Analysis of the transcriptional program governing meiosis and gametogenesis in yeast and mammals

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Ulrich Schlecht

Aus Tokyo (Japan)

Basel, 2006

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Michael Primig, Prof. Peter Philippsen und Prof. Urs Meyer.

Basel, den 10.07.2006

Prof. Hans-Jakob Wirz

Dekan

'Everything in its right place' - Radiohead

Table of contents

Abstract

During meiosis a competent diploid cell replicates its DNA once and then undergoes two consecutive divisions followed by haploid gamete differentiation. Important aspects of meiotic development that distinguish it from mitotic growth include a highly increased rate of recombination, formation of the synaptonemal complex that aligns the homologous chromosomes, as well as separation of the homologues and sister chromatids during meiosis I and II without an intervening S-phase. Budding yeast is an excellent model organism to study meiosis and gametogenesis and accordingly, to date it belongs to the best studied eukaryotic systems in this context. Knowledge coming from these studies has provided important insights into meiotic development in higher eukaryotes. This was possible because sporulation in yeast and spermatogenesis in higher eukaryotes are analogous developmental pathways that involve conserved genes.

For budding yeast a huge amount of data from numerous genome-scale studies on gene expression and deletion phenotypes of meiotic development and sporulation are available. In contrast, mammalian gametogenesis has not been studied on a large-scale until recently. It was unclear if an expression profiling study using germ cells and testicular somatic control cells that underwent lengthy purification procedures would yield interpretable results. We have therefore carried out a pioneering expression profiling study of male germ cells from *Rattus norvegicus* using Affymetrix U34A and B GeneChips. This work resulted in the first comprehensive large-scale expression profiling analysis of mammalian male germ cells undergoing mitotic growth, meiosis and gametogenesis. We have identified 1268 differentially expressed genes in germ cells at different developmental stages, which were organized into four distinct expression clusters that reflect somatic, mitotic, meiotic and post-meiotic cell types. This included 293 yet uncharacterized transcripts whose expression pattern suggests that they are involved in spermatogenesis and fertility. A group of 121 transcripts were only expressed in meiotic (spermatocytes) and post-meiotic germ cells (round spermatids) but not in dividing germ cells (spermatogonia), Sertoli cells or two somatic control tissues (brain and skeletal muscle). Functional analysis reveals that most of the known genes in this group fulfill essential functions during meiosis, spermiogenesis (the process of sperm maturation) and fertility. Therefore it is highly possible that some of the ∼30 uncharacterized transcripts in this group also contribute to these processes. A web-accessible database (called reXbase, which was later on integrated into GermOnline) has been developed for our expression profiling study of mammalian male meiosis, which summarizes annotation information and shows a graphical display of expression profiles of every gene covered in our study.

In the budding yeast *Saccharomyces cerevisiae* entry into meiosis and subsequent progression through sporulation and gametogenesis are driven by a highly regulated transcriptional program activated by signal pathways responding to nutritional and cell-type cues. Abf1p, which is a general transcription factor, has previously been demonstrated to participate in the induction of numerous mitotic as well as early and middle meiotic genes. In the current study we have addressed the question how Abf1p transcriptionally coordinates mitotic growth and meiotic development on a genome-wide level. Because *ABF1* is an essential gene we used the temperature-sensitive allele *abf1-1*. A phenotypical analysis of mutant cells revealed that *ABF1* plays an important role in cell separation during mitosis, meiotic development, and spore formation. In order to identify genes whose expression depends on Abf1p in growing and sporulating cells we have performed expression profiling experiments using Affymetrix S98 GeneChips comparing wild-type and *abf1-1* mutant cells at both permissive and restrictive temperature. We have identified 504 genes whose normal expression depends on functional *ABF1*. By combining the expression profiling data with data from genome-wide DNA binding assays (ChIP-CHIP) and *in silico* predictions of potential Abf1p-binding sites in the yeast genome, we were able to define direct target genes. Expression of these genes decreases in the absence of functional *ABF1* and whose promotors are bound by Abf1p and/or contain a predicted binding site.

Among 352 such *bona fide* direct target genes we found many involved in ribosome biogenesis, translation, vegetative growth and meiotic developement and therefore could account for the observed growth and sporulation defects of *abf1-1* mutant cells. Furthermore, the fact that two members of the septin family (*CDC3* and *CDC10*) were found to be direct target genes suggests a novel role for Abf1p in cytokinesis. This was further substantiated by the observation that chitin localization and septin ring formation are perturbed in *abf1-1* mutant cells.

General introduction

Meiosis and gametogenesis

Meiosis is the developmental pathway by which sexually reproducing diploid organisms generate haploid germ cells. It plays an important role in promoting genetic diversity, removing recessive lethal mutations from populations and maintaining constant chromosome numbers from one generation to the next. Meiosis can be regarded as a specialized form of the mitotic cell cycle with three major differences. First, following pre-meiotic DNA synthesis, high levels of recombination occur between homologous chromosomes during meiotic prophase. Second, sister-centromere cohesion and coorientation are coupled to separation and independent assortment of parental chromosomes. Third, two consecutive nuclear divisions occur without intervening DNA replication. During the reductional division (meiosis I) homologous chromosomes separate. During the equational division (meiosis II), sister chromatid segregation occurs. After these specialized nuclear divisions morphogenic differentiation events take place that lead to the formation of mature functional gametes (spores, ova, sperm) containing rearranged genomes with half the chromosome number of the original cell.

Meiotic development in mammals

During mammalian male gametogenesis initiated after birth, spermatogonial stem cells grow mitotically before they enter the meiotic differentiation pathway. At the onset of puberty they develop into primary and secondary spermatocytes undergoing the first and second meiotic divisions, respectively, to become haploid round spermatids. These cells subsequently elongate and differentiate into mature sperm during spermiogenesis. The process depends on somatic testicular Sertoli and Leydig cells that control germ cell development (Zhao and Garbers, 2002; Sharpe *et al.*, 2003).

Unlike the situation in budding yeast for which genome-scale data for meiotic development has been published on gene expression (Chu *et al.*, 1998; Primig *et al.*, 2000) and deletion phenotypes (Deutschbauer *et al.*, 2002; Enyenihi and Saunders, 2003; Rabitsch *et al.*, 2001) mammalian gametogenesis has not been studied on a genomic scale until recently. The identification of meiosis- and germ-cell-specific transcripts in multicellular organisms is a complex task because gonads contain different types of cells, only some of which are germ cells. However, it is possible to obtain informative expression data using microarrays as recently shown by different studies using various mammalian model systems (Pang *et al.*, 2003; Schultz *et al.*, 2003; Schlecht *et al.*, 2004). These large-scale expression analyses of mammalian male reproductive tissue were based on profiling the process of spermatogenesis using total testis samples or germ cell populations purified from animals at different stages of sexual maturity (Schultz *et al.*, 2003; Almstrup *et al.*, 2004; Ellis *et al.*, 2004; Shima *et al.*, 2004; Schlecht *et al.*, 2004). Understanding the genetic program controlling the mitotic and meiotic divisions of the germ line provides insight into infertility and may help to develop novel contraceptives.

Transcriptional control of meiosis in budding yeast

Saccharomyces cerevisiae is an excellent model system for the genetic and genomic analysis of mitotic growth, meiosis, and gamete formation (sporulation) for several reasons: First, since it is a unicellular organism it is possible to synchronize and harvest large populations at similar stages of development. Second, induction of meiosis can easily be induced or abolished by changes of environmental conditions, such as nutrients or temperature. Third, many different biochemical and genetic techniques are at hand to manipulate budding yeast, and therefore it is easy to recover mutations that affect the process. Numerous yeast genes important for meiotic development and sporulation have been isolated and characterized over the past 30 years. They were organized into five categories according to the timing of their induction: very early, early, middle, mid-late and late (Kupiec *et al.*, 1997). Later, a number of studies defined the transcriptome of meiosis and spore development with spotted and high-density oligonucleotide microarrays. GeneChip analysis of two different genetic backgrounds including wild type strains and meiosis defi-

cient controls yielded nine hundred core genes that were demonstrated to be meiotically regulated in a strain-independent manner, including hundreds that were also expressed during mitotic growth. The classical expression categories were further refined into seven expression clusters (Primig *et al.*, 2000). Using a similar approach, about 500 genes were found to be up-regulated during sporulation in an analysis using PCR microarrays (Chu *et al.*, 1998). In both studies it emerged that in many cases timing of induction coincided with time of function, notably for genes involved in recombination (*SPO11*, *REC102*), pachytene checkpoint function (*PCH2*), sister chromatid separation (*REC8* or *SPO69*), regulation of meiosis I (*SPO13*), induction of middle meiotic genes (*NDT80*) and spore wall formation (*DIT1*).

Both entry into meiosis and subsequent progression are driven by a highly regulated transcriptional program activated by signal pathways responding to nutritional and cell type cues (Vershon and Pierce, 2000; Kassir *et al.*, 2003). To date only three site-specifically DNA-binding transcription factors (Abf1p, Ume6p/Ime1p, and Ndt80p) involved in meiotic gene induction are described (Gailus-Durner *et al.*, 1996; Ozsarac *et al.*, 1997; Pierce *et al.*, 1998; Bowdish *et al.*, 1995; Strich *et al.*, 1994; Chu and Herskowitz, 1998; Hepworth *et al.*, 1998).

Ume6p was first identified in a screen for mutants that expressed meiosis-specific genes during mitotic growth (Unscheduled Meiotic Expression) and later demonstrated to encode a key regulator of nitrogen repression and meiosis (Bowdish *et al.*, 1995; Strich *et al.*, 1994). Ume6p exerts its function through direct interaction with a conserved histone deacteylase complex that contains Rpd3p, Sin3p, Isw2p and the meiotic activator Ime1p (Kadosh and Struhl, 1998; Goldmark *et al.*, 2000; Washburn and Esposito, 2001). An expression profiling study of ume6Δ deletion mutants compared to a wild-type strain revealed significant mitotic derepression of almost 80 genes (Williams *et al.*, 2002). Among those whose functions are known, most participate in carbon and nitrogen metabolism or sporulation and approximately half of the genes thought to be directly regulated by Ume6p are induced during early stages of meiosis and spore development. Importantly, the majority of them are essential for various steps in the meiotic pathway including recombination (*DMC1*, *SPO11*), formation of the synaptonemal complex (*HOP1*, *HOP2*, *ZIP1*, *ZIP2*), the pachytene checkpoint (*PCH2*) control of meiosis I (*SPO13*) and prospore development (*MPC54*, *SMA1*).

Ndt80p is a transcription factor required for expression of the middle sporulation genes (Non-DiTyrosine). It activates target genes by binding to a DNA sequence found upstream of several of the middle sporulation genes, termed the Midddle Sporulation Element (MSE). This element has been shown to be both necessary and sufficient for proper meiosis-specific expression of a heterologous reporter (Hepworth *et al.*, 1995; Ozsarac *et al.*, 1997). During mitotic growth and early meiosis MSEs are kept inactive by the transcriptional repressor Sum1p (Xie *et al.*, 1999; Pierce *et al.*, 1998). Gene expression profiling of an ndt80Δ deletion mutant revealed that the magnitude of induction of most middle genes is reduced to about one-third of the level in wild-type strains. Ectopical over-expression of *NDT80* in turn induced 200 genes more than threefold in vegetative cells, half of which were also expressed according to the middle induction pattern during sporulation (Chu *et al.*, 1998). Target genes of *NDT80* include the B-type cyclins and genes whose products are required for both meiotic divisions and gamete formation (Chu and Herskowitz, 1998; Xu *et al.*, 1995).

ABF1 encodes an essential DNA-binding protein that was originally found to be necessary for the activation of a number of autonomously replicating sequences (ARS; therefore termed ARS Binding Factor 1) and the maintenance of plasmids in budding yeast (Rhode *et al.*, 1992; Eisenberg *et al.*, 1988). The protein was later shown to be involved in processes as diverse as the transcriptional regulation of mitotic and meiotic genes, chromatin remodeling, silencing at the mating type locus, nucleotide excision repair, and the export of mRNA from the nucleus (Schroeder and Weil, 1998a; Gailus-Durner *et al.*, 1996; Pierce *et al.*, 1998; Ozsarac *et al.*, 1997; Miyake *et al.*, 2002; Ishida *et al.*, 2006; Reed *et al.*, 1999; Hieronymus and Silver, 2003). Abf1p is abundantly expressed in all cell types of *Saccharomyces cerevisiae*. The protein was defined to

be a General Regulatory Factor (GRF) that, like Rap1p and Reb1p, has several characteristic features: GRFs are essential genes that are abundantly transcribed. They function as enhancers of transcriptional events. Their target sites usually have little intrinsic activating capacities but instead amplify the effect of neighboring regulatory sites (Fourel *et al.*, 2002). It has also been shown that Abf1p can function as a so-called insulator, which keeps transcriptionally silent regions close to telomers in an active state (Fourel *et al.*, 2002) and that it plays a role in the differential transcriptional regulation of bi-directional genes (Ishida *et al.*, 2006).

A mutational screen for conditional mutations of *ABF1* resulted in the identification of several temperature-sensitive alleles (Rhode *et al.*, 1992). One of these alleles, *abf1-1*, proved to be a very useful tool for studying *ABF1* in various contexts, such as DNA replication (Rhode *et al.*, 1992) or transcriptional regulation (Schroeder and Weil, 1998b; Miyake *et al.*, 2004). *abf1-1* has a point mutation in a potential zinc finger domain and displays reduced affinity to target binding sites in an *ARS* element *in vitro* (Rhode *et al.*, 1992) and in promotors *in vivo* (Schroeder and Weil, 1998b). A recent genome-wide expression profiling study which compared *abf1-1* and wild-type strains resulted in a surprisingly small number of genes under positive (50 genes) and negative control (36 genes) of Abf1p in vegetatively growing cells (Miyake *et al.*, 2004). In contrast, genome-wide transcription factor binding assays showed that Abf1p binds to the largest number of genes among all known site-specific transcription factors in yeast (Harbison *et al.*, 2004; Lee *et al.*, 2002). Promotors of approximately 500 genes have been identified to be bound by Abf1p in growing cells. The discrepancy between both studies was explained by the fact that Abf1p may function according to the *hit-and-run* model in which the factor is dispensable for a gene's activity once it has activated it (Miyake *et al.*, 2004). The *hit-and-run* model was proposed earlier by Schroeder and Weil (1998b) on the basis of expression analysis and *in vivo* analysis of promotor occupancy. Previously, Abf1p has been shown to play a role in the transcriptional activation of three meiosis-specific genes (*HOP1*, *SPR3* and *SMK1*) (Gailus-Durner *et al.*, 1996; Prinz *et al.*, 1995; Pierce *et al.*, 1998; Ozsarac *et al.*, 1997). However, a comprehensive view of Abf1p target genes essential for or involved in meiotic development and sporulation not available to date. Furthermore, it is currently not understood how Abf1p coordinates mitotic growth with meiotic development on a transcriptional level.

Knowledge databases

The huge amount of information produced by high-throughput experiments such as expression profiling and genome-wide protein-DNA binding assays or protein-protein interaction studies require novel approaches for data management and organization of knowledge. However, before it is possible to efficiently process biological information, a structured vocabulary understood by all scientists has to be created. The GeneOntology project provides such a controlled vocabulary to describe gene and gene product attributes in any organism (Ashburner *et al.*, 2000). A gene product is described by using defined keywords that determine the *Biological Process*, the *Molecular Function* as well as the *Cellular Component*. Based upon this convention a number of databases related to the gametogenesis field have been constructed, such as an organ-specific database called *The Ovarian Kaleidoscope* (OKdb; supplemental table, ref. 1), which integrates published information on individual ovarian genes (Ben-Shlomo *et al.*, 2002), or *Mammalian Reproductive Genetics* (MRG; supplemental table, ref. 2), which covers genes involved in spermatogenesis in mammals. A community-based annotation approach has been taken by the *GermOnline* project (supplemental table, refs. 3-6). This database has been developed, and is being curated and updated by life scientists in cooperation with bioinformaticists. *GermOnline* covers 11 key model organisms as well as *Homo sapiens*, providing a platform for scientists studying meiosis and gamete formation to directly contribute and update information based upon peer-reviewed publications (Primig *et al.*, 2003; Wiederkehr *et al.*, 2004a; Wiederkehr *et al.*, 2004b). GermOnline provides access to microarray expression data from *Saccharomyces cerevisiae* and from

the rat male germ cell study presented in this thesis.

Goals of this work

One goal of this doctoral thesis was to elucidate the transcriptional program of mammalian male meiosis and gametogenesis. It was unclear if such an expression profiling study using germ cells and testicular somatic control cells would be possible. We therefore carried out an expression profiling study that compared highly enriched cell populations of rat male germ cells and included somatic Sertoli as well as brain and skeletal muscle controls.

The second goal of this doctoral thesis was to understand how Abf1p coordinates mitotic growth and meiotic development. This included the characterization of mitotic and meiotic phenotypes of *abf1-1* mutant cells and the identification of *ABF1*-target genes in vegetatively growing and sporulating cells on a genome-wide level. On the basis of this data we aimed at drawing conclusions about where in the mitotic and meiotic cell cycle *ABF1* is involved in.

Chapter 1: Expression profiling of Mammalian Male Meiosis and Gametogenesis Identifies Novel Candidate Genes for Roles in the Regulation of Fertility

Ulrich Schlecht*†, Philippe Demougin*†, Reinhold Koch*†, Leandro Hermida†, Christa Wiederkehr†, Patrick Descombes‡, Charles Pineau§, Bernard Jégou§, and Michael Primig†

† Biozentrum and Swiss Institute of Bioinformatics, 4056 Basel; Switzerland; ‡ Genomics Platform, National Center of Competence in Research Frontiers in Genetics, Université de Geneve/CMU, 1211 Geneva 4, Switzerland; and § Group d'Etude de la Reproduction chez le Male-Institut National de la Santé et de la Recherche Médicale U. 435, Université de Rennes I, Campus de Beaulieu, 35042 Rennes cedex, Bretagne, France

* These authors contributed equally to this study.

The work reported in this chapter represents the result of the collaborative effort of Philippe Demougin, Reinhold Koch, Leandro Hermida, Christa Niederhauser-Wiederkehr, Patrick Descombes, Charles Pineau, Bernard Jégou, Michael Primig and myself. For clarity, all results are reported. My personal contribution particularly included data mining, interpretation, and management (by creating the web-accessible reXbase database, that was later integrated into GermOnline) of the expression profiling data. Philippe Demougin, Reinhold Koch and me have contributed equally to this work that has been published in MBC.

Abstract

We report a comprehensive large-scale expression profiling analysis of mammalian male germ cells undergoing mitotic growth, meiosis and gametogenesis using High Density Oligonucleotide Microarrays and highly enriched cell populations. Among 11955 rat loci investigated, 1268 were identified as differentially transcribed in germ cells at subsequent developmental stages as compared to total testis, somatic Sertoli cells as well as brain and skeletal muscle controls. The loci were organized into four expression clusters that correspond to somatic, mitotic, meiotic and post-meiotic cell types. This work provides information about expression patterns of approximately 200 genes known to be important during male germ cell development. Approximately 40 of those are included in a group of 121 loci for which we report germ cell expression and lack of transcription in three somatic control cell types. Moreover, we demonstrate the testicular expression and transcriptional induction in mitotic, meiotic and/or post-meiotic germ cells of 293 as yet uncharacterized loci that are likely to encode novel factors involved in spermatogenesis and fertility. This group contains numerous potential germ cell specific targets for innovative contraceptives. A graphical display of the data is conveniently accessible through the GermOnline database.

Introduction

During mammalian spermatogenesis, Primordial Germ Cells develop into spermatogonia giving rise to spermatocytes that undergo two meiotic divisions to become round spermatids. These cells differentiate into spermatozoa during spermiogenesis (Pineau *et al.*, 1993b; Griswold, 1998). The development of male germ cells is governed by testicular Sertoli cells through secretion of, among others, signaling molecules and factors involved in cell adhesion (Griswold, 1998; Jégou *et al.*, 1999). Genes important for male gametogenesis and fertility in the mouse include loci involved in highly conserved landmark events such as meiotic recombination (Keeney *et al.*, 1997; Edelmann *et al.*, 1999; Luo *et al.*, 1999; Kneitz *et al.*, 2000; Romanienko and Camerini-Otero, 2000; Lipkin *et al.*, 2002), formation of the synaptonemal complex (that holds the homologous chromosomes together), sister chromatid cohesion, checkpoints and factors required for the meiotic cell cycle (Liu *et al.*, 1998; Zick-

ler and Kleckner, 1998; Wolgemuth *et al.*, 2002; Petronczki *et al.*, 2003). Other factors essential for post-meiotic stages of spermatogenesis include protein kinases (Miki *et al.*, 2002), proteases (Adham *et al.*, 1997; Shamsadin *et al.*, 1999), chromatin condensation factors (Yu *et al.*, 2000; Adham *et al.*, 2001; Cho *et al.*, 2001) and proteins involved in cell-cell adhesion, notably of Sertoli cells and spermatids (Mannan *et al.*, 2003).

It is noteworthy that genes required for spermatogenesis and fertility often encode transcripts present in meiotic or post-meiotic germ cells but not in somatic tissue (Eddy, 2002; McLean *et al.*, 2002). The notion that developmental-stage or tissue-specific expression indicates an important function is supported by results from high-throughput gene inactivation studies (Rabitsch *et al.*, 2001; Colaiácovo *et al.*, 2002) based upon expression profiling data obtained with sporulating yeast cells (Primig *et al.*, 2000; Chu *et al.*, 1998) and germ cells in the worm (Reinke *et al.*, 2000). Transcriptome studies have increased the pace at which genes important for sexual reproduction are identified in these important model systems; for review see (Schlecht and Primig, 2003). In this context it should be emphasized that budding yeast genes important for sporulation (a process analogous to spermatogenesis) were demonstrated to be transcriptionally induced during the process in a statistically significant manner (Deutschbauer *et al.*, 2002). A number of very recent array profiling analyses have addressed the problem of gene expression in the mammalian male germ-line. However, these experiments were either limited to post-meiotic germ cells (Ostermeier *et al.*, 2002), solely based upon total testis samples (Sha *et al.*, 2002; Sadate-Ngatchou *et al.*, 2004; Schultz *et al.*, 2003) or did not include somatic controls (Pang *et al.*, 2003). Moreover, none of these studies provided a searchable database accessible via the internet.

We sought to explore gene expression in rat Sertoli and germ cells to provide clues to the mitotic, meiotic and post-meiotic functions of many hundred uncharacterized transcripts. Among 11955 loci analyzed using a statistical approach, we identified 1268 as being strongly differentially regulated in testicular somatic versus germ cells. Those loci were organized into four broad categories whose transcription levels peak in either somatic or germ cells at mitotic, meiotic and post-meiotic stages of development. We included somatic testicular Sertoli cells as well as brain and skeletal muscle samples in our analysis as negative controls for transcriptional induction/expression in germ cells. The results clearly suggest that a substantial fraction of the known genes expressed in meiotic and post-meiotic germ cells but not in the examined somatic tissues are important for meiosis, spermatogenesis and/or fertility. Finally, quantitative PCR assays were carried out to underscore the reliability and reproducibility of the microarray expression data. The results of this analysis constitute an extremely useful source of information about thousands of mammalian transcripts. Moreover, it specifically marks out more than 290 as yet unknown candidate loci as being potentially involved in mammalian male gametogenesis or fertility. Since most of these loci are preferentially or maybe even exclusively expressed in germ cells, they could encode targets for novel approaches to temporarily and reversibly inhibit male fertility.

The expression data as well as supplementary material including analysis programs are accessible at (supplemental table, ref. 10). The data can be searched at reXbase integrated into the *R. norvegicus* section of GermOnline (supplemental table, refs. 3-6), a novel cross-species community annotation knowledgebase (Primig *et al.*, 2003; Wiederkehr *et al.*, 2004a; Wiederkehr *et al.*, 2004b) that provides a graphical display of microarray expression signals relevant for germ cell differentiation.

Results

Experimental design

The aim of this large-scale microarray study was to identify for the first time genes that are expressed during the mitotic, meiotic and post-meiotic developmental stages of mammalian germ cells and to further characterize their transcription patterns in somatic control cells or tissues. To this end we have employed High Density Oligonucleotide Microarrays (U34 A and B GeneChips) from Affymetrix. The tran-

Table 1. Cell purification efficiency and sample quality

Testicular cell type	Isolation procedure	Purity (%)	Contaminating material	Reference
Sertoli cells	Sequential enzymatic digestion	96	Peritubular cells (<2%), Mitotic and meiotic germ cells (2%)	(Pineau *et al.*, 1993b)
Spermatogonium	Sequential enzymatic digestion + sedimentation at unit gravity	85	Fragments of Sertoli cell cytoplasm, peritubular cells, Leydig cells	(Guillaume *et al.*, 2001)
Pachytene spermatocyte	Mechanical dispersion + centrifugal elutriation	90	Early spermatids	(Pineau *et al.*, 1993b)
Early spermatid	Mechanical dispersion + centrifugal elutriation	90	Early spermatids, heads from elongated spermatids, spermatid residual bodies, spermatocytes	(Pineau *et al.*, 1993b)

script levels of 11955 rat loci and expressed sequence tags (ESTs) were simultaneously measured in highly enriched testicular germ and Sertoli cell populations as well as in brain and skeletal muscle tissue (Table 1). Since RNA degradation causes data to be less reproducible and makes it more difficult to detect low-abundance transcripts, we included a sample prepared from total testis using a rapid extraction protocol as a control for total RNA quality. All total RNA (Figure 1, panel a and c) and cRNA target molecule preparations (panel b and d) made from purified cells and total testis tissue are of the highest quality.

Identification of differentially expressed transcripts in somatic versus germ cells

The terms locus and gene are used as synonyms throughout the text, while probe set refers to a row of oligonucleotides present on the microarray that are complementary to one transcript or one EST. Note that the arrays can detect more than one transcript/EST for one gene; that is to say a given number of probe sets detects an equal number of transcripts/ESTs that ultimately correspond to a smaller number of actual genes. Gene filtration was carried out as follows: the PerfectMatch (PM) values were corrected for background noise, normalized and summarized to one value for each probe set using methods implemented the Robust Multi-array Analysis (RMA) package (see Material and Methods for more details). This innovative approach does not take the values obtained with the Mismatch Oligonucleotide (MM) into account. This increases the sensitivity of the measurements since the MM signal is thought not just to reflect non-specific hybridization events but to contain true expression information which is lost when the default statistical algorithm for data analysis provided by the manufacturer is applied (Bolstad *et al.*, 2003; Irizarry *et al.*, 2003a). Permutation tests based upon the F Statistic that do not depend upon normal distribution of the data were employed to identify 9864 probe sets as being differentially expressed within the cell populations analyzed (Clarke and Cooke, 1998). In these cases the null hypotheses (all experimental conditions yield the same mean expression value for a given gene) were rejected at the 5% significance level. Using an approach by Storey it was estimated that among the 9864 probe sets should be 75 false positives (Storey, 2002). Finally, a subset of 1508 probe sets whose expression signals displayed a standard deviation greater than 1 across the 10 data sets (hybridization experiments) was selected for further analysis; these transcripts displayed robust differential expression patterns. The signals obtained in duplicate were found to be highly reproducible for both U34 A and B GeneChips as summarized in the box plots shown in Figure 2, panel a. The scatter plot matrix of 1508 probe sets also indicates that expression levels are highly similar within replicates while they split up into 4 broad clusters of differentially expressed loci (shown in color) when different testicular germ and Sertoli cell populations are compared (Figure 2, panel b).

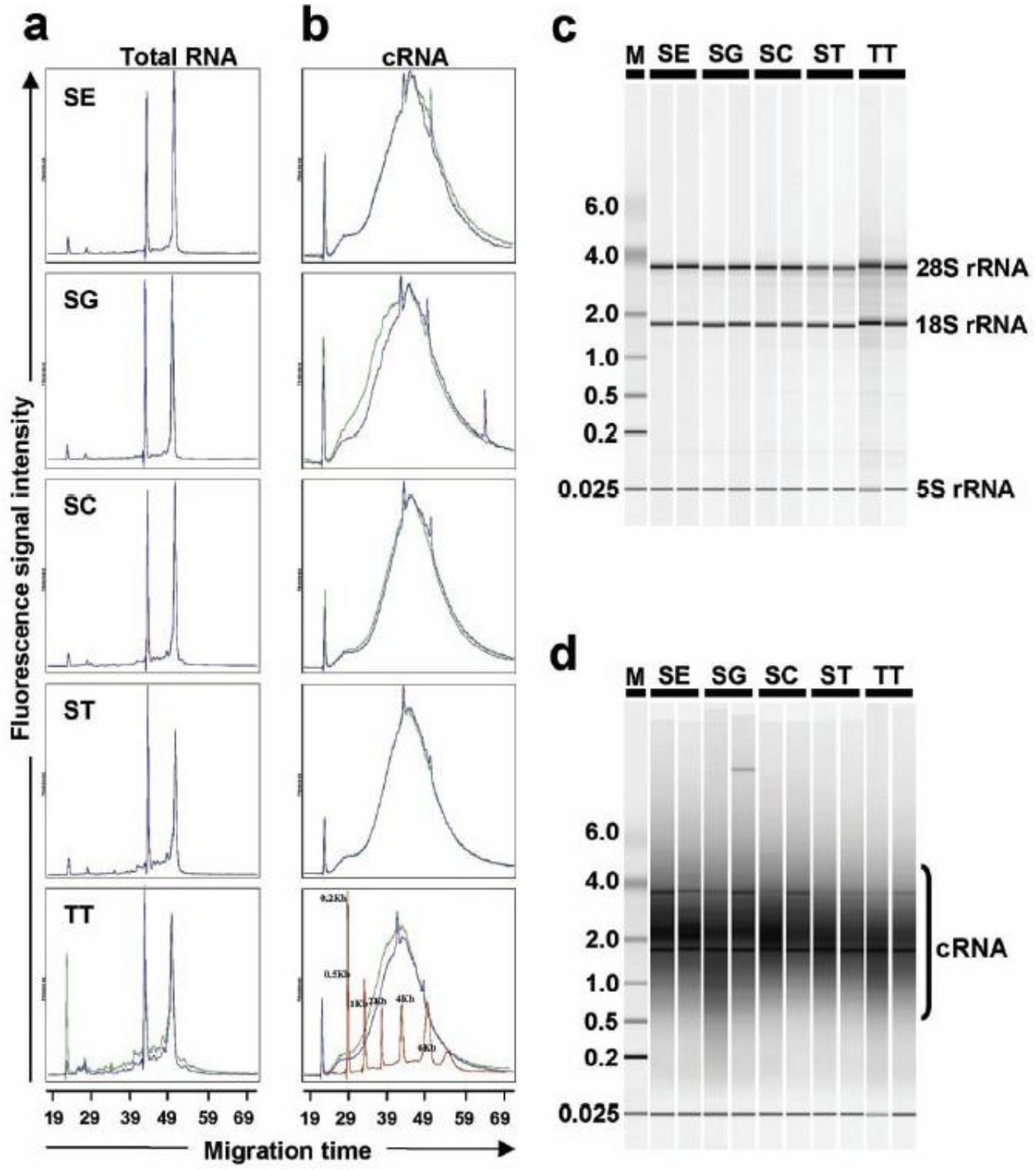

Figure 1: Total RNA and cRNA target quality assessment. Purified total RNA from Sertoli cells (SE), spermatogonia (SG), spermatocytes (SC), spermatids (ST) and total testis (TT) samples were analyzed for their concentration and overall length using the Agilent Bioanalyzer RNA Chip. Fluorescence intensities measured with total RNA and cRNA target molecules are plotted against migration time in seconds as indicated in Panel a for total RNA and Panel b for cRNA. Signals obtained with identical replicates are shown in green and blue graphs, respectively. The peaks corresponding to the molecular weight RNA ladder are displayed in the total testicle cRNA plot. Virtual gels of the total RNA and cRNA samples as indicated are shown in Panels c and d, respectively. M is molecular weight markers. rRNAs and the cRNA ladder are indicated.

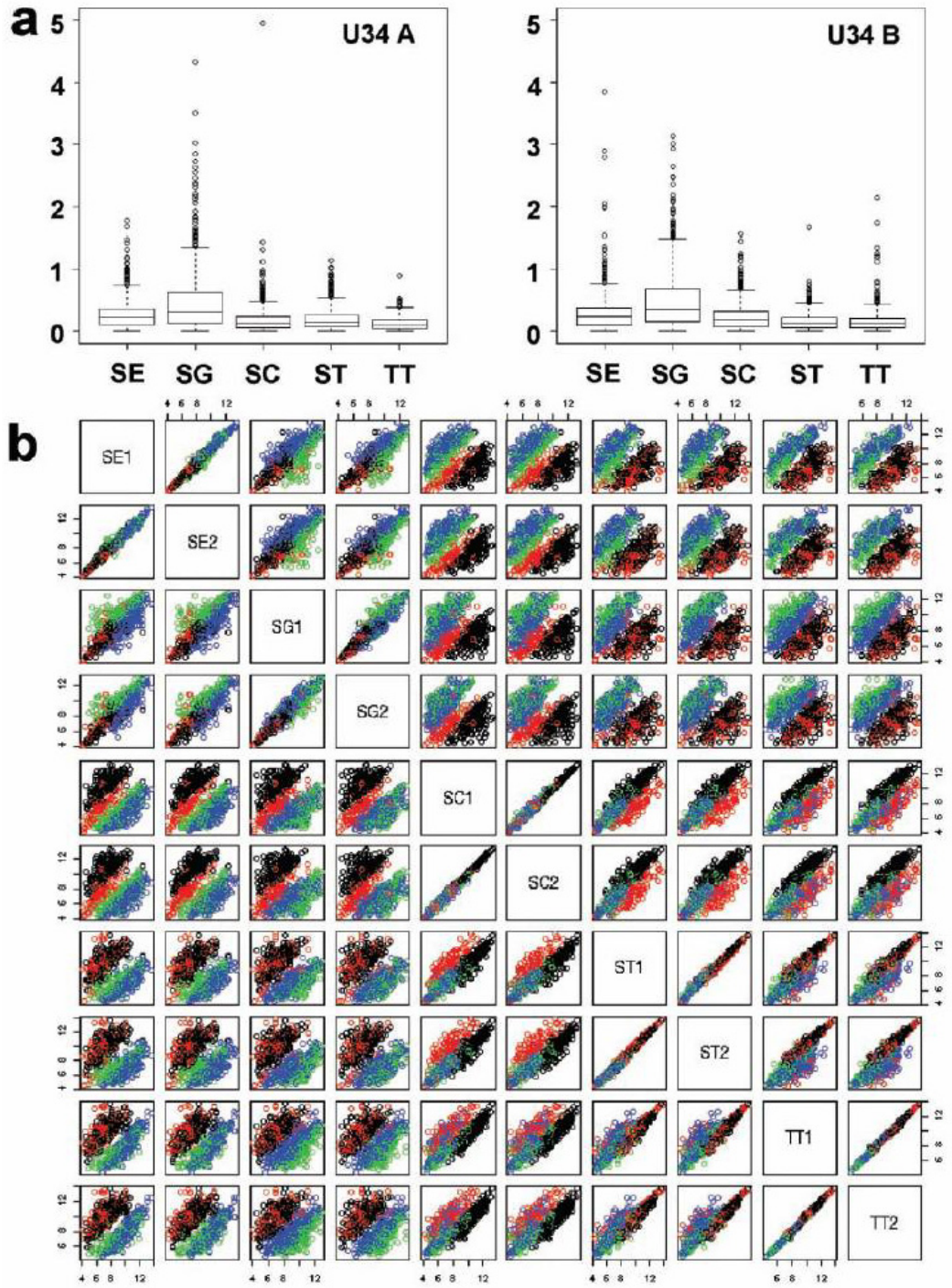

Figure 2: Comparing gene expression levels of 1508 probe sets within replicates and different cell populations. Box plots in Panel a display the overall distribution of observed changes in signal intensities obtained with two replicates from five samples hybridized to the U34 A and B GeneChips as indicated. Shown are the absolute log2 fold change values. The features of clear outliers were manually inspected on the corresponding DAT image file to exclude artifacts. The scatter plot matrix in Panel b summarizes a systematic comparison of the log2-scaled expression signals from 1508 transcripts among all samples as indicated. Genes that fall into the somatic, mitotic, meiotic and post-meiotic clusters are symbolized by circles colored in blue, green, black and red, respectively.

Identification of transcripts represented by probe sets through similarity search

The RG-U34 A and B GeneChips contain a total of 17590 probe sets that correspond to 11955 UniGene clusters, each of which is a group of sequences that belong to one unique gene. 1508 probe sets where found to display strongly changing expression signals during rat spermatogenesis (including 12 controls for rat GAPDH and Hexokinase that were not further analyzed). Since RG-U34 is based on Rat UniGene Build 34 we sought to identify the number of currently annotated loci represented by the 1496 probe sets, using Build 118 which contains 53186 distinct clusters.

The search revealed that the 1496 probe sets correspond to 1268 UniGene clusters of which 592 were annotated (Figure 3). This group comprises 487 loci that were allocated a gene symbol in the UniGene annotation and 105 that were manually processed by us. The group of annotated loci includes some that are similar to yeast genes important for the mitotic and meiotic cell cycle (Cdc42/CDC42, Ccnb1/CLB4, Ccnd2/CLB6, Stk6/IPL1, Tesk2/TPK3) and the regulation of meiosis (Mak/IME2, Gsk3β/RIM11) (Sicinski *et al.*, 1996; Gromoll *et al.*, 1997; Rosok *et al.*, 1999; Goepfert *et al.*, 2002; Shinkai *et al.*, 2002; Guo *et al.*, 2003; Lui *et al.*, 2003). See SGD (supplemental table, ref. 13) for references of the yeast genes. Among the 676 loci not yet annotated are 383 genes that have potential homologs in other species, notably *S. cerevisiae* (provided by HomoloGene). Examples of potentially important uncharacterized rat transcripts (expressed sequence tags, ESTs) include AA874887 and AI012651 that are similar to yeast genes involved in sister chromatid cohesion (SMC4) (Freeman *et al.*, 2000) and mitotic chromosome condensation (BRN1) (Ouspenski *et al.*, 2000), respectively (see reXbase for more details). Finally, we identified 62 transcripts that may have homologs not previously identified by HomoloGene and 231 transcripts expressed in germ cells that did not match any other currently known sequences in the comprehensive NCBI Protein NR database (Figure 3).

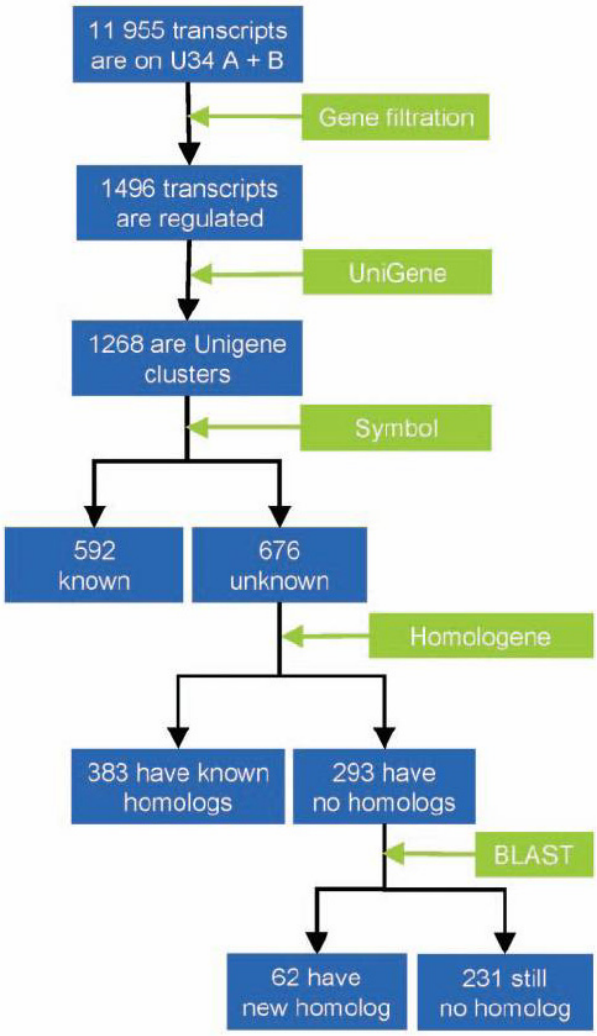

Figure 3: Mining U34 A and B chips for annotated and novel transcripts. This figure summarizes the steps taken to identify among all transcripts represented on the GeneChips those that are differentially expressed (Gene filtration; among the 1508 initial transcripts 12 were removed because they were hybridization controls), those that were in UniGene Build 118 (UniGene), those that are encoded by known genes (defined as bearing a genetic name; Symbol), those that had known homologs (Homologene) and finally those that were not known to have a homolog or those that do not appear to be similar to any currently annotated protein (BLAST).

Clustering gene expression profiles in somatic and germ cells

To verify the clustering behavior of replicated and distinct samples and to get a broad overview of the timing of gene induction, we first explored the dataset by hierarchical clustering of the 10 samples and 1508 expression patterns. The results are displayed by dendrograms (trees) and a color coded heat map. As expected, replicates clustered together in all cases (top dendrogram in Figure 4). Furthermore, we observed clustering of samples from Sertoli cells and spermatogonia indicating that differentiated somatic and mitotically growing spermatogonia share a substantial set of similarly expressed/induced genes. In contrast to that, meiotic and post-meiotic spermatocytes and spermatids are grouped together with adult total testis samples (top dendrogram). The data indicates that four broad categories of expression patterns exist in purified Sertoli and germ cells that can be summarized by high or intermediate expression peak levels in somatic/mitotic versus meiotic/post-meiotic samples, respectively (left dendrogram and indications in Figure 4).

To try and identify the minimal set of clusters that represent the observed expression signals and to assign biological functions to groups of genes whose expression peaks correlate with cell type (Sertoli cells), mitotic growth (spermatogonia) or meiotic developmental stage (spermatocytes and spermatids), the PAM algorithm was used. We defined four typical expression patterns (the medoids) within the group of 1508 transcripts (Figure 5, panel a) and then grouped all patterns around them. A graphical display of the expression patterns was produced using GeneSpring 5.1 (Panel b). These classes were verified by comparing the degree of similarity of a given pattern to those within its own cluster and to those in all other clusters using silhouette plots (panel c). Note that the somatic and mitotic expression clusters do not necessarily indicate specific expression but group genes together that display peak transcription in either of the cell-types. The results were verified using the k-means clustering algorithm implemented in GeneSpring 5.1.

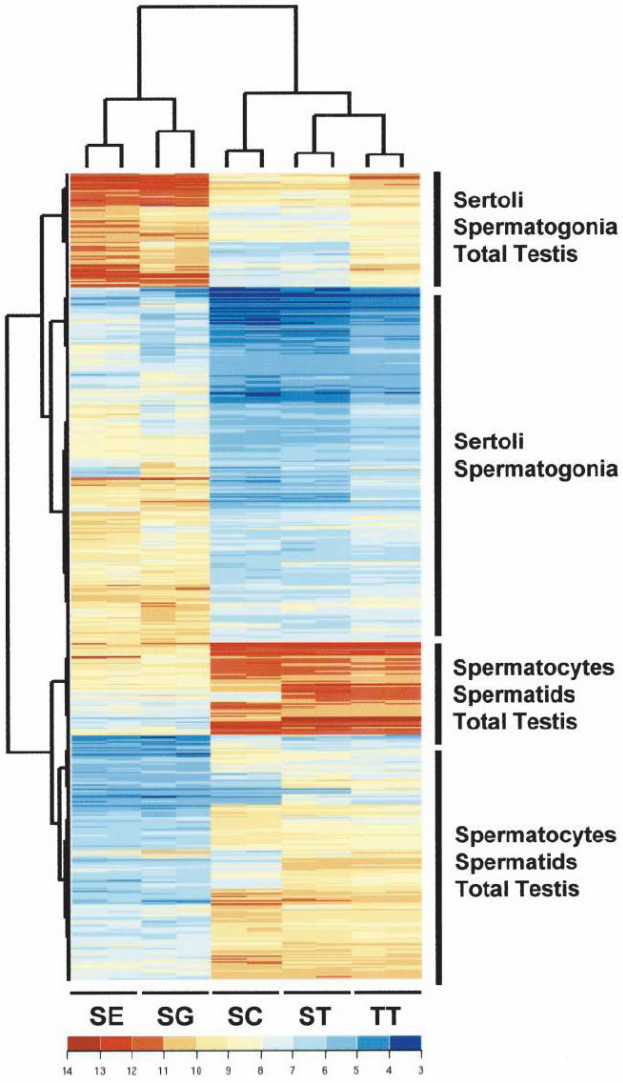

Figure 4: Hierarchical clustering of 10 samples and 1508 probe sets. A heatmap of 1508 transcripts and two dendrograms that group genes (left) and samples (top) together are shown. The samples are arranged such that somatic control Sertoli cells (SE) are followed by mitotic spermatogonia (SG), meiotic (pachytene) spermatocytes (SC), post-meiotic early spermatids (ST) and total testes (TT) tissue.Each line is a gene and each column is a sample. Clustered genes and samples as well as the color code for expression levels are given. Expression signal intensities are shown in red and blue indicating high and low expression, respectively. The cells displaying peak expression levels for a given group marked by a black bar are indicated.

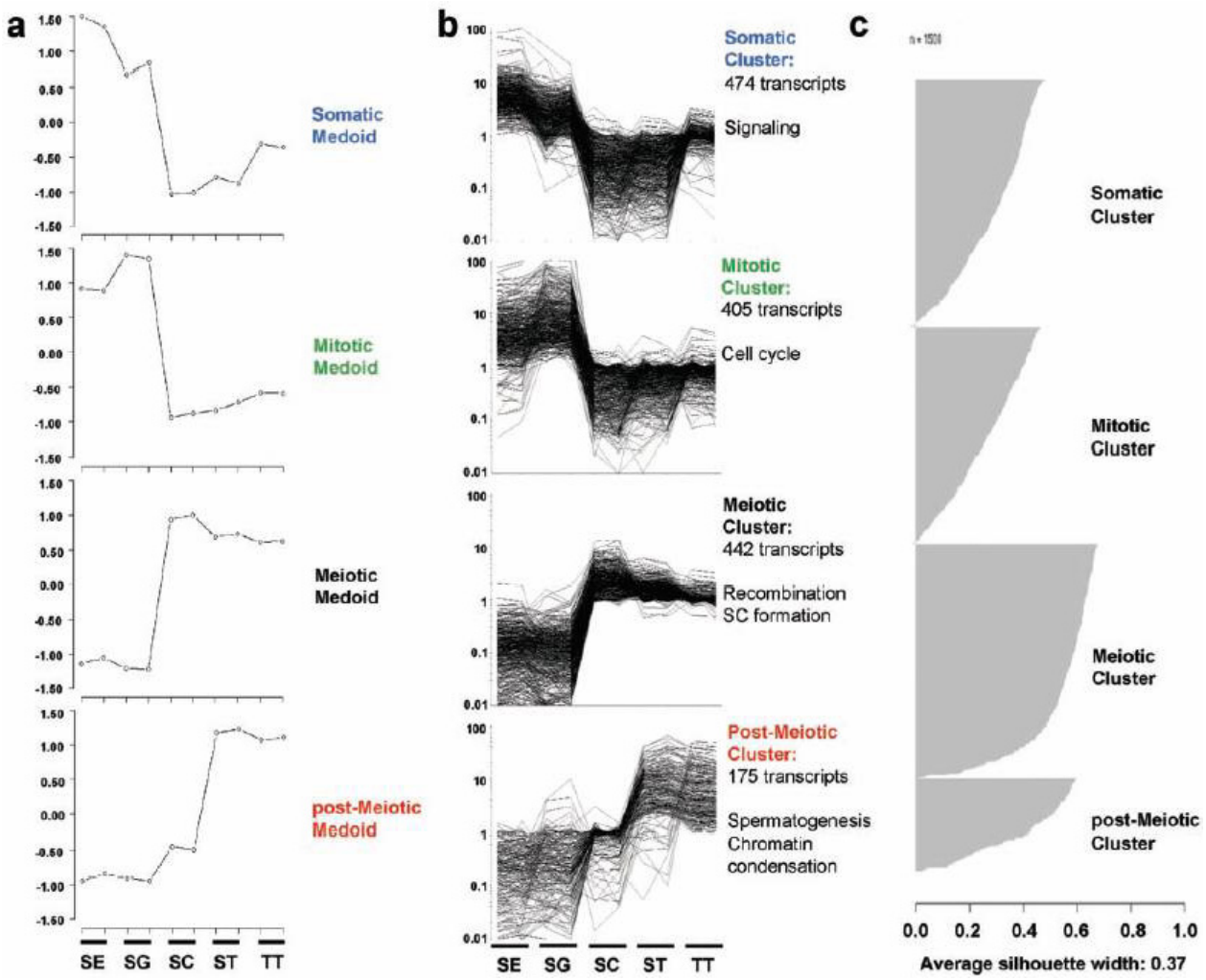

Figure 5: PAM clustering of 1508 probe sets. Panel a shows the medoids representing the four clusters covering the samples as indicated. Panel b summarizes a log-scale graphical display of the expression patterns within for the Somatic, Mitotic, Meiotic and the post-Meiotic Clusters as determined by Partitioning Around Medoids (PAM). For visualization, the data computed with MAS 5.0 were scaled and normalized using the default settings of GeneSpring 5.1. Silhouette plots of the clusters are displayed in Panel c. The scale of the silhouette is indicated and the average width is given at the bottom of the panel.

The somatic expression class contains 474 transcripts that are predominantly (but often not exclusively) expressed in Sertoli cells from adult rats. This group contains about 220 unknown transcripts. Among the annotated loci, approximately one hundred genes have previously been demonstrated to be expressed in testes and/or to play roles in testicular and Sertoli cell functions. They are involved for example in energy and fatty acid metabolism (Cycs, Ldhb, Scd2, Scp2), ubiquitin-mediated proteolysis (Ube2d3, Uchl1), detoxification (Gsta1, Gstm2, Gstp2), oxidative stress and serum response (c-Fos, Prdx1, Prdx2, Sod3) and cell-cell signaling/interaction (Ctsl, Il6st, Il6r, Nr0b1, Psap, Shbg, LOC286916 (Testin)). We also found cell adhesion factors (Cdh2, Ctnnb1) and confirmed expression in germ cells of genes involved in tumor formation/gonadogenesis (Wt1) and X-Chromosome inactivation (Tsx). The mouse homolog of Tsx was demonstrated to be transcribed in mitotic germ cells during puberty as well as in adult Sertoli cells - this pattern was correctly detected by microarray analysis (Cunningham *et al.*, 1998). Finally, we find expression in Sertoli cells and spermatogonia of a cell surface glycoprotein (Cd9) previously demonstrated to be expressed in spermatogonial stem cells (Kanatsu-Shinohara *et al.*, 2004). Cd9 is required for sperm-egg interaction and female fertility (Le Naour *et al.*, 2000).

The somatic class also contains more than 150 transcripts for known genes that were characterized in other tissues. Notably, we detect factors involved in cell cycle and growth control (Cdc42, Gas7, Rbbp7, Ywha), transcriptional regulation (Ssdp, Ureb1), signal transduction (Grb14, Jak1, Ptgfr, Rap1) and protein degradation (Adam10, Nedd4a, Prss11, Psma2, Psma3, Psme1, Psme2). Another large group consists of factors involved in energy metabolism (Aldoa, Aldr1, Atp1a1, Atp2c1, Atp5g2, Mor1, Pgam1, Pkm2). A gene encoding a factor involved in vitamin A transport (Rbp1) was found to be highly expressed both in Sertoli cells and spermatogonia. This is confirms previous reports showing that vitamin A (retinol) regulates a variety of testicular functions in rodents (Livera *et al.*, 2002). Interestingly, a factor involved in neurodegenerative conditions (Prnp, encoding a major prion protein precursor) was shown to be expressed in Sertoli cells and spermatogonia. This is in keeping with that fact that the Prnp homolog Dpl regulates male mouse fertility (Behrens *et al.*, 2002).

The mitotic expression class consists of 405 transcripts, including more than 200 uncharacterized mRNAs, which are up-regulated in mitotically growing spermatogonia. Among around 60 loci that were associated with testicular expression and/or spermatogonia in the literature, we detect the messages of genes required for cell cycle regulation (Ccnd2), components of the extracellular matrix (Col3a1, Mgp), hormone signal transduction (Ctgf, Egr1, Fgfr1, Igfbp2, Igfbp3, Pdgfa, Vegf), serum response and transcriptional regulation (Jun, JunB, Id2, Klf9, Stat3, Zfp36).

Approximately 140 genes that were not known to be expressed in germ cells include transcription factors (Fosl1, Copeb and Gata4), a protein phosphatase 2a catalytic subunit (Ppp2ca) and an AP-1 regulated serum inducible protein which is involved in inflammatory processes in the brain and pancreas (Scya2). The latter may play a role in the testicular immune response to viral infections (Le Goffic *et al.*, 2002). The expression level of about half of the genes in the mitotic cluster specifically peaks in mitotically growing spermatogonia and then substantially decreases at later meiotic stages. This is the case for e.g. histones (H1f0, H3f3b), ribosomal proteins (Rpl35, Rps3, Rps4) and motor proteins (Mrlcb, Myh9, Mylc2a, Tpm1, Tpm4).

The meiotic expression class contains 442 loci highly induced in spermatocytes. Transcripts that fall into this class, which includes 280 ESTs, are either only transiently up-regulated during meiosis or they continue to be expressed post-meiotically after meiotic induction. As expected, we found genes in this group that are involved in SC formation (Sycp2, Sycp3), DNA repair (Polβ), and chromatin condensation (Top2a). Moreover, we find metallo-proteases (Adam2, Adam3 and Adam5), factors necessary for ubiquitin-mediated protein degradation (Ube2d2, Psmc3), transcriptional regulators (Crem, Miz1), enzymes involved in energy metabolism whose somatic isoforms are expressed in Sertoli cells (Cyct, Ldhc) and a component of the flagellum ultimately required during spermiogenesis (Odf2). Previously undetected expression in germ cells was observed for a factor originally identified in Sertoli cells (Sert1), a gene involved in cell cycle regulation (Cdc20) and transcription factors involved in organogenesis of the heart, gut and lung, for example (Pitx2, Hipk3).

The post-meiotic expression class contains 175 loci, including 70 that encode uncharacterized transcripts. These transcripts are detected predominantly in spermatids and, in almost all cases, in the total testis sample. The latter provides not only a control for mRNA integrity but also a source for transcripts present in differentiated sperm that constitute approximately two thirds of the testicular cell mass. This class contains almost 60 known genes that encode proteins implicated in regulation of meiosis (Gsk3β), cell cycle control (Ccnb1), protein phosphorylation (Clk3, Mak) and degradation (Adam1, Adam4, Adam6, Cst8), motor proteins (Krp2, Krp5), water permeability (Aqp7, Aqp8) and cAMP signaling (Akap1). We also find factors involved in DNA condensation during sperm differentiation (Prm2, Prm3, Tnp1, Tnp2), sperm maturation (Pp2a, Odf1, Odf3), adhesion (Spam, Sp10) and motility (Akap4).

Among known genes from somatic or ovarian tissues we identify loci that play roles in inflammation (Aif1), signaling (Epha7) and an

estrogen-regulated factor implicated in female reproduction (Itmap1). Our results indicate that these genes may play a broader role in fertility than previously thought.

A number of loci display significant expression signals only in the total testis extract but not in any of the purified cell populations. The corresponding mRNAs could be specific for testicular cell types that were not analyzed in this study. This includes the steroidogenic Leydig cells that express e.g. Cyp17 involved in the production of corticosteroids and androgens (for review see (Payne, 1990)) or Hsd3b1 required for steroid synthesis (O'Shaughnessy *et al.*, 2002).

Somatic tissue profiling of germ cell genes

To further investigate the transcriptional patterns of loci highly expressed in germ cells, we included brain and skeletal muscle samples in the analysis of a selection of 302 transcripts from the meiotic and post-meiotic clusters (Figure 5, panel b). This approach marks out 121 loci as being expressed in germ cells but not in somatic Sertoli, nor in brain or skeletal muscle cells (Figure 6). The group comprises genes that are known to be involved in testicular transcription (Crem), cell cycle regulation (Cdc20, Ccnb1), synaptonemal complex formation (Sycp2, Sycp3), DNA repair (Polβ), chromatin condensation (Top2a, Tnp1 and Tnp2), cell-cell contact (Adam1, 2, 4, 5 and 6, Tpx1, Spam), proteolysis (Pcsk4), and water transport (Aqp8, Aqp9). Importantly, this approach identified 37 ESTs with expression signatures that suggest possible roles for their corresponding genes in spermatogenesis and fertility.

Array data validation by *Q-PCR*

To validate the microarray expression signals we carried out quantitative PCR assays (*Q-PCR*) and observed a very good correlation between both techniques in cases of three known genes and five ESTs (Figure 7): Gsk3β, an ortholog of the yeast RIM11 kinase required for meiosis and spore formation (Malathi *et al.*, 1997), was very recently shown to be expressed in testes where it regulates mammalian meiosis (Guo *et al.*, 2003). The locus encoding

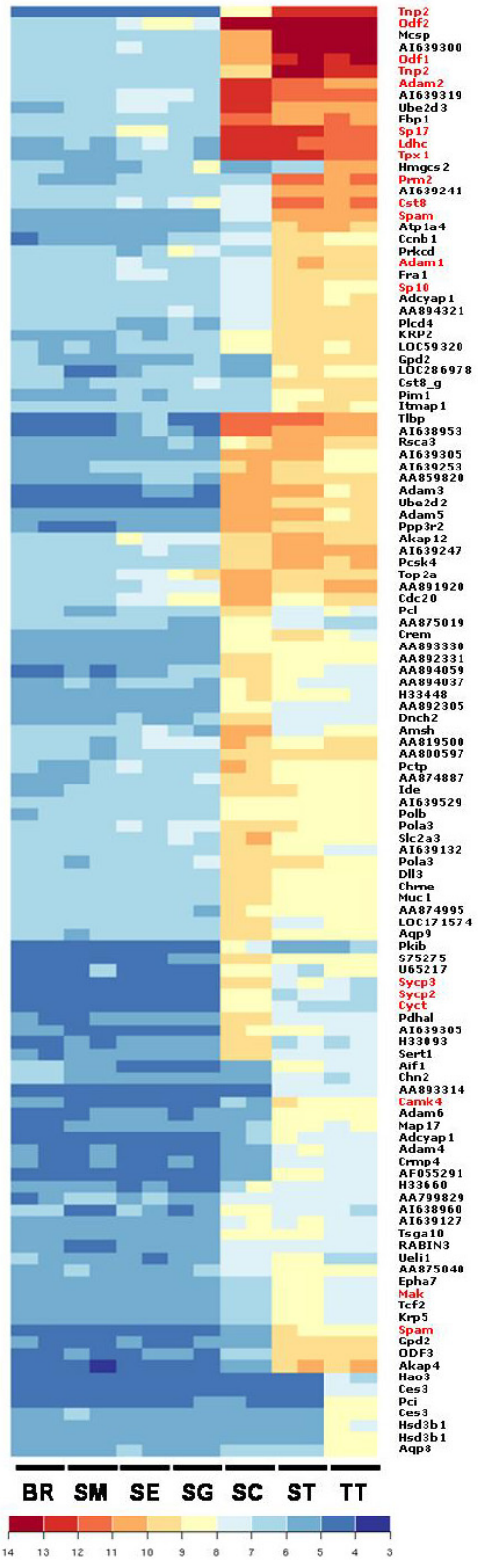

Figure 6: Clustering of somatic control cells versus germ cells. A heatmap of the expression data obtained with somatic controls from brain (BR) and skeletal muscle (SM) are compared with Sertoli and germ cells as indicated. The names of the corresponding loci and ESTs are given. Expression signal intensities are color coded as indicated.

the rat TATA box binding protein (Tbp, represented by the EST AA891928 that maps to NW_043264) also displayed the expected highly reproducible pattern of transcriptional induction in germ cells as previously reported (Persengiev *et al.*, 1996). Aqp8 was claimed to be expressed in germ cells at various stages of development (Ishibashi *et al.*, 2002). However, we detected significant expression in any of the purified cell populations neither by microarrays nor by *Q-PCR*. As opposed to that, a clear signal was observed in the total testicular extract (Figure 7). The ESTs whose germ cell expression was validated are AI030059 (no known homolog), AI639326 (similar to the human protein pir:T17234), AI639231 (similar to the human protein pir:T46480) and AI639319 (similar to the mouse locus NP_080924.1). Furthermore, we verified AI013870 which maps to the rat locus AY241457.1 that encodes a conserved histone acetyl transferase (see discussion).

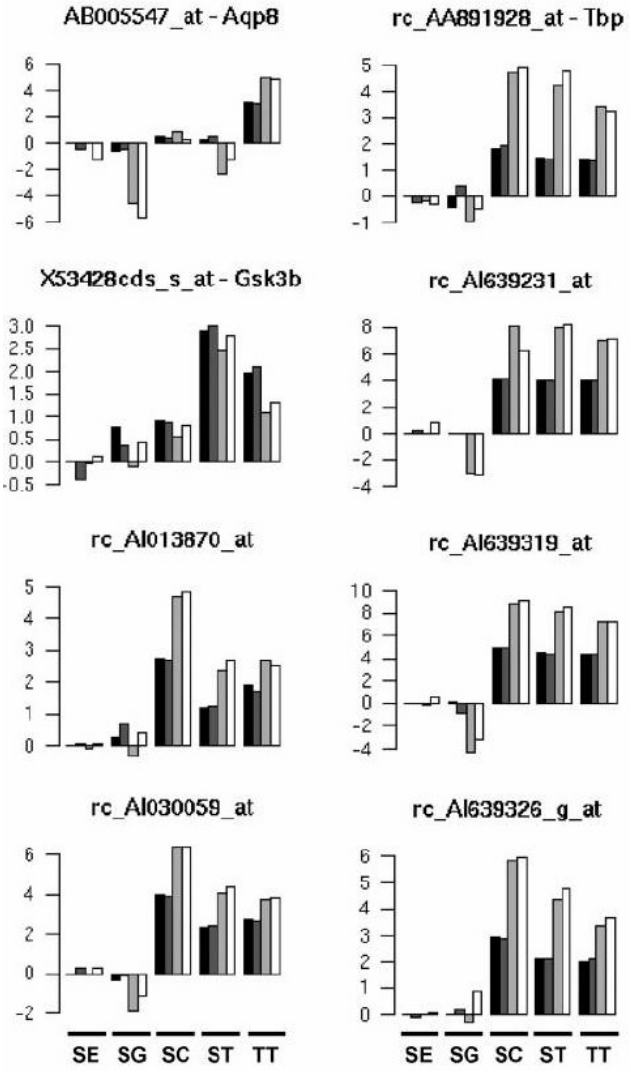

Figure 7: *Q-PCR* validation of selected transcripts. A graphical display of relative RNA concentrations as determined by quantitative PCR and microarray expression signals is shown. The samples are displayed in duplicate for array data (black and dark grey bars) and *Q-PCR* data (light gray and white bars) as indicated. Signals are given on a log2 scale.

Online data-access and visualization

An online-accessible database called reXbase (rat expression database) has been developed which provides locus information pages for every gene included in the current study. Every information page gives a brief description of the gene a provides deep-links to other databases such as Unigene, Genbank or the RGD (Rat Genome Database, supplemental table, ref. 16) (Figure 8). Each page also displays expression levels of genes as measured in the current study. reXbase is now a part of GermOnline, a cross-species community knowledgebase on germ cell differentiation. Germ-Online provides information and microarray expression data for genes involved in mitosis and meiosis, gamete formation and germ line development across species. The database has been developed, and is being curated and updated, by life scientists in cooperation with bioinformaticists. Information is contributed through an online form using free text, images and the controlled vocabulary developed by the GeneOntology Consortium. The database is governed by an international board of scientists to ensure a standardized data format and the highest quality of GermOnlines information content.

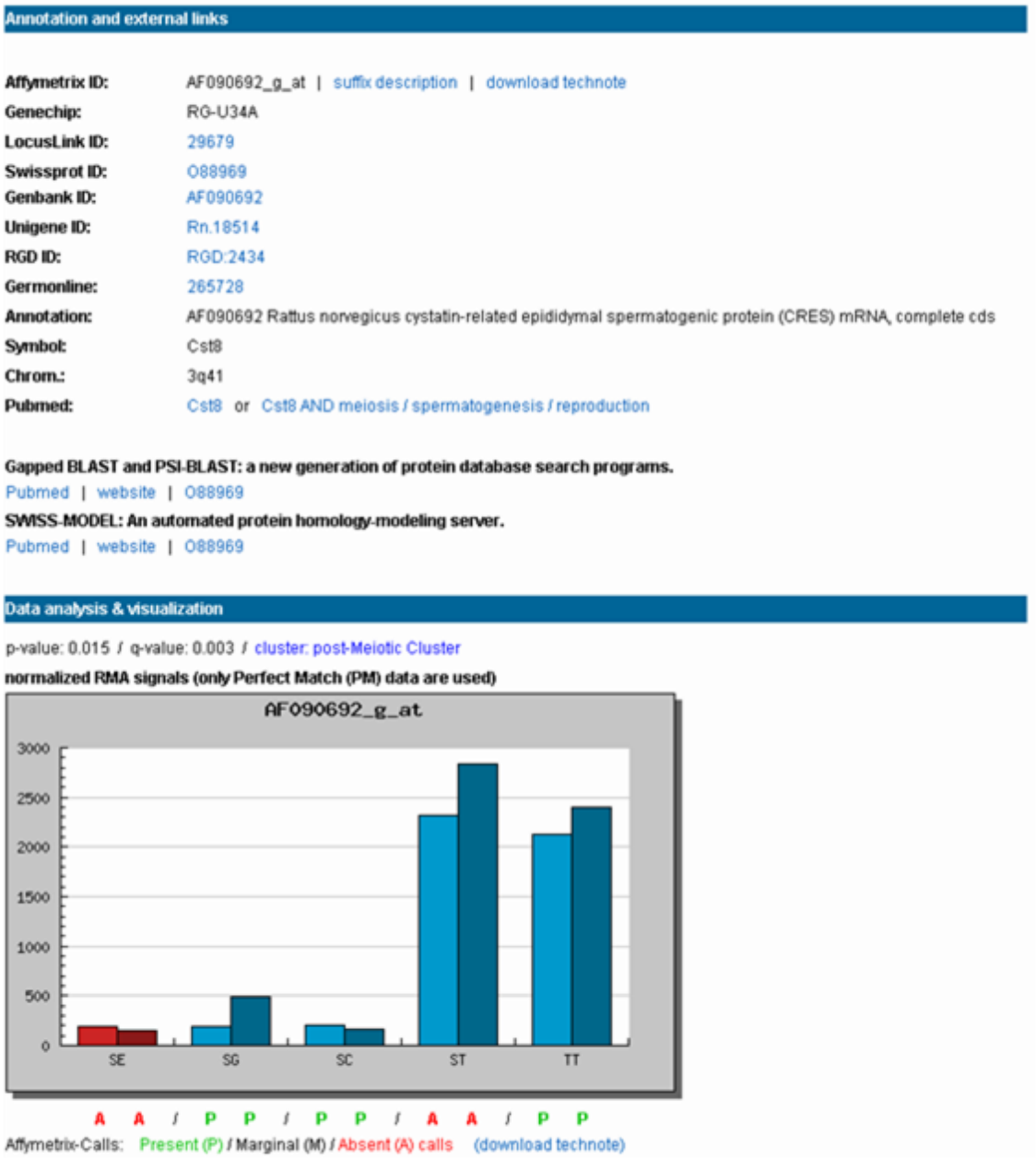

Figure 8: Example of the reXbase locus information page for the rat gene Cystatin 8 (Cst8). The annotation and external links section provides a brief description of the gene and deep-links to other relevant databases (including Unigene, Swissprot and GermOnline). The data analysis and visualization section summarizes cluster information and gives a graphical display of the expression profile of Cst8 in various testicular tissues. SE, Sertoli cells; (SG) spermatogonia; SC, spermatocytes; ST, spermatids; TT total testis

Discussion

Mining mammalian meiosis

This large scale microarray experiment explores the patterns of transcriptional regulation underlying male meiosis and gametogenesis as compared to testicular and non-testicular somatic control tissues in the rat and thereby confirms and substantially extends previous work in human and mouse (Ostermeier *et al.*, 2002; Sha *et al.*, 2002; Pang *et al.*, 2003; Schultz *et al.*, 2003). Results obtained in this study are based upon the analysis of approximately 35% of the genes expected to be present in a rodent genome (Waterston *et al.*, 2002). It therefore provides a highly valuable data set readily available for the community via the GermOnline knowledgebase and constitutes a prototype approach for comprehensive transcriptomics in the field of eukaryotic sexual reproduction. It was our aim to identify loci that display strong expression throughout meiotic and post-meiotic stages of spermatogenesis because these genes are likely to play important roles in male gametogenesis and fertility. To this end, we have partially reconstituted the mitotic, meiotic and the post-meiotic phases of rat sperm development using highly enriched spermatogonia, spermatocytes, and spermatids. As negative controls for germ cell expression/induction and mRNA stability we included purified somatic Sertoli cells, total testis samples as well as brain and skeletal muscle tissues. Among 11955 transcripts analyzed, 1268 showed substantial differences in expression levels during the mitotic, meiotic and post-meiotic phases of spermatogenesis. The corresponding probe sets were explored and organized into four expression categories based upon the cell type and timing of induction during mitotic, meiotic and post-meiotic stages of spermatogenesis using hierarchical and PAM clustering algorithms. This study confirms the expression patterns of numerous previously detected loci and provides clues about the possible roles in male meiosis and gametogenesis for 293 as yet uncharacterized loci whose transcripts are present in germ cells.

We demonstrate the reproducibility of array data by quantitative PCR which consistently matches microarray hybridization signals. Correlation coefficients of the expression profiles determined for 8 cases were found to vary between 0.83 and 0.99. The profiling data reported here confirm Gsk3β expression in Sertoli and germ cells observed by RNA in situ hybridization. Inhibition of the Gsk3β kinase a homolog of Rim11 that phosphorylates Ume6 and Ime1 (Bowdish *et al.*, 1994; Malathi *et al.*, 1997), was shown to prevent pre-meiotic DNA replication in cultured rat germ cells (Guo *et al.*, 2003). This gene is therefore a prototype case where increased mRNA concentration is concomitant to an elevated level of a protein that fulfills an important function in germ cells (Guo *et al.*, 2003). The gene encoding Tpb (TATA binding protein) was demonstrated to be highly induced in meiotic and post-meiotic rat and mouse male germ cells as compared to somatic tissue (Persengiev *et al.*, 1996); both microarrays and *Q-PCR* analysis clearly reproduce this pattern, further underlining the reliability of the array data. While we cannot confirm an earlier report claiming that Aqp8 is expressed in germ cells at various stages of development, our results are consistent with testicular expression and lack thereof in brain and skeletal muscle (Ishibashi *et al.*, 1997). Moreover, although the in situ hybridization signal published by Ishibashi *et al.* is indeed localized within the seminiferous tubules, poor resolution of the image presented makes it difficult, if not impossible, to verify germ cell expression. An EST present on the U34 B array (AI013870) lead us to identify the rat homolog of a conserved putative histone deacetylase (HDAC; AY241457.1) as being strongly expressed in spermatocytes. The yeast homolog of the rat HDAC (ESA1) is required for mitotic growth (Smith *et al.*, 1998) and Mof, its putative fly homolog, is involved in X-chromosome gene dosage compensation. This raises the intriguing possibility that the rat gene may play similar roles during spermatogenesis (Akhtar and Becker, 2000).

Experimental design and data analysis

Is it possible to reconstitute a complex developmental pathway using purified cell samples? While it is likely that lengthy cell purification procedures affect transcript concen-

trations, there are several arguments in favor of profiling experiments based upon enriched germ cells. First, microarrays confirm numerous patterns from previously studied genes whose expression was analyzed by in situ hybridization in the natural cellular environment (e.g. Tsx and Gsk3β) (Cunningham *et al.*, 1998; Guo *et al.*, 2003). Second, the timing of induction of many meiotic and post-meiotic transcripts reported here correlates well with the time of function during spermatogenesis (e.g. Sycp2 and Sycp3 which encode SC components expressed in spermatocytes) (Offenberg *et al.*, 1998; Schalk *et al.*, 1998). Third, it is a feature of germ cells to contain mRNAs that are stored after transcription for translation at later stages; these mRNA species can therefore be detected by microarrays rather efficiently; for review, see (Steger, 2001). In addition to that, signals from purified meiotic and post-meiotic germ cells were for many genes at least as strong (or even stronger) than those obtained with total testicle tissue (e.g. Ccnb1, Crem, Adam 4, Adam6, Akap4; see GermOnline). This indicates that mRNA degradation in purified cells is not a critical problem in most cases (Figures 3 and 4).

Our results are in keeping with previous findings that meiotic transcripts even from very early zygotic stages of embryological development are very stable and clearly detectable by microarrays in human sperm (Ostermeier *et al.*, 2002). Moreover, similar experiments using mouse total testis as well as enriched cell samples also yielded abundant evidence that male germ cells are an excellent source for transcripts expressed/induced during spermatogenesis (Pang *et al.*, 2003; Schultz *et al.*, 2003). A direct comparison of our results with these studies is complicated by the fact that different array technologies and utterly distinct analysis approaches were employed. However, it is noteworthy that among 22 conserved genes important for spermatogenesis, 20 (91%) were identified both by us and work based on mouse U74 GeneChips (Affymetrix) as being up-regulated in male germ cells (Table 3 in (Schultz *et al.*, 2003). A better understanding of the conserved mammalian transcriptome of spermatogenesis will have to await a genome-wide comparative profiling study in mouse, rat and human; it should be mentioned that the array technology required for such an experiment is already available.

Cross-contamination of purified cell populations is likely to be irrelevant for most transcripts identified in this study (see Figure 1 and Table 1) because many reference genes known to be very strongly expressed in Sertoli cells, spermatocytes or spermatids (e.g. Tsx, Sycp2, Sycp3, Mak, Tnp1, Tnp2, Odf1, Odf2, Odf3) are indeed detected only in these cells and not in any of the other cell populations or the controls (see GermOnline for expression patterns). It should be emphasized that this is true for both replicate samples that were analyzed for each cell type. One should bear in mind, however, that for rare cases of genes that are extremely highly expressed in certain (germ) cell types, it may be difficult to distinguish between transcription in the enriched cell population and the sub-population of contaminating cells (see Table 1).

In the successive hierarchical and PAM cluster analysis we focused on relative differential transcription patterns rather than trying to categorize loci based upon expression levels (which would have produced a larger number of clusters). This was due to the nature of our experimental approach since it is likely that mRNA concentrations change to a certain extent during germ cell purification as compared to the in vivo level. The expression signals may therefore in some cases not accurately reflect the physiological promoter activity.

Germ cell expression and function

We included brain and skeletal muscle samples in a refined analysis of 617 probe sets that yield highly expressed in germ cells to identify 121 transcripts and ESTs as being expressed in germ cells but not in at least three different types of somatic cells. This group includes 30 genes whose products are thought to fulfill meiotic or post-meiotic roles. It seems therefore reasonable to assume that 36 ESTs also identified are likely to lead to a number of genes important for spermatogenesis and fertility. It would be very interesting to compare gene expression in testicular tissue with a large number of somatic tissues in the rat to address the important question of germ cell specificity. However, a truly informative tissue profiling

approach requires comprehensive data sets that are currently available only for mouse and human (Su *et al.*, 2002). The outcome of such an experiment is bound to be extremely yielding in the light of the fact that up to approximately 4% of the genes in the mouse genome may be expressed in post-meiotic male germ cells alone (Schultz *et al.*, 2003).

Any attempt to identify tissue-specifically expressed genes has the caveat that one may fail to find important genes also transcribed in other organs. Spo11, for example, is present only in sporulating but not growing yeast cells (Atcheson *et al.*, 1987) while the mammalian Spo11 homolog, which is essential for gametogenesis in both males and females, is also weakly expressed in other tissues, including the brain. However, no neuropathology was detected in homozygote knock-out mice (Baudat *et al.*, 2000; Romanienko and Camerini-Otero, 2000). One explanation for this observation is that the Spo11 mRNA may not be translated into a functional protein in nervous tissue. A similar case was observed for the myogenic transcription factor Myf5. Although the gene is transcribed in a specific region of the brain, its mRNA is not translated into a protein detectable by immunofluorescence in that tissue (Daubas *et al.*, 2000).

The work reported here clearly indicates that expression profiling of mammalian spermatogenesis using highly enriched Sertoli and germ cells as well as testicular and somatic control tissues is feasible and very yielding. The data furthermore confirms that expression data obtained with classical molecular and cell biological methods are faithfully reproduced by microarrays. This exploratory study provides data for numerous genes previously characterized in somatic tissues (and sometimes thought to be specific for that tissue) that indicate expression in testicular germ cells. In addition to that it identifies more than 290 uncharacterized loci potentially important for male fertility. Such results, conveniently accessible through GermOnline and reXbase, will ultimately help better understand the genetic defects that lead to human reproductive pathologies such as sterility and birth defects. Finally, it is likely to contribute to the development of innovative approaches to reversible inhibition of male fertility.

Materials and Methods

Cell purification protocols Sertoli cells were isolated from 8 rats at 20 days post-partum (dpp) as previously described (Toebosch *et al.*, 1989). To isolate spermatogonia, testes of 25 Sprague-Dawley rats at 9 dpp were excised and decapsulated. Seminiferous epithelial cells were dispersed by collagenase treatment and separated at a purity of greater than 90% as previously described (Guillaume *et al.*, 2001). Pachytene spermatocytes and early spermatids were prepared by centrifugal elutriation to a purity greater than 90% from 8 rats at 90 dpp as previously described except that cells were mechanically dispersed (Pineau *et al.*, 1993a). Total testicular samples were produced by excising and snap freezing testes from three Sprague-Dawley rats at 90 dpp in liquid nitrogen. The outermost connective tissue capsule was surgically removed on the frozen organs before they were manually ground up using mortar and pestle. Total RNA was purified using the RNeasy kit (Qiagen) following the instructions of the manufacturer. Tissue samples from brain (Lewis, 60 dpp) and skeletal muscle (Wistar, 70 dpp) for microarray analysis were isolated from adult rats according to standard procedures as recommended by the manufacturer (Affymetrix).

Quantitative PCR analysis of mRNA concentration cDNA was synthesized using random hexamers and Superscript II reverse transcriptase (Invitrogen) following the instructions of the manufacturer. The amplicons were designed with the program Primer Express 2.0 (Applied Biosystems) using default parameters such that they spanned exon boundaries. The specificity of each primer pair as well as the efficiency of the amplification step were tested by assaying serial dilutions of cDNA oligonucleotides specific for target and normalization control loci and EST (see web site). PCR reactions were carried out in triplicate using a SDS 7900 HT instrument (Applied Biosystems). Raw Ct values obtained with SDS 2.0 (Applied Biosystems) were imported into Excel (Microsoft) to calculate the normalization factor and the fold changes with the geNorm script as published (Vandesompele *et al.*, 2002).

cRNA target synthesis and GeneChip hybridization 50 μg of total RNA was purified on average from the samples using RNeasy Mini-Spin columns (Qiagen) employing standard protocols provided by the manufacturer. The cell pellets were resuspended in RLT buffer and lysed by shearing in a 2 ml syringe. 600 μl of the supernatant was mixed with 600 μl of 70% Ethanol, loaded onto an RNeasy column, washed and eluted in 100 μl distilled water. Total RNA quality was monitored by loading approximately 200 ng onto an RNA Nano 6000 Chip that was processed with the 2100 Bioanalyzer (Agilent). Biotin labeling of RNA was performed as described in the Expression Analysis Technical Manual (Affymetrix) with minor modifications as indicated below. To synthesize single-stranded cDNA 13 μg of total RNA was mixed with oligo-dT and incubated with SuperScript II reverse transcriptase (Invitrogen) at 42°C for one hour. After synthesis of the second cDNA strand using the SuperScript

Double-Stranded cDNA Synthesis Kit (Invitrogen), the material was extracted with phenol-chloroform-isoamyl alcohol and precipitated with 0.5 volumes of 7.5 M ammonium acetate and 2.5 volumes of ethanol. 50% of the cDNA was used for an in vitro transcription reaction using the BioArray High Yield RNA Transcript Labeling Kit T7 (Enzo) to synthesize cRNA in the presence of biotin-conjugated UTP and CTP analogs. Approximately 50-100 μg of labeled cRNA from each reaction was purified by using RNeasy Mini-Spin columns and roughly 300 ng was analyzed on RNA Nano 6000 Chips. The cRNA targets were incubated at 94°C for 35 min and the resulting fragments of 50-150 nucleotides were again monitored using the Bioanalyzer. All synthesis reactions were carried out using a PCR machine (Biometra T1 Thermocycler) to ensure the highest possible degree of temperature control.

220 μl of the hybridization cocktail containing fragmented biotin-labeled target cRNA at a final concentration of 0.05 μg/μl was transferred into Rat U34A and U34B GeneChips (Affymetrix) and incubated at 45°C on a rotator in a Hybridization Oven 640 (Affymetrix) for 16 hours at 60 rpm. The arrays were washed and stained by using a streptavidine-phycoerythrin conjugate (SAPE; Molecular Probes). To increase the signal strength the antibody amplification protocol was employed (EukGE-WS2v4; Affymetrix Expression Analysis Manual). The GeneChips were processed with a GeneArray Scanner (Agilent) using the current default settings. DAT image files of the microarrays were generated using Microarray Analysis Suite 5.0 (MAS; Affymetrix).

Data and cluster analysis CEL data files were computed using the statistical algorithm implemented in MAS 5.0 (supplemental table, ref. 9). The data was further analyzed using programs developed in R, a programming language and developer environment for statistical computing and graphics (supplemental table, ref. 12). The Robust Multi-array Analysis (RMA) algorithm as implemented in the BioConductor package was employed for data normalization using the quantile method as implemented in the package (supplemental table, ref. 11). This method aims at making the distribution of probe intensities the same across all arrays in a given set (Bolstad *et al.*, 2003; Irizarry *et al.*, 2003b). Gene expression signal calculation was based upon the Perfect Match (PM) values from each probe set as published (Irizarry *et al.*, 2003a). The program written for data preparation and analysis in this study is available at our web site (follow the link "download software" in the navigation bar).

To identify differentially expressed genes, the F statistic was used to compute the variations between replicates as compared to those observed between two different samples. Based upon this method, a complete permutation test was carried out such that the expression changes observed when true replicates were analyzed, were compared to the changes that occur when any two samples were considered to be replicates. This approach helps distinguish true signal changes from those that occur by chance. Only the genes for which a p-value < 0.05 and a SD of > 1 was calculated were included in the list of differentially expressed genes for further analysis by clustering.

Dendrograms of the expression patterns and samples were computed by hierarchical clustering using Ward's Minimum Variance and Complete Linkage Methods, respectively, as implemented in R in the hclust function from the *mva* package. In both cases a Euclidian distance measure was employed. Dendrograms were combined with a heat map displaying color-coded expression signal intensities for each gene.

To subgroup the genes according to their overall transcription patterns but not signal intensities the values for each gene were centered on zero and scaled to a standard deviation of 1. The transformed values were then analyzed using Partitioning Around Medoids (PAM) and k-means clustering as implemented in R and GeneSpring 5.1 (Silicon Genetics, Redwood City, CA), respectively. PAM is a useful alternative to k-means because it minimizes the sum of dissimilarities rather than the sum of squared Euclidean distances which makes it more robust towards outliers (Kaufman and Rousseeuw, 1999). The significance of PAM results after several rounds of analysis using increasing numbers of clusters was verified by employing Silhouette Plots. The silhouette values are a measure of the degree of similarity of expression patterns in a given cluster as compared to all other patterns in the other clusters. Both PAM and silhouette are part of the cluster package in R.

Similarity search The UniGene Rn.data.gz file was parsed and read into a MySQL relational database. The fields from the Rn.data file used were: ID (UniGene cluster ID), TITLE (title for the UniGene cluster), GENE (gene symbol), PROTSIM (protein similarity/homology data of rat UniGene clusters to loci in rat and other model organisms), and SEQUENCE (sequence information of all sequences that belong to each UniGene cluster).

We generated the following probe set annotation groups: 1) probe sets not in UniGene 118, 2) probe sets which represent annotated genes, 3) probe sets which represent genes not yet annotated with at least one homolog, and 4) probe sets which represent genes not yet annotated with no homologs and a significant protein-NR BLAST alignment. To determine if a probe set was in UniGene 118 and which UniGene cluster it was in, the source accessions for each probe set were mapped to UniGene 118 clusters. UniGene gene symbol and BioConductor PubMed ID information were used to determine if a probe set represented an annotated gene. UniGene PROTSIM information was used to determine whether the gene represented by a probe set had a homolog. Probe sets corresponding to genes not yet annotated with no homologs were aligned to NCBI's protein-NR database by using the BLASTx program with a minimum e-value of 1×10^{-5}.

MIAMI compliance Both CEL and CHP data files corresponding to the Sertoli (SE1 and 2), germ cell (SG, SC and ST1 and 2, respectively) and total testicular samples (TT1 and 2) were uploaded to the EBI's ArrayExpress public data repository at (supplemental table, ref. 14) (Submission 280: "rat spermatogenesis"). CEL feature level data files of all samples (including brain and skeletal muscle) and *Q-PCR* raw data are

available at our website.

Acknowledgements

We thank R. Jenni, K. Potter, R. Phlmann, and D. Flanders (FMI), for excellent IT infrastructure support within the Basel Computational Biology Center. We thank I. Coiffec and M. Docquier for excellent technical assistance, V. Vallet-Erdtmann for helpful advice and R. Lindberg and M. Regg for having provided unpublished U34A brain and skeletal muscle microarray expression data. P. Demougin and the GeneChip platform at the Life Sciences Training Facility are supported by the Biozentrum, the Swiss Tropical Institute, the Department of Clinical and Biological Sciences of the Basel University Hospital and the Department of Clinical Research at the University of Bern. U. Schlecht, C. Wiederkehr and L. Hermida are supported by a grant from the Swiss Institute of Bioinformatics.

Chapter 2: Genome-wide Analysis of *ABF1*-target Genes During Mitotic Growth and Meiotic Development

The work reported in this chapter represents the result of the close collaborative effort. For clarity all contributions are reported here. Roland Züst performed the YPD part of the expression profiling experiment, Philippe Demougin has hybridized meiotic samples of the expression profiling experiment, and Christa Niederhauser-Wiederkehr has performed data analysis of the mitotic expression data. Thomas Aust has generated the polyclonal α-Abf1p antibody, and Mahamadou Faty has characterized the antibody. Nicolas Robine (from the laboratory of Alain Nicolas at the Institute Curie, Paris) has normalized the ChIP-CHIP data. Ionas Erb and Erik van Nimwegen have contributed to this work by defining the genome-wide binding sites of Abf1p *in silico* using MotEvo.

Abstract

Autonomously replicating sequence Binding Factor 1 (*ABF1*) is an essential sequence-specific DNA binding protein in *Saccharomyces cerevisiae* involved in diverse nuclear events including transcriptional regulation, gene silencing at the mating type locus, DNA replication, and chromatin remodeling. Abf1p has been demonstrated to interact with components of the nucleotide excision repair complex, the mRNA export machinery and also factors involved in chromatin remodeling, such as Esa1p, Hmo1p and Isw2p. Abf1p participates in the transcriptional control of genes involved in a broad range of cellular processes, including amino acid metabolism, ribosome biogenesis, and meiotic development. Here, we report a comprehensive whole genome approach aiming at the identification of *ABF1*-target genes in vegetatively growing and sporulating yeast cells. The approach combines whole-genome expression profiling experiments with genome-wide binding assays (ChIP-CHIP) and *in silico* prediction of potential binding sites. This study identifies a total of 504 genes that are transcriptionally dependent on *ABF1*, more than half of which are *bona fide* direct targets. Many of these genes are essential for vegetative growth or are involved in meiotic development. Several of these genes have previously been described as being essential for sporulation (such as *SPO14, SRV2* and *VMA6*) and therefore could account for the observed sporulation defect that we observe in *abf1-1* mutant cells. Furthermore, we found that two members of the septin family (*CDC3* and *CDC10*) are direct target genes of Abf1p-regulation which suggests a role for Abf1p in cytokinesis. This notion is further substantiated by the observation, that chitin localization and septin ring formation are perturbed in *abf1-1* mutants, which in turn can explain the cellular morphology of this strain at the restrictive temperature.

Introduction

Autonomously Replicating Sequence (*ARS*) Binding Factor 1 (*ABF1*) belongs to a group of proteins referred to as General Regulatory Factors (GRFs), which share a number of different characteristics. They are essential, abundantly transcribed, and function as enhancers of transcriptional events: their binding motifs usually have little intrinsic activating capacities but instead amplify the effect of neighboring regulatory sites. They can enhance both activation and repression of transcription and they are involved in diverse cellular processes. Indeed, a conserved DNA-binding motif for *ABF1* has been found at different locations throughout the yeast genome including the silenced mating type locus, various ARSs, and the regulatory sequences of around 80 genes (Loo *et al.*, 1995; Diffley and Stillman, 1988; Lee *et al.*, 2002). A number of these Abf1p-binding sites (UAS, Upstream Activating Sequence) have been shown to be essential in mediating diverse nuclear events. For example, parts of the Autonomously Replicating Sequences of *ARS1, ARS120* and *ARS121* which contain UASs have been shown to be bound by Abf1p *in vitro* (Diffley and Stillman, 1988; Francesconi and Eisenberg, 1989) and to be essential for stability of *ARS/CEN* plas-

Table 2. *Saccharomyces cerevisiae* strains used in this study

Strain	Genotype	Source
KM7	*MATa leu2-3 his4-519 ade1-100 ura3-52 GAL-ABF1::URA3*	(Gonçalves *et al.*, 1992)
JCA30	*MATa trp1Δ his3Δ200 ura3-52 lys2-801 ade2-1 gal ABF1*	(Rhode *et al.*, 1992)
JCA31	*MATa trp1Δ his3Δ200 ura3-52 lys2-801 ade2-1 gal abf1-1*	(Rhode *et al.*, 1992)
JCA40	*MATα trp1Δ his3Δ200 ura3-52 lys2-801 ade2-1 gal ABF1*	(Rhode *et al.*, 1992)
JCA41	*MATα trp1Δ his3Δ200 ura3-52 lys2-801 ade2-1 gal abf1-1*	(Rhode *et al.*, 1992)
MPY170	*SK1 MATa/MATα ho::LYS2 ura3 lys2 leu2::hisG arg4-Nsp/arg4-Bgl, his4x::LEU2-URA3 /his4B::LEU2*	(Primig *et al.*, 2000)
MPY284	*MATa/MATα trp1Δ his3Δ200 ura3-52 lys2-801 ade2-1 ABF1/ABF1*	This study
MPY283	*MATa/MATα trp1Δ his3Δ200 ura3-52 lys2-801 ade2-1 abf1-1/abf1-1*	This study
MPY285	*MATa/MATα trp1Δ his3Δ200 ura3-52 lys2-801 ade2-1 abf1-1/abf1-1 pRS416-ABF1*	This study
MPY125	*MATa/MATα trp1Δ/TRP1 his3Δ200/his4-519 ura3-52 ade2-1/ade1-100 gal ABF1/GAL-ABF1::URA3*	This study
MPY128	*MATa/MATα trp1Δ/TRP1 his3Δ200/his4-519 ura3-52 ade2-1/ade1-100 gal abf1-1/GAL-ABF1::URA3*	This study
MPY310	*MATa/MATα trp1Δ his3Δ200 ura3-52 lys2-801 ade2-1 ABF1/ABF1 CDC10-GFP::KanMX4/CDC10*	This study
MPY311	*MATa/MATα trp1Δ his3Δ200 ura3-52 lys2-801 ade2-1 abf1-1/abf1-1 CDC10-GFP::KanMX4/CDC10*	This study

mids (Rhode *et al.*, 1992). In addition, the Abf1p-binding site at the *HMRE* silent mating type locus was shown to be essential to confer gene silencing (Loo *et al.*, 1995; Diffley and Stillman, 1988). Furthermore, binding motifs within numerous promotors have been demonstrated to contribute to the transcriptional activation of genes that are involved in different cellular processes including carbon source regulation, amino acid metabolism, mitochondrial and ribosomal functions (Kratzer and Schüller, 1997; Künzler *et al.*, 1995; de Winde and Grivell, 1992; Della Seta *et al.*, 1990a; Planta, 1997; Planta *et al.*, 1995).

A number of studies have previously shown that *ABF1* participates in the activation of early and middle meiotic genes. This has been demonstrated for *HOP1*, which encodes a component of the synaptonemal complex, *SPR3*, a sporulation-specific homolog of the yeast Cdc3/10/11/12 septin family, and *SMK1*, a mitogen-activated protein kinase required for spore morphogenesis. From this it was concluded that *ABF1* plays a general role in activating early and middle meiotic genes (Vershon and Pierce, 2000). Indeed, some other sporulation-specific promoters contain Abf1p-binding sites (Ozsarac *et al.*, 1995; Gailus-Durner *et al.*, 1996). However, it was never tested whether these sites are essential for transcriptional activation, as it was shown for *HOP1, SPR3* and *SMK1* (Prinz *et al.*, 1995; Gailus-Durner *et al.*, 1996; Ozsarac *et al.*, 1995; Ozsarac *et al.*, 1997; Pierce *et al.*, 1998). Notably, promotors of these three genes also contain sequence elements which are essential for proper timing of activation. For example, the *SMK1* promotor harbors a combination of binding sites for Sum1p/Ndt80p (MSE, Middle Sporulation Element) and Ume6p (URS1, Upstream Repressor Sequence 1) which confer the gene's repression during vegetative growth, activation during mid-meiosis and repression at late stages of sporulation, respectively. When the Abf1p-bound UAS was mutated, a *SMK1*-lacZ reporter gene was not activated during sporulation anymore. In this context it has been demonstrated that the Abf1p-binding sites from the early *HOP1* and the middle *SMK1* promoters can be interchanged without altering the timing and magnitude of transcriptional activation. In addition, the early *HOP1* promoter was converted into a middle gene when its URS1 was replaced with the MSE from the *SMK1* promotor (Pierce *et al.*, 1998). This led to the notion that Abf1p contributes to the strength rather than the timing of induction.

The interplay between Abf1p and other transcription factors (i.e. Ume6p or Ndt80p as

mentioned above) to control transcriptional activity has been documented for numerous genes (Kratzer and Schüller, 1997; Künzler *et al.*, 1995; Martens and Brandl, 1994; Mulder *et al.*, 1995; de Winde and Grivell, 1992; de Boer *et al.*, 2000). That this is also connected to Abf1p's intrinsic ability to remodel the chromatin has recently been shown for two bi-directionally transcribed genes *UGA3* and *GLT1* (Ishida *et al.*, 2006). When yeast is grown on a low-quality nitrogen source, or under amino acid deprivation, the expression of both genes is induced through the action of the two global transcriptional modulators, Gln3p and Gcn4p, that bind to a region of the bi-directional promoter. Mutations in the Abf1p-binding site also located in this promotor differentially affect transcription of *UGA3* and *GLT1*, and thus alter the overall relative expression (Ishida *et al.*, 2006).

Abf1p's ability to remodel chromatin has been attributed to its C-terminal activation domain (AD), which can be further dissected into two important clusters of amino acid residues (CS1 and CS2). CS1 specifically participates in transcriptional silencing and nuclear localization of Abf1p (Miyake *et al.*, 2002; Loch *et al.*, 2004). CS2 in turn is required for activation of transcription, DNA replication, and chromatin remodeling (Miyake *et al.*, 2002). The N-terminal two-thirds of the protein harbor a DNA-binding domain (DBD), which contains a motif related to the canonical zinc finger (Rhode *et al.*, 1992; Miyake *et al.*, 2002). Sequence-specific DNA-binding activity of Abf1p to target sequences has been attributed to its DBD (Halfter *et al.*, 1989a; Rhode *et al.*, 1992) and requires zinc ions and unmodified cysteins (Diffley and Stillman, 1989). Based on *in vivo* footprinting and DNA mutation analysis earlier studies have postulated the consensus sequence 5'-CGTNNNNNRYGAY (where R, Y and N represent A or G and C or T and random bases, respectively) as an Abf1p-binding site (Halfter *et al.*, 1989b; Einerhand *et al.*, 1995; Upton *et al.*, 1995; Della Seta *et al.*, 1990b; Dorsman *et al.*, 1990). A mutational screen for conditional mutations of *ABF1* resulted in the identification of several temperature-sensitive alleles (Rhode *et al.*, 1992). One of these alleles, *abf1-1*, has a point mutation in its DNA-binding motif and abolishes the protein's ability to bind to target DNA *in vitro* (under unphysiological conditions) already at the permissive temperature and *in vivo* at the restricitive temperature. This has been demonstrated for the *ARS1* element *in vitro* and a number of promotors *in vivo* (Rhode *et al.*, 1992; Schroeder and Weil, 1998a). Notably, the latter experiments showed that occupancy of *abf1-1* at target binding sites is already lost after a few minutes upon transfer of cells from permissive to restrictive temperature (Schroeder and Weil, 1998a). The growth rate of the *abf1-1* mutants at the permissive temperature (25°C) is nearly identical to that of the isogenic *ABF1* strain. At the semi-permissive temperature (30°C) *abf1-1* mutant cells grow slowly and growth completely arrests at temperatures ≥32°C (Rhode *et al.*, 1992).

abf1-1 is a useful tool for the analysis of cellular functions governed by *ABF1*, such as DNA replication, plasmid stability (Rhode *et al.*, 1992), and transcriptional activation (Schroeder and Weil, 1998b). A recent microarray-based study compared expression profiles of wild-type and *abf1-1* mutants growing in rich medium. In this experiment 50 and 36 genes were identified as being activated and repressed, respectively, by functional *ABF1* (Miyake *et al.*, 2004) in vegetative cells. This finding was surprising because previous genome-wide binding assays had revealed that Abf1p binds to the promotors of approximately 500 genes (p-value <0.005) in vegetatively growing cells (Harbison *et al.*, 2004; Lee *et al.*, 2002). Several reasons were suggested for the disparity between the large number of Abf1p-bound loci in the genome and the relatively small number of Abf1p-controlled (Miyake *et al.*, 2004), for example, the redundancy between Abf1p and other transcription factors bound at the same promoter regions which may compensate loss of *ABF1*-functionality. Miyake *et. al.* (2004) also mentioned the *hit-and-run* model as a possible explanation. This model was postulated several years earlier (Schroeder and Weil, 1998b) on the basis of the observation that several *ABF1*-dependent genes keep on being active, even if Abf1p is not detected at their promotors by *in vivo* footprinting assays.

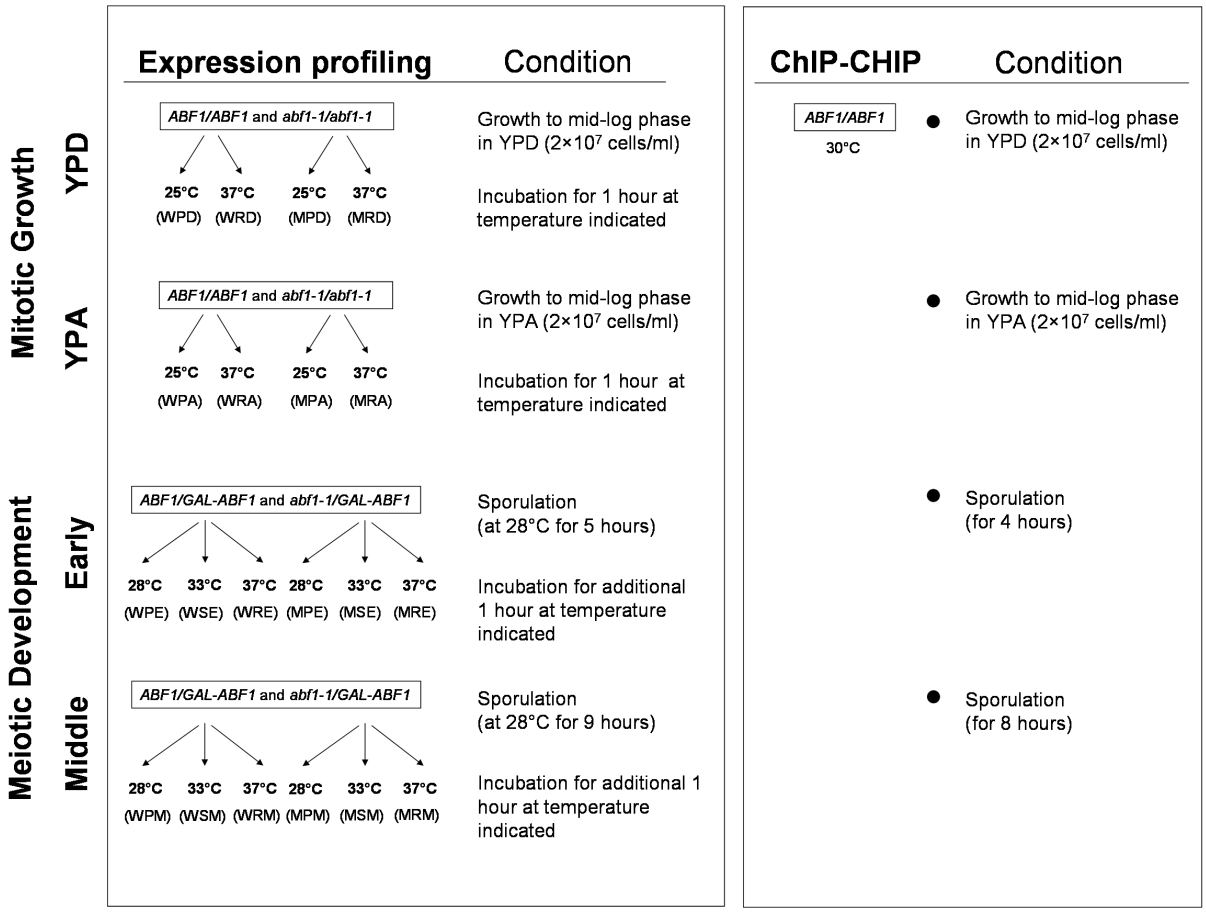

Figure 9: Flow chart of the experimental protocol. Boxes indicate the genotype of strains used for expression profiling. Arrows refer to sample splitting prior to shifting to the temperatures which are indicated. Black circles indicate timepoints at which samples for the ChIP-CHIP experiment were taken. Abbreviations were used to describe sample types: W, wild-type; M, mutant; P, permissive temperature (25°C during mitosis and 28°C during meiotic development); S, semi-permissive temperature (33°C); R, restrictive temperature (37°C); D, YPD; A, YPA; E, early sporulation stage; M, middle sporulation stage.

In this study we have examined mitotic and meiotic phenotypes of *abf1-1* mutant cells. This characterization included growth properties of vegetative cultures incubated at permissive, semi-permissive and restrictive temperature and different meiotic landmarks. Furthermore, we have combined two whole genome approaches to identify *ABF1*-target genes in vegetatively growing and sporulating yeast and included *in silico* predictions of potential binding sites. The combination of ChIP-CHIP assays with expression profiling experiments comparing wild-type and mutant *abf1-1* cells allowed us to define different categories of regulation by *ABF*. For example, genes whose transcript abundance dropped in mutant cells at the elevated temperature and whose promotors were found to be bound by Abf1p and/or that contained a predicted binding motif were defined as *bona fide* direct targets. Among the 504 genes whose expression was shown to depend on the presence of functional *ABF1* we identified 352 genes such direct target genes. This list includes numerous genes essential for vegetative growth or involved in meiotic development, such as *TAO3*, one of the recently described quantitative trait loci governing meiosis, or *SPO14* whose dependency on *ABF1* gives a possible explanation for the sporulation phenotype of *abf1-1* mutant cells. Furthermore, we also found that the activity of two septins (*CDC3* and *CDC10*) which are required for cytokinesis are dependend on Abf1p. This finding might explain the aberrant morphology, chitin mislocalization, and perturbation of septin ring formation observed in *abf1-1* mutant cells.

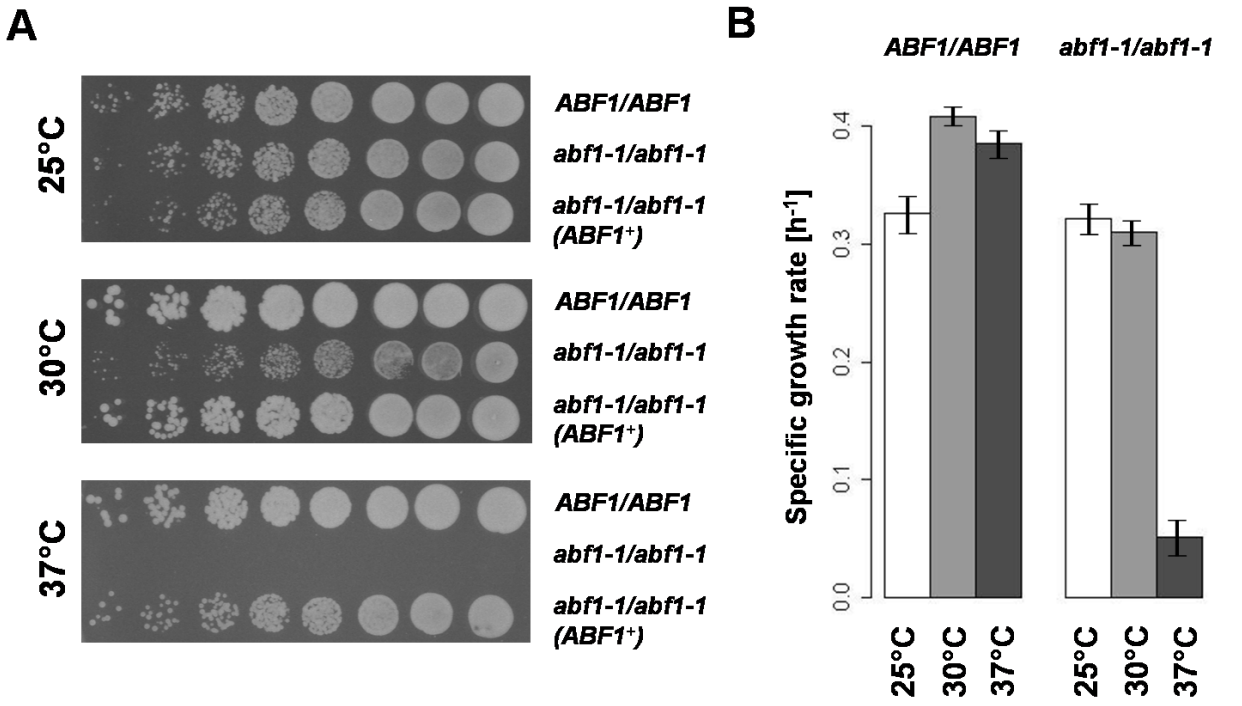

Figure 10: Growth properties of wild-type (*ABF1/ABF1*, MPY284), mutant (*abf1-1/abf1-1*, MPY283) strains and mutant strain transformed with a plasmid harboring *ABF1* (*abf1-1/abf1-1 ABF1+*, MPY285). **(A)** Serial five-fold dilutions of wild-type, mutant cells and the complemented strain were spotted on YPD plates and grown for 2 days at the temperatures indicated. **(B)** Specific growth rate of wild-type and mutant strains in liquid YPD medium at the temperatures indicated, monitored over a period of nine hours. Mean values from three independent experiments are shown. Bars represent standard deviations.

Results

Experimental design

The aim of this study was to identify target genes of *ABF1* in vegetatively growing and sporulating cells to better understand Abf1p's role in coordinating these events. To this end we have combined three experimental approaches: expression profiling experiments, genome-wide binding assays (ChIP-CHIP), and *in silico* prediction of potential binding sites of *ABF1*. The experimental setup is summarized in Figure 9. The combination of the three complementary experiments has several advantages as compared to single approaches. The most important advantage is that the data sets are complementary. This allows optimizing the filtration criteria of every single data set to gain efficiency without losing accuracy of target gene discovery. The presence of a potential binding site in a promotor (detected by *in silico* prediction algorithms) or binding of Abf1p to that promotor (monitored by protein-DNA binding assays) provides additional confidence to a putative *ABF1*-target gene identified by expression profiling. Furthermore, a key advantage of motif prediction is that it, apart from confirming a region that is found to be bound by Abf1p in the ChIP-CHIP assay, also yields the exact location of binding in this sequence region. This site can then be further validated and studied by different biochemical and genetic means, such as electro-mobility shift assays, *in vivo* footprinting, chromatin-immunoprecipitation, or mutational analysis.

In order to detect genes whose expression depends on the presence of functional *ABF1* we compared expression profiles of diploid wild-type and temperature-sensitive *abf1-1* mutant cells. These strains were derived by crossing haploid cells isolated in an earlier study on *ABF1* (Rhode *et al.*, 1992) (Table 2, MPY284 and MPY283). Cell cultures were grown to mid-log phase at permissive temperature, then split and grown at the permissive (25°C) or the restrictive temperature (37°C), respectively, for one hour prior to harvesting. As growth condi-

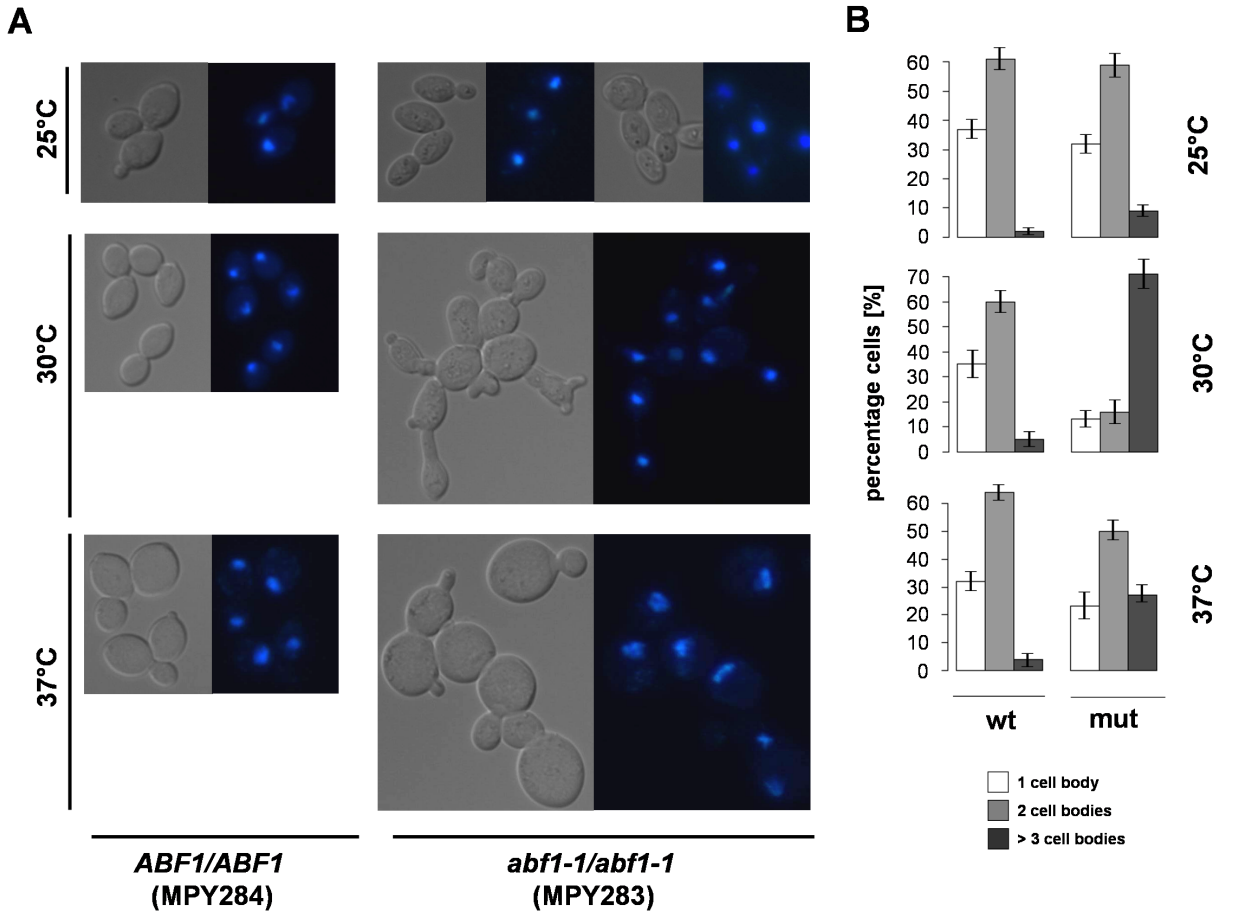

Figure 11: Arrest phenotypes of wild-type (*ABF1/ABF1*, MPY284), mutant (*abf1-1/abf1-1*, MPY283) strains. **(A)** Differential interference contrast (Nomarski) and DAPI imaging of wild-type and mutant strains that were grown at 25°C to early log phase, then divided into three cultures, and subsequently incubated at 25°C, 30°C and 37°C for 8 hours. Cells were harvested, fixed, and nuclei were DAPI-stained. **(B)** Quantification of cell separation defects of *abf1-1* mutants. Cells were treated as described above, sonicated, and counted according to their morphology. Mean values from three independent experiments are shown. Bars represent standard deviations.

tions we have chosen YPD (rich medium containing nitrogen sources and glucose) and YPA (containing nitrogen sources and acetate instead of glucose). We included the acetate experiment to detect Abf1p-regulated genes whose expression is either repressed by glucose or induced by acetate, and therefore would not be detected in an experiment based on rich medium only. There were two reasons why we have included mitotic samples (alongside with the meiotic samples) in our study. First, we wanted to gain knowledge about *ABF1*-target genes as comprehensive as possible, including different growth and differention conditions. Also, we were interested in how Abf1p coordinates mitotic growth and meiotic development. Second, many genes which are involved in or are essential for meiotic development are not meiosis-speficially expressed but are already expressed under mitotic growth conditions (Schlecht and Primig, 2003) and were therefore expected to be identified in the YPD and YPA data set.

We also aimed at the identification of genes whose activity depends on *ABF1* during meiotic development. As expected, the strains used to study mitotic roles of *ABF1* proved to progress through meiotic development and sporulation very slowly and with low efficiency, because their genetic backgrounds were close to *S288C* whose genome was sequenced (Deutschbauer and Davis, 2005). Diploid strains with one *ABF1* or *abf1-1* allele and harboring one galactose-inducible copy of *ABF1* (Table 2, MPY125 and MPY128) had better sporulation characteristics resembling that of

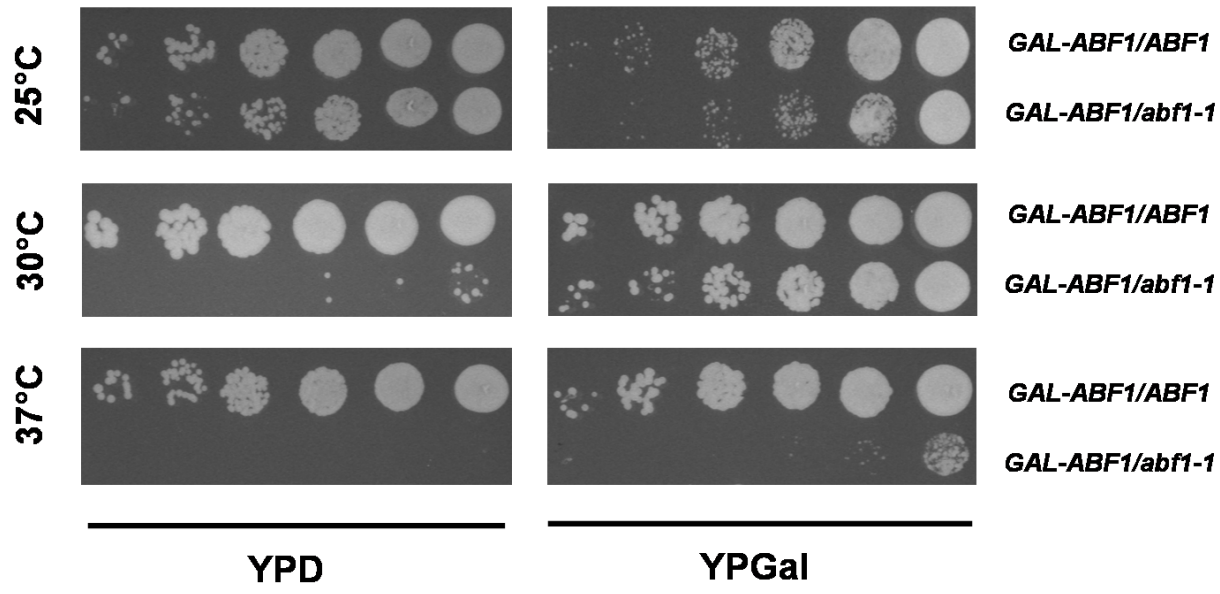

Figure 12: Growth properties of wild-type (*ABF1/GAL-ABF1*, MPY125), and mutant (*abf1-1/GAL-ABF1*, MPY128) strains. **(A)** Serial five-fold dilutions of both strains were spotted on YPD plates and grown for two days on YPD (P-GAL off) or YPGal (P-GAL on) at the temperatures indicated.

W303. For the meiotic expression profiling experiment we have chosen 28°C as a basal temperature for sporulation. After 5 and 9 hours of sporulation (corresponding to early and middle meiotic clusters in W303 as previously defined (Primig *et al.*, 2000) the cultures were shifted to 33°C and 37°C or left at 28°C for one hour prior to harvesting. The semi-permissive temperature of 33°C was chosen because temperatures higher than 34°C have previously been shown to lower the abundance of meiosis-specific transcripts during meiotic development. We further included samples shifted to 37°C as negative controls.

In parallel, we performed an *in silico* approach to predict target sequences in the yeast genome that *can* potentially be bound by Abf1p. This approach of regulatory site prediction was based on the comparison of genome sequences from five related yeast species (Cliften *et al.*, 2003; Kellis *et al.*, 2003) and led to the identification of evolutionary conserved motifs of *ABF1*. Additionally, genome-wide binding assays (ChIP-CHIP) identified those target sites that *are* actually bound by Abf1p in the living cell. Our genome-wide binding assays simultaneously measured binding of Abf1p to ∼6000 coding and ∼6000 non-coding genomic sequences (including ARS, centromer and telomer sequences) in an SK1 strain growing in rich media containing either fermentable (glucose) or non-fermentable carbon sources (acetate) or undergoing early (4 hours) or middle (8 hours) stages of meiotic development (Figure 9). We have chosen SK1 (Table 2, MPY170) for the study, because this strain sporulates synchronously and with high efficiency and therefore allowed high-resolution mapping of Abf1p-binding during sporulation. In order to exclude systematic errors in our ChIP-CHIP measurements we included three types of control experiments: a) dye swap labeling (where labeling dyes for input and precipitated DNA were interchanged), b) mock IP using an unrelated antibody, and c) input-over-input samples (where input DNA was labeled with the two dyes prior to hybridization). These control experiments allowed us to survey and minimize experimental variation caused by non-biological sources, such as a) labeling biases of input and precipitated DNA, b) preferential amplification of DNA prior to hybridization, and c) non-specific antibody interaction during immunoprecipitation.

Phenotypes of *abf1-1* mutant strains

For our study we first constructed a diploid strain homozygous for temperature-sensitive

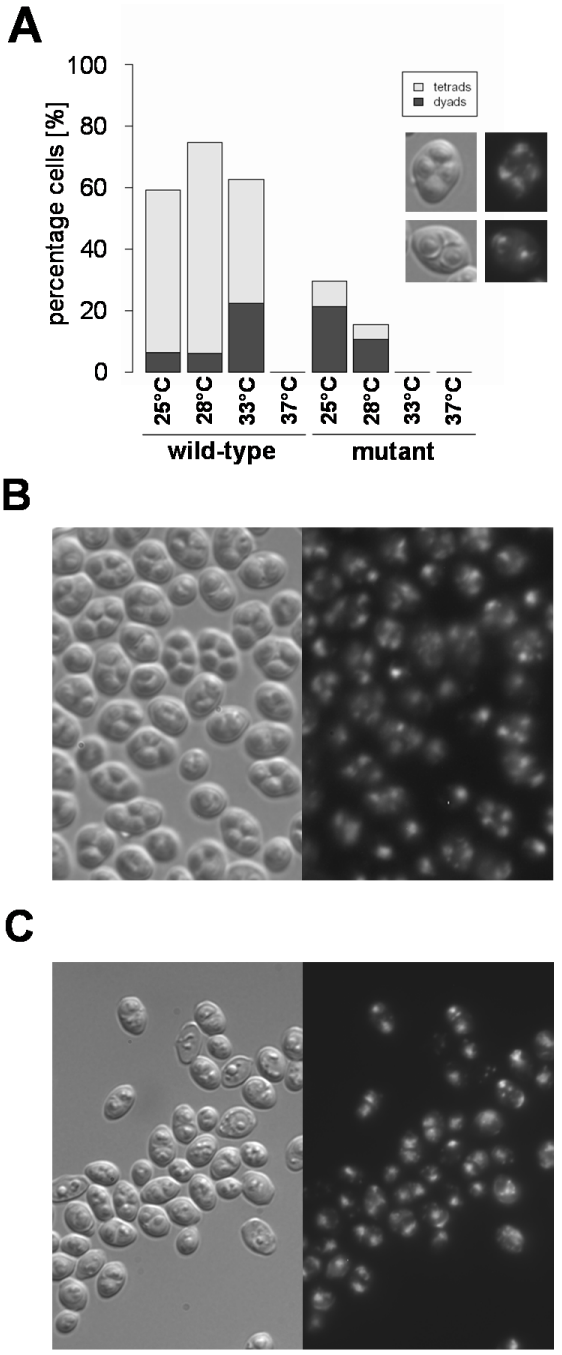

Figure 13: Sporulation properties of wild-type (*ABF1/GAL-ABF1*, MPY125) and mutant (*abf1-1/GAL-ABF1*, MPY128) strains. **(A)** Graph showing production of dyads (dark gray) and tetrads (light gray) after 5 days at the temperatures indicated. **(B)** Differential interference contrast (Nomarski) and DAPI imaging of wild-type cells after 5 days sporulation at 25°C. **(C)** Nomarski and DAPI imaging of mutant cells after 5 days sporulation at 25°C.

abf1-1 and an isogenic wild-type strain. The strains were derived from haploid JCA30, JCA31, JCA40 and JCA41 isolated in a previous study (Rhode *et al.*, 1992) (Table 2). Wild-type and mutant strains (MPY284 and MPY283, respectively) were able to grow on solid medium at the permissive (25°C) and the semi-permissive (30°C) temperatures. However, only wild-type cells were able to grow at the restrictive (37°C) temperature (Figure 10, panel A). Growth at this temperature was restored in mutant cells by inserting a plasmid harboring wild-type *ABF1* along with its promotor (MPY285).

As expected, growth rate was slowed down in mutant cells already at the semi-permissive temperature as compared to the wild-type. This has been demonstrated previously for haploid cells harboring the *abf1-1* mutation (Rhode *et al.*, 1992). We have reproduced this observation with cells growing on solid medium and in liquid culture (Figure 10, panel A and B). The morphology of the mutant was quite distinct from that of the wild-type under the same growth conditions. Cells did not separate properly forming bodies of more than 3 cells; elongated buds appeared and cell size increased (Figure 11, panel A and B). These phenotypical peculiarities of *abf1-1* mutant cells were also observed when incubated at 37°C, however only to a minor degree. This can be explained by the fact that at the restricitive temperature growth and division, which are prerequisites for the above-mentioned phenotypes, are slowed down. Upon transfer of mutant cells to 37°C the specific growth rate dramatically dropped (Figure 10, panel B).

Next, we have addressed the question whether there is a dosage-dependency of growth on *ABF1*. We therefore constructed strains with one galactose-inducible allele of *ABF1* derived from strain KM7 (Gonçalves *et al.*, 1992) (Table 2) combined with either a wild-type *ABF1* allele (from JCA40) or a mutant *abf1-1* allele (from JCA41). Growth properties of the resulting wild-type (*ABF1/GAL-ABF1*, MPY125) and mutant (*abf1-1/GAL-ABF1*, MPY128) (Table 2) cells were assessed on solid medium (Figure 12). Interestingly, growth rate of MPY128 was strongly reduced already at 30°C when grown on YPD plates. Under these conditions only the *abf1-1* allele

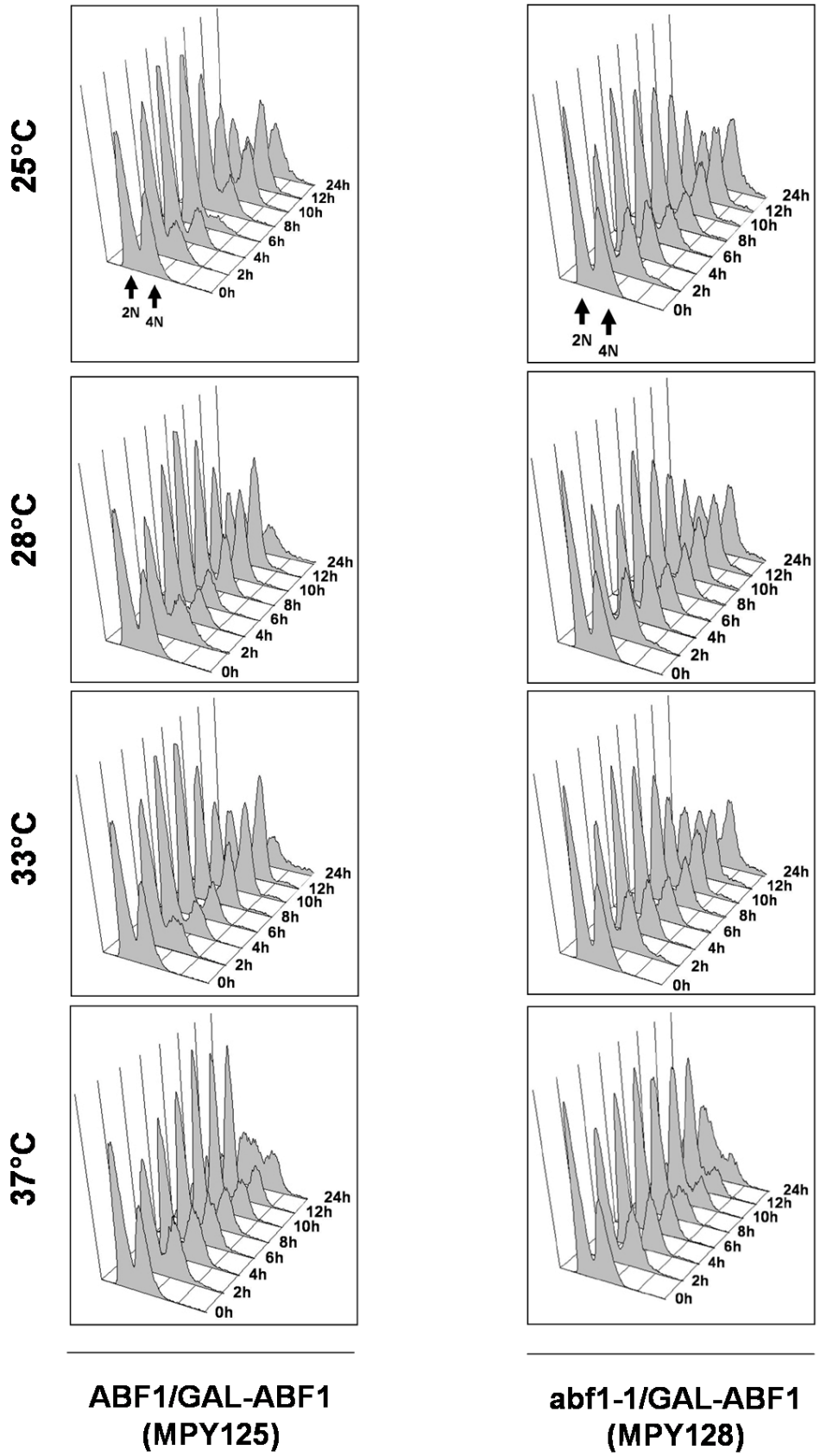

Figure 14: The kinetics of premeiotic DNA replication of wild-type (*ABF1/GAL-ABF1*, MPY125) and mutant (*abf1-1/GAL-ABF1*, MPY128) cells sporulated at the temperatures indicated as monitored by FACS analysis over a period of 24 hours. The DNA content of G1 and G2 cells is indicated as 2N and 4N, respectively.

but not the galactose-inducible allele of *ABF1* is expressed. The observed impact on growth was stronger than the one observed in mutant cells harboring two *abf1-1* alleles (MPY283) grown under similar conditions (Figure 10). Furthermore, we have observed that mutant cells grown at the restrictive temperature in the presence of galactose were also affected in their ability to grow. Under these conditions again only one of the two alleles of *ABF1* is active. In comparison, MPY128 grown at 30°C in the presence of galactose were indistinguishable from wild-type cells. This dosage-dependent effect of *ABF1* was not limited to growth only but also became apparent in the context of meiotic develoment. Already at 25°C, a temperature at which *abf1-1* is expected to be fully active, sporulation efficiency was substantially decreased in the mutant cells as compared to the corresponding wild-type (Figure 13). Notably, when the mutant strain is sporulated it tends to mostly form dyads and only few tetrads. This tendency was even more pronounced at 28°C. At 33°C meiotic development was completely abolished in the mutant cells. At 37°C neither wild-type nor mutant strains were able to sporulate (Figure 13). It is noteworthy that a diploid strain harboring two alleles of galactose-inducible *ABF1* was not able to sporulate. This is most probably because sporulation medium contains acetate but no galactose, which is needed for the expression of *ABF1* in this strain. This further underlines how important *ABF1* is for meiosis. We have also tested the ability of homozygous wild-type (*ABF1/ABF1*, MPY284) and mutant (*abf1-1/abf1-1*, MPY283) strains to undergo meiosis. Cells sporulated at 25°C for 5 days showed a very low percentage of asci in the wild-type (<20%) and a little lower in the mutant strain. This was not unexpected, because the strains' background was close to that of the sequenced *S288C* background, which has poor sporulation properties (Deutschbauer and Davis, 2005).

We asked whether the reduced ability of MPY128 to sporulate was due to a reduced rate of premeiotic DNA-replication. Previously, it was demonstrated that mitotic DNA replication in *abf1-1* strains incubated at 37°C was reduced to 50% of that seen in the wild-type strain and continued for several hours after shifting the cells to 37°C. It was hypothesized that DNA synthesis initiated prior to the temperature shift continued at a slightly slower rate and initiation at late-replicating origins was inhibited (Rhode *et al.*, 1992). In order to monitor the impact of the *abf1-1* mutation on premeiotic DNA replication we performed FACS analysis of MPY125 and MPY128 strains sporulated at 25°C, 28°C, 33°C and 37°C. Wild-type and mutant cells had similar kinetics of replication when sporulated at 25°C, 28°C, and 33°C. More than 70% of MPY125 and MPY128 had undegone DNA replication after 24 hours in sporulation medium. At 37°C DNA replication was completely blocked in both strains (Figure 14).

Expression profiling of *abf1-1* mutant cells (mitosis)

In order to identify genes that are differentially expressed in wild-type (*ABF1/ABF1*, MPY284) and mutant (*abf1-1/abf1-1*, MPY283) cells during mitotic growth we performed genome-wide expression profiling experiments. Cultures of wild-type and mutant cells were grown at 25°C to mid-log phase in the presence of glucose or acetate before they were split. One culture was shifted to 37°C for one hour, the other one was left at 25°C (Figure 9). All experiments were performed in duplicate. Quality of total RNA and cRNA targets of mitotic samples is shown in Figure 15, panel A. After preprocessing (background adjustment, normalization, and summarization of probesets) of the data using the Robust Multi-array Analysis package from Bioconductor (Bolstad *et al.*, 2003; Irizarry *et al.*, 2003a) differentially regulated genes were defined on the basis of the ANalysis Of Variance (ANOVA). This statistical method computes the variations between replicates of the same sample as compared with those observed between two different samples. We identified 1760 and 2271 differentially expressed genes in cells growing in YPD and YPA, respectively, which had a p-value <0.005 and a standard deviation >0.3. Only genes that had an abundance level of more than 100 (relative fluorescence intensity), which represents an empirical and conservative estimate of the background threshold, in

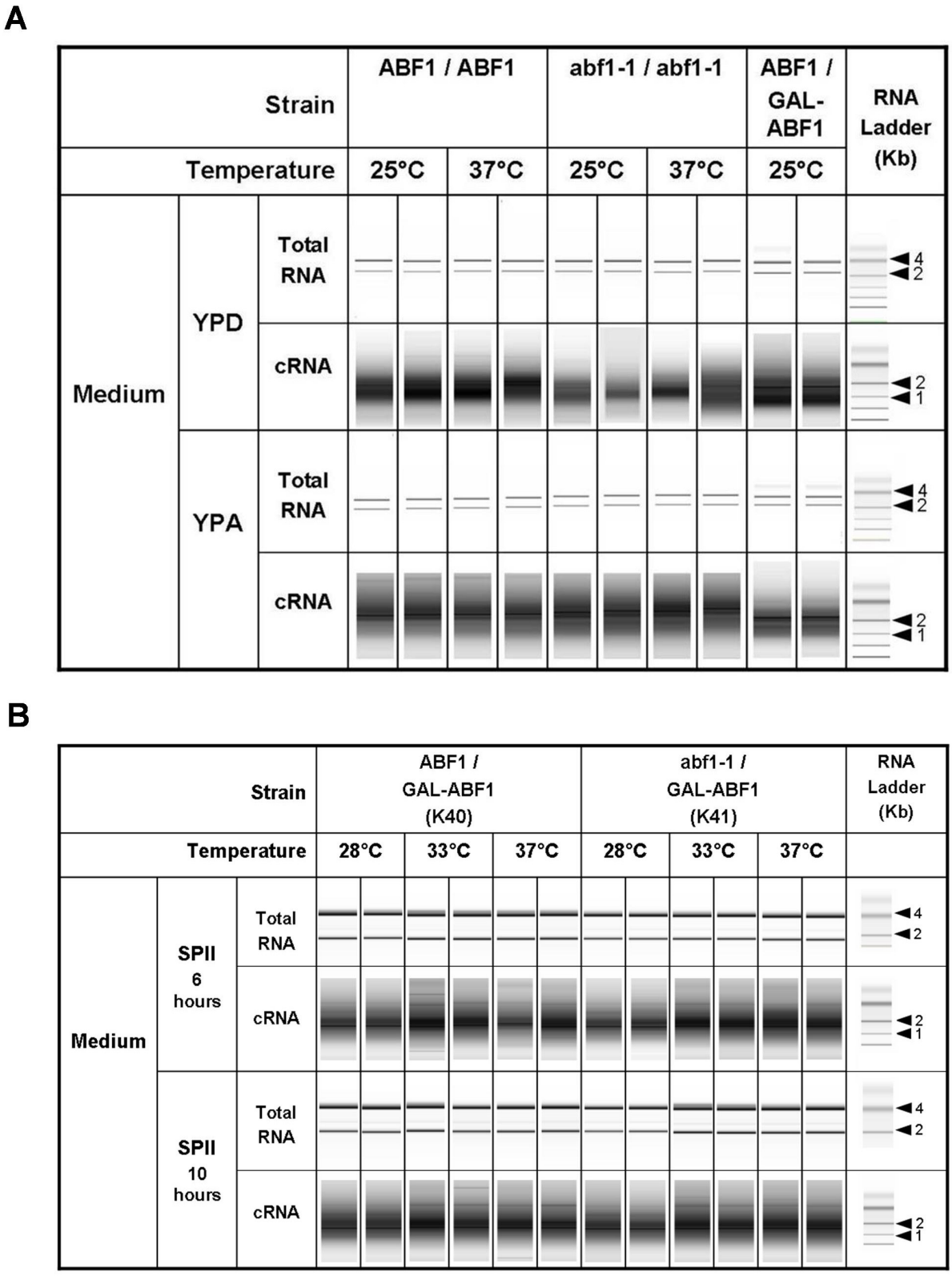

Figure 15: Total RNA and cRNA target quality assessment. **(A)** Total RNA was isolated from wild-type (*ABF1/ABF1*, MPY284) and mutant (*abf1-1/abf1-1*, MPY283) cells grown at 25°C or incubated at 37°C for one hour in the presence of YPD or YPA as indicated. **(B)** Total RNA was isolated from wild-type (*ABF1/GAL-ABF1*, MPY125) and mutant (*abf1-1/GAL-ABF1*, MPY128) cells sporulated 28°C for 5 and 9 hours prior to incubation at 28°C, 33°C, and 37°C for one additional hour as indicated. RNA was analyzed for their concentration and overall length using the Agilent Bioanalyzer RNA Chip. Virtual gels of the total RNA and cRNA samples are shown together with RNA ladder as size standard.

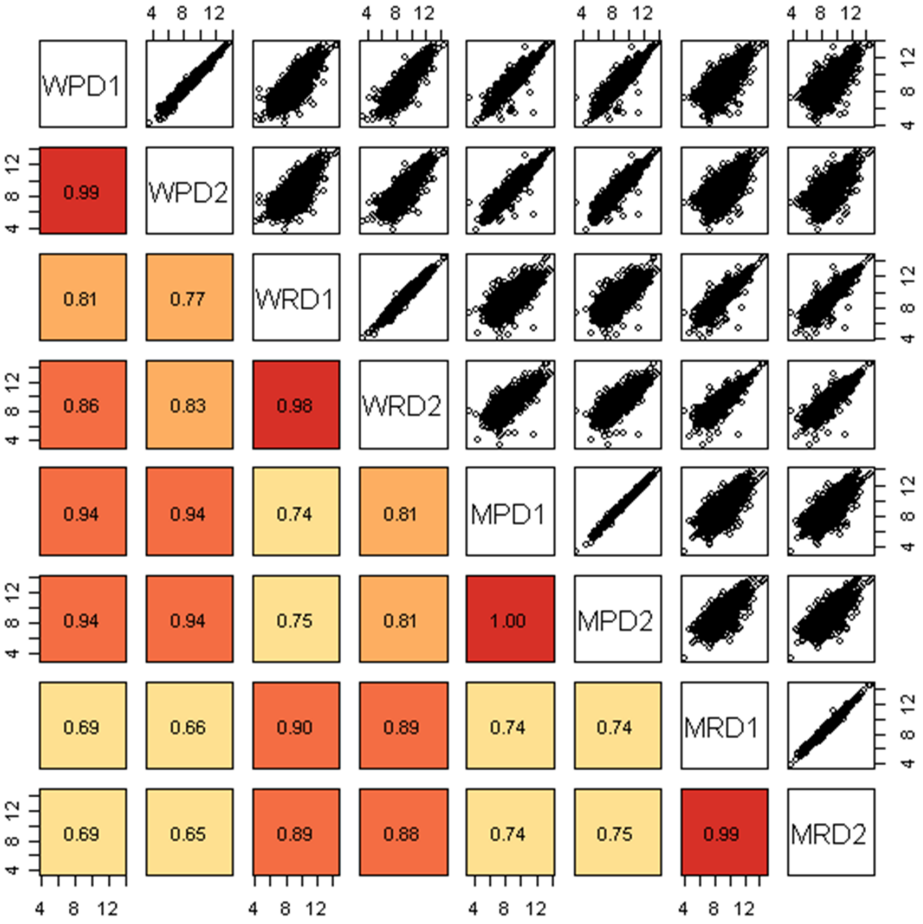

Figure 16: A scatterplot matrix summarizing a systematic comparison of the log2-scaled expression signals from 1760 differentially expressed genes in wild-type (*ABF1/ABF1*, MPY284) and mutant (*abf1-1/abf1-1*, MPY283) cells grown at the permissive or incubated at the restrictive temperature in YPD. The upper panel shows a pairwise comparison of expression values for all YPD samples as indicated (WPD, wild-type at 25°C; WRD, wild-type at 37°C; MPD, mutant at 25°C; MRD, mutant at 37°C). The lower panel depicts correlation coefficients of the pairwise comparison of expression values for all samples. The level of correlation is also represented by colors (red, high correlation; yellow, low correlation).

at least one condition were considered as expressed and included in the list. Figures 16 and 21 show scatterplot matrices with a pairwise comparison of log2-scaled expression signals of these differentially expressed genes. Correlation coefficients calculated for every sample pair showed a high reproducibility of signals in the duplicate samples (correlation coefficient $\geq$0.98). As expected, expression values of differentially expressed genes were most distinct from each other between wild-type cells grown at 25°C and mutant cells incubated at 37°C. This difference was due to a combinatorial impact of both, the *abf1-1* mutation and the temperature-shift, with the latter having a stronger effect.

Genes differentially expressed in YPD

To get an overview of the different expression patterns found among the 1760 differentially expressed genes in YPD we performed clustering on the basis of the *Partitioning Around Medoids* (PAM) algorithm (Kaufman and Rousseeuw, 1999). We decided to group genes into twelve PAM clusters (Figure 17). Although this resulted in several clusters that showed redundant patterns (clusters 7-12), it proved to be useful, because it also included clusters which revealed unexpected patterns (i.e. clusters 4 and 5). PAM clustering therefore served as an exploratory way to approach the data without having any preconception of the expected expression patterns.

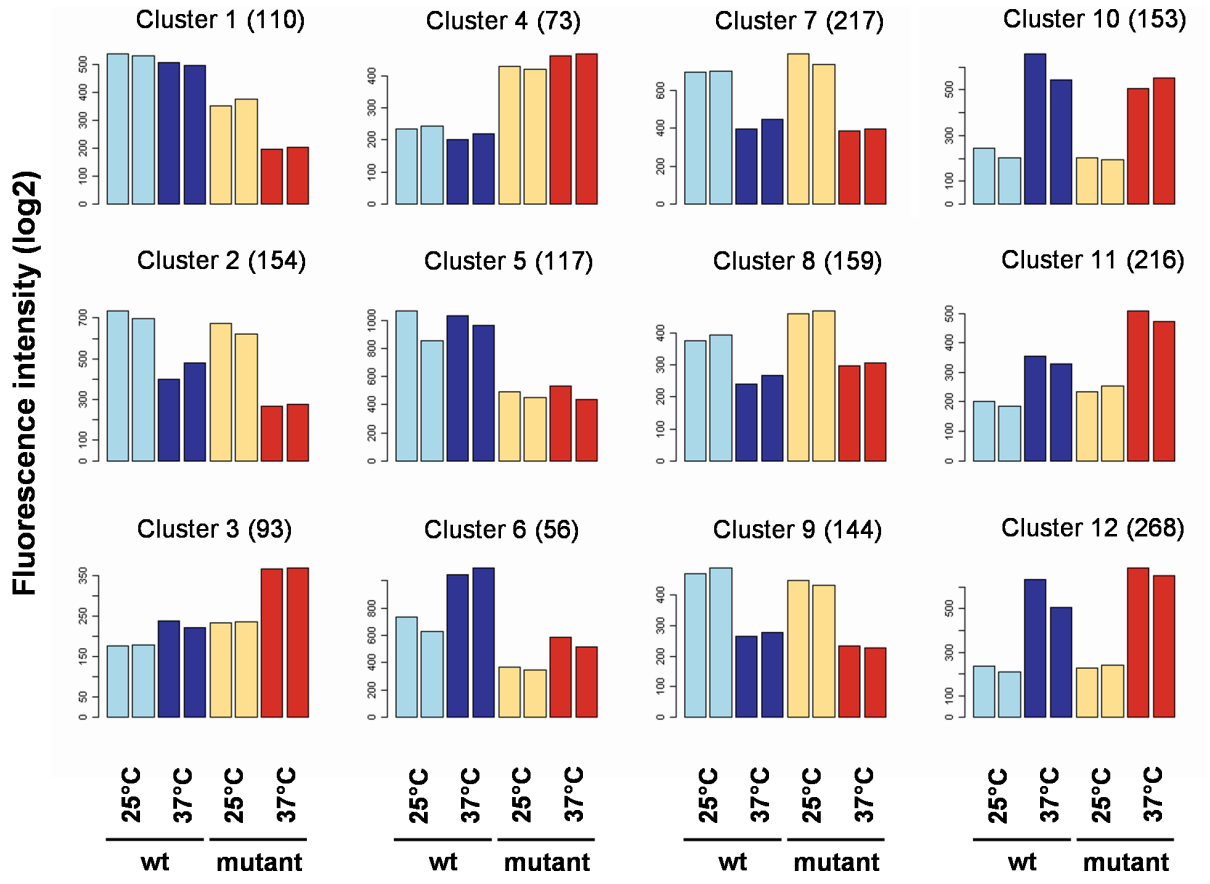

Figure 17: **(A)** Partitioning Around Medoids (PAM) clustering of 1760 differentially expressed genes in wild-type (*ABF1/ABF1*, MPY284) and mutant (*abf1-1/abf1-1*, MPY283) cells grown at the permissive or restrictive temperature in YPD into twelve expression clusters. Expression profiles of genes in the twelve clusters are represented by bars that display median expression levels of all genes in one condition.

Clusters 1-3 contained genes whose transcript levels were least (clusters 1 and 2) or most abundant (cluster 3) in mutant cells incubated at 37°C. This expression pattern suggested that these genes were positively or negatively regulated by *ABF1*, respectively. Indeed, we identified some of the previously described Abf1p-target genes in cluster 1 (*ARO3*, *HIS7*, and *TRP3*) (Künzler *et al.*, 1995; Springer *et al.*, 1997; Martens and Brandl, 1994) and in cluster 3 (*SPT2*). Furthermore, clusters 1 and 2 contained 29 of the 50 positively regulated Abf1p-target genes that were previously found in a genome-wide expression analysis of *abf1-1* mutant cells. In cluster 3 we identified 3 of the 36 genes that were previously found to be negatively regulated by Abf1p in the same genome-wide study (*GUT2, YER188W, YJR162C*) (Miyake *et al.*, 2004). An analysis of significantly enriched GeneOntology terms (p-value <0.001 with at least three genes included) revealed that cluster 1 contains numerous genes which are involved in *intracellular protein transport*, such as *ARP6, GLE2, HSE1, KAP120, MAS1, MIM1, MON2, NUP159, PAM17, SEC53, SEC72, SNF8, SRP14, TOM7, YPT52*. Additionally, genes needed for *high-affinity zinc ion transport* (*ZRT1, ZRT2, ZRT3*) were also in this cluster. Genes in cluster 2 were enriched in GO-terms related to *ribosome biogenesis*, *ribosomal RNA processing* (*DBP9, EMG1, MDN1, MRD1, MTR4, NOP53, NUG1, PWP2, RLP7, RRP1, RRP12, RRP5, RRP9, URB1, UTP18, UTP20*) and *translation* (*ALA1, GCN1, GRS1, HTS1, MIS1, MKT1, MRP51, RPP1A, TIF2, TIF3, TIF35*). This suggested that *ABF1*-regulated genes play an

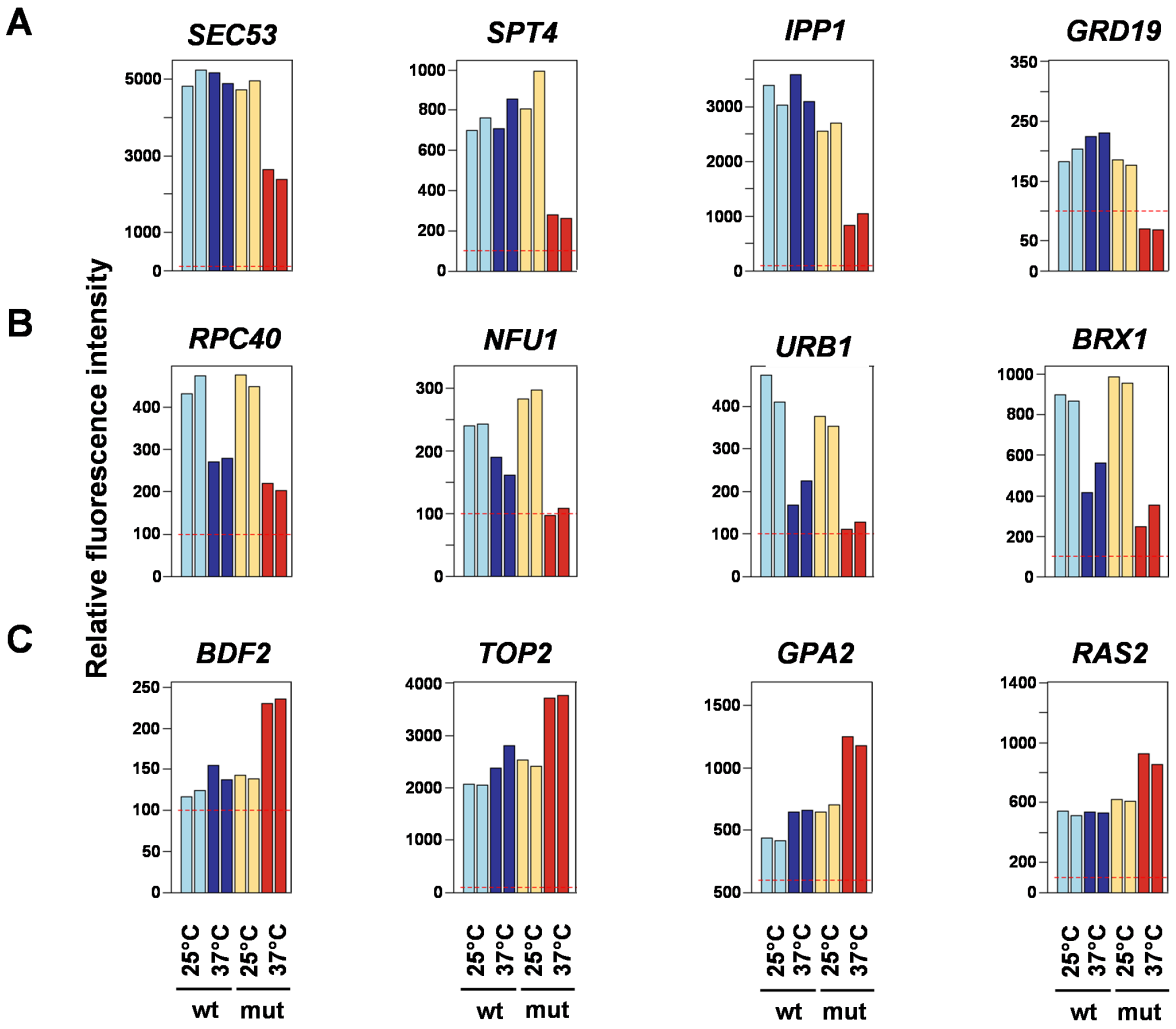

Figure 18: **(A, B and C)** Examples of expression values obtained for single genes selected from the three expression clusters of genes whose expression pattern suggests dependency on functional *ABF1* in cells growing in the presence of YPD (clusters 1-3). Barplots represent expression levels of samples as indicated. The lower detection limit is marked by the red dotted line.

important role in biogenesis of ribosomes and localization of proteins in the cell, which is in keeping with earlier observations (Planta, 1997; Planta *et al.*, 1995). Among genes in cluster 3 no enriched GO-terms were found. Figure 18 depicts expression values of four example genes from clusters 1-3 as barplots and Figure 19 shows heatmaps of expression clusters 1 and 2 with some additional information whether a given gene is essential for vegetative growth or sporulation and whether the promotor of that gene contains a predicted UAS and/or is bound by Abf1p (see below).

We identified two clusters which contained genes whose expression was higher (cluster 4) or lower (cluster 5) in the mutant as compared to the wild-type regardless of the temperature. This was unexpected, because it suggested that some genes are already affected in their transcriptional activity by the *abf1-1* mutation at a temperature at which *abf1-1* is expected to be fully functional. Interestingly, *ABF1* itself belonged to this group of genes (Figure 20), which is in keeping with a previous study that showed that *ABF1* negatively regulates its own transcription. In this study it was also found that in mutant *abf1-1* cells transcription of *ABF1* is already elevated at 25°C as compared to the wild-type (Miyake *et al.*, 2004). Furthermore, genes whose products locate to the mitochondrion and/or are involved

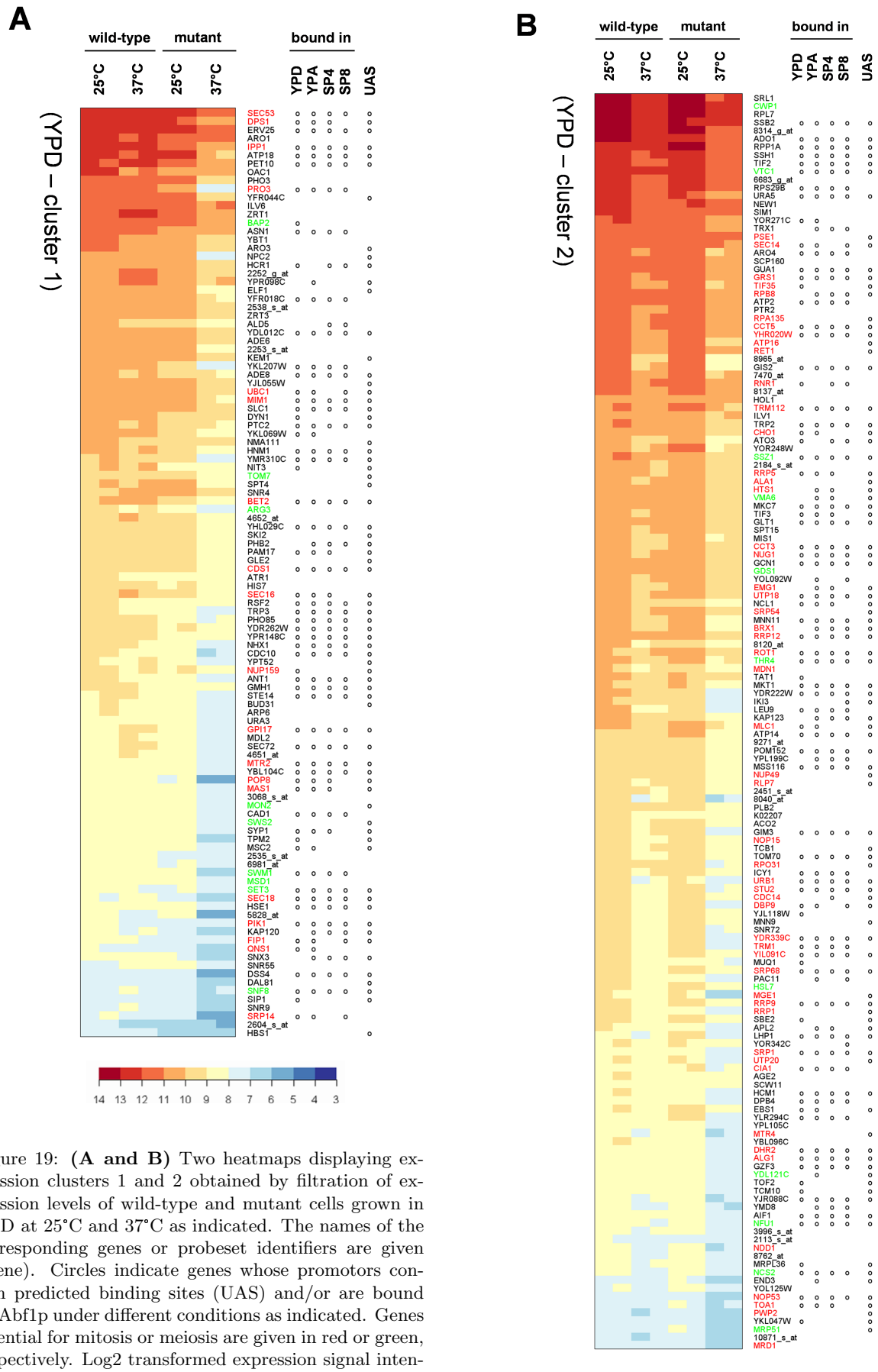

Figure 19: **(A and B)** Two heatmaps displaying expression clusters 1 and 2 obtained by filtration of expression levels of wild-type and mutant cells grown in YPD at 25°C and 37°C as indicated. The names of the corresponding genes or probeset identifiers are given (Gene). Circles indicate genes whose promotors contain predicted binding sites (UAS) and/or are bound by Abf1p under different conditions as indicated. Genes essential for mitosis or meiosis are given in red or green, respectively. Log2 transformed expression signal intensities are color coded as indicated.

in *oxydative phosphorylation* and *ATP synthesis* (e.g. *ATP17, COX7, CYT1, INH1, QCR6, QCR10, RIP1, SDH2, TIM11*) are enriched in this cluster, suggesting that mutant cells have an elevated rate of energy generation. Cluster 5 contained a large number of transposable (Ty) elements which make up a family of genetically mobile, repetitive DNA sequences in *Saccharomyces cerevisiae*. One possible explanation for this observation is that Ty-elements are in general less active in mutant cells than in wild-type. In this case there would be no connection between the activitiy of Ty-elements and *ABF1*. However, it was interesting to find *SPT2*, which is a negative regulator of Ty-controlled yeast gene expression, to be under negative transcriptional control of *ABF1* (Figure 20). The repression of transcription by *SPT2* was shown to result from an interaction between the *SPT2* gene product and Ty or delta sequences (Roeder *et al.*, 1985). However, whether there is indeed a connection between *ABF1* and transcriptional control of Ty-element has still to be elucidated. Examples of some of these genes are shown in Figure 20. Additionally, we found genes involved in leucine and arginine biosynthesis in cluster 5 (*LEU1, LEU2* and *ARG1, ARG5,6, CPA2, ECM40*, respectively).

Clusters 7-9 consisted of genes whose expression was decreased by the temperature-shift in wild-type and mutant cells to a similar extent. Previously, it has been demonstrated that cells respond to heat-stress by a substantial degradation of transcripts encoding both ribosomal proteins and rRNA synthesis and ribosome assembly factors (Grigull *et al.*, 2004). Intriguingly, we found numerous ribosomal protein mRNAs to be less abundant in wild-type and mutant cells at the restrictive temperature (e.g. *RPL7B, RPL18B, RPL36B, RPL40B, RPS1B, RPS9A, RPS22B*). The same was true for transcripts involved in ribosome biogenesis (e.g. *BRX1, DBP2, DBP9, GAR1, LRP1, NIP7, NOP1, NOP12, SMD1, URB1*). Accordingly, GeneOntology terms for *ribosome biogenesis*, *translation*, and *rRNA metabolism* were enriched among the genes of clusters 7-9.

The opposite expression pattern (elevated transcript levels in wild-type and mutant cells at 37°C) was observed in clusters (10-12).

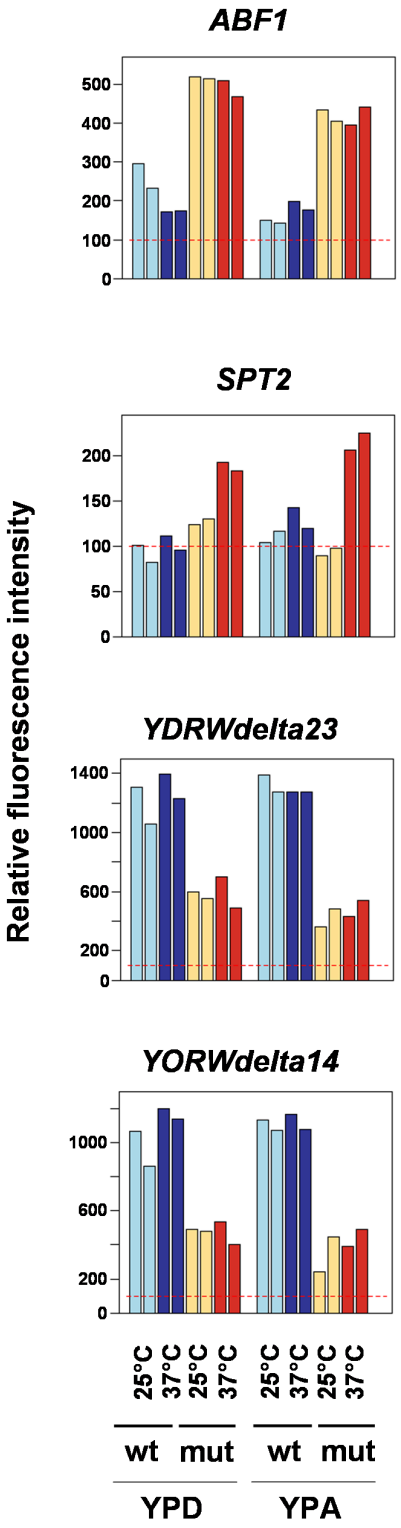

Figure 20: Examples of genes that are induced or repressed in wild-type (*ABF1/ABF1*, MPY284) and mutant (*abf1-1/abf1-1*, MPY283) cells at the restrictive temperature. Barplots represent abundance of transcripts under different growth/incubation conditions as indicated. The red dotted line indicates the empirical threshold of reliable detection.

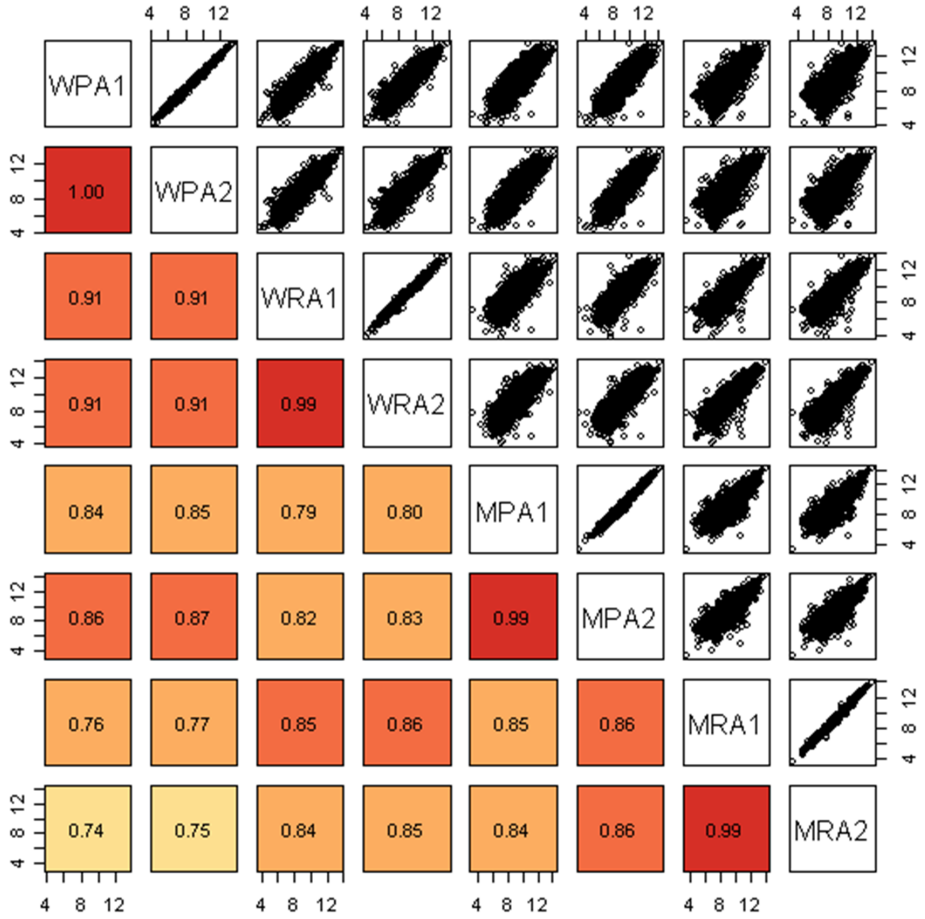

Figure 21: A scatterplot matrix summarizes a systematic comparison of the log2-scaled expression signals from 2271 differentially expressed genes in wild-type (*ABF1/ABF1*, MPY284) and mutant (*abf1-1/abf1-1*, MPY283) cells grown at the permissive or restrictive temperature in YPA. The upper panel shows a pairwise comparison of expression values for all YPA (WPA, wild-type at 25°C; WRA, wild-type at 37°C; MPA, *abf1-1* at 25°C; MRA, *abf1-1* at 37°C). The level of correlation is also represented by colors (red, high correlation; yellow, low correlation).

Strikingly, these clusters contained numerous heat-inducible genes (e.g. *GRE3, GSY2, HSF1, HSP12, HSP26, HSP30, HSP42, HSP78, HSP104, SSA1, SSA4, SYM1, TIP1*). We also found genes belonging to the GO-term *response to stress* (e.g. *DOG2, GRX1, GRX2, HAC1, HOR2, LHS1, MNN4, MPH1, MYO3, NTH2, ORM2, SIP2, SOH1, SRX1, STF2, TPS3, TRX2, UGA2, YDC1*) enriched in clusters 10-12. Taken together, genes in clusters 7-12 served as controls for the temperature-shift, because their expression pattern changed only due to the shift to 37°C. The fact that we identified groups of genes that have previously been demonstrated to react to heat-shock confirmed that our experiment was properly performed.

Genes differentially expressed in YPA

The 2271 genes differentially expressed in YPA were also grouped into twelve expression clusters using the PAM algorithm (Figure 23). Clusters 1 and 2 contained genes whose transcript levels were least abundant in mutant cells incubated at 37°C. These two clusters looked similar to the first two clusters of the YPD data set (Figure 17). Again, the expression pattern suggested that the transcriptional activity of these genes were under postitive control of *ABF1*. A GeneOntology analysis of genes in cluster 1 revealed that many genes were involved in *protein biosynthesis* and *localization* of proteins in the cell. Genes in cluster 2 were enriched in genes with the GO-term *nitrogen compound biosynthesis* (*ARO1, BAT1, GDH1, ILV2, ILV3, ILV5,*

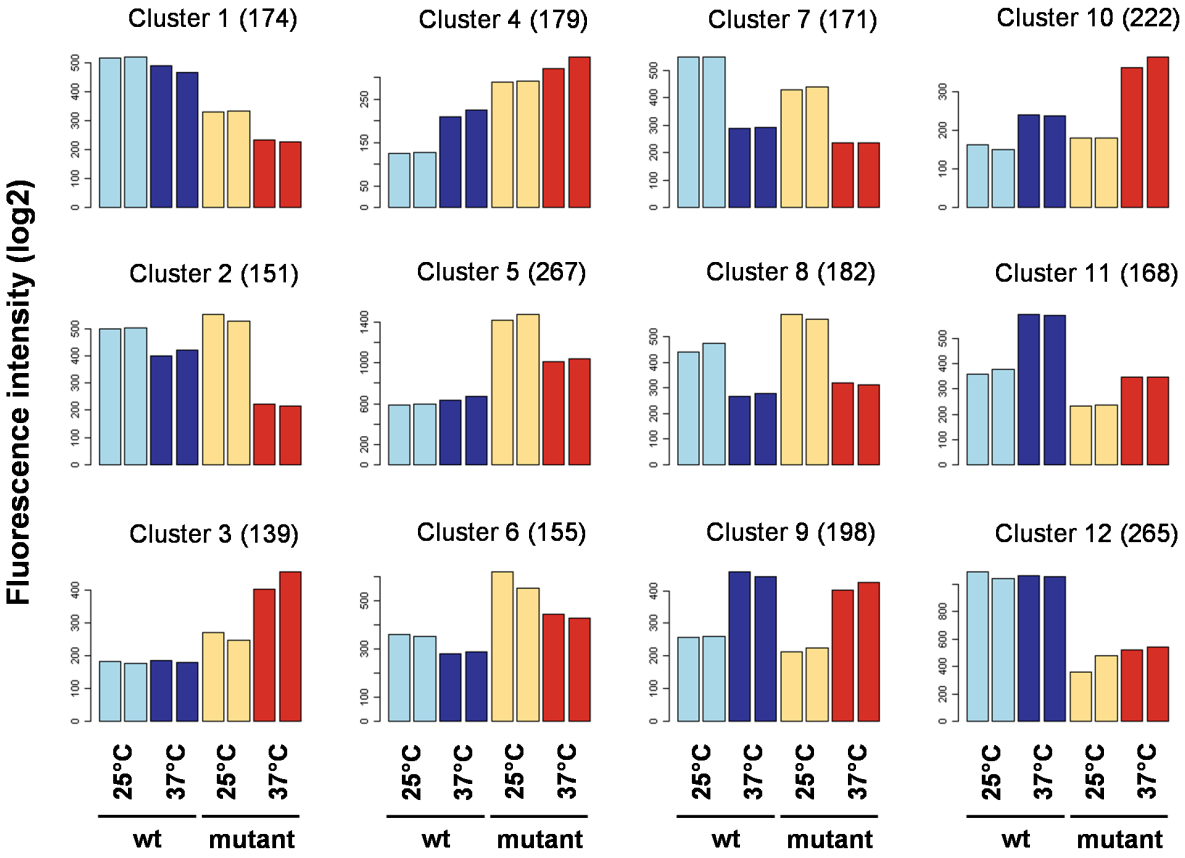

Figure 23: **(A)** Partitioning Around Medoids (PAM) clustering of 2271 differentially expressed genes in wild-type (*ABF1/ABF1*, MPY284) and mutant (*abf1-1/abf1-1*, MPY283) cells grown at the permissive or restrictive temperature in YPA. Expression profiles of genes in the twelve clusters are represented by bars that display median expression levels of all genes in one condition. Numbers in brackets indicate the number of genes belonging to the cluster.

LEU2, MUQ1, SFK1) and, interestingly, also in *cell separation during cytokinesis* (*CTS1, DSE2, SCW11*). Genes from clusters 1 and 2, that were not in the corresponding clusters in the YPD data set, are depicted in Figure 22 in heatmap format. Again a cluster with genes whose transcripts were, similar to cluster 3 in the YPD data set, upregulated in the absence of functional *ABF1* was found. Interestingly, this included numerous genes involved in *amino acid metabolism* in general and *methionine metabolism* in particular (*AAT1, CYS3, ECM17, GLT1, LYS4, MET1, MET14, MET16, MET17, MET22, MET3, MET32, MHT1, SER33, TYR1, URA2*).

In some cases cluster patterns looked very distinct from the ones observed in the YPD data set. For example, cluster 3 showed an induction of transcript abundance in wild-type and mutant cells at the elevated temperature with the mutant having a higher abundance in principle. Interestingly, many genes in this cluster were involved in *ribosome biogenesis*, *rRNA metabolism*, and *tRNA metabolism*. When comparing the transcription patterns of these genes in the presence of the two carbon sources tested, we found that the expression pattern was completely changed upon shift from YPD to YPA (Figure 24, panel A). This suggested a complex promotor architecture of these genes and underlines how important it is to assess different growth media. This was also true for genes in clusters 5 and 6, a substantial fraction of which were involved in *translation* and *ribosome biogenesis*. Again, the expression pattern of these genes was different between cells growing in the presence of YPD and YPA (Figure 24, panel B).

Clusters 7 and 8 consisted of genes whose expression was decreased by the temperature-

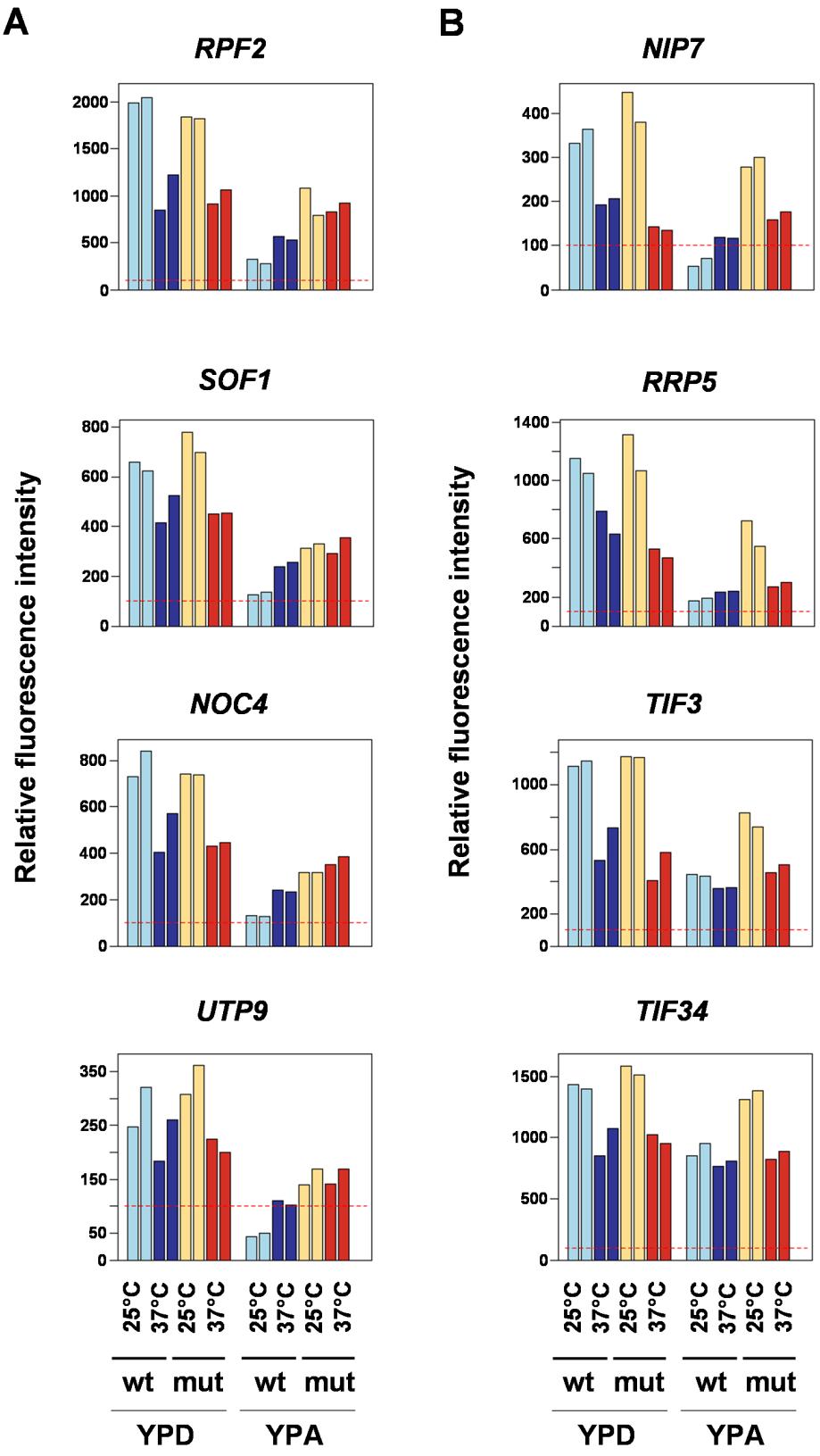

Figure 24: Examples of *ABF1* target genes that are differentially expressed in the presence of YPD and YPA. Barplots represent abundance of transcripts under different growth/incubation conditions as indicated. The red dotted line indicates the empirical threshold of reliable detection.

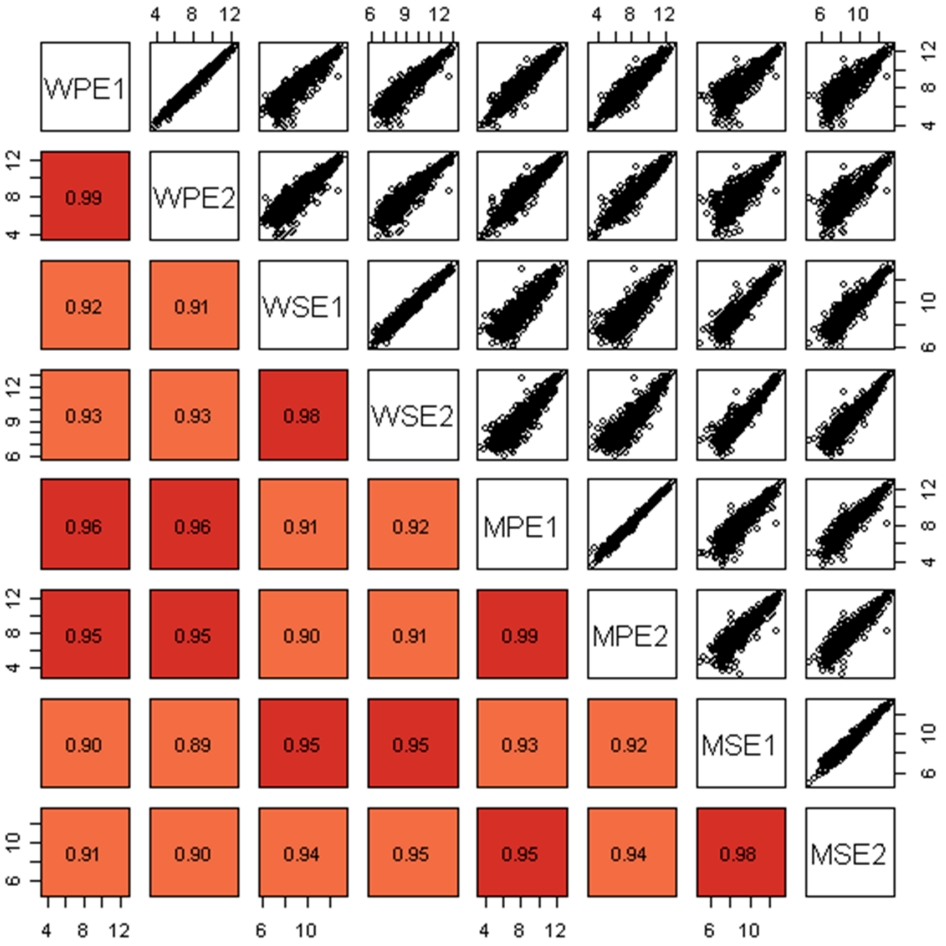

Figure 25: A scatterplot matrix summarizes a systematic comparison of the log2-scaled expression signals from 1254 differentially expressed genes in wild-type (*ABF1/GAL-ABF1*, MPY125) and mutant (*abf1-1/GAL-ABF1*, MPY128) cells sporulated at the permissive or restrictive temperature for 6 hours (early sporulation stage). The upper panel shows a pairwise comparison of expression values for all samples (WPE, wild-type at 28°C; WRE, wild-type at 33°C; MPE, mutant at 28°C; MRE, mutant at 33°C). The level of correlation is also represented by colors (red, high correlation; yellow, low correlation).

shift in wild-type and mutant cells. Genes that encode subunits of the *proteasome complex* (i.e. *PRE3, PRE8, RPN7, RPT1, UBC5, UBP6*) and genes involved in *lipid metabolism* (*ANT1, ELO1, ERG2, FAA2, FAT1, INP54*) are enriched in these two clusters. We found three clusters (9-11) with genes whose transcript levels increased in wild-type and mutant cells upon shift from the permissive to the restricitive temperature. As expected, numerous heat-inducible genes were identified in these clusters (i.e. *HSP10, HSP26, HSP30, HSP78, HSP104, HSP150*). Cluster 12 contained genes whose expression was less abundant in the mutant cells than in wild-type cells regardless of the temperature, including many Ty-elements.

Expression profiling of *abf1-1* mutant cells (meiosis)

We compared expression profiles of MPY125 and MPY128 sporulated at 28°C for 5 and 9 hours (which corresponds to early and middle meiotic stages in this strain background) with an additional incubation at 33°C for one hour. Quality of total RNA and cRNA targets of meiotic samples is shown in Figure 15, panel B. Preprocessing (background adjustment, normalization and summarization of probesets) and filtration of differentially expressed genes were carried out similarly to the mitotic experiment. This resulted in the identification of 1254 and 1768 differentially expressed genes in cells in early and middle stages of sporula-

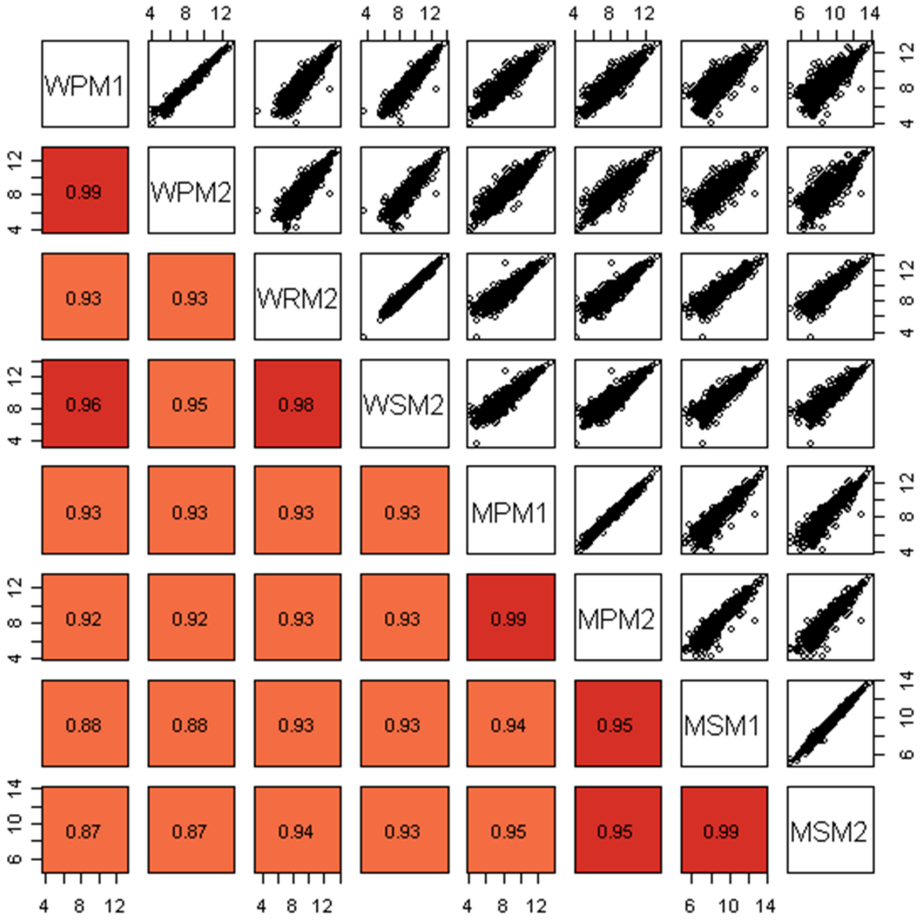

Figure 26: A scatterplot matrix summarizes a systematic comparison of the log2-scaled expression signals from 1768 differentially expressed genes in wild-type (*ABF1/GAL-ABF1*, MPY125) and mutant (*abf1-1/GAL-ABF1*, MPY128) cells sporulated at the permissive or restrictive temperature for 10 hours (middle sporulation stage). The upper panel shows a pairwise comparison of expression values for all samples (WPM, wild-type at 28°C; WSM, wild-type at 33°C; MSM, mutant at 28°C; MSM, mutant at 33°C). The level of correlation is also represented by colors (red, high correlation; yellow, low correlation).

tion, respectively. Scatterplot analysis of these genes showed that they much more correlated in their expression levels (Figures 25 and 26), suggesting that the influence of *abf1-1* mutation was less pronounced during meiosis than during mitotic growth.

The 1254 and 1768 genes that were found to be differentially expressed during early and middle stages of sporulation were subdivided into four expression clusters (Figure 27). Subdividing into more than four clusters did not yield additional distinct expression patterns, but only multiplied the already existing clusters. Notably, there was no cluster which showed least expression in the mutant cells at the elevated temperature (corresponding to clusters 1 and 2 of the two mitotic data sets - Figure 17 and 23). However, there were clusters that contained genes whose activity was lower in mutant cells than in wild-type cells regardless of the temperature. Interestingly, numerous genes that showed an *ABF1*-dependent pattern in YPD and YPA (clusters 1 and 2) were in these clusters. Figure 28 demonstrates this observation for four genes. We therefore concluded that these clusters contain genes that transcriptionally depend on *ABF1*. We found 16 and 86 genes that were in these clusters and uniquely identified in the early and middle sporulation samples, respectively. Figure 29 represents these genes as heatmaps. Cluster 1 defined for early and middle meiosis (Figure 27) were searched for enriched GeneOntology terms. The clusters included genes involved in *zinc ion transport* (*ZRT1, ZRT2*) and *nitrogen compound metabolism* (*DAL80,*

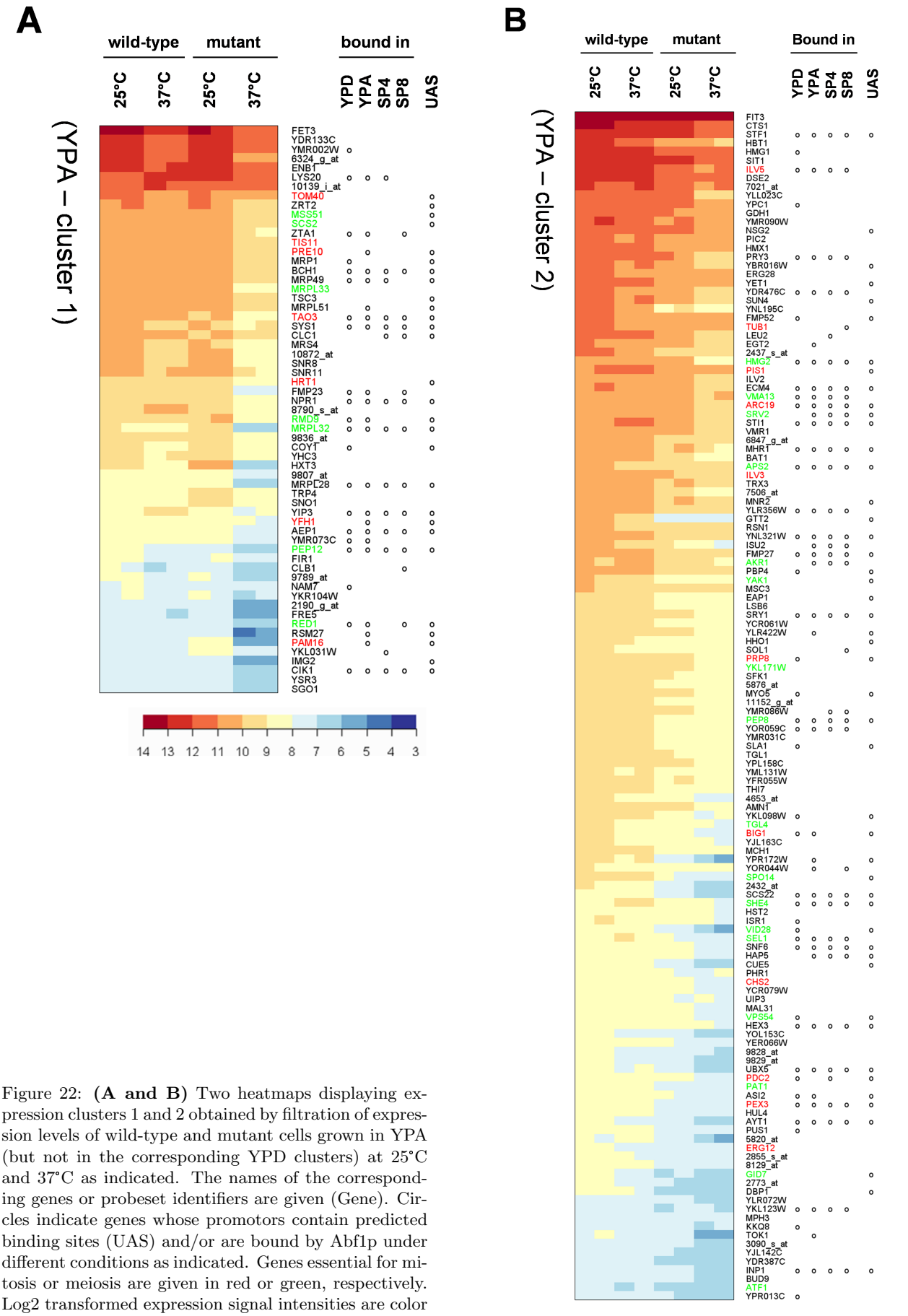

Figure 22: **(A and B)** Two heatmaps displaying expression clusters 1 and 2 obtained by filtration of expression levels of wild-type and mutant cells grown in YPA (but not in the corresponding YPD clusters) at 25°C and 37°C as indicated. The names of the corresponding genes or probeset identifiers are given (Gene). Circles indicate genes whose promotors contain predicted binding sites (UAS) and/or are bound by Abf1p under different conditions as indicated. Genes essential for mitosis or meiosis are given in red or green, respectively. Log2 transformed expression signal intensities are color coded as indicated.

DAL81, MKS1, NPR1). Numerous genes were found to be up-regulated during meiotic development, i.e. *HOP1*, *CDC3*, *CDC10*, *CDC16*, *SWM1*, and *OSW1* (Figure 41 and 42). The other 3 clusters defined for early and middle stages of sporulation included genes whose transcript abundance was up-regulated in reponse to the temperature-shift. Accordingly, these clusters involved genes with a function related to *response to stimulus* (e.g. *AGA1, AIP1, BPH1, CDC39, CDC54, CTT1, CUP1-1, DBP5, DUN1, ECM16, GLO4, GRX2, HSM3, HSP104, HSP78, HYR1, INO80, IST2, MEC1, PRP22, RAD3, RPH1, RPN4, SNQ2, SPB4, SSA4, STE11, TRX2, TRX3, XBP1, YAR1, YGK3, YOR1*), with a number of genes whose expression is known to be induced by heat-shock.

Genome-wide identification of Abf1p-bound loci

In order to monitor genomic loci bound by Abf1p in vegetatively growing and sporulating yeast we performed genome-wide ChIP-CHIP experiments. Figure 30 depicts a summary of this technique (Buck and Lieb, 2004). To this end we have raised a polyclonal antibody against Abf1p. As an antigen for immunization we used a peptide comprising amino acids 264 to 513 of Abf1p. This region has been chosen because it has minimal homology to other budding yeast proteins as confirmed by a BLAST search. Specificity of the antibody against its target protein is of crucial importance in ChIP-CHIP experiments because of the sensitivity of microarray analysis. For example, non-specific cross-reactions with chromatin will result in a high level of background hybridization to the microarray (Horak and Snyder, 2002). We have characterized the newly generated polyclonal antibody by various biochemical means in order to assure its specificity against Abf1p. Figure 31 summarizes data from western blot, chromatin immunoprecipitation and electro mobility shift assay. Immunoblot analysis showed that the antibody recognized the epitope that was used for immunization. The epitope was His-tagged and was also recognized by an α-His antibody. Further, α-Abf1p recognized a protein with a size of $\sim$130 kDa in crude extract

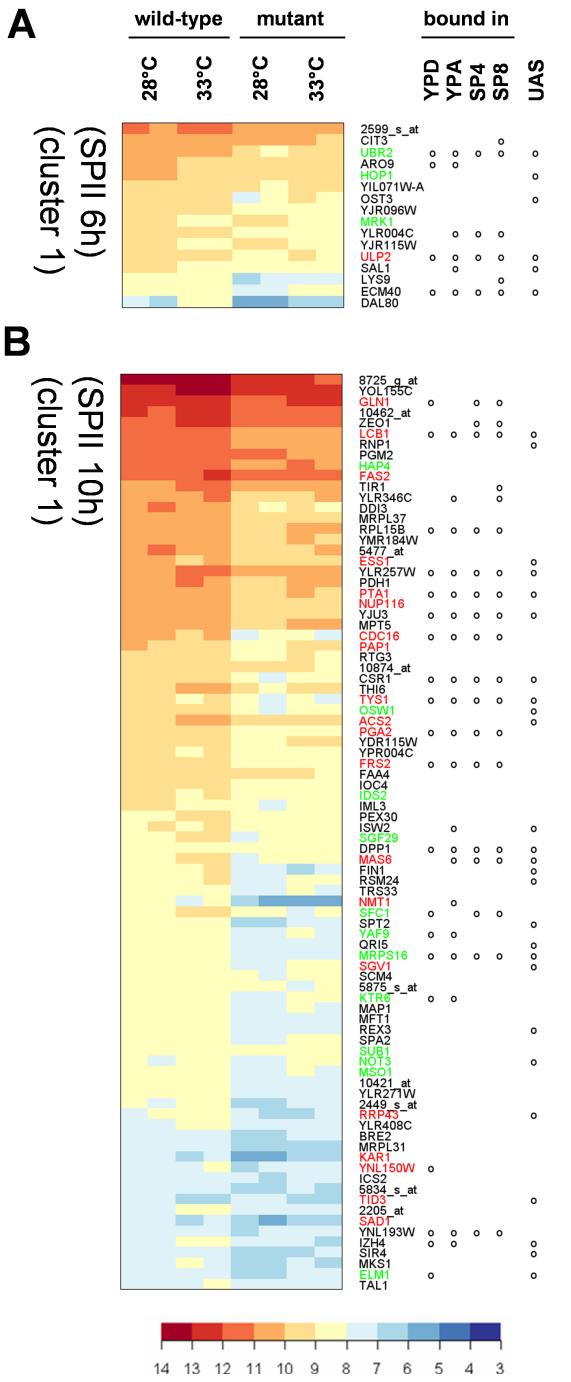

Figure 29: **(A and B)** Two heatmaps displaying expression clusters 1 and 2 obtained by filtration of expression levels of wild-type and mutant cells grown in YPA (but not in the corresponding YPD clusters) at 25°C and 37°C as indicated. The names of the corresponding genes or probeset identifiers are given (Gene). Circles indicate genes whose promotors contain predicted binding sites (UAS) and/or are bound by Abf1p under different conditions as indicated. Genes essential for mitosis or meiosis are given in redor green, respectively. Log2 transformed expression signal intensities are color coded as indicated.

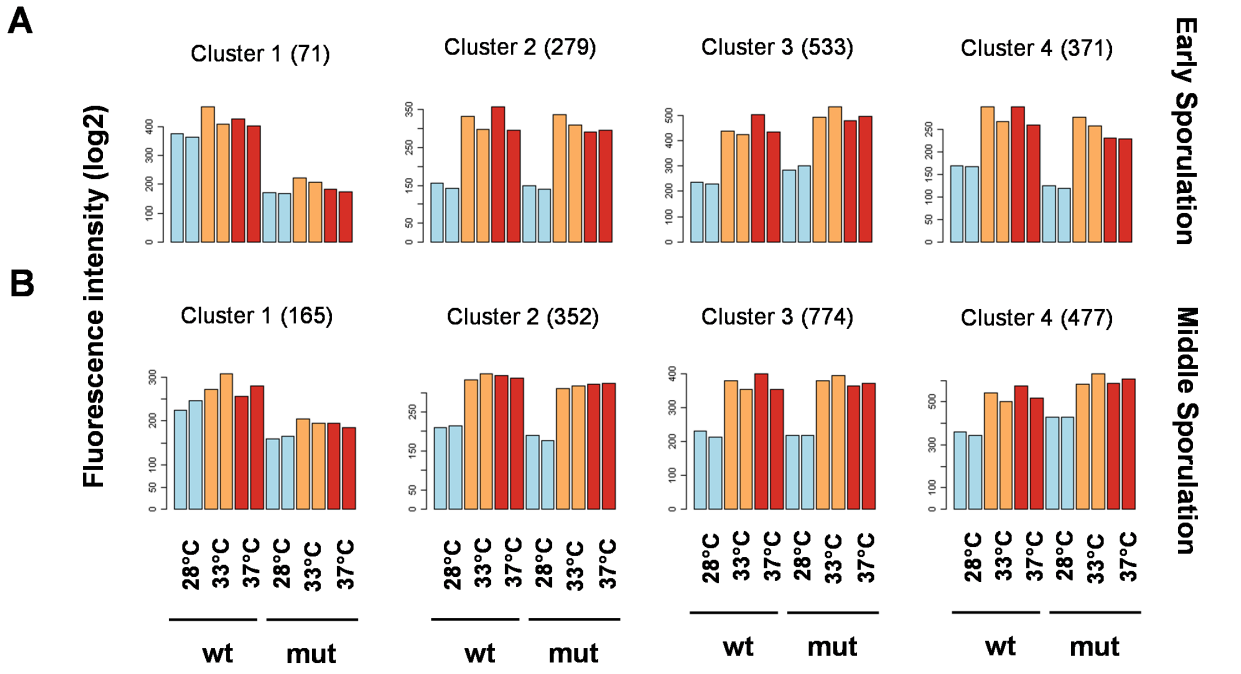

Figure 27: Partitioning Around Medoids (PAM) clustering of the 1254 and 1768 genes differentially expressed in during early **(A)** and middle **(B)** stages of sporulation, respectively. Wild-type (*ABF1/GAL-ABF1*, MPY125) and mutant (*abf1-1/GAL-ABF1*, MPY128) cells were sporulated at the 28°C for 5 and 9 hours, respectively, prior to temperature-shift to 33°C for 1 hour. Expression profiles of genes in the four clusters are represented by bars that display median expression levels of all genes in one condition. Numbers in brackets indicate the number of genes belonging to the cluster.

from wild-type and *ABF1-HA* strains. This size corresponds to that of Abf1p as previously published (Diffley and Stillman, 1989; Halfter *et al.*, 1989b; Halfter *et al.*, 1989a). The protein detected was more abundant in an *ABF1* over-expressing strain (Figure 31 panel B, lane 3). Interestingly, additional slower-migrating bands were also detected by our antibody. It has been reported that Abf1p can appear in at least four different states of phosphorylation (Silve *et al.*, 1992). The additional bands detected in western blot most probably correspond to phosphorylated forms of Abf1p (Figure 31, panel B). The antibody could be used for chromatin immunoprecipitation as demonstrated for one intergenic locus which is known to be bound by Abf1p (Figure 31, panel C). Additionally, a site located inbetween of *UME6* and *MSS4*, which contains a predicted *ABF1* but uncharacterized target site was also enriched by chromatin immunoprecipitation. In an *in vitro* DNA-binding assay the α-Abf1 antibody (but not an antibody against Ndt80p) was able to supershift a radiolabeled probe which is known to be bound by Abf1p. The results of an electrophoretic mobility shift assay (EMSA) using different probes that contain Abf1p binding sites are summarized in Figure 31 (panel D). Lane 1 shows the free labeled probe. We found a strong binding activity present in dividing diploid cells (lane 2) that is specific for the UAS_H motif (present in the promotor of *HOP1*). This activity was competed by the non-labeled UAS_H probe (lane 3) and by a non-labeled UAS_U probe (lane 5), which contained a predicted Abf1p-binding motif that is present in the promotor of *UME6*. However, the activity was not competed by the mutated non-labeled UAS_H and UAS_U probes (lanes 4 and 6). The shifted UAS_H probe was supershifted with the polyclonal α-Abf1 antibody (lanes 10-12) but not with a polyclonal antibody against Ndt80 (lanes 7-9). All these experiments gave compelling evidence that our antibody specifically recognizes Abf1p.

For our ChIP-CHIP experiments we used an SK1 strain because of its high efficiency and synchrony during sporulation. This strain completed premeiotic DNA replication after 4 hours in sporulation medium. After 12 hours

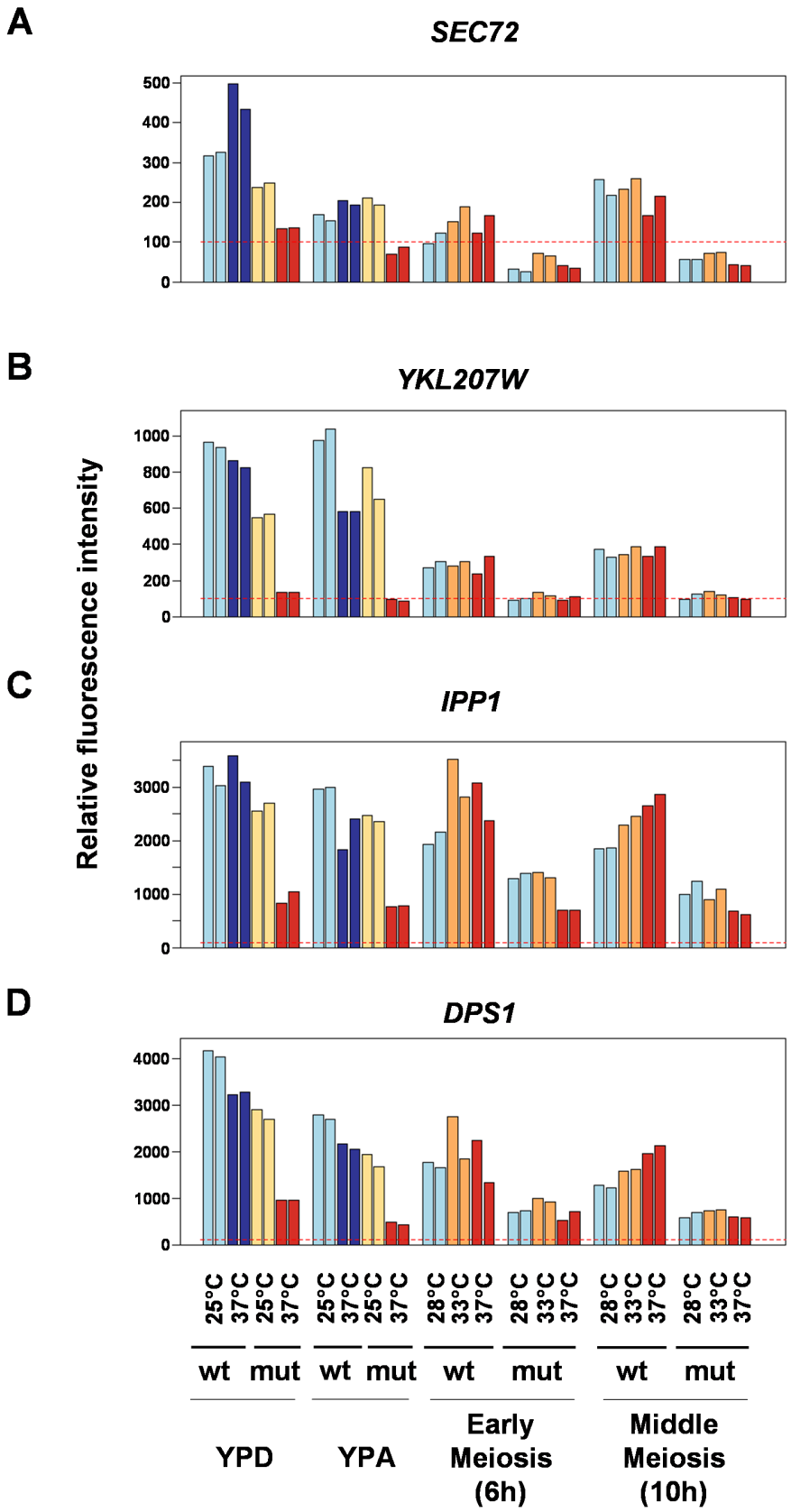

Figure 28: Examples of *ABF1* target genes and their expression during mitosis and meiosis. Barplots represent abundance of transcripts under different growth/incubation conditions as indicated. The red dotted line indicates the empirical threshold of reliable detection.

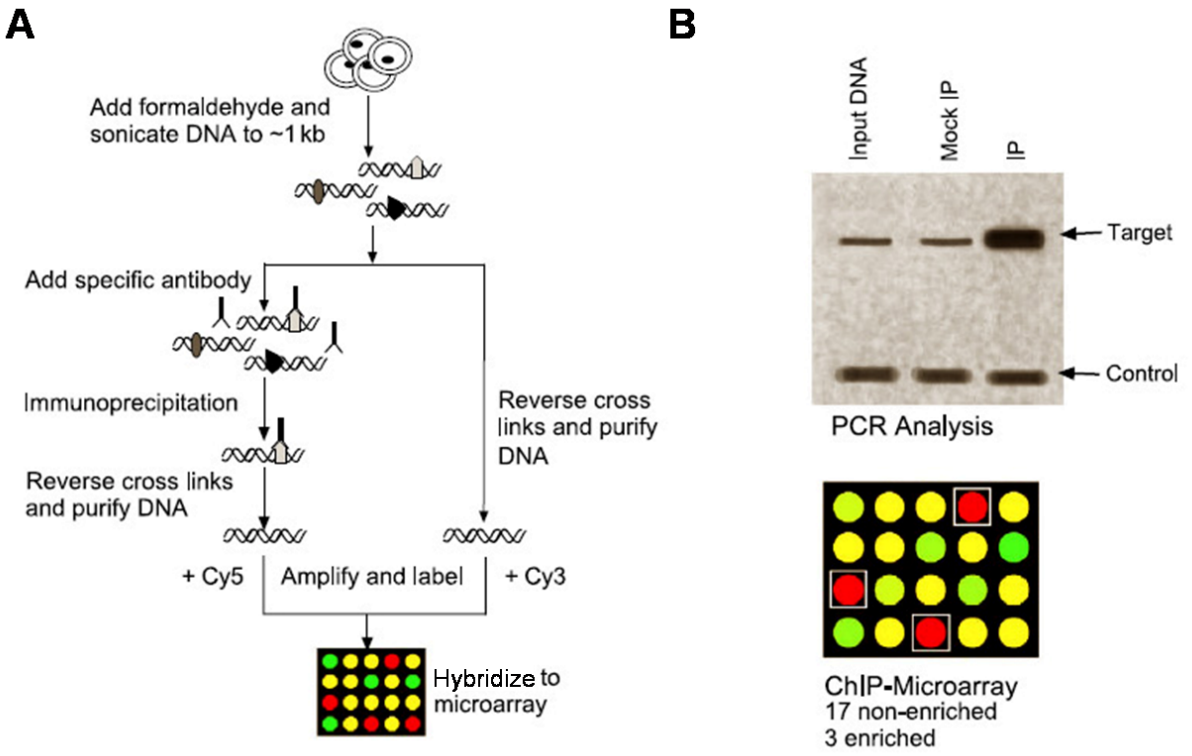

Figure 30: **(A)** A summary of the Chip-CHIP procedure. **(B)** Comparison of the controls used for single-locus, PCR-based ChIP experiments and microarray-based experiments. Single-locus experiments use a single internal control in each sample. The intensity of the target band is compared across the IP, mock IP (or control IP), and input DNA. In microarray experiments, ratios obtained for enriched elements (boxed in white) are compared to those obtained for all other elements, which are termed non-enriched. This figure was adapted from a review by Buck and Lieb (2004).

virtually all cells had undergone both meiotic divisions and had formed asci. For ChIP-CHIP assays SK1 cells were grown in the presence of glucose or acetate or sporulated for 4 and 8 hours (which corresponds to early and middle stages of sporulation) prior to treatment with formaldehyde and chromatin immunoprecipitation. We included three control experiments (dye swap, mock IP and input-over-input) in order to rule out the possibility of systemtic or technical errors. All experiments were performed in triplicate. See Figure 32 for an example of ChIP-CHIP raw data. Partial images of the scanned microarrays depict enrichment of the intergenic region between the divergently transcribed genes *MCK1* and *MRPS18*, which was previously shown to be bound by Abf1p. Signals from all experiments were found to be highly reproducible. As expected, dye swap experiments reversed the signal for this spot from red to green. Immunoprecipitation with a mock antibody did not enrich this region and the same was true for input-over-input samples. Preprocessing of ChIP-CHIP raw data included background correction and normalization based on the *Median Percentile Ranking* method (described by Buck and Lieb (2004)). This method was chosen because it avoids problems associated with ratio normalizations in ChIP-CHIP experiments. The rank of an assayed genomic element is simply the position of that element in a list sorted by ratio in descending order and therefore reflects how strong and reproducible this genomic element is bound by Abf1p. It is noteworthy, that after IP enrichment, DNA fragments bound by Abf1p will be of varying lengths, depending on the method of DNA fragmentation. In our assay average sequence size was between 300 bp and 1 kb. In many cases spots which are within 1 kb of the binding site were also enriched.

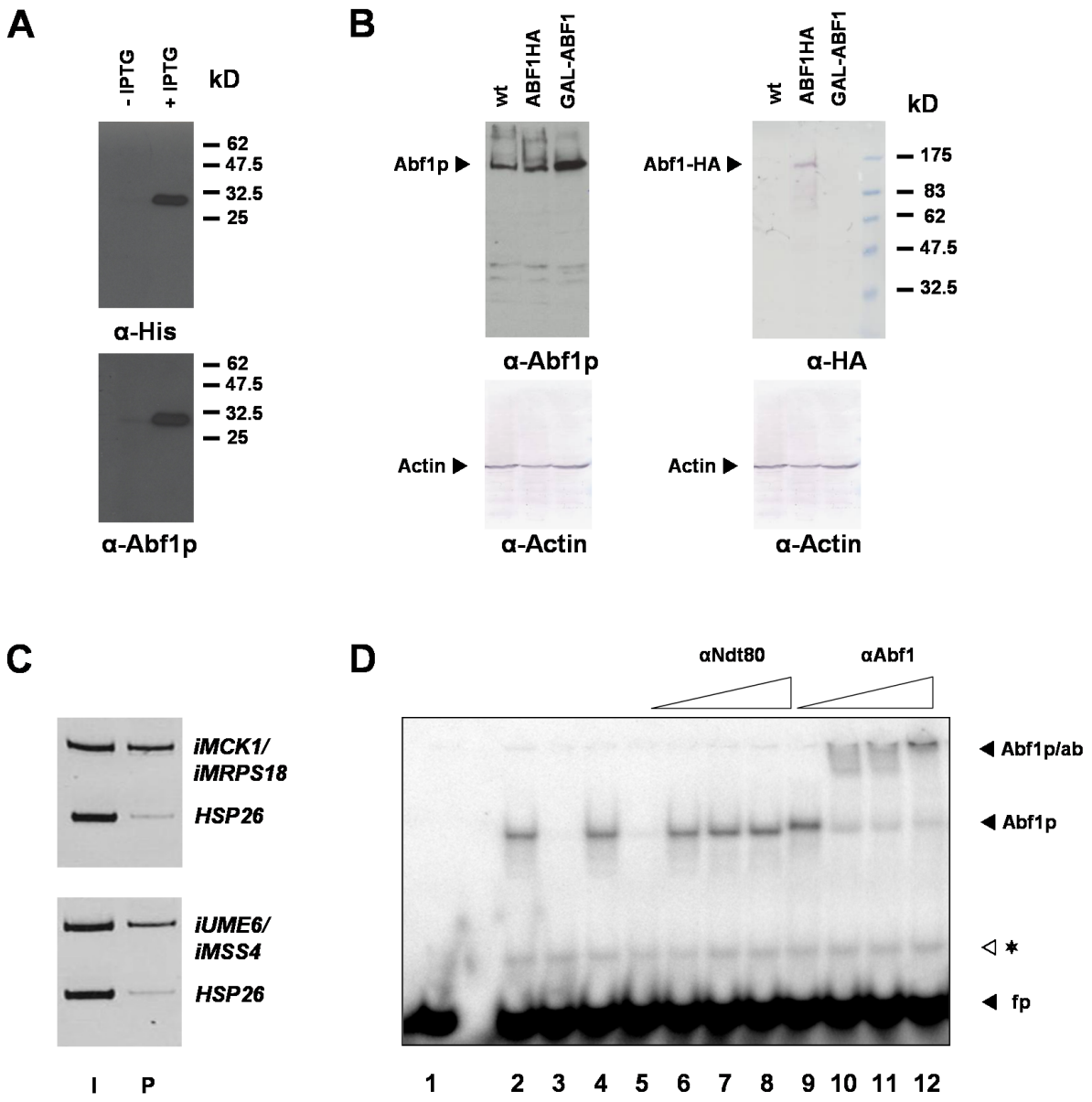

Figure 31: Confirmation of specificity of polyclonal α-Abf1p antibody. **(A)** Verification of antibody specificity by immunoblot. Protein extract of wild-type, ABF1HA-tagged, and *ABF1* over-expressing (P-GAL-ABF1) strains were analysed by immunoblot using polyclonal α-Abf1 and monoclonal α-HA antibody. Wild-type and HA-tagged Abf1p migrates at a molecular weight of about 130 kDa. **(B)** Electro-mobility shift assay with super-shift using polyclonal α-Abf1 antibody. Lane 1, free probe (radiolabeled UAS oligonucleotide of *HOP1* promotor); lane 2, binding activity of extract from vegetative cells; lane 3 and 4, competition of binding activity with unlabeled (cold) wild-type and point mutated UAS_H site (100-fold molar excess); lane 5 and 6, competition of binding activity with unlabeled (cold) wild-type and point mutated UAS_U site (100-fold molar excess); lane 7-9, supershift with α-Ndt80 antibody (2-fold, 5-fold and 10-fold dilution); lane 10-12, supershift with α-Abf1 antibody (2-fold, 5-fold and 10-fold dilution). **(C)** α-Abf1p directed chromatin IP in wild-type strain grown in rich medium, detected by multiplex PCR with primers specific for a part inbetween *MCK1* and *MRPS18* or specific for a part inbetween *UME6* and *MSS4* and part of the ORF of *HSP26*. I, input genomic DNA. P, precipitated DNA. kb, kilobases. kD, kilodalton.

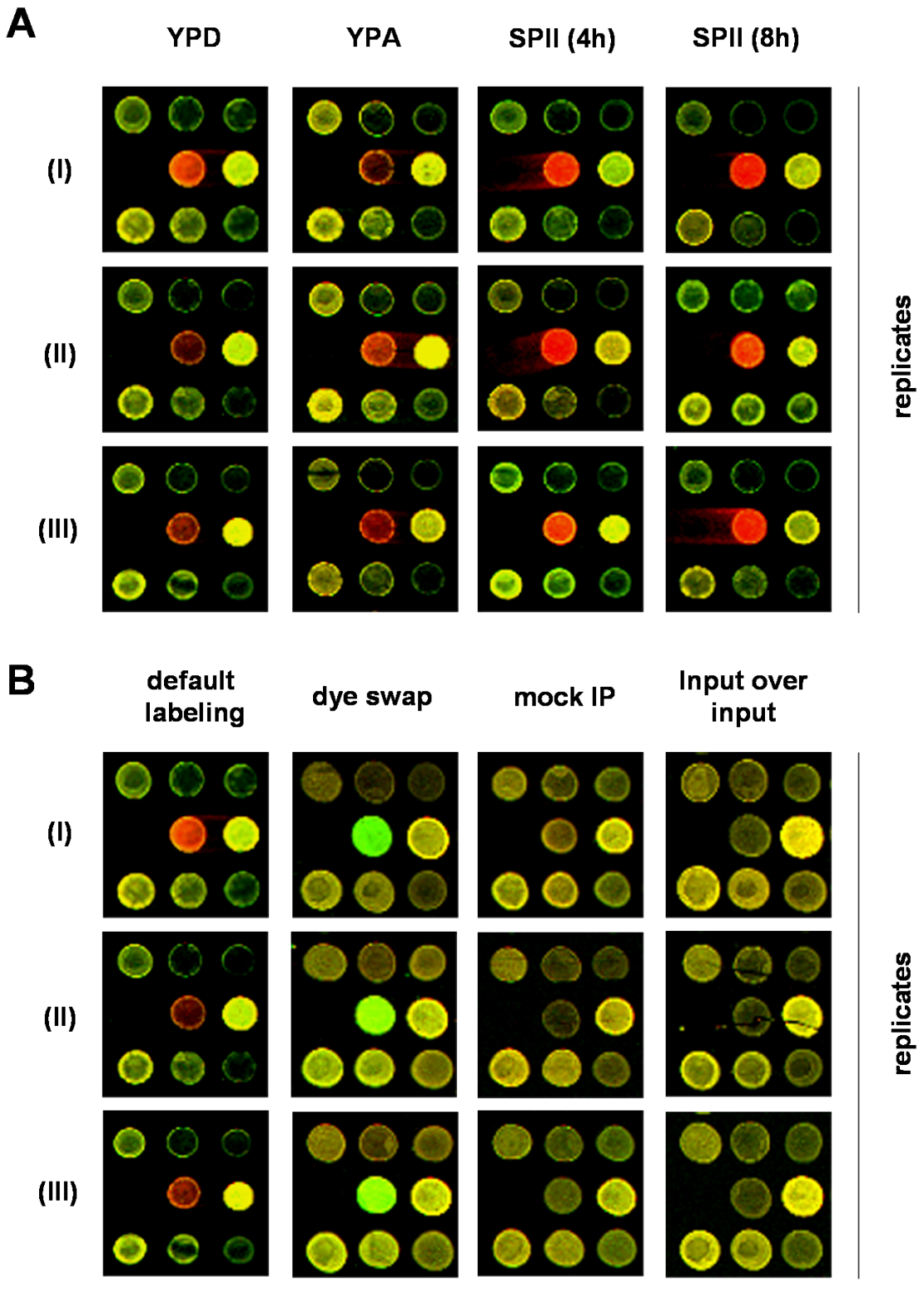

Figure 32: **(A)** Partial images of a series of microarrays hybridized with the labeled immunoprecipitated DNA (Cy5, red color) and control input (Cy3, green color), from SK1 cells (MPY170) grown in YPD or YPA and sporulated for 4h and 8h. The central red spot corresponds to the intergenic locus located inbetween *MCK1* and *MRPS18*. All three replicates are shown. **(B)** Partial images of microarrays hybridized with DNA from control experiments. Default labeling corresponds to YPD samples from panel A. Images from dye swap, mock IP and input-over-input experiments are shown for all three replicates.

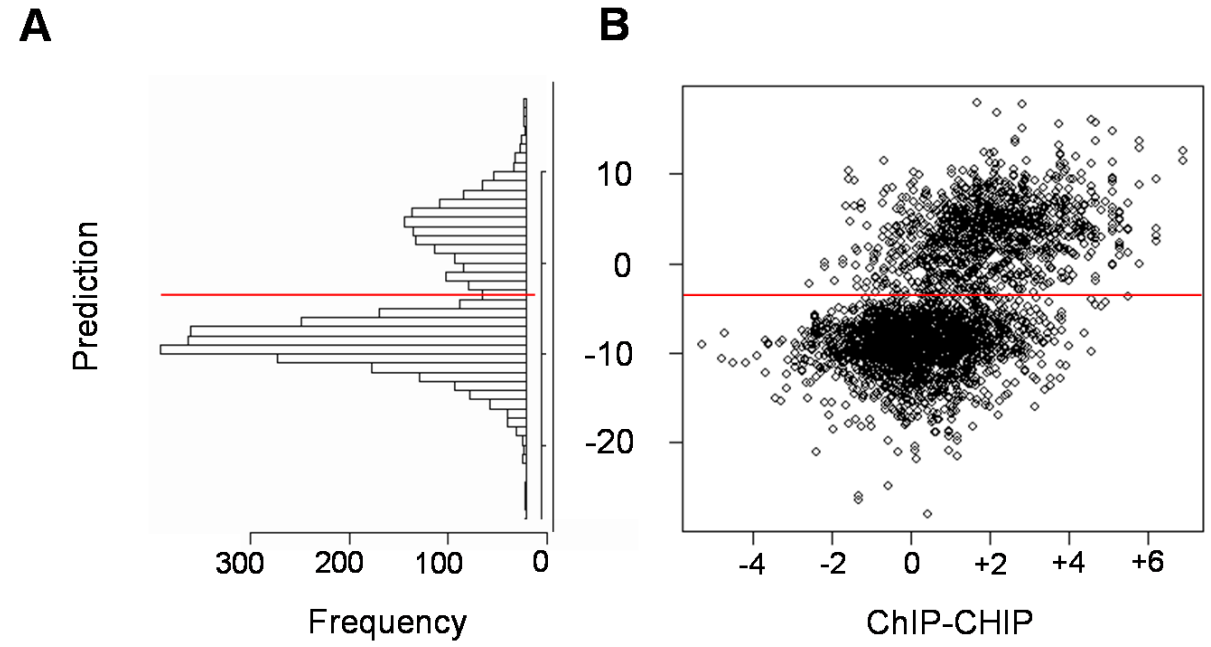

Figure 33: **(A)** Scatterplot showing correlation between the score of 3393 potential *ABF1* motifs and strength of binding in ChIP-CHIP assays (MPR levels were transformed ($z = ln(MPR/1 - MPR)$) before plotting). **(B)** Histogram displaying the frequency of predicted site scores that fall into each of 50 categories. The red line indicates the threshold of -2 that separates groups with predicted motifs with a high score from those with a low score.

In silico prediction of *ABF1*-target sites

Our genome-wide annotation of regulatory sites bound by Abf1p was produced in two steps. In the first step we built a weight matrix from experimentally determined binding sites known from the literature using the binding site clustering algorithm PROCSE (van Nimwegen *et al.*, 2002). In the second step we used a newly developed algorithm, called MotEvo (Erb and van Nimwegen, 2006), which identifies binding sites matching this weight matrix by scanning the multiple alignments of each *S. cerevisiae* intergenic region with the orthologous regions from the four other *Saccharomyces* species as detailed in the Materials and Methods section. MotEvo exhaustively reported putative locations of *ABF1*- binding sites and assigned a posterior probability to each reported site. A histogram of the probabilites calculated for every potential target site resulted in two distinct groups (Figure 33, panel A), suggesting that a given *ABF1*-motif is either well conserved across different yeast species and close to the weight matrix or its not evolutionary conserved and/or is distinct from the weight matrix. A score of -2 separated both groups from each other and was subsequently used as the threshold value for predicted sites. 1049 distinct genes have potential *ABF1*-binding sites in their promotors with a score higher than -2. We then asked whether there is a correlation of the score of a predicted site and the strength of binding of Abf1p to the region where this site is located. The scatterplot depicted in Figure 33 (panel B) shows that genomic regions which contain a potential binding site with a score bigger than -2 tend to be bound more strongly and reproducibly by Abf1p in the ChIP-CHIP assay. This was confirmed statistically with Fischer's Exact Test. The probability that a distribution such as the observed one occurs by chance is close to zero.

Combining the three data sets

Next, we hypothesized that overlapping the expression clusters determined earlier with data from genome-wide binding assays and binding site predictions for Abf1p would reveal direct Abf1p-targets in the different expression clusters. This is shown for example for the twelve clusters from the YPD dataset (Figure 34). In-

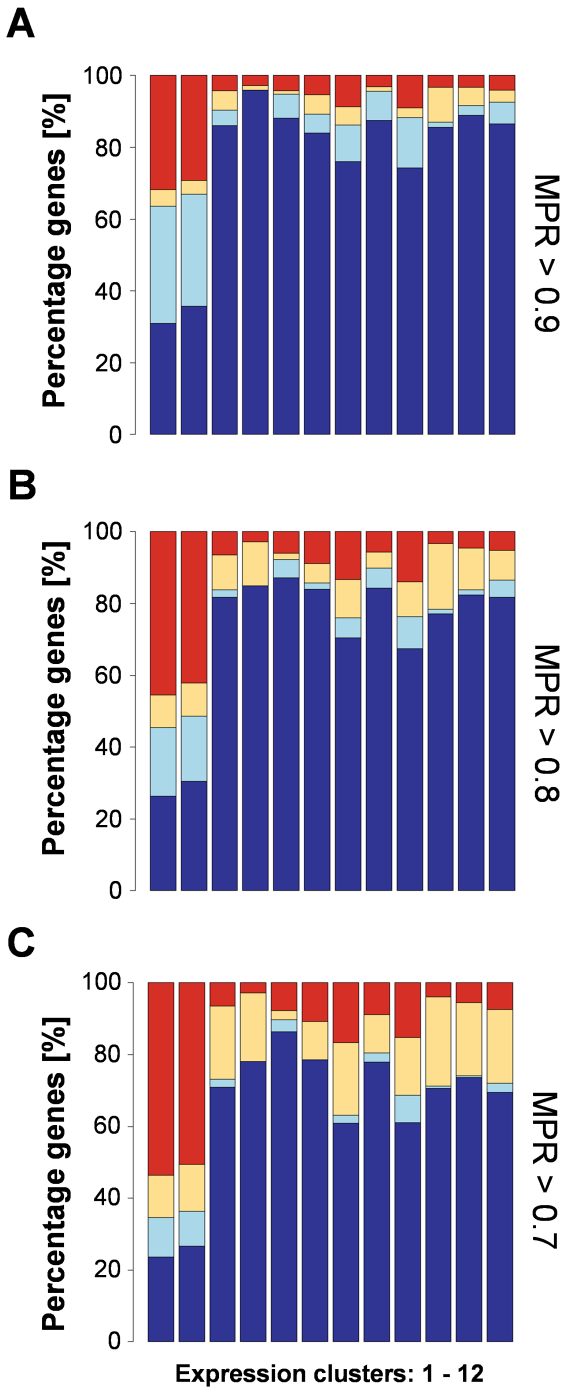

Figure 34: Analysis of twelve expression clusters (grouped into four categories) of the YPD data set. Colors represent percentage of genes whose promotors contain a predicted binding site and are bound by Abf1p (red), are bound by Abf1p but do not contain a predicted binding site (yellow), are not bound by Abf1p but harbor a predicted motif (light blue), or are not bound by Abf1p and do not contain a predicted binding site (dark blue). The three panels show the result of the overlap for three different cutoff levels for the ChIP-CHIP data (MPR 0.9 - 0.7).

deed, two groups (clusters 1 and 2) contained a substantial fraction of genes with either a predicted binding site and/or binding activity of Abf1p at their promotors. Strikingly, the expression pattern of the genes in these groups exactly displayed the anticipated profiles of being less abundant in the mutant cells at the restrictive temperature than in the other three conditions tested. In both clusters ~70% of the genes had predicted binding sites in their promotors (indicated by the red and the light blue area). Half of these sequences were also bound by Abf1p in the ChIP-CHIP experiment with an *Median Percentile Rank* higher than 0.9 (red area) (Figure 34, panel A). In all other expression clusters (3-12) there were also genes whose promotors were bound by Abf1p and contained a potential *ABF1* binding site. However, this was true only for a minor fraction of genes (<10%). By lowering the cutoff thresholds for the binding data (MPR >0.8 or MPR >0.7) the fraction of genes in clusters 1 and 2 that had a putative binding site and were bound by Abf1p gradually increased. However, this was also true for genes in the other clusters whose promotors were bound by Abf1p but did not contain a predicted binding site (indicated by the yellow area), and most importantly did not show an *ABF1*-dependent expression pattern. Thus, at lower MPR levels false positive target genes were enriched. We decided to choose an MPR of 0.8 which represented a tradeoff between identifying the largest amount of true direct target genes of Abf1p (indicated by the red area in clusters 1 and 2) and at the same time including only a small fraction of false positives. At an MPR of 0.8 Abf1p binds to 1131 intergenic regions in cells growing in YPD. The 1131 intergenic regions were located in the upstream regulatory sequences of 1428 distinct open reading frames. Intriguingly, this list included many (albeit not all) of previously identified Abf1p-regulated genes (e.g. *ACS1, BAP3, CAR1, COX6, DED1, FAS1, RPB2, RPC25, RPC40, RPS28A, SMK1, SPR3, SPT15, TRP3, TUB2, YPT1*) (Kratzer and Schüller, 1997; de Boer *et al.*, 2000; Kovari and Cooper, 1991; Trawick *et al.*, 1992; Lue *et al.*, 1989; Schüller *et al.*, 1994; Jansma *et al.*, 1996; Della Seta *et al.*, 1990b; Lascaris *et al.*, 2000; Pierce *et al.*, 1998; Ozsarac *et al.*, 1997; Schroeder and Weil, 1998b; Martens and

Brandl, 1994; Halfter *et al.*, 1989b), confirming that the ChIP-CHIP assay is capable of identifying Abf1p target sites. This notion was further substantiated by a comparison of genes whose promotors were bound by Abf1p in our assay with data from a recent genome-wide chromatin immunoprecipitation study using 204 yeast transcription factors (Harbison *et al.*, 2004). We found a good and significant correlation of our gene list with genomic loci bound by myc-tagged Abf1p in vegetatively growing yeast cells. Harbison and coworkers detected 468 genes bound by Abf1p (p-value <0.005) of which we identified 83%. A comparison of the genes whose promotors were bound by Abf1p and those with a predicted UAS (cutoff >-2) resulted in ∼700 overlapping (core) genes (Figure 35, panel A). Notably, there are genes whose promotor contains a potential binding site but is not bound or *vice cersa*. However, most of the 700 core genes were also identified in the dye swap, but not in the input-over-input or the mock IP control experiments (Figure 35, panel B-D).

A similar overlap of binding and prediction data with the expression profiles from YPA, early and middle sporulation samples is shown in Figure 36. Altogether, we identified six expression clusters (YPD, cluster 1 and 2; YPA, cluster 1 and 2; early sporulation cluster 1; middle sporulation, cluster 1) with an disproportionately high fraction of genes having predicted and bound sites in their promotors and whose expression profiles suggested dependency on Abf1p. These six expression clusters contained 563 distinct genes. Notably, 59 of these genes are non-annotated SAGE ORFs (Velculescu *et al.*, 1997) that are present on the Affymetrix S98 GeneChips used in the current study. We therefore find 504 *ABF1*-dependent genes, which is ten times more than previously identified in a similar expression profiling study (Miyake *et al.*, 2004). This underlines the power of approaches which combine data from different high-throughput techniques. The promotors of more than one third of the 504 genes had a predicted binding motif and were also bound under at least one of the growth or sporulation conditions tested (203 genes). Another 149 (71 plus 78 genes) genes either had a potential binding site in the promotor or were bound by Abf1p in their upstream regulating

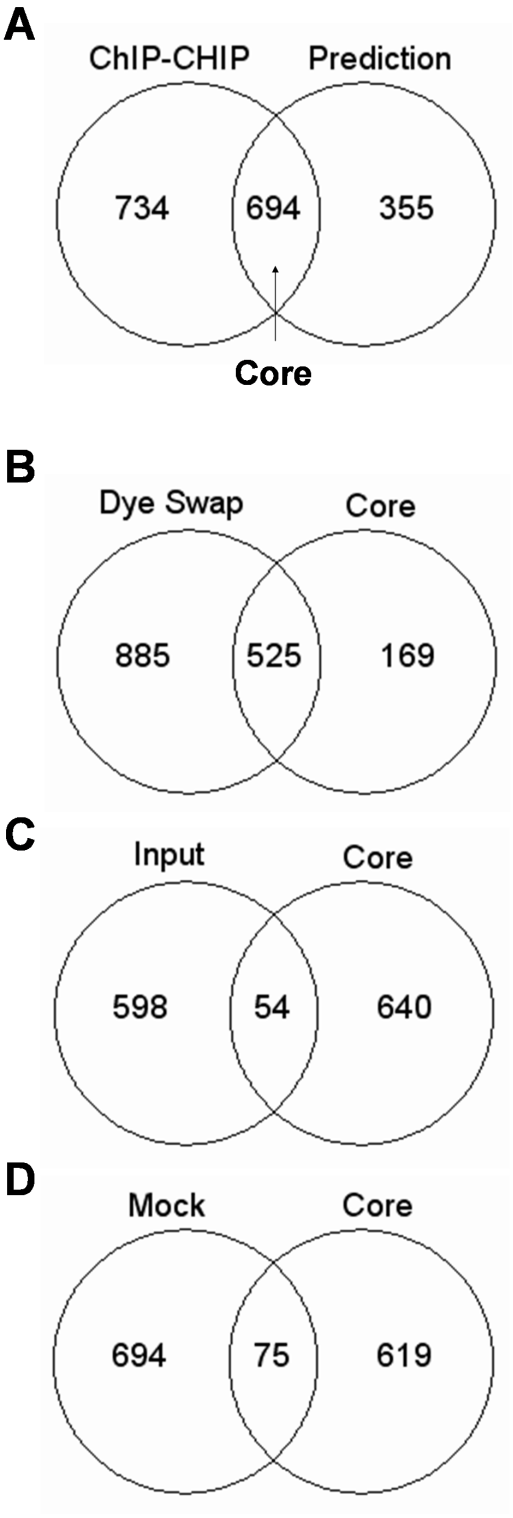

Figure 35: **(A)** Venn diagram showing the overlap between genes whose promotors are bound in the ChIP-CHIP assay (MPR >0.8) and those which harbor a predicted binding site (p-value >-2). The overlap is designated as *Core*. **(B-D)** Venn diagrams showing the overlap of the Core and genes whose promotors are bound in the three ChIP-CHIP control experiments, dye swap, input-over-input, and mock IP, respectively.

sequences. These two groups were merged (352 genes altogether) and refered to as *bona fide* direct target genes of *ABF1*, because they fulfilled at least two of the three criteria that suggests dependency on Abf1p. The remaining 152 genes were considered possible indirect target genes of *ABF1*. Figure 37 summarizes this information in a Venn diagram. Six different heatmaps represent *ABF1*-target genes from the six clusters along with additional information about presence of an UAS in the promotor and binding of Abf1p to the promotor in our ChIP-CHIP assays (Figures 19, 22, and 29).

Data mining of *ABF1*-target genes

To get a general overview of the gene functions dominating the groups of direct and indirect *ABF1* target genes we have searched for enriched GeneOntology terms. Direct target genes were enriched in *protein transport activity* (*KAP123, MAS6, PAM16, PSE1, SEC72, SRP1, SSH1, TOM40, TOM7, TOM70*) with many of the genes involved in *protein import into the mitochondrial matrix*. Furthermore, numerous genes had a function in *nitrogen compound metabolism* and *amino acid biosynthesis*, such as *ALA1, ARO3, ARO4, ARO9, ASN1, GLN1, GLT1, LEU2, LEU9, LYS9, LYS20, NPR1, PRO3, THR4, TRP2, TRP3*. Interestingly, the products of many *ABF1* target genes localized to the *mitochondrial lumen* (*ADH3, CIT3, ECM40, HMG1, HMG2, ILV5, IMG2, ISU2, MAS1, MGE1, MRP1, MRP49, MRP51, MRPL28, MRPL32, MRPL36, MRPL39, MRPL51, MRPS16, MSS116, MSS51, NFU1, RSM24, RSM27, SUN4, SWS2, YFH1*).

Among indirect target genes we found GO-terms related to *metal ion transport* (*ENB1, FET3, MRS4, PHO84, SIT1, ZRT1, ZRT3*), *amino acid biosynthsis* (*ARG3, ARO1, ATF1, BAT1, FAS2, GDH1, HIS7, ILV1, ILV2, ILV3, ILV6, MIS1, TRP4*). Numerous gene products of indirect target genes localized to the *cell wall* (*CTS1, CWP1, DSE2, EXG1, FIT3, PLB2, SCW11, SIM1, SRL1*).

Genes essential for vegetative growth

110 of the 504 *ABF1*-target genes are essential for vegetative growth and involved in disparate biological processes, such as intracellu-

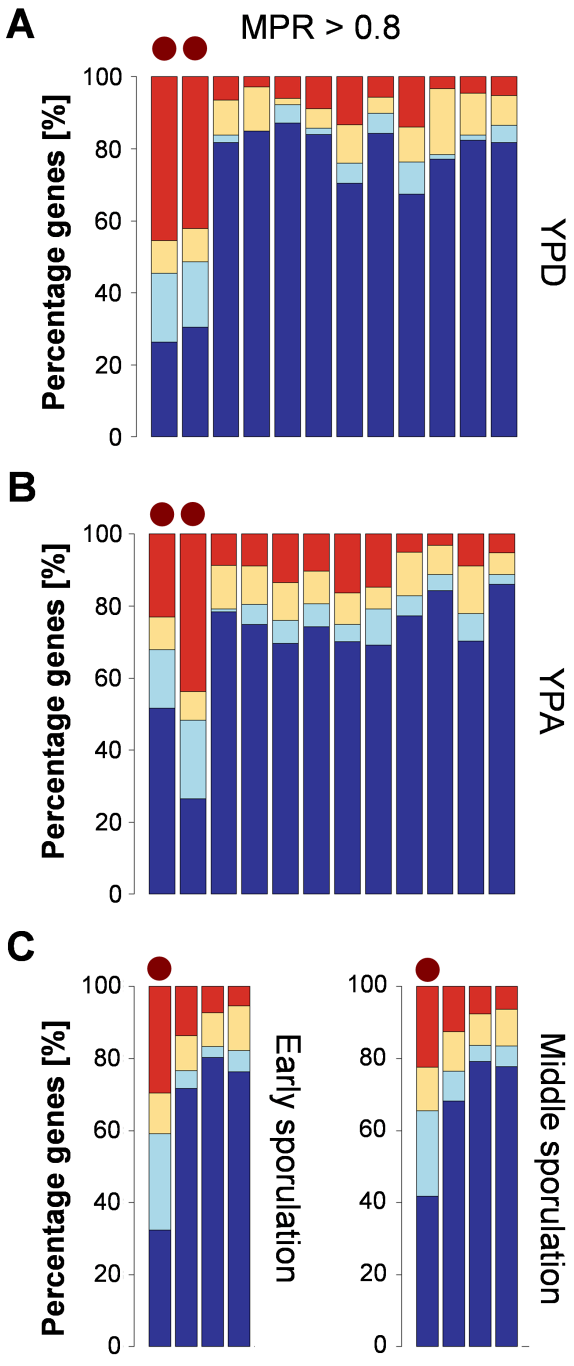

Figure 36: Analysis of twelve expression clusters of the YPD, YPA, and the four expression clusters of the early sporulation and middle sporulation data sets. Colors represent percentage of genes whose promotors contain a predicted binding site and are bound by Abf1p (red), are bound by Abf1p but do not contain a predicted binding site (yellow), are not bound by Abf1p but harbor a predicted motif (light blue), or are not bound by Abf1p and do not contain a predicted binding site (dark blue). The three panels show the result of the overlap for the MPR=0.8 cutoff level for the ChIP-CHIP data. Expression clusters with a overrepresentation of genes whose promotors are bound or have a predicted binding site are marked with a dot.

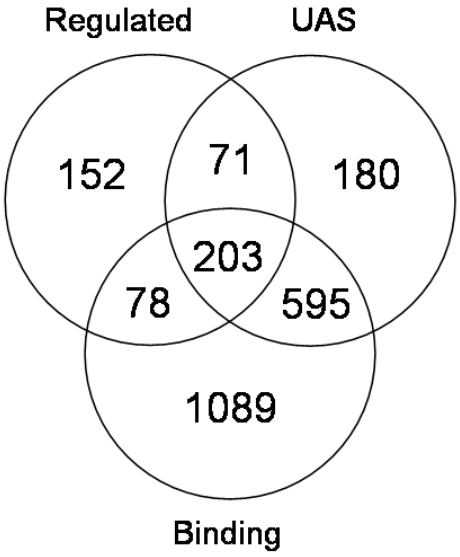

Figure 37: **(A)** Venn diagram summarizing the overlap of genes whose expression profiles suggest dependency on *ABF1* (Regulated), whose promotors contain a predicted binding site (UAS), and/or whose promotors are bound in at least one of the tested conditions (Bound). **(B)** Significance of the observed overlaps between the three different datasets. *p*-values were calculated on the basis of the hypergeometric test.

lar protein transport (*SEC14, SEC16, SEC18, SEC53*), ribosome biogenesis (*DHR2*), cell-wall biosynthesis (*BIG1*), protein modification (*GPI17*), fatty acid metabolism (*TSC13*), phosphate metabolism (*IPP1*), amino acid biosynthesis (*PRO3*), and mitochondrial protein import (*MIM1*). Interestingly, *TAO3*, which has recently been shown to be one of the three quantitative trait loci that control yeast sporulation (Deutschbauer and Davis, 2005) also belongs to this group. Many genes had a function in *rRNA metabolism* (*DBP9, EMG1, MDN1, MRD1, MTR4, NOP53, NUG1, POP8, PWP2, RLP7, RRP1, RRP5, RRP9, RRP12, RRP43, URB1, UTP18, UTP20*) and *tRNA metabolism* (*ALA1, FRS2, GRS1, HTS1, MTR4, NUP49, NUP116, POP8, PTA1, TRM1, TRM112, TYS1*). This strongly suggested that Abf1p plays an important role in translation, which is in keeping with earlier observations (Planta *et al.*, 1995). Figure 38 shows four examples of genes which are essential for vegetative growth and which are regulated by *ABF1*.

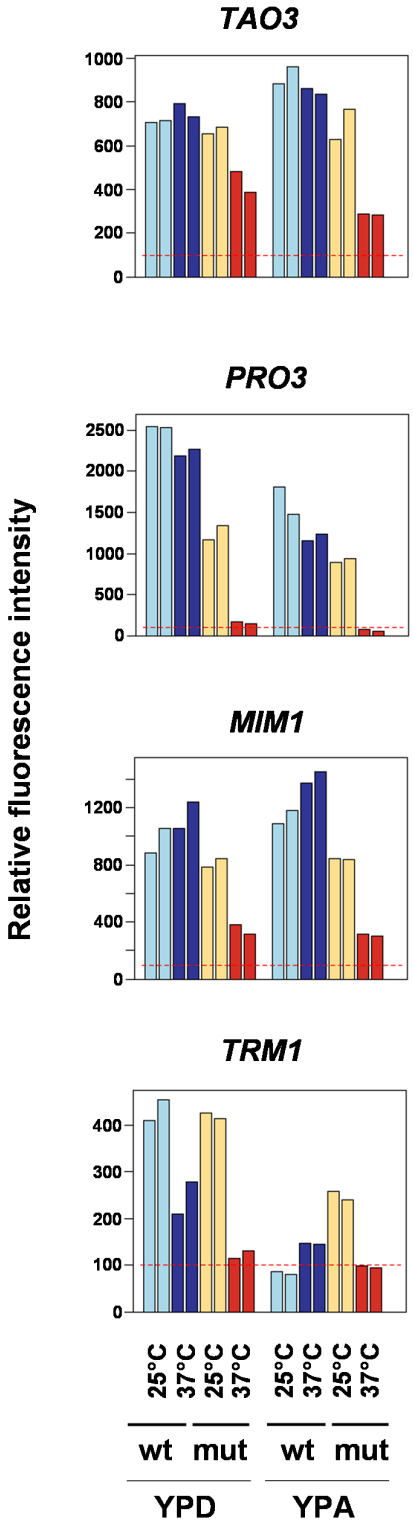

Figure 38: Examples of *ABF1* target genes that are essential for vegetative growth in budding yeast. Barplots represent abundance of transcripts under different growth/incubation conditions as indicated. The red dotted line indicates the empirical threshold of reliable detection.

Genes induced and repressed in acetate

Among the Abf1p-regulated genes identified in cells growing in the presence of a fermentable carbon source (glucose in YPD) and a non-fermentable carbon source (acetate in YPA) we found genes whose activity was activated or repressed upon shifting from YPD to YPA. Four examples for both categories are depicted in Figures 39 and 40. 89 and 90 Abf1p-target genes are up- or down-regulated two-fold in YPA, respectively. A substantial fraction of acetate-induced genes encode proteins that locate to the mitochondrion. Some of these genes are involved in ATP biosynthesis (*ATP14, APT16, ATP18*). Others are cellular components of the mitochondrial ribosomes (*RSM27, MRPL28, MRPL33, MRPL36*). This is in keeping with an earlier observation, that during diauxic shift genes encoding mitochondrial ribosomal genes were generally induced rather than repressed after glucose limitation, high-lighting the requirement for mitochondrial biogenesis (DeRisi *et al.*, 1997). Among the acetate-repressed genes we identified genes whose products are devoted to protein synthesis, including nucleolar proteins involved in rRNA synthesis (*RRP1, RRP5, RRP9, RRP12, DHR2, UTP18, UTP20*), a subunit of RNA polymerase III (*RPO31*), and proteins needed for amino acid metabolism (*ARG3, LEU9, TRP3*). Previously, it has been shown that 95% of ribosomal genes are down-regulated during the diauxic shift (DeRisi *et al.*, 1997).

Genes involved in meiotic development

Additionally, we found 59 genes which are essential for i) sporulation (*CHS5, GID7, HAP4, HMG2, MLS1, MRP51, MRPL33, MRPS16, MSD1, MSO1, MSS51, OSW1, PEP12, PKH2, SGF29, SNF8, SPO14, SWM1, THR4, TOS1*), ii) spore germination (*ARG3, HOP1, HSL7, IDS2, MRPL32, NFU1, PAT1, PEP8, RED1, SEL1, SFC1, TOM7, VTC1, YAK1*), or iii) inhibition of meiosis (*APS2, BAP2, ELM1, GDS1, KTR6, NOT3, SET3, SHE4, SUB1, UBR2, YAF9*). A number of these genes were especially interesting, because they could explain mitotic and meiotic phenotypes of *abf1-1* mutant cells. These will be briefly described in the following section:

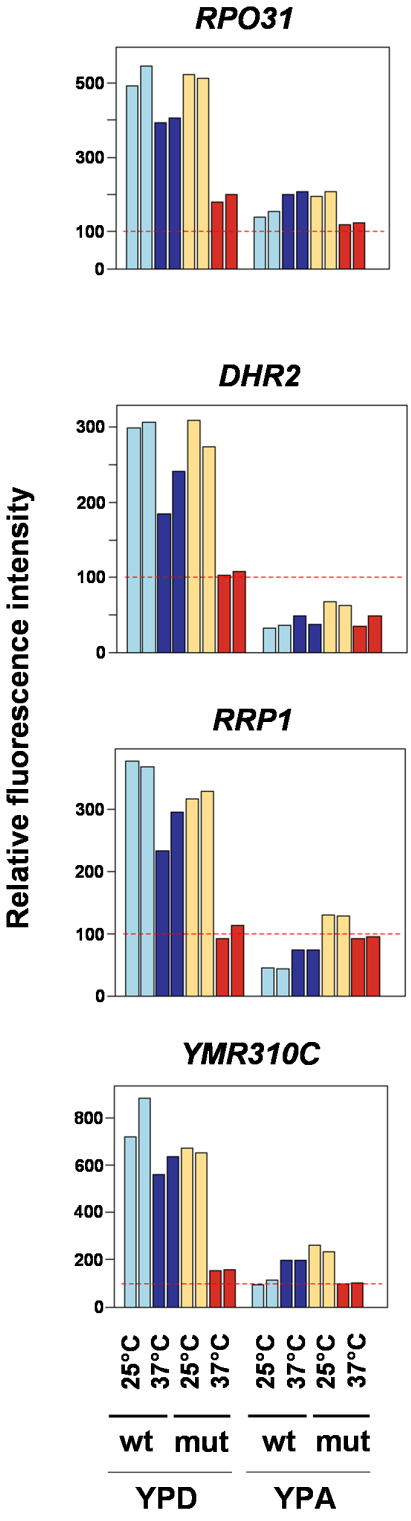

Figure 39: Examples of genes whose expression is repressed upon transfer of cells into acetate containing medium. Barplots represent abundance of transcripts under different growth/incubation conditions as indicated. The red dotted line indicates the empirical threshold of reliable detection.

SPO14 encodes a phospholipase D (PLD) essential for Golgi function (Rudge *et al.*, 2001). It is also required for commitment to meiosis and spore formation (Honigberg *et al.*, 1992; Rabitsch *et al.*, 2001). In spo14Δ mutants, the second meiotic division and spore formation are defective, while earlier meiotic events are largely unaffected. This results in the accumulation of binucleate cells during meiosis (Honigberg and Esposito, 1994), a phenomenon that is also observed in sporulating *abf1-1* mutants (Figure 13). Our expression profiling experiments showed that transcript levels of *SPO14* are affected in *abf1-1* mutant cells already at the permissive temperature (Figure 41, panel A) in growing and sporulating cells. Another *ABF1* target gene involved in spore formation is *OSW1* (Coluccio *et al.*, 2004) which is meiosis-specifically expressed (Figure 41, panel B).

SWM1 and *CDC16* are two subunits of the anaphase-promoting complex/cyclosome (APC/C), which are E3 ubiquitin ligases that regulate the metaphase-anaphase transition and exit from mitosis (Zachariae *et al.*, 1996; Zachariae and Nasmyth, 1999; Schwickart *et al.*, 2004). The promotors of both genes were bound by Abf1p but contained no consensus binding site. The expression profile showed dependency of their mRNA abundance on Abf1p (Figure 41, panel C and D) during mitotic growth and meiosis. Both genes were induced at middle stages of meiotic development in the wild-type but not in *abf1-1* mutant cells. *CDC16* is an essential member of the APC/C and several temperature-sensitive mutants of *CDC16* have been demonstrated to arrest as large-budded cells with the nucleus at the neck (Zachariae and Nasmyth, 1996). Cdc16p interacts with two other essential APC subunits, Cdc23p and Cdc27p (Lamb *et al.*, 1994). *SWM1* in turn is a non-essential gene that promotes the stable association of the APC/C-subunits Cdc16p and Cdc27p (Schwickart *et al.*, 2004). Swm1p is required for spore wall assembly during sporulation and maintainance of cell wall integrity during vegetative growth at high temperatures (Ufano *et al.*, 1999; Ufano *et al.*, 2004). It has been demonstrated that a substantial fraction (30%) of swm1Δ cells incubated at 37°C failed to produce colonies and arrested

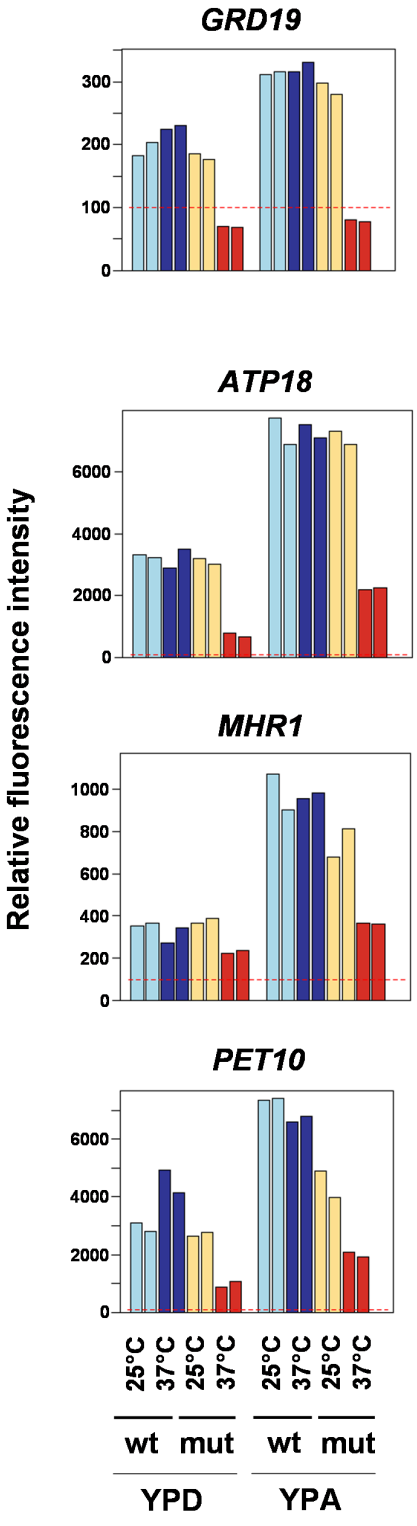

Figure 40: Examples of genes whose expression is induced upon transfer of cells into acetate containing medium. Barplots represent abundance of transcripts under different growth/incubation conditions as indicated. The red dotted line indicates the empirical threshold of reliable detection.

as large-budded cells. They also accumulated (40%) cell bodies with 4 non-separated cells (Schwickart *et al.*, 2004). These phenotypes can indirectly explain the phenotype observed in *abf1-1* mutants (see Figure 11, panel B).

We also found *SRV2, SSZ1, VMA6* and *VPS54* among direct target genes of *ABF1* (Figure 43). These genes have been isolated in a previous screen for single-deletion strains that display a visually discernable sporulation phenotype (reduced or no spore formation or formation of asci that contain less than four spores) and that were not able to activate an *IME1*-lacZ reporter gene and to undergo the two meiotic divisions (Enyenihi and Saunders, 2003). Yeast strains which were deleted in one of the four genes showed reduced levels of spore formation as tested after 3 days on acetate. Again, this could explain the reduced ability of *abf1-1* mutant cells to undergo both meiotic divisions and form asci with high efficiency (Figure 13).

Interestingly, *SET3* was identified as a direct target gene of *ABF1*. Set3p has previously been shown to form a complex with Snt1p, YIL112w, Sif2p, Cpr1p, and two putative histone deacetylases, Hos2p and Hst1p. This complex is likely to be the meiosis-specific repressor of early/middle sporulation genes (Pijnappel *et al.*, 2001). This suggests that Abf1p has not only chromatin remodeling capabilities itself but also transcriptionally regulates a gene with such a function.

Notably, we found *HOP1*, a component of the synaptonemal complex (SC) among *ABF1* regulated genes (Figure 42, panel A). This is in keeping with a previous finding that mutating the Abf1p-bound UAS in the *HOP1* promotor completely abolishes its expression (Prinz *et al.*, 1995). Another component of the SC, *RED1*, harbors a potential *ABF1*-binding site in its promotor which is also bound by Abf1p. Furthermore, its expression depends on functional *ABF1* (Figure 42, panel B). *ZIP1*, a third SC component, was not among *ABF1*-regulated genes. However, its promotor contains a well-conserved UAS and is also bound by Abf1p. Taken together this suggests that *ABF1* may play a general role in transcriptional regulation of SC components.

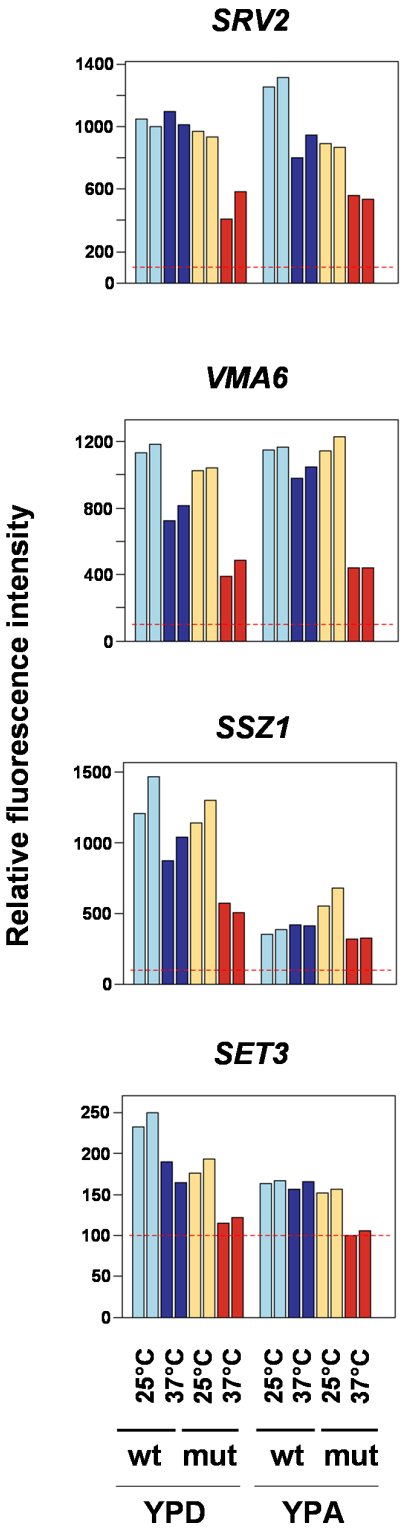

Figure 43: Examples of *ABF1* target genes that are involved in meiotic development in budding yeast. Barplots represent abundance of transcripts under different growth/incubation conditions as indicated. The red dotted line indicates the empirical threshold of reliable detection.

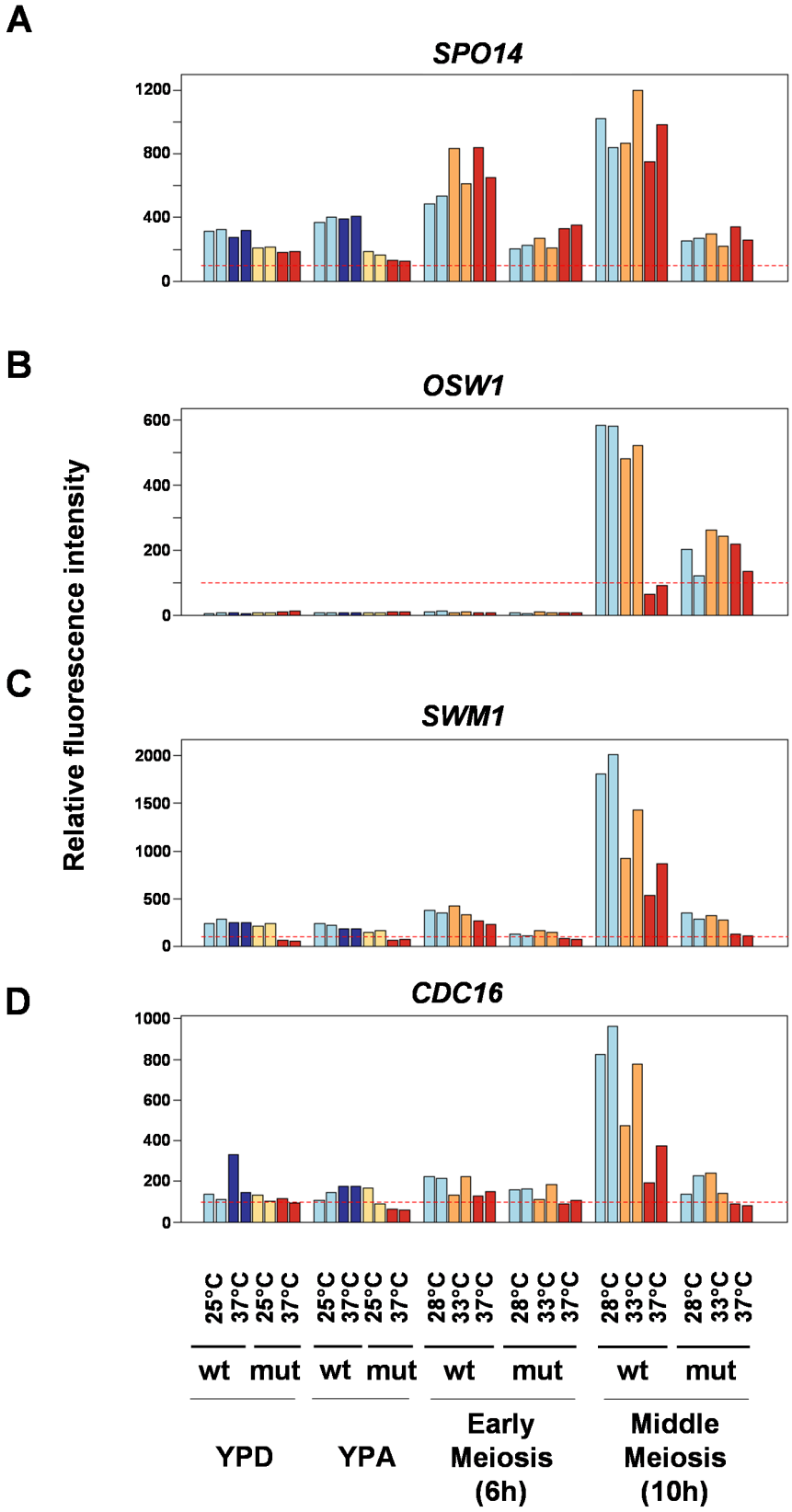

Figure 41: Examples of *ABF1* target genes that are involved in meiotic development in budding yeast. Barplots represent abundance of transcripts under different growth/incubation conditions as indicated. The red dotted line indicates the empirical threshold of reliable detection.

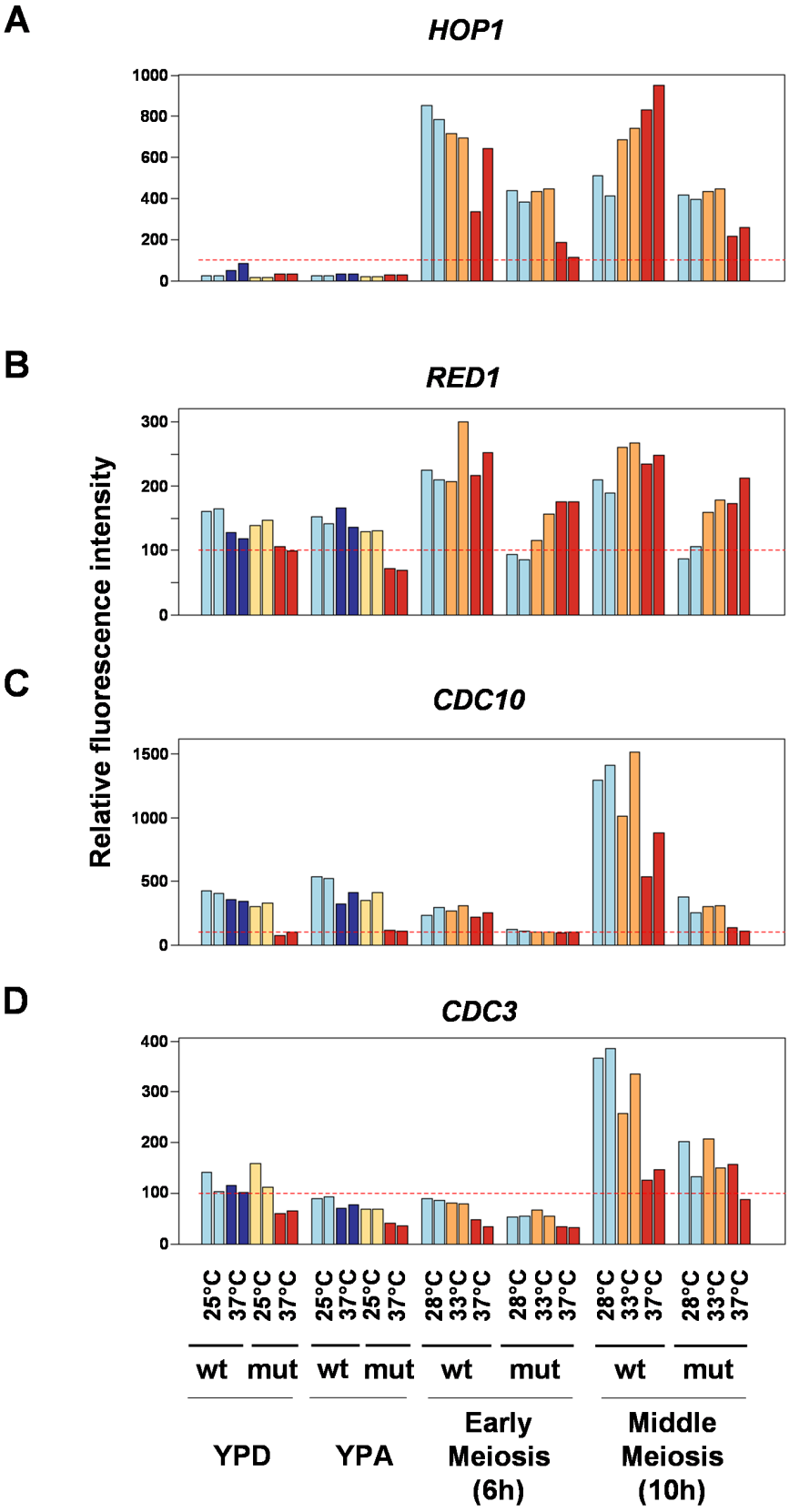

Figure 42: Examples of *ABF1* target genes that are involved in meiotic development in budding yeast. Barplots represent abundance of transcripts under different growth/incubation conditions as indicated. The red dotted line indicates the empirical threshold of reliable detection.

Abf1p and ARS functionality

Our genome-wide binding assays allowed us to examine Abf1p-binding activity at all autonomously replicating sequences known to date. Strikingly, we detected *ARS1, ARS120* and *ARS121* strongly bound by Abf1p (Figure 44), which is in keeping with earlier observations (Diffley and Stillman, 1988; Francesconi and Eisenberg, 1989). We also detected a number of additional ARS elements to be strongly bound by Abf1p (e.g. *ARS1413, ARS304, ARS603.5, ARS1002, ARS1012, ARS1016, ARS1025, ARS523, ARS1411, ARS503, ARS518, HMLI/ARS302, ARS303, ARS319).* The silent mating type loci *HMRE/ARS317, HMRI/ARS318* and *HMLE/ARS301* belonged to a group of weakly bound ARS. This is in keeping with previous observations, that these ARSs have a low binding affinity for Abf1p (Eisenberg *et al.*, 1988). Among the remaining 70 ARS sequences we detected no or only very weak binding of Abf1p.

Abf1p and cytokinesis

Our genome-wide binding assays suggested a connection between transcriptional regulation of factors essential for cytokinesis and Abf1p. The promotors of two septins (*CDC3* and *CDC10*) harbor predicted binding sites and are bound by Abf1p. Furthermore, transcriptional activity of both genes depends on *ABF1* (Figure 42, panel C and D). *CDC3* and *CDC10* are genes which belong to the Cdc3/10/11/12 septin family of bud neck filaments. Both genes are expressed and are necessary for optimal growth under vegetative conditions. Interestingly, promotors of both genes contain an MSE, which confers induction of expression during middle stages of meiotic development. Previously, the meiosis-specific septin *SPR3* has also been found to be under control of Abf1p (Ozsarac *et al.*, 1995; Ozsarac *et al.*, 1997). It is conceivable that Abf1p plays a more general role in the regulation of members of this protein family.

Temperature-sensitive mutations in *CDC3* and *CDC10* incubated at the restrictive temperature resulted in a defect in cytokinesis and developed multiple elongated buds that do not separate from the parent cell (Hartwell, 1971). Additionally, these mutants failed to

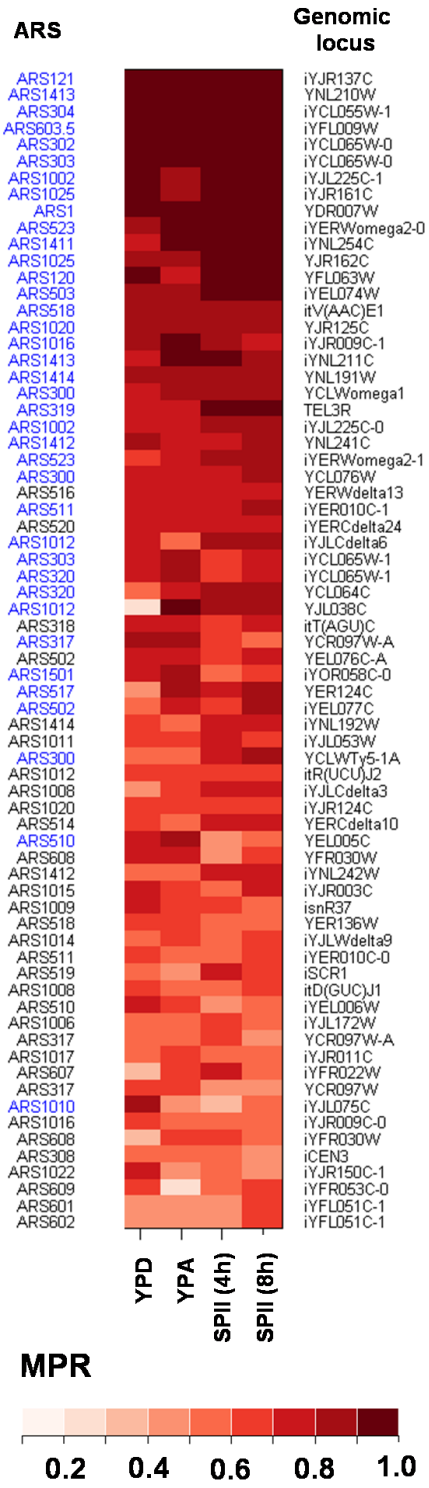

Figure 44: Abf1p binding to autonomously replicating sequences (ARSs). Strength of Abf1p-binding to 75 ARS sequences is represented by a color code. Genomic loci and corresponding ARS are indicated. ARSs whose names are colored in blue are bound by Abf1p with an Median Percentile Rank higher >0.8 in at least one condition.

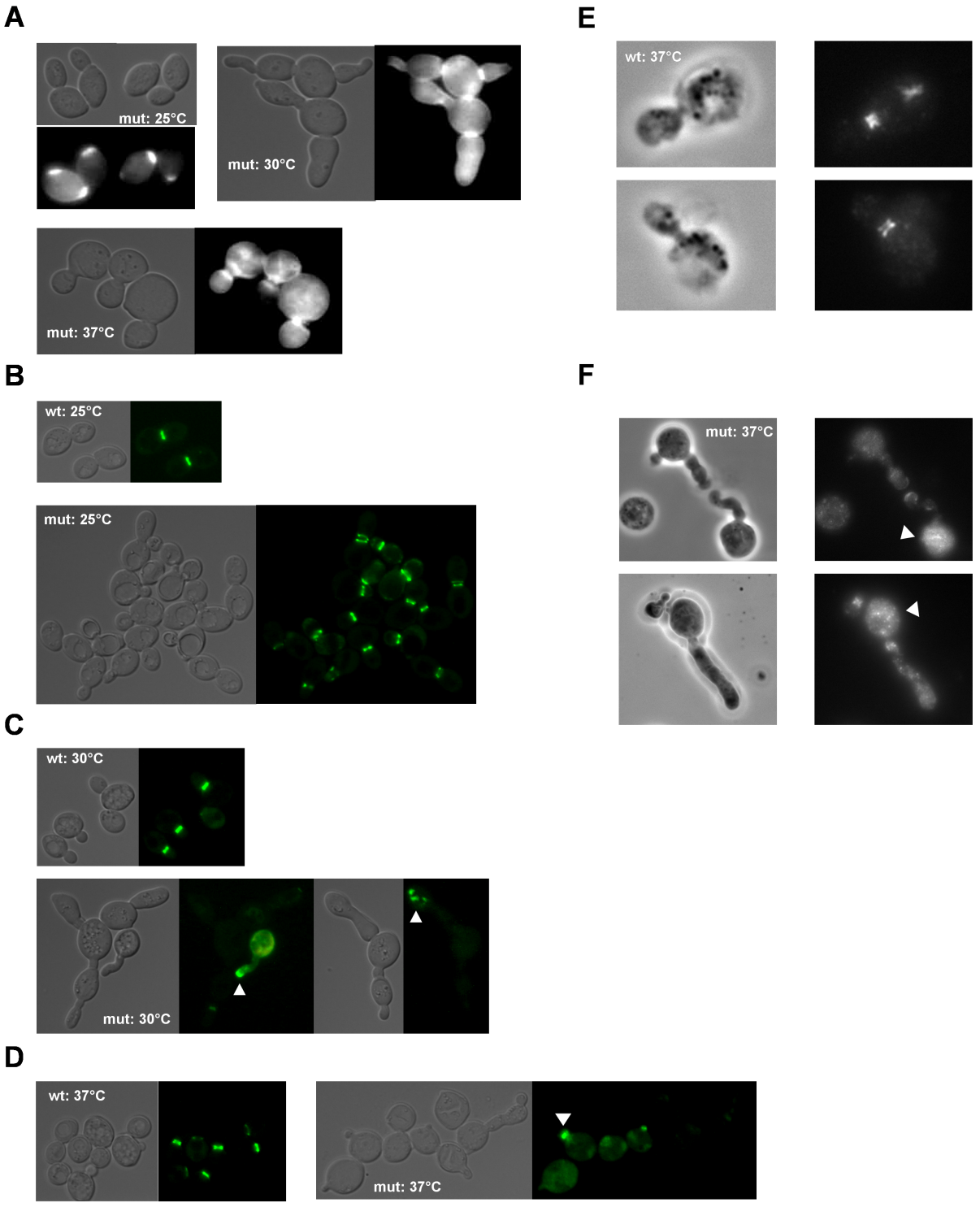

Figure 45: **(A)** Microscopic analysis of *abf1-1* mutant cells during growth at 25°C, 30°C, and 37°C. Diploid *abf1-1* cells were grown to mid-log phase before incubating for 8 hours at the indicated temperatures in liquid YPD medium. Cells were then washed and stained with Calcofluor. Photographs of differential interference contrast microscopy or Calcofluor-stained cells are shown. Arrowheads indicate the position of thickened septa. **(B - D)** Immunofluorescence of wild-type (MPY310) and mutant (MPY311) cells carrying *CDC10-GFP*. Cells were grown to mid-log phase before incubating for 8 hours at the indicated temperatures in liquid YPD medium. **(E-F)** Immunofluorescence of wild-type and mutant cells carrying with anti-Cdc11p antibodies. Cells were grown to mid-log phase before incubating overnight at 37°C in liquid YPD medium. Arrowheads mark mislocalized septin components. wt, *ABF1/ABF1* - MPY284; mut, *abf1-1/abf1-1* - MPY283

form normal chitin rings in the cell wall at the bud neck and displayed a defect in the axial budding pattern (Slater *et al.*, 1985; Flescher *et al.*, 1993; Chant *et al.*, 1995; DeMarini *et al.*, 1997). Transcriptional regulation of *CDC3* and *CDC10* by *ABF1* could explain the morphology of the *abf1-1* mutation, which showed defects in mitotic cell cycle and mother-bud separation (Figure 11). We therefore examined localization of chitin and formation of the septin ring in wild-type and *abf1-1* mutant cells. Mutant cells grown at the permissive temperature displayed chitinous bud scars, detectable by Calcofluor staining, at the poles (Figure 45, panel A). In contrast, when *abf1-1* mutant cells were grown at 30°C or incubated at 37°C for 8 hours chitin was mislocalized within the growing compartment. Wild-type cells grown at 25°C, 30°C or 37°C showed normal localization of Cdc10-GFP and hence septum formation (Figure 45, panel B-D). The same was true for mutant cells grown at 25°C. However, when mutant cells were grown at 30°C or incubated at 37°C for 8 hours they displayed a lack of septum formation between cell compartments or in some cases mislocalized to the tip of a newly forming bud. This was also shown for another component of the septin ring, Cdc11p (Figure 45, panel E-F). These results indicated that in the absence of functional *ABF1* septin rings cannot form and consequently the chitin synthase machinery cannot localize correctly.

Discussion

Experimental strategy

In the current study we have aimed at the identification of *ABF1*-target genes in a way as comprehensive as possible. This approach was based on a combinatorial approach including different high-throughput techniques: expression profiling, protein-DNA binding assays and motif prediction. The approach of identifying transcription factor target genes by examining expression profiles using microarrays is powerful as it simultaneously surveys the activity of every gene in the genome of a given organism at a given time. Using this technique one can compare expression profiles of a gene deletion strain with a corresponding wild-type strain in order to elucidate the impact of a certain gene on the transcriptional program. This has been shown for many non-essential genes (McCord *et al.*, 2003; Chu *et al.*, 1998; Williams *et al.*, 2002). Studies on essential transcription factors are more complex, because they require conditional alleles. This can be, for example, galactose-inducible mutants of a given transcription factor or temperature-sensitive alleles. The fact that these mutations only become effective under certain conditions adds another level of complication to these kinds of studies. One has to distinguish between the effect of the newly created condition (i.e. the change from glucose to galactose or from permissive to restrictive temperature). Therefore it is very important to include wild-type control cells as it was done in the work presented here. Several studies on temperature-sensitive factors have been published in the past (Grigull *et al.*, 2004; Spellman *et al.*, 1998; Shen *et al.*, 2003). Yet another level of complexity is added to a given experimental setup if a certain induced condition interferes with the process to be studied. For example, meiosis and sporulation cannot occur at temperatures higher than 34°C. This was also confirmed in the current study (Figure 13), and this observation might be based on the fact that premeiotic DNA replication is abolished at 37°C (Figure 14). Therefore we chose 33°C as the semi-permissive temperature at which *ABF1*-dependent genes would be effected in their transcriptional activity in mutant cells already without abolishing activitiy of meiotic transcripts in wild-type cells. Notably, we also included samples shifted to 37°C as controls (Figure 9).

One has to bear in mind that expression profiling methods are not direct. Alterations in transcript levels may result from secondary effects due to indirect regulatory events, which can arise when a given transcription factor regulates the activity of a second transcription factor which again regulates a number of other genes. Additionally, in an expression profiling experiment some of the target genes may not be identified because they do not exhibit changes in expression as a result of compensatory mechanisms or redundant transcription factors that exist within the cell. Direct methods that examine binding of a transcription factor to promotor sequences upstream of a gene, such as *in vivo* footprint-

ing, chromatin-immunoprecipitation or ChIP-CHIP exist. However, not every binding event to a promotor is necessarily connected to transcriptional regulation of the downstream lying gene but may, for example in the case of Abf1p, be related to DNA replication, or even be an artifact of fixation. Therefore expression profiling data and ChIP-CHIP experiments can be regarded as complementary sources of information about the regulatory capacity of a given transcription factor. A third approach aims at the identification of potential binding sites in the genome and adds another level of confidence to ChIP-CHIP and expression profiling data. However, *in silico* prediction of binding sites can only predict with a certain likelyhood where a transcription factor binds. Whether this site is indeed bound and under which conditions this is the case has to be examined experimentally.

The current study combined expression profiling and ChIP-CHIP techniques and the theoretical *in silico* approach to identify target genes of *ABF1* in vegetatively growing and sporulating cells. This resulted in 504 genes that are transcriptionally regulated by Abf1p and 352 of which were identified as targets of direct transcriptional regulation. A previous study based on a similar screen has yielded 50 genes that are under positive control of *ABF1* in vegetative cells (Rhode *et al.*, 1992). Under the same growth conditions (YPD) we found 264 target genes (Figure 19) with 29 overlapping genes from both studies. It should be emphasized that this overlap is far higher than expected by chance (p-value = 5.8×10^{-20}).

Although the overlap of both expression profiling experiments is good, only a minor fraction of genes previously found to be regulated by *ABF1* by different biochemical and molecular biological means were also identified in our current approach. Among the ∼30 *ABF1* target genes known from the literature we found only five genes (*ARO3, HIS7, HOP1, SPT2, TRP3*). One can think of different explanations for this observation. First, mutational analysis of an UAS combined with reporter gene assays (that provided experimental evidence in most studies) is distinct from using the temperature-sensitive allele *abf1-1*. This might particularly be relevant for *ABF1* for which the so-called *hit-and-run* model has been postulated in which a gene's activity is unchanged once it has been initiated even if Abf1p is removed from the gene's promotor at a later stage (Schroeder and Weil, 1998b). In contrary, in a situation in which the binding site of Abf1p is mutated it will never be bound (hit) and therefore the gene will not be activated (run). Second, one cannot exclude that *abf1-1* is still (at least partially) active after one hour of incubation at the restrictive temperature. *In vivo* footprinting assays have shown that *abf1-1* loses contact to target sequences within minutes after shifting cells to 37°C (Schroeder and Weil, 1998b). However, there still might be some remaining binding activity under these conditions not detected by the *in vivo* footprinting method. Third, the current experimental setup itself may hamper identification of *ABF1* target genes. Since *ABF1* is an essential gene one has to use conditional mutants of the gene to study it. We have made use of the temperature-sensitive allele *abf1-1* that has been isolated previously by Rhode *et. al.* (1992) and has proved to be a useful tool to characterize transcription governed by *ABF1 in vivo* (Schroeder and Weil, 1998b). However, the downside of this allele is that it has to be incubated at an elevated temperature to impair its functionality. In some cases this might have the negative consequence, that *ABF1*-target genes that are also heat-sensitive will not be detected as Abf1p-driven in an expression profile as the one performed in the current study. Examples of such cases are shown in Figure 46. All four genes have previously been shown to transcriptionally depend on *ABF1* but are down-regulated at the restrictive temperature in wild-type and mutant cells to a similar extent. Therefore these genes will be identified as temperature- but not *ABF1*-dependent. Notably, there are examples of genes that respond to both the elevated temperature and the *abf1-1* in a way so that the two effects sum up in the mutant at 37°C (Figure 18, panel B). This underlines the importance of including wild-type controls that are treated similarly to the temperature-sensitive mutants.

By combining three different experimental approaches we have three measures for whether a gene is a potential *ABF1*-target. The presence of a potential binding site and binding of Abf1p to a given promotor, without any re-

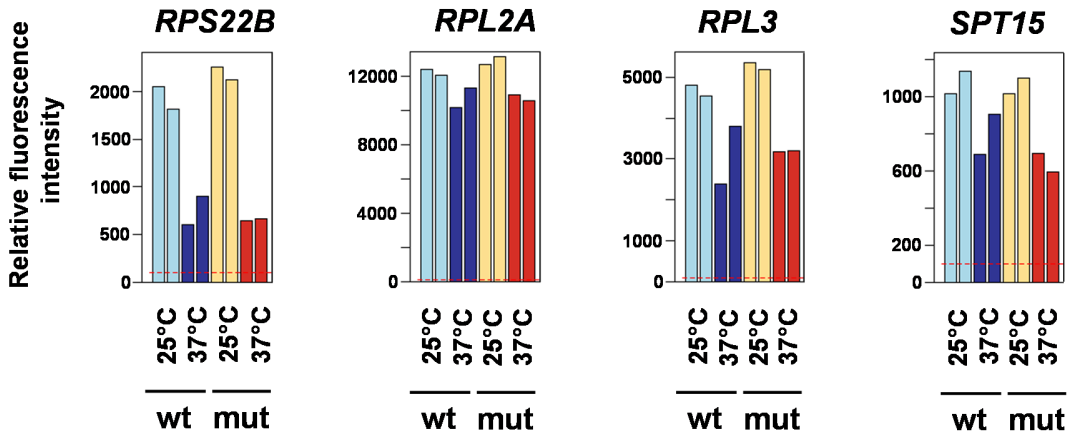

Figure 46: Examples of *ABF1* target genes that are involved in meiotic development in budding yeast. Barplots represent abundance of transcripts under different growth/incubation conditions as indicated. The red dotted line indicates the empirical threshold of reliable detection.

sulting effect on transcriptional activity on the downstream lying gene in *abf1-1* mutant cells incubated at 37°C, may still classify a gene as an *ABF1*-target. Indeed, the genome-wide binding assay and the *in silico* prediction approach find most (albeit not all) of previously identified *ABF1*-dependent genes described in the literature. Among the ∼30 *ABF1*-target genes known from the literature both techniques identified 20 genes.

Not surprisingly, the number of Abf1p-bound sequences exceeded the number of transcriptionally genes by three-fold. In the current study, we identified altogether 1642 distinct intergenic regions located upstream of 1965 genes to be enriched by Abf1p in growing and differentiating cells. However, transcriptional activity of ∼600 genes was found to be dependent upon functional *ABF1*. Several reasons might account for the observed differences between the large number of bound target sites and the comparably lower number of regulated genes. First, this discrepancy might be explained by the *hit-and-run* model. In such a case binding of Abf1p would be detected by ChIP-CHIP assays, however, regulation of corresponding genes would have escaped the detection under current expression profiling conditions. Second, some of the Abf1p-binding sites located proximal to promotors may not be related to transcription. For example, sites essential for Abf1p-mediated DNA-replication via ARSs, gene silencing or genome partitioning may belong to this group.

A comparison of predicted sites with our ChIP-CHIP data reveals a good correlation between the quality of a potential binding site (in the sense of how close it is to the initial weight matrix) and the strength and reproducibility of Abf1p binding to the genomic region in which this site is located in (Figure 33). Eight hundred of the 1049 promotors that contain predicted binding sites are also found in our ChIP-CHIP assay to be enriched. The remaining 250 sites may contain alternative binding sites for Abf1p. There are several lines of evidence that such alternative sites exist. First, in a previous study a non-consensus sequences with a strong affinity for Abf1p has been described: the promotor of *SPT15/TBP1*, the budding yeast TATA-binding protein, contains a so-called Positive Distal Element (PDE), which acts as a binding site for Abf1p *in vivo* and which differs (albeit only slightly) from the classical *ABF1* consensus site (Schroeder and Weil, 1998a). Second, a recent study based on a newly developed technique called *Protein Binding Microarrays* (PBM) has identified 256 genomic loci that are bound by Abf1p *in vitro*. Most of these intergenic regions (94%) contain a predicted binding motif for *ABF1* and a substantial fraction was also identified (88%) in our ChIP-CHIP assays. However, there are also cases of loci which are enriched by Abf1p *in vitro* but do not contain a predicted binding site (Mukherjee *et al.*, 2004).

Third, another study based on an *in vitro* SELEX approach identified a group of sequences which are strongly bound *in vivo* by Abf1p but did not contain the canonical binding motif (Beinoraviciute-Kellner *et al.*, 2005) indicating that a novel target motif may exist.

However, an attempt to identify an alternative binding sequence based on sequence alignments of target loci bound by Abf1p did not yield any significant motif. A possible explanation for this is that there might be not only one but several different binding motifs for Abf1p which complicates the identification. One could also think of a situation in which it is not Abf1p that binds to DNA but in which Abf1p interacts with a second DNA-binding protein that determines the targeted locus. In ChIP-CHIP-assays using formaldehyde as a cross-linking reagent (which covalently connects not only DNA with associated proteins but also proteins with interacting proteins) these cases would be indistiguishable from cases where Abf1p itself interacts with a genomic locus. However, such cases are currently not known for Abf1p. Last but not least, the polyclonal antibody generated and used in the current study could cross-react with another DNA-binding protein resulting in the enrichment of sequences that are targeted by this other factor. However, we proved in several biochemical assays (western blot, chromatin-immunoprecipitation and supershifts in an electro-mobility shift assay) that the polyclonal antibody is likely to be specific for Abf1p. Even more compelling is the fact that a comparison with a previous ChIP-CHIP study which monitored genome-wide binding of 204 transcription factors, including Abf1p, (Harbison *et al.*, 2004) shows a very high overlap of loci enriched by Abf1p (as detected in our experiments) and myc-tagged Abf1p.

Abf1p and DNA replication

In budding yeast, specific sequences known as autonomously replicating sequences (ARSs) can function both on plasmids and in their normal chromosomal location as origins of bidirectional DNA replication (Umek *et al.*, 1989; Brewer and Fangman, 1991; Campbell and Newlon, 1991). The archetypical ARS, *ARS1*, has been the subject of considerable analysis (Campbell and Newlon, 1991). Deletion studies have defined different functional domains of *ARS1* one of which contains the 11 bp ARS consensus sequence (ACS) which is bound by the origin recognition complex (ORC) and which is found within all yeast ARSs (Celniker *et al.*, 1984; Marahrens and Stillman, 1992). The observation that the origin recognition complex and Abf1p are bound at replication origins after DNA replication and before chromosomes segregate has led to the notion that DNA replication is not restricted to S phase simply by regulating either the availability of these proteins or their ability to bind DNA (Diffley *et al.*, 1994). Instead it was suggested that Abf1p represents the first level of protein-DNA interactions at replication origins. This ground state is laid down immediately after origin use in S phase and lasts throughout the life of the cell, even during periods of quiescence (Diffley *et al.*, 1994). Our ChIP-CHIP data confirms this notion. *ARS1*, along with a number of additional ARSs, is found to be strongly bound in a unsynchronized population of vegetatively growing cells. If Abf1p bound *ARS1* only during a short time frame of the cell cycle, this binding event could not be monitored in such a heterogenous population. This is further substantiated by the fact that also in highly synchroneously sporulating populations binding of *ARS1* is unchanged before and after meiotic S-phase (Figure 44).

Besides *ARS1*, Abf1p binding sites have been located in a number of additional ARS-elements (Buchman *et al.*, 1988; Diffley and Stillman, 1988; Eisenberg *et al.*, 1988; Francesconi and Eisenberg, 1989; Sweder *et al.*, 1988). Electro-mobility shift assays showed that Abf1p binds to *ARS1*, *ARS120*, *ARS121* with high affinity (Rhode *et al.*, 1992; Eisenberg *et al.*, 1988). Site-directed mutagenesis of the *ABF1* binding site at *ARS121* resulted in a large decrease in the stability of plasmids bearing this ARS (Walker *et al.*, 1989), while deletions of this site at *ARS1* (Diffley and Stillman, 1988; Srienc *et al.*, 1985) and the silent mating type *HMRE* ARS (Brand *et al.*, 1987) had a smaller but detectable effect on ARS plasmid stability. Competition analysis based on DNA binding assays of seven well-characterized ARSs revealed that Abf1p only binds *ARS1* and *HMRE* and therefore is not a univer-

sal ARS-binding factor (Diffley and Stillman, 1988). Our genome-wide analysis monitored binding of Abf1p to *ARS1*, *ARS120*, *ARS121*, but also showed that among the approximately 80 known ARSs (described in SGD) only 20 are bound by Abf1 strongly and with high confidence (MPR >0.8). This suggests that Abf1p is likely not to be a general ARS-binding factor in the sense of binding to every ARS element in the yeast genome, at least not under the conditions assessed in the current study.

Abf1p acts via interaction with different regulatory complexes

ABF1 is a very versatile gene involved in numerous processes, such as chromatin remodeling, positive and negative transcriptional regulation, DNA replication, nucleotide excision repair, and mRNA transport from the nucleus. How is it possible that a single protein can fulfill these different tasks? Answers to this question are provided by several reports, including three recent high-throughput studies (Gavin *et al.*, 2006; Krogan *et al.*, 2006; Reguly *et al.*, 2006), that demonstrate interaction of Abf1p with other proteins that are involved in specific processes. The following section briefly discusses complexes in which Abf1p has been detected:

Chromatin remodeling: Hta2p, Htb1p, and Hhf2p belong to the core histones required for chromatin assembly and chromatin remodeling. Interestingly, Hhf2p also contributes to telomeric silencing (Wyrick *et al.*, 1999). Hmo1p is a chromatin associated member of the High Mobility Group (HMG) family which is involved in genome maintenance (Lu *et al.*, 1996). Two components of the RSC (Remodel the Structure of Chromatin) complex, Rsc8p and Sth1p (Cairns *et al.*, 1996), were also found to be in a complex together with Abf1p. The latter protein is required for kinetochore function in chromosome segregation (Hsu *et al.*, 2003) and the induction of early meiotic genes (Yukawa *et al.*, 1999). Isw2p in turn is an ATP-dependent chromatin remodeling complex required for repression of early meiotic genes during mitotic growth upon recruitment by Ume6p (Goldmark *et al.*, 2000). It is conceivable that Isw2p functions as a linker between Abf1p and Ume6p to confer repression of for example the early meiotic genes *HOP1* or *ZIP1* during mitotic growth or late meiosis. Rvb1p and Rvb2p are two components of chromatin remodeling complexes and are required for maintaining the active state of many inducible promoters (Jónsson *et al.*, 2001). Furthermore, Abf1p has previously been shown to be involved in the recruitment of Esa1p, a member of the NuA4 histone acetylase complex, to ribosomal protein (RP) promoters (Reid *et al.*, 2000). Interactions mentioned in this paragraph are summarized in Figure 47 in subcomplex A.1 and B. Notably, not only the protein-protein interactions mentioned above may explain Abf1p's role in chromatin silencing. Our current study shows that *ABF1* transcriptionally regulates *SET3*, which is the defining member of the *SET3* histone deacetylase complex, a repressor of sporulation genes necessary for efficient transcription by RNA Polymerase II (Pijnappel *et al.*, 2001).

Transcriptional regulation: The way how Abf1p contributes to transcriptional regulation may be (at least partially) defined by interacting partners, such as *TAF5*, which is involved in RNA polymerase II transcription initiation (Lee and Young, 2000), and *MOT1*, which can activate basal transcription by regulating the distribution of Spt15p (the yeast TATA-binding protein) (Muldrow *et al.*, 1999) (Figure 47, subcomplex A.2). Sth1p, which was already mentioned in the context of its chromatin remodeling capacities, is required for the maximal expression of early meiotic genes (Yukawa *et al.*, 1999). Interaction with several subunits of the RNA polymerase I (*HMO1, RPC40*) and RNA polymerase III (*RPC34, RPC40, RPC53, RPC82, RPO31*) complexes (Figure 47, subcomplex A.3) may explain how Abf1p contributes to the transcription of ribosomal and transfer RNAs. *CKA1, CKA2, CKB1* and *CKB2* are four casein kinases with roles in growth and proliferation (Ahmed *et al.*, 2002). Notably, substrates of these kinases include transcription factors and all RNA polymerases (Ackermann *et al.*, 2001) (Figure 47, subcomplex C.1).

DNA replication: Cdc6p has been shown to stimulate Abf1p DNA binding activities to

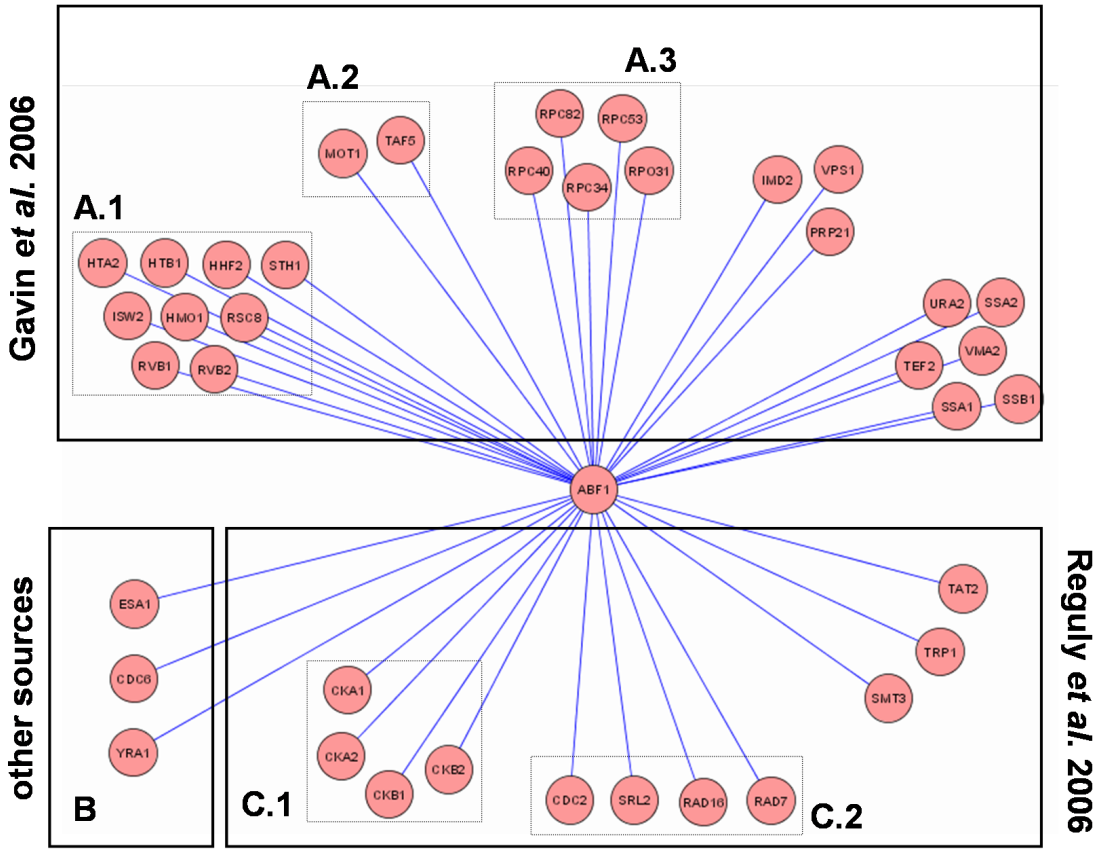

Figure 47: Protein-protein interactions of Abf1p with other proteins. Different data sources from a high-throughput study (Gavin *et al.* 2006), and a literature curated dataset (Reguly *et al. 2006*, also see Table 3, ref. 15) were supplemented with interactons found by our own literature search (other sources). The three sources are separated by solid boxes. Subcomplexes which are discussed in the text are boxed with dotted lines.

an ARS element (Feng *et al.*, 1998). Cdc6p is an essential ATP-binding protein and a component of the pre-replicative complex (pre-RC) required for DNA replication (Piatti *et al.*, 1995; Cocker *et al.*, 1996; Detweiler and Li, 1997). No direct physical interaction between Cdc6p and Abf1p could be shown. However, binding of Abf1p to *ARS1* was demonstrated to be increased *in vitro* by addition of recombinant Cdc6p (Feng *et al.*, 1998) (Figure 47, subcomplex B). Furthermore, a binding mutant of Cdc6p (with the four basic amino acid residues, ^{29}KRKK, mutated to the neutral amino acids, ^{29}AGAS) showed increased rates of plasmid loss indicating defects in DNA synthesis. This mutant also did not stimulate Abf1p-binding activity to *ARS1*, which suggested that Cdc6p DNA-binding activity may play a role in the initiation of DNA replication (Feng *et al.*, 2000).

mRNA export: Abf1p was demonstrated to co-immunoprecipitate with Yra1p (Hieronymus and Silver, 2003), which is a nuclear protein that binds to RNA and is required for export of poly(A)+ mRNA from the nucleus (Portman *et al.*, 1997; Zenklusen *et al.*, 2001; Strässer and Hurt, 2000) (Figure 47, subcomplex B). The interaction between both proteins is RNase-sensitive and knockdown of Abf1p in an Abf1p degron strain resulted in a slight nuclear accumulation of poly(A)+ RNA. This suggested that mRNA export from the nucleus is connected to the transcriptional machinery.

Nucleotide excision repair: Abf1p's contribution to the removal of photoproducts has been demonstrated *in vitro* and *in vivo*. Notably, *abf1-1* mutant cells are very sensitive to killing by ultraviolet radiation (Reed *et al.*, 1999). Abf1p has been found to form a subcomplex with Rad7p and Rad16p (Reed *et al.*, 1999; Yu *et al.*, 2004). Furthermore, it has

been found to interact with Srl2p (Krogan *et al.*, 2006), a protein of unknown function whose overexpression suppresses the lethality caused by a rad53Δ mutation (Desany *et al.*, 1998). *RAD53* is an essential gene required for the transcriptional and cell cycle responses to DNA damage and DNA replication blocks. This suggests that also Srl2p is related to DNA damage repair. Additionally, *ABF1* was found in a genetic screen for genes that are synthetic lethal with mutations in *CDC2/POL3*, which encodes the proofreading DNA polymerase delta and plays a role in base excision repair of exogenous DNA methylation damage (Blank *et al.*, 1994). Interactions mentioned in this paragraph are summarized in Figure 47, subcomplex C.2.

ABF1-target genes and phenotypes of *abf1-1*

Our current study involved phenotypical analysis of diploid *abf1-1* mutants during mitotic growth and meiotic development. Previously, it has been shown that growth of haploid mutant cells arrested in a single cell cycle at the non-permissive temperature. Arrested cells were enlarged and had small elongated buds. The growth rate of the *abf1-1* mutants at the permissive temperature was found to be nearly identical to that of the isogenic *ABF1* strain. Growth at 30°C was slowed down in the mutant and completely abolished at temperatures higher than 32°C (Rhode *et al.*, 1992). In the current study we have reproduced all of these phenotypes in diploid *abf1-1* (Figures 10 and 11). Additionally, we observed a novel phenomenon in mutant cells grown at the semi-permissive temperature that has not been reported previously. Cells grown at 30°C for several hours formed elongated buds and did not separate properly forming cell bodies of 4 and more cells (Figure 11).

Our high-troughput data suggested different explanations for this phenotypic peculiarity. For example, we found *SWM1* to be under transcriptional control of *ABF1*. *SWM1* was initially described to be a non-essential and meiosis-specific gene required for spore wall assembly (Ufano *et al.*, 1999). Later on it was shown that *SWM1* is also required to maintain cell wall integrity during vegetative growth at high temperatures (Ufano *et al.*, 2004). Strains with deletions in this genes and incubated at 37°C for several hours either partially arrested as large-budded cells or accumulated cell bodies with 4 non-separated cells (Schwickart *et al.*, 2004). These phenotypes are also observed in *abf1-1* mutant cells incubated at 37°for several hours (Figure 11, panel B). Notably, already one hour after shifting mutant cells to the restrictive temperature *SWM1* transcript levels were below the detection treshold (Figure 41, panel C) and might therefore have the same effect as a null mutation of *SWM1*.

Furthermore, our data suggested that Abf1p plays a role in cytokinesis and the transcriptional activation of a subset of septins. We found that *abf1-1* mutant cells form elongated buds, display aberrant chitin localization and septum formation at the semi-permissive and the restricitive temperature (Figure 45). These phenomena could be explained by the fact that transcriptional levels of *CDC10* and *CDC3* decrease in *abf1-1* mutants at the restricitve temperature. This is in keeping with earlier observations that temperature-sensitive septin mutants arrest with elongated buds that fail to undergo cytokinesis (Hartwell, 1971; Adams and Pringle, 1984). Consequently, these proteins became known as septins for their suspected role in forming the septum between the mother and the daughter cells. Additionally, these mutants were demonstrated to be unable to form normal chitin rings in the cell wall at the bud neck and display a defect in the axial budding pattern (Slater *et al.*, 1985; Flescher *et al.*, 1993; Chant *et al.*, 1995; DeMarini *et al.*, 1997). The formation of the septin ring apparently requires interdependent interactions between the septin proteins. Shifting any of the septin mutants to the non-permissive temperature resulted not only in the disappearance of the mutated septin but also of the other septins from the bud neck and the loss of the 10-nm filaments (Haarer and Pringle, 1987; Ford and Pringle, 1991; Kim *et al.*, 1991).

We have also studied effects of the *abf1-1* mutation on meiosis and spore formation. We found, that the impact of this mutation on meiotic development is much stronger than on vegetative growth. Already at the permissive temperature (25°C) mutant cells were compro-

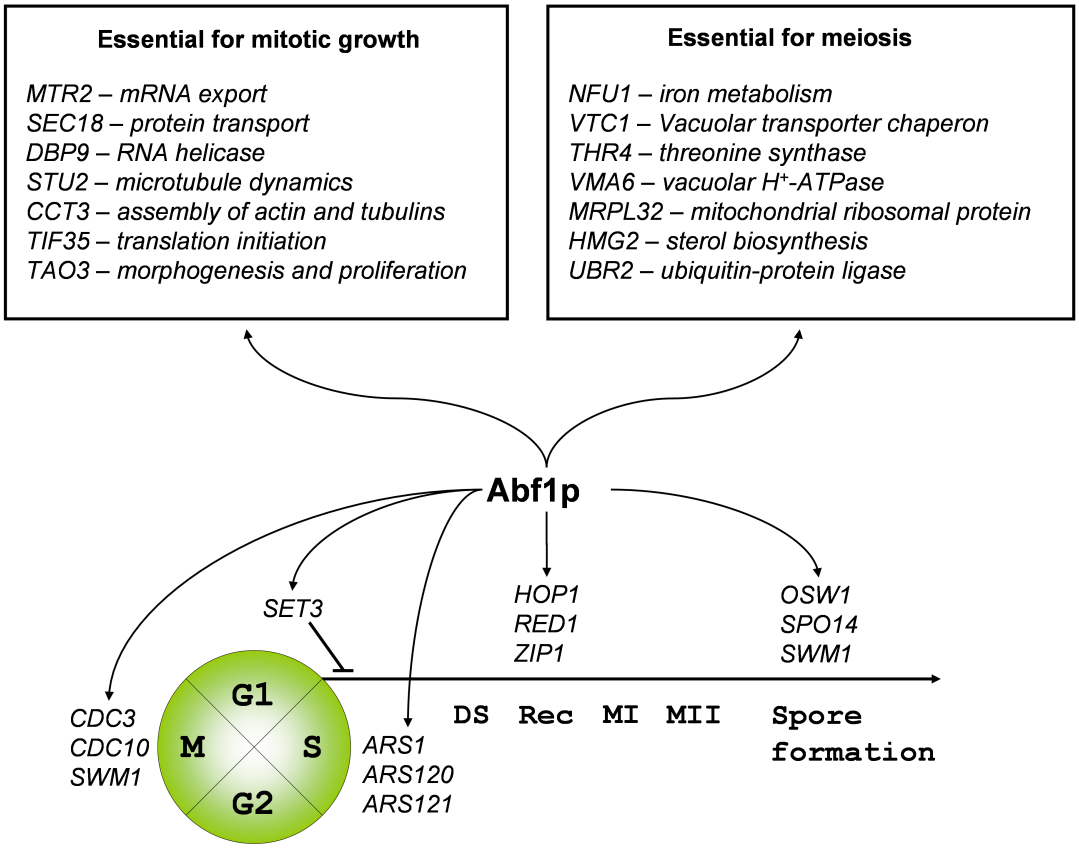

Figure 48: **(A)** Coordination of mitotic growth and meiotic development by *ABF1*. Transcriptional targets of *ABF1*-regulation involved in both processes are indicated by arrows. Boxes show example genes which are essential and involved in diverse processes. Lower panel depicts genes that are required at certain stages of the mitotic and meiotic cell cycle. DS, premeiotic DNA synthesis; Rec, recombination; MI, meiosis I; MII, meiosis II.

mised in undergoing sporulation efficiently. After 5 days in sporulation medium only 30% of the cells had formed asci (as compared to 60% of the reference wild-type strain), two-thirds of which were dyads.

Our high-throughput data suggested different explanations for the observed meiotic phenotypes. First, *SPO14*, a phospholipase D (PLD), which is essential for Golgi function (Rudge *et al.*, 2001) and is also required for commitment to meiosis and spore formation (Honigberg *et al.*, 1992; Rabitsch *et al.*, 2001) was among genes that are less expressed in mutant cells regardless of the temperature in growing and sporulating cells (Figure 41, panel A). Previously, spo14Δ mutants were demonstrated to be defective in performing the second meiotic division and spore formation, while earlier meiotic events were largely unaffected which resulted in the accumulation of bi-nucleate cells during meiosis (Honigberg and Esposito, 1994). Strikingly, this phenomenon was also observed in sporulating *abf1-1* mutants (Figure 13). The fact that transcriptional activitiy of *SPO14* is already reduced in mutant cells at the permissive temperature might explain the low efficiency of *abf1-1* mutants to undergo meiotic development already under these conditions.

Several other genes involved in meiotic development could also account for the low ability of *abf1-1* to sporulate. For example, we found four genes (*SRV2, SSZ1, VMA6, VPS54*) to be direct targets of *ABF1* (Figure 43, panels A-C). These genes were previoulsy demonstrated to have defects in undergoing the two meiotic divisions and to activate an *IME1*-lacZ reporter gene (Enyenihi and Saunders, 2003). Yeast strains which were deleted in one of the four genes showed reduced level of spore formation as tested after 3 days on acetate. Importantly, the strains were able to grow on a non-

fermentable carbon source (glycerol), suggesting that the observed meiotic defect was not due to the cells' inability to respirate. Again, this might explain the reduced ability of *abf1-1* mutant cells to undergo both meiotic divisions and form asci with high efficiency (Figure 13).

We also found two components of the synaptonemal complex (*HOP1, RED1*) to be transcriptionally dependent on *ABF1*. The promotor of a third component of the SC (*ZIP1*) was identified in our ChIP-CHIP assay to be bound by Abf1p and to harbor a well-conserved binding site for *ABF1*. Notably, *HOP1* was still expressed in the mutant, although not to wild-type levels (Figure 42, panel A). In comparison, the message abundance of *RED1* was below the treshold in mutant cells (panel B). Previously, it has been reported that red1Δ mutant cells fail to make any SC or any obvious precursors to the SC. Additionally, Hop1 protein localizes only in areas that also contain Red1p, and this localization depends on Red1p function (Smith and Roeder, 1997).

Figure 48 depicts in an overview of how *ABF1* transcriptionally contributes to mitotic growth and meiotic development in budding yeast.

Materials and Methods

Yeast strains and culture conditions The genotypes and sources of yeast strains used in this study are shown in table (Table 2). Vegetative cultures were porpagated in either YPD (1% yeast extract, 2% bacto peptone, 2% glucose), YPA (1% yeast extract, 2% bacto peptone, 1% potassium acetate). Wild-type and mutant strains were cultured in YPD medium at 25°C until a cell density of 2 $\times 10^7$ cells/ml. The culture was split into two, and one half was rapidly heated to 37°C in a water bath and subequently incubated in the shaker at 37°C. The other half was incubated in the shaker at 25°C. After an 1 hour incubation, the cultures were harvested and total RNA was isolated using the hot phenol method. Logarithmically growing YPD cultures were also used to inoculate YPA at a cell density of 2 $\times 10^5$ cells/ml. The cultures were expanded overnight at 25°C to a density of 1.5 $\times 10^7$ cells/ml. Again the culture was split into two, and one half was rapidly heated to 37°C in a water bath and subequently incubated in the shaker at 37°C. The other half was incubated in the shaker at 25°C. After 1 hour incubation, the cultures were harvested and total RNA was isolated using the hot phenol method. Total RNA quality was measured using the Bioanalyzer from Agilent.

Antibodies Rabbit antibodies were raised against His6-Abf1p (amino acids 264 to 513) expressed in bacteria. Commercial antibodies from Roche (12CA5) were used to detect HA-tagged proteins.

Electro mobility shift assay For elecotromobility shift assays (EMSA) oligonucleotides were annealed by mixing equal amounts of strands together with NaCl, incubating at 95°C for 15 min, and slowly cooling to 25°C overnight. Doublestranded DNA was then end labeled with [γ-^{32}P] ATP by using klenow enzyme and subsequently purified with Quick Spin Columns for radiolabeled DNA purification (Sephadex G-50) from Roche columns. Binding reactions were acrried out in a solution containing 10 mM Tris-HCl (pH 7.5), 40 mM NaCl, 4 mM MgCl, 6% glycerol, 200 μg/ml dI-dC polymer from Amersham, and labeled oligonucleotide (10000 cpm) in a total volume of 20 μl at room temperature for 20 min. Crude extracts were prepared as described above. Samples were analyzed on a 6% polyacrylamide gel (run in 0.5 x Tris-borate-EDTA buffer for 120 min at 200 V). Gels were dried after electrophoresis and exposed to a phosphor screen.

ChIP and ChIP-CHIP For ChIP experiments 20 ml cultures containing about 4 $\times 10^8$ cells were treated with 1% formaldehyde for 15 min at room temperature and ChIP was performed in 1 M NaCl-containing lysis buffer as described (Lieb *et al.*, 2001). After cell lysis and chromatin shearing by sonication (to 0.5-1 kb), each sample was divided in two halves and added to reactions in the presence and absence of antibody. Abf1p was immuno-precipitated using a poly-clonal antibody against amino acids 264-513. For PCR analysis, DNA is eluted from protein G beads in 20 μl TE buffer for 20 minutes at 95°C and the crosslinks are reversed over night at 65°C. Multiplex PCR is performed under excess of primers and nucleotides with HiFi polymerase (Roche). 200 μM dNTPs and 50 pmoles of each primer are incubated at 94°C for 2 minutes. Then ten cycles at 94°C 15 sec, 57°C for 15 sec, 72°C for 30 sec and 20 cycles at 94°C 15 sec, 57°C for 15 sec, 72°C for 30 sec plus an additional 5 sec at each cycle were run. Final incubation is at 72°C for 5 min. PCR products were analyzed using polyacrylamide/TBE gels (6%) and SYBR Green staining.

For ChIP-CHIP analysis, DNA was eluted from the protein G beads, purified, and amplified as described (Lieb *et al.*, 2001) with modifications: the immunoprecipitate is eluted from the beads in 100 μl elution buffer (50 mM Tris pH 8.0, 10 mM EDTA, 1% SDS) for 30 min at 65°C. When appropriate, the eluate and an aliquot of the whole-cell extract input (WCE) diluted in elution buffer are incubated overnight at 65°C to reverse crosslinks. One volume of TE, 100 μg/ml Proteinase K and 20 μg glycogen are added for 1 hour at 37°C. DNA is recovered using a Qiaquick PCR purification kit (Qiagen), ethanol-precipitated, and resuspended in 20 μl TE and 100 μg/ml RNase. One third of the immuno-precipitated DNA or 1/20000 of DNA from the WCE is amplified by random primer extension followed by PCR amplification, incorporating amino-allyl dUTP for subsequent dye coupling. PCR Microarray printing and hybridization is carried out as follows: Yeast ORFs and intergenic regions are amplified by PCR using the complete set of Yeast GenePairs Primers from Research

Genetics. After isopropanol precipitation, PCR products are re-suspended in 3 x SSC. Robotic microarray printing is performed on polylysine-coated glass slides or Corning Ultra Gaps II slides using a Microgrid TAS arraying robot (Biorobotics). Post-processing of the slides, labeling of 3-6 μg of DNA with Cy3 or Cy5 fluorophors, microarray hybridization at 63°C, and array washing steps are done as described at (supplemental table, ref. 7)

cRNA target synthesis and GeneChip hybridization Culture conditions and array hybridization protocols were carried out as published with a few modifications (Primig *et al.*, 2000). cRNA target molecules were prepared from 50 ml cultures at 3-5 $\times 10^7$ cells/ml. Samples were hybridized to yeast S98 GeneChips (Affymetrix) that contain probes for 6400 transcripts (supplemental table, ref. 8). 15 μg of yeast total RNA was purified using RNeasy Mini-Spin columns (Qiagen) employing standard protocols provided by the manufacturer. The cell pellets were re-suspended in RLT buffer and lysed by shearing in a 2 ml syringe. 600 μl of the supernatant were mixed with 600 μl of 70% ethanol, loaded onto an RNeasy column, washed and eluted in 100 μl of double-distilled water. Quality of total RNA was monitored using RNA Nano 6000 Chips processed with the 2100 Bioanalyzer (Agilent). Biotin labeling of RNA was done as outlined in the Expression Analysis Technical Manual (supplemental table, ref. 9) with minor modifications. Single-stranded cDNA was synthesized by mixing 13 μg of total RNA with oligo-dT and incubating the reaction mix with Superscript II reverse transcriptase (Invitrogen) at 42°C for one hour. After synthesis of the second cDNA strand using the Superscript Double-Stranded cDNA Synthesis Kit (Invitrogen), the material was extracted with phenol-chloroform-isoamyl alcohol and precipitated with 0.5 volumes of 7.5 M ammonium acetate and 2.5 volumes of ethanol. The cDNA was employed for an *in vitro* transcription (IVT) reaction using the BioArray High Yield RNA Transcript Labeling Kit T7 (Enzo) to synthesize cRNA in the presence of biotin-conjugated UTP and CTP analogs. Approximately 50 μg of cRNA was purified over RNeasy Mini-Spin columns and analyzed again on RNA Nano 6000 Chips. The cRNA targets were incubated at 94°C for 35 minutes and the resulting fragments of 50-150 nucleotides were again analyzed with RNA Nano 6000 Chips. All synthesis reactions were carried out in a PCR machine (Biometra T1 Thermocycler) to ensure optimal temperature control. 220 μl hybridization cocktail containing heat-fragmented and biotin-labeled cRNA at a concentration of 0.05 μg/μl were injected into GeneChips and incubated at 45°C on a rotator in a Hybridization Oven 640 (Affymetrix) over night at 60 rpm. The arrays are washed and stained with a streptavidine-phycoerythrin conjugate (SAPE; Molecular Probes). To increase signal strength a standard antibody amplification protocol was employed (EukGE-WS2v4; Affymetrix Expression Analysis Manual). The GeneChips were processed with a GeneArray Scanner (Agilent). A detailed description of current cRNA synthesis protocols is available in (Schlecht *et al.*, 2004) and the corresponding web portal (supplemental table, ref. 10).

Abf1p motif prediction For each intergenic region in *S. cerevisiae* orthologous intergenic regions from the 4 other *sensu stricto Saccharomyces* species were obtained using the ORF annotations of (2003) and (2003). All groups of orthologous intergenic regions were aligned with MLagan (Brudno *et al.*, 2003). MotEvo scanned these alignments with the .... Sites were scored in a two-step process: First, we obtain the following two probabilities. The *cerevisiae* sequence segment $s = (s_j)$ is scored under the model for the weight matrix $mu = (\mu_j)$:

$$p_\mu(s) = \prod_j \frac{\mu^j_{s_j} n + 0.5}{n+2}. \quad (1)$$

Here, n denotes the number of sequences used to construct the weight matrix. Then s is scored under a third order Markov background model

$$p_b(s) = \prod_j b(s_j|s_{j+1}, s_{j+2}, s_{j+3}), \quad (2)$$

where b denotes the background probability for a base given the bases to the right of it. Only sequence seqments that score better to the weight matrix ($p_\mu(s)/p_b(s) > 1$) are kept for further investigation. For these segments we then collected all sequence segments from the multiple alignment that have the same gap pattern as the *cerevisiae* sequence at that location. In step two, for the gapless multiple alignment so obtained, the following probabilities were calculated. P_1: The probability that the sequences in the alignment contain binding sites for the WM and have evolved under the constraints set by this weight matrix. To calculate this probability, MotEvo uses the same probabilistic evolutionary model as PhyloGibbs (Siddhartan *et al.*, 2005). In this model, the transition probability (under model x) from base α to β is given by

$$p_x(\beta|\alpha) = \delta_{\alpha\beta} q + p(\beta|x)(1-q). \quad (3)$$

where δ is 1 if the indices are equal and 0 otherwise, q is the probability that no mutation takes place, and $p(\beta|x)$ is either $\frac{\mu^j_{s_j} n+0.5}{n+2}$ or $b(s_j|s_{j+1}, s_{j+2}, s_{j+3})$, depending on whether x stands for μ or b (c.f. equations (1) and (2)). This means β evolved from α either according to the WM model or according to the background model. To obtain P_1, we do the product of $p_\mu(\beta|\alpha)$ over those species scoring better to the WM multiplied by the product of $p_b(\beta|\alpha)$ over those species scoring better to background. This expression is summed over all ancestors weighed by (1) and finally multiplied over all column positions. P_2: The probability that the sequences have evolved according to the background evolution model, i.e. the same expression as for P_1 but all $p_\mu(\beta|\alpha)$ replaced by $p_b(\beta|\alpha)$. P_3: the probability that the sequences that contain sites for another regulatory factor have evolved according to the constraints set by another regulatory factor. We found that it is especially important to include this third possibility since otherwise conserved segments that show a vague similarity to the WM may be mistaken for binding sites of this transcription factor, whereas in reality

these are conserved sites for another transcription factor. We obtain this probability by replacing $p_b(\beta|\alpha)$ by $p_\omega(\beta|\alpha)$ in P_2 and integrating over all possible WMs ω. We now took as the 'background' probability P_b of the alignment, given that no site for our weight matrix occurs, the maximum of the two probabilities P_2 and P_3, i.e. $P_b = \max(P_2, P_3)$. Finally, with a prior probability π that a binding site for the weight matrix in question occurs at a randomly chosen position in an intergenic region the posterior probability p that a site for the weight matrix occurs becomes:

$$p = \frac{p_1\pi}{p_1\pi + p_b(1-\pi)}, \tag{4}$$

Based on previous estimates that the total number of binding sites in yeast intergenic regions (Chin *et al.*, 2005; Siddhartan *et al.*, 2005) is about 10 per intergenic region, there are about 200 sequences specific transcription factors in yeast, and that intergenic regions are on average about 700 bp wide we use $\pi = 10/(200 * 700)$ as a reasonable guess for the *a priori* probability that a site for a particular weight matrix occurs at a particular position.

Acknowledgements

We thank V. Borde for excellent technical assistance for establishment of the ChIP-CHIP technique. P. Demougin and the GeneChip platform at the Life Sciences Training Facility are supported by the Biozentrum, the Swiss Tropical Institute, the Department of Clinical and Biological Sciences of the Basel University Hospital and the Department of Clinical Research at the University of Bern.

General conclusion and outlook

This work contributes to the understanding of the transcriptional program that underlies mitotic growth, meiosis, and gametogenesis. In the first part regulation of these processes are explored in the model organisms *Rattus norvegicus* on a genome-wide level, leading to the first comprehensive large-scale expression profiling analysis of mammalian male gametogenesis. As a result our knowledge about genes which are transcriptionally regulated in germ cells undergoing mitotic growth, meiosis, and gametogenesis was substantially enlarged. Our study yielded numerous yet uncharacterized transcripts that are likely to be involved in spermatogenesis and fertility and therefore provides the scientific community with interesting candidate genes for further analysis. This work also served as a pilot feasibility study and is currently followed up by an even larger comparative approach exploring the transcriptonal program governing gametogenesis in two mammalian model organisms *Mus musculus*, *Rattus norvegicus*, and *Homo sapiens* in our laboratory.

In the second part we have defined the transcriptional target genes of *ABF1* in *Saccharomyces cerevisiae* and thereby enlarged our knowledge of its contributions to mitotic growth and meiotic development. In this work we have explored the power of combining different high-throughput techniques for target gene discovery. This led to the identification of roles for Abf1p in the transcriptional activation of components required for cytokinesis, the formation of the synaptonemal complex, and spore formation. In this context it will be important to assess the impact of *ABF1*-regulation on different stages of meiosis and sporulation by additional experimental means, such as microscopic analysis of components of the SC or the septins during initial steps of spore wall formation. Apart from the identification of direct target genes of Abf1p, we have also defined a group of genes for which it is currently not clear whether they are true *ABF1*-targets. For example there are numerous cases of genes whose promotors were bound in our ChIP-CHIP assay and contain a potential binding motif, but were not transcriptionally regulated in the anticipated way. It will be interesting to study such cases by mutational anaylsis of the binding site and to assess the impact on transcription. Notably, this is a different approach from that we have taken in this study, because it creates situations where a promotor is never bound by Abf1p and therefore never has the chance to contribute to transcriptional activation of the corresponding gene. This approach will also elucidate how important the earlier proposed *hit-and-run* model is, and will enlarge our machanistic understanding of transcriptional regulation governed by Abf1p. A detailed characterization of *in vitro* and *in vivo* binding activities of *abf1-1* to target sites will also contribute to this understanding. In this context it will be interesting to determine whether *abf1-1* is only deficient in binding to target sites or also in interacting with other proteins. Ultimately, we would like to unravel the complete regulatory network of Abf1p, Ume6p, Ndt80p, and novel yet unknown transcriptional factors governing transition from mitotic growth to meiosis and the progression through the meiotic developmental pathway.

Supplemental table. Web addresses mentioned in this thesis

Reference	Name of database	Web address
1	The Ovarian Kaleidoscope	http://ovary.stanford.edu/
2	Mammalian Reproductive Genetics	http://mrg.genetics.washington.edu/
3	Germonline	http://www.germonline.org/
4	European mirror	http://germonline.igh.cnrs.fr/
5	Japanese mirror	http://germonline.biochem.s.utokyo.ac.jp/
6	American mirror	http://germonline.yeastgenome.org/
7	Microarray protocols and software	http://www.microarrays.org/
8	Affymetrix	http://www.affymetrix.com/
9	Affymetrix Technical Note	http://www.affymetrix.com/support/technical/technotes/statistical_algorithms_technote.pdf
10	Rat study	http://www.bioz.unibas.ch/personal/rat_spermatogenesis/
11	Bioconductor	http://bioconductor.org/
12	R project	http://www.r-project.org/
13	*Saccharomyces* Genome Database	http://yeastgenome.org/
14	Arrayexpress	http://www.ebi.ac.uk/arrayexpress/
15	BioGRID	http://thebiogrid.org/
16	Rat Genome Database	http://rgd.mcw.edu/

References

Ackermann, K., Waxmann, A., Glover, C. V., and Pyerin, W. (2001). Genes targeted by protein kinase CK2: a genome-wide expression array analysis in yeast. *Mol Cell Biochem* **227(1-2)**, 59–66.

Adams, A. E. and Pringle, J. R. (1984). Relationship of actin and tubulin distribution to bud growth in wild-type and morphogenetic-mutant Saccharomyces cerevisiae. *J Cell Biol* **98(3)**, 934–945.

Adham, I. M., Nayernia, K., Burkhardt-Göttges, E., O, T., Dixkens, C., Holstein, A. F., and Engel, W. (2001). Teratozoospermia in mice lacking the transition protein 2 (Tnp2). *Mol Hum Reprod* **7(6)**, 513–520.

Adham, I. M., Nayernia, K., and Engel, W. (1997). Spermatozoa lacking acrosin protein show delayed fertilization. *Mol Reprod Dev* **46(3)**, 370–376.

Ahmed, K., Gerber, D. A., and Cochet, C. (2002). Joining the cell survival squad: an emerging role for protein kinase CK2. *Trends Cell Biol* **12(5)**, 226–230.

Akhtar, A. and Becker, P. B. (2000). Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in Drosophila. *Mol Cell* **5(2)**, 367–375.

Almstrup, K., Nielsen, J. E., Hansen, M. A., Tanaka, M., Skakkebaek, N. E., and Leffers, H. (2004). Analysis of cell-type-specific gene expression during mouse spermatogenesis. *Biol Reprod* **70(6)**, 1751–1761.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25(1)**, 25–29.

Atcheson, C. L., DiDomenico, B., Frackman, S., Esposito, R. E., and Elder, R. T. (1987). Isolation, DNA sequence, and regulation of a meiosis-specific eukaryotic recombination gene. *Proc Natl Acad Sci U S A* **84(22)**, 8035–8039.

Baudat, F., Manova, K., Yuen, J. P., Jasin, M., and Keeney, S. (2000). Chromosome synapsis defects and sexually dimorphic meiotic progression in mice lacking Spo11. *Mol Cell* **6(5)**, 989–998.

Behrens, A., Genoud, N., Naumann, H., Rülicke, T., Janett, F., Heppner, F. L., Ledermann, B., and Aguzzi, A. (2002). Absence of the prion protein homologue Doppel causes male sterility. *EMBO J* **21(14)**, 3652–3658.

Beinoraviciute-Kellner, R., Lipps, G., and Krauss, G. (2005). In vitro selection of DNA binding sites for ABF1 protein from Saccharomyces cerevisiae. *FEBS Lett* **579(20)**, 4535–4540.

Ben-Shlomo, I., Vitt, U. A., and Hsueh, A. J. (2002). Perspective: the ovarian kaleidoscope database-II. Functional genomic analysis of an organ-specific database. *Endocrinology* **143(6)**, 2041–2044.

Blank, A., Kim, B., and Loeb, L. A. (1994). DNA polymerase delta is required for base excision repair of DNA methylation damage in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **91(19)**, 9047–9051.

Bolstad, B. M., Irizarry, R. A., Astrand, M., and Speed, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19(2)**, 185–193.

Bowdish, K. S., Yuan, H. E., and Mitchell, A. P. (1994). Analysis of RIM11, a yeast protein kinase that phosphorylates the meiotic activator IME1. *Mol Cell Biol* **14(12)**, 7909–7919.

Bowdish, K. S., Yuan, H. E., and Mitchell, A. P. (1995). Positive control of yeast meiotic genes by the negative regulator UME6. *Mol Cell Biol* **15(6)**, 2955–2961.

Brand, A. H., Micklem, G., and Nasmyth, K. (1987). A yeast silencer contains sequences that can promote autonomous plasmid replication and transcriptional activation. *Cell* **51(5)**, 709–719.

Brewer, B. J. and Fangman, W. L. (1991). Mapping replication origins in yeast chromosomes. *Bioessays* **13(7)**, 317–322.

Brudno, M., Do, C. B., Cooper, G. M., Kim, M. F., Davydov, E., Green, E. D., Sidow, A., Batzoglou, S., and Sidow, A. (2003). LAGAN and Multi-LAGAN: efficient tools

for large-scale multiple alignment of genomic DNA. *Genome Res* **13(4)**, 721–731.

Buchman, A. R., Kimmerly, W. J., Rine, J., and Kornberg, R. D. (1988). Two DNA-binding factors recognize specific sequences at silencers, upstream activating sequences, autonomously replicating sequences, and telomeres in Saccharomyces cerevisiae. *Mol Cell Biol* **8(1)**, 210–225.

Buck, M. J. and Lieb, J. D. (2004). ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. *Genomics* **83(3)**, 349–360.

Cairns, B. R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B., and Kornberg, R. D. (1996). RSC, an essential, abundant chromatin-remodeling complex. *Cell* **87(7)**, 1249–1260.

Campbell, J. L. and Newlon, C. S. (1991). Chromosomal DNA replication. *Cold Spring Harbor, NY: Cold Spring Harbor Laboratory* **1**, 41–146

Celniker, S. E., Sweder, K., Srienc, F., Bailey, J. E., and Campbell, J. L. (1984). Deletion mutations affecting autonomously replicating sequence ARS1 of Saccharomyces cerevisiae. *Mol Cell Biol* **4(11)**, 2455–2466.

Chant, J., Mischke, M., Mitchell, E., Herskowitz, I., and Pringle, J. R. (1995). Role of Bud3p in producing the axial budding pattern of yeast. *J Cell Biol* **129(3)**, 767–778.

Chin, C. S., Chuang, J. H., and Li, H. (2005). Genome-wide regulatory complexity in yeast promoters: separation of functionally conserved and neutral sequence. *Genome Res* **15(2)**, 205–213.

Cho, C., Willis, W. D., Goulding, E. H., Jung-Ha, H., Choi, Y. C., Hecht, N. B., and Eddy, E. M. (2001). Haploinsufficiency of protamine-1 or -2 causes infertility in mice. *Nat Genet* **28(1)**, 82–86.

Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P. O., and Herskowitz, I. (1998). The transcriptional program of sporulation in budding yeast. *Science* **282(5389)**, 699–705.

Chu, S. and Herskowitz, I. (1998). Gametogenesis in yeast is regulated by a transcriptional cascade dependent on Ndt80. *Mol Cell* **1(5)**, 685–696.

Clarke, G. and Cooke, D. (1998). A Basic Course in Statistics. *Oxford, NY: Oxford University Press*

Cliften, P., Sudarsanam, P., Desikan, A., Fulton, L., Fulton, B., Majors, J., Waterston, R., Cohen, B. A., and Johnston, M. (2003). Finding functional features in Saccharomyces genomes by phylogenetic footprinting. *Science* **301(5629)**, 71–76.

Cocker, J. H., Piatti, S., Santocanale, C., Nasmyth, K., and Diffley, J. F. (1996). An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast. *Nature* **379(6561)**, 180–182.

Colaiácovo, M. P., Stanfield, G. M., Reddy, K. C., Reinke, V., Kim, S. K., and Villeneuve, A. M. (2002). A targeted RNAi screen for genes involved in chromosome morphogenesis and nuclear organization in the Caenorhabditis elegans germline. *Genetics* **162(1)**, 113–128.

Coluccio, A., Bogengruber, E., Conrad, M. N., Dresser, M. E., Briza, P., and Neiman, A. M. (2004). Morphogenetic pathway of spore wall assembly in Saccharomyces cerevisiae. *Eukaryot Cell* **3(6)**, 1464–1475.

Cunningham, D. B., Segretain, D., Arnaud, D., Rogner, U. C., and Avner, P. (1998). The mouse Tsx gene is expressed in Sertoli cells of the adult testis and transiently in premeiotic germ cells during puberty. *Dev Biol* **204(2)**, 345–360.

Daubas, P., Tajbakhsh, S., Hadchouel, J., Primig, M., and Buckingham, M. (2000). Myf5 is a novel early axonal marker in the mouse brain and is subjected to post-transcriptional regulation in neurons. *Development* **127(2)**, 319–331.

de Boer, M., Nielsen, P. S., Bebelman, J. P., Heerikhuizen, H., Andersen, H. A., and Planta, R. J. (2000). Stp1p, Stp2p and Abf1p are involved in regulation of expression of the amino acid transporter gene BAP3 of Saccharomyces cerevisiae. *Nucleic Acids Res* **28(4)**, 974–981.

de Winde, J. H. and Grivell, L. A. (1992). Global regulation of mitochondrial biogenesis in Saccharomyces cerevisiae: ABF1 and CPF1 play opposite roles in regulating expression of the QCR8 gene, which encodes subunit VIII of the mitochon-

drial ubiquinol-cytochrome c oxidoreductase. *Mol Cell Biol* **12(6)**, 2872–2883.

Della Seta, F., Ciafré, S. A., Marck, C., Santoro, B., Presutti, C., Sentenac, A., and Bozzoni, I. (1990). The ABF1 factor is the transcriptional activator of the L2 ribosomal protein genes in Saccharomyces cerevisiae. *Mol Cell Biol* **10(5)**, 2437–2441.

Della Seta, F., Treich, I., Buhler, J. M., and Sentenac, A. (1990). ABF1 binding sites in yeast RNA polymerase genes. *J Biol Chem* **265(25)**, 15168–15175.

DeMarini, D. J., Adams, A. E., Fares, H., De Virgilio, C., Valle, G., Chuang, J. S., and Pringle, J. R. (1997). A septin-based hierarchy of proteins required for localized deposition of chitin in the Saccharomyces cerevisiae cell wall. *J Cell Biol* **139(1)**, 75–93.

DeRisi, J. L., Iyer, V. R., and Brown, P. O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278(5338)**, 680–686.

Desany, B. A., Alcasabas, A. A., Bachant, J. B., and Elledge, S. J. (1998). Recovery from DNA replicational stress is the essential function of the S-phase checkpoint pathway. *Genes Dev* **12(18)**, 2956–2970.

Detweiler, C. S. and Li, J. J. (1997). Cdc6p establishes and maintains a state of replication competence during G1 phase. *J Cell Sci* **110 (Pt 6)**, 753–763.

Deutschbauer, A. M. and Davis, R. W. (2005). Quantitative trait loci mapped to single-nucleotide resolution in yeast. *Nat Genet* **37(12)**, 1333–1340.

Deutschbauer, A. M., Williams, R. M., Chu, A. M., and Davis, R. W. (2002). Parallel phenotypic analysis of sporulation and postgermination growth in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **99(24)**, 15530–15535.

Diffley, J. F., Cocker, J. H., Dowell, S. J., and Rowley, A. (1994). Two steps in the assembly of complexes at yeast replication origins in vivo. *Cell* **78(2)**, 303–316.

Diffley, J. F. and Stillman, B. (1988). Purification of a yeast protein that binds to origins of DNA replication and a transcriptional silencer. *Proc Natl Acad Sci U S A* **85(7)**, 2120–2124.

Diffley, J. F. and Stillman, B. (1989). Similarity between the transcriptional silencer binding proteins ABF1 and RAP1. *Science* **246(4933)**, 1034–1038.

Dorsman, J. C., van Heeswijk, W. C., and Grivell, L. A. (1990). Yeast general transcription factor GFI: sequence requirements for binding to DNA and evolutionary conservation. *Nucleic Acids Res* **18(9)**, 2769–2776.

Eddy, E. M. (2002). Male germ cell gene expression. *Recent Prog Horm Res* **57**, 103–128.

Edelmann, W., Cohen, P. E., Kneitz, B., Winand, N., Lia, M., Heyer, J., Kolodner, R., Pollard, J. W., and Kucherlapati, R. (1999). Mammalian MutS homologue 5 is required for chromosome pairing in meiosis. *Nat Genet* **21(1)**, 123–127.

Einerhand, A. W., Kos, W., Smart, W. C., Kal, A. J., Tabak, H. F., and Cooper, T. G. (1995). The upstream region of the FOX3 gene encoding peroxisomal 3-oxoacyl-coenzyme A thiolase in Saccharomyces cerevisiae contains ABF1- and replication protein A-binding sites that participate in its regulation by glucose repression. *Mol Cell Biol* **15(6)**, 3405–3414.

Eisenberg, S., Civalier, C., and Tye, B. K. (1988). Specific interaction between a Saccharomyces cerevisiae protein and a DNA element associated with certain autonomously replicating sequences. *Proc Natl Acad Sci U S A* **85(3)**, 743–746.

Ellis, P. J., Furlong, R. A., Wilson, A., Morris, S., Carter, D., Oliver, G., Print, C., Burgoyne, P. S., Loveland, K. L., and Affara, N. A. (2004). Modulation of the mouse testis transcriptome during postnatal development and in selected models of male infertility. *Mol Hum Reprod* **10(4)**, 271–281.

Enyenihi, A. H. and Saunders, W. S. (2003). Large-scale functional genomic analysis of sporulation and meiosis in Saccharomyces cerevisiae. *Genetics* **163(1)**, 47–54.

Erb, I. and van Nimwegen, E. (2006). Statistical Features of Yeast's Transcriptional Regulatory Code. *submitted*

Feng, L., Wang, B., Driscoll, B., and Jong, A. (2000). Identification and characterization of Saccharomyces cerevisiae Cdc6 DNA-binding properties. *Mol Biol Cell* **11(5)**, 1673–1685.

Feng, L., Wang, B., and Jong, A. (1998). Saccharomyces cerevisiae Cdc6 stimulates Abf1 DNA binding activity. *J Biol Chem* **273(3)**, 1298–1302.

Flescher, E. G., Madden, K., and Snyder, M. (1993). Components required for cytokinesis are important for bud site selection in yeast. *J Cell Biol* **122(2)**, 373–386.

Ford, S. K. and Pringle, J. R. (1991). Cellular morphogenesis in the Saccharomyces cerevisiae cell cycle: localization of the CDC11 gene product and the timing of events at the budding site. *Dev Genet* **12(4)**, 281–292.

Fourel, G., Miyake, T., Defossez, P. A., Li, R., and Gilson, E. (2002). General regulatory factors (GRFs) as genome partitioners. *J Biol Chem* **277(44)**, 41736–41743.

Francesconi, S. C. and Eisenberg, S. (1989). Purification and characterization of OBF1: a Saccharomyces cerevisiae protein that binds to autonomously replicating sequences. *Mol Cell Biol* **9(7)**, 2906–2913.

Freeman, L., Aragon-Alcaide, L., and Strunnikov, A. (2000). The condensin complex governs chromosome condensation and mitotic transmission of rDNA. *J Cell Biol* **149(4)**, 811–824.

Gailus-Durner, V., Xie, J., Chintamaneni, C., and Vershon, A. K. (1996). Participation of the yeast activator Abf1 in meiosis-specific expression of the HOP1 gene. *Mol Cell Biol* **16(6)**, 2777–2786.

Gavin, A. C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L. J., Bastuck, S., Dümpelfeld, B., Edelmann, A., Heurtier, M. A., Hoffman, V., Hoefert, C., Klein, K., Hudak, M., Michon, A. M., Schelder, M., Schirle, M., Remor, M., Rudi, T., Hooper, S., Bauer, A., Bouwmeester, T., Casari, G., Drewes, G., Neubauer, G., Rick, J. M., Kuster, B., Bork, P., Russell, R. B., and Superti-Furga, G. (2006). Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440(7084)**, 631–636.

Goepfert, T. M., Adigun, Y. E., Zhong, L., Gay, J., Medina, D., and Brinkley, W. R. (2002). Centrosome amplification and overexpression of aurora A are early events in rat mammary carcinogenesis. *Cancer Res* **62(14)**, 4115–4122.

Goldmark, J. P., Fazzio, T. G., Estep, P. W., Church, G. M., and Tsukiyama, T. (2000). The Isw2 chromatin remodeling complex represses early meiotic genes upon recruitment by Ume6p. *Cell* **103(3)**, 423–433.

Gonçalves, P. M., Maurer, K., Mager, W. H., and Planta, R. J. (1992). Kluyveromyces contains a functional ABF1-homologue. *Nucleic Acids Res* **20(9)**, 2211–2215.

Grigull, J., Mnaimneh, S., Pootoolal, J., Robinson, M. D., and Hughes, T. R. (2004). Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors. *Mol Cell Biol* **24(12)**, 5534–5547.

Griswold, M. D. (1998). The central role of Sertoli cells in spermatogenesis. *Semin Cell Dev Biol* **9(4)**, 411–416.

Gromoll, J., Wessels, J., Rosiepen, G., Brinkworth, M. H., and Weinbauer, G. F. (1997). Expression of mitotic cyclin B1 is not confined to proliferating cells in the rat testis. *Biol Reprod* **57(6)**, 1312–1319.

Guillaume, E., Evrard, B., Com, E., Moertz, E., Jégou, B., and Pineau, C. (2001). Proteome analysis of rat spermatogonia: reinvestigation of stathmin spatio-temporal expression within the testis. *Mol Reprod Dev* **60(4)**, 439–445.

Guo, T. B., Chan, K. C., Hakovirta, H., Xiao, Y., Toppari, J., Mitchell, A. P., and Salameh, W. A. (2003). Evidence for a role of glycogen synthase kinase-3 beta in rodent spermatogenesis. *J Androl* **24(3)**, 332–342.

Haarer, B. K. and Pringle, J. R. (1987). Immunofluorescence localization of the Saccharomyces cerevisiae CDC12 gene product to the vicinity of the 10-nm filaments in the mother-bud neck. *Mol Cell Biol* **7(10)**, 3678–3687.

Halfter, H., Kavety, B., Vandekerckhove, J., Kiefer, F., and Gallwitz, D. (1989). Sequence, expression and mutational analysis of BAF1, a transcriptional activator and ARS1-binding protein of the yeast Saccharomyces cerevisiae. *EMBO J* **8(13)**, 4265–4272.

Halfter, H., Müller, U., Winnacker, E. L., and Gallwitz, D. (1989). Isolation and DNA-binding characteristics of a protein involved in transcription activation of two divergently transcribed, essential yeast genes.

EMBO J **8(10)**, 3029–3037.

Harbison, C. T., Gordon, D. B., Lee, T. I., Rinaldi, N. J., Macisaac, K. D., Danford, T. W., Hannett, N. M., Tagne, J. B., Reynolds, D. B., Yoo, J., Jennings, E. G., Zeitlinger, J., Pokholok, D. K., Kellis, M., Rolfe, P. A., Takusagawa, K. T., Lander, E. S., Gifford, D. K., Fraenkel, E., and Young, R. A. (2004). Transcriptional regulatory code of a eukaryotic genome. *Nature* **431(7004)**, 99–104.

Hartwell, L. H. (1971). Genetic control of the cell division cycle in yeast. IV. Genes controlling bud emergence and cytokinesis. *Exp Cell Res* **69(2)**, 265–276.

Hepworth, S. R., Ebisuzaki, L. K., and Segall, J. (1995). A 15-base-pair element activates the SPS4 gene midway through sporulation in Saccharomyces cerevisiae. *Mol Cell Biol* **15(7)**, 3934–3944.

Hepworth, S. R., Friesen, H., and Segall, J. (1998). NDT80 and the meiotic recombination checkpoint regulate expression of middle sporulation-specific genes in Saccharomyces cerevisiae. *Mol Cell Biol* **18(10)**, 5750–5761.

Hieronymus, H. and Silver, P. A. (2003). Genome-wide analysis of RNA-protein interactions illustrates specificity of the mRNA export machinery. *Nat Genet* **33(2)**, 155–161.

Honigberg, S. M., Conicella, C., and Esposito, R. E. (1992). Commitment to meiosis in Saccharomyces cerevisiae: involvement of the SPO14 gene. *Genetics* **130(4)**, 703–716.

Honigberg, S. M. and Esposito, R. E. (1994). Reversal of cell determination in yeast meiosis: postcommitment arrest allows return to mitotic growth. *Proc Natl Acad Sci U S A* **91(14)**, 6559–6563.

Horak, C. E. and Snyder, M. (2002). ChIP-chip: a genomic approach for identifying transcription factor binding sites. *Methods Enzymol* **350**, 469–483.

Hsu, J. M., Huang, J., Meluh, P. B., and Laurent, B. C. (2003). The yeast RSC chromatin-remodeling complex is required for kinetochore function in chromosome segregation. *Mol Cell Biol* **23(9)**, 3202–3215.

Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B., and Speed, T. P. (2003). Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 31(4)

Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., and Speed, T. P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4(2)**, 249–264.

Ishibashi, K., Kuwahara, M., Kageyama, Y., Tohsaka, A., Marumo, F., and Sasaki, S. (1997). Cloning and functional expression of a second new aquaporin abundantly expressed in testis. *Biochem Biophys Res Commun* **237(3)**, 714–718.

Ishibashi, K., Morinaga, T., Kuwahara, M., Sasaki, S., and Imai, M. (2002). Cloning and identification of a new member of water channel (AQP10) as an aquaglyceroporin. *Biochim Biophys Acta* **1576(3)**, 335–340.

Ishida, C., Aranda, C., Valenzuela, L., Riego, L., Deluna, A., Recillas-Targa, F., Filetici, P., López-Revilla, R., and González, A. (2006). The UGA3-GLT1 intergenic region constitutes a promoter whose bidirectional nature is determined by chromatin organization in Saccharomyces cerevisiae. *Mol Microbiol* **59(6)**, 1790–1806.

Jansma, D. B., Archambault, J., Mostachfi, O., and Friesen, J. D. (1996). Similar upstream regulatory elements of genes that encode the two largest subunits of RNA polymerase II in Saccharomyces cerevisiae. *Nucleic Acids Res* **24(22)**, 4543–4551.

Jégou, B., Pineau, C., and Dupaix, A. (1999). Male Reproductive Function. *Boston, MA: Kluwer Academic Publishers*

Jónsson, Z. O., Dhar, S. K., Narlikar, G. J., Auty, R., Wagle, N., Pellman, D., Pratt, R. E., Kingston, R., and Dutta, A. (2001). Rvb1p and Rvb2p are essential components of a chromatin remodeling complex that regulates transcription of over 5 *J Biol Chem* **276(19)**, 16279–16288.

Kadosh, D. and Struhl, K. (1998). Targeted recruitment of the Sin3-Rpd3 histone deacetylase complex generates a highly localized domain of repressed chromatin in vivo. *Mol Cell Biol* **18(9)**, 5121–5127.

Kanatsu-Shinohara, M., Toyokuni, S., and Shinohara, T. (2004). CD9 is a surface marker on mouse and rat male germline stem cells. *Biol Reprod* **70(1)**, 70–75.

Kassir, Y., Adir, N., Boger-Nadjar, E., Raviv,

N. G., Rubin-Bejerano, I., Sagee, S., and Shenhar, G. (2003). Transcriptional regulation of meiosis in budding yeast. *Int Rev Cytol* **224**, 111–171.

Kaufman, L. and Rousseeuw, P. (1999). Finding Groups in Data: An Introduction to Cluster Analysis. *Brussels, Belgium: John Wiley & Sons*

Keeney, S., Giroux, C. N., and Kleckner, N. (1997). Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* **88(3)**, 375–384.

Kellis, M., Patterson, N., Endrizzi, M., Birren, B., and Lander, E. S. (2003). Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423(6937)**, 241–254.

Kim, H. B., Haarer, B. K., and Pringle, J. R. (1991). Cellular morphogenesis in the Saccharomyces cerevisiae cell cycle: localization of the CDC3 gene product and the timing of events at the budding site. *J Cell Biol* **112(4)**, 535–544.

Kneitz, B., Cohen, P. E., Avdievich, E., Zhu, L., Kane, M. F., Hou, H., Kolodner, R. D., Kucherlapati, R., Pollard, J. W., and Edelmann, W. (2000). MutS homolog 4 localization to meiotic chromosomes is required for chromosome pairing during meiosis in male and female mice. *Genes Dev* **14(9)**, 1085–1097.

Kovari, L. Z. and Cooper, T. G. (1991). Participation of ABF-1 protein in expression of the Saccharomyces cerevisiae CAR1 gene. *J Bacteriol* **173(20)**, 6332–6338.

Kratzer, S. and Schüller, H. J. (1997). Transcriptional control of the yeast acetyl-CoA synthetase gene, ACS1, by the positive regulators CAT8 and ADR1 and the pleiotropic repressor UME6. *Mol Microbiol* **26(4)**, 631–641.

Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A. P., Punna, T., Peregrín-Alvarez, J. M., Shales, M., Zhang, X., Davey, M., Robinson, M. D., Paccanaro, A., Bray, J. E., Sheung, A., Beattie, B., Richards, D. P., Canadien, V., Lalev, A., Mena, F., Wong, P., Starostine, A., Canete, M. M., Vlasblom, J., Wu, S., Orsi, C., Collins, S. R., Chandran, S., Haw, R., Rilstone, J. J., Gandi, K., Thompson, N. J., Musso, G., St Onge, P., Ghanny, S., Lam, M. H., Butland, G., Altaf-Ul, A. M., Kanaya, S., Shilatifard, A., O'Shea, E., Weissman, J. S., Ingles, C. J., Hughes, T. R., Parkinson, J., Gerstein, M., Wodak, S. J., Emili, A., and Greenblatt, J. F. (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature* **440(7084)**, 637–643.

Künzler, M., Springer, C., and Braus, G. H. (1995). Activation and repression of the yeast ARO3 gene by global transcription factors. *Mol Microbiol* **15(1)**, 167–178.

Kupiec, M., Byers, B., Esposito, R. E., and Mitchell, A. P. (1997). Meiosis and sporulation in Saccharomyces cerevisiae. *Cold Spring Harbor, NY: Cold Spring Harbor Laboratory* **3**, 889–1036

Lamb, J. R., Michaud, W. A., Sikorski, R. S., and Hieter, P. A. (1994). Cdc16p, Cdc23p and Cdc27p form a complex essential for mitosis. *EMBO J* **13(18)**, 4321–4328.

Lascaris, R. F., Groot, E., Hoen, P. B., Mager, W. H., and Planta, R. J. (2000). Different roles for abf1p and a T-rich promoter element in nucleosome organization of the yeast RPS28A gene. *Nucleic Acids Res* **28(6)**, 1390–1396.

Le Goffic, R., Mouchel, T., Aubry, F., Patard, J. J., Ruffault, A., Jégou, B., and Samson, M. (2002). Production of the chemokines monocyte chemotactic protein-1, regulated on activation normal T cell expressed and secreted protein, growth-related oncogene, and interferon-gamma-inducible protein-10 is induced by the Sendai virus in human and rat testicular cells. *Endocrinology* **143(4)**, 1434–1440.

Le Naour, F., Rubinstein, E., Jasmin, C., Prenant, M., and Boucheix, C. (2000). Severely reduced female fertility in CD9-deficient mice. *Science* **287(5451)**, 319–321.

Lee, T. I., Rinaldi, N. J., Robert, F., Odom, D. T., Bar-Joseph, Z., Gerber, G. K., Hannett, N. M., Harbison, C. T., Thompson, C. M., Simon, I., Zeitlinger, J., Jennings, E. G., Murray, H. L., Gordon, D. B., Ren, B., Wyrick, J. J., Tagne, J. B., Volkert, T. L., Fraenkel, E., Gifford, D. K., and Young, R. A. (2002). Transcriptional reg-

ulatory networks in Saccharomyces cerevisiae. *Science* **298(5594)**, 799–804.

Lee, T. I. and Young, R. A. (2000). Transcription of eukaryotic protein-coding genes. *Annu Rev Genet* **34**, 77–137.

Lieb, J. D., Liu, X., Botstein, D., and Brown, P. O. (2001). Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nat Genet* **28(4)**, 327–334.

Lipkin, S. M., Moens, P. B., Wang, V., Lenzi, M., Shanmugarajah, D., Gilgeous, A., Thomas, J., Cheng, J., Touchman, J. W., Green, E. D., Schwartzberg, P., Collins, F. S., and Cohen, P. E. (2002). Meiotic arrest and aneuploidy in MLH3-deficient mice. *Nat Genet* **31(4)**, 385–390.

Liu, D., Matzuk, M. M., Sung, W. K., Guo, Q., Wang, P., and Wolgemuth, D. J. (1998). Cyclin A1 is required for meiosis in the male mouse. *Nat Genet* **20(4)**, 377–380.

Livera, G., Rouiller-Fabre, V., Pairault, C., Levacher, C., and Habert, R. (2002). Regulation and perturbation of testicular functions by vitamin A. *Reproduction* **124(2)**, 173–180.

Loch, C. M., Mosammaparast, N., Miyake, T., Pemberton, L. F., and Li, R. (2004). Functional and physical interactions between autonomously replicating sequence-binding factor 1 and the nuclear transport machinery. *Traffic* **5(12)**, 925–935.

Loo, S., Laurenson, P., Foss, M., Dillin, A., and Rine, J. (1995). Roles of ABF1, NPL3, and YCL54 in silencing in Saccharomyces cerevisiae. *Genetics* **141(3)**, 889–902.

Lu, J., Kobayashi, R., and Brill, S. J. (1996). Characterization of a high mobility group 1/2 homolog in yeast. *J Biol Chem* **271(52)**, 33678–33685.

Lue, N. F., Buchman, A. R., and Kornberg, R. D. (1989). Activation of yeast RNA polymerase II transcription by a thymidine-rich upstream element in vitro. *Proc Natl Acad Sci U S A* **86(2)**, 486–490.

Lui, W. Y., Lee, W. M., and Cheng, C. Y. (2003). Rho GTPases and spermatogenesis. *Biochim Biophys Acta* **1593(2-3)**, 121–129.

Luo, G., Yao, M. S., Bender, C. F., Mills, M., Bladl, A. R., Bradley, A., and Petrini, J. H. (1999). Disruption of mRad50 causes embryonic stem cell lethality, abnormal embryonic development, and sensitivity to ionizing radiation. *Proc Natl Acad Sci U S A* **96(13)**, 7376–7381.

Malathi, K., Xiao, Y., and Mitchell, A. P. (1997). Interaction of yeast repressor-activator protein Ume6p with glycogen synthase kinase 3 homolog Rim11p. *Mol Cell Biol* **17(12)**, 7230–7236.

Mannan, A. U., Nayernia, K., Mueller, C., Burfeind, P., Adham, I. M., and Engel, W. (2003). Male mice lacking the Theg (testicular haploid expressed gene) protein undergo normal spermatogenesis and are fertile. *Biol Reprod* **69(3)**, 788–796.

Marahrens, Y. and Stillman, B. (1992). A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* **255(5046)**, 817–823.

Martens, J. A. and Brandl, C. J. (1994). GCN4p activation of the yeast TRP3 gene is enhanced by ABF1p and uses a suboptimal TATA element. *J Biol Chem* **269(22)**, 15661–15667.

McCord, R., Pierce, M., Xie, J., Wonkatal, S., Mickel, C., and Vershon, A. K. (2003). Rfm1, a novel tethering factor required to recruit the Hst1 histone deacetylase for repression of middle sporulation genes. *Mol Cell Biol* **23(6)**, 2009–2016.

McLean, D. J., Friel, P. J., Pouchnik, D., and Griswold, M. D. (2002). Oligonucleotide microarray analysis of gene expression in follicle-stimulating hormone-treated rat Sertoli cells. *Mol Endocrinol* **16(12)**, 2780–2792.

Miki, K., Willis, W. D., Brown, P. R., Goulding, E. H., Fulcher, K. D., and Eddy, E. M. (2002). Targeted disruption of the Akap4 gene causes defects in sperm flagellum and motility. *Dev Biol* **248(2)**, 331–342.

Miyake, T., Loch, C. M., and Li, R. (2002). Identification of a multifunctional domain in autonomously replicating sequence-binding factor 1 required for transcriptional activation, DNA replication, and gene silencing. *Mol Cell Biol* **22(2)**, 505–516.

Miyake, T., Reese, J., Loch, C. M., Auble, D. T., and Li, R. (2004). Genome-wide analysis of ARS (autonomously replicating sequence) binding factor 1 (Abf1p)-mediated transcriptional regulation in Saccharomyces cerevisiae. *J Biol Chem* **279(33)**, 34865–

34872.

Mukherjee, S., Berger, M. F., Jona, G., Wang, X. S., Muzzey, D., Snyder, M., Young, R. A., and Bulyk, M. L. (2004). Rapid analysis of the DNA-binding specificities of transcription factors with DNA microarrays. *Nat Genet* **36(12)**, 1331–1339.

Mulder, W., Scholten, I. H., and Grivell, L. A. (1995). Distinct transcriptional regulation of a gene coding for a mitochondrial protein in the yeasts Saccharomyces cerevisiae and Kluyveromyces lactis despite similar promoter structures. *Mol Microbiol* **17(5)**, 813–824.

Muldrow, T. A., Campbell, A. M., Weil, P. A., and Auble, D. T. (1999). MOT1 can activate basal transcription in vitro by regulating the distribution of TATA binding protein between promoter and nonpromoter sites. *Mol Cell Biol* **19(4)**, 2835–2845.

Offenberg, H. H., Schalk, J. A., Meuwissen, R. L., van Aalderen, M., Kester, H. A., Dietrich, A. J., and Heyting, C. (1998). SCP2: a major protein component of the axial elements of synaptonemal complexes of the rat. *Nucleic Acids Res* **26(11)**, 2572–2579.

O'Shaughnessy, P. J., Willerton, L., and Baker, P. J. (2002). Changes in Leydig cell gene expression during development in the mouse. *Biol Reprod* **66(4)**, 966–975.

Ostermeier, G. C., Dix, D. J., Miller, D., Khatri, P., and Krawetz, S. A. (2002). Spermatozoal RNA profiles of normal fertile men. *Lancet* **360(9335)**, 772–777.

Ouspenski, I. I., Cabello, O. A., and Brinkley, B. R. (2000). Chromosome condensation factor Brn1p is required for chromatid separation in mitosis. *Mol Biol Cell* **11(4)**, 1305–1313.

Ozsarac, N., Bhattacharyya, M., Dawes, I. W., and Clancy, M. J. (1995). The SPR3 gene encodes a sporulation-specific homologue of the yeast CDC3/10/11/12 family of bud neck microfilaments and is regulated by ABFI. *Gene* **164(1)**, 157–162.

Ozsarac, N., Straffon, M. J., Dalton, H. E., and Dawes, I. W. (1997). Regulation of gene expression during meiosis in Saccharomyces cerevisiae: SPR3 is controlled by both ABFI and a new sporulation control element. *Mol Cell Biol* **17(3)**, 1152–1159.

Pang, A. L., Taylor, H. C., Johnson, W., Alexander, S., Chen, Y., Su, Y. A., Li, X., Ravindranath, N., Dym, M., Rennert, O. M., and Chan, W. Y. (2003). Identification of differentially expressed genes in mouse spermatogenesis. *J Androl* **24(6)**, 899–911.

Payne, A. H. (1990). Hormonal regulation of cytochrome P450 enzymes, cholesterol side-chain cleavage and 17 alpha-hydroxylase/C17-20 lyase in Leydig cells. *Biol Reprod* **42(3)**, 399–404.

Persengiev, S. P., Robert, S., and Kilpatrick, D. L. (1996). Transcription of the TATA binding protein gene is highly up-regulated during spermatogenesis. *Mol Endocrinol* **10(6)**, 742–747.

Petronczki, M., Siomos, M. F., and Nasmyth, K. (2003). Un ménage à quatre: the molecular biology of chromosome segregation in meiosis. *Cell* **112(4)**, 423–440.

Piatti, S., Lengauer, C., and Nasmyth, K. (1995). Cdc6 is an unstable protein whose de novo synthesis in G1 is important for the onset of S phase and for preventing a 'reductional' anaphase in the budding yeast Saccharomyces cerevisiae. *EMBO J* **14(15)**, 3788–3799.

Pierce, M., Wagner, M., Xie, J., Gailus-Durner, V., Six, J., Vershon, A. K., and Winter, E. (1998). Transcriptional regulation of the SMK1 mitogen-activated protein kinase gene during meiotic development in Saccharomyces cerevisiae. *Mol Cell Biol* **18(10)**, 5970–5980.

Pijnappel, W. W., Schaft, D., Roguev, A., Shevchenko, A., Tekotte, H., Wilm, M., Rigaut, G., Séraphin, B., Aasland, R., and Stewart, A. F. (2001). The S. cerevisiae SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. *Genes Dev* **15(22)**, 2991–3004.

Pineau, C., Syed, V., Bardin, C. W., Jégou, B., and Cheng, C. Y. (1993). Germ cell-conditioned medium contains multiple factors that modulate the secretion of testins, clusterin, and transferrin by Sertoli cells. *J Androl* **14(2)**, 87–98.

Pineau, C., Syed, V., Bardin, C. W., Jégou, B., and Cheng, C. Y. (1993). Identification and partial purification of a germ cell factor

that stimulates transferrin secretion by Sertoli cells. *Recent Prog Horm Res* **48**, 539–542.

Planta, R. J. (1997). Regulation of ribosome synthesis in yeast. *Yeast* **13(16)**, 1505–1518.

Planta, R. J., Gonçalves, P. M., and Mager, W. H. (1995). Global regulators of ribosome biosynthesis in yeast. *Biochem Cell Biol* **73(11-12)**, 825–834.

Portman, D. S., O'Connor, J. P., and Dreyfuss, G. (1997). YRA1, an essential Saccharomyces cerevisiae gene, encodes a novel nuclear protein with RNA annealing activity. *RNA* **3(5)**, 527–537.

Primig, M., Wiederkehr, C., Basavaraj, R., Sarrauste de Menthière, C., Hermida, L., Koch, R., Schlecht, U., Dickinson, H. G., Fellous, M., Grootegoed, J. A., Hawley, R. S., Jégou, B., Maro, B., Nicolas, A., Orr-Weaver, T., Schedl, T., Villeneuve, A., Wolgemuth, D. J., Yamamoto, M., Zickler, D., Lamb, N., and Esposito, R. E. (2003). GermOnline, a new cross-species community annotation database on germ-line development and gametogenesis. *Nat Genet* **35(4)**, 291–292.

Primig, M., Williams, R. M., Winzeler, E. A., Tevzadze, G. G., Conway, A. R., Hwang, S. Y., Davis, R. W., and Esposito, R. E. (2000). The core meiotic transcriptome in budding yeasts. *Nat Genet* **26(4)**, 415–423.

Prinz, S., Klein, F., Auer, H., Schweizer, D., and Primig, M. (1995). A DNA binding factor (UBF) interacts with a positive regulatory element in the promoters of genes expressed during meiosis and vegetative growth in yeast. *Nucleic Acids Res* **23(17)**, 3449–3456.

Rabitsch, K. P., Tóth, A., Gálová, M., Schleiffer, A., Schaffner, G., Aigner, E., Rupp, C., Penkner, A. M., Moreno-Borchart, A. C., Primig, M., Esposito, R. E., Klein, F., Knop, M., and Nasmyth, K. (2001). A screen for genes required for meiosis and spore formation based on whole-genome expression. *Curr Biol* **11(13)**, 1001–1009.

Reed, S. H., Akiyama, M., Stillman, B., and Friedberg, E. C. (1999). Yeast autonomously replicating sequence binding factor is involved in nucleotide excision repair. *Genes Dev* **13(23)**, 3052–3058.

Reguly, T., Breitkreutz, A., Boucher, L., Breitkreutz, B. J., Hon, G. C., Myers, C. L., Parsons, A., Friesen, H., Oughtred, R., Tong, A., Stark, C., Ho, Y., Botstein, D., Andrews, B., Boone, C., Troyanskya, O. G., Ideker, T., Dolinski, K., Batada, N. N., and Tyers, M. (2006). Comprehensive curation and analysis of global interaction networks in Saccharomyces cerevisiae. *J Biol* **5(4)**, 11–11.

Reid, J. L., Iyer, V. R., Brown, P. O., and Struhl, K. (2000). Coordinate regulation of yeast ribosomal protein genes is associated with targeted recruitment of Esa1 histone acetylase. *Mol Cell* **6(6)**, 1297–1307.

Reinke, V., Smith, H. E., Nance, J., Wang, J., Van Doren, C., Begley, R., Jones, S. J., Davis, E. B., Scherer, S., Ward, S., and Kim, S. K. (2000). A global profile of germline gene expression in C. elegans. *Mol Cell* **6(3)**, 605–616.

Rhode, P. R., Elsasser, S., and Campbell, J. L. (1992). Role of multifunctional autonomously replicating sequence binding factor 1 in the initiation of DNA replication and transcriptional control in Saccharomyces cerevisiae. *Mol Cell Biol* **12(3)**, 1064–1077.

Roeder, G. S., Beard, C., Smith, M., and Keranen, S. (1985). Isolation and characterization of the SPT2 gene, a negative regulator of Ty-controlled yeast gene expression. *Mol Cell Biol* **5(7)**, 1543–1553.

Romanienko, P. J. and Camerini-Otero, R. D. (2000). The mouse Spo11 gene is required for meiotic chromosome synapsis. *Mol Cell* **6(5)**, 975–987.

Rosok, O., Pedeutour, F., Ree, A. H., and Aasheim, H. C. (1999). Identification and characterization of TESK2, a novel member of the LIMK/TESK family of protein kinases, predominantly expressed in testis. *Genomics* **61(1)**, 44–54.

Rudge, S. A., Pettitt, T. R., Zhou, C., Wakelam, M. J., and Engebrecht, J. A. (2001). SPO14 separation-of-function mutations define unique roles for phospholipase D in secretion and cellular differentiation in Saccharomyces cerevisiae. *Genetics* **158(4)**, 1431–1444.

Sadate-Ngatchou, P. I., Pouchnik, D. J., and Griswold, M. D. (2004). Identification of

testosterone-regulated genes in testes of hypogonadal mice using oligonucleotide microarray. *Mol Endocrinol* **18(2)**, 422–433.

Schalk, J. A., Dietrich, A. J., Vink, A. C., Offenberg, H. H., van Aalderen, M., and Heyting, C. (1998). Localization of SCP2 and SCP3 protein molecules within synaptonemal complexes of the rat. *Chromosoma* **107(8)**, 540–548.

Schlecht, U., Demougin, P., Koch, R., Hermida, L., Wiederkehr, C., Descombes, P., Pineau, C., Jégou, B., and Primig, M. (2004). Expression profiling of mammalian male meiosis and gametogenesis identifies novel candidate genes for roles in the regulation of fertility. *Mol Biol Cell* **15(3)**, 1031–1043.

Schlecht, U. and Primig, M. (2003). Mining meiosis and gametogenesis with DNA microarrays. *Reproduction* **125(4)**, 447–456.

Schroeder, S. C. and Weil, P. A. (1998). Biochemical and genetic characterization of the dominant positive element driving transcription ofthe yeast TBP-encoding gene, SPT15. *Nucleic Acids Res* **26(18)**, 4186–4195.

Schroeder, S. C. and Weil, P. A. (1998). Genetic tests of the role of Abf1p in driving transcription of the yeast TATA box bindng protein-encoding gene, SPT15. *J Biol Chem* **273(31)**, 19884–19891.

Schüller, H. J., Schütz, A., Knab, S., Hoffmann, B., and Schweizer, E. (1994). Importance of general regulatory factors Rap1p, Abf1p and Reb1p for the activation of yeast fatty acid synthase genes FAS1 and FAS2. *Eur J Biochem* **225(1)**, 213–222.

Schultz, N., Hamra, F. K., and Garbers, D. L. (2003). A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets. *Proc Natl Acad Sci U S A* **100(21)**, 12201–12206.

Schwickart, M., Havlis, J., Habermann, B., Bogdanova, A., Camasses, A., Oelschlaegel, T., Shevchenko, A., and Zachariae, W. (2004). Swm1/Apc13 is an evolutionarily conserved subunit of the anaphase-promoting complex stabilizing the association of Cdc16 and Cdc27. *Mol Cell Biol* **24(8)**, 3562–3576.

Sha, J., Zhou, Z., Li, J., Yin, L., Yang, H., Hu, G., Luo, M., Chan, H. C., Zhou, K., and (2002). Identification of testis development and spermatogenesis-related genes in human and mouse testes using cDNA arrays. *Mol Hum Reprod* **8(6)**, 511–517.

Shamsadin, R., Adham, I. M., Nayernia, K., Heinlein, U. A., Oberwinkler, H., and Engel, W. (1999). Male mice deficient for germ-cell cyritestin are infertile. *Biol Reprod* **61(6)**, 1445–1451.

Sharpe, R. M., McKinnell, C., Kivlin, C., and Fisher, J. S. (2003). Proliferation and functional maturation of Sertoli cells, and their relevance to disorders of testis function in adulthood. *Reproduction* **125(6)**, 769–784.

Shen, W. C., Bhaumik, S. R., Causton, H. C., Simon, I., Zhu, X., Jennings, E. G., Wang, T. H., Young, R. A., and Green, M. R. (2003). Systematic analysis of essential yeast TAFs in genome-wide transcription and preinitiation complex assembly. *EMBO J* **22(13)**, 3395–3402.

Shima, J. E., McLean, D. J., McCarrey, J. R., and Griswold, M. D. (2004). The murine testicular transcriptome: characterizing gene expression in the testis during the progression of spermatogenesis. *Biol Reprod* **71(1)**, 319–330.

Shinkai, Y., Satoh, H., Takeda, N., Fukuda, M., Chiba, E., Kato, T., Kuramochi, T., and Araki, Y. (2002). A testicular germ cell-associated serine-threonine kinase, MAK, is dispensable for sperm formation. *Mol Cell Biol* **22(10)**, 3276–3280.

Sicinski, P., Donaher, J. L., Geng, Y., Parker, S. B., Gardner, H., Park, M. Y., Robker, R. L., Richards, J. S., McGinnis, L. K., Biggers, J. D., Eppig, J. J., Bronson, R. T., Elledge, S. J., and Weinberg, R. A. (1996). Cyclin D2 is an FSH-responsive gene involved in gonadal cell proliferation and oncogenesis. *Nature* **384(6608)**, 470–474.

Siddhartan, R., Siggia, E. D., and van Nimwegen, E. (2005). PhyloGibbs: A Gibbs Sampling Motif Finder That Incorporates Phylogeny. *PLoS Computational Biology* **1(7)**, e67

Silve, S., Rhode, P. R., Coll, B., Campbell, J., and Poyton, R. O. (1992). ABF1 is a phosphoprotein and plays a role in carbon source control of COX6 transcription in Saccharomyces cerevisiae. *Mol Cell Biol* **12(9)**,

4197–4208.

Slater, M. L., Bowers, B., and Cabib, E. (1985). Formation of septum-like structures at locations remote from the budding sites in cytokinesis-defective mutants of Saccharomyces cerevisiae. *J Bacteriol* **162(2)**, 763–767.

Smith, A. V. and Roeder, G. S. (1997). The yeast Red1 protein localizes to the cores of meiotic chromosomes. *J Cell Biol* **136(5)**, 957–967.

Smith, E. R., Eisen, A., Gu, W., Sattah, M., Pannuti, A., Zhou, J., Cook, R. G., Lucchesi, J. C., and Allis, C. D. (1998). ESA1 is a histone acetyltransferase that is essential for growth in yeast. *Proc Natl Acad Sci U S A* **95(7)**, 3561–3565.

Spellman, P. T., Sherlock, G., Zhang, M. Q., Iyer, V. R., Anders, K., Eisen, M. B., Brown, P. O., Botstein, D., and Futcher, B. (1998). Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. *Mol Biol Cell* **9(12)**, 3273–3297.

Springer, C., Krappmann, S., Künzler, M., Zmasek, C., and Braus, G. H. (1997). Regulation of the yeast HIS7 gene by the global transcription factor Abf1p. *Mol Gen Genet* **256(2)**, 136–146.

Srienc, F., Bailey, J. E., and Campbell, J. L. (1985). Effect of ARS1 mutations on chromosome stability in Saccharomyces cerevisiae. *Mol Cell Biol* **5(7)**, 1676–1684.

Steger, K. (2001). Haploid spermatids exhibit translationally repressed mRNAs. *Anat Embryol (Berl)* **203(5)**, 323–334.

Storey, J. (2002). A direct approach to false discovery rates. *J R Stat Soc Ser B* **64**, 479–498

Strässer, K. and Hurt, E. (2000). Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. *EMBO J* **19(3)**, 410–420.

Strich, R., Surosky, R. T., Steber, C., Dubois, E., Messenguy, F., and Esposito, R. E. (1994). UME6 is a key regulator of nitrogen repression and meiotic development. *Genes Dev* **8(7)**, 796–810.

Su, A. I., Cooke, M. P., Ching, K. A., Hakak, Y., Walker, J. R., Wiltshire, T., Orth, A. P., Vega, R. G., Sapinoso, L. M., Moqrich, A., Patapoutian, A., Hampton, G. M., Schultz, P. G., and Hogenesch, J. B. (2002). Large-scale analysis of the human and mouse transcriptomes. *Proc Natl Acad Sci U S A* **99(7)**, 4465–4470.

Sweder, K. S., Rhode, P. R., and Campbell, J. L. (1988). Purification and characterization of proteins that bind to yeast ARSs. *J Biol Chem* **263(33)**, 17270–17277.

Toebosch, A. M., Robertson, D. M., Klaij, I. A., de Jong, F. H., and Grootegoed, J. A. (1989). Effects of FSH and testosterone on highly purified rat Sertoli cells: inhibin alpha-subunit mRNA expression and inhibin secretion are enhanced by FSH but not by testosterone. *J Endocrinol* **122(3)**, 757–762.

Trawick, J. D., Kraut, N., Simon, F. R., and Poyton, R. O. (1992). Regulation of yeast COX6 by the general transcription factor ABF1 and separate HAP2- and heme-responsive elements. *Mol Cell Biol* **12(5)**, 2302–2314.

Ufano, S., del Rey, F., and Vázquez de Aldana, C. R. (2004). Swm1p, a subunit of the APC/cyclosome, is required to maintain cell wall integrity during growth at high temperature in Saccharomyces cerevisiae. *FEMS Microbiol Lett* **234(2)**, 371–378.

Ufano, S., San-Segundo, P., del Rey, F., and Vázquez de Aldana, C. R. (1999). SWM1, a developmentally regulated gene, is required for spore wall assembly in Saccharomyces cerevisiae. *Mol Cell Biol* **19(3)**, 2118–2129.

Umek, R. M., Linskens, M. H., Kowalski, D., and Huberman, J. A. (1989). New beginnings in studies of eukaryotic DNA replication origins. *Biochim Biophys Acta* **1007(1)**, 1–14.

Upton, T., Wiltshire, S., Francesconi, S., and Eisenberg, S. (1995). ABF1 Ser-720 is a predominant phosphorylation site for casein kinase II of Saccharomyces cerevisiae. *J Biol Chem* **270(27)**, 16153–16159.

van Nimwegen, E., Zavolan, M., Rajewsky, N., and Siggia, E. D. (2002). Probabilistic clustering of sequences: inferring new bacterial regulons by comparative genomics. *Proc Natl Acad Sci U S A* **99(11)**, 7323–7328.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normaliza-

tion of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3(7)

Velculescu, V. E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M. A., Bassett, D. E., Hieter, P., Vogelstein, B., and Kinzler, K. W. (1997). Characterization of the yeast transcriptome. *Cell* **88(2)**, 243–251.

Vershon, A. K. and Pierce, M. (2000). Transcriptional regulation of meiosis in yeast. *Curr Opin Cell Biol* **12(3)**, 334–339.

Walker, S. S., Francesconi, S. C., Tye, B. K., and Eisenberg, S. (1989). The OBF1 protein and its DNA-binding site are important for the function of an autonomously replicating sequence in Saccharomyces cerevisiae. *Mol Cell Biol* **9(7)**, 2914–2921.

Washburn, B. K. and Esposito, R. E. (2001). Identification of the Sin3-binding site in Ume6 defines a two-step process for conversion of Ume6 from a transcriptional repressor to an activator in yeast. *Mol Cell Biol* **21(6)**, 2057–2069.

Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S. E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., Bloom, T., Bork, P., Botcherby, M., Bray, N., Brent, M. R., Brown, D. G., Brown, S. D., Bult, C., Burton, J., Butler, J., Campbell, R. D., Carninci, P., Cawley, S., Chiaromonte, F., Chinwalla, A. T., Church, D. M., Clamp, M., Clee, C., Collins, F. S., Cook, L. L., Copley, R. R., Coulson, A., Couronne, O., Cuff, J., Curwen, V., Cutts, T., Daly, M., David, R., Davies, J., Delehaunty, K. D., Deri, J., Dermitzakis, E. T., Dewey, C., Dickens, N. J., Diekhans, M., Dodge, S., Dubchak, I., Dunn, D. M., Eddy, S. R., Elnitski, L., Emes, R. D., Eswara, P., Eyras, E., Felsenfeld, A., Fewell, G. A., Flicek, P., Foley, K., Frankel, W. N., Fulton, L. A., Fulton, R. S., Furey, T. S., Gage, D., Gibbs, R. A., Glusman, G., Gnerre, S., Goldman, N., Goodstadt, L., Grafham, D., Graves, T. A., Green, E. D., Gregory, S., Guigó, R., Guyer, M., Hardison, R. C., Haussler, D., Hayashizaki, Y., Hillier, L. W., Hinrichs, A., Hlavina, W., Holzer, T., Hsu, F., Hua, A., Hubbard, T., Hunt, A., Jackson, I., Jaffe, D. B., Johnson, L. S., Jones, M., Jones, T. A., Joy, A., Kamal, M., Karlsson, E. K., Karolchik, D., Kasprzyk, A., Kawai, J., Keibler, E., Kells, C., Kent, W. J., Kirby, A., Kolbe, D. L., Korf, I., Kucherlapati, R. S., Kulbokas, E. J., Kulp, D., Landers, T., Leger, J. P., Leonard, S., Letunic, I., Levine, R., Li, J., Li, M., Lloyd, C., Lucas, S., Ma, B., Maglott, D. R., Mardis, E. R., Matthews, L., Mauceli, E., Mayer, J. H., McCarthy, M., McCombie, W. R., McLaren, S., McLay, K., McPherson, J. D., Meldrim, J., Meredith, B., Mesirov, J. P., Miller, W., Miner, T. L., Mongin, E., Montgomery, K. T., Morgan, M., Mott, R., Mullikin, J. C., Muzny, D. M., Nash, W. E., Nelson, J. O., Nhan, M. N., Nicol, R., Ning, Z., Nusbaum, C., O'Connor, M. J., Okazaki, Y., Oliver, K., Overton-Larty, E., Pachter, L., Parra, G., Pepin, K. H., Peterson, J., Pevzner, P., Plumb, R., Pohl, C. S., Poliakov, A., Ponce, T. C., Ponting, C. P., Potter, S., Quail, M., Reymond, A., Roe, B. A., Roskin, K. M., Rubin, E. M., Rust, A. G., Santos, R., Sapojnikov, V., Schultz, B., Schultz, J., Schwartz, M. S., Schwartz, S., Scott, C., Seaman, S., Searle, S., Sharpe, T., Sheridan, A., Shownkeen, R., Sims, S., Singer, J. B., Slater, G., Smit, A., Smith, D. R., Spencer, B., Stabenau, A., Stange-Thomann, N., Sugnet, C., Suyama, M., Tesler, G., Thompson, J., Torrents, D., Trevaskis, E., Tromp, J., Ucla, C., Ureta-Vidal, A., Vinson, J. P., Von Niederhausern, A. C., Wade, C. M., Wall, M., Weber, R. J., Weiss, R. B., Wendl, M. C., West, A. P., Wetterstrand, K., Wheeler, R., Whelan, S., Wierzbowski, J., Willey, D., Williams, S., Wilson, R. K., Winter, E., Worley, K. C., Wyman, D., Yang, S., Yang, S. P., Zdobnov, E. M., Zody, M. C., Lander, E. S., and (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* **420(6915)**, 520–562.

Wiederkehr, C., Basavaraj, R., Sarrauste de Menthière, C., Hermida, L., Koch, R., Schlecht, U., Amon, A., Brachat, S., Breitenbach, M., Briza, P., Caburet, S., Cherry, M., Davis, R., Deutschbauer, A., Dickinson, H. G., Dumitrescu, T., Fellous, M., Goldman, A., Grootegoed, J. A., Hawley,

R., Ishii, R., Jégou, B., Kaufman, R. J., Klein, F., Lamb, N., Maro, B., Nasmyth, K., Nicolas, A., Orr-Weaver, T., Philippsen, P., Pineau, C., Rabitsch, K. P., Reinke, V., Roest, H., Saunders, W., Schröder, M., Schedl, T., Siep, M., Villeneuve, A., Wolgemuth, D. J., Yamamoto, M., Zickler, D., Esposito, R. E., and Primig, M. (2004). GermOnline, a cross-species community knowledgebase on germ cell differentiation. *Nucleic Acids Res* **32(Database issue)**, 560–567.

Wiederkehr, C., Basavaraj, R., Sarrauste de Menthière, C., Koch, R., Schlecht, U., Hermida, L., Masdoua, B., Ishii, R., Cassen, V., Yamamoto, M., Lane, C., Cherry, M., Lamb, N., and Primig, M. (2004). Database model and specification of GermOnline Release 2.0, a cross-species community annotation knowledgebase on germ cell differentiation. *Bioinformatics* **20(5)**, 808–811.

Williams, R. M., Primig, M., Washburn, B. K., Winzeler, E. A., Bellis, M., Sarrauste de Menthiere, C., Davis, R. W., and Esposito, R. E. (2002). The Ume6 regulon coordinates metabolic and meiotic gene expression in yeast. *Proc Natl Acad Sci U S A* **99(21)**, 13431–13436.

Wolgemuth, D. J., Laurion, E., and Lele, K. M. (2002). Regulation of the mitotic and meiotic cell cycles in the male germ line. *Recent Prog Horm Res* **57**, 75–101.

Wyrick, J. J., Holstege, F. C., Jennings, E. G., Causton, H. C., Shore, D., Grunstein, M., Lander, E. S., and Young, R. A. (1999). Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast. *Nature* **402(6760)**, 418–421.

Xie, J., Pierce, M., Gailus-Durner, V., Wagner, M., Winter, E., and Vershon, A. K. (1999). Sum1 and Hst1 repress middle sporulation-specific gene expression during mitosis in Saccharomyces cerevisiae. *EMBO J* **18(22)**, 6448–6454.

Xu, L., Ajimura, M., Padmore, R., Klein, C., and Kleckner, N. (1995). NDT80, a meiosis-specific gene required for exit from pachytene in Saccharomyces cerevisiae. *Mol Cell Biol* **15(12)**, 6572–6581.

Yu, S., Owen-Hughes, T., Friedberg, E. C., Waters, R., and Reed, S. H. (2004). The yeast Rad7/Rad16/Abf1 complex generates superhelical torsion in DNA that is required for nucleotide excision repair. *DNA Repair (Amst)* **3(3)**, 277–287.

Yu, Y. E., Zhang, Y., Unni, E., Shirley, C. R., Deng, J. M., Russell, L. D., Weil, M. M., Behringer, R. R., and Meistrich, M. L. (2000). Abnormal spermatogenesis and reduced fertility in transition nuclear protein 1-deficient mice. *Proc Natl Acad Sci U S A* **97(9)**, 4683–4688.

Yukawa, M., Katoh, S., Miyakawa, T., and Tsuchiya, E. (1999). Nps1/Sth1p, a component of an essential chromatin-remodeling complex of Saccharomyces cerevisiae, is required for the maximal expression of early meiotic genes. *Genes Cells* **4(2)**, 99–110.

Zachariae, W. and Nasmyth, K. (1996). TPR proteins required for anaphase progression mediate ubiquitination of mitotic B-type cyclins in yeast. *Mol Biol Cell* **7(5)**, 791–801.

Zachariae, W. and Nasmyth, K. (1999). Whose end is destruction: cell division and the anaphase-promoting complex. *Genes Dev* **13(16)**, 2039–2058.

Zachariae, W., Shin, T. H., Galova, M., Obermaier, B., and Nasmyth, K. (1996). Identification of subunits of the anaphase-promoting complex of Saccharomyces cerevisiae. *Science* **274(5290)**, 1201–1204.

Zenklusen, D., Vinciguerra, P., Strahm, Y., and Stutz, F. (2001). The yeast hnRNP-Like proteins Yra1p and Yra2p participate in mRNA export through interaction with Mex67p. *Mol Cell Biol* **21(13)**, 4219–4232.

Zhao, G. Q. and Garbers, D. L. (2002). Male germ cell specification and differentiation. *Dev Cell* **2(5)**, 537–547.

Zickler, D. and Kleckner, N. (1998). The leptotene-zygotene transition of meiosis. *Annu Rev Genet* **32**, 619–697.

ABF1	autonomously replicating sequence binding factor 1	**MRG**	Mammalian Reproductive Genetics
AD	activation domain	**mRNA**	messenger ribonucleic acid
ANOVA	analysis of variance	**MSE**	middle sporulation element
APC/C	anaphase promoting complex / cyclosome	**mut**	mutant
ARS	autonomously replicating sequence	**NCBI**	National Center for Biotechnology Information
ATP	Adenosine triphosphate	**NER**	nucleotide excision repair
BLAST	basic local alignment search tool	**Okdb**	The Ovarian Kaleidoscope
bp	base pairs	**ORC**	origin recognition complex
cDNA	complementary DNA	**ORF**	open reading frame
CEN	centromer	**PAM**	partitioning around medoids
ChIP	chromatin immunoprecipitation	**PBM**	protein binding microarray
ChIP-CHIP	ChIP followed by microarray hybridization	**PCR**	polymerase chain reaction
cRNA	complementary ribonucleic acid	**PDE**	positive distal element
CS	carboxy-terminal sequence	**pH**	pondus hydrogenii
CTP	cytidine triphosphate	**PLD**	phospholipase D
DAPI	4',6-Diamidino-2-phenylindole	**PM**	perfect match
DBD	DNA binding domain	**PROCSE**	probabilistic clustering of sequences
DNA	deoxyribonucleic acid	**Q-PCR**	quantitative polymerase chain reaction
dNTP	deoxynucleotide triphosphate	**RGD**	Rat Genome Database
dpp	days post-partum	**RMA**	Robust Multi-array Analysis package from Bioconductor
dUTP	deoxyuridine triphosphate	**RNA**	ribonucleic acid
EMSA	electro-mobility shift assay	**rpm**	revolutions per minute
EST	expressed sequence tags	**rRNA**	ribosomal ribonucleic acid
FACS	fluorescence activated cell sorting	**SAGE**	serial analysis of gene expression
GAL	galactose inducible	**SAPE**	streptavidine phycoerythrine conjugate
GAPDH	glyseraldehyde-3-phosphate dehydrogenase	**SC**	synaptonemal complex
GFP	green fluorescence protein	**sd**	standard deviation
GO	GeneOntology	**sec**	second
GRF	general regulatory factor	**SGD**	Saccharomyces Genome Database
HA	hemagglutinin	**SPII**	sporulation medium
HiFi	high fidelity	**SQL**	Structured Query Language
HMG	high mobility group	**TBP**	TATA-binding protein
IF	immunofluorescence	**tRNA**	transfer ribonucleic acid
IP	immunoprecipitation	**Ty**	transposable element
IVT	in vitro transcription	**UAS**	upstream activating sequence
kb	kilobases	**URS1**	upstream repressing sequence 1
kDa	kilodalton	**UTP**	uridine triphosphate
M	molar	**V**	volt
MAS	microarray suite software	**WCE**	whole cell extract
MAT	mating-type locus	**WM**	weight matrix
min	minutes	**wt**	wild-type
ml	mililiter	**YPA**	yeast extract/peptone/acetate
mM	milimolar	**YPD**	yeast extract/peptone/glucose
MM	mismatch	**μg**	microgram
MPR	median percentile ranking	**μl**	microliter

Acknowledgements

First of all I would like to thank Prof. Michael Primig for giving me the opportunity to work in his laboratory, for providing me with a great working environment and for sharing his scientific expertise. Especially, I enjoyed the exciting wet-lab/dry-lab interface in his laboratory. I had a chance to work on very interesting projects and I very much enjoyed the freedom that I had in developing and contributing own ideas. I also greatly appreciated the opportunity to participate in stimulating meetings and courses.

I would like to thank Prof. Alain Nicolas for giving me the opportunity to perform important parts of my experiments in his laboratory at the Institut Curie in Paris. I very much enjoyed my trips to the Curie not only because of the excellent assistance of the experimental work that was supervised by Dr. Valérie Borde, but also because I was always warmly welcomed by Gaël, Hajime, Waka, Norio, Nelly, Cyril and all other lab members. Marie-Françoise was a big help for organizing housing in Paris.

I wish to acknowledge Prof. Peter Philippsen for helpful discussions especially during the last stages of my thesis. Prof. Peter Philippsen and Prof. Alain Nicolas were members of my thesis board.

My warm thanks to all former and current members of the Primig lab for the good discussions, the great support and the many good times: Thomas Aust, Roland Züst, Guenda Bertholt, Philippe Demougin, Marie-Astrid Cartron, Gunnar Wrobel, Reinhold Koch, Christa Niederhauser-Wiederkehr, Frédéric Chalmel. Leandro Hermida is a skilled computer scientist who was always eager to help me with my PERL and PHP scripts with amazing patience.

I am very grateful for the special friendship that I share with Thomas Walther and Mahamadou Faty. I will miss our extended discussions about science, the universe and (most importantly) tango. The great atmosphere in our laboratory was a strong stimulus for the projects and the completion of this thesis.

Thanks to my family and my dearest friends who have accompanied me on my way during all these years. Many thanks to my parents, my brother and grandmother for their love and support. Thanks to Volker, Miriam, Maria, Seraina and Jan. You know how much your friendship means to me.

Ulrich Schlecht

General Information

Contact	Klingelbergstrasse 50-70 Divisions of Biochemistry Swiss Institute of Bioinformatics Biozentrum, University of Basel CH-4056 Basel, Switzerland
	Voice: +41 (0)61 267 2152 *Fax:* +41 (0)61 267 33 98 *E-mail:* ulrich.schlecht@unibas.ch
Date of birth	April 27^{th} 1975
Place of birth	Tokyo (Japan)
Marital status	Single
Nationality	German

Education

Biozentrum, University of Basel, Switzerland

Ph.D., Biochemistry, 2002 - 2006

- Dissertation Topic: "Analysis of the transcriptional program governing meiosis and gametogenesis in yeast and mammals"
- Supervision: Prof. Dr. Michael Primig

University of Stuttgart, Germany

Diploma work, Cell Biology, 2001 - 2002

- Research Topic: "Molecular Cloning, expression and characterization of soluble tumor necrosis factor receptor constructs"
- Supervision: Prof. Dr. Harald Wajant

University of Tokyo, Japan

Student research project, Biology, 2000

- Research Topic: "Reversible surface thiol immobilization of carboxyl group containing haptens to a BIAcore biosensor chip enabling repeated usage of a single sensor surface"
- Supervision: Dr. Yoko Nomura / Dr. Till Bachmann

Honors and Awards

German Academic Exchange Service (DAAD) fellowship, 2000

Publications

Schlecht, U., Demougin, P., Koch, R., Hermida, L., Wiederkehr, C., Descombes, P., Pineau, C., Jégou, B., and Primig, M. (2004). Expression profiling of mammalian male meiosis and gametogenesis identifies novel candidate genes for roles in the regulation of fertility. *Mol Biol Cell* 15(3), 1031-1043.

Wiederkehr, C., Basavaraj, R., Sarrauste de Menthière, C., Hermida, L., Koch, R., Schlecht, U., Amon, A., Brachat, S., Breitenbach, M., Briza, P., Caburet, S., Cherry, M., Davis, R., Deutschbauer, A., Dickinson, H. G., Dumitrescu, T., Fellous, M., Goldman, A., Grootegoed, J. A., Hawley, R., Ishii, R., Jegou, B., Kaufman, R. J., Klein, F., Lamb, N., Maro, B., Nasmyth, K., Nicolas, A., Orr-Weaver, T., Philippsen, P., Pineau, C., Rabitsch, K. P., Reinke, V., Roest, H., Saunders, W., Schroder, M., Schedl, T., Siep, M., Villeneuve, A., Wolgemuth, D. J., Yamamoto, M., Zickler, D., Esposito, R. E., and Primig, M. (2004). GermOnline, a cross-species community knowledgebase on germ cell differentiation. *Nucleic Acids Res* 32(Database issue), 560-567.

Wiederkehr, C., Basavaraj, R., Sarrauste de Menthière, C., Koch, R., Schlecht, U., Hermida, L., Masdoua, B., Ishii, R., Cassen, V., Yamamoto, M., Lane, C., Cherry, M., Lamb, N., and Primig, M. (2004). Database model and specification of GermOnline Release 2.0, a cross-species community annotation knowledgebase on germ cell differentiation. *Bioinformatics* 20(5), 808-811.

Schlecht, U. and Primig, M. (2003). Mining meiosis and gametogenesis with DNA microarrays. *Reproduction* 125(4), 447-456.

Primig, M., Wiederkehr, C., Basavaraj, R., Sarrauste de Menthière, C., Hermida, L., Koch, R., Schlecht, U., Dickinson, H. G., Fellous, M., Grootegoed, J. A., Hawley, R. S., Jégou, B., Maro, B., Nicolas, A., Orr- Weaver, T., Schedl, T., Villeneuve, A., Wolgemuth, D. J., Yamamoto, M., Zickler, D., Lamb, N., and Esposito, R. E. (2003). GermOnline, a new cross-species community annotation database on germ-line development and gametogenesis. *Nat Genet* 35(4), 291-292.

Weingärtner, M., Siegmund, D., Schlecht, U., Fotin-Mleczek, M., Scheurich, P., and Wajant, H. (2002). Endogenous membrane tumor necrosis factor (TNF) is a potent amplifier of TNF receptor 1-mediated apoptosis. *J Biol Chem* 277(38), 34853-34859.

Schlecht, U., Nomura, Y., Bachmann, T., and Karube, I. (2002). Reversible surface thiol immobilization of carboxyl group containing haptens to a BIAcore biosensor chip enabling repeated usage of a single sensor surface. *Bioconjug Chem* 13(2), 188-193.

Manuscripts in preparation

Schlecht, U., Niederhauser-Wiederkehr, C., Erb, I., Demougin, P., Robine, N., Borde, V., van Nimwegen, E., Alain Nicolas and Primig, M.: "Genome-wide Analysis of ABF1-target Genes During Mitotic Growth and Meiotic Development"

Erklärung

Ich erkläre hiermit, dass ich die Dissertation

"Analysis of the transcriptional program governing meiosis and gametogenesis in yeast and mammals"

nur mit der darin angegebenen Hilfe verfasst und bei keiner anderen Fakultät eingereicht habe.

Basel, June 2006

Ulrich Schlecht