intercellular communication (Mihalcescu et al., 2004). Such persistence of regular oscillations suggests that the circadian network is strongly resistant to biochemical noise. Recent work has asserted that this oscillator does not require transcription or translation, but instead relies on posttranslational molecular events (Tomita et al., 2005). It is interesting to speculate that a core posttranslational oscillator, which can rely on large numbers of molecules and avoid the small-number stochasticity of gene expression, may be required for robust oscillations.

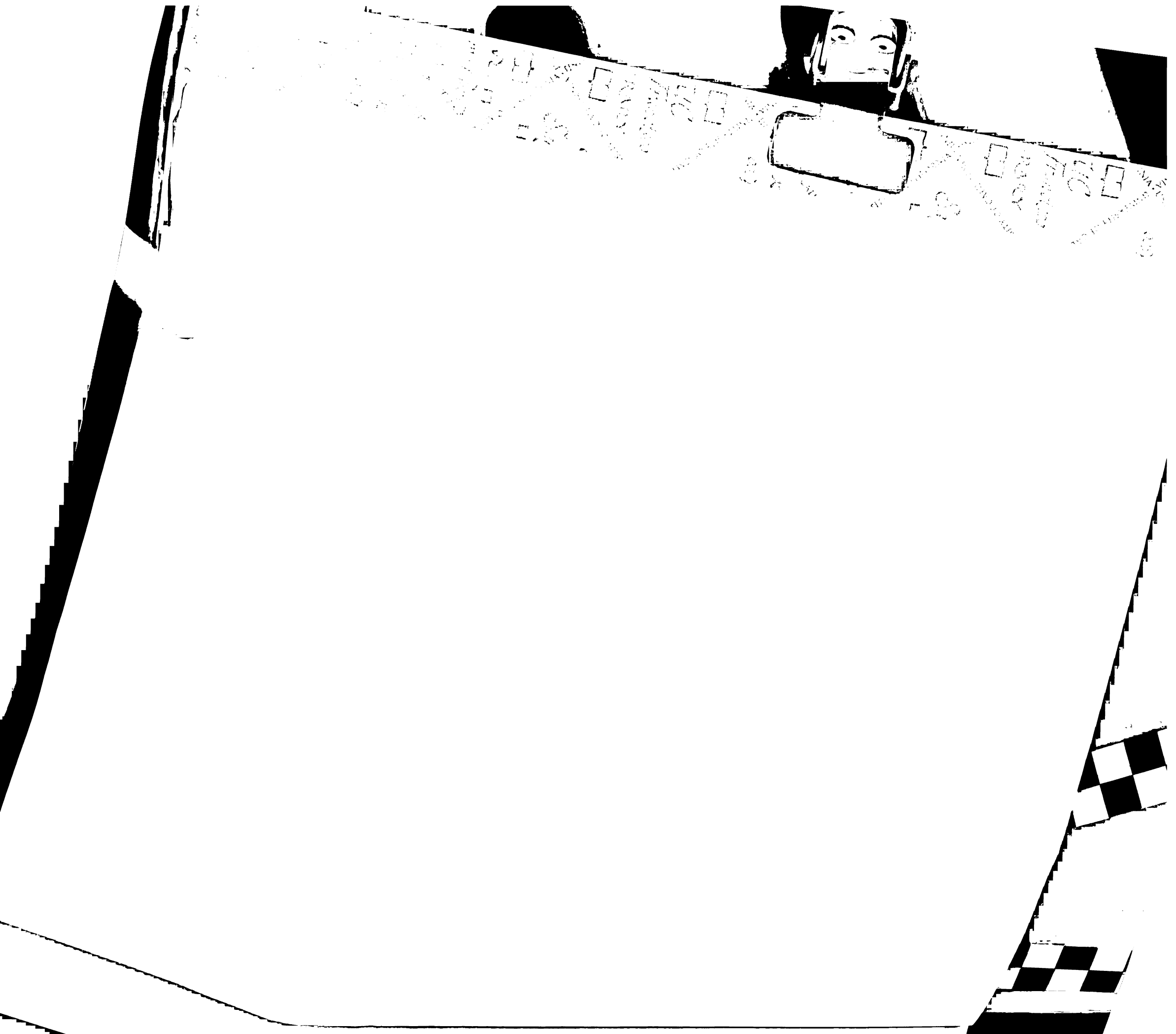
It is possible that cellular control mechanisms exist which enable the switch between globally noisy or globally "quiet" states of gene expression. Elowitz *et al.* showed that different genetic backgrounds produce different levels of intrinsic noise in *E. coli* and identified the *recA* gene as a genetic modifier of intrinsic noise, though this is probably due to heterogeneity in gene copy number in *recA-* cells, rather than the stochastic nature of gene expression (Elowitz et al., 2002). In budding yeast cells, mutation of a component of the SWI/SNF chromatin-remodeling complex not only increases both extrinsic and intrinsic noise at the *PHO5* gene (Raser and O'Shea, 2004) but also results in variability in cellular morphology (Raser and O'Shea, unpublished data). As the SWI/SNF complex facilitates expression of a large percentage of the yeast genome, changes in SWI/SNF activity, either genetic or signal-dependent, could affect the noise for a large subset of the yeast proteome. Queitsch *et al.* have demonstrated that reduction of heat shock protein 90 (Hsp90) chaperone activity in *Arabidopsis thaliana*

increases diversity in morphology in inbred lines, in addition to revealing otherwise silent genetic variation among different lines (Queitsch et al., 2002). Hsp90 chaperone activity is therefore hypothesized to reduce the effect of stochastic molecular events that might otherwise result in developmental variability.

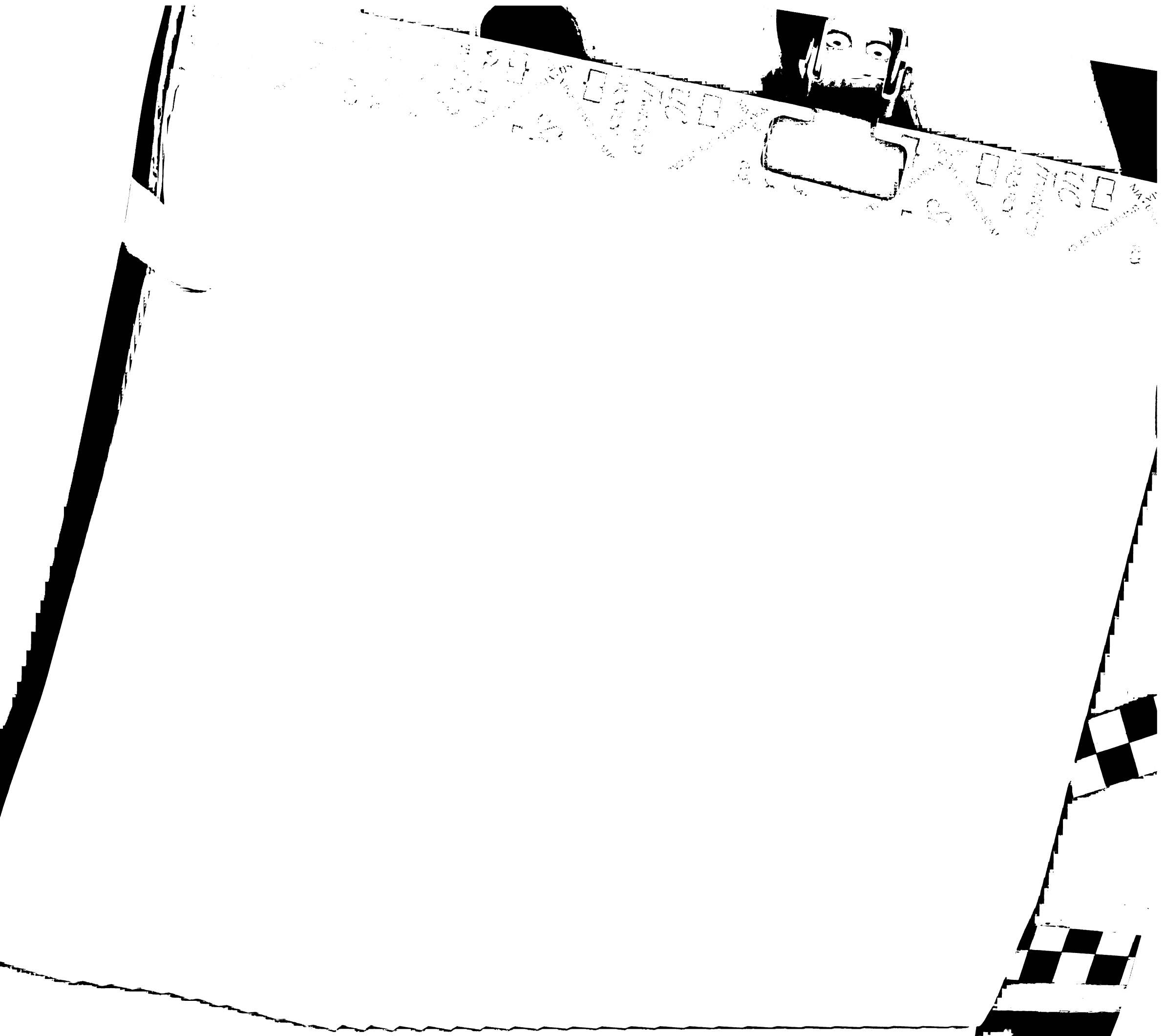
There may exist a large number of buffering agents that reduce either the magnitude of noise in gene expression or the impact of such noise on cellular or organismal phenotype. These buffering agents may be regulated, especially in times of stress, in order to produce a phenotypically diverse population. Waddington's theories of canalization and genetic assimilation propose that wasteful phenotypic variability in a population is suppressed when the population is well-adapted to its environment (Waddington, 1959). However, if environmental conditions shift, phenotypic noise becomes advantageous because a noisy population will produce some members that are better adapted to the new environment. The advantageous phenotype conferred by noise on a subset of the population will be rapidly fixed genetically. The new strain is well-adapted to the new environmental condition, and phenotypic variability will again be suppressed. Much recent work supports the idea that it is advantageous to increase variability in times of stress and decrease variability when organisms are well-adapted to the environment (Pigliucci and Murren, 2003). Regulation of global noise factors could provide a molecular basis for such evolutionary flexibility.

Concluding remarks

Many questions remain concerning the generation of noise in gene expression and its consequences for cellular behavior. The presence of stochasticity in gene expression has been confirmed to result in noise in protein abundance, but other sources

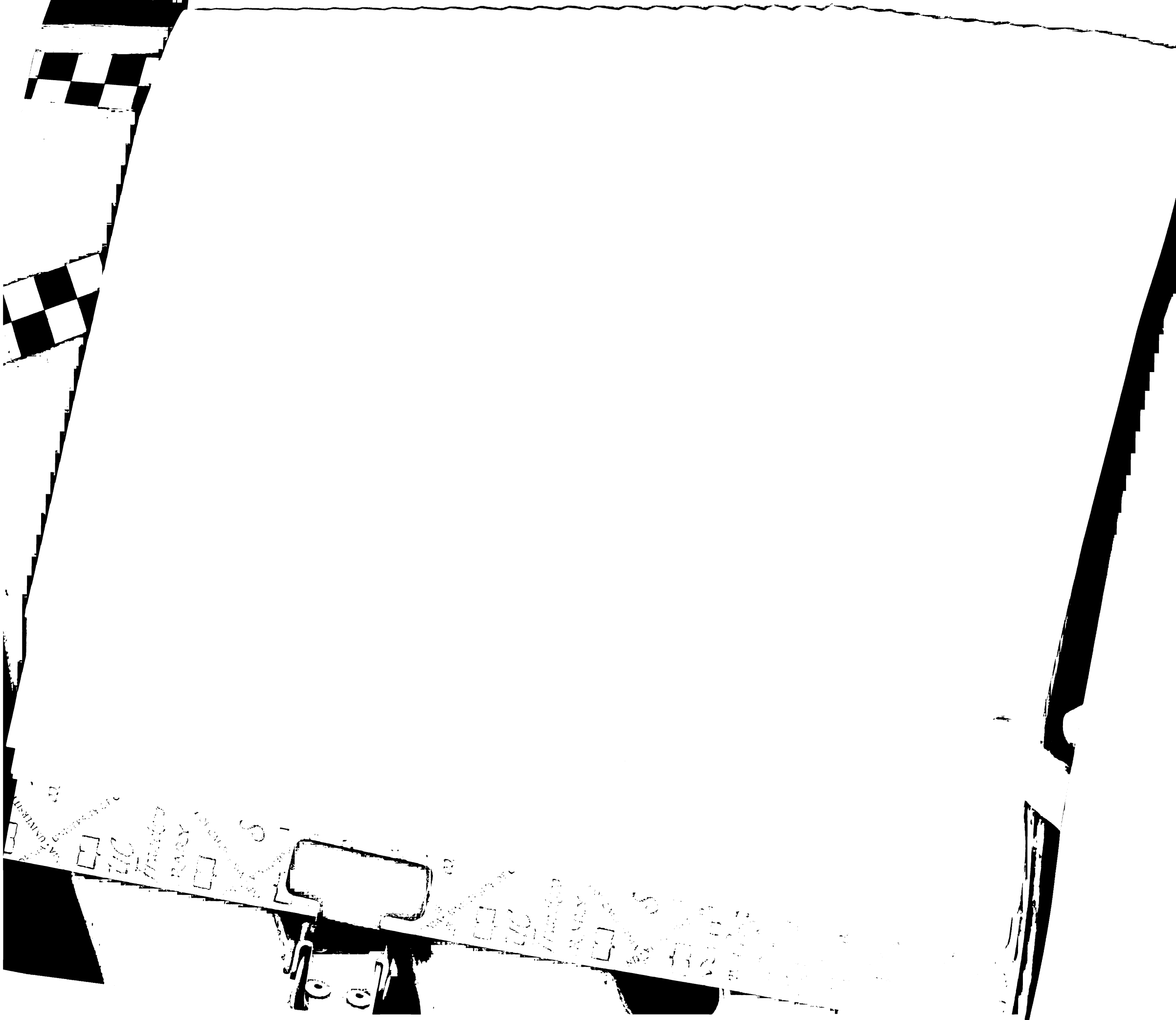


of noise may result in phenotypic variability. In many cases, a particular source of noise has been hypothesized but not confirmed. For example, there exists more variation in fingerprint patterns among identical quadruplets than between identical twins, suggesting that local environmental factors in the uterus may be the source of variation rather than any stochastic developmental process (Steinman, 2001). Beyond the identification of true examples of phenotypic consequence, much work must be done to understand how cellular processes behave robustly in the presence of underlying stochasticity. Such work often requires a nontraditional collaboration between mathematicians, physicists, and *in vivo* experimentalists. Fortunately, many biologists are beginning to focus on the limitations and benefits that stochasticity creates for biological systems, and we expect that future investigations will reveal results both unexpected and, well, unpredictable.



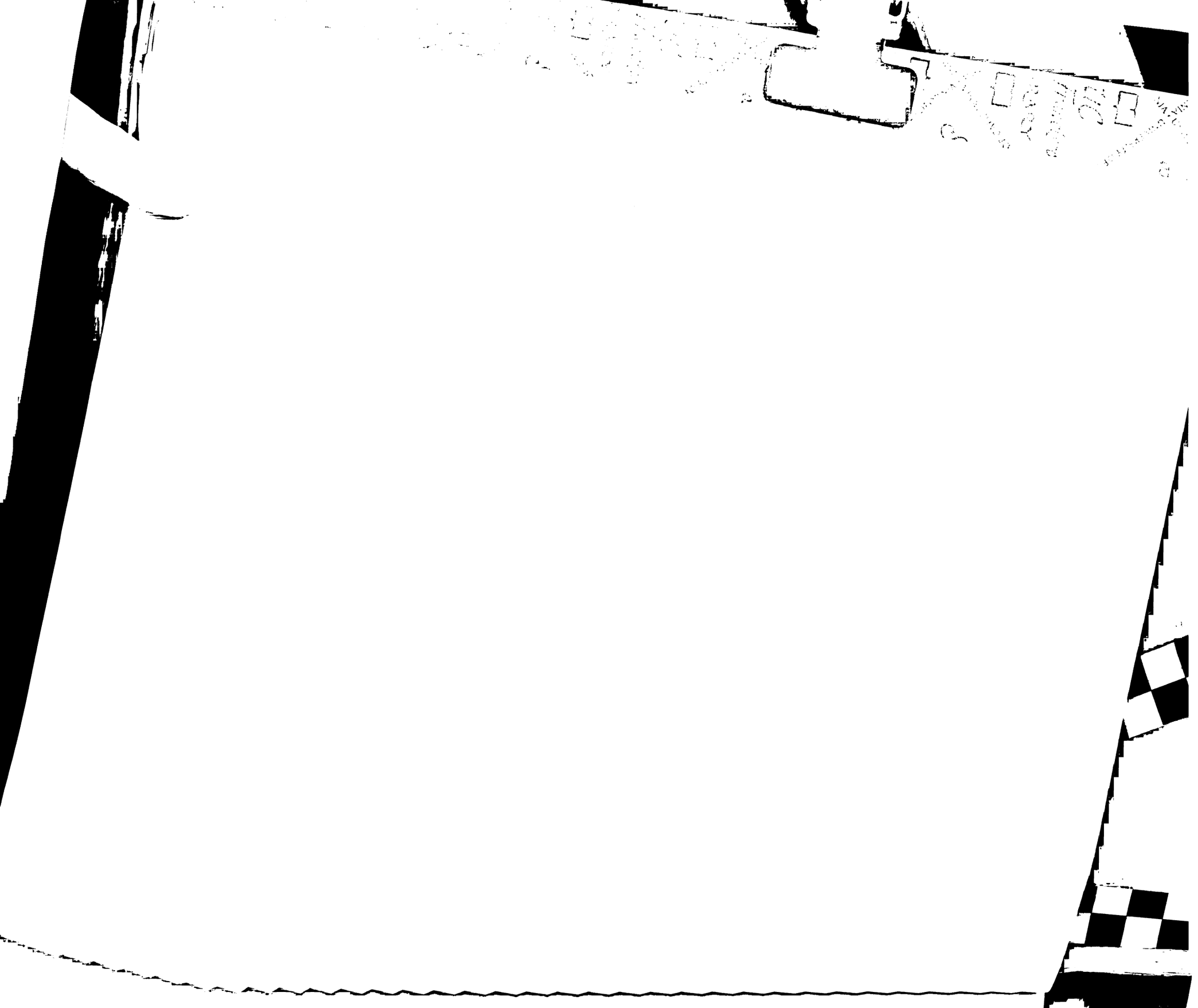
PREFACE TO CHAPTERS THREE AND FOUR

One useful goal of the experimental characterization of signal transduction pathway is to achieve a comprehensive understanding of not only the qualitative, but also the quantitative aspects of signal transduction at every step. Many signal transduction pathways result in feedback that is predicted to act both positively and negatively on upstream signaling but is of unknown significance in vivo. Most genetic and biochemical approaches in the past have revealed the roles of the components essential for signal transduction, but these approaches usually only result in a qualitative understanding of the activities of these components. In chapters three and four, we emphasize a quantitative approach to understanding nutrient-responsive signal transduction pathways in the budding yeast *S. cerevisiae*.



CHAPTER THREE:

Quantitative analysis of zinc-responsive signal transduction



Introduction

Zinc in its divalent cation form is an essential nutrient for all living cells. Within the cell, zinc is required for a number of enzymatic activities, as well as the zinc finger protein structural motif. Zinc(II) is not active in oxidation-reduction reactions in the conditions found intracellularly, but is important for protein structure and catalytic activity. In addition to being essential for growth and cell division, zinc is also toxic when present in excess. Zinc toxicity is thought to be due to competitive inhibition of the binding of other essential metal cations, especially copper and iron, which are required for cellular oxidation-reduction reactions.

Budding yeast cells thrive in environments that contain free divalent zinc concentrations over a range of nine orders of magnitude (MacDiarmid et al., 2000; Zhao and Eide, 1997). A zinc-specific signal transduction pathway is indispensable for this remarkable adaptability. In the absence of components of the pathway, cells can survive over a range of concentrations of six orders of magnitude but can grow at near wild-type rates over only two orders of magnitude of zinc concentration.

Genes important for cellular homeostasis that have been identified in budding yeast cells include multiple different activities (Fig. 3.1). Within the cell, zinc is thought to exist in multiple pools: a nucleocytoplasmic pool, which can be divided into both protein-bound and labile forms, and a vacuolar pool. There exist multiple proteins that transport zinc specifically across the zinc-impermeable lipid bilayers, including plasma membrane zinc transporters and the vacuolar membrane zinc transporters that control the flux between vacuolar and cytoplasmic compartments. At the plasma membrane, there exist at least two low-affinity zinc transporters, Zrt2 (Zhao and Eide, 1996a) and Fet4 (Waters and Eide, 2002), and one high-affinity transporter, Zrt1 (Zhao and Eide,

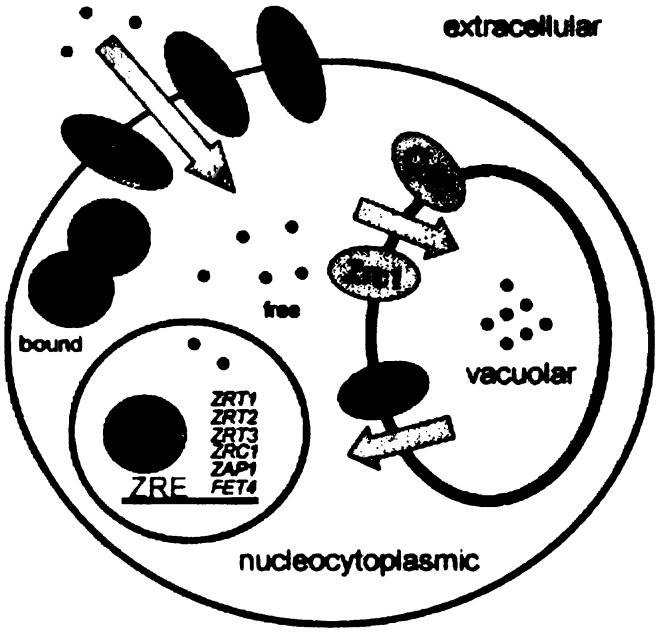
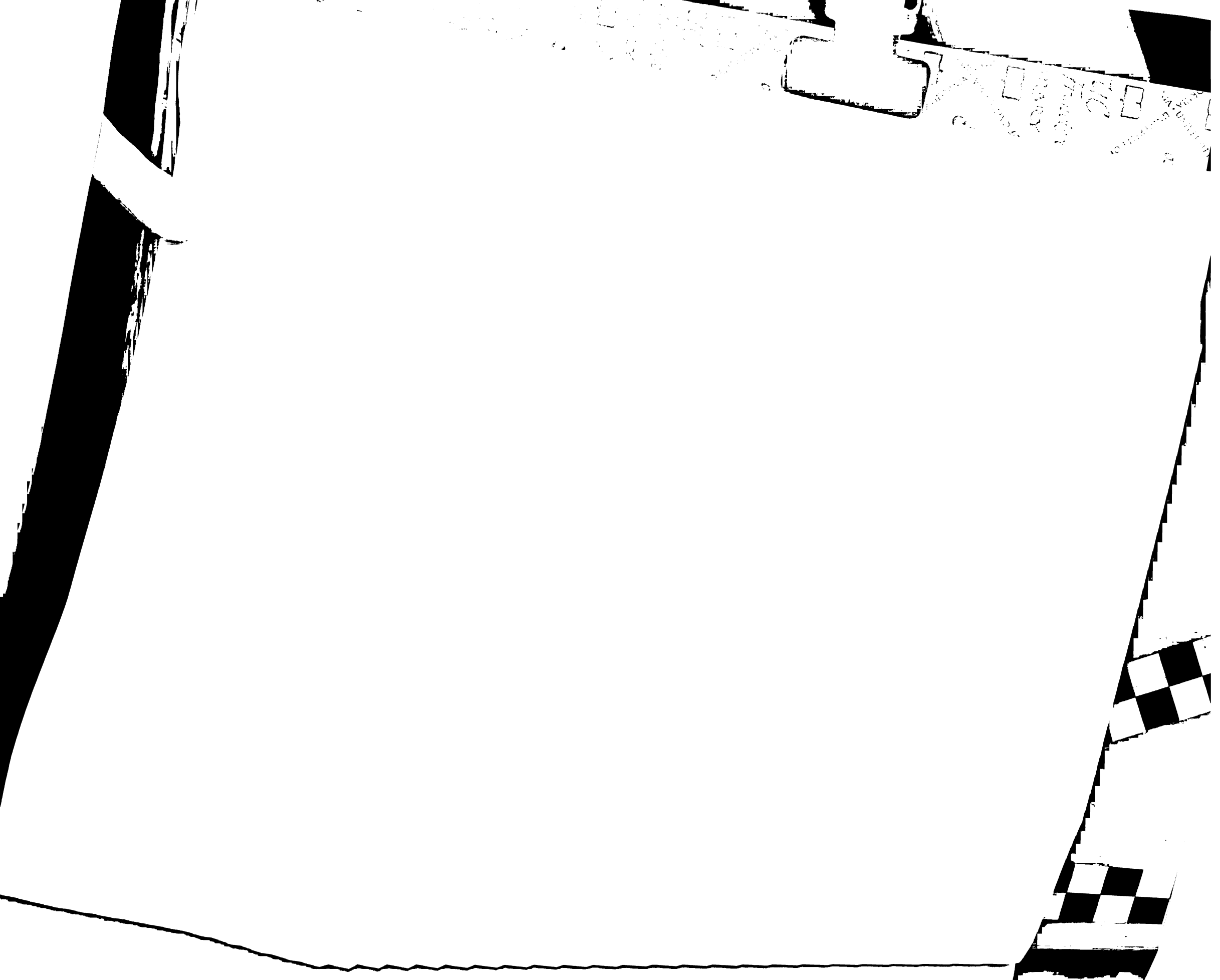


Figure 3.1. Schematic of zinc signaling. This diagrams details the multiple cellular zinc pools, the transporters that are responsible for zinc flux into the cell and between subcellular compartments, and the zinc-dependent signal transduction pathway consisting of the nuclear zinc-inhibited transcription factor (Zap1) and zinc-responsive elements (ZREs) in the promoters of zinc-regulated genes.

1996b). Vacuolar flux is dependent on two influx transporters, Zrc1 and Cot1 (MacDiarmid et al., 2000; MacDiarmid et al., 2002), and a single efflux transporter, Zrt3 (MacDiarmid et al., 2000). A zinc-specific signal transduction pathway that regulates the transcription of a number of genes consists solely of a transcription factor, Zap1 (Zhao and Eide, 1997), whose activity is directly modulated by binding to intracellular Zn^{++} (Bird et al., 2003; Bird et al., 2000). The products of zinc-responsive genes include proteins involved in zinc transport, zinc usage, and zinc signaling(Lyons et al., 2000; Yuan, 2000), including Zrt1, Zrt2, Fet4, Zrt3, Zrc1, and Zap1 itself. In addition, a second, Zap1-independent signal transduction pathway regulates the rate of endocytosis of the high-affinity plasma membrane zinc transporter (Gitan and Eide, 2000; Gitan et al., 1998; Gitan et al., 2003).

The zinc-dependent signal transduction pathway represents a model system for the study of nutrient homeostasis. As zinc is both required and toxic in excess, it is reasonable to hypothesize that zinc-dependent signaling minimizes fluctuations in intracellular free zinc levels. This hypothesis is further supported by the observation that deletion of particular pathway components results in a restriction of viability from a wide to a limited range of extracellular zinc concentrations. Previous work has

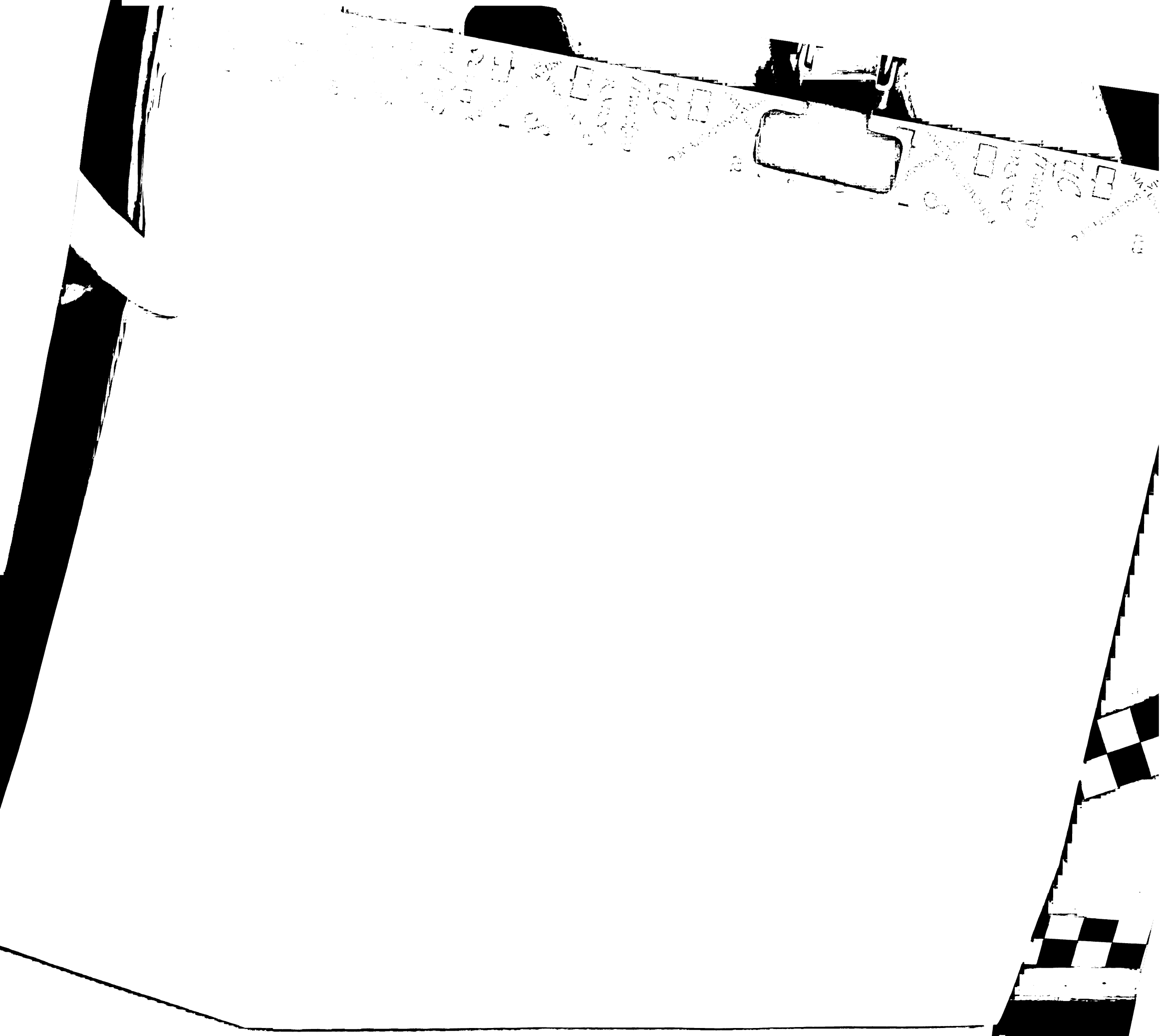


demonstrated that the average zinc response is quantitatively different in a population of cells over various concentrations of extracellular zinc (Zhao and Eide, 1996a). We set out to obtain a quantitative characterization of zinc-dependent signaling. We can compare these observations with the qualitative predictions from an integrated model of intracellular zinc flux and signaling. Eventually, it will be possible to construct a quantitative integrated model that confirms our understanding of the zinc signal transduction pathway.

Results

Construction of a qualitative integrated model of zinc sensing and signal transduction

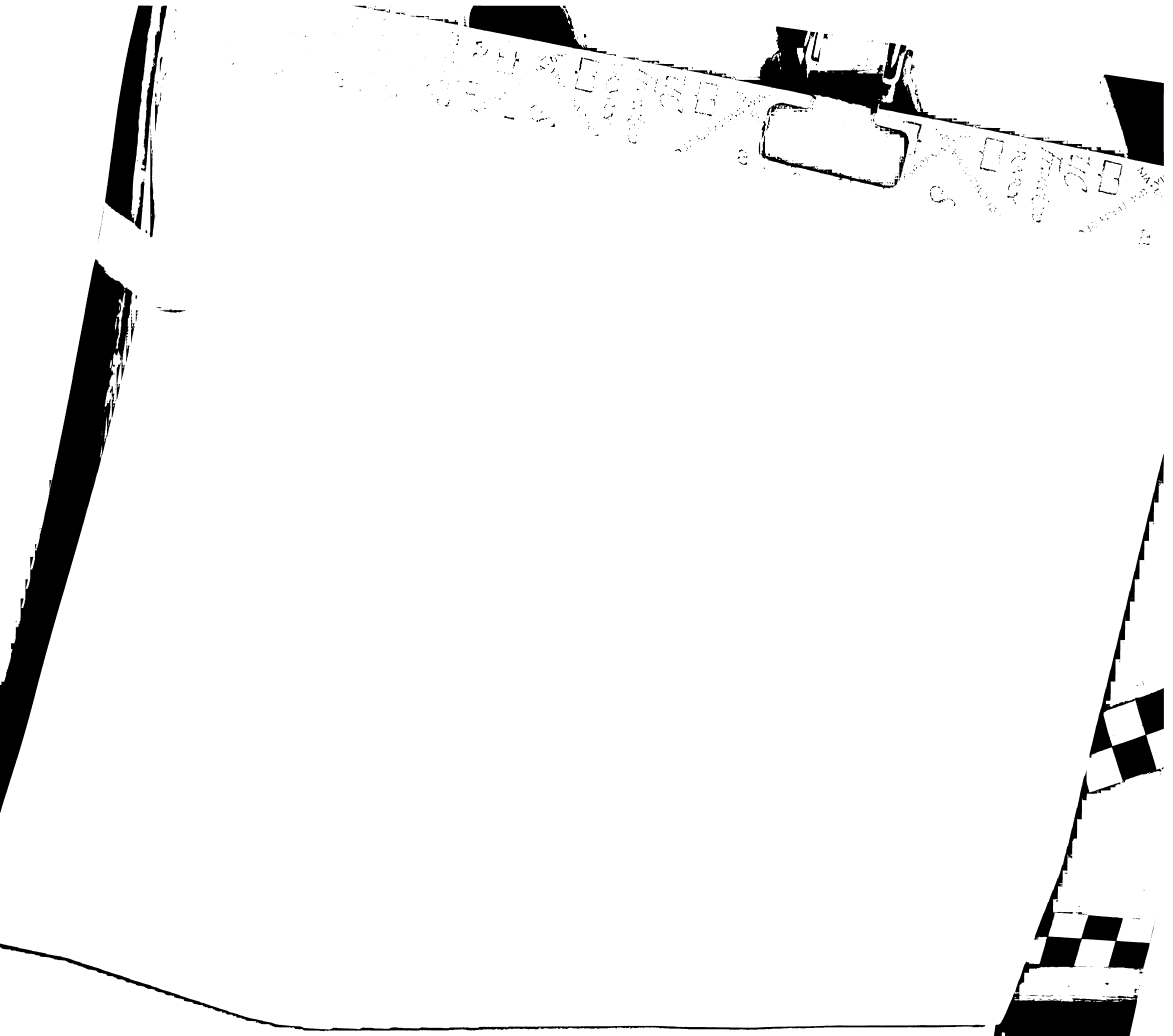
We wished to construct an integrated model of zinc signaling to be able to understand how each individual component contributes to the behavior of the zinc signal transduction pathway (Fig. 3.2A). We base our model on the hypothesis that the transcription factor Zap1 binds directly to and is inhibited by free intracellular zinc ions. Zap1, in a zinc-dependent manner, controls the transcription of the zinc-responsive genes and therefore the abundance of the gene products Zrt1, Zrt2, Fet4, Zrc1, Zrt3, and Zap1 itself. Intracellular zinc is produced by transport from the extracellular environment through the zinc-specific plasma membrane transporters Zrt1, Zrt2, and Fet4. Intracellular zinc levels are diminished due to usage during cellular growth and through transport into the vacuolar by Cot1 and Zrc1. When zinc is present in the vacuole, it can contribute to free intracellular levels by transport through the Zrt3 vacuolar membrane transport. Because we lack detailed, quantitative information about nearly all of the relationships displayed in the model, we cannot yet construct an accurate quantitative model of zinc signaling. We can, however, make hypotheses about



the qualitative effects of perturbations in individual components and test these predictions experimentally. Our qualitative model includes multiple feedback loops, both positive and negative in nature (Fig. 3.2B), and we expect that these feedback loops will contribute to the behavior of the zinc signal transduction pathway.

Measurement of Zap1-dependent gene expression in single cells

To quantify the amount of Zap1-dependent gene expression in a single cell, we placed the yeast codon-optimized gene encoding the Venus yellow fluorescent protein (yVYFP) under the control of the promoter of the zinc-dependent gene *ZRT1* and integrated the promoter-fluorescent gene fusion (*ZRT1prYFP*) at the chromosomal *leu2* locus in the haploid budding yeast. Population measurements of *ZRT1* mRNA levels have demonstrated previously that the average level of *ZRT1* expression changes over a range of zinc concentrations in medium containing the divalent metal chelator EDTA (Zhao and Eide, 1996a). We confirmed by flow cytometric measurements of yVYFP fluorescence that our strain expressed increasing levels of yVYFP as the zinc concentration in the medium decreased (Fig. 3.3A). We confirmed that the expression observed was dependent on both the presence of the sequence coding for the *yVYFP* gene and the *ZAP1* gene (Fig. 3.3A and data not shown). In addition, we were able to differentiate between two models that equivalently account for the previous observations regarding *ZRT1* expression. One model suggests that the *ZRT1* promoter exists in two states, completely off or fully active. In this model, the change in average expression over the range of zinc concentration is due to changes in the proportion of cells in the active state. The other model, which is supported by our measurements of



ZRT1prYFP in single yeast cells, suggests that the *ZRT1* promoter is capable of a wide range of rates of expression (Fig. 3.3A).

Measurement of Zap1-dependent gene expression in perturbed strains

The qualitative model shown in figure 3.2A makes testable predictions about the effect of deletion of various genes involved in zinc transport. In particular, the deletion of the each of the genes encoding the zinc-specific plasma membrane transporters Zrt1, Zrt2, or Fet4 should reduce the levels of intracellular zinc and therefore increase the Zap1-dependent transcription. This increase in Zap1-dependent transcription in the perturbed strains should occur only at extracellular zinc concentrations at which the transporter in question contributes to zinc

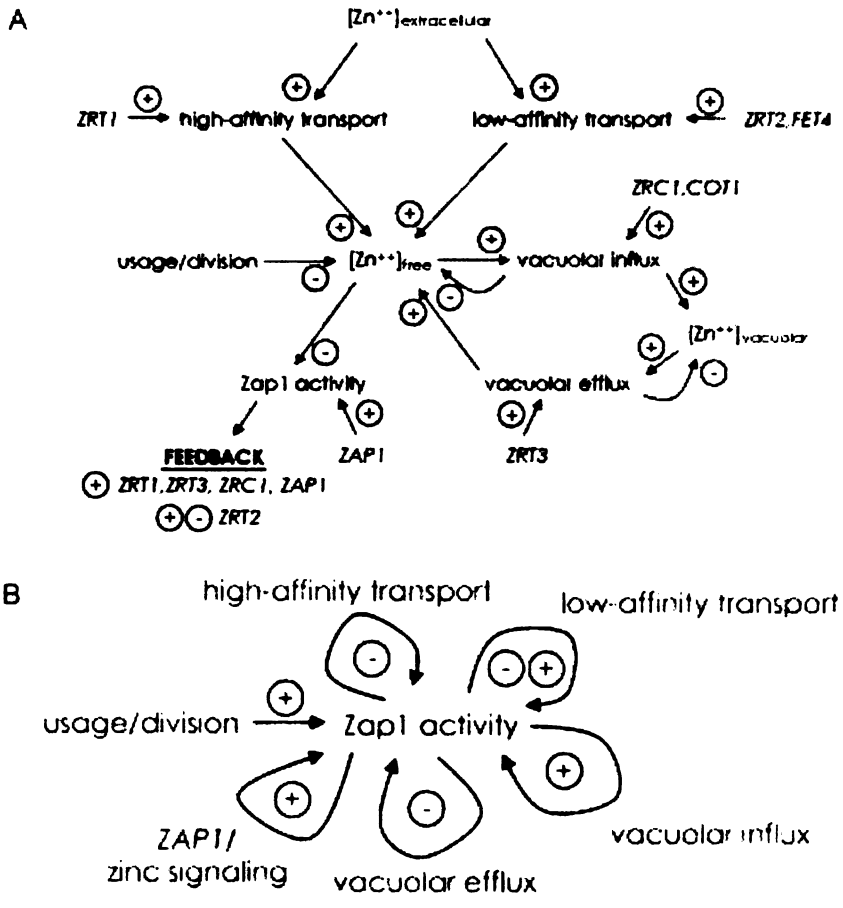
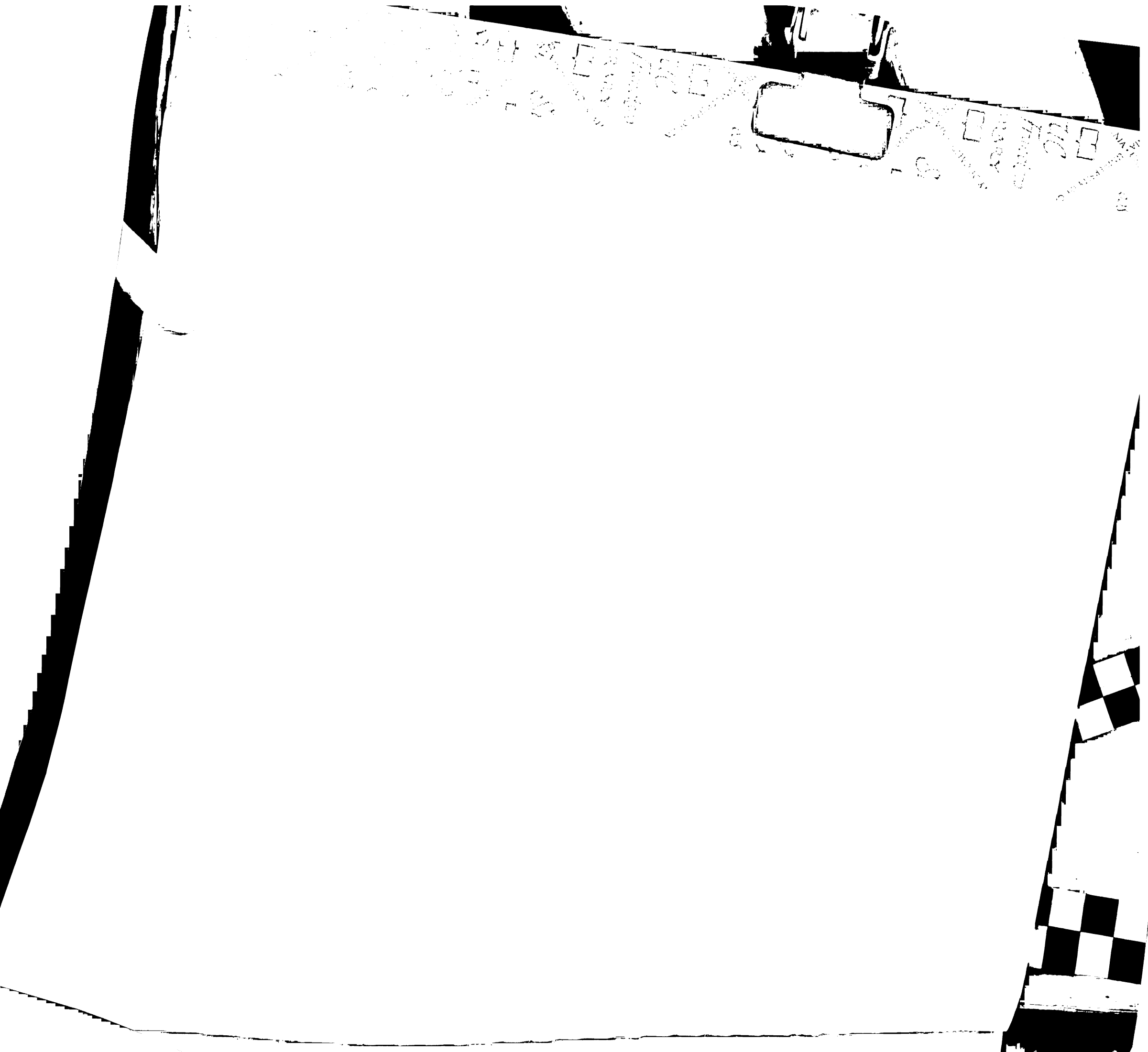


Figure 3.2. Qualitative model of zinc signaling. **(A)** Gene products and processes that affect intracellular zinc levels and Zap1 activity are diagrammed. The directionality of the relationships is indicated by + for an positive effect and - for a negative effect. **(B)** Feedback loops in the zinc-dependent signal transduction pathway and their predicted effects on Zap1 activity. Recall that Zap1 activity increases with decreasing intracellular zinc levels. The low-affinity transporter is both upregulated and downregulated by Zap1 activity (Bird et al., 2004).



uptake. As predicted, the levels of Zap1-dependent yVYFP expression are increased in the *zrt1Δ* strain and the *zrt2Δ* strain relative to the wild-type strain (Fig. 3.3B-C). Significantly, the deletion of the *ZRT1* gene affects Zap1-dependent gene expression at lower levels of extracellular zinc, while the deletion of the *ZRT2* gene affects gene expression at higher levels of gene expression.

Furthermore, the simultaneous deletion of both low-affinity transporters, *ZRT2* and *FET4*, results in an increase in Zap1-dependent gene expression no more substantial than that observed in the *zrt2Δ* strain (Fig.3.3F) . Consistent with this observation, the *fet4Δ* strain shows no increase in gene expression relative to wild-type levels of expression (data not shown). We conclude that *FET4* does not play a substantial role in zinc uptake under these conditions. The strain lacking both *ZRT1* and *ZRT2* showed a striking increase in Zap1-dependent gene expression (Fig. 3.3H). In fact, this strain was unable to continue growing at lower levels of extracellular zinc.

The integrated model predicts that the deletion of the *ZRT3* gene, whose gene product is responsible for the release of zinc from the vacuole into the cytoplasm, will result in an increase in the level of Zap1-dependent gene expression. Measurements of *ZRT1* promoter-dependent yVYFP expression in the *zrt3Δ* strain demonstrate that this prediction is grossly true at all concentrations of extracellular zinc tested (Fig. 3.3D). Unexpectedly, the *zrt3Δ* strain shows a bimodal pattern of expression across all extracellular zinc concentrations employed. Each of the two populations induces a higher level of yVYFP than the single population found in wild-type cells. The total population of *zrt3Δ* cells is split approximately equally between the two discrete subpopulation across all extracellular concentrations of zinc. This bimodality is not

necessarily predicted by the qualitative model, but can be explained. There are several positive feedback loops, especially in the Zap1-dependent induction of *ZAP1* or *ZRC1*, that are sufficient to provoke multistability and explain the presence of two different populations of cells. Alternatively, the bimodality may reflect some reversible extrinsic heterogeneous factor, such as cell cycle or the change in gene copy number that accompanies cell cycle progression. The fact that the bimodality is displayed at all concentrations of zinc, and the relative ratio of the populations remains fairly constant across all concentrations, is more consistent with an extrinsic heterogeneity.

The qualitative model predicts that the simultaneous deletion of the *ZRC1* and *COT1* genes, whose gene products are responsible for the movement of zinc from the cytoplasm into the vacuole, will result in an increase in cytoplasmic zinc levels and therefore a decrease in Zap1-dependent gene expression. This decrease in gene expression should only be apparent at steady-state concentrations of zinc at which the vacuole is zinc-replete. Previous measurements of intracellular zinc levels suggest that these vacuolar stores of zinc are present only at higher levels of extracellular zinc (MacDiarmid et al., 2000). As predicted, the *zrc1Δ cot1Δ* strain displays less Zap1-dependent gene expression than wild-type cells across intermediate but not low zinc concentrations (Fig. 3.3G). At high concentrations of extracellular zinc, the cells did not grow and therefore the increase in gene expression seen for the highest two zinc concentrations is likely irrelevant. The absence of growth in high concentrations of zinc in the *zrc1Δ cot1Δ* strain is consistent with the previously suggested importance of

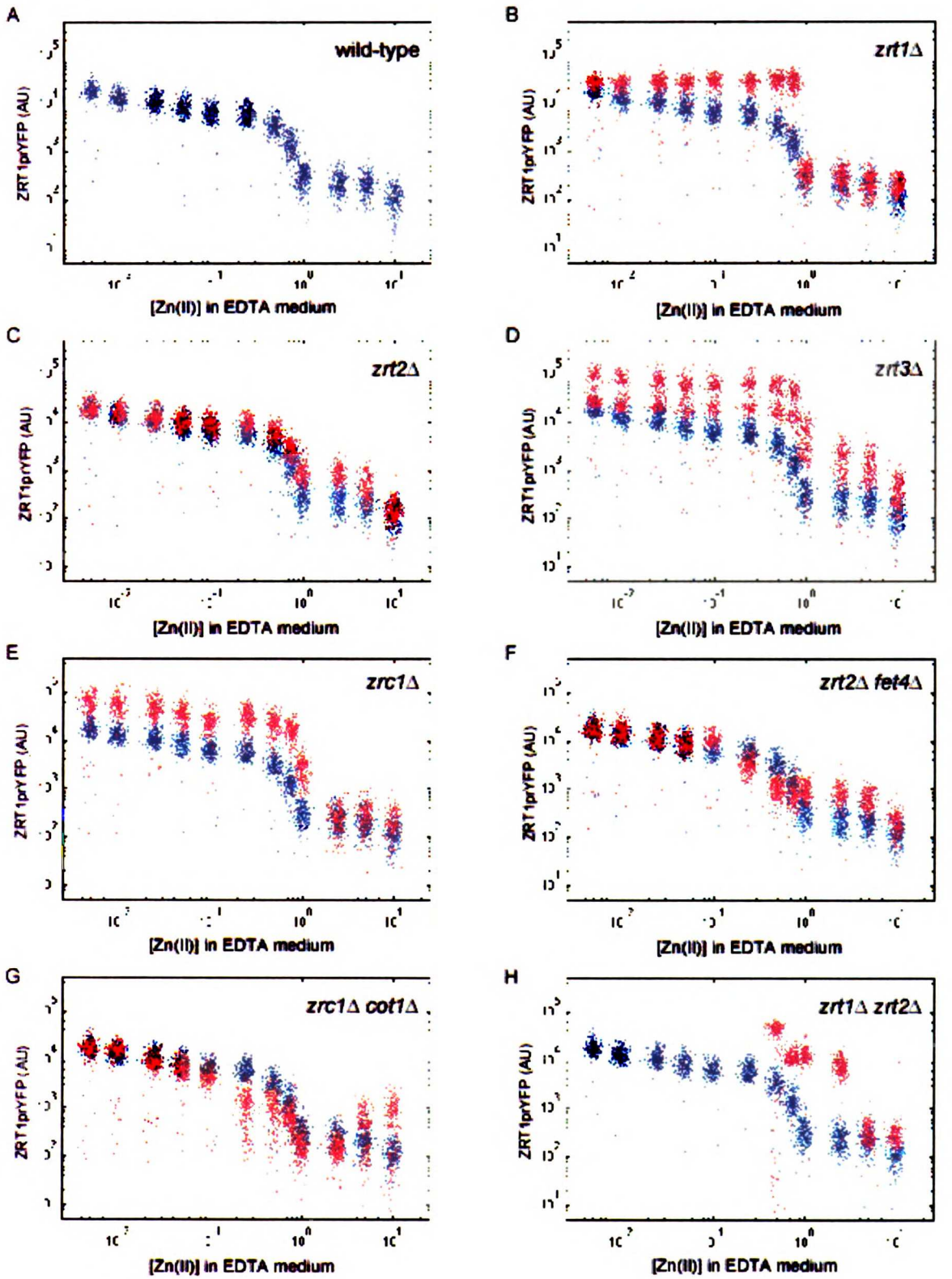

A
wild-type
B
zrt1Δ
C
zrt2Δ
D
zrt3Δ
E
zrc1Δ
F
zrt2Δ fet4Δ
G
zrc1Δ cot1Δ
H
zrt1Δ zrt2Δ
ZRT1prYFP (AU)
[Zn(II)] in EDTA medium

Figure 3.3 on previous page. ZRT1prYFP gene expression across a range of extracellular zinc concentration in strains lacking various zinc transporters. For all panels, single-cell FACS measurements are represented by individual points; the spread in the y dimension within each subpopulation is present for visualization purposes and does not signify small amounts of variability in zinc concentration. The wild-type cells are displayed in blue for all panels. (A) Wild-type cells containing ZRT1prYFP in blue. (B) ZRT1prYFP expression in *zrt1Δ* cells. (C) ZRT1prYFP expression in *zrt2Δ* cells. (D) ZRT1prYFP expression in *zrt3Δ* cells. (E) ZRT1prYFP expression in *zrc1Δ* cells. (F) ZRT1prYFP expression in *zrt2Δfet4Δ* cells. (G) ZRT1prYFP expression in *zrc1Δcot1Δ* cells. (H) ZRT1prYFP expression in *zrt1Δzrt2Δ* cells.

vacuolar sequestration of zinc to avoid toxicity in the face of high levels of extracellular zinc (MacDiarmid et al., 2003).

The individual removal of the *COT1* gene is predicted to be have one of two results: either no effect, due to redundancy with the *ZRC1* gene; or to have an effect similar to the removal of both *COT1* and *ZRC1*. In fact, the *cot1Δ* strain displays no significant difference from wild-type in Zap1-dependent gene expression (data not shown). Similarly, the individual deletion of the *ZRC1* gene should have one of these two effects according to the integrated model. Strikingly, the *zrc1Δ* strain displays more gene expression than wild-type at low and intermediate levels of extracellular zinc. This observation can be explained by the integrated model only by the invocation of a complex, multistable system with nonintuitive properties. Such a system requires a very specific set of quantitative relationships. It is also possible that the Zrc1 gene product has activities unanticipated by previous experiment and the proposed model. For example, if the model were adapted to include the possibility that Zrc1 transports zinc bidirectionally, the phenotype observed would be consistent with the model.

Measurement of a time-course of Zap1-dependent gene expression

In order to understand the time scales over which the zinc signal transduction pathway responds to changes in extracellular zinc, we measured the approach of

Zap1-dependent gene expression to steady-state values over a range of zinc concentrations. We measured this approach starting from two different concentrations of exogenous zinc added to the EDTA medium: 10 μM representing low zinc conditions, and 1 mM representing zinc repletion. The data are presented in figure 3.4 for both conditions. It is apparent that three hours after the change in medium zinc concentrations, cells are only beginning to respond. It takes the cells nine hours to approach the steady-state distribution, which is not realized until between twelve and fifteen hours. For cells moved from low zinc to higher zinc concentrations, even fifteen hours is a not sufficient time period for cells to fully suppress Zap1-dependent gene expression. This observation may reflect a lack of substantial growth under these conditions, which may be due to zinc toxicity that is caused by exposure to high zinc when Zap1 is active, the so-called "zinc shock" phenomenon (MacDiarmid et al., 2003).

Measurement of hysteresis in Zap1-dependent gene expression

The Zap1-dependent gene expression profile across all zinc concentrations tested in the EDTA medium displays a unimodal response in the wild-type strain. However, at least two hypothesized positive feedback loops can be identified: the induction of *ZAP1* in a Zap1-dependent manner, and the induction of *ZRC1* by Zap1. The autocatalytic property of Zap1 may result in increasing expression of the transcription factor. If the increasing expression allows for more activity, this represents a potentially strong feedback loop. Also, the Zrc1 protein is a vacuolar membrane transporter that is hypothesized to transport zinc from the cytoplasm into the vacuole (Li and Kaplan, 1998). Therefore the induction of Zrc1 during zinc starvation may result in further depletion of the cytoplasmic zinc pool, another possible positive feedback loop.

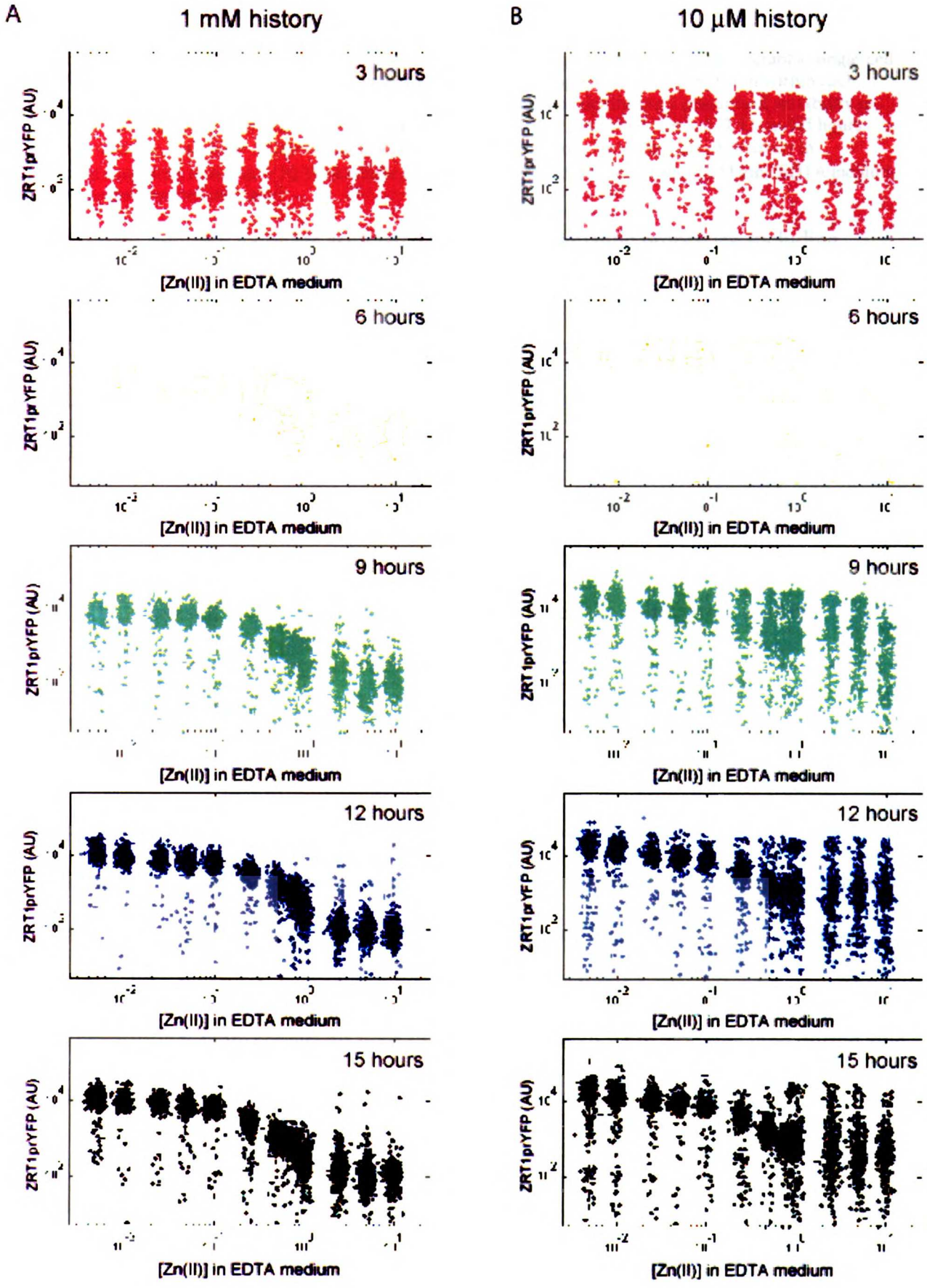
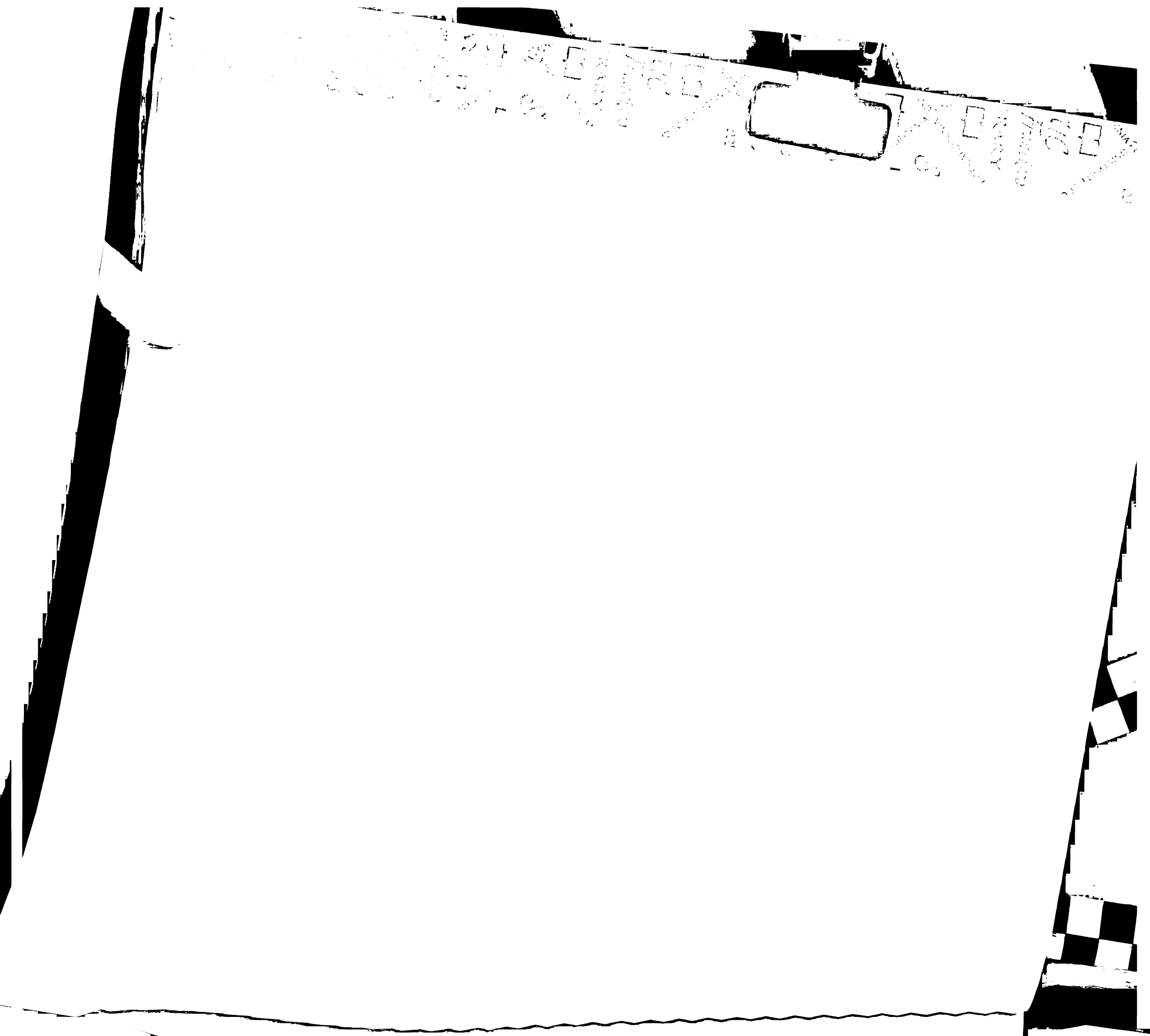

A
1 mM history
B
10 µM history
3 hours
6 hours
9 hours
12 hours
15 hours
ZRT1prYFP (AU)
[Zn(II)] in EDTA medium

Figure 3.4 on previous page. Time courses of Zap1-dependent gene expression. For all graphs, single-cell FACS measurements are represented by individual points; the spread in the y dimension within each subpopulation is present for visualization purposes and does not signify small amounts of variability in zinc concentration. The color scheme is red, yellow, green, blue, and black, for 3, 6, 9, 12, and 15 hours, respectively. **(A)** Time course of Zap1-dependent expression from 1 mM zinc EDTA medium to concentrations indicated. **(B)** Time course of Zap1-dependent expression from 10 μM zinc EDTA medium to concentrations indicated.

Positive feedback loops are associated with certain signaling properties, including increased sensitivity of the response to ligand concentration; multistability in the response; and hysteresis, or a history-dependence wherein induced cells are more difficult to repress at intermediate signal than repressed cells (Angeli et al., 2004; Ferrell, 2002; Ferrell and Xiong, 2001). While we cannot assess the sensitivity of the zinc input-output response without direct measurement of intracellular zinc concentrations, we can test if hysteresis is a property of the signal transduction pathway.

To test if hysteresis is detectable, we compared the steady-state level of Zap1-dependent gene expression over a range of concentrations in the wild-type strain subjected to two different histories, a low zinc (10 μM) and a high zinc (5 mM) EDTA medium history (Fig. 3.5). If hysteresis exists, it is not dramatic; however, the intermediate concentrations of zinc do show a tendency to higher levels of Zap1-dependent gene expression in the low zinc history, but this observation is subtle and would require further repetition. Interestingly, at high zinc concentrations there seems to be less Zap1-dependent gene expression in cells with a low zinc history. This observation is more consistent with the idea of an overshoot in zinc signaling during zinc shock than hysteretic behavior.

We repeated the same hysteresis experiment in the *zrt3Δ* strain background, as this strain had demonstrated both bimodality and an increased level of Zap1-dependent expression. We hypothesized that the deletion of the *ZRT3* negative feedback element



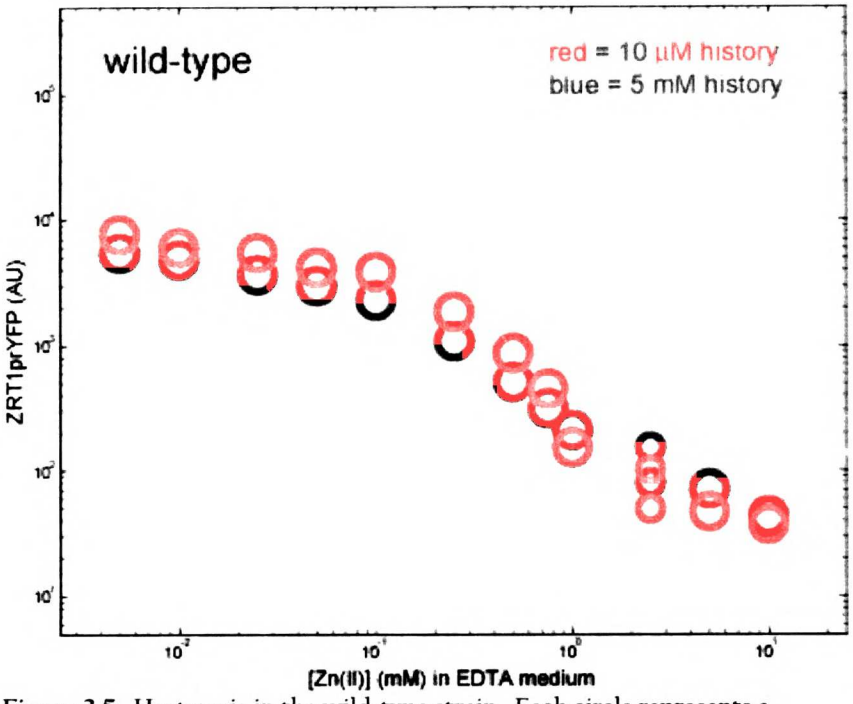


Figure 3.5. Hysteresis in the wild-type strain. Each circle represents a discrete population of cells; the diameter of the circle is proportional to the percentage of cells in each subpopulation. Red circles are populations with a low zinc (10 µM) EDTA medium history, and blue circles are populations with a high zinc (5 mM) EDTA medium history.

would possibly potentiate the hysteretic effects of any positive feedback. The data are not consistent with any stronger hysteresis than displayed by the wild-type strain (Fig. 3.6). The low zinc history cells do display less tendency to bimodality, particularly at the high-intermediate zinc concentrations. For both the wild-type and *zrt3Δ* strains, these experiments require repetition and confirmation that they do not represent kinetic intermediates before the differences in Zap1-dependent gene expression between the low and high zinc histories could be further investigated.

The impact of Zap1 concentration and autoregulation on gene expression

Previous work had identified *ZAP1* itself as a Zap1-regulated gene (Zhao et al., 1998). However, the difference in Zap1 abundance between low and high zinc conditions that results from this autocatalysis is neither necessary nor sufficient for the difference in Zap1-dependent gene expression in those two conditions (Bird et al., 2000). We wished to examine the role of Zap1 autoregulation in the quantitative aspect of zinc signal transduction, as this previous work had not suggested any role for such

autoregulation. We constructed two additional strains that expressed Zap1 in a zinc-insensitive manner. In the *ZAP1mZREprZAP1* strain, the zinc responsive element (ZRE) in the chromosomal *ZAP1* promoter has been mutated to eliminate Zap1 binding and any Zap1-dependence in gene expression (Zhao et al., 1998). In the *PGK1prZAP1* strain, the native *ZAP1* promoter has been replaced by the high-level, zinc-insensitive promoter from the *PGK1* gene. We then measured the Zap1-dependent expression of yVYFP across a range of zinc concentrations in each of the mutated strain as well as the wild-type strain (Fig. 3.7). It is clear that at high levels of zinc, there is no effect on Zap1-dependent gene expression of perturbation of the Zap1 autoregulation in the *ZAP1mZREprZAP1* strain. Neither is there any effect by the predicted substantial overexpression of Zap1 in the *PGK1prZAP1* strain. However, at low levels of extracellular zinc, the level of Zap1-dependent gene expression is significantly dependent on the expression levels of *ZAP1*. Overexpression of Zap1 results in the maximal level of Zap1-dependent gene expression seen in the wild-type strain at much higher extracellular zinc levels. The lack of autoregulation results in an approximately two-fold decrease in Zap1-dependent gene expression across a range of intermediate to low extracellular zinc concentrations.

Development of a growth medium capable of a wider range of steady-state zinc levels

All of the experiments detailed so far were conducted in a synthetic complete yeast medium containing EDTA to chelate the contaminating zinc that is present in trace amounts in the laboratory environment. There are sufficient amounts of this contaminating zinc to sustain a low level of growth in supposedly zinc-free medium, but the addition of EDTA eliminates this growth (data not shown). The use of

EDTA-containing medium presents a problem in exploring a wide range of zinc concentrations, as there is a dramatic change in the calculated free zinc concentrations when zinc is increased from slightly below to slightly above the chelator concentration. Therefore, two different chelation media were developed. One contains both EDTA and citrate as zinc chelators and is used for low zinc levels. The second contains only citrate, which is a much weaker zinc chelator than EDTA. This citrate-containing medium can be used to access a wider range of free zinc concentrations. The free zinc concentrations in each of the formulated media were calculated by the MAXCHELATOR program (Patton et al., 2004). All subsequent experiments employ these two media to enable a wider range of free extracellular zinc concentrations.

Measurement of zinc dependence of all transporter gene promoters

One of the barriers to the development of a quantitative integrated model of zinc signal transduction is the lack of knowledge about the Zap1-dependent

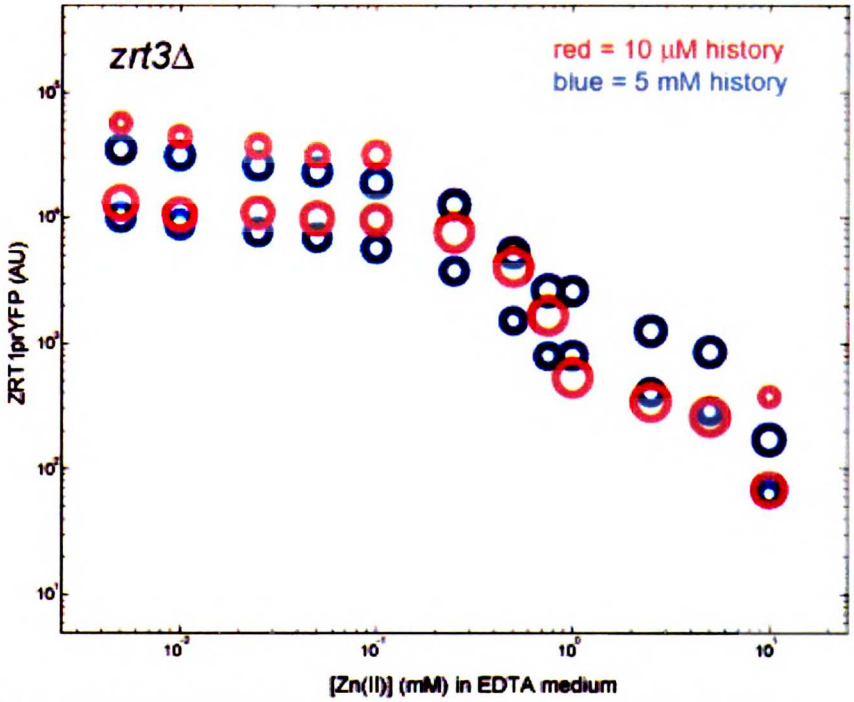


Figure 3.6. Hysteresis in the *zrt3Δ* strain. Each circle represents a discrete population of cells; the diameter of the circle is proportional to the percentage of cells in each subpopulation. Red circles are populations with a low zinc (10 μM) EDTA medium history, and blue circles are populations with a high zinc (5 mM) EDTA medium history.

expression of all the transporter genes and the Zap1 transcription factor itself. Therefore, we measured the expression of yVYFP from the promoters of the *ZAP1*, *ZRT2*, *ZRT3*, *ZRC1*, *COT1*, *FET4*, and *PGK1* genes over the expanded range of predicted extracellular free zinc concentrations (Fig. 3.8). The *PGK1* gene promoter was included as a zinc-insensitive control as its expression has been reported to be Zap1-insensitive (Lyons et al., 2000).

As expected, the *ZRT1*, *ZAP1*, *ZRT2*, *ZRT3*, and *ZRC1* promoters display zinc-dependence in gene expression, though each promoter has a distinct basal level of gene

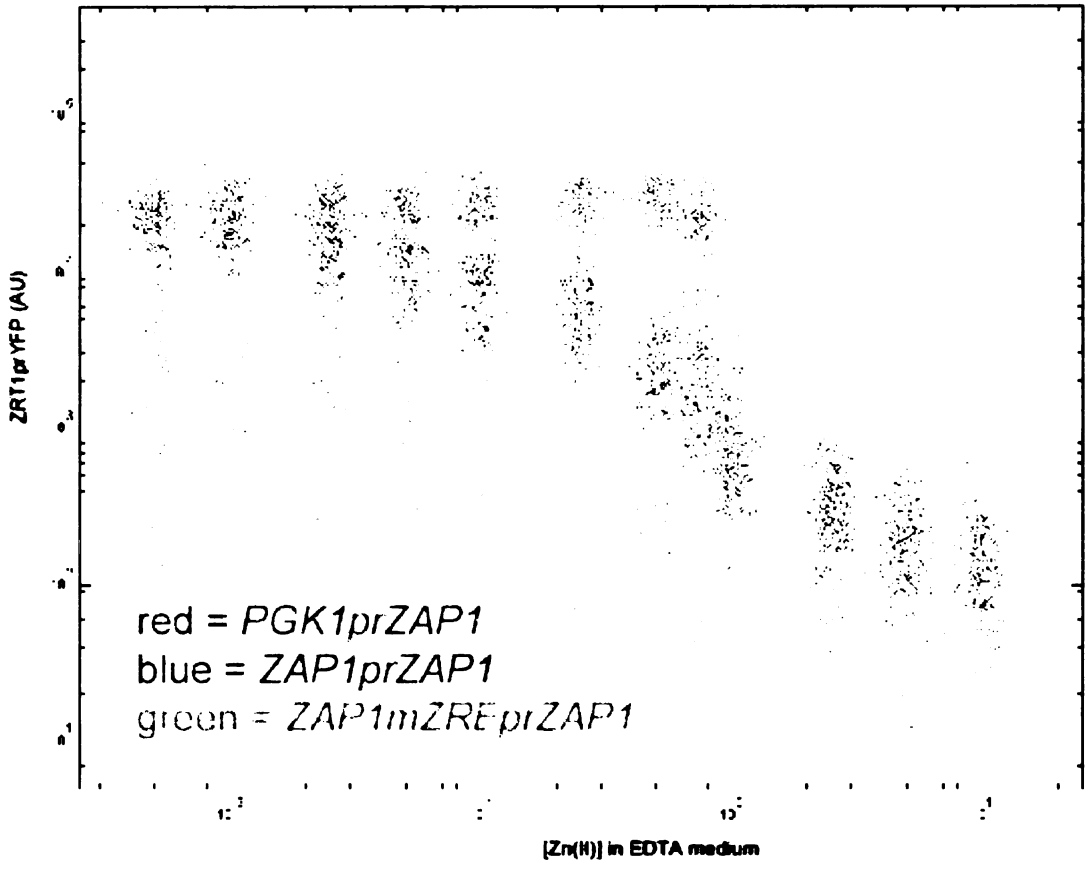
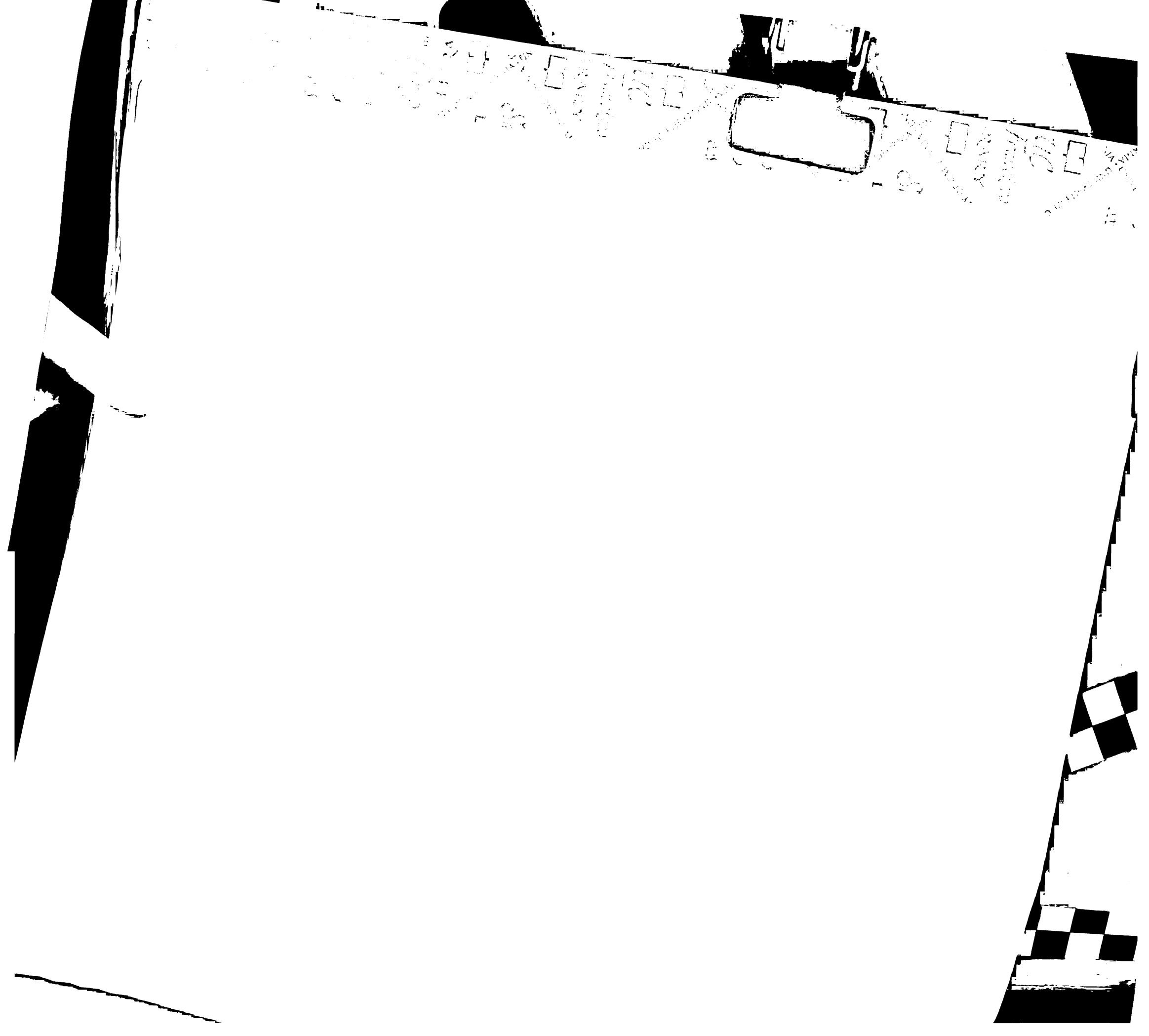
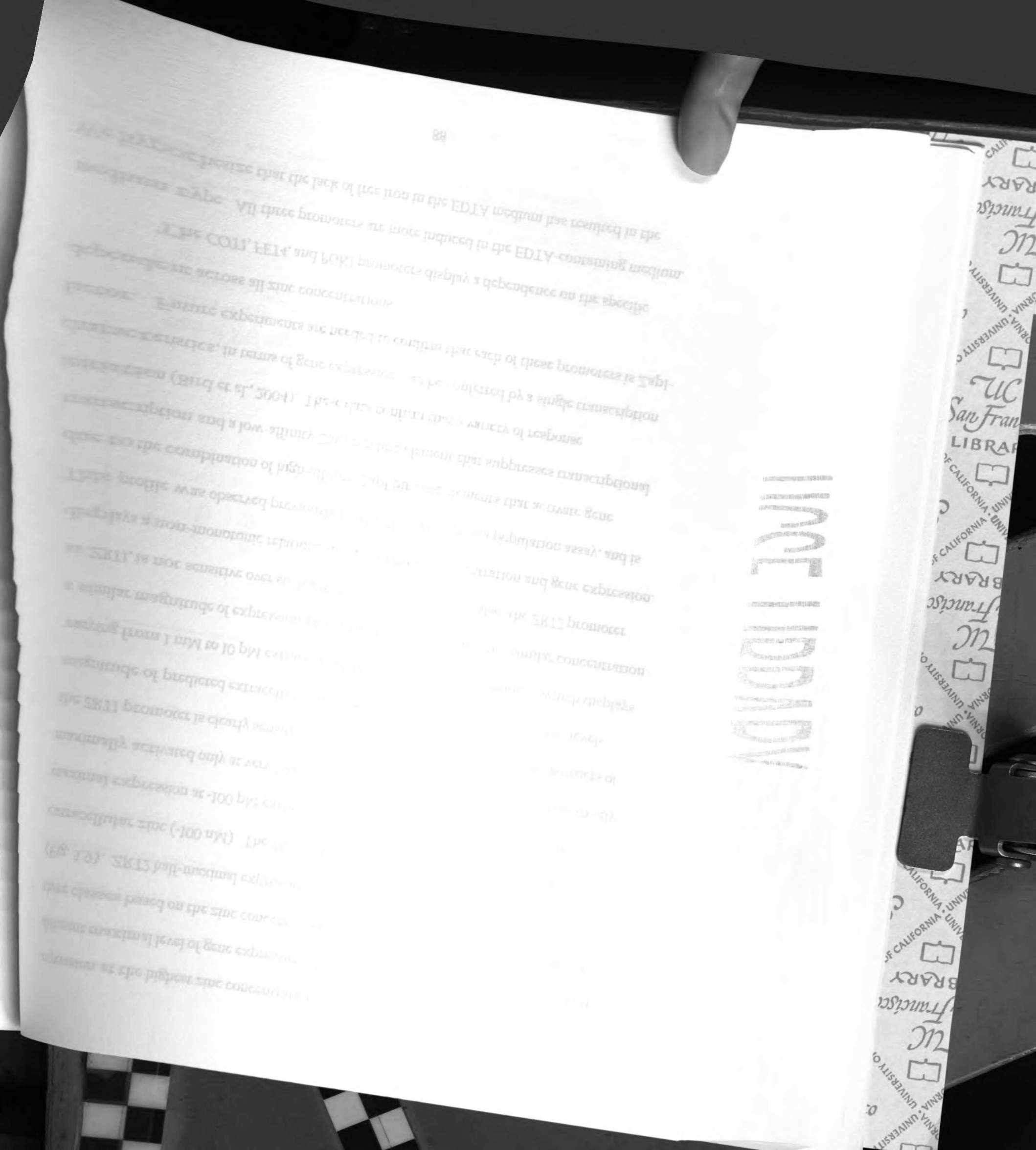


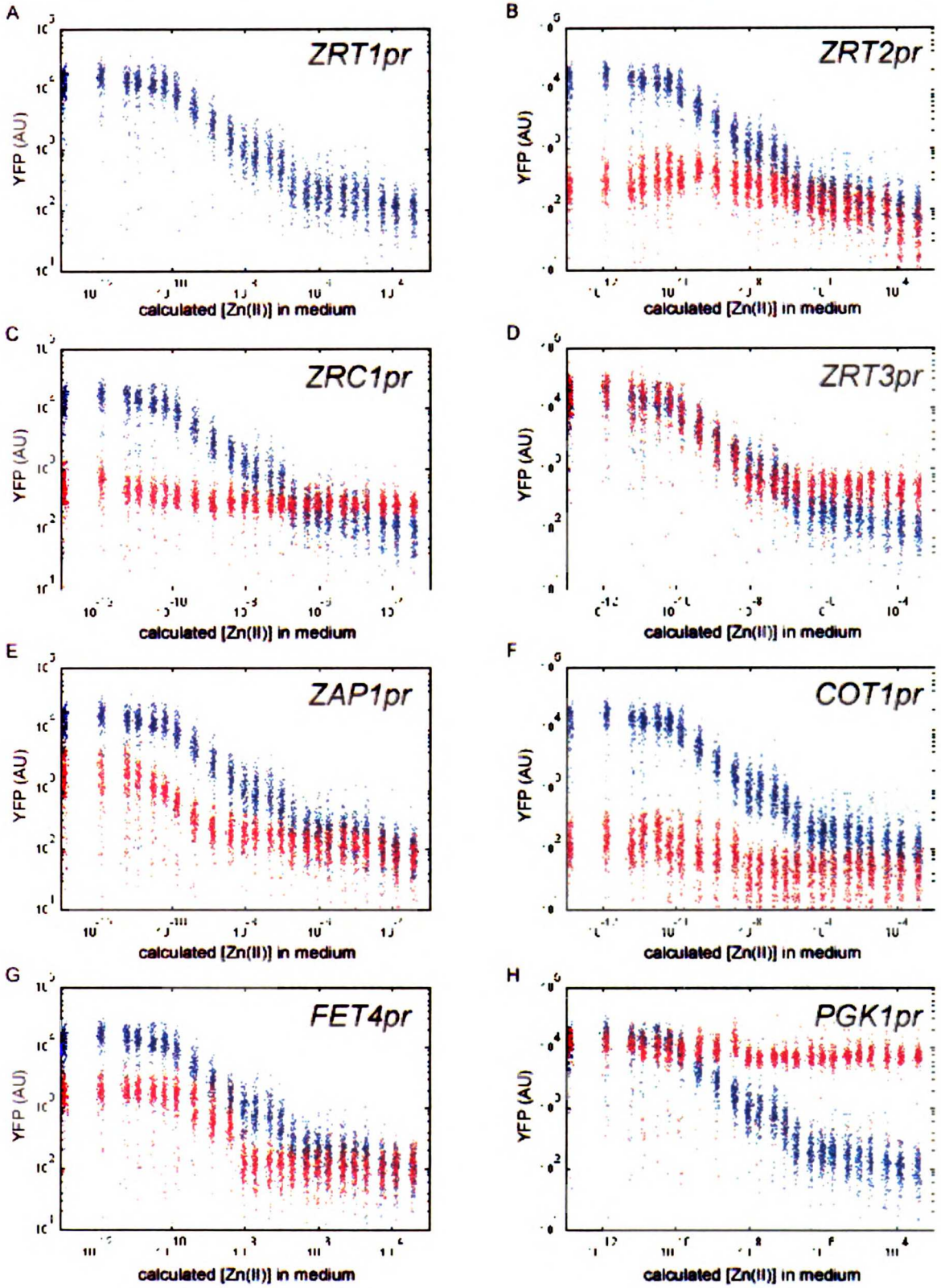
Figure 3.7. Zap1 levels affect the Zap1-dependent gene expression at lower but not high levels of extracellular zinc. Single-cell FACS measurements are represented by individual points; the spread in the y dimension within each subpopulation is present for visualization purposes and does not signify small amounts of variability in zinc concentration. Strains are represented by colors as indicated.



expression at the highest zinc concentration. In addition, each promoter displays a different maximal level of gene expression. The five promoters can also be separated into three classes based on the zinc concentration at which half-maximal expression occurs (Fig. 3.9). *ZRT2* half-maximal expression occurs at higher levels of calculated extracellular zinc (~100 nM). The *ZRT1*, *ZAP1*, and *ZRT3* promoter all display half-maximal expression at ~100 pM extracellular zinc, and the *ZRC1* promoter is half-maximally activated only at very low levels of extracellular zinc (~10 pM). Interestingly, the *ZRT1* promoter is clearly sensitive to the extracellular zinc levels across ~8 orders of magnitude of predicted extracellular free zinc concentration, with expression levels varying from 1 mM to 10 pM extracellular free zinc. The *ZRT3* promoter, which displays a similar magnitude of expression and half-maximal activation at a similar concentration as *ZRT1*, is not sensitive over such a wide range, in contrast. Also, the *ZRT2* promoter displays a non-monotonic relationship between zinc concentration and gene expression. This profile was observed previously for the *ZRT2* mRNA in a population assay, and is due to the combination of high-affinity Zap1 binding elements that activate gene transcription and a low-affinity Zap1 binding element that suppresses transcriptional initiation (Bird et al., 2004). These data confirm that a variety of response characteristics, in terms of gene expression, can be conferred by a single transcription factor. Future experiments are needed to confirm that each of these promoters is Zap1-dependent across all zinc concentrations.

The *COT1*, *FET4*, and *PGK1* promoters display a dependence on the specific medium type. All three promoters are more induced in the EDTA-containing medium. We hypothesize that the lack of free iron in the EDTA medium has resulted in the




A
ZRT1pr
B
ZRT2pr
C
ZRC1pr
D
ZRT3pr
E
ZAP1pr
F
COT1pr
G
FET4pr
H
PGK1pr
YFP (AU)
calculated [Zn(II)] in medium

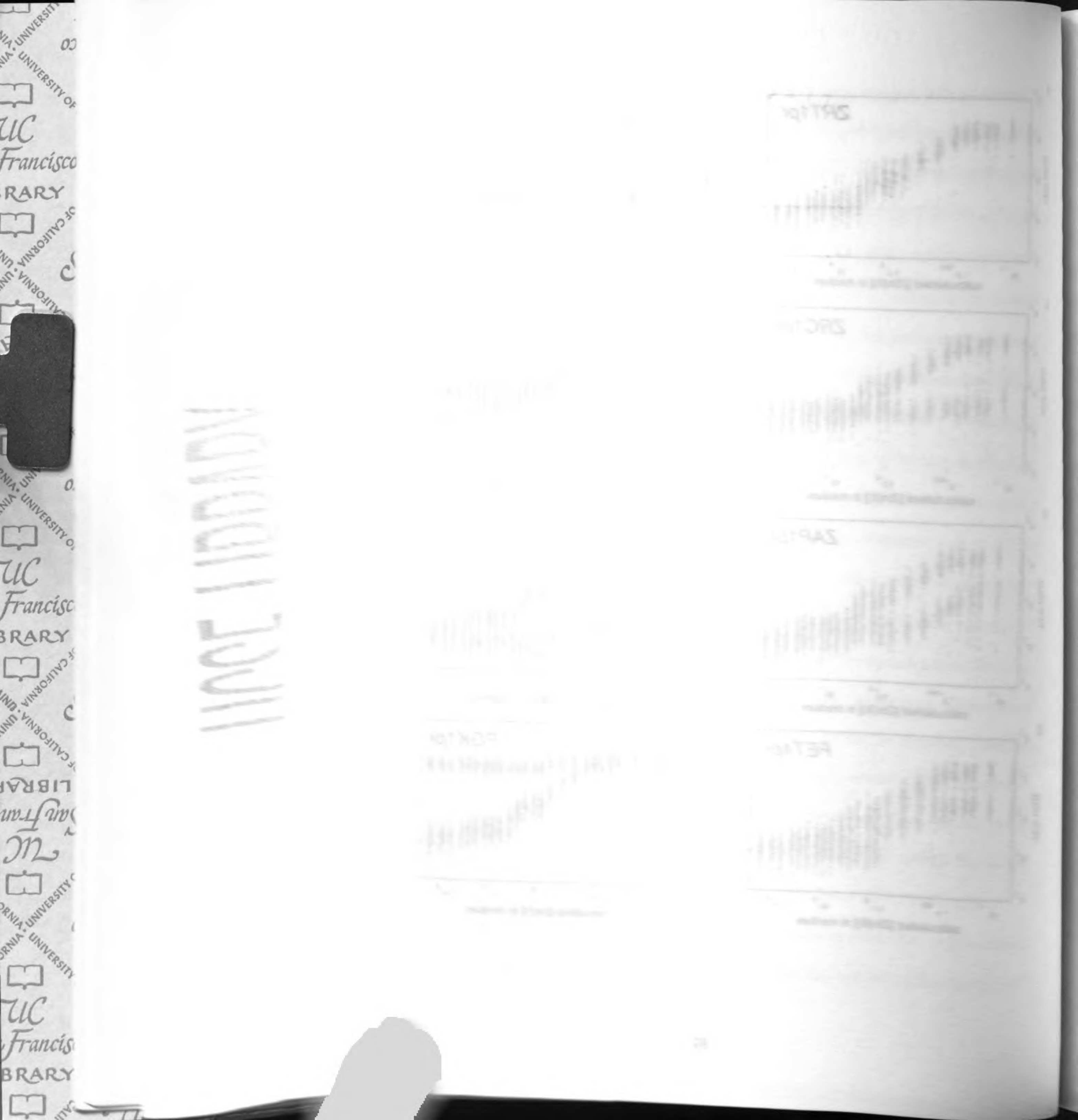


Figure 3.8 on previous page. YFP expression across a range of zinc concentrations for eight different promoters on a log-log scale. For all panels, single-cell FACS measurements are represented by individual points; the spread in the y dimension within each subpopulation is present for visualization purposes and does not signify small amounts of variability in zinc concentration. For all panels, the ZRT1 promoter is indicated in blue for comparative purposes. **(A)** The *ZRT1* promoter in blue. **(B)** The *ZRT2* promoter in red. **(C)** The *ZRC1* promoter in red. **(D)** The *ZRT3* promoter in red. **(E)** The *ZAP1* promoter in red. **(F)** The *COT1* promoter in red. **(G)** The *FET4* promoter in red. **(H)** The *PGK1* promoter in red.

induction of *COT1* and *FET4*, as both these genes are known to be upregulated by low iron levels (Waters and Eide, 2002). It will be important to reformulate the medium to provide equivalent levels of free ionized iron across all zinc concentrations before repeating these experiments in order to determine the zinc-dependence of these promoters. Additionally, the *COT1* promoter displays a heterogeneity in induction in the EDTA-containing medium. This heterogeneity may be the underlying reason for the heterogeneity in *ZRT1* expression seen in the cells lacking *ZRT3* (Figs. 3.3D and 3.6).

Discussion

The emphasis on the quantitative measurement of signal transduction pathway characteristics has allowed us to accumulate a much finer resolution understanding of zinc signaling in budding yeast. By testing the predictions of an initial qualitative integrated model, we have uncovered novel and unexpected aspects of zinc signal transduction. Below, we comment on some of these discoveries and provide explanatory hypotheses that can be tested in future experiments. Furthermore, we have contributed a substantial set of data toward the goal of a quantitative integrated model of zinc signaling and intracellular zinc flux.

We have observed that the zinc signaling pathway is sensitive and responsive to extracellular free zinc concentrations over a striking range – from 10 pM to 1 mM.

Moreover, at each different concentration of zinc, the response is homogeneous within the cellular population, but quantitatively differs from one concentration of zinc to the next. The observation of a graded, unimodal response over 8 orders of magnitude of extracellular nutrient concentration has not been documented previously, to our knowledge. As we observe that the zinc response is specific for each concentration of zinc, we can speculate that small changes in gene expression of the zinc-dependent genes help the cell adapt to small differences in the level of extracellular zinc. This graded response is consistent with tight intracellular control of zinc concentration, though clearly the levels of intracellular zinc must differ slightly at every extracellular concentration in order to produce a different level of Zap1 activity. This one-to-one correspondence of extracellular nutrient concentration with a discrete intracellular state

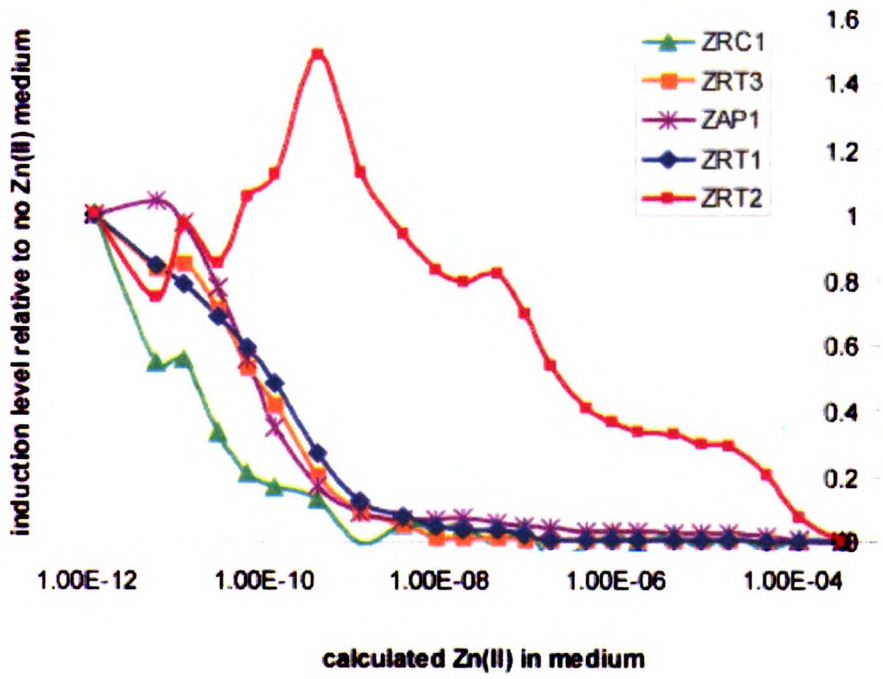
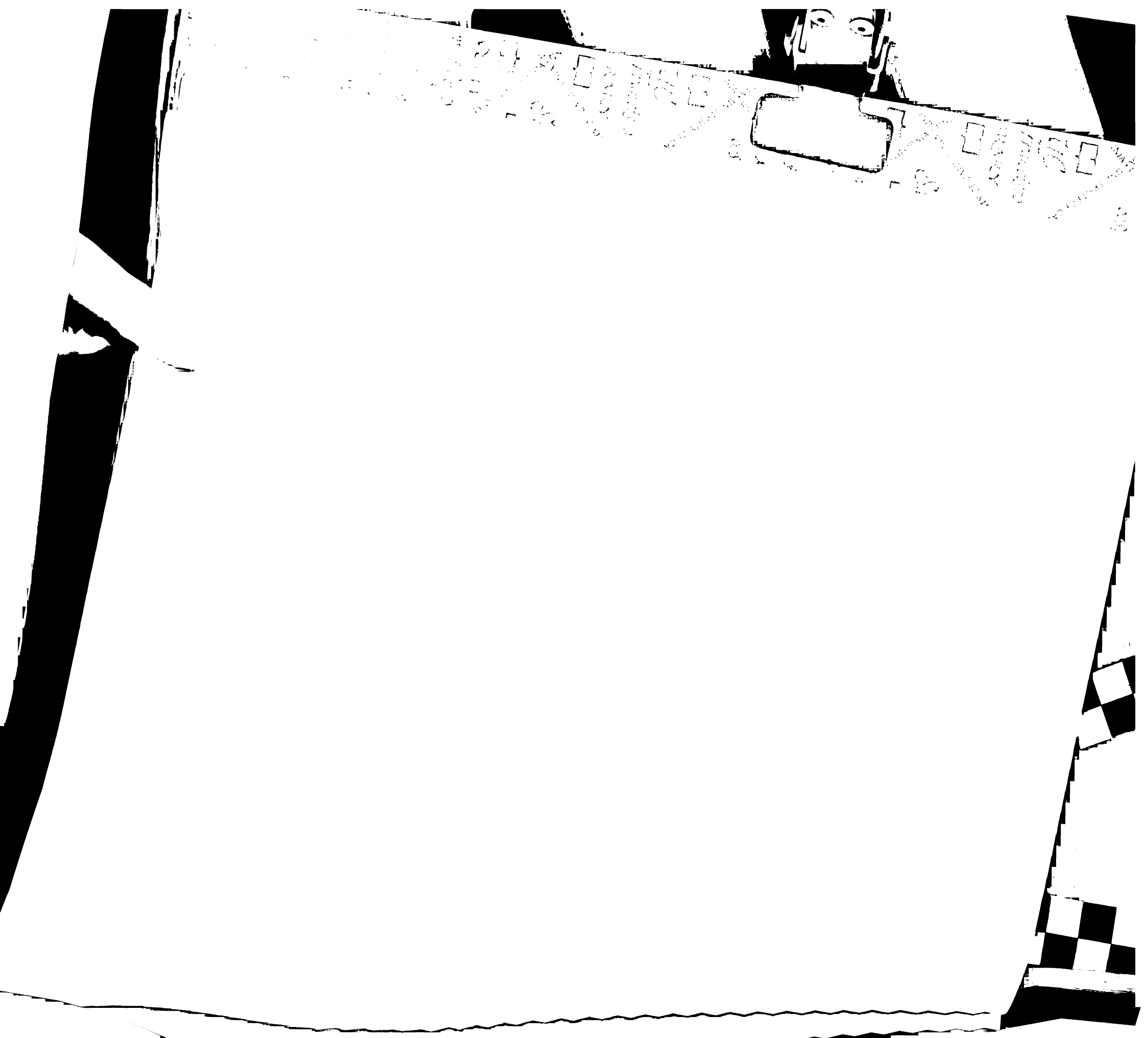


Figure 3.9. A linear representation of the gene expression for select promoters across a range of calculated extracellular zinc concentrations, relative to gene expression at 1 pM extracellular zinc.



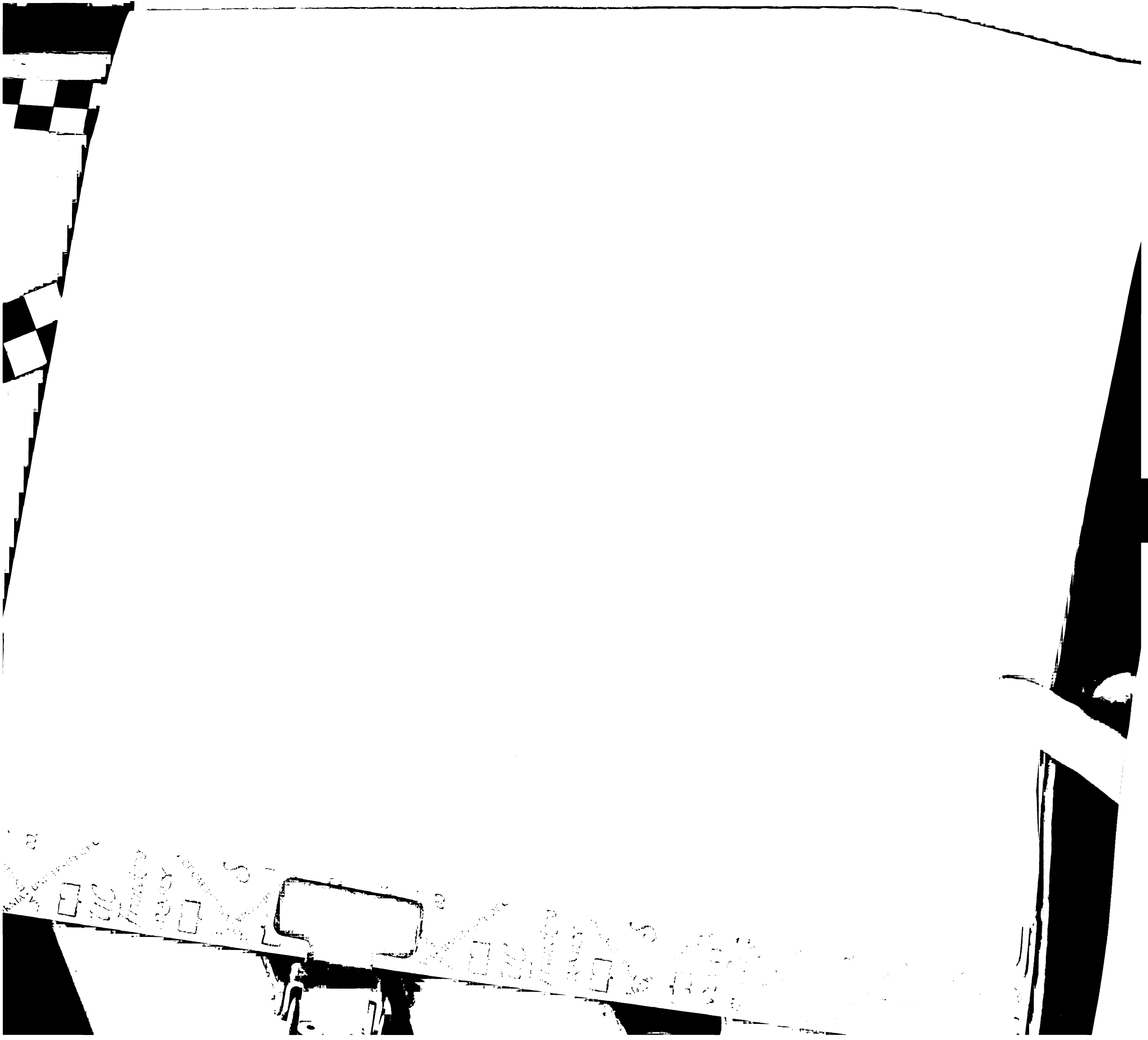
is not necessarily a predicted component of a nutrient response. Future work must confirm the actual concentrations of zinc in the chelation media employed in order to substantiate our assertions. In addition, the medium recipe must be reformulated to provide relatively constant concentrations of other divalent and trivalent metal ions in the presence of the chelators. It would be extremely helpful to be able to assay intracellular zinc levels directly, in order to confirm our hypothesis that intracellular levels are different at every different extracellular concentration, and in order to measure how substantially intracellular levels vary.

We have reported some unexpected results when the vacuolar membrane zinc transporter gene products are eliminated. Three observations in particular are puzzling. First, the *zrt3Δ* strain displays a bimodal pattern of gene expression over a range of extracellular zinc concentrations. This bimodality could represent multistability, perhaps due to the induction of *ZRC1* or *ZAP1* in the higher-expressing subpopulation. Alternatively, this bimodality could represent a consequence of a normally "silent" heterogeneity in the population, such as the bimodality in gene copy number that occurs due to cell cycle progression. The bimodality could be due directly to the absence of the Zrt3 gene product and affect all gene expression nonspecifically. Measurement of the expression from the zinc-insensitive promoter *PGK1* could confirm or disprove this hypothesis.

The second puzzling observation involves the heightened expression in the strain lacking the Zrc1 gene product, putatively a vacuolar influx transporter. Heightened expression contradicts the straight-forward prediction of the qualitative model. This contradiction is especially glaring in the face of evidence that the strain lacking both the

Zrc1 and Cot1 gene products behaves as expected, with reduced expression when compared to wild-type cells. We can hypothesize that the Zrc1 transporter acts as a bidirectional transport, allowing zinc to flow into or out of the vacuole, and the Cot1 transporter a unidirectional influx transporter. However, this hypothesis is not supported by any biochemical evidence nor the membrane topology of Zrc1, and would require further experimentation to confirm or disprove. Alternatively, it is possible that the *ZRC1* gene is required for the activity of another component of the signal transduction pathway component such as the Zrt3 vacuolar efflux transporter. In this case, the *zrc1Δzrt3Δ* strain may behave similarly to the *zrc1Δ* strain. Finally, it remains possible but unlikely that the heightened level of expression in the *zrc1Δ* strain is a kinetic intermediate, due to increased lability in intracellular zinc concentration in a strain lacking substantial vacuolar influx capability. This hypothesis could be eliminated by showing that the data presented do in fact represent steady-state values.

The third puzzling observation is the fact that both the *zrt3Δ* strain and the *zrc1Δ* strain display a level of Zap1-dependent gene expression that is substantially (>10-fold) higher than the maximal level of expression in wild-type. We know that we have achieved the maximal expression in wild-type that can occur, as even lower levels of extracellular zinc starvation do not increase gene expression but do result in the cessation of cellular growth. If both Zap1 activity and cellular viability are related to intracellular zinc levels, we must postulate that the *zrt3Δ* strain and the *zrc1Δ* strain display intracellular levels high enough to sustain growth but lower than the wild-type intracellular levels at the time of zinc starvation at which the wild-type strain stops growing. These assertions are contradictory, and suggest that either the growth



requirement for intracellular zinc has decreased in the mutant strains, which is arguably unlikely, or that Zap1 activity is not exclusively related to intracellular zinc levels. One appealing alternative hypothesis is that Zap1 activity is regulated by the zinc flux from the extracellular environment through the cytoplasm into the vacuole. In this model, it is not the cytoplasmic zinc concentration but the residency time of each zinc ion in the cytoplasm that determines Zap1 activity. This model is reasonable if we imagine that both essential cellular proteins and the vacuolar zinc transporters bind zinc very quickly, but Zap1 binds zinc on a slower time-scale. In the wild-type strain, zinc-binding to cellular proteins is irreversible, but zinc is constantly flowing into and out of the vacuole. Therefore Zap1 has the opportunity to bind zinc and be inhibited. In the *zrt3Δ* strain, zinc cannot return to the cytoplasm from the vacuole, so the residency time in the cytoplasm is extremely low. Cellular proteins bind zinc quickly, enabling cell growth to continue, but Zap1 remains unbound and fully active. This model is highly speculative but the kinetic parameters are supported by previous in vitro work (Bird et al., 2003). This model could be tested directly by construction of strains that have the same ratio of vacuolar influx to vacuolar efflux transporters, but different total amounts of vacuolar transporter. In all these strains, the steady-state concentration of vacuolar and cytoplasmic zinc should be constant, but the residency time will change. If Zap1 activity changes in the absence of a change in zinc concentration, then it is likely that kinetic factors are important in controlling Zap1 activity. Alternatively, a simpler hypothesis is that the vacuolar transporters may be directly involved in zinc signaling.

We have made quantitative measurements that show different thresholds of zinc-sensitive expression for the distinct zinc membrane transporters. These measurements will be useful directly for the construction of a quantitative integrated

model of zinc signaling. Also, these measurements provide a data set for comparison of promoter structure, specifically in regards to zinc responsive element number, position, and sequence. It would be interesting to identify the motifs that confer substantially different basal levels of expression on *ZRT1* and *ZRT3*, whose promoters display nearly identical maximal rates of expression.

Much work remains to be done to understand the individual relationships that are integrated to form the zinc-responsive signal transduction pathway in budding yeast. However, we have begun a quantitative approach whose promise is the detailed understanding of the behavior of a relatively simple but essential signal transduction pathway. The lessons learned from the analysis of the zinc signal transduction pathway in budding yeast will be applicable to a wider set of nutrient-responsive signaling pathways in all eukaryotic cells. The mathematical structures at the base of zinc signaling will likely be found in the diverse set of all cellular signal transduction pathways.

Materials and Methods

Unless otherwise indicated, yeast strains were grown at 30°C in synthetic complete medium containing glucose with addition of 0.1 mg/mL adenine and 0.1 mg/mL tryptophan to suppress autofluorescence. Chelation media containing different concentrations of zinc(II) sulfate was made from yeast nitrogen base without amino acids, ammonium sulfate, and zinc (BioLabs 101) as described previously (Zhao and Eide, 1996a). FACS measurements were performed on Becton-Dickenson LSR II machine. All yeast strains expressing YFP were constructed by insertion at the *leu2* locus into the EY0690 strain background of a pJRL2 derivative (Raser and O'Shea, 2004) containing

the 1000 base pairs of promoter sequence immediately 5' to the start codon of interest, fused to the yVYFP coding sequence. Deletion strains were constructed with the *Candida glabrata* system as described previously with elimination of the entire coding sequence from the start codon to the stop codon (Steger et al., 2003). The *ZAP1* overexpression allele was constructed with a Pringle promoter vector containing 700 base pairs of the *PGK1* promoter, and the uninducible *ZAP1* allele was constructed by a loop-in loop-out strategy employing a *ZAP1* promoter fragment in which the ZRE was mutated by transversion (Zhao et al., 1998). Steady-state experiments were conducted by switching cells grown at OD_{600} of less than 0.5 in high zinc medium overnight to the media containing different concentrations of zinc for at least 15 hours growth, at OD_{600} of less than 0.5. Hysteresis experiments were conducted by growth in the initial condition for at least 24 hours, followed by growth for an additional 18 hours in the new condition.

CHAPTER FOUR:

Quantitative analysis of phosphate-responsive signal transduction

The work reported in this chapter has been performed in general collaboration with Dr. Dennis D. Wykoff; all reported data was collected by the author himself.

Introduction

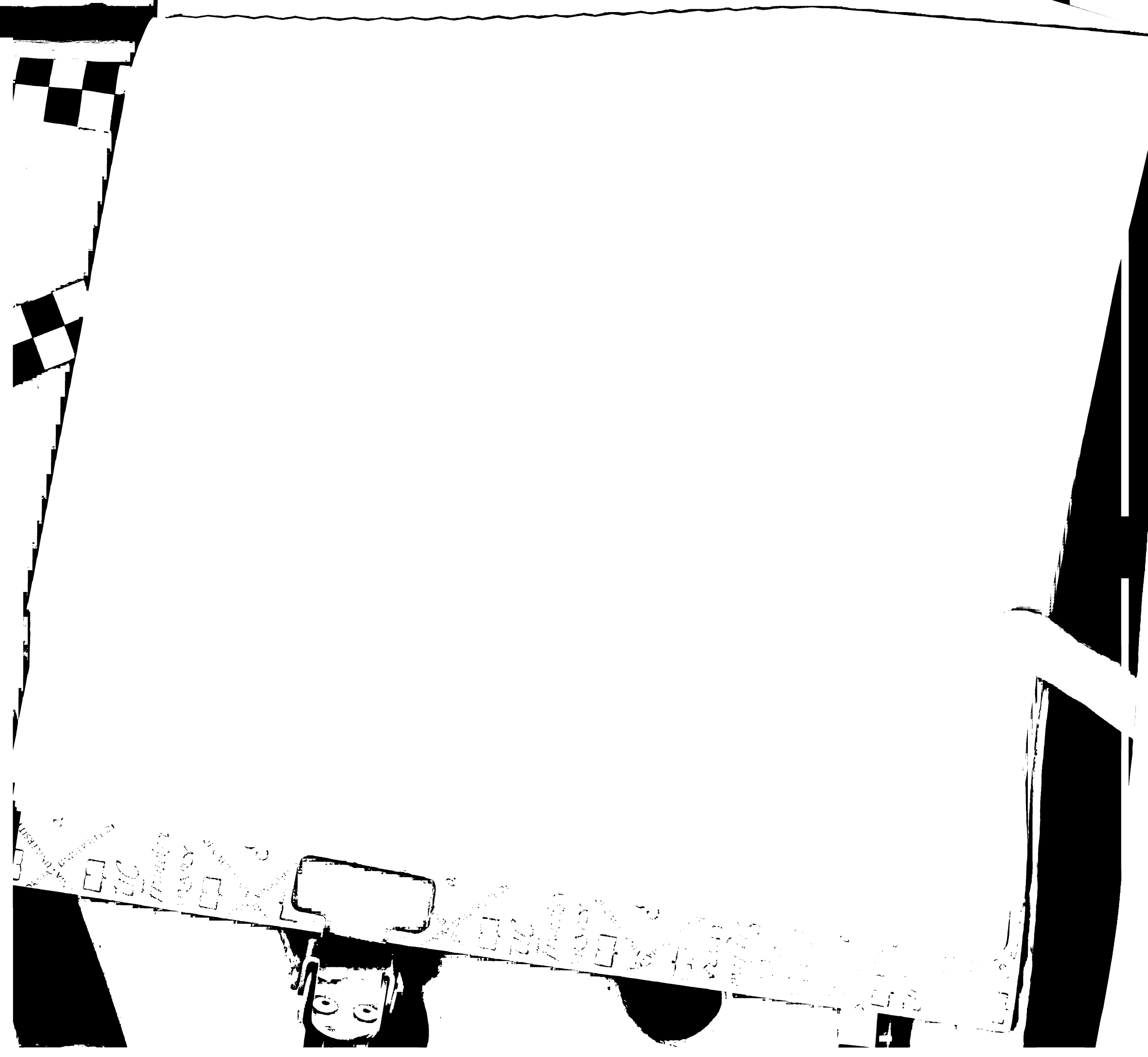
Budding yeast cells respond to environmental phosphate deprivation with the induction of a set of phosphate-responsive genes, including the repressible acid phosphatase gene *PHO5* (for review, see (Lenburg and O'Shea, 1996)). In low phosphate conditions, cells express the acid phosphatase Pho5; in high phosphate conditions, Pho5 is transcriptionally repressed (Toh-e et al., 1981). The transcription factor Pho4 is responsible for this phosphate-dependent regulation (Vogel et al., 1989). Pho4 is regulated by phosphorylation on multiple serine residues in high phosphate-grown cells (Kaffman et al., 1994). This phosphorylation results in its inactivation by multiple mechanisms, including the stimulation of nuclear export, the inhibition of nuclear import, and the inhibition of binding to its transcriptional coactivator, Pho2 (Komeili and O'Shea, 1999). The kinase responsible for Pho4 phosphorylation is the cyclin-dependent kinase (CDK) Pho85, which is active for Pho4 phosphorylation only in association with the cyclin Pho80 (Kaffman et al., 1994). The nuclear Pho85/Pho80 complex associates with the CDK inhibitor Pho81 in both high and low phosphate conditions; however, Pho81 is required for inhibition of the kinase, leading to activation of Pho4, only in low phosphate conditions (Schneider et al., 1994). Together, the Pho81/Pho80/Pho85 kinase and the Pho4 and Pho2 transcription factors constitute the core Pho pathway, a phosphate-responsive signal transduction pathway.

These core Pho pathway components that regulate *PHO5* expression were originally identified genetically (To et al., 1973; Ueda et al., 1975). As expected, loss of function mutations in *PHO4*, *PHO2*, or *PHO81* result in the elimination of Pho5 expression in low phosphate, or the Pho-uninducible (Pho^{U}) phenotype. Conversely, loss of

function mutations in *PHO80* or *PHO85*, or dominant gain of function mutations in *PHO4* or *PHO81* result in the constitutive expression of the Pho5 acid phosphatase, the so-called Pho^C phenotype. While the biochemical interactions between the Pho80/Pho85 kinase, the Pho4 transcription factor, and the subsequent steps in transcriptional activation are fairly well-understood, there is no definitive biochemical evidence that explains the connection between environmental phosphate levels and regulation of Pho81/Pho80/Pho85 kinase activity. A number of genes have been identified whose mutation results in Pho5 expression, including the genes that encode the plasma membrane protein Pho84, a high-affinity phosphate/proton symporter (Bun-Ya et al., 1991); the adenosine kinase Ado1 and the adenylate kinase Adk1 (Huang and O'Shea E, 2005); and the Pho84 endoplasmic reticulum chaperone protein Pho86 (Bun-ya et al., 1996; Lau et al., 2000).

Components of the Pho pathway also play a role in phospholipid metabolism and signaling (Flick and Thorner, 1998). Deletion of the budding yeast phospholipase C gene *PLC1* results in a yeast strain unable to grow at 37°. This temperature sensitivity is suppressed by growth on a phosphate-depleted medium, by deletion of the *PHO80* gene, or by overexpression of either the CDK inhibitor *PHO81* or a second gene *SPL2* (suppressor of phospholipase C) that may be a CDK inhibitor homolog. This suppression of temperature sensitivity is not dependent on a functioning copy of the *PHO4* gene. The authors of this study note that a common factor in the suppression of temperature sensitivity is the inhibition of the Pho80/Pho85 kinase, though the role of *SPL2* in the Pho pathway is not known (Flick and Thorner, 1998).

Recent work has clarified the genetic components of the plasma membrane



phosphate transport system (Wykoff and O'Shea, 2001). There exist three genes that encode low-affinity plasma membrane phosphate transporters, *PHO87*, *PHO90*, and *PHO91*. Additionally, the *PHO84* gene encodes a high-affinity plasma membrane phosphate/proton symporter; the *PHO89* gene encodes a putative plasma membrane phosphate/sodium symporter; and the *GIT1* glycerophosphoinositol plasma membrane transporter is capable of transport of inorganic phosphate as well.

Phosphate is stored intracellularly in budding yeast in the vacuole in the form of polyphosphate (Katchman and Van Wazer, 1954). Though the biochemical identity of the polyphosphate synthetase is not clear in budding yeast, the four genes *PHM1*, *PHM2*, *PHM3*, and *PHM4* are required for normal vacuolar stores of polyphosphate (Ogawa et al., 2000). The *PHM* genes have been shown to be involved in vacuolar membrane fusion (Cohen et al., 1999); this function may be their primary biochemical activity, or it may be separable from the polyphosphate synthesis phenotype. The *PHM5* gene encodes an vacuolar endopolyphosphatase; in its absence, polyphosphate chains are not rapidly cleaved upon phosphate starvation (Ogawa et al., 2000).

In addition to the *PHO5* gene, a number of other genes are highly induced in a Pho4-dependent manner during phosphate starvation (Ogawa et al., 2000), or during ectopic activation of Pho4 by chemical inhibition of an inhibitor-sensitive allele of the Pho85 kinase (Carroll et al., 2001). These genes include three of the plasma membrane phosphate transporters, *PHO84*, *PHO89*, and *GIT1*. The low-affinity transporter genes *PHO87*, *PHO90*, and *PHO91* are not transcriptionally regulated during phosphate starvation (Auesukaree et al., 2003). The putative polyphosphate metabolism genes *PHM1*, *PHM2*, *PHM3*, *PHM4*, and *PHM5* are all induced substantially by phosphate

starvation. Also, the CDK inhibitor *PHO81* is itself induced by phosphate starvation, though its induction is not necessary for Pho4 activation (Ogawa et al., 1993). The *SPL2* gene is also highly induced during phosphate starvation. The *PHO5*, *PHO84*, and *SPL2* genes are the three genes that display the most Pho4-dependent transcriptional activation during phosphate starvation (Melissa R. Thomas, personal communication).

The Pho81/Pho80/Pho85 kinase is hypothesized to respond, whether directly or indirectly, to the intracellular levels of inorganic phosphate. This model is based on a number of independent observations. An alternative model had proposed that the product of the *PHO84* gene acts both as an extracellular phosphate sensor and a phosphate transporter. Though the deletion of the phosphate transporter gene *PHO84* results in the Pho^C phenotype, this phenotype is suppressed by overexpression of a number of other unrelated phosphate transporters (Wykoff and O'Shea, 2001). This suppression suggests that *PHO84* is not necessary for the signaling of high levels of phosphate. Secondly, the simultaneous deletion of the *PHO87*, *PHO90*, and *PHO91* genes results in an increase in *PHO5* expression relative to wild-type (Auesukaree et al., 2003), confirming that the presence of *PHO84* is not sufficient for the signaling of high extracellular phosphate. These observations are then consistent with a model in which any defect in phosphate uptake triggers a deficit in intracellular phosphate levels and therefore activation of the Pho pathway. Finally, further evidence for this model is found in the report that alterations in polyphosphate metabolism change the kinetic parameters of *PHO5* activation in response to phosphate starvation (Melissa R. Thomas, manuscript in press). These kinetic changes are consistent with vacuolar polyphosphate acting as a buffer of intracellular phosphate levels and preventing activation of the Pho

pathway until intracellular levels have dropped by a certain threshold. It remains formally possible that extracellular phosphate is signaling through the combined activities of the high-affinity transporter Pho84 and the low-affinity transporters, and that vacuolar phosphate levels are contributing a concurrent signal through the Phm proteins. However, we strongly advocate for the parsimonious intracellular sensor model (Fig. 4.1).

The toxin arsenate is a molecular analog of inorganic phosphate, and therefore it is not surprising that mutations in the Pho signal transduction pathway affect the ability of budding yeast to survive in a medium containing sodium arsenate. It was reported that any reduction in phosphate transport capability, either by deletion of the *PHO84* gene (Bun-Ya et al., 1991), or by deletion of the low-affinity phosphate transporter *PHO87*

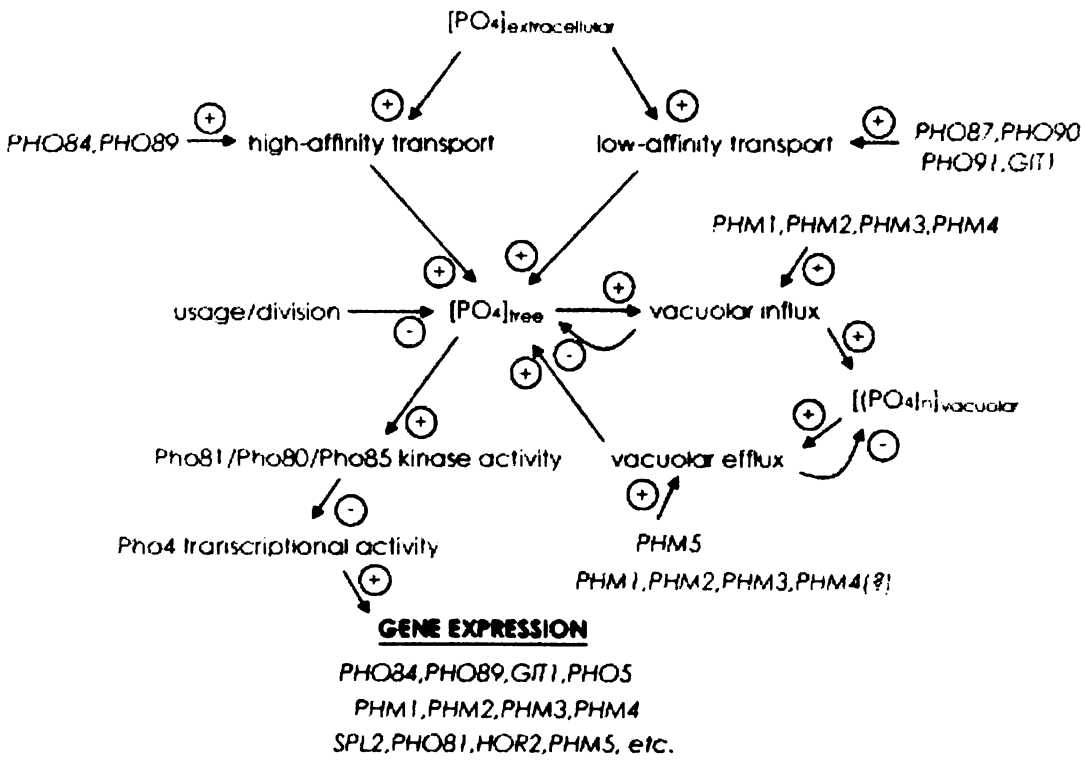


Figure 4.1. Qualitative relationships between phosphate signal transduction pathway in the intracellular phosphate sensor model. Plus signs include direct relationships and minus signs indicate inverse relationships.

(Bun-ya et al., 1996), results in an increased resistance to arsenate exposure (Ars^R). Surprisingly, it was discovered that the inactivation of Pho pathway signaling by mutation of either *PHO81* or *PHO4* in the *pho84Δ* background resulted in suppression of the arsenate resistance phenotype, or arsenate sensitivity (Ars^S). The Ars^R phenotype of the *pho84Δ* requires *PHO4* but not *PHO81* and does not require *PHO5* (Table 4.1 and Terry Schroyer, personal communication). The Ars^R phenotype correlates with poor growth on media formulations that contains low levels of inorganic phosphate, and the Ars^S phenotype correlates with the ability to grow well on low phosphate medium. These correlations support the assertion that the Ars^R phenotype is a cellular state of low phosphate uptake ability, and the Ars^S phenotype is a cellular state of increased phosphate uptake potential. Interestingly, this model suggests that the *pho84Δ* strain is deficient in phosphate uptake only due to the activity of the Pho pathway.

Recent work has also suggested that the level of *PHO84* expression is negligible in high phosphate medium (Raser and O'Shea, 2004). This observation presents a paradox: why does the *pho84Δ* strain display such a strong Pho^C phenotype and substantial defect in phosphate transport, if *PHO84* is expressed at very low levels? In light of the

Genotype	Ars phenotype	Pho phenotype	Low Pi growth
pho84Δ	Ars^R	Pho^C	0
pho84Δ pho81Δ	Ars^S	Pho^U	+++
pho84Δ pho81Δ pho80Δ	Ars^R	Pho^C	0
pho84Δ pho4Δ	Ars^S	Pho^U	+++
pho84Δ pho5Δ	Ars^R	Pho^C	0

Table 4.1. The arsenate (Ars), Pho5 acid phosphatase (Pho), and low phosphate medium growth phenotype of selected *pho84Δ* strains. Ars^R and Ars^S refer to resistance and sensitivity, respectively. Pho^C and Pho^U refer to constitutive and uninducible Pho5 acid phosphatase expression, respectively. +++ indicates robust growth, and 0 indicates poor growth.

accumulated data on the impact of phosphate transport on Pho pathway signaling, we reject the hypothesis that *PHO84* displays a catalytic signaling activity. Rather, we propose that the *pho84Δ* strain displays a defect in the inactivation of the Pho pathway due to a previously unrecognized positive feedback loop, in which the activity of the Pho pathway in the absence of the *PHO84* gene suppresses the cell's capability for phosphate transport. This positive feedback loop model implies that slight fluctuations in the intracellular phosphate level in the *pho84Δ* strain will be sufficient to induce the Pho pathway. Once induced, it is difficult to inactivate the pathway due to the suppression of phosphate transport.

Results

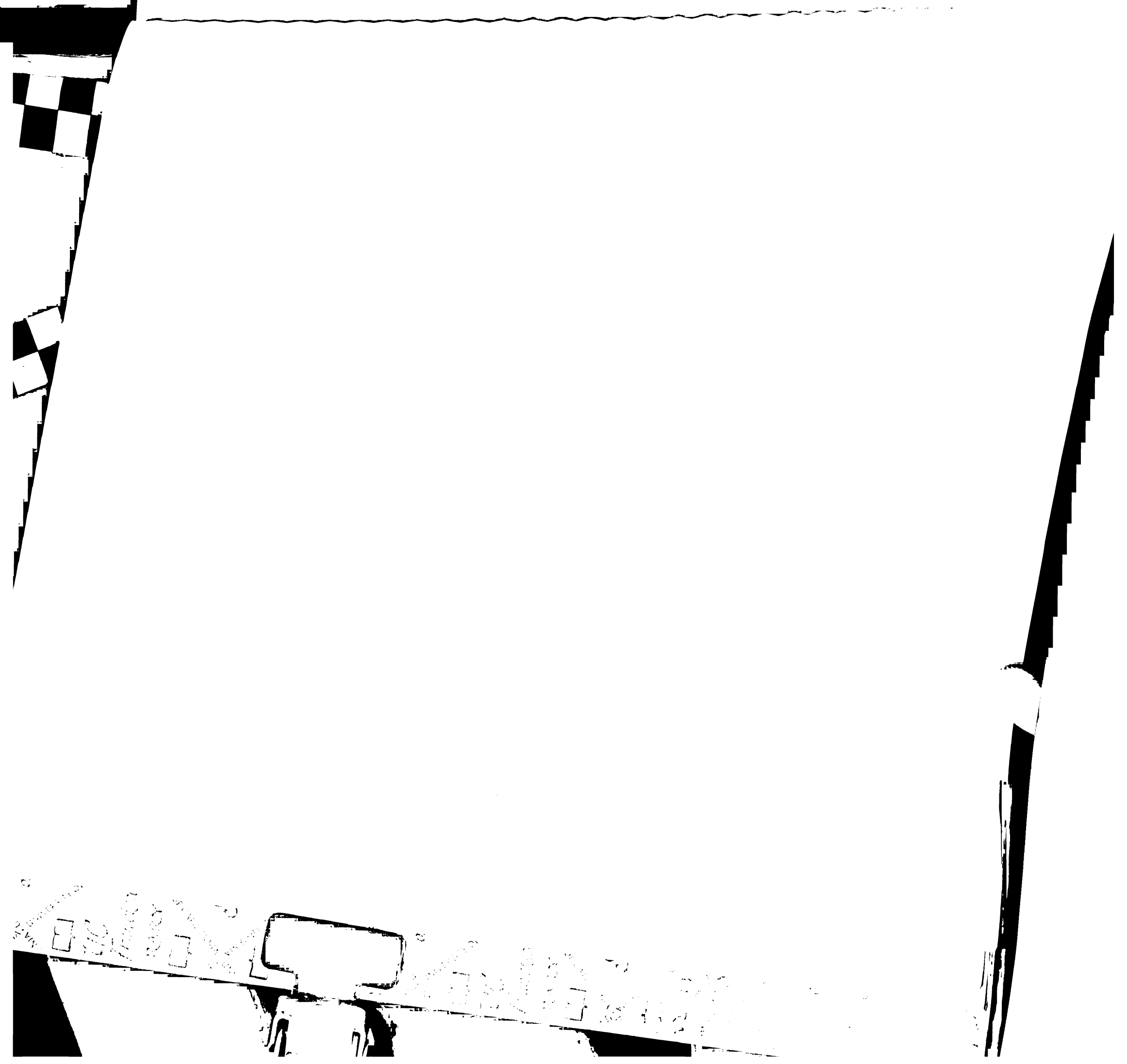
A positive feedback model of the Pho pathway

Based on the data on the *PHO4*-dependence of the Ars^R phenotype in the *pho84Δ* strain, we hypothesize that Pho4 activity suppresses a non-Pho84 component of the phosphate transport system. As Pho4 activates the transcription of the *PHO89* and *GIT1* plasma membrane phosphate transporter genes, it is reasonable to propose that the transporters whose activity is suppressed are one or more of the low-affinity phosphate transporters Pho87, Pho90, and Pho91. Active Pho4 is a nuclear-localized transcription factor, and therefore we further hypothesize that the Pho4-dependent inactivation of the low-affinity phosphate transporter occurs indirectly through the transcriptional regulation of one or more genes, rather than directly through an unrecognized non-transcription activity on the part of Pho4. Pho4 has only been shown to activate transcription, never inhibit, so we suggest that the gene in question is activated by Pho4

during phosphate starvation. Figure 4.2 details our positive feedback hypothesis explicitly.

Pho5 expression in the *pho84Δ* strain is suppressed by phosphate repletion

One possibility suggested by our positive feedback model was that the Pho^C phenotype of the *pho84Δ* strain was not a constitutive phenotype, but rather represented the increased sensitivity to phosphate starvation due to unrestrained positive feedback in the absence of *PHO84*. Pho5 expression has traditionally been assayed by two methods: a plate-based assay that reports the expression of a colony or patch of cells; and a liquid-based assay of a population of cells in suspension. Yeast colonies on plates are likely to undergo nutrient deprivation as cells deplete the local environment, whereas cells in suspension are more uniformly exposed to the liquid medium. Also, if cells are grown with care to the liquid culture density, the phosphate concentration in the liquid can be maintained at a known level, which is not possible on a plate. We have noticed in the past that the *pho84Δ* strain on a plate assay sometimes displays the "ring of white" plate assay phenotype (data not shown), in which the outermost cells in the colony, or the ring around the edge, appear white due to the lack of expression of Pho5. In contrast, the *pho80Δ* strain in which Pho4 is constitutively active does not display the ring of white plate assay phenotype. We therefore measured the Pho5 activity of the *pho84Δ* strain compared to a wild-type and *pho80Δ* strain in liquid culture in which great care was taken to maintain the cells in logarithmic growth and at a low density. With an A_{420}/A_{600} ratio of 1.61, the *pho84Δ* strain displays a low level of Pho5 induction in these conditions. The ratio was 0.89 for the uninduced wild-type cells and 4.96 for the



constitutively induced *pho80Δ* strain.

Pho4 nuclear localization in the *pho84Δ* strain is suppressed by phosphate repletion

If long periods of growth in high phosphate medium lower the production of Pho5 in the *pho84Δ* strain, incubation in high phosphate medium should also result in an increase in the Pho81/Pho80/Pho85 kinase activity and the inactivation of the transcription factor Pho4. As Pho4 is regulated by subcellular localization, we can measure the localization of a Pho4 green fluorescent protein (Pho4-GFP) fusion by epifluorescence microscopy as a proxy of kinase activity. Cytoplasmic localization of the

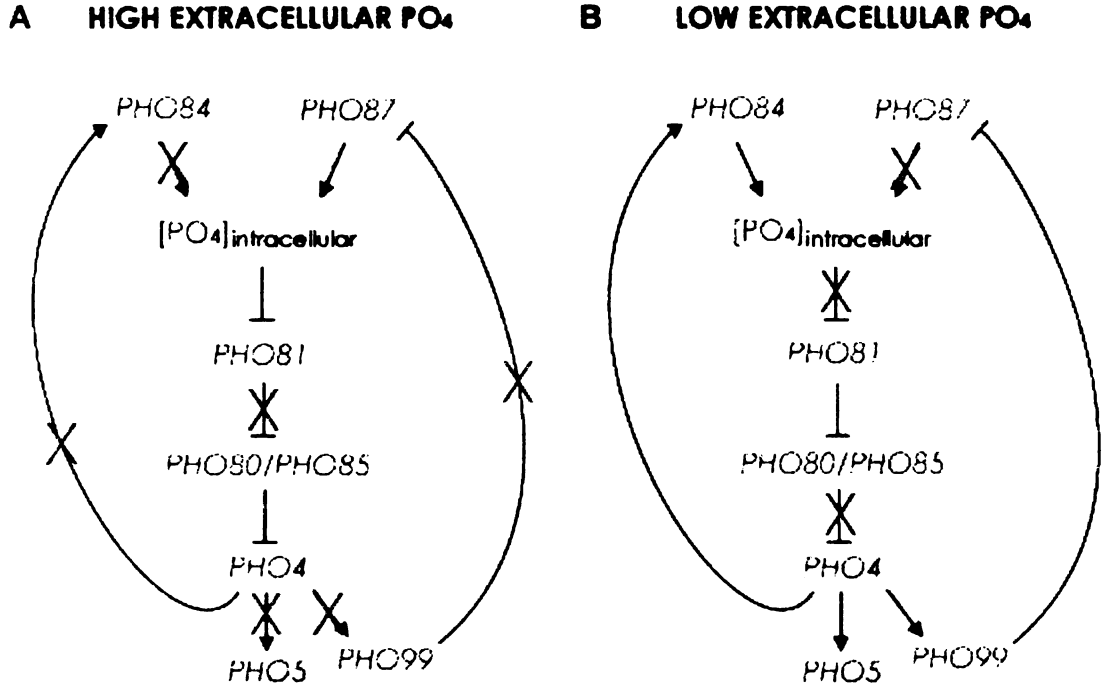


Figure 4.2. The simplified positive feedback model. Lines with pointed heads indicate activation, lines with flat heads indicate inhibition, and the red X indicates that a relationship is not active in the specified condition. *PHO99* indicates the hypothesized gene responsible for the inhibition of low-affinity phosphate transport. **(A)** In high extracellular phosphate, low-affinity Pho87 transport results in high levels of intracellular phosphate, which inhibit Pho81, allowing the Pho80/Pho85 kinase to inactivate Pho4, preventing gene expression from *PHO5*, *PHO84*, and *PHO99*. **(B)** In low extracellular phosphate, inefficient phosphate transport through Pho84 results in low intracellular phosphate levels which are unable to inhibit the CDK inhibitor Pho81. The active Pho81 inhibits the Pho80/Pho85 kinase, resulting in uninhibited Pho4. Active Pho4 results in the transcription of *PHO84*, *PHO5*, and *PHO99*. Active Pho99 inhibits the low-affinity transporter Pho87.

Pho4-GFP will indicate a higher kinase activity, whereas nuclear localization will be consistent with lower kinase activity. We therefore examined the localization of the Pho4-GFP gene product expressed from a plasmid in the wild-type yeast, in a strain lacking the gene *PHO80* whose gene product is required for kinase activity, and in a strain lacking the PHO84 gene (Fig. 4.3). We observed after incubation in high phosphate medium for 24 hours the expected cytoplasmic and nuclear localization of Pho4-GFP in the wild-type and *pho80Δ* strains, respectively. The Pho4-GFP in the *pho84Δ* strain is predominantly cytoplasmic as well, consistent with our hypothesis that extended growth in phosphate-replete medium is able to suppress the constitutive activation of Pho4 in the *pho84Δ* strain.

Our positive feedback model suggests that the *pho84Δ* strain will be especially sensitive to small fluctuations in environment phosphate levels. Specifically, we predict that Pho4-GFP in the *pho84Δ* strain will become nuclear more rapidly than in the wild-type strain during a short exposure to phosphate-limited medium. Additionally, the positive feedback may result in the retention of Pho4-GFP in the nucleus upon subsequent exposure to phosphate-replete conditions. We were able to observe both these phenotypes (Fig. 4.3).

Pho4 nuclear localization in the *pho84Δ* strain is suppressed by inactivation of *PHO4*

An alternative hypothesis that does not require a positive feedback can explain all of the observations concerning Pho5 expression and Pho4-GFP localization. It is plausible to suggest that the *pho84Δ* strain is chronically starved for intracellular phosphate due to a low uptake capability. The suppression of Pho4 nuclear localization by extended periods of exposure to phosphate-replete medium is due to the eventual

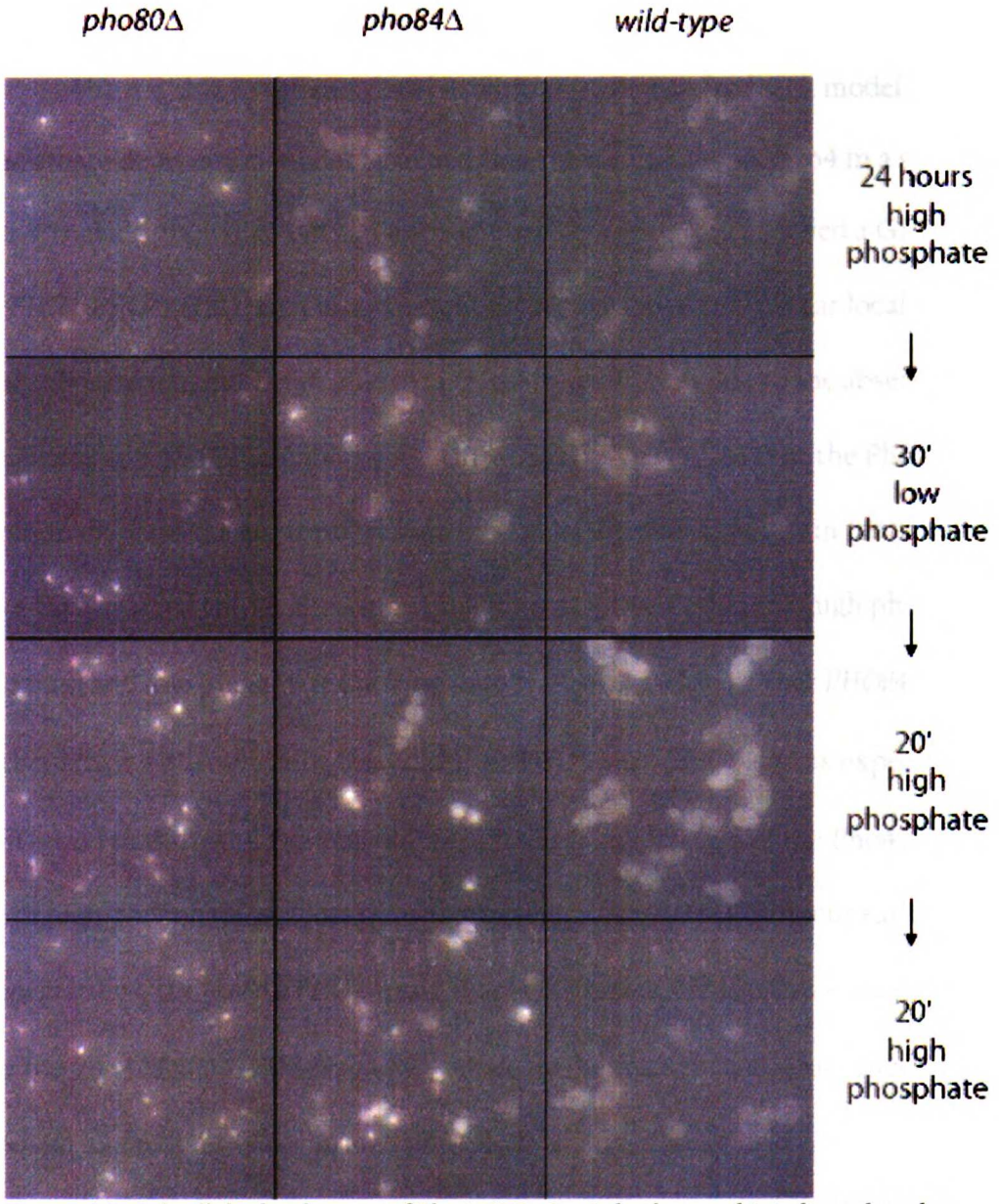
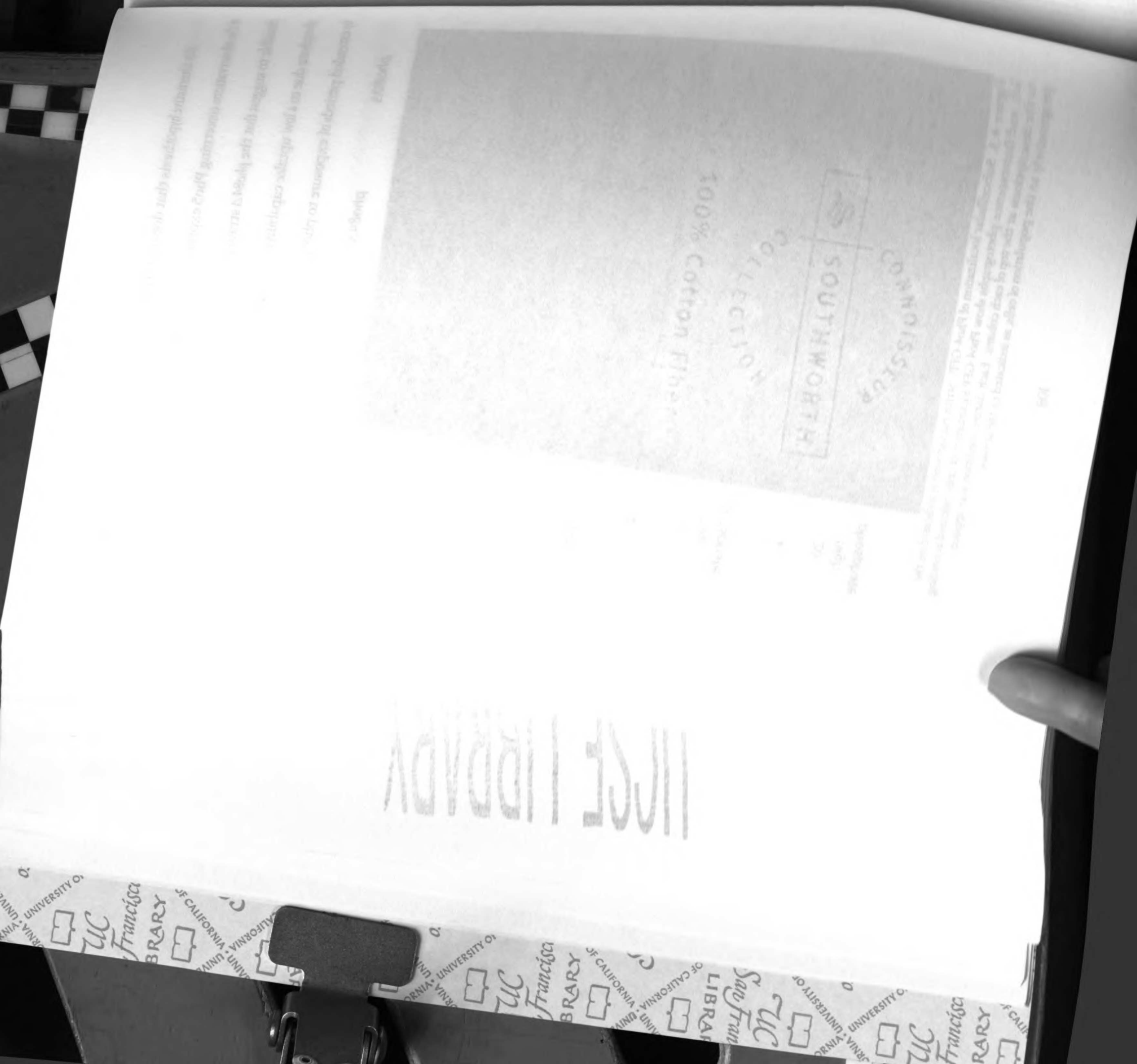


Figure 4.3. Subcellular localization of Pho4-GFP. Strain background is indicated on the left, and fluorescence photographs show Pho4-GFP localization of cells cultured according to the descriptions at the top of each column. Each culture condition was applied progressively to the population of cells, as indicated by the arrows.


CONNOISSEUR
SOUTHWORTH
COLLECTION
100% Cotton Fiber

repletion of intracellular phosphate levels. Since phosphate uptake is diminished, it is not surprising that the *pho84Δ* strain would display an increased sensitivity to short periods of starvation, nor that Pho4-GFP in the *pho84Δ* strain would remain nuclear for much longer. In this model, a chronic deficit in phosphate transport capability due to the absence of the *PHO84* gene results in a strain that is generally always more starved for phosphate than the wild-type strain.

We are able to differentiate between our positive feedback model and the chronic phosphate deprivation model by examining the localization of Pho4 in a strain that lacks Pho4 transcriptional activity. To accomplish this goal, we employed a GFP-tagged allele of *PHO4*, *PHO4ΔDBD*, that is competent for regulation of subcellular localization by phosphorylation, but is not able to activate transcription due to the absence of the sequence-specific DNA-binding domain (DBD). We verified that the Pho4ΔDBD-GFP protein displays the appropriate localization in the wild-type strain (data not shown). We then assayed the localization of the Pho4ΔDBD-GFP in both high phosphate medium and low phosphate medium in four strains, lacking either *PHO84* or *PHO80*, and either lacking or containing the wild-type *PHO4* gene (Fig. 4.4). As expected, the *pho80Δ PHO4* and *pho80Δ pho4Δ* strains display a nuclear localization of the Pho4ΔDBD-GFP in both high phosphate and low phosphate media. Consistent with our earlier observations, the *pho84Δ PHO4* strain displays a mixed nuclear/cytoplasmic localization for the Pho4ΔDBD-GFP in high phosphate medium and a nuclear localization in low phosphate medium. The chronic phosphate deprivation model predicts that the Pho4ΔDBD-GFP will behave equivalently in both the *pho84Δ PHO4* strain and the *pho84Δ pho4Δ* strain. In contrast, our positive feedback model mandates that the

Pho4ΔDBD-GFP will be cytoplasmic in the *pho84Δ pho4Δ* strain. We observed this phenotype and assert that the positive feedback model is consistent with these measurements.

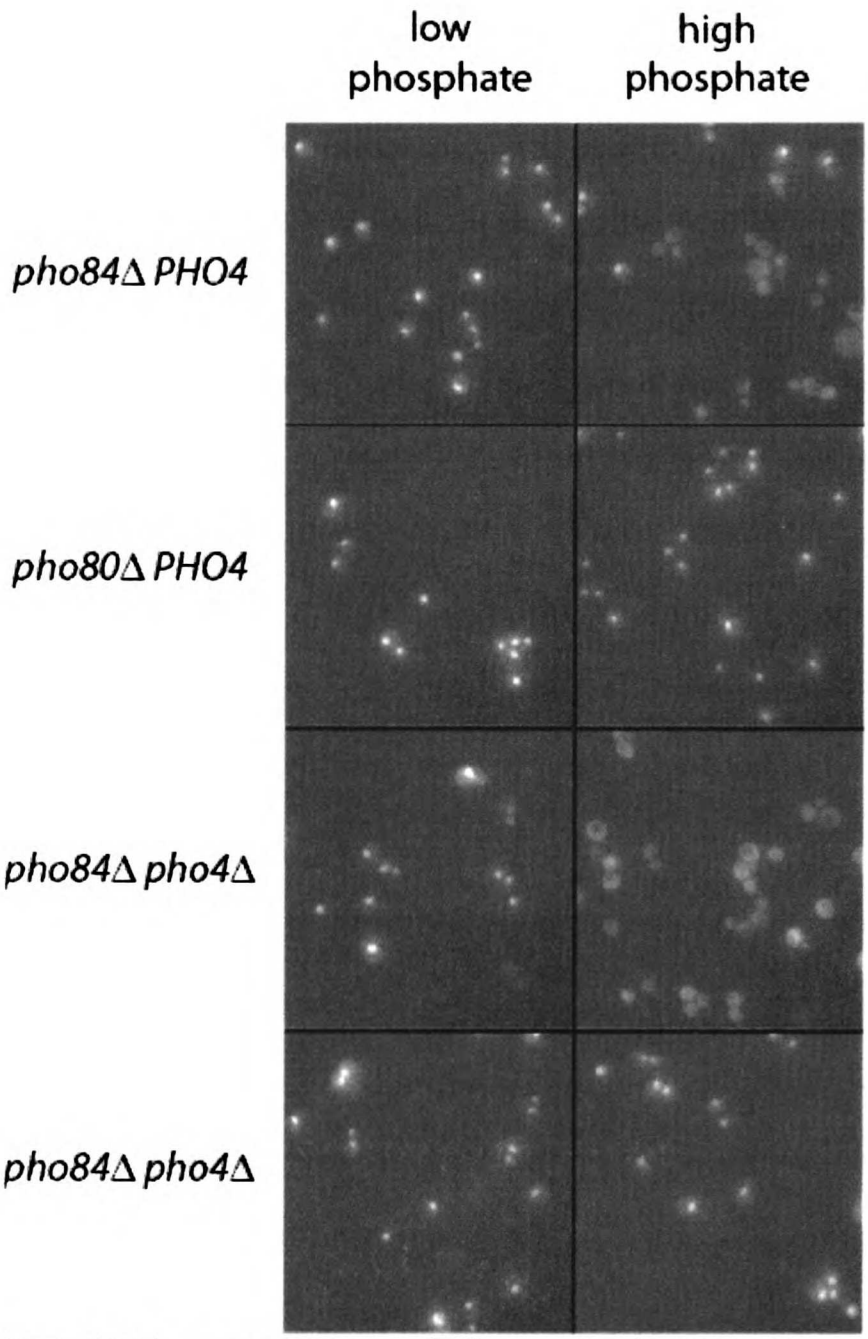


Figure 4.4. Fluorescence microscopy of cells expressing Pho4ΔDBD-GFP. The medium is indicated at the top of each column, and the genetic background of the cells is indicated to the left of each row.


% Cotton Fiber
LIBRARY
UC San Francisco
UNIVERSITY OF CALIFORNIA

A systematic deletion screen of the Pho4-dependent genes

Our model proposes that there exist one or more genes whose activation in low phosphate medium is Pho4-dependent and results in the inhibition of low-affinity phosphate transport. Using previously published genome-wide surveys of gene expression (Carroll et al., 2001; Ogawa et al., 2000) and unpublished data (Melissa R. Thomas, personal communication), we compiled a list of the genes whose activation in low phosphate medium is both substantial (>2-fold) and highly Pho4-dependent. Specifically, our list consisted of *PHO5*, *PHO8*, *PHM1*, *PHM2*, *PHM3*, *PHM4*, *PHM5*, *PHM6*, *PHM7*, *HOR2*, *SDT1*, *SPL2*, *PHO11*, *PHO12*, *PHM8*, and *ATG22*. We systematically deleted each of these genes in the *pho84Δ* strain, and tested the ability of each double deletion strain to grow on low phosphate medium (data not shown). The deletion of only one gene, *SPL2*, coincided with the suppression of the growth defect of the *pho84Δ* strain on low phosphate medium (data not shown). Tetrad analysis confirmed that this phenotype segregated 2:2 with the genetic marker linked with the deletion (data not shown).

Pho5 expression in the *pho84Δ* strain is suppressed by inactivation of *SPL2*

We wished to test if the deletion of *SPL2* in the *pho84Δ* strain is capable of suppressing the PhoC phenotype of the *pho84Δ* strain. We expect that if *SPL2* is required for the inhibition of the low-affinity transporters, the *pho84Δ spl2Δ* strain will not express the Pho5 acid phosphatase in high phosphate medium, but will express Pho5 in low phosphate conditions. We verified that the *pho84Δ spl2Δ* strain displays the expected phenotype (Fig. 4.5).

The arsenate resistance of the *pho84Δ* strain is suppressed by inactivation of *SPL2*

We expect that the ability of the *pho84Δ spl2Δ* strain to grow on low phosphate medium will be accompanied by an increased sensitivity to arsenate, similar to the *pho84Δ pho4Δ* strain. We simultaneously compared the growth of the *pho84Δ spl2Δ* strain, the *pho84Δ pho4Δ* strain, and the *pho84Δ* strain on low phosphate medium, on arsenate medium, on high phosphate medium, and measured the expression of Pho5 acid phosphatase activity to confirm that the deletion of the *SPL2* gene confers the expected phenotypes (Fig. 4.6).

Pho signaling in the *pho84Δ* strain is suppressed by inactivation of *SPL2*

We expect that Pho4 localization in the *pho84Δ spl2Δ* strain will be cytoplasmic in high phosphate medium, consistent with the positive feedback model that also explains the effect of the deletion of *SPL2* on growth in low phosphate medium, growth in

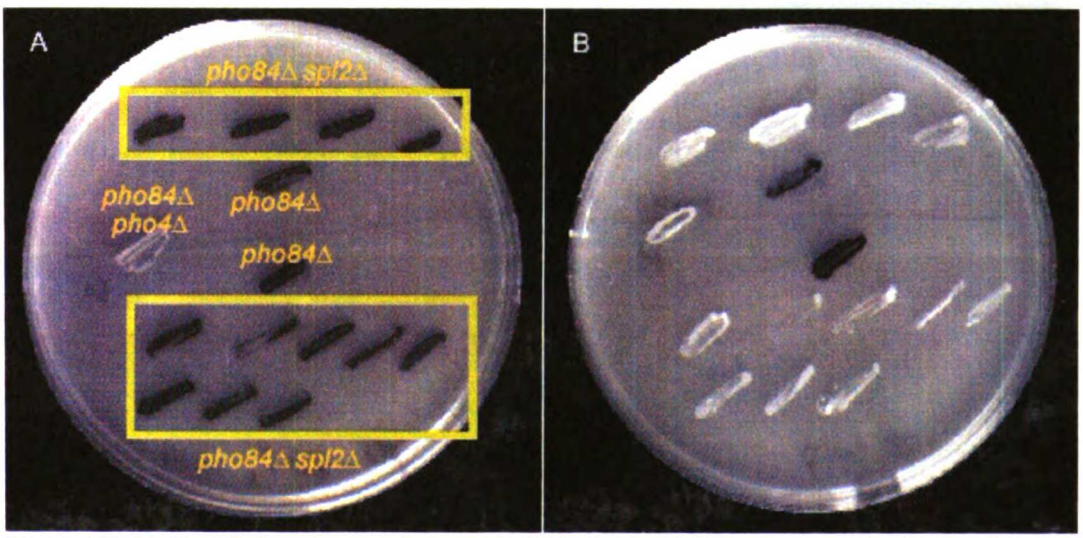


Figure 4.5. Pho5 acid phosphatase plate assay of the *pho84Δ spl2Δ* strain. Strain background for both panels is indicated in yellow on the left. A dark color indicates the presence of Pho5 activity. **(A)** Cells grown on low phosphate medium. **(B)** Cells grown on high phosphate medium.

arsenate medium, and the Pho5 acid phosphatase plate assay in high phosphate medium. We therefore employed a chromosomally integrated copy of Pho4 yellow fluorescent protein fusion (Pho4-YFP) to assay the localization of Pho4 in high phosphate in the *pho84Δ* strain, the *pho84Δ pho2Δ* strain, and the *pho84Δ spl2Δ* strain (Fig. 4.7). Additionally, we examined the localization of a *pho84Δ* strain in which the wild-type *PHO4* allele has been replaced by a chromosomal *PHO4ΔDBD-YFP* variant. These strains should express near-native levels of Pho4 in contrast to the strains used for the fluorescence microscopy described above, which employed plasmid-based fluorescent *PHO4* fusion genes. As expected, we observed that deletion of either *SPL2*, the DNA-binding domain sequence of *PHO4*, or *PHO2* results in cytoplasmic localization of the Pho4-YFP protein encoded by a single gene copy in the *pho84Δ* strain.

A genetic selection for mutations that suppress the *pho84Δ* growth defect

To identify the genes involved in the inhibition of the low-affinity transporters in

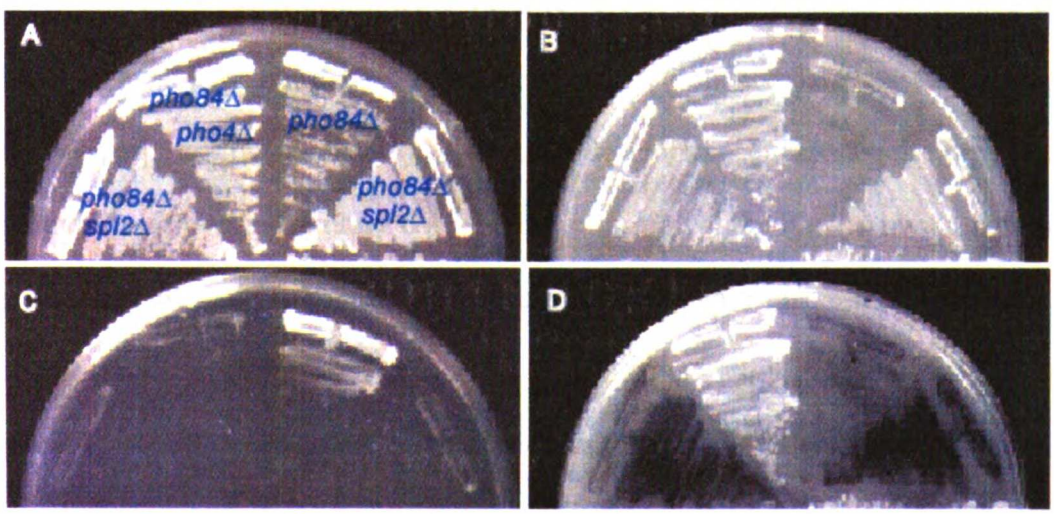
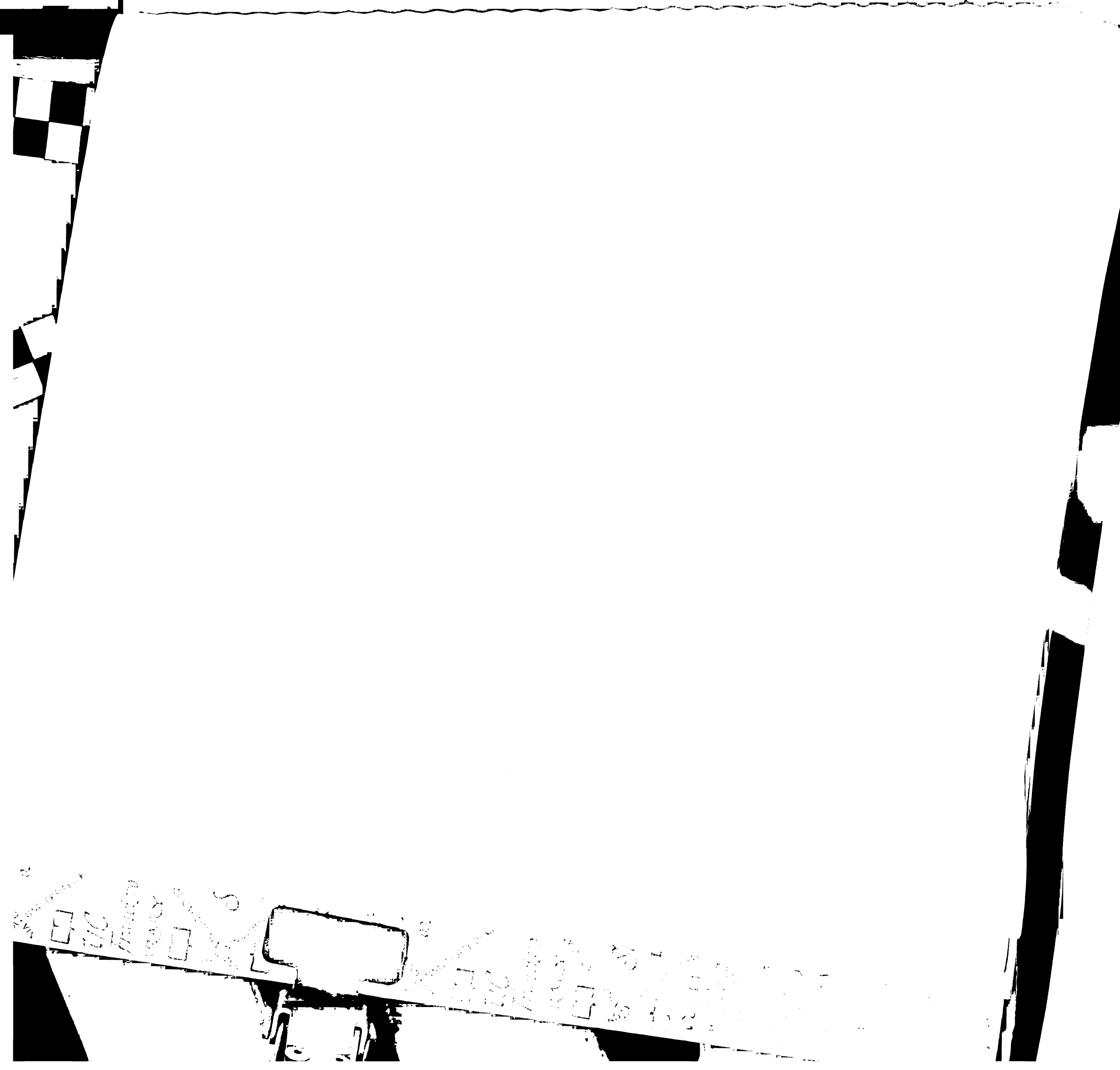


Figure 4.6. Phenotypes of the *pho84Δ*, *pho84Δ spl2Δ*, and *pho84Δ pho4Δ* strains. Strain identities are indicated in blue on the upper left panel. **(A)** Growth on high phosphate medium. (A) Growth on low phosphate medium. **(C)** Growth on 12 mM arsenate medium. **(D)** Pho5 acid phosphatase plate assay on low phosphate medium. A dark color indicates the presence of Pho5 activity.



the *pho84Δ* strain, we conducted a genetic screen concurrently with the systematic deletion of the known Pho4-dependent genes described above. The *pho84Δ* strain grows poorly on plates containing low phosphate medium but acquires spontaneous suppressor mutations that allow for rapid growth. Previously, it was determined that the majority of such mutations consist of loss of function mutations in the *PHO4*, *PHO2*, and *PHO81* genes (D.D.W., unpublished results). These mutations confer a Pho-uninducible plate assay phenotype. We wished to isolate a distinct category of suppressor mutations: mutations that suppressed the poor growth phenotype of the *pho84Δ* strain while still allowing Pho5 induction by the plate assay. These mutations must be downstream of *PHO81*, *PHO4*, and *PHO2*, and therefore should represent mutations in the genes specific to the inhibition of low-affinity phosphate transporter. To add to the stringency of our selection, we created a strain that grows more poorly than the *pho84Δ* strain on low phosphate medium. This strain lacks all three Pho4-dependent plasma membrane phosphate transporters genes, *PHO84*, *PHO89*, and *GIT1*. This triple deletion strain grows extremely poorly on medium containing low levels of phosphate, but demonstrates a dramatic level of arsenate resistance compared to the wild-type strain (see below). This strain develops suppressors quite rapidly due to the extreme growth advantage conferred by mutation of the Pho pathway, and must be maintained on medium containing arsenate in order to prevent overgrowth of the original strain by suppressed derivatives. We grew the triple deletion strain overnight in intermediate phosphate medium to allow spontaneous mutations to occur and transferred the cells to low phosphate medium plates. Colonies formed on the low phosphate medium plates after several days.

A secondary screen eliminates mutations in the Pho pathway

To assess whether we were successful in isolating the two types of mutant strains we expected, we performed the Pho5 acid phosphatase plate assay on the colonies that grew on the low phosphate medium plates (Fig. 4.8). As anticipated, two types of colonies were readily identified: red colonies that express Pho5 and white colonies that do not. However, the plate assay is toxic to the colonies, and we therefore isolated a number of the suppressor strains and subsequently assayed their Pho5 activity by plate assay on both low phosphate medium plates and high phosphate medium plates (Fig. 4.9). Three types of colonies were apparent: colonies that are white on both types

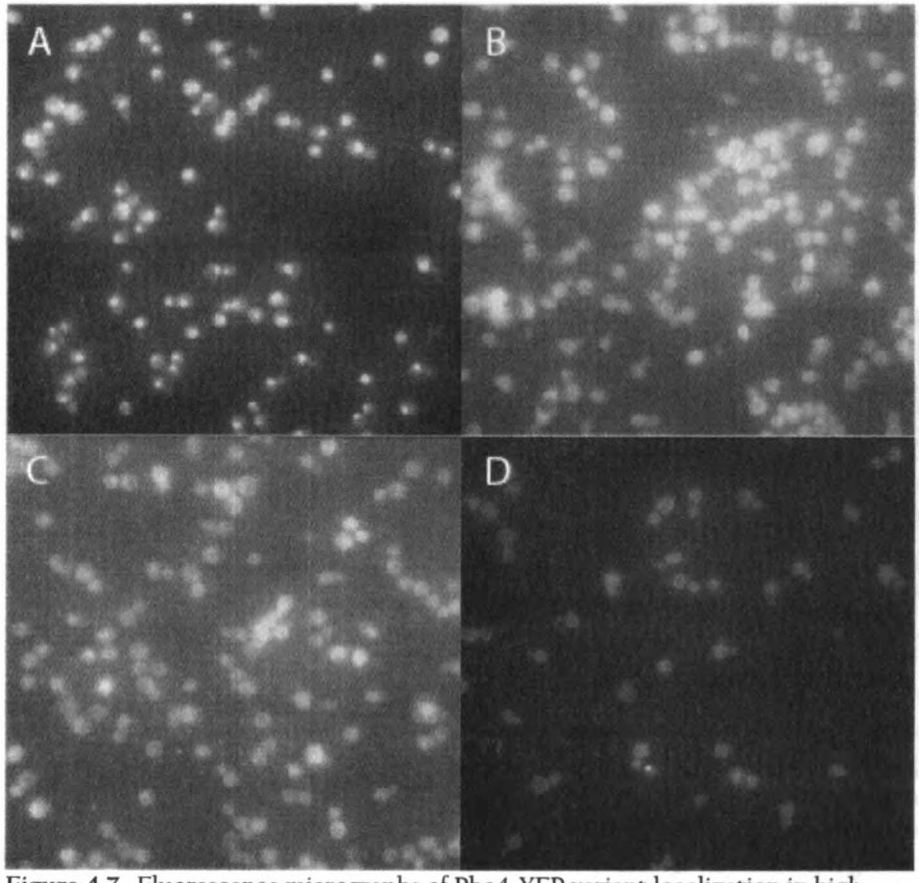
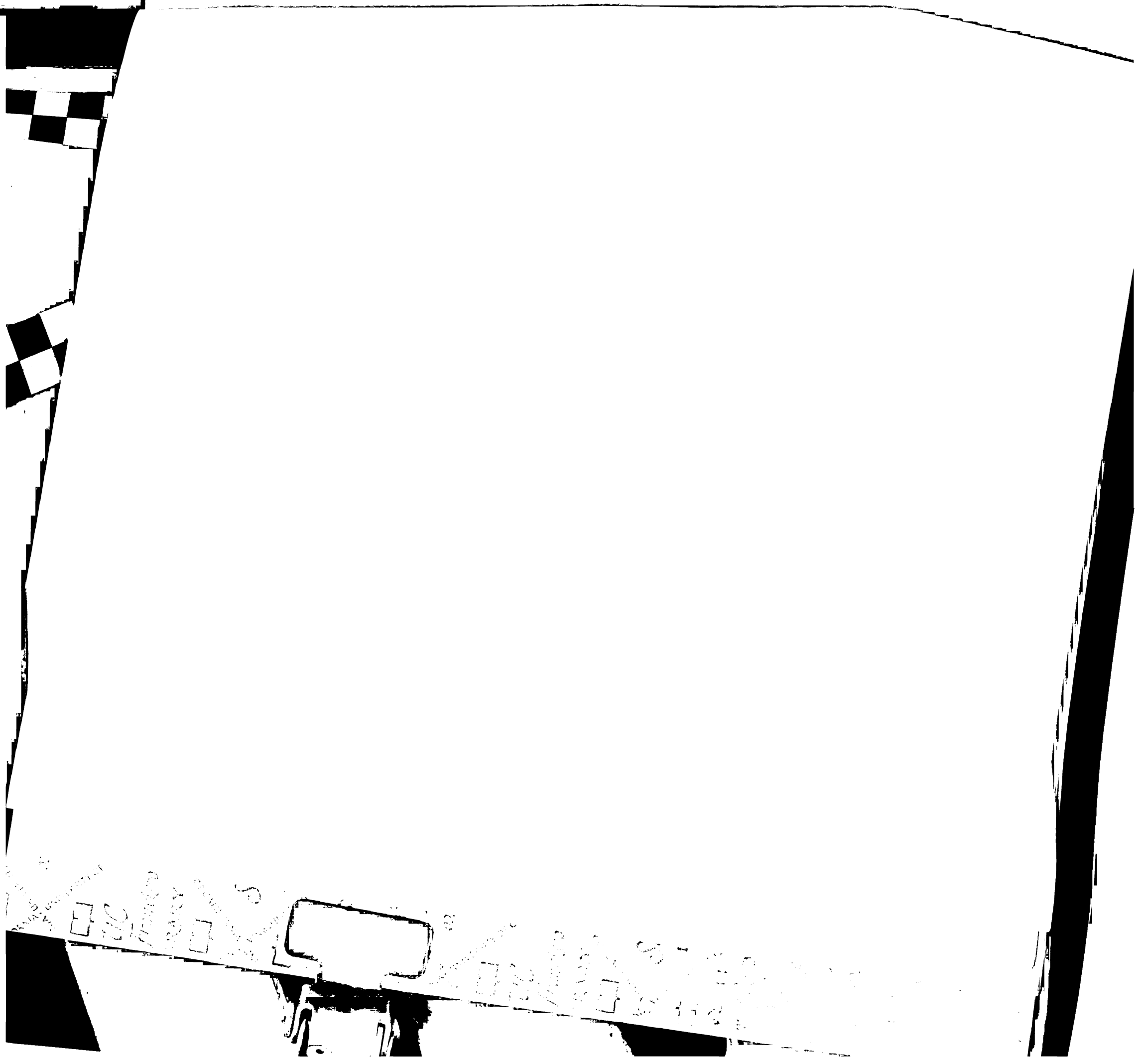


Figure 4.7. Fluorescence micrographs of Pho4-YFP variant localization in high phosphate growth medium. **(A)** Pho4-YFP in the *pho84Δ* strain. **(B)** Pho4-YFP in the *pho84Δ spl2Δ* strain. **(C)** Pho4ΔDBD-YFP in the *pho84Δ* strain. **(D)** Pho4-YFP in the *pho84Δ pho2Δ* strain.



of media, colonies that are white in high phosphate conditions but red in low phosphate conditions, and colonies that are red on both types of media. We assumed that the constitutively white colonies represent mutations in upstream components of the Pho pathway and therefore discard these strains. We assume that all those colonies that are red on low phosphate medium plates have acquired mutations that prevent the downregulation of the low-affinity phosphate transport system. We suspect that those strains that are red in high phosphate conditions represent weaker suppression than those that are white in high phosphate conditions.

We chose the strains that are red in low phosphate medium and white in high phosphate medium for further study. We named these strains *jd1* (for Jonathan and Dennis mutants) to *jd45*. We preserved two mutant strains, *jd100* and *jd101*, that are white on both low and high phosphate conditions for comparative purposes.

The *jd* mutations confer arsenate sensitivity

We expect that mutations in the *pho84Δ pho89Δ git1Δ* strain that confer the ability to grow on low phosphate medium will also confer arsenate sensitivity due to increased arsenate transport capability. We simultaneously measured the growth on low phosphate medium, the growth on high phosphate medium, the expression of Pho5 on low phosphate medium, and the arsenate sensitivity of the *jd* mutant

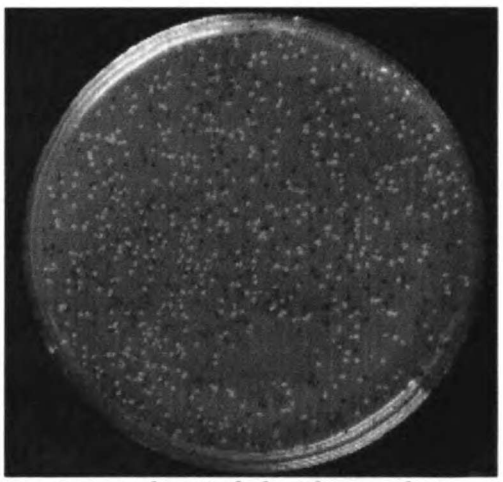


Figure 4.8. Pho5 acid phosphatase plate assay performed on colonies of the *pho84Δ pho89Δ git1Δ* strain that grew on low phosphate medium plates. Dark colonies express Pho5 and therefore have functional *PHO81*, *PHO4*, and *PHO2* genes; the light colonies do not express Pho5.

strains (Fig. 4.10). We found that all of the *jd* strains are arsenate sensitive regardless of their ability to induce the Pho5 acid phosphatase on low phosphate medium. In the context of our hypothesis that any increase in phosphate transport is accompanied by an increase in sensitivity to the toxic phosphate analog arsenate, these independent observations provide further evidence that the *jd* strains contain mutations that increase the cellular phosphate uptake capability.

Testing the *jd* mutations for dominance

It is plausible that the *jd* mutant strains will contain both recessive and dominant mutations. For example, loss of function mutations in the low-affinity transport inhibitor gene hypothesized to be induced by Pho4 would be likely to be recessive, and gain of function mutations in the low-affinity transporters that prevent downregulation are likely to act in a dominant fashion. To assess the dominant or recessive status of the *jd* mutations, we mated the *jd* strains to a *pho84Δ* strain. The resulting diploids should contain homozygous deletions of the *PHO84* locus but be heterozygous for each individual *jd* mutation.

We measured the arsenate sensitivity of the *pho84Δ/pho84Δ jd/+* **diploid** strains. We expect that if the *jd* allele is dominant, the strain will be arsenate sensitive. If the *jd* mutant allele is recessive, the diploid should be arsenate resistant. All of the *pho84Δ/pho84Δ jd/+*diploid strains are arsenate resistant (data not shown), indicating that all of the *jd* mutations are recessive. Future work will test which if any of the *jd* mutations have occurred in the *SPL2* gene, and will identify complementation groups for those *jd* mutations that are not allelic with *SPL2*. It will be possible to use the arsenate

sensitivity of these *jd* mutant strains to select for complementation by a plasmid genomic library, to identify the *jd* genomic loci rapidly. If all the *jd* mutations are allelic with *SPL2*, an additional mutagenesis and selection could be performed. To identify other genes that function downstream of *SPL2* in the inhibition of the low-affinity phosphate transporters, it is possible to repeat the selection on low phosphate medium with a mutagenized strain containing multiple plasmid-born copies of the *SPL2* gene.

Other Pho^C strains are not suppressed by loss of function of *PHO4* or *SPL2*

Deletion of the *PHO84* gene is not the only genetic modification that confers a Pho^C phenotype. We wished to divide the other mutations that confer the PhoC phenotype into two categories: Pho^C mutations that are suppressed by deletion of either *SPL2* or the sequence encoding the DNA-binding domain of Pho4, and Pho^C mutations, that are not suppressed by deletion of either *SPL2* or the sequence encoding the DNA-

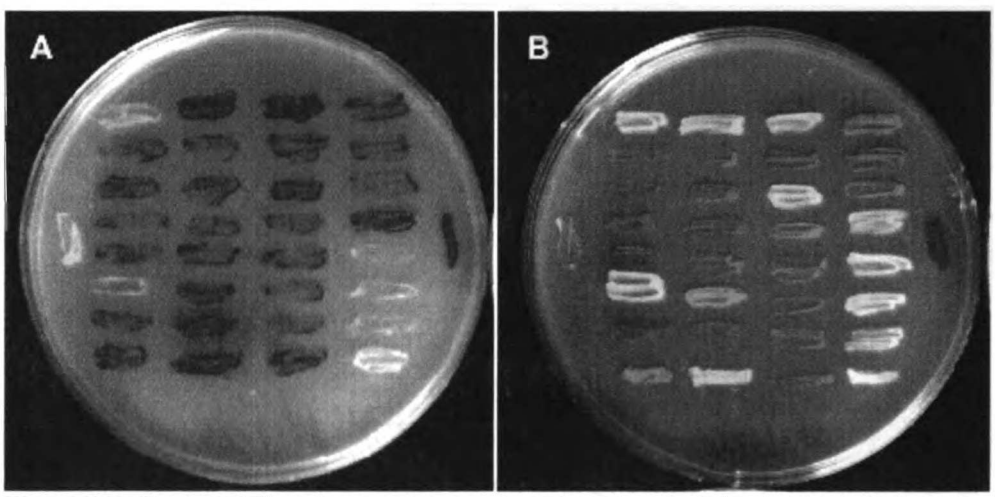


Figure 4.9. Pho4 acid phosphatase plate assay on the colonies of the *pho84Δ pho89Δ git1Δ* strain that grew on low phosphate medium plates. (A) Plate assay on a low phosphate medium plate. (B) Plate assay of the same strains as in (A) grown on a high phosphate medium plate. For both plates, on the far left is the *pho84Δ pho4Δ* strain, and on the right is the parent *pho84Δ pho89Δ git1Δ* strain.

binding domain of Pho4. We believe that the former set of mutations are selectively defective in high-affinity phosphate transport, resulting in a Pho^C phenotype due to feedback-based suppression of low-affinity transport. The latter class of mutations should contain mutations in the intracellular signaling pathway itself; mutations that selectively affect the low-affinity transporters; and mutations that affect both high- and low-affinity phosphate transport.

We therefore examined the localization of Pho4-YFP and Pho4DDBD-YFP in

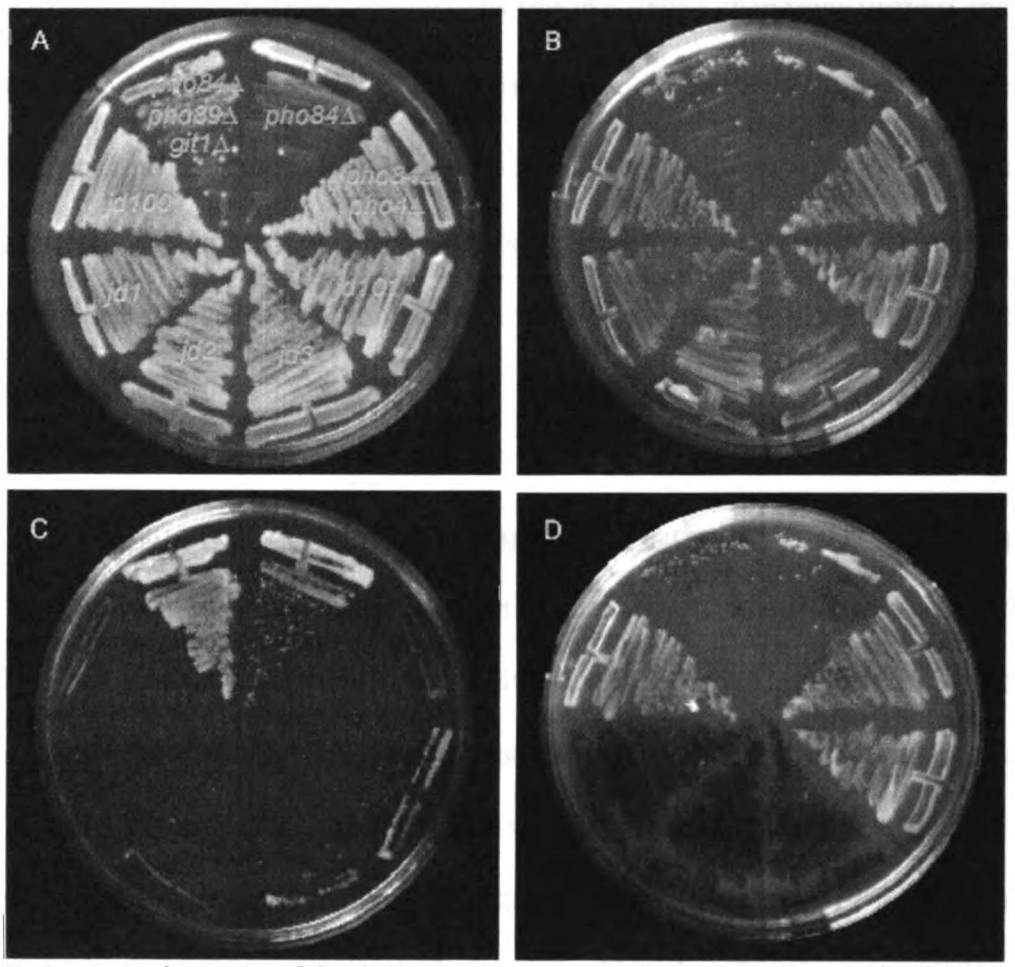


Figure 4.10. Phenotypes of the *jd* mutant strains compared to the parent strain, the *pho84Δ* strain, and the *pho84Δ pho4Δ* strain. The strain genotypes are labeled in yellow in panel A. **(A)** Growth of the strains on medium containing high levels of phosphate. **(B)** Growth of the strains on medium containing low levels of phosphate. **(C)** Growth of the strains on medium containing 12 mM arsenate. **(D)** Pho4 acid phosphatase plate assay of the strains grown on media containing low phosphate. Dark coloring indicates the presence of Pho5 activity.

cells grown in high phosphate medium for four Pho^C strains, the *pho80Δ* strain, the *pho85Δ* strain, the *PHO81*^{C} strain, and the *ado1Δ* strain (data not shown). Preliminarily, it appears that all four strains display nuclear Pho4-YFP in high phosphate medium, nuclear Pho4ΔDBD-YFP localization in high phosphate medium, and nuclear Pho4-YFP in high phosphate medium when the *SPL2* gene has been deleted. Therefore none of these mutations represent selective defects in high-affinity phosphate transport. It will be interesting to confirm these observations, and then conduct the same experiments on additional strains: the *pho86Δ* strain defective in a Pho84 chaperone protein, the *acc1* strain defective in lipid synthesis, the *pma1* strain defective in plasma membrane proton transport, and the *pho87Δpho90Δpho91Δ* strain defective in low-affinity phosphate transport. We expect that the strain lacking low-affinity phosphate transport will fall into the second class of mutations, but it is unclear to which class the other three mutations belong.

Pho4 localization as a quantitative measure of Pho pathway signal transduction

As Pho4 localization varies with the Pho81/Pho80/Pho85 kinase activity, it is likely that smaller changes in kinase activity are reflected in less dramatic localization changes. We proposed to quantify the localization of Pho4 using a technique developed previously (Melissa R. Thomas and James V. Falvo, personal communication). This technique relies on simultaneous visualization of both the Pho4-YFP protein and the nucleus itself, by detection of a nuclear protein, Hrp1, fused to the monomeric red fluorescent protein mRFP (Campbell et al., 2002). We tested this quantitative localization technique in several strains: wild-type cells, cells lacking all three

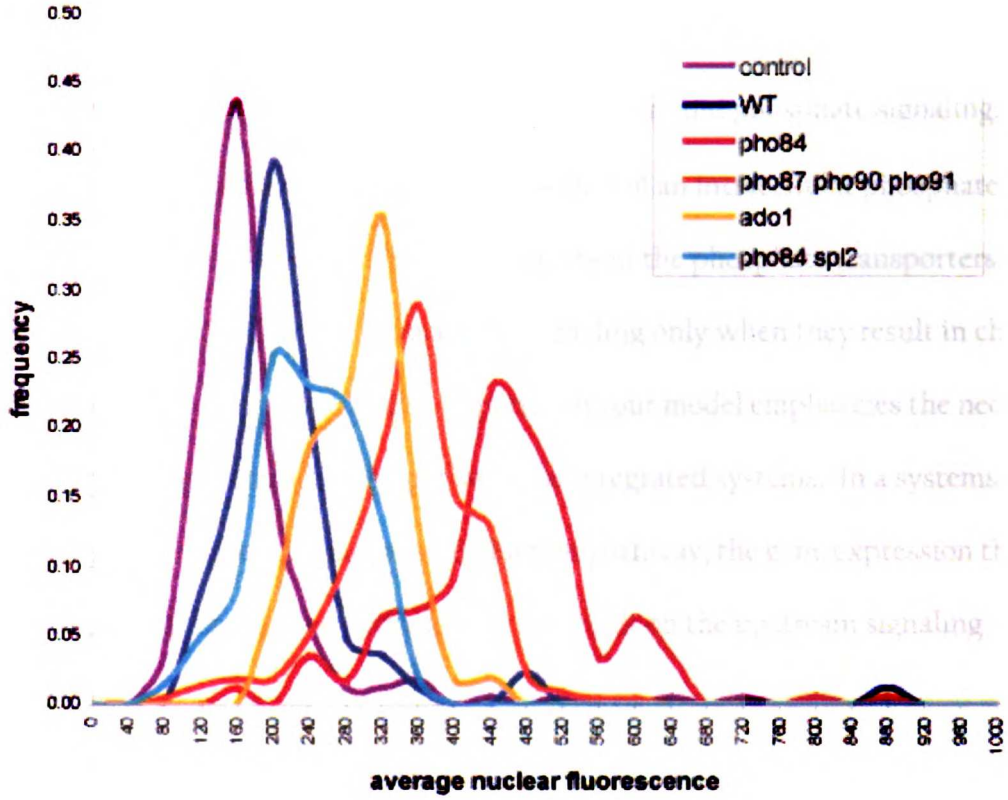


Figure 4.11. Quantitation of Pho4-YFP localization. A histogram of individual cellular nuclear Pho4-YFP fluorescence in the strains indicated. The frequencies have been normalized for to the total population measured.

low-affinity transporters, cells lacking Pho84, and cells lacking both Pho84 and Spl2 (Fig. 4.11). It is clear that the technique is sensitive to differences in Pho4 localization than the standard cytoplasmic or nuclear classification. It will be interesting to apply this technique to quantify phosphate signal transduction over time courses and at different levels of extracellular phosphate.

Discussion

The phosphate signal transduction pathway in budding yeast has been studied for decades as a model for signal transduction and, in particular, the activation of gene expression in response to environmental conditions. We have proposed a novel model to

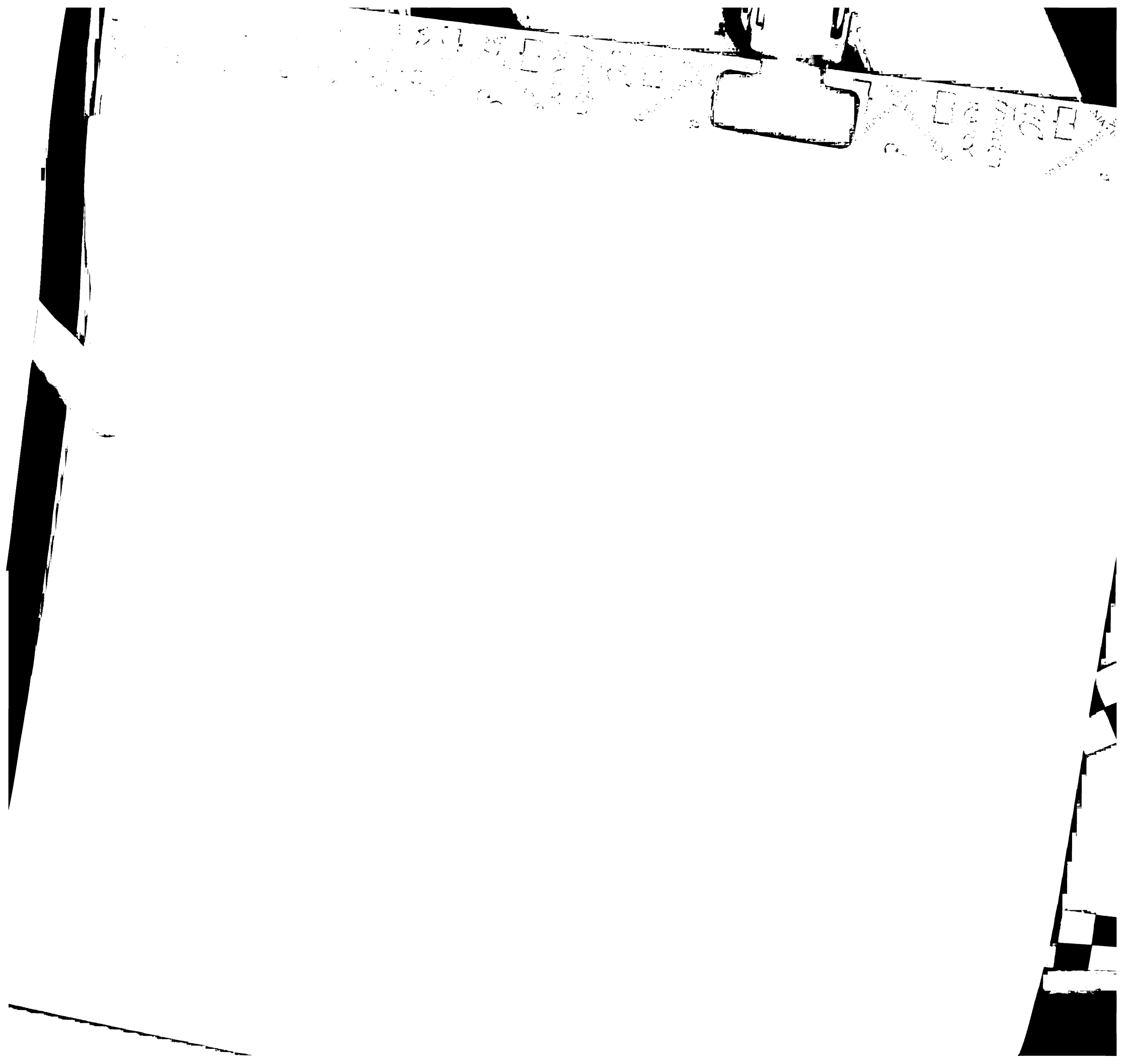
explain the phenotype of cells lacking the *PHO84* gene, encoding a high-affinity phosphate transporter. Our model accounts for several puzzling observations concerning the role of *PHO84* in phosphate transport and phosphate signaling. The model provides further support for the existence of an intracellular phosphate sensor that is only indirectly affected by perturbations in the phosphate transporters. These perturbations will affect the intracellular signaling only when they result in changes in the intracellular phosphate levels. Importantly, our model emphasizes the necessity of considering signal transduction pathways as integrated systems. In a systems view of the phosphate-responsive signal transduction pathway, the gene expression that results from signaling through the pathway can feed back on the upstream signaling components.

The positive feedback model predicted that there existed a specific Pho4-dependent gene with a novel function. We were able to screen the Pho4-dependent genes and identify at least one gene whose function is consistent with the model's prediction. The discovery of a novel and quite significant function of a gene induced by the Pho pathway emphasizes the importance of understanding the role of each individual gene induced by the Pho pathway during phosphate starvation. It is likely that every gene activated in a Pho4-dependent manner plays some role in the phosphate starvation response. The functions of a number of genes, such as *PHM6* and *PHM7*, remain unclear, and more experimentation will be required to identify the functional consequences of the induction of these genes..

While we were able to identify *SPL2* as a gene required for the inhibition of the low-affinity transport system during phosphate starvation, we have still not provided

formal proof that *SPL2* represents the positive feedback element. It remains to be demonstrated that the low-affinity transporters are inhibited in a manner dependent on the Pho4-dependent induction of *SPL2* gene expression. For instance, it is possible that small amounts of Spl2 are required to facilitate the inhibitory process, but another gene confers the Pho4-dependence on the process. Simple *SPL2* overexpression experiments in a strain lacking *PHO4* will aid in eliminating this possibility. Also, the function of Spl2 and the mechanism of inhibition of the low-affinity transporters have not yet been uncovered. Of immediate concern are the identification of the specific low-affinity transporters affected, whether it is the products of one, two, or all three genes. A second concern is whether we are able to detect changes in either the abundance or the plasma membrane localization of the transporters. Finally, a more detailed mechanistic understanding of the inhibition of the transporters will be a future experimental goal. Spl2 may display a novel activity in the regulation of a plasma membrane transporter that is applicable to a wide range of transporters. Spl2 may act as an adaptor protein that causes retention of the low-affinity transporters in the ER/Golgi vesicular system and prevents plasma membrane localization.

While the positive feedback model provides a coarse explanation of why the strain lacking the *PHO84* gene displays a Pho^C phenotype, it does not definitively answer the question of the role of *PHO84* in high phosphate conditions. Several explanations remain possibilities, and will require additionally experimentation to confirm or eliminate. The problem remains that the *pho84Δ* strain, even when grown continuously for long time periods in high phosphate medium, does not appear to revert to wild-type levels of Pho signaling as assayed by Pho4 nuclear localization. One explanation is that



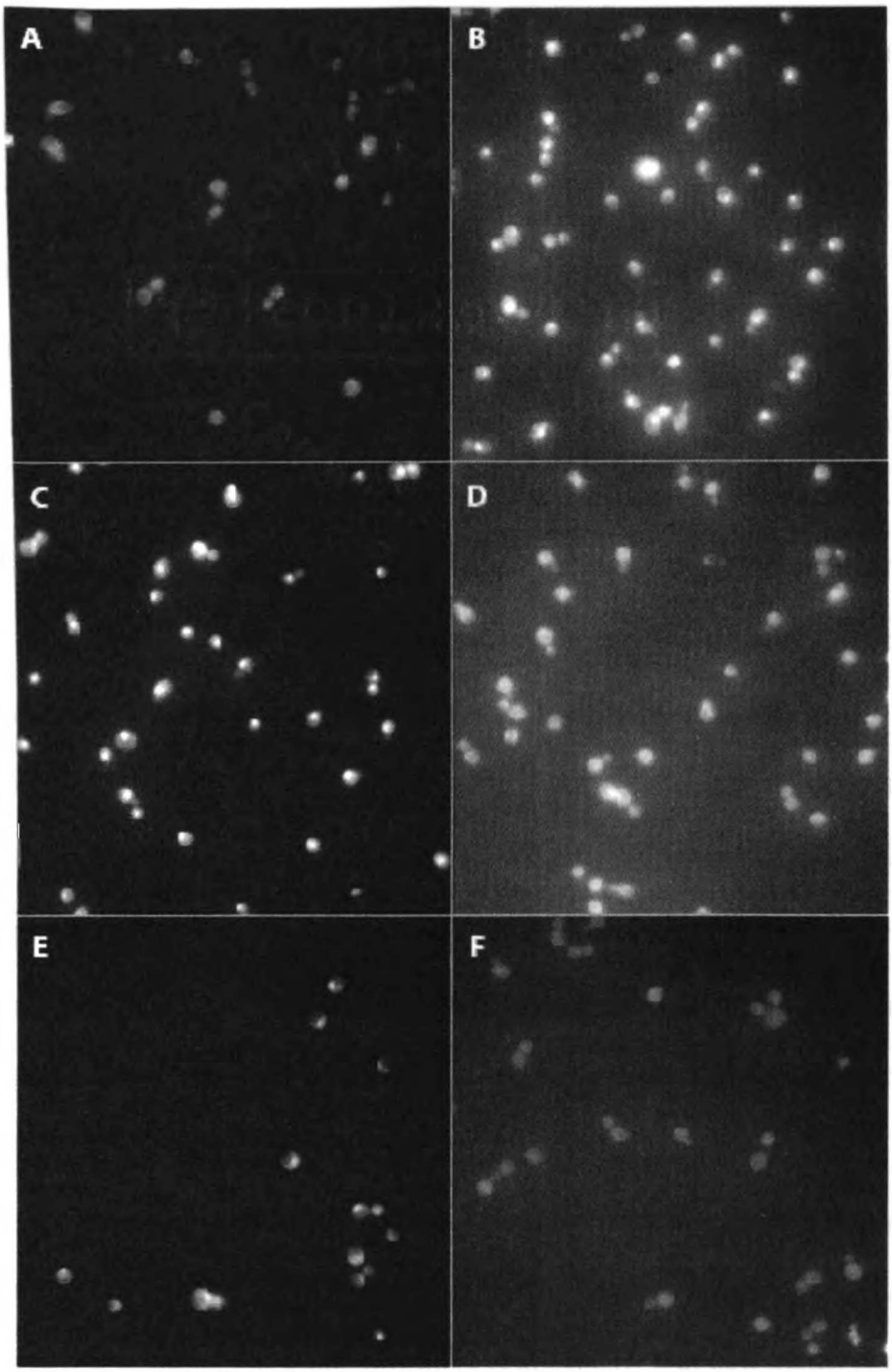
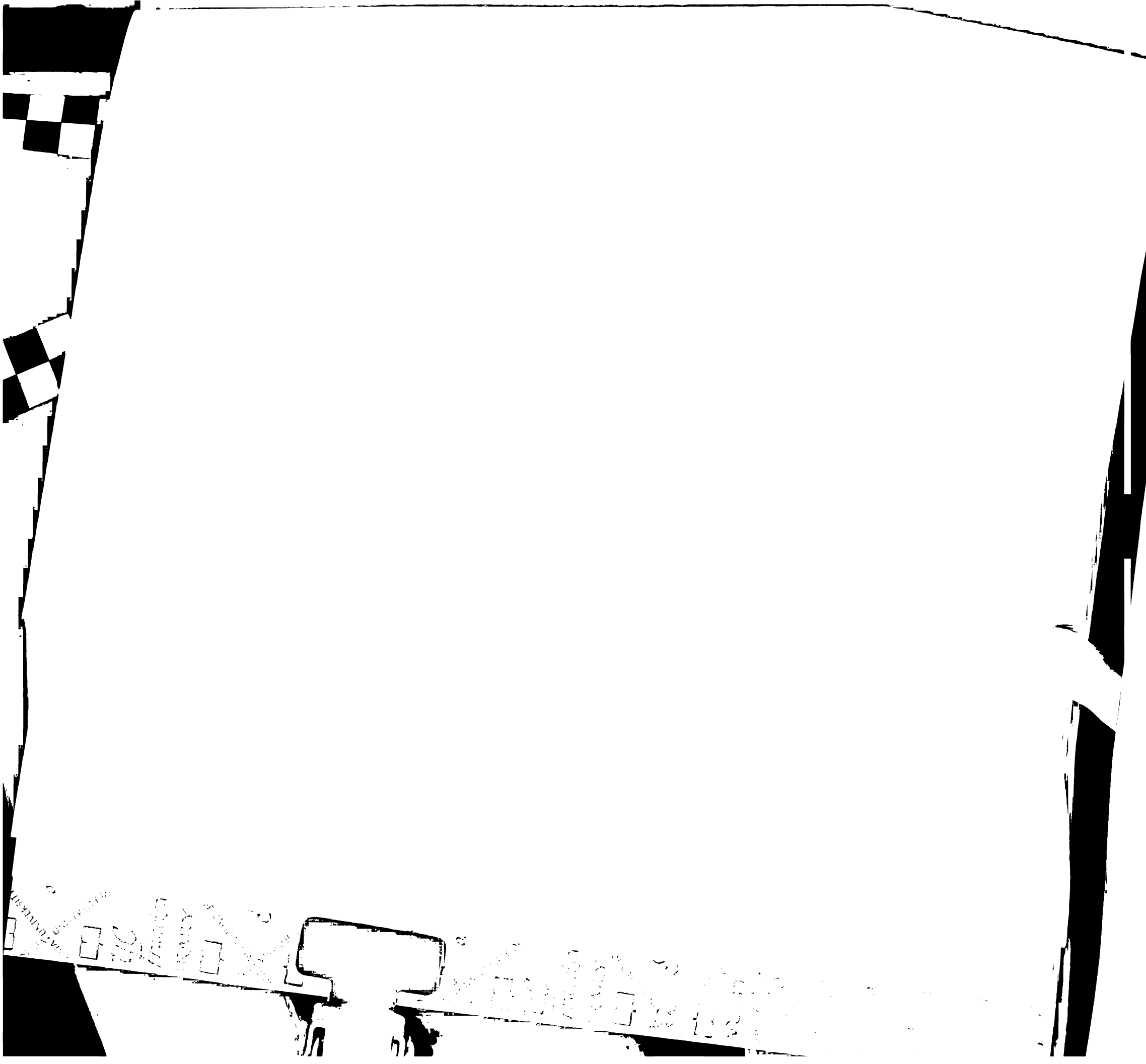


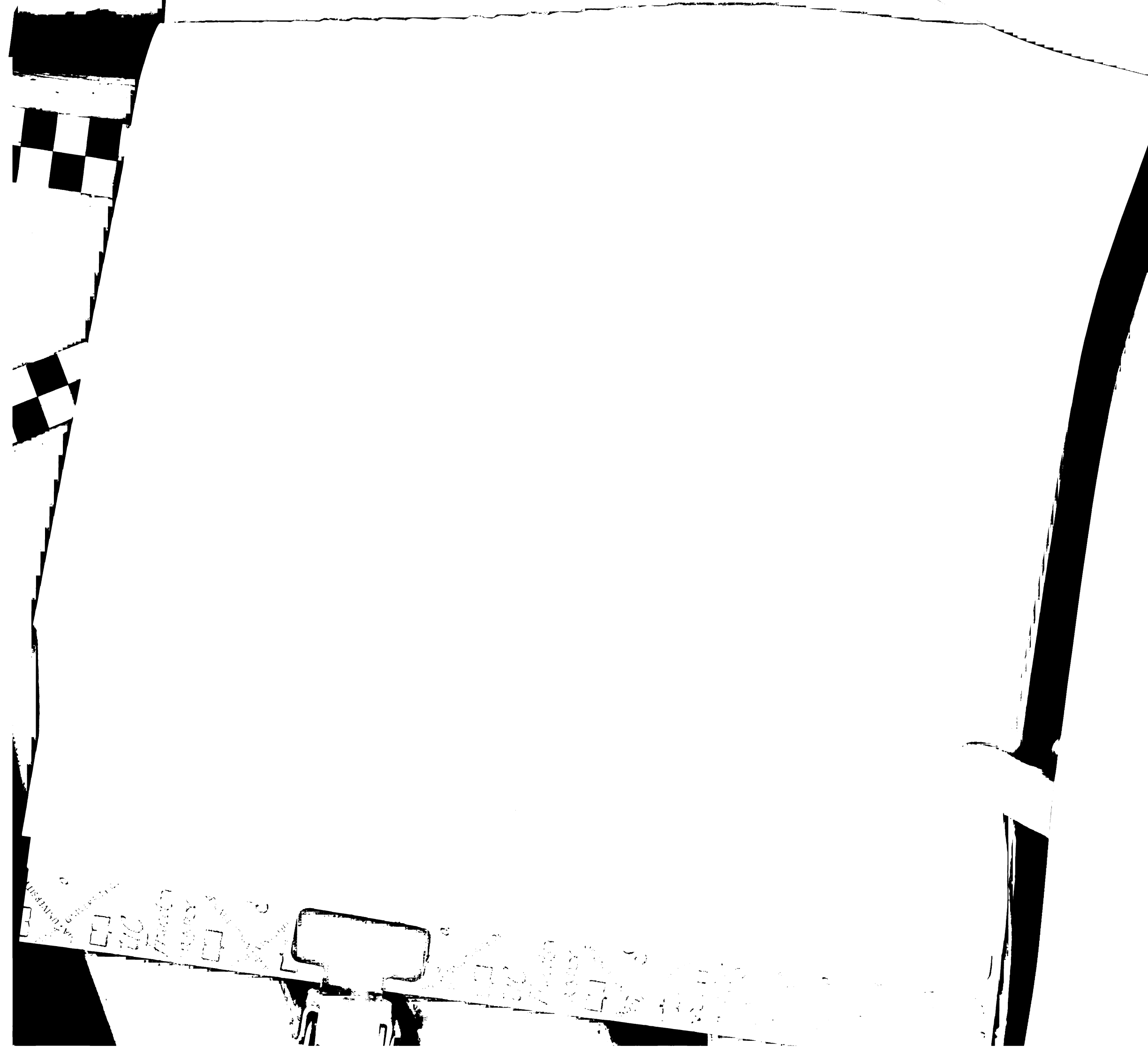
Figure 1.12. Examples of fluorescence micrographs for quantitation of Pho4-YFP localization. Different strains are displayed in each panel: (A) wild-type (B) *pho87Δ pho90Δ pho91Δ* (C) *pho84Δ* (D) *pho84Δ spl2Δ* (E) *ado1Δ* (F) control strain lacking Pho4-YFP.



the cells frequently experience intracellular phosphate fluctuations that require small but significant amounts of *PHO84* expression to preserve the high phosphate state. It may be possible to observe these fluctuations by continuous monitoring of Pho4 localization in individual cells, or by a more sensitive method of detection of expression from the *PHO84* promoter (such as luciferase). A second explanation is that in the *pho84Δ* strain, the low-level expression of *SPL2* in high phosphate conditions stands unopposed by the low level of *PHO84* expression normally present in the wild-type strain in high phosphate conditions. This unopposed Spl2 activity triggers a low-level of inhibition of the low-affinity phosphate transport system, resulting in lower intracellular phosphate levels in the *pho84Δ* cells. If this explanation is true, the deletion of *SPL2* in the *pho84Δ* strain should allow a full return to wild-type intracellular phosphate levels. Preliminary data suggests that this is not the case (see Fig. 4.11). A third, intriguing possibility is that there is no function whatsoever for *PHO84* if the gene is deleted while the cell is in a high-phosphate condition, but that once the cell starves for the first time, *PHO84* is required to achieve the high phosphate state of fully active low-affinity phosphate transport. This explanation could be tested by use of a plasmid loss experiment employing a plasmid-based copy of *PHO84*.

Previous views of the Pho pathway envisioned its function as a rheostatic response that induced high-affinity phosphate transport in proportion to the perceived deficit in intracellular phosphate (Springer et al., 2003). Our new model of the Pho pathway suggests that the pathway acts as a switch between two mutually exclusive, discrete intracellular states. In the low phosphate state, Pho4 activates high-affinity phosphate transport and suppresses low-affinity transport. In the high phosphate state,

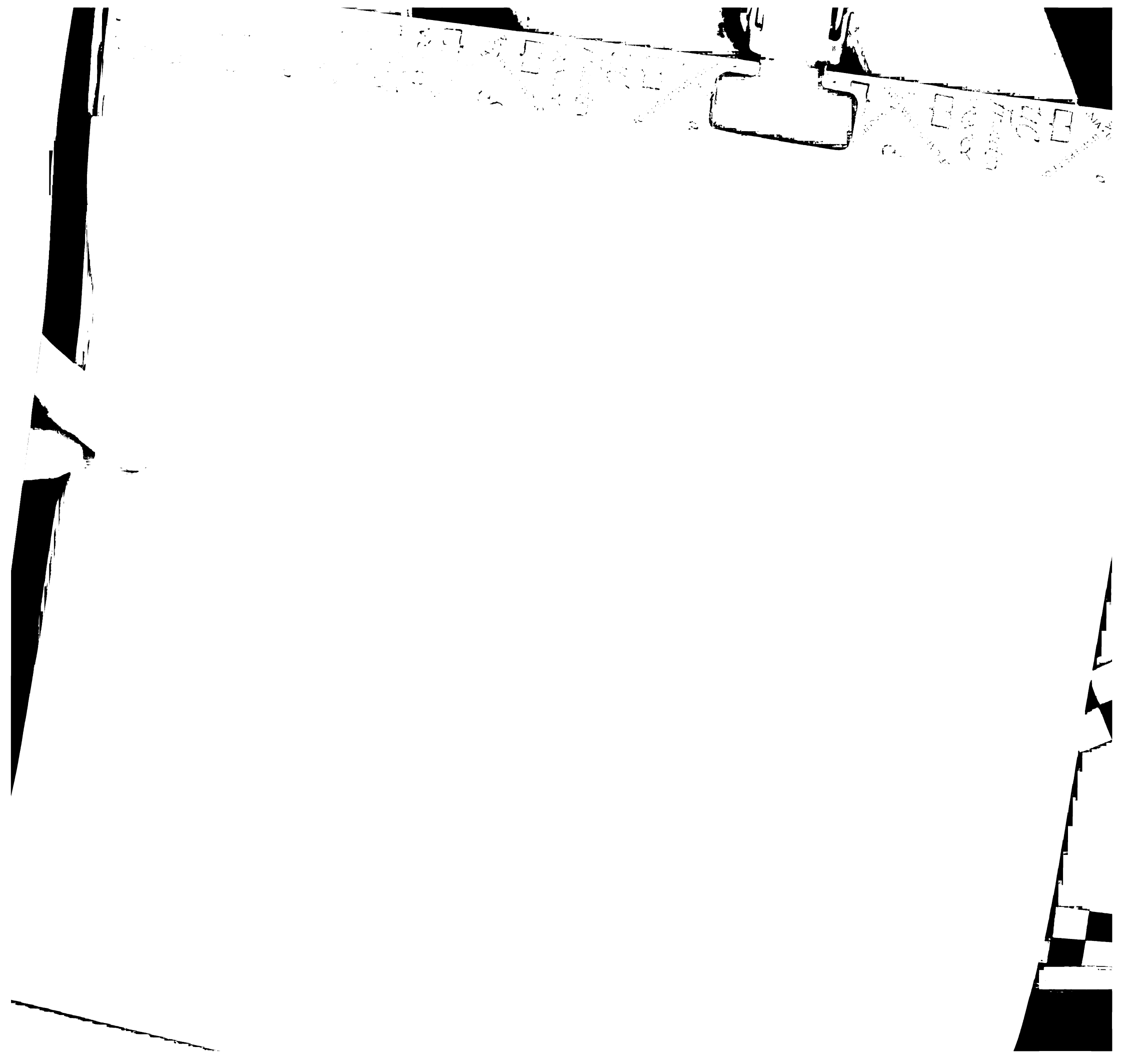
high-affinity transport is not substantial and the low-affinity transport system is fully active. It is notable that nutrient signal transduction pathways that behave in such a switch-like manner often respond to alternative forms of nutrients, such as the galactose pathway (Acar et al., 2005). The threshold for full activation of all the Pho4-dependent genes is at a very low extracellular inorganic phosphate concentration (Springer et al., 2003). In fact, the concentration threshold is so low that any substantial local density of cells, such as those found in a colony or in a liquid suspension, would rapidly deplete the medium of extracellular inorganic phosphate at this concentration. This observation seems counterintuitive: what is the utility of activating an inorganic phosphate scavenging gene program only immediately before the disappearance of inorganic phosphate from the environment? One rationalization for this behavior suggests a different function for the Pho pathway. In this view, the Pho pathway is not useful for surviving periods of low extracellular inorganic phosphate. Rather, the Pho pathway is useful for switching from utilization of extracellular inorganic phosphate to other organic phosphate sources present in the environment. In this case, the induction of the acid and alkaline phosphatases, enzymes that are capable of hydrolysis of inorganic phosphate from conjugated organic sources, are as important for survival as the induction of a high-affinity phosphate transporter. Evidence for this hypothesis is provided by the observation that the deletion of the *PHO4* gene does not affect cellular growth or viability in low inorganic phosphate medium, but is highly deleterious in medium containing only organic phosphate sources (data not shown). It will be interesting to study the effect of the induction of the extracellular hydrolytic enzymes like the acid phosphatase encoded by *PHO5* on intracellular signaling in medium containing only organic phosphate sources. It is likely that the *PHO5* gene will act as a



negative feedback element in those conditions.

While the switch from an inorganic to an organic source provides an explanation for the threshold of induction of the Pho4-dependent genes, it still does not suggest an answer to the question of why the low-affinity transporters are inhibited during phosphate starvation. There exist at least five different plausible arguments for the regulation of the low-affinity phosphate transport system by the Pho pathway. In fact, it is likely that more than one reason applies *in vivo*. One reason for the mutually exclusive control of the two transport systems would be if the different systems display different susceptibilities to toxic phosphate analogs such as arsenate and selenite. In this case, the Pho pathway activity might itself by regulated by trace amounts of toxic analogues. It would be advantageous for the cell to be able to switch to an alternative phosphate transport system if the switch enabled survival in the presence of an environmental toxin. A second reason would be if the low-affinity transporters are capable of bidirectional transport. In this case, any export of intracellular phosphate during periods of low environmental phosphate would be extremely deleterious for growth and survival.

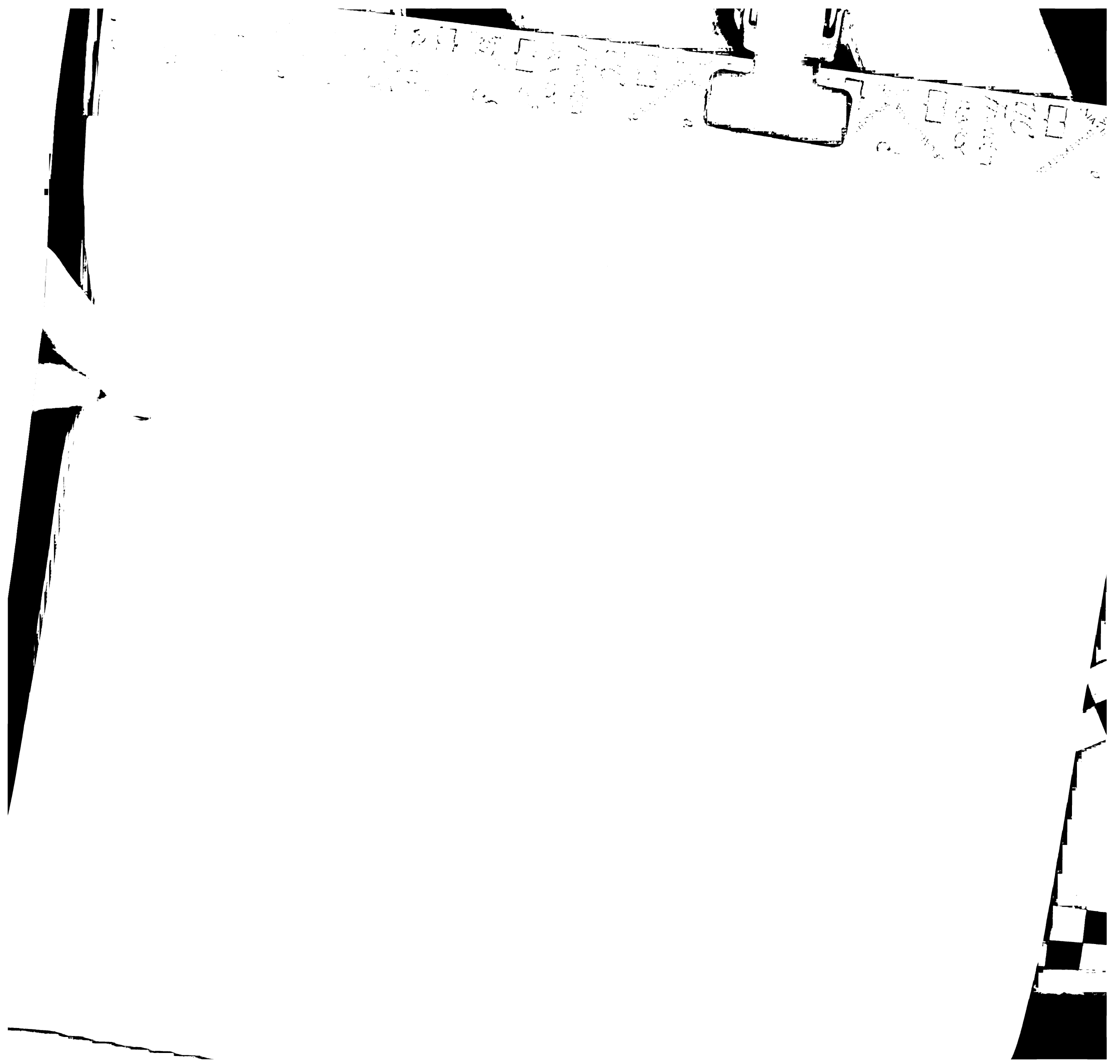
A third explanation for the downregulation of the low-affinity transporters is to prevent the intracellular phosphate overload that could occur when phosphate-depleted yeast, expressing both low-affinity and Pho4-dependent high-affinity phosphate transporters, suddenly encountered high levels of environmental phosphate. While inorganic phosphate is not generally recognized as toxic when present in excess, it is reasonable to postulate that cells would want to avoid "phosphate shock." In fact, the high-affinity phosphate transporter Pho84 is rapidly endocytosed when cells are moved from low to high phosphate media (Lau et al., 2000). While it is not known if the



low-affinity transporters are subject to the same endocytosis-based regulation, it is possible that the low-affinity transporters are inhibited by *SPL2* expression during low phosphate conditions to preempt the possibility of phosphate shock.

A fourth explanation concerns the energetic cost to the cell of allowing transporters of little functionality to remain active on the plasma membrane. The estimated K_m of the low-affinity transporters for inorganic phosphate (~100 μM) suggests that at very low concentrations of extracellular inorganic phosphate these transporters are highly inefficient. The transporters occupy some portion of the finite plasma membrane, and it may be beneficial for the cell to eliminate the nonfunctional transporters from the plasma membrane surface. (Of course, this argument relies on an unproven mechanism of inhibition involving removal from the plasma membrane.) In addition, and perhaps even more significantly, these transporters may act as proton or sodium symporters. It is possible that in the absence of sufficient extracellular phosphate the transporters result in an energetically costly sodium or proton "leak" into the cell.

A fifth and final explanation for the inhibition of the low-affinity transporters suggests that there is no direct physiological reason for transporter regulation. Instead, the transporters are downregulated solely for the effect that positive feedback has on the signaling characteristics of the Pho pathway. Positive feedback can confer a number of interesting signaling behaviors, such as kinetic anticipation of starvation, ultrasensitivity of the pathway response to the signal, multistability, and hysteresis or memory of past environmental states. These behaviors may prove beneficial in relation to the typical pattern of environmental phosphate changes over time. All five hypotheses suggest a number of potential experimental approaches to the question of why the low-affinity

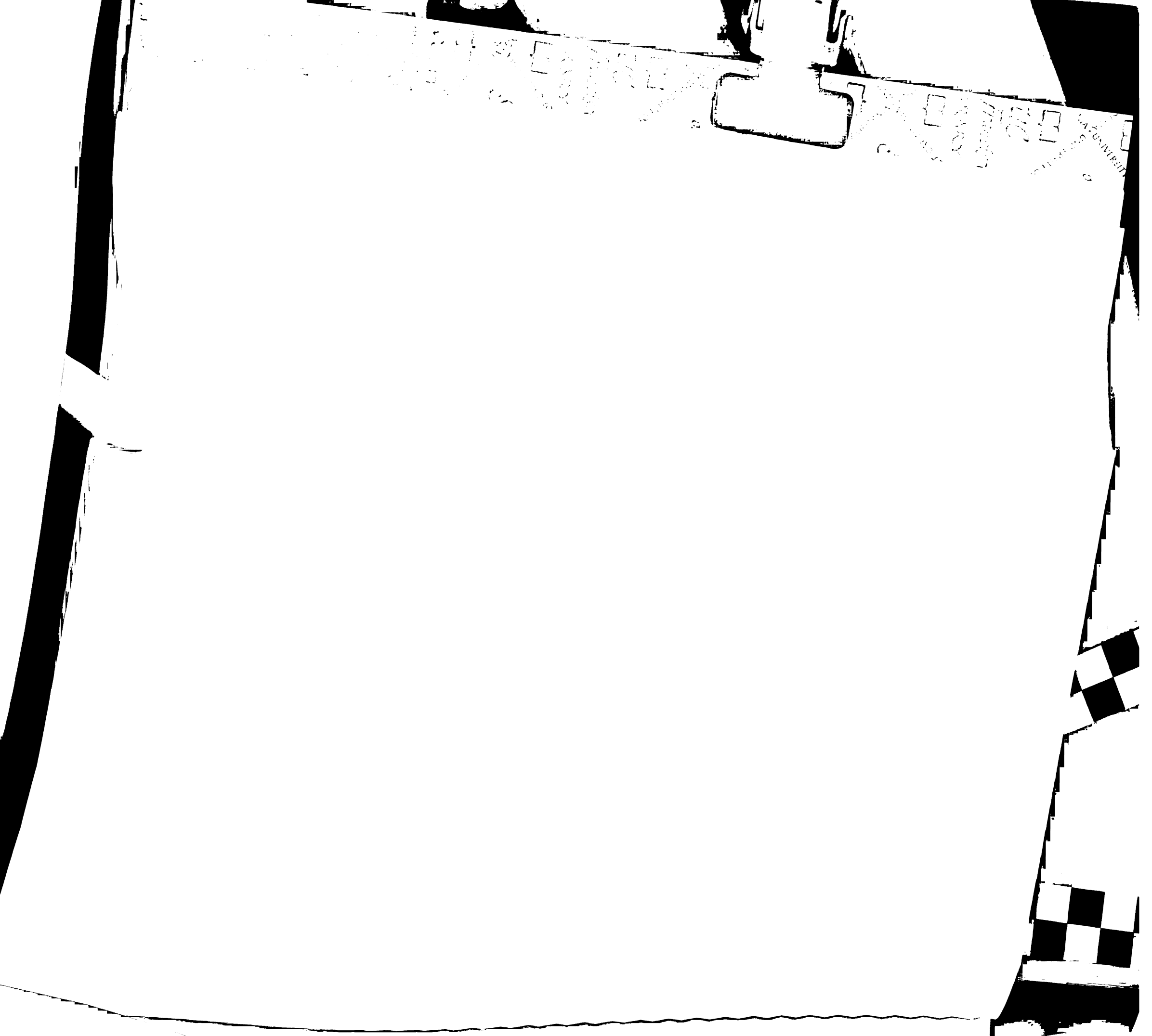


transporters are regulated, and what utility exists in a mutually exclusive regulation scheme.

The positive feedback model and the supporting data presented provide a basis for an integrated qualitative model of the Pho signal transduction pathway. Much more information about the quantitative relationships between the many components of the system remain to be gathered; however, we have provided new insight into the quantitative behavior of the Pho pathway. The proposed model represents an intermediate step between the linear pathways constructed by geneticists and the computational models that will inevitably be used to describe signal transduction pathways in more detail in the future.

Materials and Methods

Unless otherwise indicated, yeast strains were grown at 30°C in synthetic complete medium containing glucose with addition of 0.1 mg/mL adenine and 0.1 mg/mL tryptophan to suppress autofluorescence. No phosphate and low phosphate media were made as previously described (Wykoff and O'Shea, 2001). Liquid high phosphate medium contained ~19 mM phosphate. Deletion strains were constructed with the *Candida glabrata* system as described previously with elimination of the entire coding sequence from the start codon to the stop codon (Steger et al., 2003). Plate assay and liquid phosphatase assays were performed as previously described (Steger et al., 2003; Wykoff and O'Shea, 2001). Microscopy was performed using a digital imaging-capable Nikon TE200/300 inverted microscope with an oil-immersed objective at X 60 magnification. Image analysis was performed with the MetaMorph version 4.6r8 imaging software. Further details on materials and methods will be found elsewhere

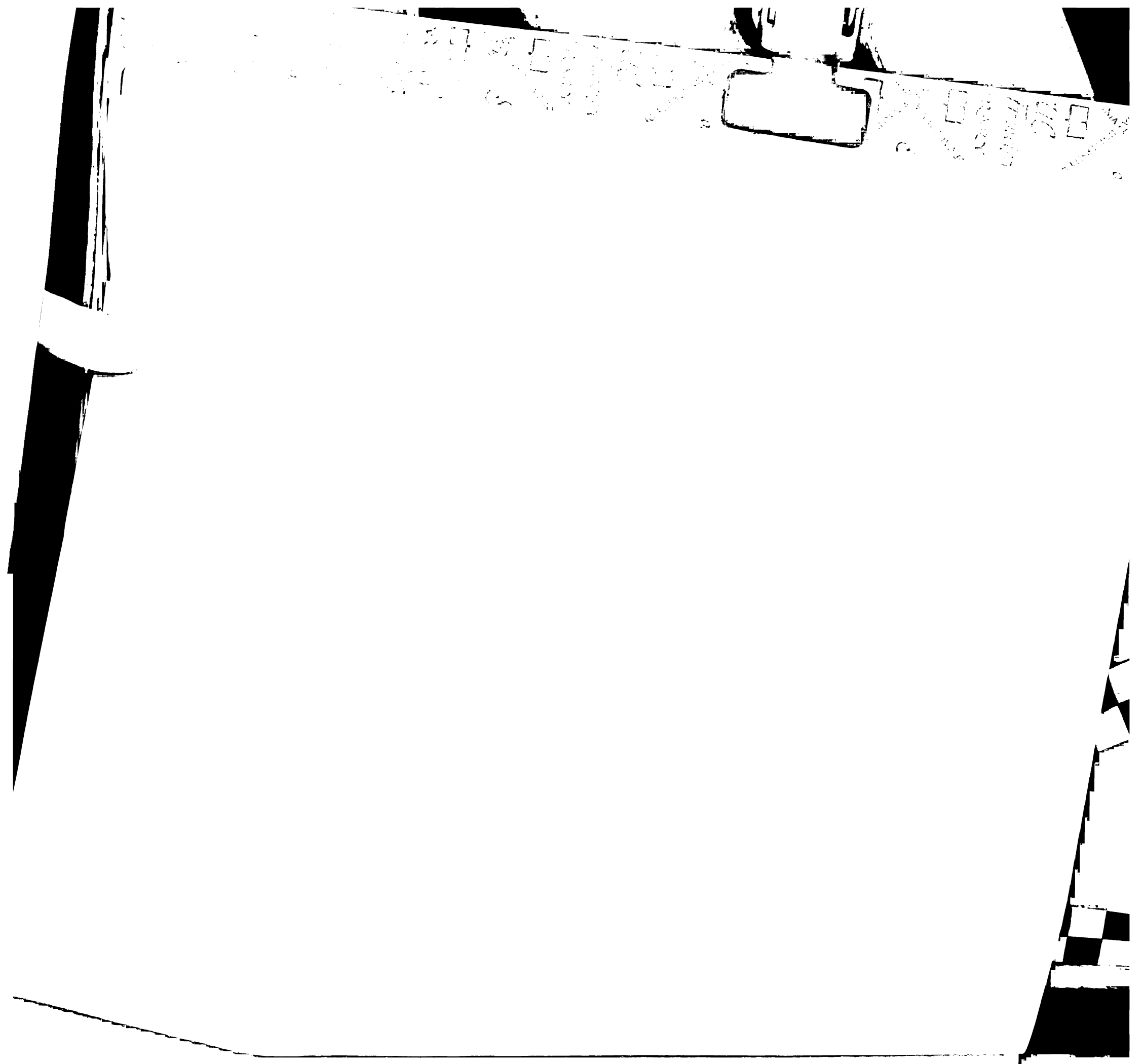


CONCLUSION

Our investigations have increased the understanding of the quantitative workings of nutrient-sensitive signal transduction pathways in the budding yeast *S. cerevisiae*. We have demonstrated that noise, or variability, is a seemingly universal characteristic of gene expression in budding yeast. Noise in gene expression has multiple sources and is specific to each gene and signal transduction pathway. It is likely that future work in other model organisms, especially in mammalian cells, will demonstrate that noise in signal transduction and gene expression displays similar characteristics in eukaryotic cells. It will be interesting to uncover the ways in which noise is suppressed or capitalized upon in cellular signal transduction.

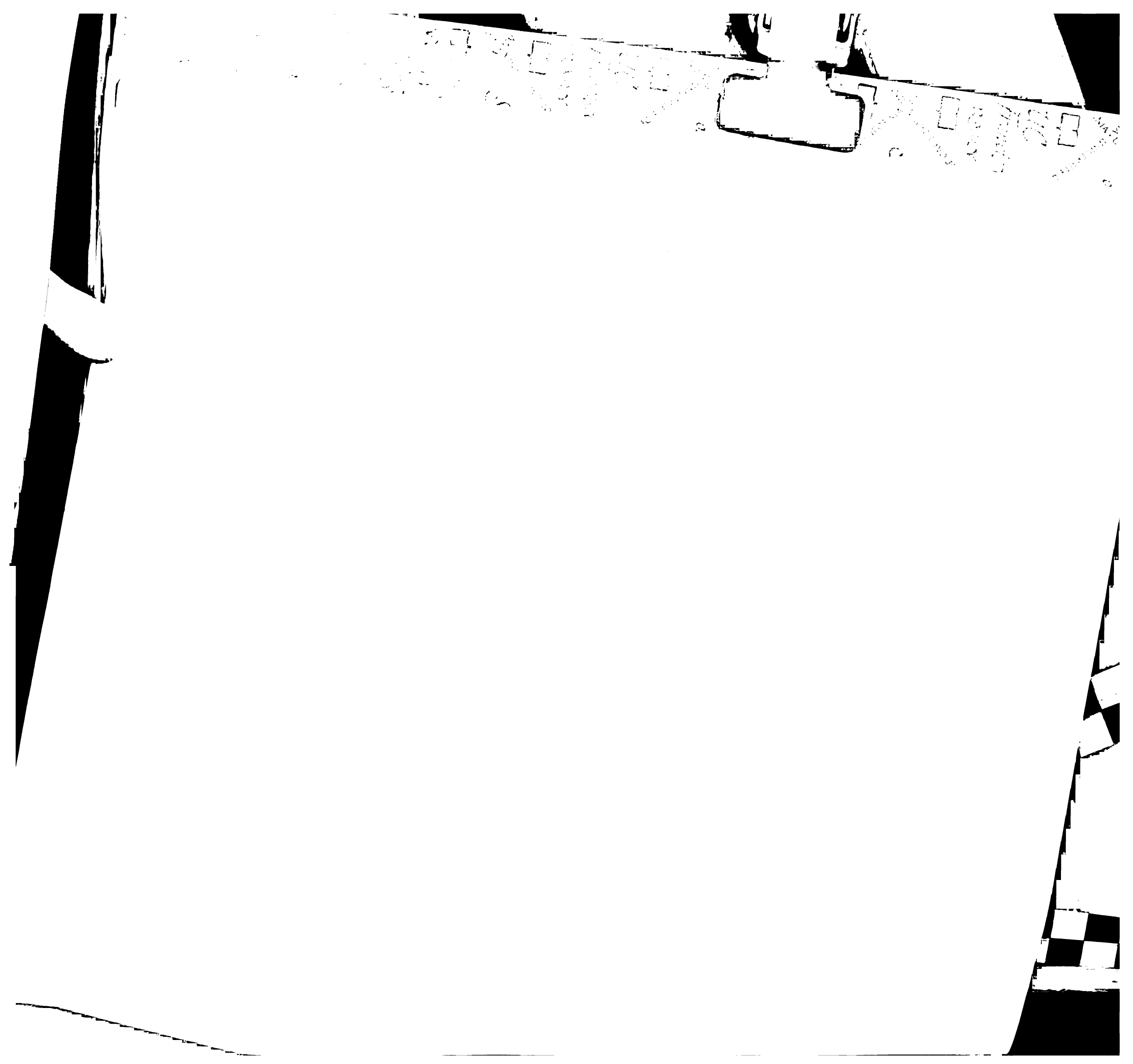
We have investigated the quantitative behavior of a the zinc-responsive signal transduction pathway in budding yeast. The zinc signaling pathway represents a system with a potentially simple relationship between the putative ligand, free intracellular zinc, and the zinc-sensitive transcription factor that controls the cellular zinc response. However, the system also represents a common complexity found in nutrient-sensing signal transduction pathway, namely the effect of multiple feedback loops on signal levels. In the case of zinc, as for many intracellular nutrients, feedback occurs through the manipulation of the rate of transport of the intracellular nutrient, both from the extracellular environment into the cell, and into and out of the intracellular stores. It is clear from our initial characterization that the behavior of the zinc signal transduction pathway is not consistent with our current understanding of the activities of and relationships between its components. We anticipate future experimental work that will allow a quantitative model in which the behavior of the signal transduction pathway is truly the sum of its component parts.

The phosphate-responsive signal transduction pathway in budding yeast



represents a model system that has been extensively characterized by genetic and biochemical investigations. By considering the behavior of the signaling pathway in the context of feedback loops, we have hypothesized and confirmed the existence of a significant missing component, a positive feedback loop. Our new model of the pathway behavior alters our understanding of the physiology and the mechanism of the phosphate response in budding yeast. Our quantitative approach emphasizes the importance of considering integrated models of signal transduction pathways rather than individual relationships between supposedly independent components.

All of our investigations have been conducted in the model unicellular eukaryote *S. cerevisiae*. We expect that the insights derived from studies in this organism will apply to diverse signal transduction pathways in other organisms. It is likely that the mathematical control structure that governs the pathways that we have examined will be conserved across cellular nutrient responses in highly diverged organisms, and may illuminate the processes that occur in all human cells. Regardless, we believe our work has confirmed the value of a quantitative approach to such investigations, and hope that this work will assist others in the future.



APPENDICES

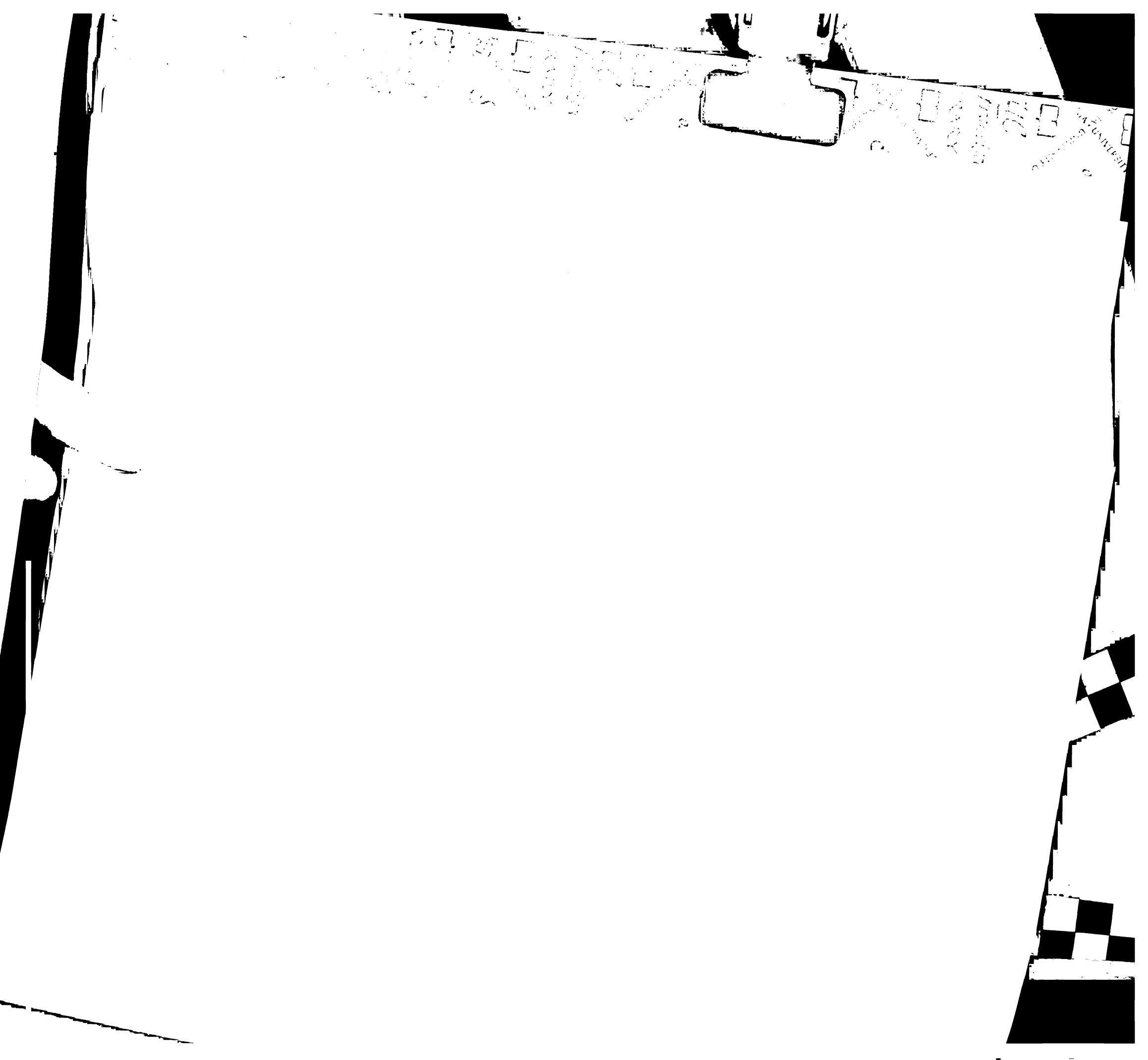
APPENDIX 1: Efficient exact stochastic simulator

In order to simulate the consequences of the stochastic nature of gene expression, I constructed a MatLab-compatible efficient exact stochastic simulator (EESS) as described previously (Gibson and Bruck, 2000; Gillespie, 1977). MatLab version 6.1.0.450 (May 18, 2001) was used. The code employed for the work described in Chapter One follows (EESSq.m):

```
% efficient exact stochastic simulator, in manner of Gibson and Bruck (2000)
% uses next reaction method [does not retain information from entire simulation!]

function [Q] = EESSq(I,P,R,K,T,n,it);

% user-defined input
% R is the reactant matrix (r by s) with stoichiometry
% P is the product matrix (r by s) with stoichiometry
% K is the constant matrix (r by 1), all constants in units per time
% I is the initial condition matrix (1 by s)
% T is the final time
% n is the number of cells to simulate

% it is the maximum number of iterations (if T is very large, may limit computation)
%it = 100000;

% the output variable, Q, is a three-dimensional matrix of dim it, s+1, n
% which has [0 I1 I2 I3;
%            tau1 S1 S2 S3
% and so on, for the n repetitions

% gather information about reactions and products
[r,s] = size(R);

% D is the dependency matrix, equal to abs(P-R)*R', without the diagonals
D = (abs(P-R))*R';
D = D - eye(r);

% building a tree with r nodes; tree matrix is u
% parent matrix is p; child matrix is c; index matrix is i
u = zeros(r,2);
p = zeros(r,1);
i = zeros(r,1);
c = zeros(r,2);

% parent matrix always starts with 1
p(1) = 1;

% special case for child matrix: 2 reactions
if r==2
    c = [2 2; 2 2];
else
    c(1,1) = 2;
    c(1,2) = 3;
end

% generate child matrix for >2 reactions
if r>2
for tb=2:r
    if c(tb-1,2)<r
        c(tb,1)=c(tb-1,2)+1;
        if  c(tb,1)<r
```

```
            c(tb,2)=c(tb,1)+1;
        else c(tb:r,2)=(tb:r)';
            c((tb+1):r,1)=((tb+1):r)';
        break
        end
    else c(tb:r,1)=(tb:r)';
        c(tb:r,2)=(tb:r)';
    break
    end
end
end

% generate parent matrix for 2 or more reactions
for tb=1:r
    p(c(tb,1))=tb;
    if c(tb,1)==r
        break
    end
    p(c(tb,2))=tb;
    if c(tb,2)==r
        break
    end
end

Q = zeros((s+1),n);
%FT = zeros(1,n);

for q=1:n

% initialization: generate the initial a matrix and tau matrix (r by 1)
a=zeros(r,1);
tau=zeros(r,1);

for x=1:r
    a(x)=K(x,1)*sum(R(x,1:s).*(I(1,1:s)));
    if a(x)<=0
        tau(x)=Inf;
    else
        tau(x)=(1/a(x))*log(1/rand);
    end
end

% sort the initial tau matrix and place in the tree
[A1 A2] = sort(tau);
u = [A1 A2];

% build the tree index
i(u(1:r,2))=(1:r)';

% estimate the empty t matrix size based on the number of iterations
X = zeros(it,(s+1));
X(1,1:(s+1))=[0 I];
t1 = 0;
t2 = 0;

% the actual loop
for step=2:it

% stop 2 reactions after tau reaches the time constant
    if t2>T
      break
    end

% for systems that can reach an irreversible state
% if all tau's are infinite, stop the simulation
    if u(1:r,1)==Inf
        break
    end

% identify the next reaction (mu) and the next time step t(step)
    mu=u(1,2);
```

```
    t0=u(1,1);

% record the tau and the new state in the X matrix
    X(step,1)=t0;
    X(step,2:(s+1))=[X((step-1),2:(s+1))] - [R(mu,1:s)] + [P(mu,1:s)];

% generate a new tau and new a for the current reaction
    anew=K(mu,1)*sum(R(mu,(1:s)).*(X(step,2:(s+1))));
    if anew<=0
        taunew=Inf;
    else
        taunew=(1/anew)*log(1/rand) + t0;
    end
    u(1,1) = taunew;
    a(mu,1) = anew;
    [u,i] = resort(u,1,c,p,i);

% recalculate a for the dependent reactions (excepting current reaction)
   for y=1:r
      if D(mu,y)>0
         anew=K(y,1)*sum(R(y,1:s).*(X(step,2:(s+1))));
         if anew<=0          % if reaction is now impossible
             taunew=Inf;
         elseif u(i(y),1)==Inf    % if impossible reaction is now possible
             taunew=(1/anew)*log(1/rand) + t0;
         else
             taunew=(u(i(y),1) - t0)*(a(y,1)/anew) + t0;
         end
         u(i(y),1)=taunew;
         a(y,1)=anew;
         [u,i] = resort(u,i(y),c,p,i);
      end
   end

% move times forward
   t2 = t1;
   t1 = t0;

end

%FT(q) = step;
Q(:,q)= X((step-1),:)';

end

% swap creates a temp matrix and then changes input matrix as well as index

function [z,i] = swap(z,n1,n2,i);
    ut = [z(n1,1:2);z(n2,1:2)];
    z(n1,1:2)=ut(2,1:2);
    z(n2,1:2)=ut(1,1:2);
    i(ut(1,2))=n2;
    i(ut(2,2))=n1;

% resort reorders the tree (seems to work well and quickly)

function [z,i] = resort(z,n1,c,p,i);
if z(c(n1,1),1)<z(c(n1,2),1)
    minc = z(c(n1,1),1);
    mincID = c(n1,1);
else
    minc = z(c(n1,2),1);
    mincID = c(n1,2);
end

if z(n1,1) < z(p(n1),1)
    [z,i] = swap(z,n1,p(n1),i);
    [z,i] = resort(z,p(n1),c,p,i);
elseif z(n1,1) > minc
    [z,i] = swap(z,n1,mincID,i);
    [z,i] = resort(z,mincID,c,p,i);
```



```
else
end
i = i;
```

To implement the stochastic simulator for a particular reaction scheme, the following code was employed, varying the k_a constant (see Chapter One):

```
function out = provar

% first section: user-defined specific reactions
% R is the reactant matrix (r by s)
% P is the product matrix (r by s)
% K is the constant matrix (r by 1)
% I is the initial condition matrix (1 by s)
% T is the final minimum time

kcount = 33;
kmax = 10;
kmin = 0.001;
f = exp(log(kmax/kmin)/(kcount-1));

for ttt=1:4

for k=1:kcount
    k1 = kmin*f^(k-1);

    R1 = [1 0 0 0; 0 1 0 0; 0 1 0 0; 0 0 1 0; 0 0 1 0; 0 0 0 1];
    P1 = [0 1 0 0; 0 1 1 0; 1 0 0 0; 0 0 1 1; 0 0 0 0; 0 0 0 0];
    K1 = [k1; 10; 0.1; 10; 5; 0.1];
    I1 = [1 0 0 0];
    R = [R1];
    P = [P1];
    K = [K1];
    I = [I1];
    T = ttt*100;
    n = 100;
    it = 1000000;
    [Q] = EESSq(I,P,R,K,T,n,it);
    Q(2:4,:) = [];

    Qfin = zeros(n,1);
    Qfin(:,1) = Q(2,:)';

    Qexp = [Qfin];
    meanQ = mean(Qexp,1);
    mean(k) = meanQ;
    constant(k) = k1;
    stdQ = std(Qexp,1);
    NS(k) = stdQ*stdQ./mean(k);

    rate(k) = mean(k)/T;
    %dlmwrite('temp.txt',[constant; mean; NS]',',')

end

X1 = mean;
Y1 = NS;
rootname = 'caseIka';
extension = '.txt';
filename = [rootname, int2str(T), extension];
dlmwrite(filename,[constant; X1; Y1]',',')

end
```

APPENDIX 2: Modeling the evolution of nutrient homeostasis

Complex feedback is a common characteristic of nutrient signal transduction pathways. In order to explore how such complexity may have evolved, I developed an in silico evolutionary simulator. The simulator explored the fitness of various regulatory schema, based on the assumption that the cell would grow the faster when intracellular nutrient fluctuations were minimized, regardless of the extracellular environment. In particular, the simulator focused on the regulation of plasma membrane import, with independent low-affinity and high-affinity transporter activites, and vacuolar import and export. Values were taken from the literature whenever possible. The MatLab-compatible programs evolve3.m and Zn4varevol3.m follow. MatLab version 6.1.0.450 (May 18, 2001) was used.

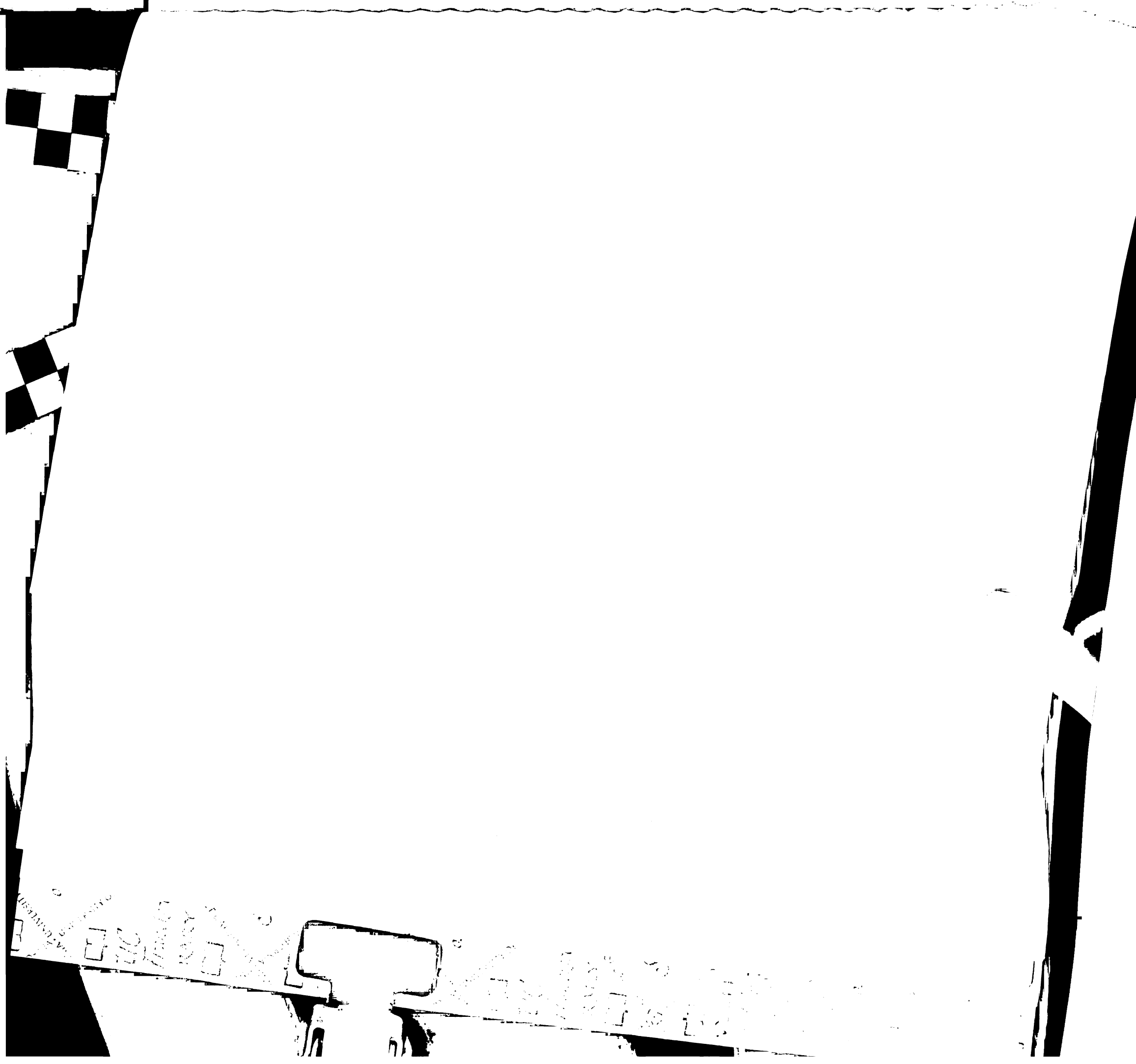
evolve3.m

```
%function S = evolve3(cellnumber,generationnumber,generationtime)

% the in silico evolution m-file v2
% the goal is to see if there are relative thresholds in gene expression
% that are more or less advantageous in maintaining intracellular nutrient
% concentrations relatively constant
% the m-file Zn4varevol3.m contains the differential equation set with fixed constants

% the state of the system S is a cellnumber by (variablenumber + odevariable +
flagnumber) matrix that determines each cell absolutely

% note on time to run program: 1 cell x 100 generations is ~10-25 s
% implying that 100 cells x 10000 generations will be 10000 times longer = ~28 hours
% new run: 2 cells x 1000 generations = 4.8 min
% 100 cells x 10000 generations is then 40 hours
% as of 12-18-02 0006 plan:
%       DONEdebug the loop generator 0 duration issue
%       DONEadd the fitness calculation
%       DONEadd the sort algorithm (from EESS?)
%       DONEadd the selection procedure
%       DONEadd the mutagenesis procedure
%       add the data save commands
%       write a display m-file for the saved data
%       DONEdecide on 'real' numbers for BASAL, MAXIMAL, THRESHOLD
%       translate the random BMT numbers into real numbers

cellnumber = 100;
% variablenumber is the number of attributes each cell has (4 transporters x 3
attributes) = 12
variablenumber = 12;
% odevariable is the number of internal state variables: 4 transporters, Zi, Zv (not
including Ze)
odevariable = 6;
flagnumber = 3;     % at the moment, flag1 = duplications, flag2 = mutations, flag3 =
parentage
```

```
generationnumber = 100;
generationtime = 20;
recordfreq = 1;
% each cell's chance per generation of mutation (only 1 of 12 attributes is mutated at a
time)
mutfreq = 0.04;
% maximum step size of mutagenesis (a nonnegative integer)
maxmutation = 3;

% the maximal length of each individual stimulus
maxstimulus = 270;

% the number of cells to be duplicated
fitnessnumber = 1;

% threshold for toxicity of intracellular zinc in uM
toxicthresh = 120;
% Km of fitness function at toxicity end
toxicsteep = 10;
% Km of fitness function at starvation end
starvsteep = 10;

totaltime = generationnumber*generationtime

S = zeros(cellnumber,variablenumber+odevariable+flagnumber);

% figure

% establish the random set of initial conditions
% BASAL level of expression is a number between 0 and 10 that linearly scales with the
expression
% THRESHOLD level of expression is a number between 0 and 10
%    will convert this to a log proportional concentration scale
%    0 = never turns on (Km = Inf); 10 = constitutive (Km is 0);
%    the rest are scattered across the range of Zap1 sensitivity (0 to 100 uM)
% MAXIMAL level of expression is a number between 0 and 10 that linearly scales with the
expression;

% rand(m,n) generates a matrix of random over interval [0 1];
% fix(x) rounds down toward zero

S(1:cellnumber,1:variablenumber) = fix(11*rand(cellnumber,variablenumber));

S(:,variablenumber+odevariable+3) = [1:cellnumber]';

% start with none of anything and allow cells to equilibrate at an intermediate Ze before
looping

% make conditions matrix C; each entry is [Zeflag Ze1 Ze2 Ze3 time]

C = zeros(1,5);
Cind = 1;
tt = 0;

while sum(C(:,5)) < totaltime

    c1 = 0;
    %c2 = 50;
    c2 = 10^(12*rand - 8);
    %c2 = rand;
    c3 = 0;
    c4 = 0;
    c5 = max(fix((maxstimulus+1)*rand),1);
    %c5 = totaltime;
    C(Cind,:) = [c1 c2 c3 c4 c5];
    tt = tt + c5;


    while tt > generationtime

        originterval = c5;
```

```
        previnterval = tt - originterval;

        c5 = generationtime - previnterval;

        C(Cind,:) = [c1 c2 c3 c4 c5];

        Cind = Cind + 1;

        c5 = originterval - c5;

        C(Cind,:) = [c1 c2 c3 c4 c5];

        tt = c5;
    end

    % without this, will get zero after tt = 90 exactly

    if tt == generationtime

        tt = 0;

    end

    Cind = Cind+1;

end

loopnumber = size(C,1)

% equilibration of initial cells

for k=1:cellnumber

    Ze = [0 50 0 0];

    y0 = [Ze(2); 100; S(k,(variablenumber+2):(variablenumber+odevariable))'];

    %y0 = [Ze(2); S(k,(variablenumber+1):(variablenumber+odevariable))']

    V = S(k,1:variablenumber);

    tspan = [0 2.5 5];

    % numerically solve the ODEs
    % t is a time column vector with values from 0 to looptime
    % y is a looptime + 1 x odevariable matrix with the odevariable values at each time

    [t,y] = ode23(@Zn4varevol3,tspan,y0,[],V,Ze);

    % transfer the state of the cell at the end of the loop into the S matrix

    S(k,(variablenumber+1):(variablenumber+odevariable)) = y(3,2:odevariable+1);

end

check = 'equilibrated'

% giant loop: take matrix of random cells, run through ODE solver over generated stimuli

timemark = 0;

M = zeros(cellnumber,odevariable+1,generationtime);

E = zeros(generationnumber,1);
Em = zeros(generationnumber,1);
G = [1:generationnumber]';

for i=1:loopnumber

        Ze = C(i,1:4);
```

```
        tspan = 0:C(i,5);

        ttt = mod(timemark,generationtime);

        for k=1:cellnumber

            y0 = [Ze(2); S(k,(variablenumber+1):(variablenumber+odevariable))'];

            V = S(k,1:variablenumber);

            % numerically solve the ODEs
            % t is a time column vector with values from 0 to looptime
            % y is a looptime + 1 x odevariable matrix with the odevariable values at
each time

            [t,y] = ode23(@Zn4varevol3,tspan,y0,[],V,Ze);

            % save the state of the cell (just the odevariables) over time into M matrix
            % M matrix has dimensions cellnumber x odevariable x 90
            % time runs in the third dimension from +1 to +90

            M(k,1:odevariable+1,ttt + 1:ttt + C(i,5)) = y(2:C(i,5)+1,1:odevariable+1)';

            % transfer the state of the cell at the end of the loop into the S matrix

            S(k,(variablenumber+1):(variablenumber+odevariable)) =
y(C(i,5)+1,2:odevariable+1);

        end

        timemark = timemark + C(i,5)

        % calculate aggregate fitness if a full generation has passed

        F = zeros(cellnumber,generationtime);

        F2 = zeros(cellnumber,1);

        F3 = zeros(cellnumber,1);

        if mod(timemark,generationtime) == 0

            % calculate fitness

            for m=1:cellnumber

                for n=1:generationtime

                    Zint = M(m,2,n);

                    if Zint < 0

                        F(m,n) = 0;

                    elseif Zint < toxicthresh

                        F(m,n) = Zint*(-(Zint-toxicthresh))/((Zint +
starvsteep)*(toxicsteep - (Zint - toxicthresh)));

                    else

                        F(m,n) = 1/(Zint^2);

                    end

                end

                F2(m) = sum(F(m,:))/generationtime;

            end
```



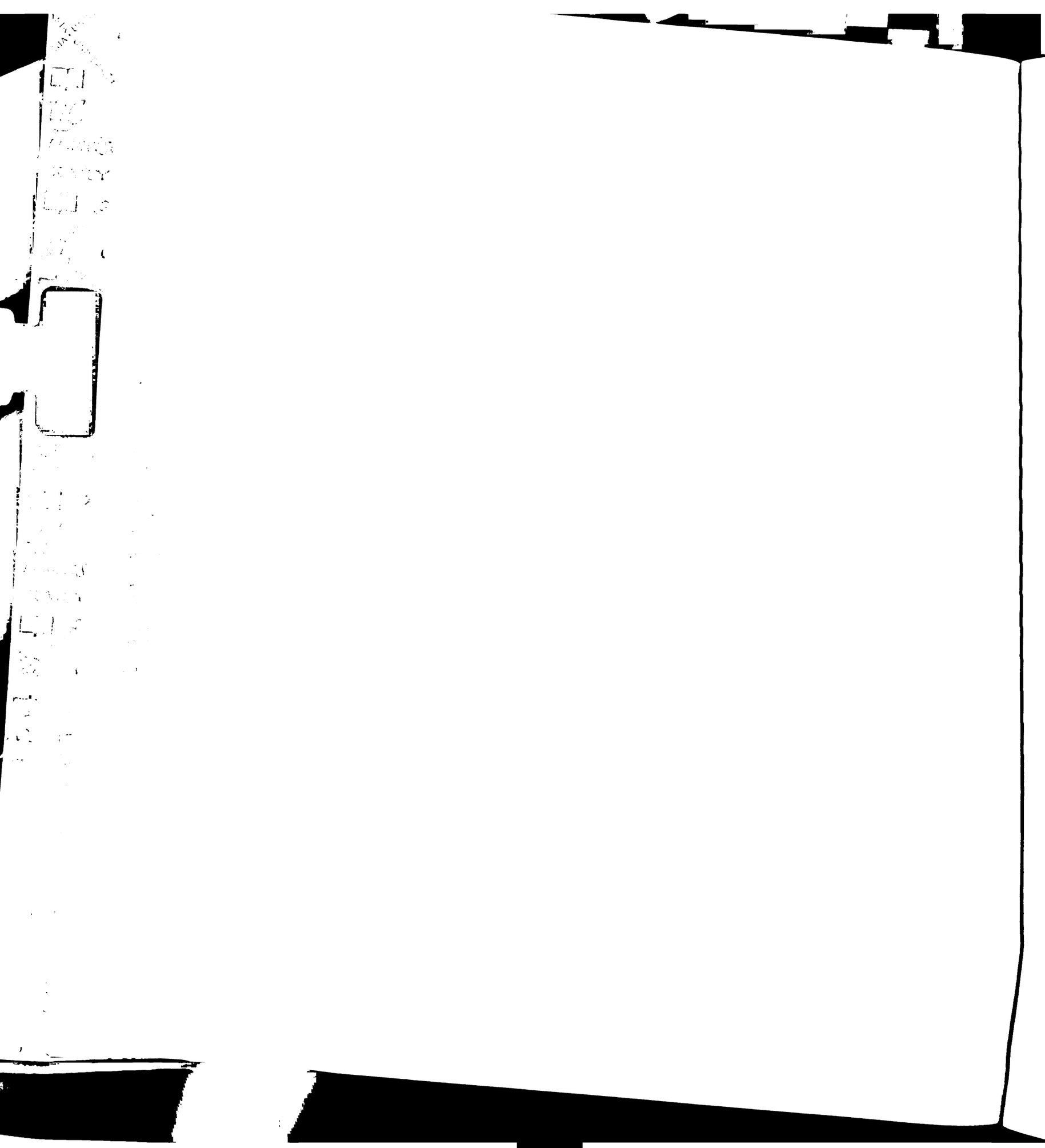
```
% obtain index of order

[F3,IF] = sort(F2);

% record best fitness into matrix E

E(timemark/generationtime) = F3(cellnumber);

Em(timemark/generationtime) = mean(F3);
% record average concentrations

Cn(timemark-generationtime+1:timemark) = M(1,1,:);

% replace the bottom fitnessnumber # with the top fitnessnumber #

if fitnessnumber > 0

for q=1:fitnessnumber

    S(IF(q),:) = S(IF(cellnumber-q+1),:);

    % flag that a cell has been duplicated

    S(IF(cellnumber-q+1),variablenumber+odevariable+1) = S(IF(cellnumber-
q+1),variablenumber+odevariable+1) + 1;

    % flag the parentage of the new cell

    %S(IF(q),variablenumber+odevariable+3) = IF(cellnumber-q+1);

end

else

end

% each cell has a random chance of mutagenesis;

for k=1:cellnumber

    mutchance = rand;

    if mutchance <= mutfreq

        % mutagenize the cell

        mutindex = 1 + fix(rand*variablenumber);

        % this mutagenesis protocol moves from the current value,
        % either up or down, by a random step (maximum is fixed by
        % the maxmutation parameter

        mutdir = randn;

        mutsize = rand*maxmutation;

        if randn > 0

            S(k,mutindex) = min(10,(S(k,mutindex)+mutsize));

        else

            S(k,mutindex) = max(0,(S(k,mutindex)-mutsize));

        end

        % ! an alternative would be to simply replace the current parameter
        % by a random number between 0 and 10;

        %S(k,mutindex) = rand*10;
```

```
                    % add to mutagenic flag

                    S(k,variablenumber+odevariable+2) = S(k,variablenumber+odevariable+2)
+ 1;

                end

            end

        end


        % temporary plotting device: Ze vs t, Zi vs t, Zv vs t

        %subplot(3,1,1),semilogy(t,y(:,1))
        %hold on

        %subplot(3,1,2),plot(t,y(:,2))
        %hold on

        %subplot(3,1,3),plot(t,y(:,5))
        %hold on

    end

end

% display final cellular state
figure
E = E(1:generationnumber);
Em = Em(1:generationnumber);
Cn = Cn(1:totaltime);
subplot(2,1,1), plot(G,Em,'.r',G,E,'.b')
hold on
subplot(2,1,2), semilogy([1:totaltime]',Cn,'.g')
hold off
S1 = [S(:,1:12) S(:,19:21)]
S2 = [S(:,13) S(:,15)]
S2(IF(cellnumber),:)
S1(IF(cellnumber),:)
```

Zn4varevol3.m

```
function dydt = Zn4varevol(t,y,V,Zeinfo)

% this function sets up differential equations relating changes in intracellular Zn and
transporter abundance,
% in response to extracellular Zn over a certain time-span

% some variables are fixed, such as the Km and Vmax of each transporter, the Ki of
% vacuolar import, the Km of Zap1 for Zn, the degradation rate of each component,
% and the usage rate of Zn; the source for each number is indicated

% v1, v2, v3 are the high-affinity plasma membrane transporter variables;
% v4, v5, v6 are the low-affinity plasma membrane transporter variables;
% v7, v8, v9 are the vacuolar influx transporter variables;
% v10, v11, v12 are the vacuolar efflux transporter variables;
% in each triple, the first is BASAL expression rate,
% the second the THRESHOLD for Zap1-dependent activation,
% and the third the MAXIMAL expression rate


transportermax = 100000;
% the maximum acculumalation of transporter at the cell surface
% theoretically, allows for a maximum of

thrfactor = 20;

expfac = 6;
```

```
if V(1) > 0
v1 = exp(V(1)-expfac);
else v1 = 0;
end
if V(3) > 0
v3 = exp(V(3)-expfac);
else v3 = 0;
end
if V(4) > 0
v4 = exp(V(4)-expfac);
else v4 = 0;
end
if V(6) > 0
v6 = exp(V(6)-expfac);
else v6 = 0;
end
if V(7) > 0
v7 = exp(V(7)-expfac);
else v7 = 0;
end
if V(9) > 0
v9 = exp(V(9)-expfac);
else v9 = 0;
end
if V(10) > 0
v10 = exp(V(10)-expfac);
else v10 = 0;
end
if V(12) > 0
v12 = exp(V(12)-expfac);
else v12 = 0;
end

%v3 = exp(V(3)-expfac);

%v4 = exp(V(4)-expfac);

%v6 = exp(V(6)-expfac);

%v7 = exp(V(7)-expfac);

%v9 = exp(V(9)-expfac);

%v10 = exp(V(10)-expfac);

%v12 = exp(V(12)-expfac);

v2 = V(2)*thrfactor;
v5 = V(5)*thrfactor;
v8 = V(8)*thrfactor;
v11 = V(11)*thrfactor;

% the following variables are set by documented data or realistic estimates
% abbreviations: HA = high-affinity; LA = low-affinity; VI and VE = vacuolar influx and
efflux

% all concentration values in micromoles; all time values in minutes
% all abundances and rates are per cell

% thought it was 1e-14; that doesn't work
transportrate = 1e-15;

VmaxHA = transportrate;      %each individual transporter can transport at a rate of x
micromoles per minute
VmaxLA = transportrate;
VmaxVI = transportrate;
VmaxVE = transportrate;

% vacuolar volume is estimated at 1/2 the cytoplasm
vacratio = 0.5;
```



```
%concratio = 1/(1.25e-13); %just messing

concratio = 1/(5e-14); %just messing


% cellular cytoplasmic volume = (50 um)^3; this means that every mole transported in
results in
% 8 * 10^12 molar concentration

KmHA = 0.5e-6;
% ~0.5 nM based on uptake measurements WT vs zrt1 delete

KmLA = 5e-4;
% ~10 nM based on uptake measurements WT, zrt1 delete

KmVI = 0.1;
% ~0.1 uM based on in vitro vacuolar Zn uptake measurements

KmVE = 1e-5;
% ~10 nM speculative based on homology of ZRT3 to ZRT1 and ZRT2

KiVI = 10000;
% ~10 mM based on resting vacuolar concentration of ~5 mM

setpoint = 100;
% based on calculations of free Zn/cytoplasmic volume at high Zn
% setpoint is the Zi at which Zap1 is OFF

% the degradation rate is due to division for all transporters and for intracellular and
extracellular zinc as well:
% calculated as gamma = ln(2) / 90 minutes

gamma = 0.007701635339554947882 4136902384242;

% usage = 50 pmol/10^6 cells/90 minutes

usage = (5e-11)/90*1e11;

% the VmaxT sets the maximal level of accumulation of transporter
% ~100,000 per cell

transportermaxfactor = transportermax/exp(10-expfac);

VmaxTHA = gamma*transportermaxfactor;
VmaxTLA = gamma*transportermaxfactor;
VmaxTVI = gamma*transportermaxfactor;
VmaxTVE = gamma*transportermaxfactor;

% y1 = [Ze];

% the if/elseif setup gives the derivative of the [Ze] based on the flags below:

% Zeflag indicates which type of extracellular Zn stimulus is indicated;
% Zeflag = 0 indicates a constant concentration; Ze1 is that concentration; others are
null
% Zeflag = 1 indicates an exponential decay (a*exp(-bt)); Ze1 is a, Ze2 is b, Ze3 is null
% Zeflag = 2 indicates an extracellular conc. of form a sin (bt + c); Ze1 a, Ze2 b, Ze3 c

Zeflag = Zeinfo(1);
Ze1 = Zeinfo(2);
Ze2 = Zeinfo(3);
Ze3 = Zeinfo(4);

% flags section

if Zeflag == 0

    y1 = 0;

elseif Zeflag == 1
```

```
    y1 = -Ze1*Ze2*exp(-Ze2*t);

elseif Zeflag == 2

    y1 = Ze1 * Ze2 * cos (Ze2*t + Ze3);

else

    break

end

% y2 = [Zi]; unclear if concratio should be applied to usage term as well

y2 = concratio*y(1)*VmaxHA*y(3)/(y(1)+KmHA) + concratio*y(1)*VmaxLA*y(4)/(y(1)+KmLA) +
y(5)*VmaxVE*y(7)/(y(5)+KmVE) - y(2)*VmaxVI*y(6)/(KmVI+y(2)+KmVI*y(5)/KiVI) - gamma*y(2) -
usage*y(2)/(y(2)+1e-12);

% active Zap1 as a function of Zn

Zap1 = max(setpoint-y(2),0.0001);

% y3 = HA

y3 = -gamma*y(3) + VmaxTHA*v1 + VmaxTHA*Zap1*v3/(Zap1+v2);

% y4 = LA

y4 = -gamma*y(4) + VmaxTLA*v4 + VmaxTLA*Zap1*v6/(Zap1+v5);

% y5 = [Zv]

y5 = y(2)*VmaxVI*y(6)/(KmVI+y(2)+KmVI*y(5)/KiVI)/vacratio -
y(5)*VmaxVE*y(7)/(y(5)+KmVE)/vacratio - gamma*y(5) ;

% y6 = VI

y6 = -gamma*y(6) + VmaxTVI*v7 + VmaxTVI*Zap1*v9/(Zap1+v8);

% y7 = VE

y7 = -gamma*y(7) + VmaxTVE*v10 + VmaxTVE*Zap1*v12/(Zap1+v11);

dydt = [y1; y2; y3; y4; y5; y6; y7];
```

APPENDIX 3: Identification of the Pho4 phosphatase

The phosphate signal transduction pathway employs posttranslational modification to regulate the activity of the transcription factor Pho4 (see Chapter Four). Pho4 is phosphorylated on multiple serine residues by the kinase complex Pho80/Pho85 (Kaffman et al., 1994). It is believed that the dephosphorylation of Pho4 also occurs posttranslationally. While the identity of the kinase responsible for phosphorylation has been known for years, the protein responsible for catalysis of Pho4 dephosphorylation remains unidentified. I believe that the essential serine/threonine type 1 protein phosphatase Glc7 represents a plausible candidate for the Pho4 phosphatase for the reasons that follow.

First, Glc7 displays the appropriate enzymatic activity, as it is capable of dephosphorylation of phosphoserine residues. It is localized to the cytoplasm as well as the nucleus and numerous other specific subcellular compartments (Bloecher and Tatchell, 2000). It is involved in nutrient sensing and signal transduction, as it is responsible for the dephosphorylation of glycogen metabolic enzymes (Ramaswamy et al., 1998) as well as regulation of the Snf1/AMP kinase (Sanz et al., 2000). These observations provide circumstantial evidence that the identification of Glc7 as the Pho4 phosphatase would not be surprising.

Secondly, phosphorylated Pho4 is rapidly dephosphorylated when incubated with cellular extract (J.M.R., unpublished data). The responsible phosphatase activity is strongly inhibited by calyculin A and beta-glycerophosphate but not by EDTA, EGTA, sodium vanadate or sodium tungstate. (J.M.R., unpublished data). These observation suggest that the phosphatase belongs to either the type 1 (Glc7) or type 2A (Pph21, Pph22, Pph3, Sit4) in budding yeast. Simultaneous deletion of *PPH21* and *PPH22* or the

deletion of *PPH3* or *SIT4* alone did not affect the rate of relocalization of Pho4-GFP from the cytoplasm to the nucleus (J.M.R., unpublished data). Cellular extracts prepared from these deletion strains showed no defects in phosphatase activity *in vitro* (J.M.R., unpublished data). One temperature-sensitive allele of the *GLC7* gene, *glc7-5*, conferred a constitutively cytoplasmic localization on Pho4-GFP at the nonpermissive temperature of 37°C (J.M.R., unpublished data). However, this cytoplasmic localization occurred at a temperature much higher than that at which the *glc7-5* cells lost viability (J.M.R., unpublished data). At the 30°C nonpermissive temperature, Pho4-GFP was more cytoplasmic in the *glc7-5* strain than in wild-type cells (J.M.R., unpublished data). Deletions of the genes encoding all known regulatory subunits of Glc7 except for the essential nuclear subunit Sds22 were screened for defects in Pho4-GFP nuclear localization, and no strains displayed detectable defects in localization (J.M.R., unpublished data). Cellular extracts from the glc7-5 strain prepared either at the permissive or nonpermissive temperature showed no defects in Pho4 dephosphorylation *in vitro* (J.M.R., unpublished data).

Thirdly, singular deletion of the genes encoding each of the other serine/threonine phosphatases did not affect the localization from cytoplasm to nucleus of Pho4-GFP during phosphate starvation (J.M.R., unpublished data). The genes encoding the dual-specificity phosphatases were not screened. It remains possible that either a dual-specificity phosphatase or a redundant set of serine/threonine protein phosphatases are responsible for Pho4 dephosphorylation, but some evidence supports the hypothesis that Glc7 is the Pho4 phosphatase.

REFERENCES

Acar, M., Becskei, A. and van Oudenaarden, A. (2005) Enhancement of cellular memory by reducing stochastic transitions. *Nature*, **435**, 228-232.

Ahmad, K. and Henikoff, S. (2001) Modulation of a transcription factor counteracts heterochromatic gene silencing in Drosophila. *Cell*, **104**, 839-847.

Alani, E., Cao, L. and Kleckner, N. (1987) A method for gene disruption that allows repeated use of URA3 selection in the construction of multiply disrupted yeast strains. *Genetics*, **116**, 541-545.

Almer, A. and Horz, W. (1986) Nuclease hypersensitive regions with adjacent positioned nucleosomes mark the gene boundaries of the PHO5/PHO3 locus in yeast. *Embo J*, **5**, 2681-2687.

Angeli, D., Ferrell, J.E., Jr. and Sontag, E.D. (2004) Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proc Natl Acad Sci U S A*, **101**, 1822-1827.

Arkin, A., Ross, J. and McAdams, H.H. (1998) Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected Escherichia coli cells. *Genetics*, **149**, 1633-1648.

Attfield, P.V., Choi, H.Y., Veal, D.A. and Bell, P.J. (2001) Heterogeneity of stress gene expression and stress resistance among individual cells of Saccharomyces cerevisiae. *Mol Microbiol*, **40**, 1000-1008.

Auesukaree, C., Homma, T., Kaneko, Y. and Harashima, S. (2003) Transcriptional regulation of phosphate-responsive genes in low-affinity phosphate-transporter-defective mutants in Saccharomyces cerevisiae. *Biochem Biophys Res Commun*, **306**, 843-850.

Barbaric, S., Reinke, H. and Horz, W. (2003) Multiple mechanistically distinct functions of SAGA at the PHO5 promoter. *Mol Cell Biol*, **23**, 3468-3476.

Barbaric, S., Walker, J., Schmid, A., Svejstrup, J.Q. and Horz, W. (2001) Increasing the rate of chromatin remodeling and gene activation--a novel role for the histone acetyltransferase Gcn5. *Embo J*, **20**, 4944-4951.

Becskei, A., Seraphin, B. and Serrano, L. (2001) Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *Embo J*, **20**, 2528-2535.

Becskei, A. and Serrano, L. (2000) Engineering stability in gene networks by autoregulation. *Nature*, **405**, 590-593.

Berridge, M.J. (2005) Unlocking the secrets of cell signaling. *Annu Rev Physiol*, **67**, 1-21.

Biggar, S.R. and Crabtree, G.R. (2001) Cell signaling can direct either binary or graded transcriptional responses. *Embo J*, **20**, 3167-3176.

Bird, A.J., Blankman, E., Stillman, D.J., Eide, D.J. and Winge, D.R. (2004) The Zap1 transcriptional activator also acts as a repressor by binding downstream of the TATA box in ZRT2. *Embo J*, **23**, 1123-1132.

Bird, A.J., McCall, K., Kramer, M., Blankman, E., Winge, D.R. and Eide, D.J. (2003) Zinc fingers can act as Zn2+ sensors to regulate transcriptional activation domain function. *Embo J*, **22**, 5137-5146.

Bird, A.J., Zhao, H., Luo, H., Jensen, L.T., Srinivasan, C., Evans-Galea, M., Winge, D.R. and Eide, D.J. (2000) A dual role for zinc fingers in both DNA binding and zinc sensing by the Zap1 transcriptional activator. *Embo J*, **19**, 3704-3713.

Bix, M. and Locksley, R.M. (1998) Independent and epigenetic regulation of the interleukin-4 alleles in CD4+ T cells. *Science*, **281**, 1352-1354.

Blake, W.J., Kaern, M., Cantor, C.R. and Collins, J.J. (2003) Noise in eukaryotic gene expression. *Nature*, **422**, 633-637.

Bloecher, A. and Tatchell, K. (2000) Dynamic localization of protein phosphatase type 1 in the mitotic cell cycle of Saccharomyces cerevisiae. *J Cell Biol*, **149**, 125-140.

Blondel, M.O., Morvan, J., Dupre, S., Urban-Grimal, D., Haguenauer-Tsapis, R. and Volland, C. (2004) Direct sorting of the yeast uracil permease to the endosomal system is controlled by uracil binding and Rsp5p-dependent ubiquitylation. *Mol Biol Cell*, **15**, 883-895.

Boeger, H., Griesenbeck, J., Strattan, J.S. and Kornberg, R.D. (2003) Nucleosomes unfold completely at a transcriptionally active promoter. *Mol Cell*, **11**, 1587-1598.

Bun-Ya, M., Nishimura, M., Harashima, S. and Oshima, Y. (1991) The PHO84 gene of Saccharomyces cerevisiae encodes an inorganic phosphate transporter. *Mol Cell Biol*, **11**, 3229-3238.

Bun-ya, M., Shikata, K., Nakade, S., Yompakdee, C., Harashima, S. and Oshima, Y. (1996) Two new genes, PHO86 and PHO87, involved in inorganic phosphate uptake in Saccharomyces cerevisiae. *Curr Genet*, **29**, 344-351.

Campbell, R.E., Tour, O., Palmer, A.E., Steinbach, P.A., Baird, G.S., Zacharias, D.A. and Tsien, R.Y. (2002) A monomeric red fluorescent protein. *Proc Natl Acad Sci U S A*, **99**, 7877-7882.

Carroll, A.S., Bishop, A.C., DeRisi, J.L., Shokat, K.M. and O'Shea, E.K. (2001) Chemical inhibition of the Pho85 cyclin-dependent kinase reveals a role in the environmental stress response. *Proc Natl Acad Sci U S A*, **98**, 12578-12583.

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

Chess, A., Simon, I., Cedar, H. and Axel, R. (1994) Allelic inactivation regulates olfactory receptor gene expression. *Cell*, **78**, 823-834.

Cohen, A., Perzov, N., Nelson, H. and Nelson, N. (1999) A novel family of yeast chaperons involved in the distribution of V-ATPase and other membrane proteins. *J Biol Chem*, **274**, 26885-26893.

Cormack, B.P., Bertram, G., Egerton, M., Gow, N.A., Falkow, S. and Brown, A.J. (1997) Yeast-enhanced green fluorescent protein (yEGFP)a reporter of gene expression in Candida albicans. *Microbiology*, **143 (Pt 2)**, 303-311.

Crameri, A., Whitehorn, E.A., Tate, E. and Stemmer, W.P. (1996) Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol*, **14**, 315-319.

Dupre, S., Urban-Grimal, D. and Haguenauer-Tsapis, R. (2004) Ubiquitin and endocytic internalization in yeast and animal cells. *Biochim Biophys Acta*, **1695**, 89-111.

Ebbert, R., Birkmann, A. and Schuller, H.J. (1999) The product of the SNF2/SWI2 paralogue INO80 of Saccharomyces cerevisiae required for efficient expression of various yeast structural genes is part of a high-molecular-weight protein complex. *Mol Microbiol*, **32**, 741-751.

Elowitz, M.B. and Leibler, S. (2000) A synthetic oscillatory network of transcriptional regulators. *Nature*, **403**, 335-338.

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

Fascher, K.D., Schmitz, J. and Horz, W. (1990) Role of trans-activating proteins in the generation of active chromatin at the PHO5 promoter in S. cerevisiae. *Embo J*, **9**, 2523-2528.

Fascher, K.D., Schmitz, J. and Horz, W. (1993) Structural and functional requirements for the chromatin transition at the PHO5 promoter in Saccharomyces cerevisiae upon PHO5 activation. *J Mol Biol*, **231**, 658-667.

Felice, M.R., De Domenico, I., Li, L., Ward, D.M., Bartok, B., Musci, G. and Kaplan, J. (2005) Post-transcriptional Regulation of the Yeast High Affinity Iron Transport System. *J Biol Chem*, **280**, 22181-22190.

Femino, A.M., Fay, F.S., Fogarty, K. and Singer, R.H. (1998) Visualization of single RNA transcripts in situ. *Science*, **280**, 585-590.

Ferrell, J.E., Jr. (2002) Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr Opin Cell Biol*, **14**, 140-148.

Ferrell, J.E. and Xiong, W. (2001) Bistability in cell signaling: How to make continuous processes discontinuous, and reversible processes irreversible. *Chaos*, **11**, 227-236.

Fiering, S., Whitelaw, E. and Martin, D.I. (2000) To be or not to be active: the stochastic nature of enhancer action. *Bioessays*, **22**, 381-387.

Finnegan, E.J. (2002) Epialleles - a source of random variation in times of stress. *Curr Opin Plant Biol*, **5**, 101-106.

Flick, J.S. and Thorner, J. (1998) An essential function of a phosphoinositide-specific phospholipase C is relieved by inhibition of a cyclin-dependent protein kinase in the yeast Saccharomyces cerevisiae. *Genetics*, **148**, 33-47.

Fraser, H.B., Hirsh, A.E., Giaever, G., Kumm, J. and Eisen, M.B. (2004) Noise minimization in eukaryotic gene expression. *PLoS Biol*, **2**, advance online publication.

Giardina, C. and Lis, J.T. (1993) DNA melting on yeast RNA polymerase II promoters. *Science*, **261**, 759-762.

Gibson, M.A. and Bruck, J. (2000) Efficient exact stochastic simulation of chemical systems with many species and many channels. *Journal of Physical Chemistry A*, **104**, 1876-1889.

Gillespie, D.T. (1977) Exact stochastic simulation of coupled kinetic reactions. *J. Phys. Chem.*, **81**, 2340-2361.

Gitan, R.S. and Eide, D.J. (2000) Zinc-regulated ubiquitin conjugation signals endocytosis of the yeast ZRT1 zinc transporter. *Biochem J*, **346 Pt 2**, 329-336.

Gitan, R.S., Luo, H., Rodgers, J., Broderius, M. and Eide, D. (1998) Zinc-induced inactivation of the yeast ZRT1 zinc transporter occurs through endocytosis and vacuolar degradation. *J Biol Chem*, **273**, 28617-28624.

Gitan, R.S., Shababi, M., Kramer, M. and Eide, D.J. (2003) A cytosolic domain of the yeast Zrt1 zinc transporter is required for its post-translational inactivation in response to zinc and cadmium. *J Biol Chem*, **278**, 39558-39564.

Griesbeck, O., Baird, G.S., Campbell, R.E., Zacharias, D.A. and Tsien, R.Y. (2001) Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. *J Biol Chem*, **276**, 29188-29194.

Hansson, B. and Westerberg, L. (2002) On the correlation between heterozygosity and fitness in natural populations. *Mol Ecol*, **11**, 2467-2474.

Hasty, J., Pradines, J., Dolnik, M. and Collins, J.J. (2000) Noise-based switches and amplifiers for gene expression. *Proc Natl Acad Sci U S A*, **97**, 2075-2080.

Heim, R. and Tsien, R.Y. (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr Biol*, **6**, 178-182.

Holland, M.J. (2002) Transcript abundance in yeast varies over six orders of magnitude. *J Biol Chem*, **277**, 14363-14366.

Hooshangi, S., Thiberge, S. and Weiss, R. (2005) Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc Natl Acad Sci U S A*, **102**, 3581-3586.

Huang, S. and O'Shea E, K. (2005) A Systematic High-Throughput Screen of a Yeast Deletion Collection for Mutants Defective in PHO5 Regulation. *Genetics*, **169**, 1859-1871.

Isaacs, F.J., Hasty, J., Cantor, C.R. and Collins, J.J. (2003) Prediction and measurement of an autoregulatory genetic module. *Proc Natl Acad Sci U S A*, **100**, 7714-7719.

Johnston, M. (1987) A model fungal gene regulatory mechanism: the GAL genes of Saccharomyces cerevisiae. *Microbiol Rev*, **51**, 458-476.

Johnston, M. and Kim, J.H. (2005) Glucose as a hormone: receptor-mediated glucose sensing in the yeast Saccharomyces cerevisiae. *Biochem Soc Trans*, **33**, 247-252.

Kaffman, A., Herskowitz, I., Tjian, R. and O'Shea, E.K. (1994) Phosphorylation of the transcription factor PHO4 by a cyclin-CDK complex, PHO80-PHO85. *Science*, **263**, 1153-1156.

Katchman, B.J. and Van Wazer, J.R. (1954) The soluble and insoluble polyphosphates of yeast. *Biochim Biophys Acta*, **14**, 445-446.

Kellis, M., Birren, B.W. and Lander, E.S. (2004) Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae. *Nature*, **428**, 617-624.

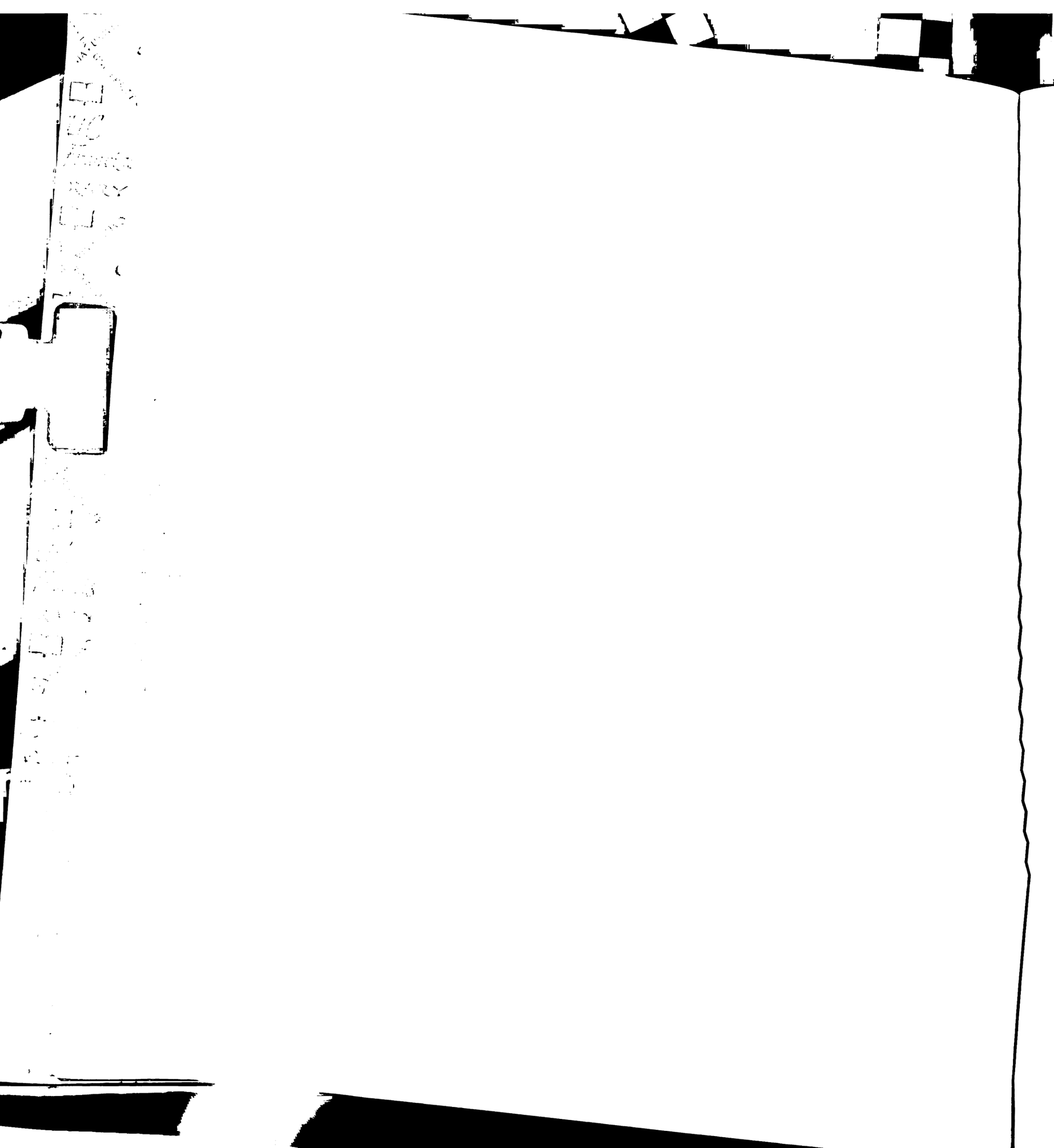
Kemkemer, R., Schrank, S., Vogel, W., Gruler, H. and Kaufmann, D. (2002) Increased noise as an effect of haploinsufficiency of the tumor-suppressor gene neurofibromatosis type 1 in vitro. *Proc Natl Acad Sci U S A*, **99**, 13783-13788.

Kepler, T.B. and Elston, T.C. (2001) Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys J*, **81**, 3116-3136.

Kierzek, A.M., Zaim, J. and Zielenkiewicz, P. (2001) The effect of transcription and translation initiation frequencies on the stochastic fluctuations in prokaryotic gene expression. *J Biol Chem*, **276**, 8165-8172.

Ko, M.S. (1991) A stochastic model for gene induction. *J Theor Biol*, **153**, 181-194.

Komeili, A. and O'Shea, E.K. (1999) Roles of phosphorylation sites in regulating activity of the transcription factor Pho4. *Science*, **284**, 977-980.



Koshland, D.E., Jr., Goldbeter, A. and Stock, J.B. (1982) Amplification and adaptation in regulatory and sensory systems. *Science*, **217**, 220-225.

Larschan, E. and Winston, F. (2001) The S. cerevisiae SAGA complex functions in vivo as a coactivator for transcriptional activation by Gal4. *Genes Dev*, **15**, 1946-1956.

Lau, W.T., Howson, R.W., Malkus, P., Schekman, R. and O'Shea, E.K. (2000) Pho86p, an endoplasmic reticulum (ER) resident protein in Saccharomyces cerevisiae, is required for ER exit of the high-affinity phosphate transporter Pho84p. *Proc Natl Acad Sci U S A*, **97**, 1107-1112.

Lemaire, K., Van de Velde, S., Van Dijck, P. and Thevelein, J.M. (2004) Glucose and sucrose act as agonist and mannose as antagonist ligands of the G protein-coupled receptor Gpr1 in the yeast Saccharomyces cerevisiae. *Mol Cell*, **16**, 293-299.

Lenburg, M.E. and O'Shea, E.K. (1996) Signaling phosphate starvation. *Trends Biochem Sci*, **21**, 383-387.

Li, L. and Kaplan, J. (1998) Defects in the yeast high affinity iron transport system result in increased metal sensitivity because of the increased expression of transporters with a broad transition metal specificity. *J Biol Chem*, **273**, 22181-22187.

Liu, X., Ma, L., Zhang, J.F. and Lu, Y.T. (2004) Determination of single-cell gene expression in Arabidopsis by capillary electrophoresis with laser induced fluorescence detection. *J Chromatogr B Analyt Technol Biomed Life Sci*, **808**, 241-247.

Lyons, T.J., Gasch, A.P., Gaither, L.A., Botstein, D., Brown, P.O. and Eide, D.J. (2000) Genome-wide characterization of the Zap1p zinc-responsive regulon in yeast. *Proc Natl Acad Sci U S A*, **97**, 7957-7962.

MacDiarmid, C.W., Gaither, L.A. and Eide, D. (2000) Zinc transporters that regulate vacuolar zinc storage in Saccharomyces cerevisiae. *Embo J*, **19**, 2845-2855.

MacDiarmid, C.W., Milanick, M.A. and Eide, D.J. (2002) Biochemical properties of vacuolar zinc transport systems of Saccharomyces cerevisiae. *J Biol Chem*, **277**, 39187-39194.

MacDiarmid, C.W., Milanick, M.A. and Eide, D.J. (2003) Induction of the ZRC1 metal tolerance gene in zinc-limited yeast confers resistance to zinc shock. *J Biol Chem*, **278**, 15065-15072.

Magbanua, J.P., Ogawa, N., Harashima, S. and Oshima, Y. (1997) The transcriptional activators of the PHO regulon, Pho4p and Pho2p, interact directly with each other and with components of the basal transcription machinery in Saccharomyces cerevisiae. *J Biochem (Tokyo)*, **121**, 1182-1189.

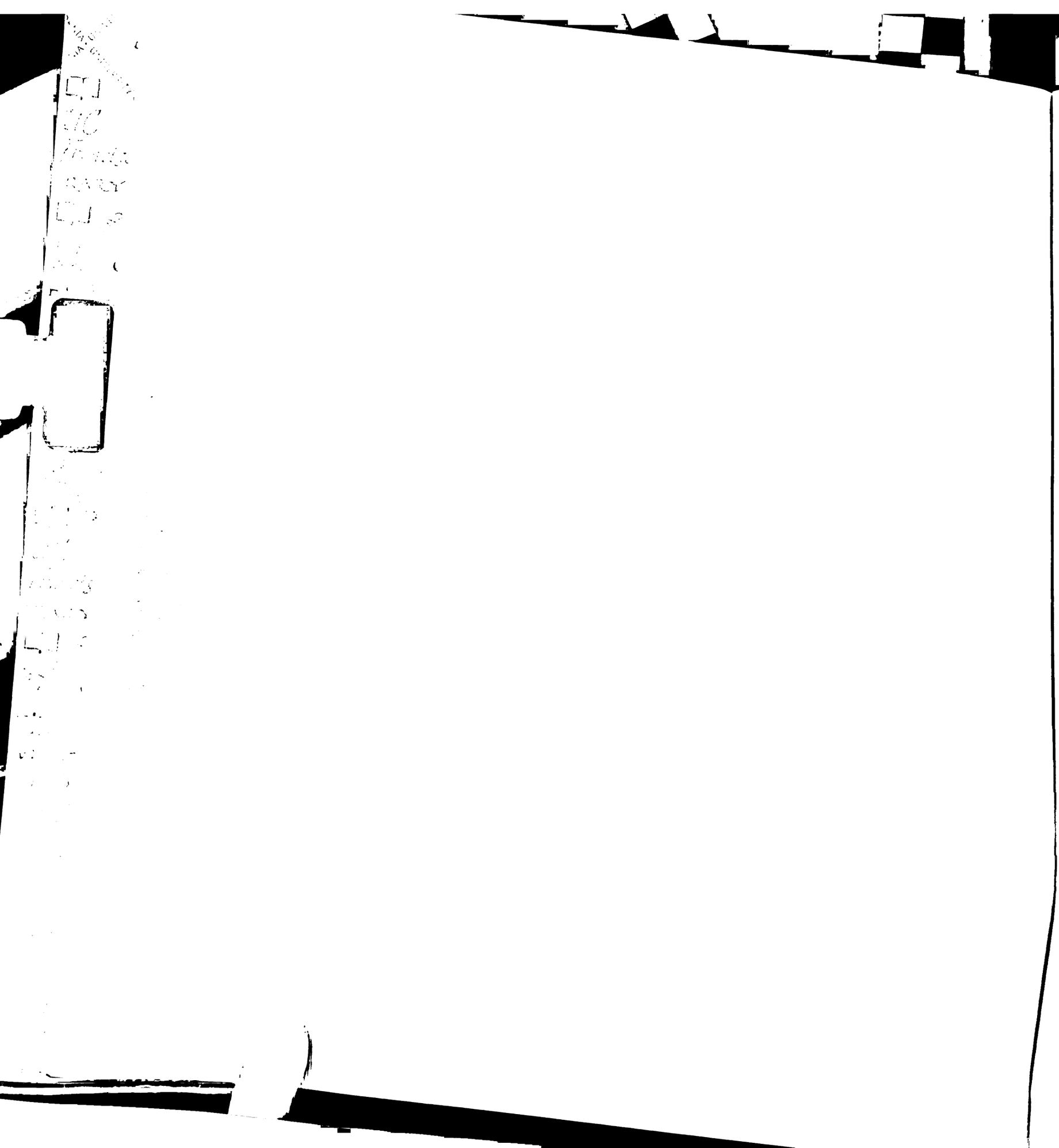
Magee, J.A., Abdulkadir, S.A. and Milbrandt, J. (2003) Haploinsufficiency at the Nkx3.1 locus. A paradigm for stochastic, dosage-sensitive gene regulation during tumor initiation. *Cancer Cell*, **3**, 273-283.

McAdams, H.H. and Arkin, A. (1997) Stochastic mechanisms in gene expression. *Proc Natl Acad Sci U S A*, **94**, 814-819.

McAdams, H.H. and Arkin, A. (1999) It's a noisy business! Genetic regulation at the nanomolar scale. *Trends Genet*, **15**, 65-69.

Meyers, L.A. and Bull, J.J. (2002) Fighting change with change: adaptive variation in an uncertain world. *Trends in Ecology & Evolution*, **17**, 551-557.

Mihalcescu, I., Hsing, W. and Leibler, S. (2004) Resilient circadian oscillator revealed in individual cyanobacteria. *Nature*, **430**, 81-85.
Miyasaka, H. (1999) The positive relationship between codon usage bias and translation initiation AUG context in Saccharomyces cerevisiae. *Yeast*, **15**, 633-637.
Mostoslavsky, R., Singh, N., Tenzen, T., Goldmit, M., Gabay, C., Elizur, S., Qi, P., Reubinoff, B.E., Chess, A., Cedar, H. and Bergman, Y. (2001) Asynchronous replication and allelic exclusion in the immune system. *Nature*, **414**, 221-225.
Nagai, T., Ibata, K., Park, E.S., Kubota, M., Mikoshiba, K. and Miyawaki, A. (2002) A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat Biotechnol*, **20**, 87-90.
Ng, H.H., Robert, F., Young, R.A. and Struhl, K. (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell*, **11**, 709-719.
Ogawa, N., DeRisi, J. and Brown, P.O. (2000) New components of a system for phosphate accumulation and polyphosphate metabolism in Saccharomyces cerevisiae revealed by genomic expression analysis. *Mol Biol Cell*, **11**, 4309-4321.
Ogawa, N., Noguchi, K., Yamashita, Y., Yasuhara, T., Hayashi, N., Yoshida, K. and Oshima, Y. (1993) Promoter analysis of the PHO81 gene encoding a 134 kDa protein bearing ankyrin repeats in the phosphatase regulon of Saccharomyces cerevisiae. *Mol Gen Genet*, **238**, 444-454.
Oshima, Y. (1997) The phosphatase system in Saccharomyces cerevisiae. *Genes Genet Syst*, **72**, 323-334.
Ozbudak, E.M., Thattai, M., Kurtser, I., Grossman, A.D. and van Oudenaarden, A. (2002) Regulation of noise in the expression of a single gene. *Nat Genet*, **31**, 69-73.
Ozbudak, E.M., Thattai, M., Lim, H.N., Shraiman, B.I. and van Oudenaarden, A. (2004) Multistability in the lactose utilization network of Escherichia coli. *Nature*, **427**, 737-740.
Pastinen, T., Sladek, R., Gurd, S., Sammak, A., Ge, B., Lepage, P., Lavergne, K., Villeneuve, A., Gaudin, T., Brandstrom, H., Beck, A., Verner, A., Kingsley, J., Harmsen, E., Labuda, D., Morgan, K., Vohl, M.C., Naumova, A.K., Sinnett, D. and Hudson, T.J. (2004) A survey of genetic and epigenetic variation affecting human gene expression. *Physiol Genomics*, **16**, 184-193.
Patton, C., Thompson, S. and Epel, D. (2004) Some precautions in using chelators to buffer metals in biological solutions. *Cell Calcium*, **35**, 427-431.
Paulsson, J. (2004) Summing up the noise in gene networks. *Nature*, **427**, 415-418.
Paulsson, J., Berg, O.G. and Ehrenberg, M. (2000) Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proc Natl Acad Sci U S A*, **97**, 7148-7153.
Peccoud, J. and Ycart, B. (1995) Markovian modeling of gene product synthesis. *Theor. Pop. Biol.*, **48**, 222-234.
Pedraza, J.M. and van Oudenaarden, A. (2005) Noise propagation in gene networks. *Science*, **307**, 1965-1969.
Piggot, P.J. and Hilbert, D.W. (2004) Sporulation of Bacillus subtilis. *Curr Opin Microbiol*, **7**, 579-586.



Pigliucci, M. and Murren, C.J. (2003) Perspective: Genetic assimilation and a possible evolutionary paradox: Can macroevolution sometimes be so fast as to pass us by? *Evolution*, **57**, 1455-1464.

Poulsen, P., Wu, B., Gaber, R.F., Ottow, K., Andersen, H.A. and Kielland-Brandt, M.C. (2005) Amino acid sensing by Ssy1. *Biochem Soc Trans*, **33**, 261-264.

Queitsch, C., Sangster, T.A. and Lindquist, S. (2002) Hsp90 as a capacitor of phenotypic variation. *Nature*, **417**, 618-624.

Rakyan, V.K., Preis, J., Morgan, H.D. and Whitelaw, E. (2001) The marks, mechanisms and memory of epigenetic states in mammals. *Biochem J*, **356**, 1-10.

Ramaswamy, N.T., Li, L., Khalil, M. and Cannon, J.F. (1998) Regulation of yeast glycogen metabolism and sporulation by Glc7p protein phosphatase. *Genetics*, **149**, 57-72.

Rao, C.V., Wolf, D.M. and Arkin, A.P. (2002) Control, exploitation and tolerance of intracellular noise. *Nature*, **420**, 231-237.

Raser, J.M. and O'Shea, E.K. (2004) Control of stochasticity in eukaryotic gene expression. *Science*, **304**, 1811-1814.

Riviere, I., Sunshine, M.J. and Littman, D.R. (1998) Regulation of IL-4 expression by activation of individual alleles. *Immunity*, **9**, 217-228.

Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S. and Elowitz, M.B. (2005) Gene regulation at the single-cell level. *Science*, **307**, 1962-1965.

Samoilov, M., Plyasunov, S. and Arkin, A.P. (2005) Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proc Natl Acad Sci U S A*, **102**, 2310-2315.

Santisteban, M.S., Kalashnikova, T. and Smith, M.M. (2000) Histone H2A.Z regulats transcription and is partially redundant with nucleosome remodeling complexes. *Cell*, **103**, 411-422.

Sanz, P., Alms, G.R., Haystead, T.A. and Carlson, M. (2000) Regulatory interactions between the Reg1-Glc7 protein phosphatase and the Snf1 protein kinase. *Mol Cell Biol*, **20**, 1321-1328.

Schneider, K.R., Smith, R.L. and O'Shea, E.K. (1994) Phosphate-regulated inactivation of the kinase PHO80-PHO85 by the CDK inhibitor PHO81. *Science*, **266**, 122-126.

Schneper, L., Duvel, K. and Broach, J.R. (2004) Sense and sensibility: nutritional response and signal integration in yeast. *Curr Opin Microbiol*, **7**, 624-630.

Schroedinger, E. (1944) *What is Life?* Cambridge Univ. Press, Cambridge.

Serizawa, S., Miyamichi, K., Nakatani, H., Suzuki, M., Saito, M., Yoshihara, Y. and Sakano, H. (2003) Negative feedback regulation ensures the one receptor-one olfactory neuron rule in mouse. *Science*, **302**, 2088-2094.

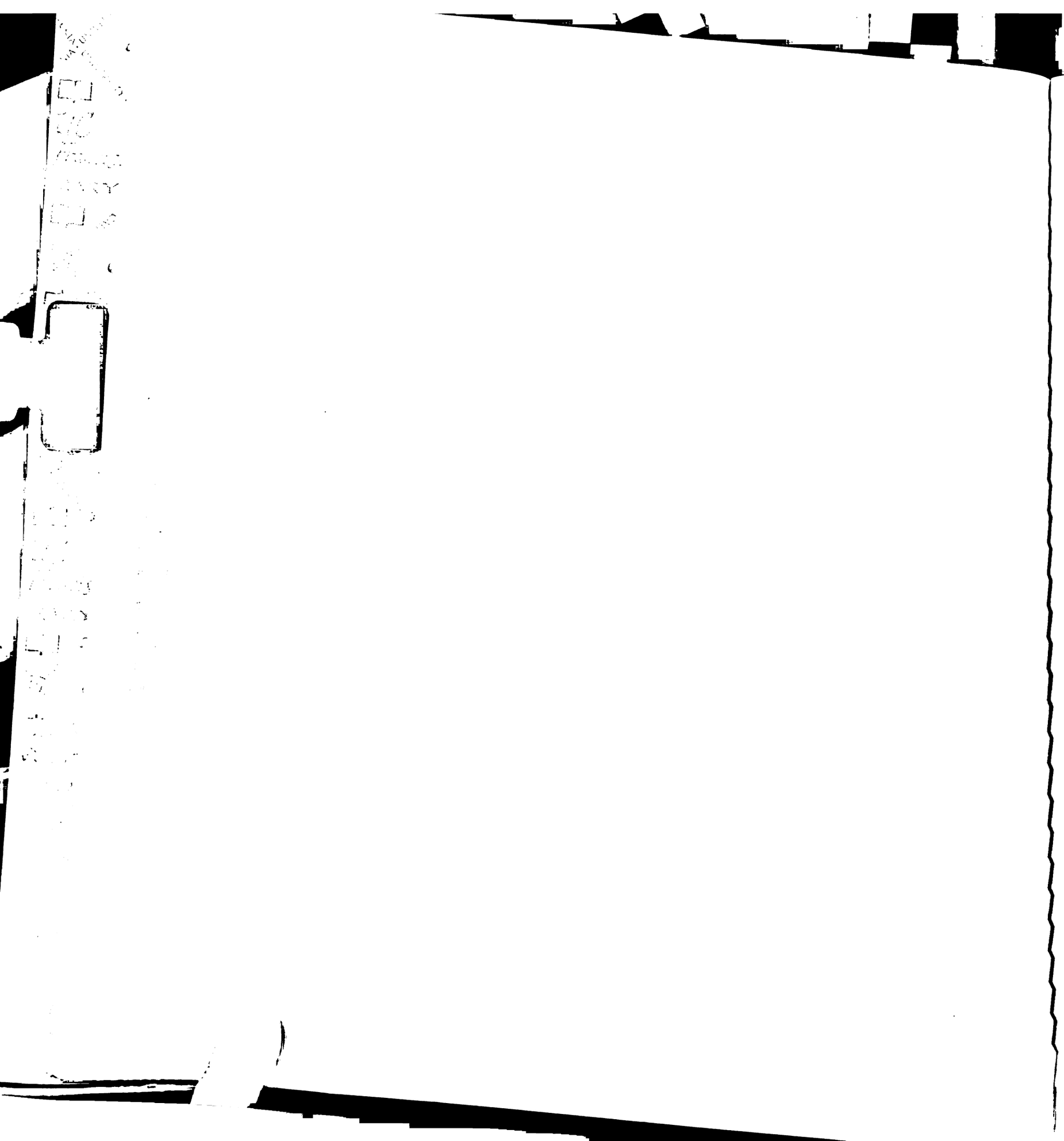
Serizawa, S., Miyamichi, K. and Sakano, H. (2004) One neuron-one receptor rule in the mouse olfactory system. *Trends Genet*, **20**, 648-653.

Siemering, K.R., Golbik, R., Sever, R. and Haseloff, J. (1996) Mutations that suppress the thermosensitivity of green fluorescent protein. *Curr Biol*, **6**, 1653-1663.

Springer, M., Wykoff, D.D., Miller, N. and O'Shea, E.K. (2003) Partially Phosphorylated Pho4 Activates Transcription of a Subset of Phosphate-Responsive Genes. *PLoS Biol*, **1**, E28.

Steger, D.J., Haswell, E.S., Miller, A.L., Wente, S.R. and O'Shea, E.K. (2003) Regulation of chromatin remodeling by inositol polyphosphates. *Science*, **299**, 114-116.

Steinman, G. (2001) Mechanisms of twinning. I. Effect of environmental diversity on genetic expression in monozygotic multifetal pregnancies. *J Reprod Med*, **46**, 467-472.
Svaren, J. and Horz, W. (1997) Transcription factors vs nucleosomes: regulation of the PHO5 promoter in yeast. *Trends Biochem Sci*, **22**, 93-97.
Swain, P.S., Elowitz, M.B. and Siggia, E.D. (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci U S A*, **99**, 12795-12800.
Thattai, M. and van Oudenaarden, A. (2001) Intrinsic noise in gene regulatory networks. *Proc Natl Acad Sci U S A*, **98**, 8614-8619.
Thattai, M. and van Oudenaarden, A. (2004) Stochastic gene expression in fluctuating environments. *Genetics*, **167**, 523-530.
To, E.A., Ueda, Y., Kakimoto, S.I. and Oshima, Y. (1973) Isolation and characterization of acid phosphatase mutants in Saccharomyces cerevisiae. *J Bacteriol*, **113**, 727-738.
Toh-e, A., Inouye, S. and Oshima, Y. (1981) Structure and function of the PHO82-pho4 locus controlling the synthesis of repressible acid phosphatase of Saccharomyces cerevisiae. *J Bacteriol*, **145**, 221-232.
Tomita, J., Nakajima, M., Kondo, T. and Iwasaki, H. (2005) No transcription-translation feedback in circadian rhythm of KaiC phosphorylation. *Science*, **307**, 251-254.
Ueda, Y., To, E.A. and Oshima, Y. (1975) Isolation and characterization of recessive, constitutive mutations for repressible acid phosphatase synthesis in Saccharomyces cerevisiae. *J Bacteriol*, **122**, 911-922.
van Kampen, N.G. (1992) *Stochastic processes in physics and chemistry*. North-Holland, Amsterdam.
Vanfleteren, J.R. and Braeckman, B.P. (1999) Mechanisms of life span determination in Caenorhabditis elegans. *Neurobiol Aging*, **20**, 487-502.
Vilar, J.M., Kueh, H.Y., Barkai, N. and Leibler, S. (2002) Mechanisms of noise-resistance in genetic oscillators. *Proc Natl Acad Sci U S A*, **99**, 5988-5992.
Vogel, K., Horz, W. and Hinnen, A. (1989) The two positively acting regulatory proteins PHO2 and PHO4 physically interact with PHO5 upstream activation regions. *Mol Cell Biol*, **9**, 2050-2057.
Waddington, C.H. (1959) Canalization of development and genetic assimilation of acquired characters. *Nature*, **183**, 1654-1655.
Waters, B.M. and Eide, D.J. (2002) Combinatorial control of yeast FET4 gene expression by iron, zinc, and oxygen. *J Biol Chem*, **277**, 33749-33757.
Wolf, D.M., Vazirani, V.V. and Arkin, A.P. (2005) Diversity in times of adversity: probabilistic strategies in microbial survival games. *J Theor Biol*, **234**, 227-253.
Wu, W.H. and Hampsey, M. (1999) An activation-specific role for transcription factor TFIIB in vivo. *Proc Natl Acad Sci U S A*, **96**, 2764-2769.
Wykoff, D.D. and O'Shea, E.K. (2001) Phosphate transport and sensing in Saccharomyces cerevisiae. *Genetics*, **159**, 1491-1499.
Yuan, D.S. (2000) Zinc-regulated genes in Saccharomyces cerevisiae revealed by transposon tagging. *Genetics*, **156**, 45-58.
Zhao, H., Butler, E., Rodgers, J., Spizzo, T., Duesterhoeft, S. and Eide, D. (1998) Regulation of zinc homeostasis in yeast by binding of the ZAP1 transcriptional activator to zinc-responsive promoter elements. *J Biol Chem*, **273**, 28713-28720.



Zhao, H. and Eide, D. (1996a) The yeast ZRT1 gene encodes the zinc transporter protein of a high-affinity uptake system induced by zinc limitation. *Proc Natl Acad Sci U S A*, **93**, 2454-2458.

Zhao, H. and Eide, D. (1996b) The ZRT2 gene encodes the low affinity zinc transporter in Saccharomyces cerevisiae. *J Biol Chem*, **271**, 23203-23210.

Zhao, H. and Eide, D.J. (1997) Zap1p, a metalloregulatory protein involved in zinc-responsive transcriptional regulation in Saccharomyces cerevisiae. *Mol Cell Biol*, **17**, 5044-5052.