Impact of ploidy and cell size on genome expression in fission yeast

Nazia Bibi

This dissertation is submitted for the

Degree Doctor of Philosophy

July 2013

University College London

London, UK

Declaration

I, hereby, confirm that the work presented in this thesis is entirely my own and has not been submitted for a degree, diploma or any other qualification at any other university. Where information has been derived from other sources, I confirm that this has been indicated in this thesis.

Nazia Bibi

1st July 2013

To

My Father and Mother! For all their efforts to give me the best in my life

My dear Husband! For always being very supportive

&

My lovely daughter! For bringing joys to my life

Acknowledgements

Firstly, I would like to thank my supervisor **Jürg Bähler** for his continuous support throughout this project during the last 4 years, his ideas, constructive discussions, his encouragement and understanding at critical moments. I value the fact that he never left me alone in moments of hesitation and always promoted my independence to develop my own ideas and fulfil them. I appreciate the freedom of being able to walk into his office and receive the endless feedback and suggestions on any problem, no matter how big or small. I would particularly like to thank him for proof reading/editing my thesis on a weekend, despite being extremely tired after organising/hosting a conference. Without his continuous support and patience, I could not have completed this thesis. **This has helped a lot - thanks very much!**

Working in the fission yeast functional genomics lab, has been an unforgettable scientific and personal experience for me and I want to thank all the past and present lab members who made it as such and taught me the many different techniques that were implemented throughout my PhD. Working with all of you has always been an invaluable intellectual experience and a source of much enjoyment! A very special thanks to Samuel Marguerat without whom it would have been impossible to learn the various fission yeast techniques, especially "Normalisation of the Microarrays". Thanks for being so patient, while teaching me again and again and again... and again! I would also like to thank Tania for introducing me to Western Blotting, Sanjay for helping me in using the GeneSpring and Luis for all the tips and tricks to getting stable polyploids.

A huge thanks to Sandra Codlin who has been superb in organising the lab and who has always been extremely helpful, no matter which problem had to be solved. Thanks a lot for all your support and for proof reading this thesis! I would like to thank Xavi, who took a lot of time especially in the beginning to explain the basics of polyploidy in fission yeast. Babis and Martin, I would like to thank you for being very kind and helpful in introducing me to the various techniques especially to the pheno-typing assay and answering all my queries. I cannot forget to thank Bruno and Danny for all the laughs and teasers, although you never quite…. succeeded in mimicking Jürg's voice. I owe thanks to Maria for being so generous and supportive in showing me the analysis of FACS data. I would also like to thank Sadaf Sheikh, Dora and Sara for proofreading parts of this thesis.

A very big thank you goes to my wonderful friend Shafaq and her family including Uncle, Aunty, Faisal Bhai, Sadaf Sheikh, Waleed, Hisham and all the crazy kids for keeping doors open for me, whenever I missed my family; for helping me to shape up my thesis and above all for providing yummy curries while I was writing the thesis. I would also like to thank my lovely friend Rehana for keeping me motivated and always being there to help during the stressful times.

Thank you to COMSATS Institute of Information Technology for funding my studies at UCL. A special thanks to Rector Dr. Junaid Zaidi and project director Tahir Naeem for their kind support. I would also like to thank all my teachers, mentors and supervisors from the start of my school life; without all your immense efforts, I could not have come this far.

Unending thanks to my parents, who gave me the greatest start of life, confidence, motivation and all the support I needed. A big, big thanks to my father who always wanted his daughter to go for higher education even when there was no such trend in our family and encouraged and helped me in my studies during my entire school life. Thanks to my brothers and sisters for all the love, support and the fun in my life, especially Bhai jan for always taking care of me.

There was a time when I stopped thinking about pursuing my dream to do a PhD, but my father-in-law kept it alive and persuaded me to compete for the overseas scholarships. Thank you Taya ji for motivating me and giving me the strength and confidence to go forward. Also, without the help of my mother- in-law, it would not have been possible to finish my PhD; she made it all possible for me by taking care of my baby daughter Maheen better than any mother could ask for. You gave me the peace of mind to carry on. Thanks a lot Bubo ji for being so kind and supportive. I cannot forget to thank Haleema for all the support in all these years. A big thanks to my daughter Maheen for behaving so well and being so patient; mummy loves you and can't wait to see you.

Last but not least, the greatest contribution towards my PhD is from my loving husband, who went against the norms of our traditions and our society by sending me to the UK to achieve my goal. Thank you very, very much for all the love, support and patience - **I appreciate every single bit of it!**

Abstract

Cells are classified based on their ploidy into haploids, containing a single chromosome set, diploids, containing two chromosome sets, polyploids, containing more than two chromosome sets, and aneuploids, containing abnormal chromosome numbers. Polyploidy is typically accompanied by increased cell size. Polyploid cells are found in most tumors and exhibit chromosomal instability that leads to aneuploidy. The effects of aberrant ploidy on genome regulation and on cell size are not well understood.

I used fission yeast as a model to analyse impacts of altered ploidy and cell size on gene expression. Using aneuploids that are disomic or trisomic for a portion of chromosome III, I find that total mRNA levels scale with DNA copy numbers. Aneuploidy also affects the transcription of some genes present in monosomic areas, possibly reflecting associated regulatory genes in disomic or trisomic areas.

I also analysed the effect of polyploidy on genome expression by constructing diploid and tetraploid strains. Diploids were stable with normal cell shape, while tetraploids showed irregular morphologies and often lost chromosomes. Increased ploidy resulted in increased cell size, and also in a linear increase in cellular RNA levels. Using spike-in controls and normalization, we showed that increased transcription in polyploids does not affect ratios between total RNA and mRNA. Cells kept a tight control on genome-wide transcription which generally scaled with the copy numbers of genes, a few genes were differentially regulated as a function of polyploidy and/or cell size. These genes were present in multiple copies close to telomeres and may function at the cell surface. They were also differentially regulated in haploid cell-size mutants, indicating a role of cell size, rather than ploidy, in controlling these genes. Intriguingly, deletion and overexpression of these genes in turn resulted in a significant decrease or increase in cell size, respectively, raising the possibility that the genes are involved in size control.

Table of Contents

List of Figures

List of Tables

Chapter 1

1 Introduction

Higher eukaryotes exhibit incredible diversity in phenotypic traits such as size, colour, or body structure. Genetic studies have revealed that these variations map to the level of the genome. Genomes of organisms vary both in terms of size and chromosomal content. An important but poorly understood variable of the genome, and hence of phenotypic traits of living organisms, is their ploidy. Ploidy refers to the number of chromosomal sets contained in the nucleus of a eukaryotic cell. Ploidy of organisms accounts for an important characteristic that differs not only among species but even within a species (Lewis, 1980). Cells can be classified on the basis of their ploidy into: haploid - containing a single set of chromosomes; diploid - containing two sets of chromosomes; polyploid - containing more than two sets of chromosomes; and an aneuploid - containing an irregular number of chromosomes. Human body (somatic) cells are diploid. During meiosis, these cells give rise to germs cells (gametes), eggs or sperm, in which the original number of chromosomes is reduced to half, thus making these gametes haploid cells. Diploid chromosomal content is restored in a zygote after fusion of two gametes during fertilisation. The condition of a cell being haploid, diploid, polyploid and aneuploid is known as haploidy, diploidy, polyploidy and aneuploidy. In higher eukaryotes, diploidy is a common phenomenon, probably due to the fact that it not only allows sexual reproduction but also assists in genetic recombination (Storchova and Pellman, 2004a).

Although many living species on earth are either haploids (such as bacteria, many species of fungi, and moss.), or diploids (such as ferns, several fungi and vertebrates), there are many species that show polyploidy (such as coffee plants, wheat, and other plants). Polyploids represent a special type of class in nature, as they can live with three or more full sets of homologous chromosomes. In mammals, polyploidy can occur in specific tissues due to a variety of reasons and can result in chromosomal instabilities leading to aneuploidy. Both chromosomal instabilities and aneuploidy are common features of cancers. Variation in ploidy is not only a common phenomenon of plants and some animals but also occurs in certain cell lines within an organism.

This chapter begins with a review of literature on polyploidy, mechanisms of polyploidy, and its advantages and dis-advantages in plants, vertebrates and yeast. The consequences of polyploidy on growth and genome regulation of different organisms, along with its specific role in regulating the cell size, are then explained. The relationship between polyploidy and aneuploidy and the role of aneuploidy on different cellular aspects of living organism is discussed with an emphasis on its role in gene expression regulation drawn from previous studies conducted in plants, mammals and yeast. The last section describes the aims and objectives of this study using fission yeast as a model organism.

1.1 Polyploidy

Polyploidy or Whole Genome Duplication (WGD) refers to the state of having more than two sets of homologous chromosomes in the nucleus of a cell. A polyploid with three sets of chromosomes is called a triploid, with four sets a tetraploid, and so on. Analysis based on complete genome sequence has revealed that many existing eukaryotes have undergone at least one round of ancient genome duplication.

This type of polyploid formation is observed in nurse cells of *Drosophila melanogaster,* and also in megakaryocytes and hepatocytes (Zimmet and Ravid, 2000; Edgar and Orr-Weaver, 2001; Ravid et al., 2002). Many tissues, including salivary glands in *Drosophila,* undergo endoreplication during the process of embryogenesis and become autotetraploids (Smith and Orr-Weaver, 1991). Polyploid nurse and follicle cells are essential for oocyte development in *Drosophila* adult females. If the endoreplication in these cells is disrupted, maturation of egg chambers is severely delayed and results in sterility (Maines et al., 2004).

Polyploids are categorized into two main groups on the basis of their origin, autopolyploids and allopolyploids (Ramsey and Schemske, 1998). Autopolyploids result from the doubling of homologous genomes from a single, or closely related, species (e.g., autotetraploid, AAAA). These can be formed by the fusion of two gametes from a single species or by endoreplication (Edgar and Orr-Weaver, 2001). On the other hand, allopolyploids result from the hybridization of genomes between different parental species (figure 1.1). Consequently, they contain a genome that is a

combination of two or more genomes derived from different species (e.g., allotetraploid, AABB). A large number of plant species have been proposed to arise as a result of allopolyploidization (Otto and Whitton, 2000), such as wheat, oat, canola, coffee, banana and wheat (Jiang et al., 1998).

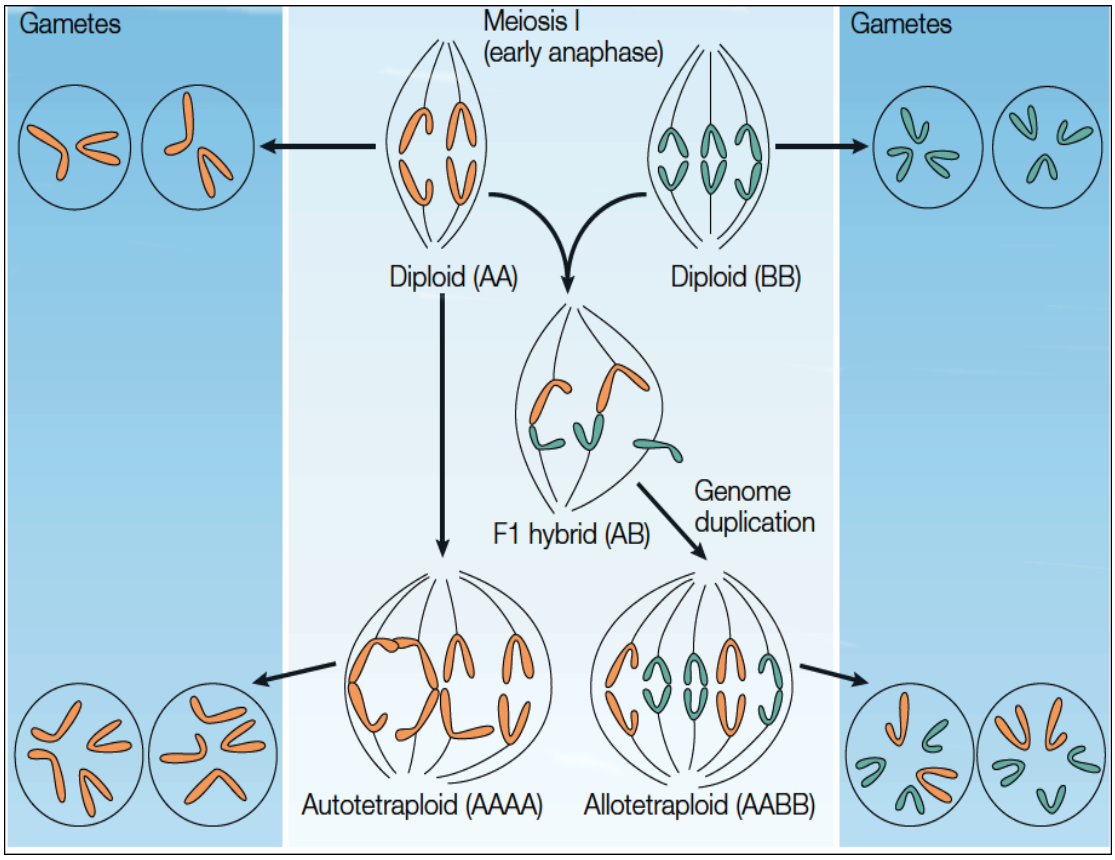


Figure 1.1: Schematic illustration of polyploid formation. This figure explains the formation of autotetraploidy and allotetraploidy. Defective pairing of homologous chromosomes in the F1 is restored by one round of genome duplication. The duplication process results in an allotetraploid where homologous chromosomes are derived from two parents, while an autotetraploid has chromosomes from one parent only. Adapted from (Comai, 2005).

1.2 Polyploidy as a mechanism of evolution

Large scale duplication in the genome of contemporary diploid organisms provides evidence of a past polyploidization event in the origin of these organisms. Ancient polyploids that underwent diploidization by means of chromosomal rearrangements in duplicated genomes are considered as paleopolyploids such as plants, *Sacharomyces cerevisiae*, zebrafish and humans (Wolfe, 2001). Many diploid plants have also been identified as paleopolyploids (Blanc and Wolfe, 2004).

The relationship between polyploidy and the evolution of vertebrates was initiated by Susumu Ohno who first proposed the idea that whole genome duplication/polyploidization gives rise to substrates for evolution (Ohno, 1970). He put forward the theory that is widely known as **2R hypothesis**. According to this theory, two rounds (2R) of whole genome duplications had shaped the evolutionary history of vertebrates (figure 1.2). This hypothesis was mainly based on his findings of apparently duplicated chromosomal segments in the human genome. These segments were found to be delineated by two pairs of duplicated genes present on chromosome 11 and 12. He suggested that these segments were the remainder of an ancient polyploidy.

The 2R hypothesis is best supported by the analysis of the homeobox (HOX) genes that are involved in controlling the development of body plans. Mammals possess four identical clusters of HOX genes, while *Drosophila* and other metazoans have only one copy of this cluster. The order of genes in these clusters has remained identical in both vertebrates and invertebrates (Duboule and Dollé, 1989; Graham et al., 1989; Schughart et al., 1989). These findings strongly suggest that the four large HOX gene clusters in mammals arose from two rounds of duplication of one ancestral cluster related to the one in *Drosophila* (Schughart et al., 1989). Prevalence of high degree of conservation of structural organization between mammalian and insect HOX clusters suggests that the duplicated chromosomal regions might be quite large. Since duplications of individual chromosomes tend to be deleterious due to imbalances in gene dosage, whole genome duplication was proposed as a plausible mechanism.

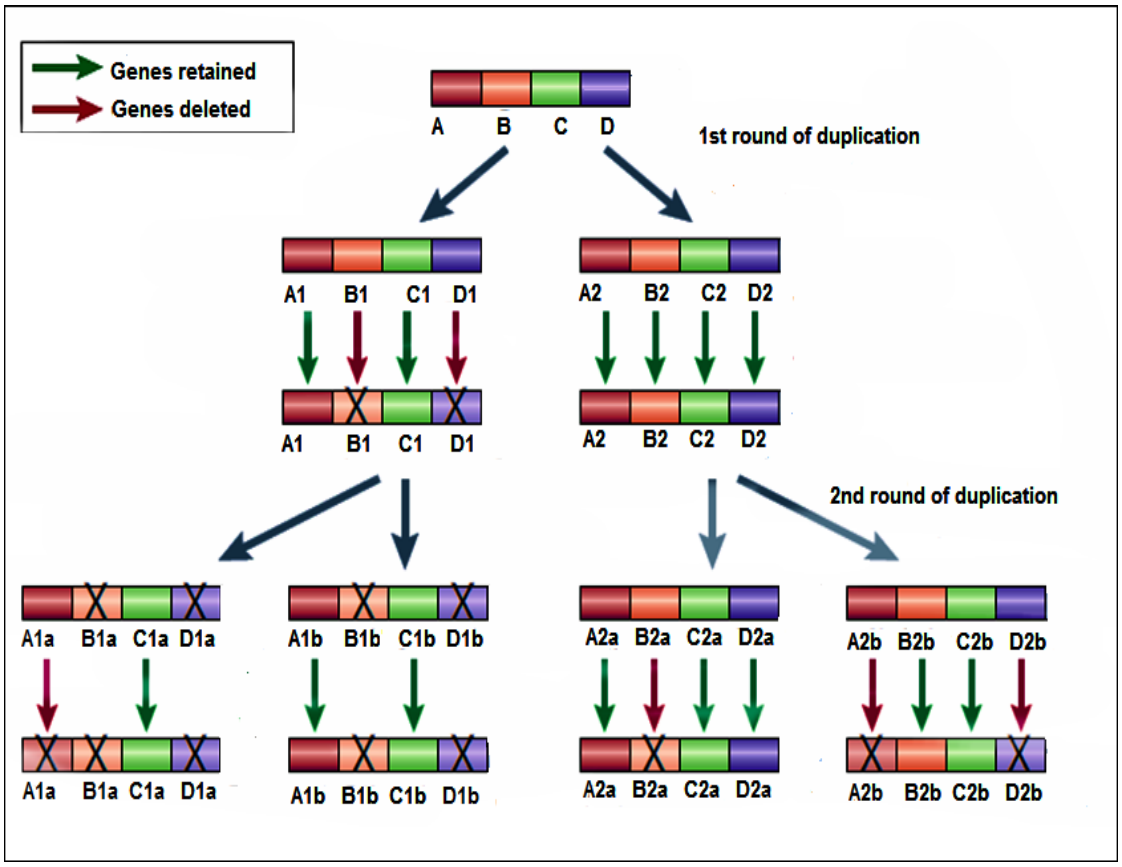


Figure 1.2: Model of the 2R hypothesis. This model shows how 5 genes are doubled after 2 rounds of duplication events. A–D represents 5 genes on a chromosome in an ancestral organism. Some copies of genes were retained, while others were deleted over generations. Adapted from (Wolfe, 2001).

Further evidence in support of Ohno's hypothesis comes from the fungal kingdom. Species of the *Saccharomyces* genus seem to have evolved from an ancestor that ascended from whole genome duplication. In one study, Wolfe and Shields (1997) find a sequence homology within the *S. cerevisiae* genome and discovered 55 pairs of large (>50Kb) duplicated regions, which were dispersed throughout the genome. Each pair of duplicated regions contained homologous genes that maintained their gene order and orientation with respect to the centromeres, thus suggesting the tetraploid origin for *Saccharomyces* (Wolfe and Shields, 1997). The apparent link

between WGD and evolution of *Saccharomyces* species demonstrates the evolutionary potential of polyploidy.

It is worth mentioning that the WGD events occurred as fruiting plants populated the earth's flora (Wolfe and Shields, 1997). The improved evolutionary potential of polyploidy and the increased availability of fruits probably gave rise to the emergence of fermentative yeast.

1.3 Polyploidy in plants

"*Most of the world's biota is polyploid, and life on earth is predominantly a polyploid plant phenomenon*" (Bennett, 2004). Plant species, specifically angiosperms are more prone to polyploidization. Estimates show that up to 70% of flowering plants are polyploid (Masterson, 1994). Polyploidy plays a central role in plant speciation. Many crop species at present are polyploids such as wheat, oat, coffee, cotton, soybean and banana (Jiang et al., 1998).

Polyploid plants show differences not only in morphology and physiology but also in cellular and biochemical aspects when compared to diploid plants. Polyploid plants have bigger cells, stomatas, larger leaves, as well as bigger flowers and fruits (figure 1.3). Tetraploids also show higher vegetative volume and seed weight. Their shoots are thicker and have shorter internodes. On the other hand, these tetraploids are less fertile and show delayed flowering and fruit formation (Yildiz, 2013). In a grass (*Panicum virgatum*) octaploid variants show higher DNA concentration, photosynthetic rate, and chlorophyll abundance as compared to tetraploid variants. These octaploids also show 20-70% increase in enzyme activity (Warner et al., 1987).

Arabidopsis thaliana – an example of whole genome duplication in plants

The flowering plant *Arabidopsis thaliana* provides the best studied example of whole genome duplication in plants. The Arabidopsis Genome Initiative published the complete sequence of this plant in 2000. It was shown that 67.9 Mb, about 60% of its genome, was duplicated and that the co-linear arrangement of genes within duplicated regions was not disturbed (Arabidopsis Genome Initiative, 2000). Later,

protein sequence alignments also confirmed these findings and showed that as much as 70% of the *Arabidopsis* genome exists in duplicates (Blanc et al., 2003).

Both studies could not find any overlap of duplicated regions, or any triplication of genome segments that could arise from random and independent duplication of smaller portions of chromosomes. The presence of duplicated regions in this organism strongly suggests that its genome might have evolved from large-scale duplication events.

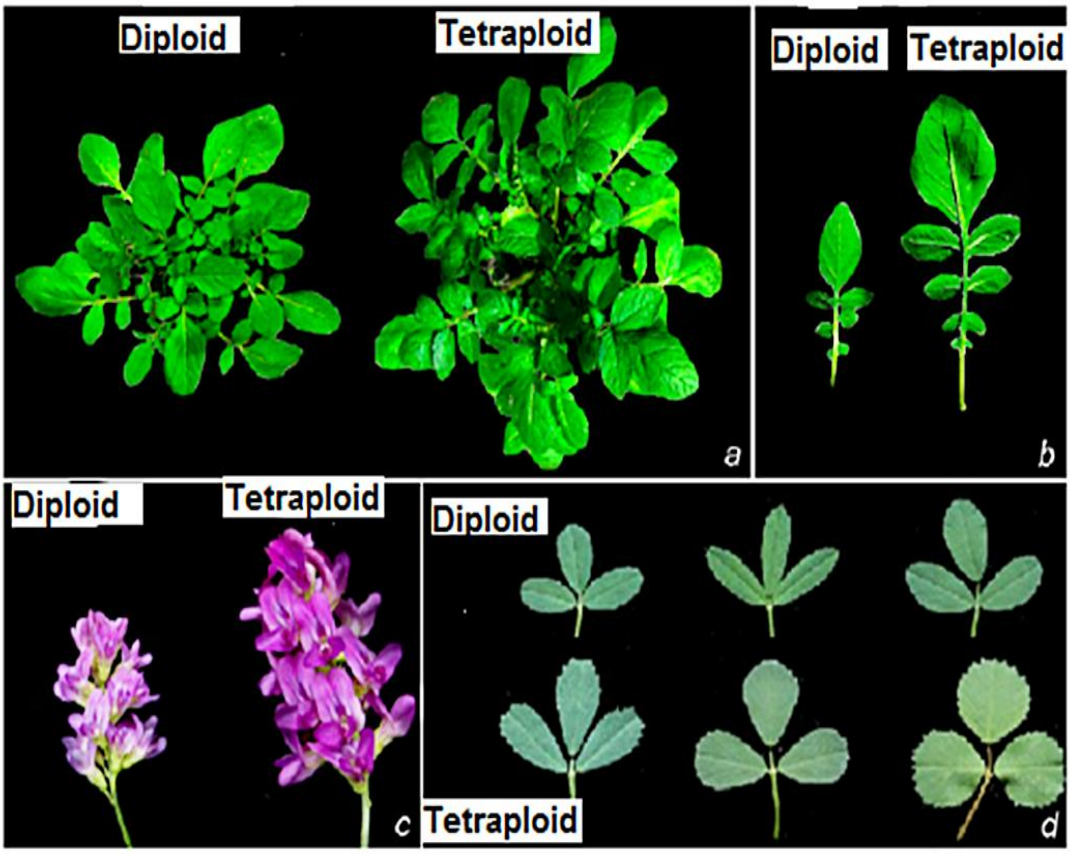


Figure 1.3: phenotypic variations in response to increase in ploidy in plants. Difference in phenotype between diploid and tetraploid plants of *Solanum commersonii (a, b) and in Medicago sativa (c,d).* Adapted from (Aversano et al., 2012).

1.4 Polyploidy in fungi

There are more than 100,000 species in the fungal kingdom (Albertin and Marullo, 2012). Several species among them have been identified as natural polyploids. Evidence of polyploidy has been observed mainly among aquatic fungi that exhibit a series of polyploidy such as autotriploidy, autotetraploidy and allotetraploidy (Albertin and Marullo, 2012). Natural polyploids have also been identified in the phylum *Basidiomycota*. *Cyathus stercoreus* and numerous edible mushrooms of this phylum are polyploids. For instance, *Cyathus stercoreus*, known as the dung-loving bird's nest, shows tetraploidy. It shows tetravalent formation during meiosis between homologous chromosome (Lu, 1964). Although, many polyploids have been identified within the largest phylum *Ascomycota* (Nielsen and Yohalem, 2001), the *Saccharomyces* genus still provides the best studied evidence of polyploidy in the fungi kingdom. Two such examples of polyploids in fungi are discussed below.

Rhizopus oryzae

A human pathogen, known as *Rhizopus oryzae,* is from the early sample of fungi whose whole genome is now sequenced. Its genome analysis has revealed the occurrence of an ancient WGD (Ma et al., 2009). These duplicated regions cover ~12% of the genome and cover 45% of those genes that have paralogs inside the genome. These duplications have resulted in an increase in copy number for proteins that are required for respiratory and proteasome functions. The ancient WGD have led to an increase of 2- to 10-fold in gene families that are required for pathogen virulence, synthesis of fungal specific cell walls, and signal transduction. These duplications provide *R. oryzae* with the genetic flexibility for rapid adaptations to severe conditions, and also to the host immune system.

R. oryzae shares many ancestral genes with Metazoans. The homologs that are common between *R. oryzae* and Metazoans include genes that are required for important cellular functions, such as developmental processes, transcriptional regulation and signal transduction. Analysis of the duplicated regions has also shown that the order and orientation of these genes is conserved, thus providing the evidence of an ancestral duplication event for these regions (Ma et al., 2009).

Saccharomyces cerevisiae

S. cerevisiae is widely used as a eukaryote model organism in molecular and cell biology. The emergence of this fermentative yeast species is mentioned as one of the best known examples for the evolutionary consequences of WGD. The analysis of its genome sequence has revealed that the *Saccharomyces* genus has undergone WGD in its evolutionary history (Wolfe and Shields, 1997). The event of WGD has been proposed to occur after the divergence of *Saccharomyces* from *Kluyveromyces*. After the WGD, their genome underwent a diploidization event, by which a polyploid genome was reduced to a diploid one (Wolfe, 2001). Many of the gene pairs that have been retained after diploidization are involved in carbohydrate metabolism and show specific expression behaviour in response to glucose or oxygen availability. Genes present in the duplicated regions have escaped deletion by attaining new functions that help the organism to cope with conditions related to fermentation (Wolfe and Shields, 1997). Another study has shown that a significant number of yeasts have recently become polyploids, e.g. different strains of *S. cerevisiae* that are used in food processing have shown evidence of autotetraploidy in 10 out of 26 strains (Naumov et al., 2000; Albertin et al., 2009).

Polyploid cells of *S. cerevisiae* remain viable during exponential growth, but they die very quickly after entering the stationary phase. Haploid cells have been found to stay viable for a few weeks in stationary phase, but tetraploids are completely inviable after just a few days (Andalis et al., 2004).

The above mentioned examples provide evidence that polyploidy shapes the evolution of species, helping them for rapid adaptations by changing fundamental cellular processes (such as resistance against host immune system or metabolism).

1.5 Polyploidy in vertebrates

Although polyploidy is thought to be fairly common in plants, it is also observed in a wide range of tissues in diploid organisms. Polyploids in mammals can arise due to an aberrant mitotic or meiotic cycle, which results in the gametes having two sets of chromosomes. These gametes, on successful fusion with another gamete (of haploid

or diploid origin), can produce a polyploid zygote that could be triploid or tetraploid. Polyploidy has been reported for megakaryocytes (ploidy ranging from 16n to 128n) (Ravid et al., 2002) , cardiomyocytes (tetraploid) [3], trophoblast giant cells (from 8n to 64n) (Zybina and Zybina, 2005) and retinal polyploid cells. Polyploid cells can arise in normal diploid organisms due to endoreplication, cell – cell fusion or an abortive cell cycle.

1.5.1 Mechanisms of polyploidy

In diploid organisms, a large number of tissues can undergo polyploidy through the process of endoreplication/cytokinesis failure (figure 1.4). This is a specialized cell cycle in which DNA replication is uncoupled from the subsequent cell division. In this type of cycle, cells undergo several rounds of DNA replication but skip mitosis, resulting in an autotetraploid (Zimmet and Ravid, 2000; Edgar and Orr-Weaver, 2001). In mammals, trophoblasts and megakaryocytes are well-known examples of polyploid cell types that arise due to endoreplication (Edgar and Orr-Weaver, 2001; Ravid et al., 2002). Polyploidy through endoreplication is typically related to terminal differentiation, and is considered as a suitable mechanism to produce cells with increased metabolic capability and specialization in mass production/storage of macromolecules (Edgar and Orr-Weaver, 2001). Trophoblasts in the mammalian placenta show a DNA content of greater than 1000C. These cells facilitate the implantation of an embryo and to meet the high demand of molecular transport between mother and the developing foetus (Zybina and Zybina, 2005). In certain cell types, during the process of development and differentiation, cell fusion leads to the formation of polyploid cells/tissues such as, osteoclasts and skeletal muscle cells (Vignery, 2000). During this process, cells can fuse either their nuclei or membranes giving rise to a mononuclear or multinuclear cell, respectively.

Polyploidy in vertebrates can also arise due to an abortive cell cycle/mitotic slippage that could be caused by a wide variety of abnormalities at different stages in the cell cycle such as replication, segregation of chromosomes, or cytokinesis. Although cells can trigger a checkpoint response that will arrest this aberrant cell cycle, some cells can skip this arrest, leading to the formation of tetraploids (Andreassen et al., 1996; Minn et al., 1996).

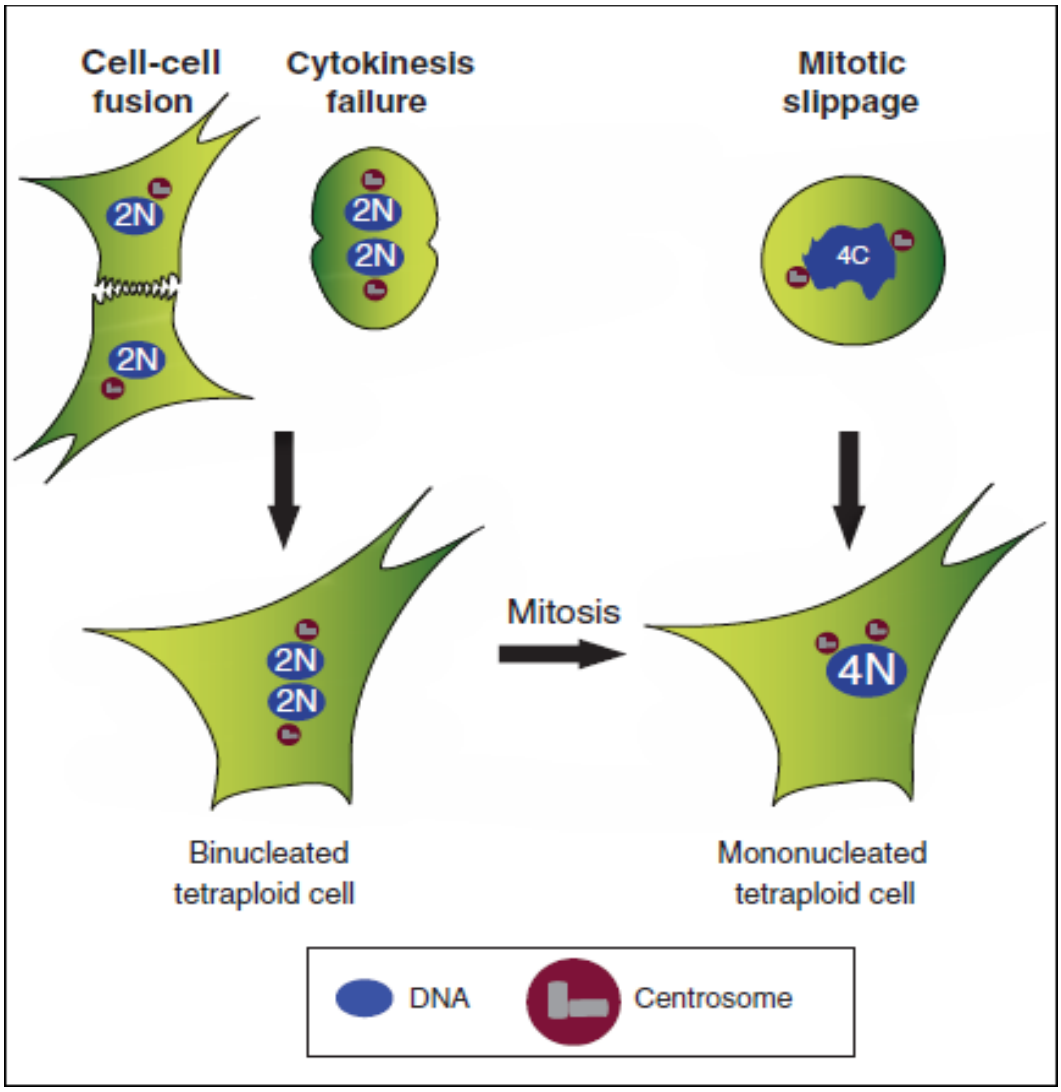


Figure 1.4: Mechanism of polyploid formation in vertebrates. Cell to cell fusion and endoreplication/cytokinesis failure results in a binucleated cell, while mitotic slippage results in a mononucleated tetraploid cell. Taken from (Storchova and Kuffer, 2008).

1.5.2 Polyploidy in vertebrates in response to stress and aging

Although diploidy is the preferred state for mammalian cells, various studies have mentioned a major role of diploid to polyploid transformation during different pathological conditions in several tissues. Polyploid cells have been observed in hepatocytes and cardiomyocytes in response to an injury or various stresses.

Hepatocytes with dysfunctional telomeres do not go through apoptosis or senescence like normal mitotically active cells. When these cells are challenged by the demand to regenerate liver after partial hepatectomy, they undergo several cycles of replication but skip anaphase and telophase. This results in an increase in ploidy and cell size without altering the cell number. Such polyploidization events account for the regeneration of liver mass and function (Lazzerini Denchi et al., 2006).

Similarly, when exposed to oxidative stress, hepatocytes show a significant increase in polyploid cells. It has been shown that on exposure to the radiation, hepatocytes show a signature of oxidative injury that is accompanied by a deletion of intracellular antioxidants and an increase in polyploidy (Gorla et al., 2001). Moreover, a study on rats has also shown the role of aging in the increase of reactive oxygen species. This state induces the proliferative potential of liver cells, thus increasing in the production of octaploid population of the cells (Sanz et al., 1997).

In humans, polyploid cells in the liver start appearing during postnatal liver development. Their rate of accumulation stays steady during a mature phase but a significant rise in polyploid cells is observed during the process of ageing (Kudryavtsev and Kudryavtseva, 1993). Polyploid cardiomyocytes have also been reported in the human heart after myocardial infarction (Herget et al., 1997). A shift to polyploidy may be a suitable way to cope with stress and injuries when cells cannot undergo apoptosis or proliferation.

All these above mentioned examples provide evidence of a significant association between a rise of polyploidy and various stressors, whether they are mechanical injuries, oxidative stress or ageing. Currently, there is no clear estimate of how frequently this occurs in the lifetime of a normal individual. There are few observations that suggest that polyploidy might have higher rates than expected, for

example, the identification of tetraploid cells during tumour development and several defects that can result in an abortive cell cycle (Storchova and Pellman, 2004b).

All these examples establish the role of polyploidy in conferring specific properties to these cells. For example, polyploidy in megakaryocytes is related to differentiation. That is required for regulation of platelets production and function (Raslova et al., 2007). Similarly, polyploidy might also be involved in protecting the hepatocytes from genotoxic damage by simply increasing the copy number of functional genes. This is particularly important for the liver which is mainly involved in metabolism and elimination of toxic compounds. It could also be an economic way out to tackle growth issues when an organ works within its capabilities, thus keeping the energy needed for cell division to a minimum (Gentric et al., 2012).

1.6 Role of polyploidy in regulation of gene expression

Polyploidy has long been considered as an important factor in shaping the evolution of higher eukaryotes, particularly of flowering plants. Polyploidy results in instant phenotypic variations, e.g. enlarged cell size, increased biomass, slow growth, as mentioned previously in sections 1.2 and 1.4. The appearances of these morphological and functional phenotypes in response to varying ploidy indicate that ploidy alters the levels of gene expression. Polyploidy results in altered gene dosage that could disturb normal patterns of gene expression by affecting regulatory interactions. Although much is known about the presence of polyploid cell types in eukaryotes, transcriptional profiles of very few cell types have been characterized.

Examples from plants

In plants, a few studies have focused on exploring the role of polyploidy on the regulation of gene expression. Guo et al. (1996) studied the effect of dosage change on gene expression in the maize plant, *Zea mays*. He used a polyploid series consisting of monoploids, diploids, triploids and tetraploids and analysed the expression of 18 genes. For most genes, their expression patterns showed linear relationships with ploidy/gene dosage, but one gene, the thiol protease like gene *csu5*, showed a decrease in expression with increasing ploidy. This study suggested

that with few exceptions, an increase in ploidy results in a corresponding increase in gene expression per cell (Guo et al., 1996).

In another study Albertin et al. (2006) conducted large scale analyses of the proteome, in the stem and root, of *Brassica napus* that is derived from *B. rapa* and *B. oleracea*, to study the modifications of gene expression. They found that a large number of proteins show non additive expression patterns similar to that of the paternal parent. They suggested that in allopolyploids hybridization of different genomes, rather than polyploidization per species, causes large-scale changes (Albertin et al., 2006). Similarly, it has been found that deletions and rearrangements in the genome can also bring changes in gene expression in Allopolyploids (reviewed in Osborn et al., 2003).

Arabidopsis has also been examined for the effects of polyploidy on transcriptional regulation. Gene expression profiles were analysed and compared for several successive generations for *A. thaliana*, as well as for *A. arenosa* and *A. suecica*, where the first two are autotetraploid and the latter is allotetraploid, generated by hybridization between *A. thaliana* and *A. arenosa*. A total of 2430 genes were analysed for their expression patterns. In newly formed allotetraploids, some genes were silenced and remained silenced over many next generations. Similarly, some genes that were silent in autotetraploid parental populations were reactivated in this allopolyploid. A small number of genes showed ploidy related changes in expression (Wang et al., 2004). Differentially regulated genes in autotetraploids, showing ploidy dependent repression, were ones coding for malate dehydrogenase, cellulose synthase subunit and a DNA-binding protein. Overexpressed genes include those encoding a protein for vacuolar transport, a nuclear matrix component, Rad54 – involved in DNA repair and recombination and a kinesin-related protein. Interestingly, two DNA-methylases involved in modification of RAD54 chromatin silenced its expression in allotetraploids (Wang et al., 2004).

Besides genetic changes, regulatory networks such as transcription factors encoded by genomes derived from an individual parent may also result in novel or unbalanced molecular function. Factors encoded by the maternal/paternal genome might control the chromatin on the other genome. All these mechanisms in

allopolyploids could result in epigenetic changes, which often show novel phenotypes and thus better adaptation when compared with their parental species (Osborn et al., 2003; Wang et al., 2004).

Although few studies in plants have explored the changes in gene expression patterns caused by polyploidy, most of them focused on allopolyploids that masked the effects on gene expression caused by the increase in ploidy. Differential expression of genes found in these studies is caused mainly by the two different parental types rather than ploidy. Many genes of either maternal or paternal origin were silenced in the newly formed allopolyploids (Kashkush et al., 2002; Adams et al., 2004; Wang et al., 2004; Albertin et al., 2006).

Examples from mammals

In mammals, polyploidy is frequently observed in megakaryocytes, hepatocytes and cardiomyocytes. Polyploidy in megakaryocytes is associated with differentiation and enlargement of cells required for platelet biogenesis (Raslova et al., 2007). To check whether it has any role in modification of gene expression in megakaryocytes, Raslova et al. (2007) conducted a microarray based study in a series of polyploids (2N, 4N, 8N, 16N) obtained from human $CD34^{+}$ cell lines. They found two differentially regulated clusters of 105 genes, one corresponding to 2N and 4N while the others to 8N and 16N, suggesting that cells with higher ploidies have different properties than those with lower ploidies (Raslova et al., 2007). Significant numbers of the up-regulated genes in these polyploids overlapped with the ones previously observed to be up-regulated during different stages of differentiation of megakaryocytes (Kim et al., 2002; Shim et al., 2004), while down-regulated genes were involved in DNA replication (CDC6 and MCM), proliferation and DNA repair. Ploidy dependent overexpression of platelet biogenesis genes is consistent with the role of polyploidy in megakaryocyte differentiation, but the down-regulation of DNA replication genes in the polyploids remained unclear. This study suggests a role for polyploidization in terminal differentiation of megakaryocytes and platelet biogenesis.

Authors have reported a few technical challenges that might have affected the clear establishment of the role of polyploidy on gene expression in megakaryocytes. As megakaryocytes in bone marrow are not abundant enough to proceed for expression analysis, they were cultured *in vitro*. This provided a limitation because these cultured megakaryocytes could not reach high ploidy as they do in bone marrow and were mixed with progenitor cells. Therefore, these observations obtained *in vitro* may not be the true representative of the *in vivo* physiology of these cells, and could create a totally unique megakaryocyte transcriptome.

In another study, the role of polyploidization events on regulation of gene expression was explored in hepatocytes in both humans and mouse tissue (Anatskaya and Vinogradov, 2007). A large scale gene expression analysis in these tissues showed that genes involved in stress response are overexpressed by polyploidy, e.g. regulators of apoptosis, DNA damage and repair, hypoxia and protein turnover and transcription (Anatskaya and Vinogradov, 2007). Similarly, they also found ploidy related de-repressed genes that encode proteins for aerobic respiration and metabolism of carbohydrates and fatty acids. However, they could not establish whether these differentially expressed genes were strictly regulated by polyploidy, or whether other factors might also be involved. Their data indicate a link between a change in ploidy and modification of gene expression to prolong the cell survival by protecting against apoptosis, hypoxia and DNA damage. These findings are in agreement with the previous study that also observed the up-regulation of genes involved in stress response due to mechanical injury in polyploid cardiomyocytes and hepatocytes (Herget et al., 1997; Lazzerini Denchi et al., 2006). Activation of the stress responsive genes might help them to cope with these stresses.

Example from yeast - Saccharomyces cerevisiae

The budding yeast *S. cerevisiae* is generally haploid or diploid but its polyploid states can be easily generated in the laboratory. Haploid cells of *S. cerevisiae* are either *MATa* or *MATα* and can mate with cells of the opposite mating type to form diploid cells that are heterologous at the mating type locus MATa/ MATα. Although these diploids stay stable, due to the heterozygosity at the mating type locus they cannot mate with each other to produce tetraploids, but they can be easily

manipulated to create mating type homozygotes. They can be further mated to produce cells of higher ploidies (Andalis et al., 2004). *S. cerevisiae* cells show an increase in cell size in response to an increase in ploidy (figure 1.5 a) (Andalis et al., 2004; Galitski et al., 1999; Storchova et al., 2006). Polyploid cells also show different growth patterns and morphology as compared to their haploid and diploid counterparts.

In a microarray based study, Galitski et al. (1999) used an isogenic series of polyploids to identify the pattern of gene expression in response to increasing ploidy (figure 1.5b). They identified a small number of genes (17) that were differentially regulated in polyploids (Galitski, 1999). Some of the differentially regulated genes provide an explanation for the novel phenotypes of the polyploids. For example, CTS1 that codes for endochitinase, shows ploidy dependent up regulation, and is likely that its overexpression contributes to the separation of mother-daughter cells at cytokinesis. Null mutants for this gene form larger clumps as daughter cells cannot separate from mother cells, while tetraploids show much reduced clumps. Similarly, down regulation of *FLO11*, a gene functioning in adhesion, affects the adhesion of polyploids to agar. In this study, they could not find enrichment for any particular functional category of genes that could explain a relationship between ploidy and differential regulation of genes.

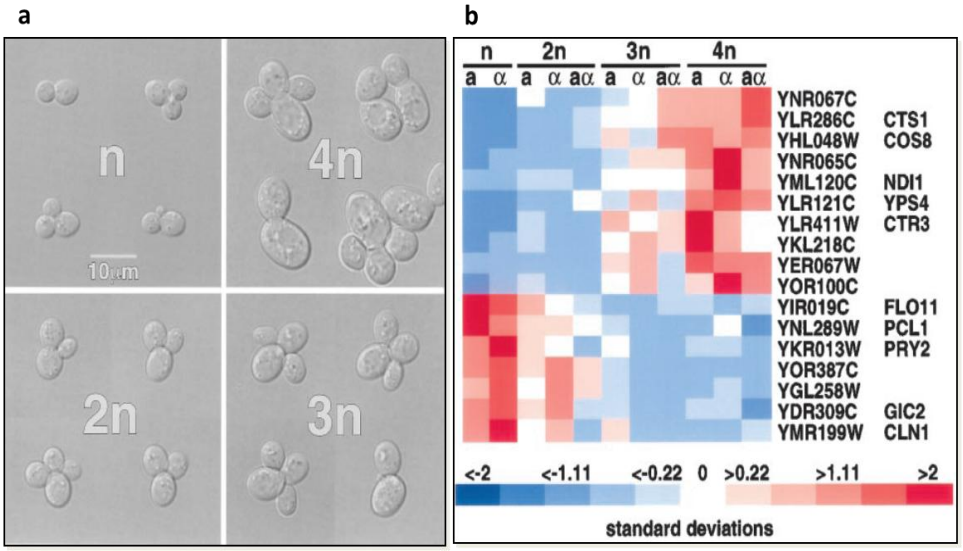


Figure 1.5: Effect of polyploidy on phenotype and gene expression in budding yeast. Cell size increases with ploidy (a). Gene expression profiles for differentially

expressed genes (b), where blue color represents down-regulated and red represents up-regulated genes. Taken from (Galitski, 1999).

Recently, Wu et al. (2010) used RNA-seq to identify the differentially regulated genes in tetraploid cells of *S. cerevisiae*. They identified a larger set of genes (65) that showed differential expression patterns in tetraploids as compared to haploids. Among these, 30 genes were induced while 35 were repressed in tetraploid cells. The majority of these genes were found to encode proteins localised to the cell wall and plasma membrane, which led to a hypothesis that their expression might be driven by larger cell size rather than an increase of ploidy (Wu et al., 2010). They further analysed the expression pattern of these differentially regulated genes in haploid mutants of larger cells and found similar expression behaviour as in polyploids. Their findings suggested a relationship between cell size and control of transcription such that cells adjust their gene expression in response to changes in size.

Findings from the study conducted by Wu et al. are in agreement with a study in *S. pombe*: Zhurinsky et al. (2010), by using temperature sensitive mutants of larger size, showed that the rate of total RNA synthesis increases with an increase in cell size (Zhurinsky et al., 2010). They also observed that the rate of transcription for individual genes increases linearly with an increase in cell size. However, after a 4- to 5-fold increase in cell size, the rate of transcription starts to level off as the gene dosage becomes limiting at this stage. These findings shed light on the importance of cell size, along with ploidy, in the control of global cellular transcription.

1.7 Influence of polyploidy on cell size

One instantly noticeable effect of polyploidy on phenotype is an increased cell size: size scales linearly with ploidy, ranging up to a million-fold, both at inter- and intra-species levels (Cavalier-Smith, 2005; Turner et al., 2012). For example, diploid and tetraploid cells in yeast are almost 2-fold and 4-fold larger than haploid cells (Cook and Tyers, 2007). Correlations between cell size and ploidy have been observed in many organisms. An increase in ploidy results in larger nuclear volume, more chromosomal content and altered genes expression, which in turn can affect the

cytoplasmic content (Jorgensen and Tyers, 2004). Increased size in response to ploidy also explains the special requirements of specific cell types such as salivary glands of *Drosophila* larvae, megakaryocytes and trophoblasts in mammals (Edgar and Orr-Weaver, 2001). Despite the huge range of cell size among organisms in response to polyploidy, actively proliferating cells within a species or a tissue show very little variation in terms of size. This shows that cells have a size sensing mechanism required for size homeostasis, which could involve many parallel and interrelated processes. One of these processes is the coordination between cell growth and cell cycle that seems to apply to all organisms (Jorgensen and Tyers, 2004). It is widely considered that a minimum threshold is established for a certain cell size at division, due to coordination between cell growth and cell division (Cook and Tyers, 2007). During cell cycle progression, the ploidy of cells also alternates between G1 and G2, therefore it can be argued that it could be ploidy rather than cell growth that helps in determining the cell size threshold. Whether and how this cell size threshold is altered in response to an increase in ploidy or cell growth is still unknown, as the precise mechanism by which cells regulate their size is not yet clear.

So far, studies in fission and budding yeast have provided much of the understanding about the molecular basis of cell size control, as both of these organisms possess potentially simple regulatory pathways, simpler cell geometry and powerful tools for genetic analysis (Mitchison, 2003).

1.7.1 Mechanism of cell size control in *S. cerevisiae*

S. cerevisiae has been used extensively as a unicellular model to study the molecular mechanisms that governs cell size, due to two most important properties concerning cell size control. Firstly, a constant mass to ploidy ratio, and secondly, a critical size threshold required for cell cycle progression (Ferrezuelo et al., 2012).

In *S. cerevisiae*, a step during G1phase known as ‘start’ (transition from G1-S) is considered to determine the size control (Johnston et al., 1977), which involves the G1 cyclin Cln3 (Rupes, 2002; Turner et al., 2012). Cln3 acts as an activator of ‘start’ in a dose dependent manner (Nash et al., 1988; Tyers et al., 1993), and is degraded quickly (Tyers et al., 1992). Cln3 targets the transcription factor SBF and its

inhibitor Whi5. It is thought that it activates the SBF by Cln3-CDK-mediated inhibition of Whi5 (Costanzo et al., 2004; de Bruin et al., 2004). Previously, Jorgenson et al. (2002), in a systematic screen to identify mutants with varying cell size, showed that the mutation that affect ribosome assembly also affect the threshold for cell size at division. This finding suggested that regulators of ribosome synthesis, such as Sfp1 and Sch9, also play a role in cell size regulation (Jorgensen et al., 2002). The mechanism that links ribosome biogenesis with the control of the cell size threshold is not yet clear. Ydj1 chaperone dependent release of Cln3 from the periphery of the rough endoplasmic reticulum has been suggested to be an important factor in regulating the threshold size as a function of the individual growth rate in response to ribosome biogenesis (Ferrezuelo et al., 2012) (figure 1.6).

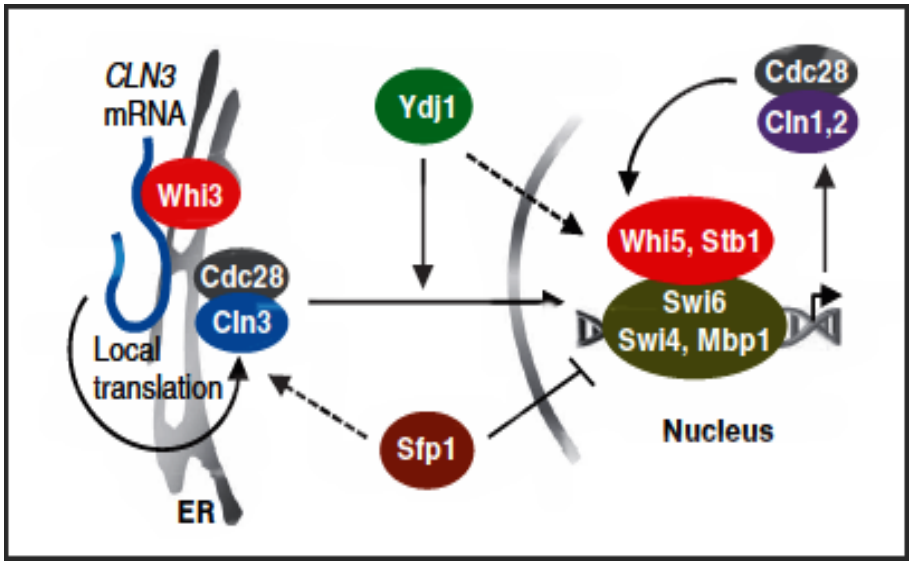


Figure 1.6: Schematic of cell size control in budding yeast. Cln3 and Cdc28 phosphorylate Whi5 to activate SBF and MBF TFs, and induce G1-S regulators; whereas Whi3 and Ydji1 regulate the release of Cln3 from ER. Taken from (Ferrezuelo et al., 2012)

This study tried to explain the relationship between the role of ribosome biogenesis and growth rate by proposing that higher levels of chaperones will be incorporated for ribosome synthesis, and synthesis and translocation of proteins to meet the requirements of increased growth rate. This sudden demand of chaperones for these

processes will result in an overall decrease in their availability for other processes, thus causing a release of Cln3 from the endoplasmic reticulum that will in turn activate cell cycle entry (Ferrezuelo et al., 2012).

1.7.2 Mechanism of cell size control in *S. pombe*

In *S. pombe*, activity of Cdc2 kinase determines the length of the G2 phase (Nurse, 1990), and inhibitors of Wee1, Cdr1 and Cdr2 are involved in the coordination of cell growth and cell division (Moseley et al., 2009). A gradient of Pom1 kinase originating from the cell tips is required for inhibition of G2 - M entry, until cells have reached a threshold length (Martin and Berthelot-Grosjean, 2009; Moseley et al., 2009). As cells are rod shaped in fission yeast, their length correlates with cell size. Pom1 is associated with the cell membrane and its autophosphorylation results in its dissociation from the membrane, thus generating a Pom1 gradient such that it has maximum concentration at the tips and minimum in the cell centre (Hachet et al., 2011). When cells reach their maximum length, the concentration of the Pom1 decreases at the centre resulting in the inhibition of Wee1 by Cdr1 and Cdr2, resulting in entry to M phase (Martin and Berthelot-Grosjean, 2009; Moseley et al., 2009) (Figure 1.7). In this way the fission yeast cells sense alterations in the size of the cell and thus regulate the mitotic entry.

Although, the models mentioned above provide some clues about how cells sense their size and regulate cell division, still these pathways cannot account for the whole story of cell size regulation, especially when taking into consideration that the increase in cell size at division is almost proportional with the increase in ploidy. Cells, therefore, must have some mechanism to monitor their ploidy and then incorporate the information obtained into the cell size sensing mechanism (Marshall et al., 2012).

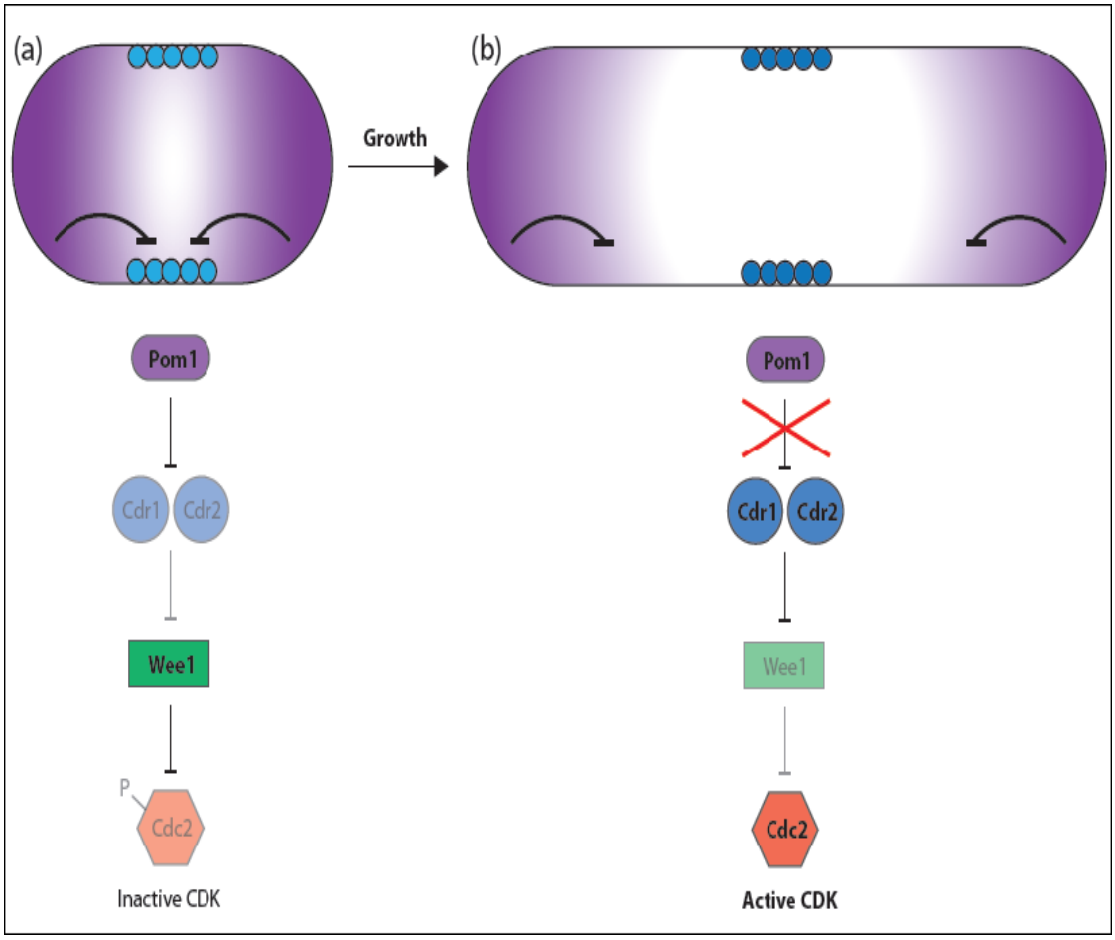


Figure 1.7: Schematic of cell size control in fission yeast. Pom1 gradient keeps a check on size threshold. At start of G2, Pom1 (purple) originates from cell tips and deactivates Cdr1 and Cdr2 that activate Wee1 to phosphorylate Cdc2 (a). When cell grows up to a certain length, the Pom1 concentration decreases in the middle, thus activating Cdr1/Cdr2 to lift the inhibition from Cdc2 that initiates cell division. Light and dark colours represent the active and inactive states of proteins, respectively. Taken from (Marshall et al., 2012).

1.8 Polyploidy-a route to-aneuploidy

Many mammalian cells and tissues become polyploid due to a requirement of differentiation and proliferation processes during development, but unscheduled polyploidization is not tolerated in vertebrates. For example, triploidy and tetraploidy have been estimated to account for about 20% of all miscarriages (Hassold et al., 1980; Eiben et al., 1990; Neuber et al., 1993). Karyotyping of many tumors have revealed the presence of highly variable chromosomal content in cancer cells ranging from hypo-diploidy (slightly less than 46) to tetraploidy and hyper-tetraploidy (more than 200). Polyploid cells found in almost all stages of tumors show higher levels of chromosomal loss and gain, known as 'chromosomal instability' (Lengauer et al., 1997), which results in aneuploidy. Chromosomal instability is a common feature of tumors. Chromosomal instability refers to the rate at which cells undergo chromosomal loss and gains, i.e. the rate of karyotype change, while aneuploidy refers to the state of having an abnormal karyotype. Chromosomal instability gives rise to aneuploidy, but not all aneuploid cells show a feature of karyotype instability as some aneuploids stay stable (Gordon et al., 2012). Diploid cells with chromosomal instability have less chances of survival and die sooner, whilst tetraploid cells, due to higher chromosomal content, can follow multipolar mitosis to give rise to a viable progeny of aneuploid cells (Storchova and Pellman, 2004).

The exact mechanisms by which tetraploid cells promote chromosomal instability and aneuploidy is not known yet, but for vertebrates the presence of multiple centrosomes have been anticipated as a cause of this instability (Nigg, 2002; Boveri, 2008). These multiple centrosomes can form multipolar spindles, resulting in abnormal chromosomal segregation (figure 1.8). This whole process can ultimately lead to a multipolar mitosis that is a cause of chromosomal instability (Gisselsson et al., 2008). Ganem et al. (2009) have shown that multiple centrosomes also increase the occurrence of merotelic attachment, where a single kinetochore can attach to the microtubules arising from two poles of spindle, resulting in erroneous chromosomal segregation (Ganem et al., 2009). Presence of multiple centrosomes and merotelic attachments have been reported for several human tumours and are considered to be a cause of chromosomal instability (Gordon et al., 2012).

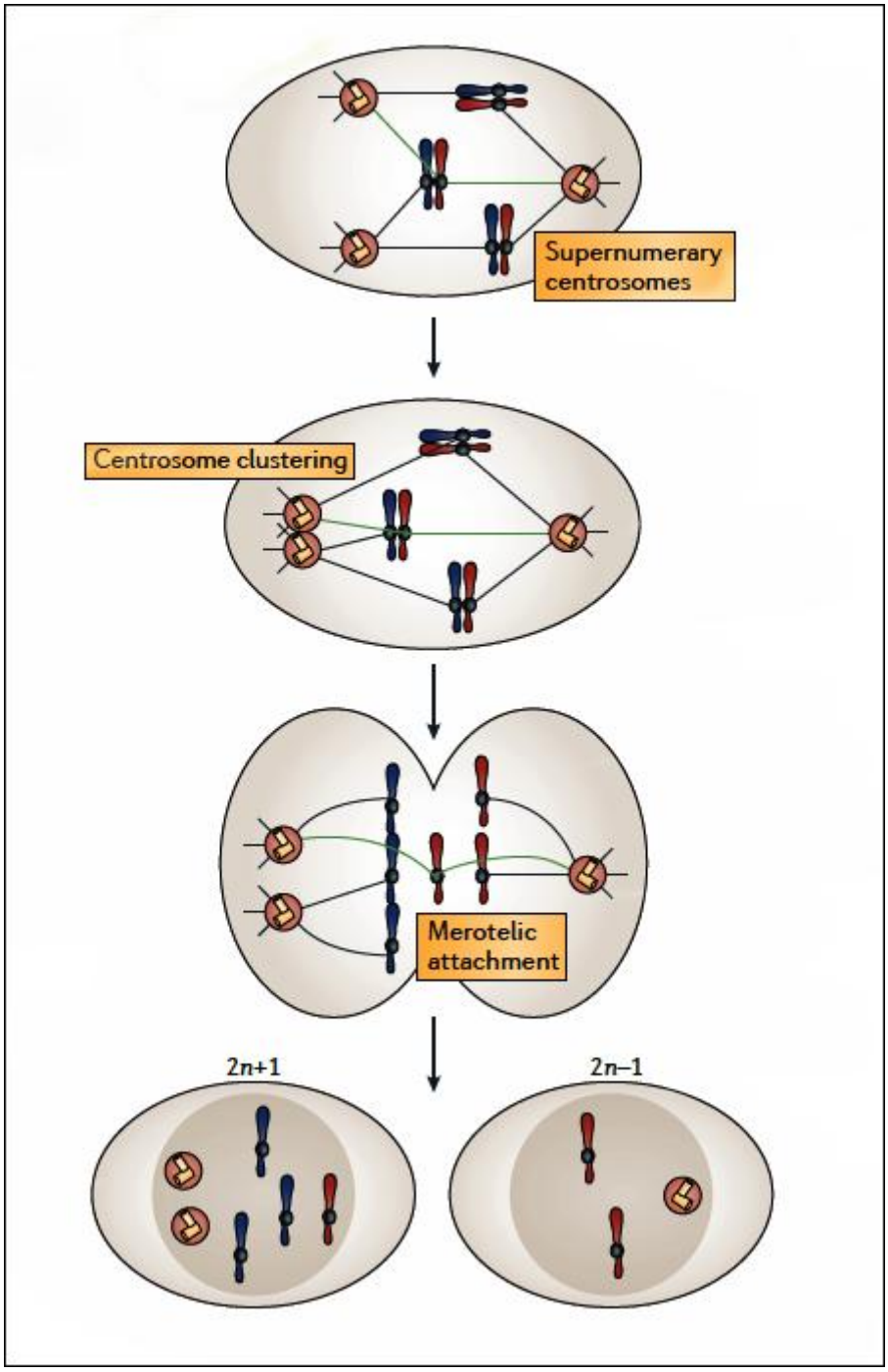


Figure 1.8: Schematic representation of aneuploid formation. Multiple centrosomes can give rise to chromosomal segregation defects, either by centrosome clustering or by merotelic attachment. Adapted from (Gordon et al., 2012).

1.8.1 Detrimental effects of aneuploidy

Aneuploidy – the presence or absence of one or more extra chromosome, is lethal in many organisms such as humans, mice, *Drosophila* and *C. elegans* (Torres et al., 2008). Here we review some studies that have reported the effects of aneuploidy on various aspects of cellular morphology and physiology.

Mouse

William et al. (2008) studied the effects of aneuploidy in cell lines derived from mouse embryonic fibroblasts that showed trisomy for four different chromosomes, namely chromosomes 1, 13, 16 and 19. Except for the mouse embryo that was trisomic for chromosome 19, all the others died. Aneuploid embryos showed slow growth, reduced size and developmental defects (figure 1.9). It was observed that cell proliferation was severely affected in trisomic cells as compared to the diploid cells. Trisomic cells showed larger size and altered metabolism that could affect the physiology of these cell lines (Williams et al., 2008).

Humans

Aneuploidy is a major cause of frequent abortions, developmental impairments and is observed in more than 90 % of tumours in humans (Hassold et al., 1980, 1996; Weaver and Cleveland, 2006; Brown, 2008). In humans, aneuploids that are trisomic for chromosome 21 provide the only viable examples, but trisomy of this autosomal chromosome results in Down's syndrome. 5 to 10 % of individuals exhibiting autosomal trisomy of chromosomes 13 and 18 can also survive to birth but mostly they die very young (Rasmussen et al., 2003). *In vitro* studies have shown that cell lines trisomic for chromosome 21 exhibit slow growth, increased protein content and have defective cell proliferation as compared to diploid cells (Segal and Mccoy, 1973). It has been found that these three human trisomic aneuploids (for chromosome 13, 18, and 21) are tolerated *in utero*, because these chromosomes contain very few protein coding genes and thus they cause fewer disturbances in net gene dosage.

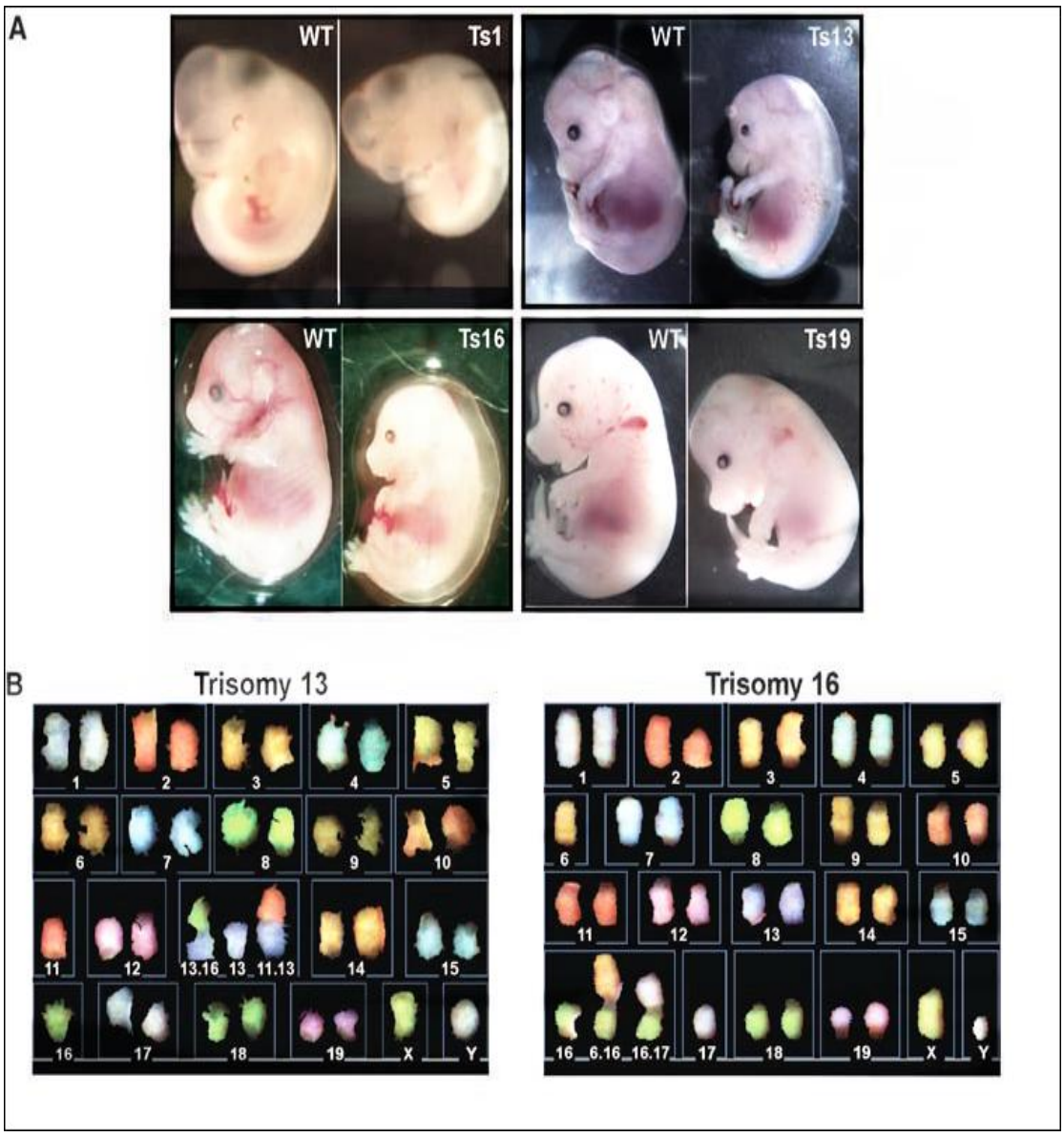


Figure 1.9: Growth defects in aneuploid mouse embryos and karyotype of aneuploids. Photographs showing severe developmental abnormalities associated with trisomy (Ts) of chromosome 1, 13, 16 and 19 (a), while (b) shows karyotype for Ts 13 and 19. Taken from (Williams et al., 2008).

Fission yeast

Aneuploidy in *S. pombe* causes growth arrest along with chromosomal instability (Niwa et al., 2006). It was found that culture conditions severely affect the sustainability and growth behaviour of aneuploid strains. Many individual spores of aneuploid strains converted to euploids via irregular cell divisions that in many instances resulted in cell death. Many aneuploids of this organism also showed cell cycle arrest. Aneuploids that are disomic for chromosome I and II are highly unstable and they undergo structural modification to form a circular chromosome (Niwa et al., 2006).

Budding yeast

Torres et al. (2007) conducted a study to observe the differences in growth, cell size, metabolism and cell cycle progression in budding yeast where they used the chromosomal transfer strategy to generate aneuploid cells with one extra chromosome. These aneuploids showed slow growth due to delayed G1 stage of the cell cycle as compared to haploid wild type cells. Aneuploid yeast cells demonstrated altered metabolism with higher uptake of glucose and increased sensitivity to different drugs that target protein synthesis (Torres et al., 2007).

In another study, aneuploid strains of *S. cerevisisae* showed defects in chromosomal recombination during mitosis, defective DNA repair pathway and whole chromosome loss. These strains had increased rate of accumulation of point mutations along with aneuploidy induced increase in chromosomal instability as compared to wild type haploid cells (Sheltzer et al., 2011).

Recently, Oromendia et al. (2012) have shown that aneuploidy in *S. cerevisiae* affects protein homeostasis. All aneuploid strains exhibited aggregation of endogenous and ectopic proteins. Although, they could not find particular proteins involved in aggregate formation, protein quality control systems comprised of proteasome and Hsp90 chaperone was severely impaired, thus leading to reduced protein quality control and protein aggregation. Aggregation of proteins was decreased when proteasome function was increased. It might be possible that excessive production of proteins, due to the presence of an extra chromosome, makes

use of all available chaperones, leaving none or very few for the correct protein folding process (Oromendia et al., 2012).

These studies have not only demonstrated the detrimental nature of aneuploidy but have also provided the basis of mechanisms that lead to growth and proliferation defects in aneuploid yeast cells as compared to normal haploid cells.

1.8.2 Benefits of aneuploidy

Even though the presence of an extra chromosome can result in defective growth and chromosomal segregation as mentioned above, a single gene present on this extra chromosome can provide adaptive advantages under unfavourable conditions.

In a study conducted in *Candida albicans* aneuploids, containing an isochromosome consisting of two copies of left arm of chromosome 5, it was found that these aneuploids are resistant to fluconazole. This resistance was established to be associated with the presence of this isochromosome (Selmecki et al., 2006). Later they found that the two genes ERG11 (ergosterol biosynthesis) and TAC1 (transcriptional regulator of drug efflux pumps) present on the isochromosome are involved in this drug resistance. These genes act independently but additively in a copy number dependent manner to provide drug resistance to aneuploid strains of *C. albicans* (Selmecki et al., 2008).

Similar observations were made by Pavelka et al. (2011) when they conducted a study to compare phenotypic profiles of aneuploid strains of *S. cerevisiae* to euploid controls. When observed for growth after exposing them to various stressors and drugs, they found that several aneuploid strains grew better than control strains at conditions that were not suitable for euploid control strains, such as in the presence of rapamycin, bleomycin or flucanzole. Also, aneuploid strains with similar karyotypes exhibited similar growth rates under different stress conditions. Analysis by whole genome sequencing showed that no other mutations had occurred in these aneuploid strains; therefore the improved growth under stress conditions was conferred by the presence of an extra chromosome in these aneuploid strains. Further, they also showed that the presence of an extra copy of ATR1 gene on

chromosome XIII provides resistance against 4-nitroquinoline-1-oxide (tumorigenic compound) to aneuploid strains disomic for chromosome XIII (Pavelka et al., 2010).

These studies show that the presence of even one or a few beneficial genes on an extra chromosome in aneuploids can result in a novel phenotype**,** which can help these organisms to adapt to adverse growth conditions.

1.8.3 Effects of aneuploidy on gene expression regulation

Aneuploidy affects the genome by altering the copy number of several genes and can result in novel phenotypes. These novel phenotypes can be a result of the collective changes of copy number variations on global transcription and translation of aneuploid cells. The presence of an extra chromosome has been shown to affect the transcriptional response in many organisms as mentioned below.

Arabidopsis thaliana

In a study of transcription regulation in response to aneuploidy, *A. thaliana* was used with an extra copy of chromosome 5. This study showed that the presence of an extra chromosome in this organism results in an increase of mRNA levels in a dosage dependent manner. Interestingly, they also found differential expression for mRNA levels of many other genes located on four other chromosomes. The majority of these up-regulated genes were found to be stress responsive. Also, genes under the control of the same transcription regulator showed different mRNA levels in response to trisomy. These changes might be the cause of: alteration in gene dosage for those encoding regulatory proteins; sensitivity to these alterations; and response of silencing mechanism (Huettel et al., 2008).

Ts65Dn mouse

The Ts65Dn mouse is used as a model for studies of Down Syndrome in humans, as it contains an extra copy for the orthologs of about 142 genes present on a chromosome 21 of humans. Also, this mouse model shares several developmental abnormalities of Down syndrome. A comparison of transcript levels was made for 136 genes that are mouse orthologs of chromosome 21, in 9 tissues of the trisomic

and euploid mice (Kahlem et al., 2004). Almost all genes which were at dosage imbalance in Ts65Dn showed higher mRNA levels in all the tested tissues. Many genes showed altered expression, independently of gene dosage, suggesting that their expression might be dependent on some other regulatory mechanisms present in the trisomic area. These findings were in complete agreement with another study which made use of the same mouse model (Lyle et al., 2004).

Budding yeast

Microarray based studies of gene expression regulation in response to aneuploidy, in budding yeast, have shown that expression of genes increases or decreases in response to gene dosage (Torres et al., 2007; Pavelka et al., 2010; Sheltzer et al., 2012). Torres et al. (2007), by using 16 aneuploid strains, found that genes contained in the disomic region showed 2 fold increases in gene expression, indicating that transcription of genes corresponds to the DNA copy number. Many aneuploid strains also showed an increase in expression of those genes involved in the stress response (ESR) that might be the response of the cell to defective proliferation (Torres et al., 2007). These ESR genes were also found to be induced in a later study (Pavelka et al., 2010). Torres et al. (2007) showed that these ESR genes are regulated by defective growth, as many aneuploid strains showed prolonged G1. They normalised this effect by growing all strains in a chemostat in phosphate limited media. When they compared the transcriptome profiles of aneuploids after normalising growth rate, they found functional categories that showed higher expression and which were enriched for ribosomal biogenesis genes and nucleic acid metabolism. While genes that showed decrease in gene expression were enriched for carbohydrate metabolism functions (Torres et al., 2007).

Both studies used different approaches for the generation of aneuploid strains in budding yeast and used different growth conditions, which could explain why the latter study could not find genes enriched for ribosomal biogenesis and nucleic acid metabolism.

All the above mentioned examples clearly indicate that expression levels of genes are proportional to the DNA copy number in an aneuploid organism. Some other

genes also show altered expression levels independent of the copy number changes. Altered expression of these genes could be a result of some regulatory changes caused by aneuploidy. All these findings suggest that aneuploidy results in altered gene expression patterns, as compared to normal haploid or euploid individuals. These changes are also manifested by impaired cellular growth and lowered fitness.

1.9 Fission yeast as a model organism to study the impact of ploidy on gene regulation

Fission yeast (*Schizosaccharomyces pombe*) is a unicellular eukaryote and serves as an attractive model organism for studying various aspects of cellular physiology. Its genome is about 14.1 Mbp; it is fully annotated and is estimated to contain ~5000 protein-coding genes. A wealth of information is now available about its physiology, molecular biology and genetics. Due to its short cell cycle (2-4 hours), easy handling and rather simple genome (presence of only three chromosomes, 2 large and 1 small), it is amenable to precise genetic manipulations. *S. pombe,* under normal growth conditions, stays haploid during most of its life cycle, but upon nitrogen starvation, cells of opposite mating type can fuse with each other to form a diploid zygote (figure 1.10). These diploid cells undergo meiosis and acquire a normal haploid life cycle as soon as they are returned to normal media (Molnar and Sipiczki, 1993).

Polyploid strains of fission yeast can be generated by the process of protoplast fusion. In a previous study, ploidy of fission yeast has been increased up to pentaploidy by using this technique (Molnar and Sipiczki, 1993). Polyploid strains generated by crossing cells of either opposite mating types (h^+/h^-) or homothallic (h^{90}/h^{90}) strains are very unstable and accumulate cells of lower ploidies over generations. Whereas, polyploid cells generated by protoplast fusion of cells of the same mating type (h^+/h^+ or h^-/h^-) stay stable and can propagate vegetatively. Similar to higher eukaryotes, polyploid cells in *S. pombe*, also exhibit larger cells size and poor tolerance to higher ploidies (Molnar and Sipiczki, 1993). Therefore, tetraploid cells frequently accumulate cells of lower ploidies, a characteristic feature of polyploid cells in mammals and especially of cancer cells. Although, many phenotypic characteristics of fission yeast polyploid cells are similar to that of higher

eukaryotes, how they affect the transcriptional regulation in this organism was still unknown.

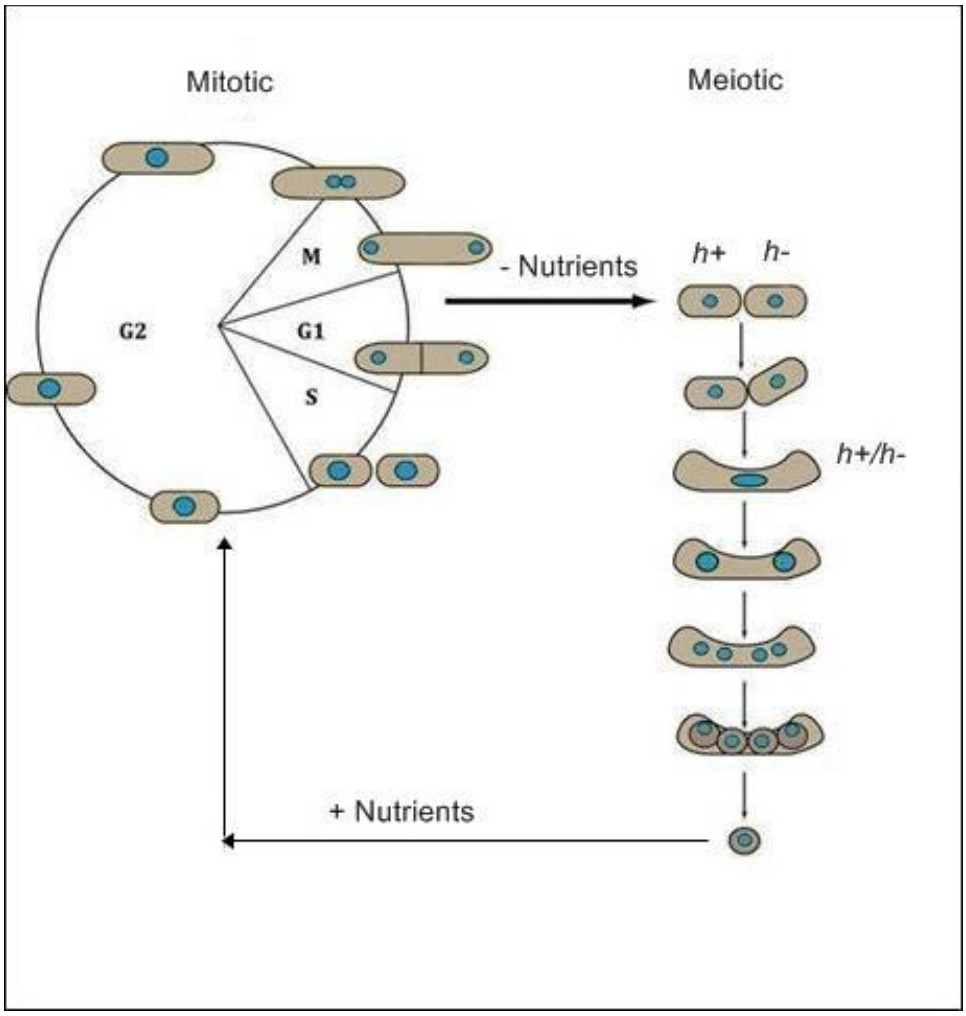


Figure 1.10: Schematic representation of the fission yeast cell cycle. In normal growth conditions, cells undergo the mitotic cell cycle and stay haploid. Under nutrient limited conditions, cells of opposite mating type mate to form diploid cells that then undergo the meiotic cell cycle. Diploid cells thus become haploid spores during meiosis, which can germinate and proliferate again once nutrient conditions are returned to normal.

In addition to polyploid strains, aneuploid strains have also been reported in *S. pombe* (Niwa et al., 2006). Aneuploid strains in this organism contain a smaller portion of any of the three chromosomes (minichromosome) rather than a full extra chromosome; therefore they are referred to as partial aneuploids at many instances in this thesis. Aneuploid strains that are disomic for chromosome I and II are shown to be unstable and can be maintained only when selected for different markers. These aneuploids become stable only after various structural rearrangements in the extra chromosome that results in a circular minichromosome. Fission yeast cells that are disomic for a smaller portion of chromosome III are the only stable and maintainable aneuploids that can be propagated easily (Niwa et al., 2006).

In a microarray based study, Chikashige et al. (2007) showed that in aneuploid strains the expression patterns of genes contained in the disomic or trisomic regions correlates with the DNA copy number. Many genes in the monosomic regions, however, also exhibited altered expression, which they attributed to the increased expression levels of genes contained within the minichromosome. This study has shown that binding of Swi6 was decreased in the telomeric regions of chromosome I and II, which resulted in increased expression levels of genes in these areas (Chikashige et al., 2007). Although, this study has demonstrated the role of aneuploidy on the regulation of transcription in *S. pombe*, it could not find an enrichment of any particular functional category of genes that could explain the altered expression of genes in a monosomic area other than on telomeres.

The presence of both polyploid and aneuploid strains makes *S. pombe* an ideal organism to study the effects of ploidy on regulation of gene expression by using microarrays.

1.10 Aims and objectives

1. A wealth of information is now available about defects caused by aneuploidy on cellular growth, developmental impairments and adaptive progression. However, it still remains to be fully explored how aneuploidy affects gene expression and thus cellular fitness. Previous studies in higher eukaryotes and two yeast models have shown that expression levels of genes scale with copy number, but how expression of genes in monosomic regions is altered remains unresolved. Therefore the aim was to further explore the role of aneuploidy on fission yeast gene expression and the differential regulation of genes contained in monosomic regions. The objectives of this study were:

a) To establish the effects of aneuploidy on fission yeast gene expression.
b) To analyse functional characteristics of any particular gene category showing differential regulation in monosomic regions.

2. Despite the presence of large number of polyploid cell types, how ploidy affects the cellular physiology is still not well understood. Although a recent study in *S. cerevisiae* has established the effect of altered cell size rather than polyploidy on differential regulation of gene expression, previous studies in higher eukaryotes and yeast could not find any such relationship. Therefore, it would be interesting to use *S. pombe* to explore the effect of polyploidy on transcription regulation. Therefore we set out to:

a) To generate a series of strains that is stable and differs only in their ploidy.
b) To quantify the effect of polyploidy on cell size.
c) To determine the relationship between polyploidy and global RNA levels, including both total and mRNAs.
d) To determine effects of polyploidy on the control of gene expression
e) To explore the relationship between polyploidy, gene expression, and cell size by following up selected genes that are differentially regulated as a function of ploidy and/or cell size.

Chapter 2

2 Materials and Methods

This chapter explains the nomenclature used to denote fission yeast gene and protein names. Then, I describe methods and materials used to construct polyploid strains required to generate the results in chapters 4 and 5. It also contains cell biological, genetic and molecular biology related techniques used in this thesis to achieve the results described in coming chapters. List of strains used in this study is provided in Appendix I, while all the primers used in this thesis are mentioned in Appendix II.

2.1 Yeast gene and protein nomenclature

In this thesis the nomenclature used for genes and proteins is based on (Kohli, 1987). A gene that has got its common name is presented by a three-letter-code followed by a specific number, and written in lower case italics (e.g. *ade6*). Sometimes, to emphasise a wild type gene, a '+' is added at the end of a gene name (e.g. *ade6+*). Mutant alleles are specified by the gene name followed by the allele number or a combination of letters and numbers (e.g. *ade6-210* or *ade6-M210*), where allele is separated from the gene name by a hyphen. Mating phenotypes are presented as italicized following the genotype of that strain. For example, h^{90} represents homothallic strains, h^{+} heterothallic plus strains, and h^{-} heterothallic minus strains (e.g. *ade6-M210* h^{-}). Deletion of a specific gene is marked with a 'Δ' before the gene name (e.g. Δ*ade6*). To indicate a replacement of a particular gene with another, the original gene name is followed by two colons and name of a gene that substitutes it (e.g. *ade6::ura4*). To distinguish alleles that confer resistance and sensitivity, superscripts r and s, respectively are added to the gene name, e.g. Kan^{r}. *S. pombe* plasmids are designated by non-italic letters and they always start with a lower case p, e.g. pFA6. When a gene is fused to another DNA sequence, then the name of that individual DNA sequences appear in the genome with a hyphen separating the two sequences. For example, *ade6-GFP-kanMX6*, that means, *ade6* gene is C-terminally fused to a *GFP* gene followed by the kanamycin (Kan) resistance gene (*kanMX6* is a selective marker).

A protein encoded by a particular gene is represented with the same three letters and number code in a regular style (non-italic) but the first letter is written in upper-case (e.g. Cbp1).

2.2 Propagation and storage of fission yeast strains

All strains used were maintained in solid or liquid Yeast Extract plus Supplements (YES). Selection of diploids was performed on Malt Extract Agar (MEA) or YE with low adenine (YEA) media, while for tetraploids SD agar plates were used. All cultures were grown in YES or Edinburgh Minimal Media (EMM) at 32°C (unless specified otherwise) and shaken at 130 rpm. Cells were normally grown to optical density (OD_{595}) of 0.5 corresponding to $5 \times 10^6 – 6.5 \times 10^6$ cells/ml that corresponds to mid log-phase. For long term storage strains were stored at -80 °C in freezing mix.

YES: For 1 Litre: 5 g Difco Yeast Extract, 30 g glucose, PH 5.6. Supplements: 225 mg of each of histidine, leucine, adenine, uracil and lysine.

YES Agar: For 1 Litre: 5 g Difco Yeast Extract, 30 g glucose, PH 5.6. Supplements: 225 mg of each of histidine, leucine, adenine, uracil and lysine, 2% agar (w/v)

MEA: Bacto-malt extracts 50 g, 225 mg of each of adenine, histidine, leucine and uracil, 20 g/l agar.

SD agar: 0.67% Difco Yeast Nitrogen base without amino acid, 2% dextrose (glucose), 2% agar (w/v).

Freezing mix: 50 % (w/v) YES, 50 % (v/v) glycerol

EMM: For 1 Litre: 32.3 g EMM broth, 225 mg of adenine, 5 g/l NH4Cl

EMM Agar: For 1 Litre: 32.3 g, 225 mg of adenine, 5 g/l NH4Cl, 2% agar (w/v).

2.3 Selection of diploid strains

2.3.1 Type 1: Prototrophic diploids

To obtain prototrophic diploids, two strains auxotrophic for adenine (*ade6-M210* h^+ and *ade6-M216* h^-) were streaked on MEA plate and incubated at 32 °C for 1 day. After 1 day, I streaked the loop full of cells on YEA plates and selected for prototrophic diploids that formed white colonies.

2.3.2 Type 2: Auxotrophic diploids

To obtain the diploids that are auxotrophic for adenine, *ade6-M210* h^- and *ade6-M216* h^- were plated separately on YES agar containing phloxin B and incubated at 32°C for three days. Dark red colonies were selected and streaked on to YES plates, incubated for three days and examined under the microscope for cells size. Ploidy was confirmed by FACS analysis.

YES phloxin B plates: YES agar + 5 mg/l phloxin B (Sigma).

2.3.3 Construction of tetraploids by protoplast fusion

Tetraploids were generated by using a protocol for protoplast fusion except for below mentioned changes. Diploids *ade6-M210* h^-*/ade6-M210* h^- and ade6-*M216* h^-*/ade6-M216* h^- selected as above were used to make tetraploids.

10 ml of both diploid strains grown to OD_{595} 0.5 were harvested by centrifugation at 3000 rpm. Cells were washed once with dH_2O and re-suspended in 0.5 ml buffer A. After mixing equal cell numbers of both diploids to final volume of 1 ml, 0.3 ml glyconyx solution was added and incubated at 30°C with rotation for 5-6 hrs. Samples were then centrifuged at 2000 rpm for 1 min, and the pellet was re-suspended in 1 ml of buffer B. After washing twice with buffer B, 0.3 ml PEG buffer was added to pellet and incubated at 30°C for up to 3 hrs. Supernatant was discarded and cells were plated on agar plates (containing 0.67 5 SD and 1 % glucose) and incubated at 32 °C for 5 days. Tetraploid cells (*ade6-M210* h^- *ade6-M210* h^-*/ade6-*

M216 h⁻ade6-M216 h⁻) appearing as white colonies were selected and streaked on YEA plates at 32 ℃ for 3 days.

Buffer A: 1.2 M sorbitol, 50 mM citrate phosphate (pH 5.6)

Glyconyx solution: buffer A + 5mg/ml glyconyx enzyme

Buffer B: 1.2 M sorbitol, 30 mM Tris-HCl (pH 7.6)

PEG buffer: 30% PEG 4000, 10 mM Tris HCl pH 7.6, 10 mM $CaCl_2$

2.4 Fission yeast molecular genetics

2.4.1 PCR based gene deletions

The one step PCR-based approach was used for deletion of specific genes in fission yeast (Bähler et al., 1998). The PCR primers used were 100 bp long (see appendix for the list of primers used). Each primer contains an 80 bp gene specific sequence on its 5′-end and 18 bp sequence homologous to plasmid multiple cloning site on its 3′ end. DNA fragments were amplified using LongExpand PCR template system (Roche) and pFa6-KanMX6 as a template in 96 well plates. PCR products were checked on 1% (w/v) agarose gels before proceeding. The resulting PCR product, approximately of 1.6 Kb in size contains is the Kan cassette flanked by regions homologous to the gene of interest. When fission yeast is transformed with such a construct, Kan cassette replaces the gene whose flanking sequences are next to the Kan^r cassette, due to homologous recombination, resulting in a specific gene deletion. Once checked on agarose gel, products from four PCR reactions were pooled together, precipitated (to clean up) by adding 1/10 volume of 3M sodium acetate (NaAc) and 2.5 volumes of ice-cold absolute ethanol. After, incubating at -20°C for at least 30 minutes, DNA was pelleted at 13,000 rpm for 15 minutes, washed with 70 % (v/v) ethanol and dried at room temperature. Pellet was then resuspended in 15-20 µl dH_2O, and the resulting concentrated DNA was used directly for the transformation of fission yeast cells as described (section 2.4.4). As PCR ingrediants and PCR programme used for gene deletion, C- terminal and over expression were the same, they are mentioned separately in tables 2.1 and 2.2

respectively. The PCR reaction was carried out on a PTC-225 thermal cycler (MJ Research).

2.4.2 Endogenous C-terminal tagging

Endogenous C-terminal tagging of yeast proteins was performed by 1-step PCR tagging (Bähler et al., 1998). Forward primers were designed such that they contain approximately 80 bp of sequence homologous to 3′ end of the gene of interest before its STOP codon, followed by an 18 nucleotides long sequence, homologous to the respective vector. The reverse primer consisted of an 80 bp sequence homologous to the 3′ UTR region of the gene, followed by a 20-mer sequence that facilitated priming to the vector. The resulting PCR product was flanked by regions homologous to the gene of interest, facilitating the in-frame integration of an epitope containing Kan cassette within the 3′ end of the desired gene. Rest of the procedure for pooling and clean-up is same as mentioned above in section 2.4.1. This purified and concentrated PCR product was used directly to transform fission yeast cells as mentioned in section 2.4.4.

2.4.3 Overexpression of specific genes

Genes can be overexpressed by placing their ORF under the control of a regulatable promoter, which we achieved by using one-step PCR based approach (Bähler et al., 1998). The plasmid construct used as a template contains an *S. pombe* regulatable promoter known as P3nmt1 that can be repressed or induced, respectively, by addition or removal of thiamine in the media. The Forward primer contained an 80 bp sequence homologous to the 5′ UTR of the gene and a stretch of sequence homologous to the plasmid multiple cloning site on its 3′ end. The Reverse primer contains an 80 bp sequence showing homology to the 5′ end of the specific gene after its START CODON and 23 bp sequences priming to plasmid. PCR ingredients and programme used to prepare the constructs are mentioned in tables 2.1 and 2.2, while the rest of procedure was the same as the one mentioned in section 2.4.1.

Table 2.1 PCR cocktail for gene deletion, C-terminal tagging and N-terminal tagging

Ingrediants	**Amount**
pFA6a-KanMx6 template (25 ng/µl)	1 µ
10 X long expand buffer 3	5 µl
dNTPs (10 mM)	2 µl
Forward primer (10 µM)	1 µl
Reverse primer (10 µM)	1 µl
Polymerase mix	1 µl
H_2O	38 µl

Table 2.2 PCR programme for gene deletions, C-terminal and N-terminal tagging

Step	**Temperature**	**Time**
1	94ºC	5 minutes
2	94ºC	1 minute
3	55-57ºC (depending on Tm of a primer)	2 minutes
4	72ºC	2-3 minutes (depending on the length of a PCR product)
5	go to step 2 and repeat 7 times	
6	94ºC	1 minute
7	55-57ºC (depending on Tm of a primer)	2 minute
8	72ºC	2-3 minutes (depending on the length of a PCR product)
9	go to step 6 and repeat 30 times	
10	72ºC	7 minutes
11	4ºC	forever

2.4.4 *S. pombe* transformation

20 ml of mid-log phase growing cells were harvested by centrifugation (8000 rpm for 4 minutes) and washed once with distilled water. The resulting cell pellet was washed with 1 ml of LiAc/TE buffer before resuspending in a final volume of 100 µl LiAc/TE. Up to 10 µg of PCR product was added to the cell suspension and mixed carefully. After 10 minutes of incubation at room temperature, I added 280 µl of LiAc-TE-PEG and incubated at 30°C for 1-2 hour without shaking. 48 µl of DMSO were then added to the cell suspension and heat-shocked at 42°C for 5 minutes. Cells were pelleted by centrifugation for 15 seconds at 4000 rpm at room temperature. After washing once with distilled water, cells were resuspended in 200 µl of YES, plated onto YES agar and incubated at 32°C over night. In case of selection by kan^r and Nourseothricin (Nat^r), I replica plated these cells onto YES agar plates, containing either 100 µg/ml Geneticin 418 (Gibco) or 100 µg/ml nourseothricin (Roche). For selection of the auxotrophic transformants, EMM agar plates lacking the appropriate amino acid were used for selection. Transformed fission yeast cells having correct integration of the PCR cassette at right place were selected by colony PCR analysis (section 2.7.1).

LiAc-TE: 0.1M lithium acetate, 10 mM Tris/HCl pH 7.5, 1 mM EDTA

LiAc-TE-PEG: 0.1M lithium acetate, 10 mM Tris/HCl pH 7.5, 1 mM EDTA, 40% PEG4000

2.5 Fission yeast microscopy

Cells were visualised under the microscope after staining with DAPI and calcofluor as described below. To perform microscopy, cells were fixed either by adding ethanol or formaldehyde. For fixation with ethanol, 1 ml of exponentially growing cells was centrifuged at 4000 rpm for 5 minutes. 1 ml of cold 70% ethanol was added drop by drop to the pellet, while vortexing the cells slowly, and stored at 4°C. For cell length and width measurements, cells were fixed using a formaldehyde solution. For this fixation, 1/10 volume of 37% (v/v) formaldehyde solution was added to 1 ml of a culture and mixed well. After 15 minutes, cells were pelleted by

centrifuging at 8000 rpm for 1 minute and washed twice with Phosphate buffer saline (PBS). These cells were stored at 4°C after resuspending in 100 µl of PBS.

2.5.1 DNA visualisation (DAPI staining)

For DNA visualisation cells were stained with DAPI. To stain with DAPI, 100 µl of the ethanol fixed cells were washed twice with PBS. The cell pellet was resuspended in PBS for 1-2 hours and then centrifuged at 8000 rpm for 1 minute. 50 µl of pellet was resuspended in distilled water containing DAPI at 1 µg/ml final concentration. 2 µl of cell suspension was then placed on a slide, sealed with a cover slip and visualized by fluorescence microscopy. Images were recorded using a Hamamatsu digital camera C4742-95 fitted to a Zeiss Axioskop microscope with plan-Apochromat X 63 1.25 oil objective.

PBS for 1 litre: 8 g NaCl, 0.2 g KH_2PO_4, 1.44 g Na_2HPO_4, 0.2 g KCl, pH 7.4

2.5.2 Septum visualisation (Calcofluor staining)

For septum and cell wall visualisation, formaldehyde fixed cells were centrifuged at 8000 rpm for 1 minute. After washing twice with PBS, cells were pelleted by centrifuging at 8000 rpm for 1 minute. After resuspending in PBS, 2 µl of cell suspension was mixed with 1 µl of a 10 mg/ml calcofluor solution.

2.5.3 Cell Length and width measurement

For measurements of cell length and width, cells were stained with calcofluor. Images were taken by using a 63 X objective under Zeiss microscope. The objectJ plugin of ImageJ software (NIH) was used to measure the cell length and width by drawing a line from one end to the other of a cell and converting length of line from pixels to µm.

2.6 Fission yeast cell biology

2.6.1 DNA content measurement

For DNA content measurement, samples were prepared using Sytox Green staining. 50 µl of ethanol fixed cells were washed twice with 10 mM EDTA and centrifuged at 4000 rpm for 3 minutes. Pellet was then resuspended in 250 µl of 10 mM EDTA containing RNAse A at 200 µg/ml final concentration and incubated at 37°C overnight. After incubation, 250 µl of 10 mM EDTA containing 2 µM Sytox Green was added to each sample and incubated at room temperature for at least 2 hours. Samples were sonicated for 10-15 seconds and DNA content was measured using the CyAn ADP flow cytometer (Becton Dickinson). FlowJo software was used for subsequent image acquisition and analysis.

2.6.2 Cell number determination

1.4 ml of a mid-log phase growing culture was fixed by adding 2.8 ml of formal saline solution. Samples were diluted 10 times in ISOTON II (Beckman coulter) to get an average of 25000 counts. Each sample was processed in triplicate for cell counting and then the average was taken. Cells were counted by using a Beckman Coulter z series instrument.

2.7 Fission yeast molecular biology

2.7.1 Colony PCR

To confirm the deletion, tagging or overexpression of genes of interest, colony PCR was performed to amplify both the left and right junctions of genes. Primers were designed such that one lies in the UTR of the gene and the other lies inside the deletion cassette. The presence of the expected size of PCR product confirmed the integration of Kan^r or Nat^r cassette at the right place. Cells from desired colony were picked and boiled with 0.2% SDS at 100°C for 10 minutes. After placing them on ice for 3-4 minutes, 10 µl of dH_2O was added to cell suspension, mixed well and centrifuged for 1 minute at 13000 rpm. 1 µl of resulting supernatant was used as a

template for PCR run in a PTC-225 thermal cycler (MJ Research). The recipe for PCR mixture and PCR programme used are given in tables 2.3 and 2.4, respectively.

Table 2.3 PCR cocktail for colony PCR

Ingredients	**Amount**
Template	1 µl
10 X NH_4 buffer	5 µl
$MgCl_2$	1.5 µl
dNTPs (10 mM)	1.5 µl
Forward primer (10 µM)	1 µl
Reverse primer (10 µM)	1 µl
Taq polymerase (Bioline)	0.2 µl
H_2O	38.8 µl

Table 2.4 PCR programme

Step	**Temperature**	**Time**
1	95ºC	2 minutes
2	94ºC	1 minute
3	55-57ºC	30 seconds
4	72ºC	30 seconds (depending on the length of a PCR product)
5	go to step 2 and repeat 30 times	
6	72ºC	5 minute
7	4ºC	forever

2.7.2 Agarose gel electrophoresis

Depending on the size of PCR products to be analysed, agar was added to 1/2 X TBE buffer at 1-2 % (w/v). Agarose was solubilised by boiling for 2 minutes, cooled to below 60ºC, before ethidium bromide was added to a final concentration of 0.5 µg/ml. Then, 5 µl of PCR product was mixed with 2 µl of 6 x loading buffer (Bioline) to facilitate efficient loading of DNA samples into the wells of the gels. A DNA marker (Bioline) was run in parallel to determine the size of DNA within the gel. Gels were then run at 100 volts for 45 minutes, and visualisation of DNA fragments was performed under a UV trans-illuminator (BioDoc-It).

2.7.3 RNA extraction from cells

25 ml of cells (OD_{595} 0.5) were harvested by centrifugation for 3 minutes at 3000 rpm, supernatant was discarded and cells were stored at -70°C. Cells were then thawed on ice and washed with 1ml of DEPC treated water. 750 µl TES and 750 µl acidic phenol-chloroform (Sigma) were added to each sample. The mixture was vortexed for 15 seconds and incubated at 65°C for 1 hour, vortexing every 10 minutes. Samples were then placed on ice for 1 minute, mixed, and centrifuged for 15 minutes at 14,000 rpm at 4°C. 700 µl of upper aqueous-phase from each sample was then transferred to Maxtract high density phase lock tubes (Qiagen) together with an equal volume of acidic phenol-chloroform (Sigma) and centrifuged at 14000 rpm for 5 minutes at 4°C. The upper phase from these samples was transferred to phase lock tubes (Qiagen) and spun again after adding 700 µl chloroform:isoamyl alcohol (24:1 Sigma). RNA was precipitated from the upper phase with 1.5 ml 100% ethanol and 50 µl 3M NaAc (pH 5.2) overnight at -20°C or for 30 minutes at -70°C. The sample RNA was then collected by centrifugation (15 minutes at 14,000 rpm at 4°C) and washed once with 70% ethanol. The sample RNA was dissolved in 200 µl of DEPC treated water by incubation at 65°C for at least 1 minute at room temperature.

To assess RNA integrity, 500 ng of the sample was loaded on 2% (w/v) agarose gel together with DNA marker Hyperladder™ IV (Bioline) at 40V in TBE buffer. RNA

samples were then purified using the RNeasy Mini Kit (Qiagen) after manufacturer instructions.

TES: 10 mM Tris pH 7.5, 10 mM EDTA pH 8, 0.5 % SDS

2.7.4 DNA extraction

50 ml of OD 0.5 cells were harvested by centrifugation at 3000 rpm for 3 minutes. DNA extraction was performed using the ZR fungal DNA isolation kit (Zymo Research Corp) after manufacturer instructions. The quality of DNA was checked by loading 500 ng on 1% (w/v) agarose gel.

2.7.5 Strand-specific reverse transcription and PCR

RNA samples were digested with 2U DNaseI (TURBO DNase™, Ambion) for 25 minutes at 37ºC. Reaction was terminated by incubation with 10 µl inactivation mix (TURBO DNA-free Kit, Ambion) at room temperature for 2 minutes. 1 µg of RNA was then used to set up the reverse transcription mixture together with 3.33 µM of either the forward or reverse primer, 1 x First Strand buffer, 0.01 M DTT, 40 U RNAse OUT (RNAse inhibitor), 1.66 mM dNTP. This mixture was incubated at 42ºC for 2 hours. The RNA was then degraded by incubation with 33 mM NaOH at 70ºC for 15 minutes, and then 33 mM HCl was added to neutralise the pH of sample. The RT product was then extracted using MiniElute PCR purification Kit (Qiagen) after manufacturer instructions. For PCR analysis the BIOTAQ™ DNA polymerase system (Bioline) was used after manufacturer instructions. The reaction mixture was exactly as shown in table 2.3. Water was used as a negative control for cDNA whilst genomic DNA was used as a positive control. The reaction was run on PCR program (table 2.4) on PTC-225 thermal cycler (MJ Research), and the product was assessed on 1.5% (w/v) agarose gel together with DNA marker Hyperladder™ IV (Bioline).

2.7.6 Real time -PCR

Reverse transcription was performed on RNA samples using the same method given for strand-specific RT-PCR. Real-time PCR was then performed in duplicate in the

7500 Fast Real-Time PCR System (Applied Biosystems). The 20 µl reaction mixtures contained 10 ng cDNA, 0.2 µM each of the forward and reverse primers, and 10µl Fast SYBR Green PCR master mix (Applied Biosystems). An external standard curve was also set up using a suitable range of genomic DNA dilutions. The amplification program was as follows: 1 cycle of 95ºC for 20 seconds, 40 cycles of 95ºC for 3 seconds, 60ºC for 20 seconds.

2.8 Microarray techniques

2.8.1 Spikes synthesis

Spikes are exogenous RNAs that are used as external controls for normalisation of microarray data. To amplify the spikes, *E.coli* competent cells (Invitrogen) were transformed with plasmids, containing genes from *B. subtilis* (*ycxA, yceg, ybdO, ybbR, ybaS, ybaF, ybaC, yacK, yabQ, Trp, Thr, Dap, Phe* and *Lys)* and *E.coli* (*BioB, BioC* and *BioD*). These transformed *E.coli* cells were selected on LB plates containing ampicillin (50 µg/ml). Single colony from each plate was picked to inoculate the starter culture of 5 ml LB plus ampicillin. After growing them overnight at 37˚C with vigorous shaking, I inoculated 100 ml LB medium containing ampicillin and grew them at same conditions. Plasmids were extracted and purified by using QIAfilter plasmid Midi Protocol (QIAGEN). Plasmid DNA was digested with XhoI (spikes 1 to 9) and NotI (spikes 10 to 18) to linearize them for the production of polyA RNA. After precipitation of digested plasmids, *in-vitro* transcription was performed by using MEGAscript^{TM}, after manufacturer's instructions, to synthesize polyA RNA, which were mixed in different known concentrations (Appendix III).

2.8.2 Labelling

2.8.2.1 Genomic DNA labelling

40 µl 2.5 X random primer solution (Bio prime labelling kit, Invitrogen) was added to 0.6 µg DNA in 45 µl water and denatured for 10 min at 100°C. After cooling on ice for 1 min, 10 µl 10 X dNTP mix (0.5 mM dCTP, 2 mM dATP, 2 mM dGTP, 2

mM dTTP), 3 µl Cy3 or Cy5 and 2 µl Klenow fragment was added. Samples were incubated at 37 °C overnight and reaction was stopped by adding 10 µl stop buffer. Removal of labelled nucleotides from DNA labelling reaction was done by using Superscript III direct labelling purification kit (Invitrogen).

2.8.2.2 Synthesis and labelling of cDNA

For synthesis and labelling of cDNA, 10-20 µg of total RNA was mixed with 2 µg anchored oligo-dT, 1 µg random hexamer primers, 1 µl spikes in a total volume of 15 µl and incubated at 70°C for 10 minutes. After cooling on ice for 1 minute, I added 6 µl first strand buffer, 3 µl 0.1 M DTT, 2 µl dNTPs mix, 1 µl alexa-647 or alexa-555 conjugated dUTPs (Invitrogen), 1 µl RNAse out and 2 µl of Superscript III reverse transcriptase (Invitrogen). In case of hybridisation to in-house microarrays, 2µl of Cy5 or Cy3 dyes conjugated dCTPs were used instead of Alexa dyes. These samples were incubated at 42°C for 3 hours to generate labelled cDNA. Starting RNA was hydrolysed by incubating at 70°C for 30 minutes after adding 15 µl of 1 M NaOH. Samples were then neutralized by the addition of 15 µl 1 M HCl. Labelled cDNA was then purified by using Superscript III direct labelling purification kit (Invitrogen).

2.8.3 Hybridisation and post hybridisation

2.8.3.1 In house microarrays

Labelled cDNA resuspended in 27 µl hybridisation buffer was mixed with 3 µl/reaction of polyA DNA (2 µg/µl, Sigma). Hybridisation mixture was then denatured at 95°C for 5 minutes, cooled down to room temperature and hybridised to microarrays.

Microarrays covered with cover slip were placed inside the hybridisation chamber. 2 ml of 15 X SSC was added to the hybridisation chamber to keep it humid and placed at 49°C in an oven. After 15-18 hours incubation, microarrays were removed from hybridisation chamber and processed for washings. They were washed once with wash solution I for 5 minutes, twice for 10 min in wash solution II and once for 5 minutes in wash solution III.

Wash solution I: 2 X SSC

Wash solution II: 0.1 X SSC, 0.1 % SDS

Wash solution III: 0.1 X SSC

2.8.3.2 Agilent expression microarrays

For hybridisation of samples to Agilent 4x44 expression arrays, labelled cDNA samples were eluted together from the same purification column in 44 µl H_2O. Samples were boiled at 95°C for 5 minutes after adding 11 µl of 10 X blocking agent and 55 µl of 2 X GEX hybridisation buffer and were cooled to room temperature. After cleaning the microarray and cover slip, 100 µl of sample was added to each array. The tightened hybridisation chamber was then placed in an oven and incubated at 65°C for 12-17 hours.

After incubation, I unscrewed the hybridisation chamber, washed in wash buffers I and II for 1 minute each. Arrays were then rinsed in acetonitrile for 10 seconds and in stabilization and drying solution for 30 seconds.

2.8.3.3 Image acquisition

Microarrays were scanned using Genepix 4000B (Axon). Each image obtained at wavelengths 635 nm and 532 nm for Alexa-647 and Alexa-555, respectively, was saved as a separate TIFF file. These data were then normalized to calculate the relative expression levels of genes and to identify differently expressed genes.

2.8.4 Microarray data analysis

2.8.4.1 Image processing

GenePix Pro 6.0 software was used for image processing (Axon instruments). A preformed grid containing information about spots is placed over array blocks. After identifying regions corresponding to each spot, each array block was manually

inspected and damaged spots were removed. Then the program applies an algorithm to separate spots from background: local background is calculated and background subtracted intensities are calculated for both Alexa-647or Alexa-555 channels.

2.8.4.2 Data normalizations

For the in house arrays, the data generated after GenePix Pro 6 processing must be further filtered and normalised before proceeding to data analysis. For this purpose, a Perl script developed by our group was used (Lyne et al., 2003). This script applies cut off based criteria to discard data from weak signals. All spots that show >50% pixel intensities with a standard deviation >2 above median local background and spots that have >95 % of pixel intensities with a standard deviation > 2 above local background in one or both channel were kept for normalization. All other spots were discarded. To find whether the genome of polyploid cells is globally regulated, samples were normalised with spikes by using our perl script. The major step in this process is the normalisation of relative signal intensities in both scanned channels. Normalisation is also required for the Agilent array data to adjust the differences in labelling efficiencies for both labelling dyes and for differences in the quantity of starting RNA from the two samples used in experiments. These differences can cause a shift in the average ratio of Alexa-647 to Alexa-555, so intensities must be adjusted before the experiment can be analysed properly. For this purpose, data obtained by Agilent arrays were imported in GeneSpring (Agilent) and applied "Per spot and per chip: intensity dependent (Lowess) normalisation" to remove intensity dependent effects.

2.8.4.3 Identification of differentially expressed genes

First step in analysis of microarray data is to identify genes that are differentially expressed between two samples. In aneuploid and sporulating diploids, differentially expressed genes were selected by using the GeneSpring "Filter on expression level" tool. This tool allows finding genes that show up- or down-regulation in experimental samples above a certain threshold, relative to the control. Not many genes passed 2 fold threshold criterions for selection in sporulating polyploid cells. Therefore, for selection of genes that are significantly up or down regulated in non sporulating polyploids, the statistical programme Significant Analysis of

Microarrays (SAM) was used (Tusher et al., 2001). This programme was used as an Excel add-in for 6 experimental repeats. SAM determines a d-statistic and uses permutations to find out significant genes that change between two samples. Gene lists were selected with False Discovery Rate (FDR) equals to zero. This gene list was then imported to GeneSpring to select genes that showed 1.2-fold up- or down-regulation in at least 2 out of 3 polyploid samples.

Chapter 3

3 Impact of aneuploidy on fission yeast gene expression

In this chapter, I discuss the use of fission yeast aneuploid strains to explore the effects of aneuploidy on gene expression regulation. Two aneuploid strains were used, both containing an extra portion (minichromosome) of chromosome III such that they provided the disomy and trisomy for some genes present on chromosome III. The presence of an extra chromosome, location of mini-chromosomes, genes contained in disomic and trisomic area and microarray-based expression analysis was analysed by using the Agilent microarray system.

3.1 Introduction

Previously, by using aneuploid strains of fission yeast, Chikashige et al. (2007) have shown that all those genes that are duplicated or triplicated in a mini-chromosome exhibit corresponding fold up-regulation on microarrays. They used three different aneuploid strains containing mini-chromosomes named Ch16, S28 and Ch10. Although they found that aneuploidy affects the gene expression pattern in a dose dependent manner, they could not find enrichment for any particular functional category of genes that would establish the effects of aneuploidy on cellular fitness (Chikashige et al., 2007).

This was the first time the Agilent microarray platform was used in our lab; therefore it was important to see how efficiently these microarrays can detect the changes in expression of relatively few genes containing multiple genomic copies. The question was whether the presence of more than 2 copies of some genes will also show corresponding 2-fold up-regulation, or whether the Agilent platform presents limitations in detecting changes in gene expression.

Also, this was the first attempt to study the role of polyploidy in regulation of fission yeast gene expression (see chapter 4). As polyploid cells contain multiple copies of whole genomes, it was necessary to understand how these absolute changes in genome copy number will be reflected in terms of expression patterns on Agilent microarrays. Also, it was important to understand how methods used to normalise microarray data will affect gene expression changes associated with multiple copies

of whole genomes, when normalising the data with a haploid control containing only one copy of the genome.

In order to find answers to the above mentioned questions, we used two partial aneuploid strains carrying an extra copy of smaller portion of chromosome III (figure 2.1). For ease of use, these partial aneuploids will be named as follows:

Aneuploid 1: these are the haploid cells containing mini-chromosome Ch16, which is a smaller portion of chromosome III (figure 3.1A). It is about 530 kb long, spanning the centromeric sequences (Niwa et al., 1986). These haploids present partial aneuploidy in which 3% of the genome is duplicated, hence creating a disomic region in chromosome III.

Aneuploid 2: these haploid cells contain a smaller portion of mini-chromosome Ch16, mentioned in this study as IsoCh16. IsoCh16 was generated due to extensive end processing during repair of double strand breaks on Ch16 (Tinline-Purvis et al., 2009).

During repair of a double strand break on the right arm of Ch16, the DNA repair mechanism removed the broken chromosome arm and caused break induced replication of the intact arm from the centromere, thus resulting in duplication of the left arm and loss of the right arm of Ch16. This resulting IsoCh16 is 388 kb long, containing 2 copies of the left arm and centromere of Ch16, while its right arm is missing (figure 3.1B) (Tinline-Purvis et al., 2009). These partial aneuploid cells contain 3 copies of a small portion of chromosome III mentioned occasionally as trisomic area. All those chromosomal areas having only one copy of genome are mention as monosomic areas.

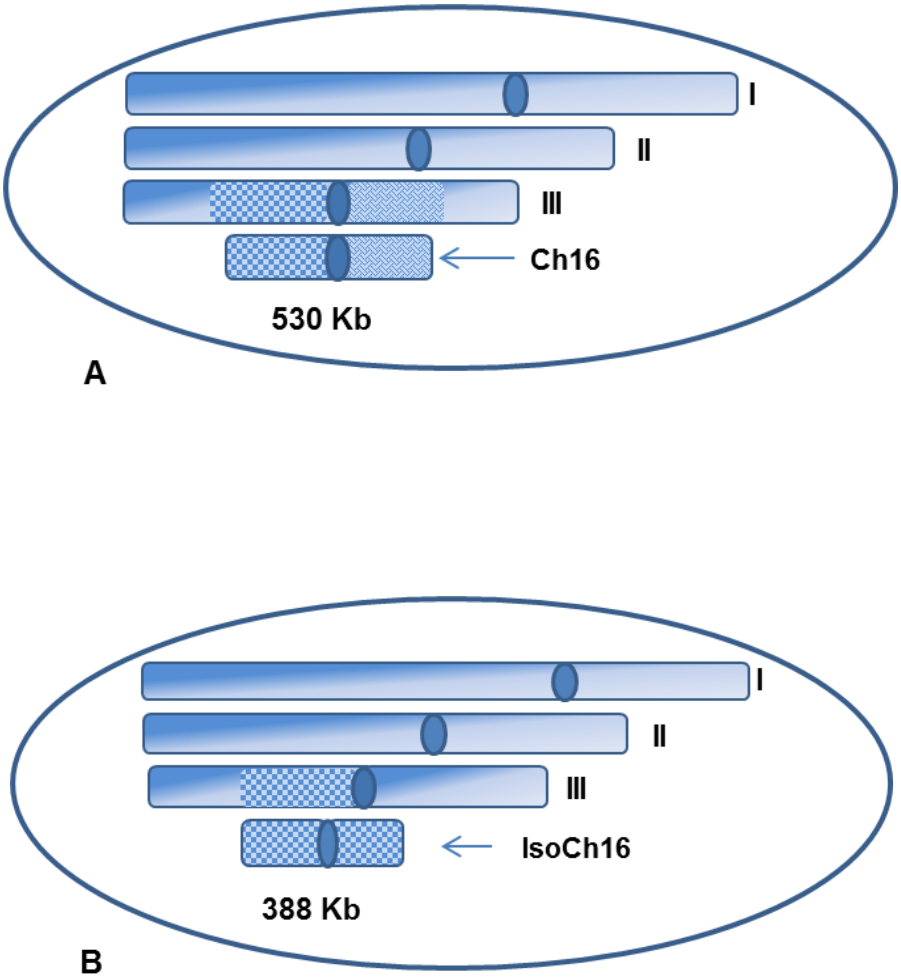


Figure 3.1: Schematic representation of aneuploids 1 and 2 of fission yeast. Both of these aneuploids contain an additional chromosome which is derived from chromosome III. The Ch16 containing aneuploid 1 contains 2 copies of the peri-centromeric region (A). The IsoCh16 containing aneuploid 2 contains 3 copies of a small portion of the left arm of chromosome III (B). Shaded areas on chromosome III represent the part that is present in Ch16 and IsoCh16, whereas dark blue ovals represent centromeres.

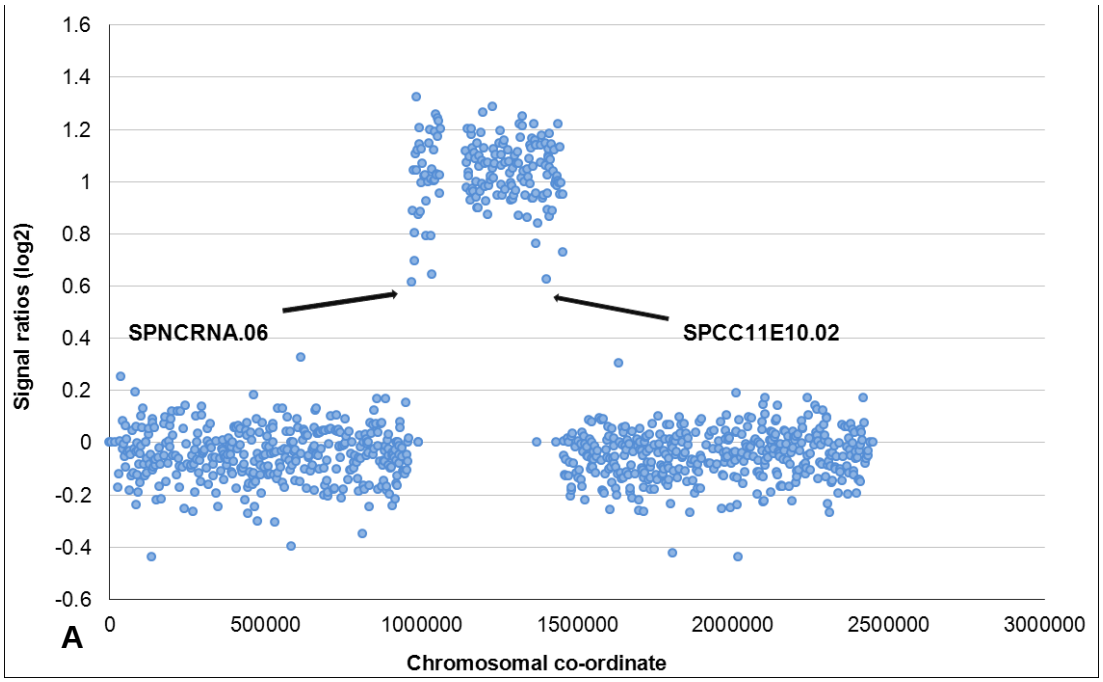
3.2 Determination of copy number and break points for Ch16 and IsoCh16

Although breakpoints of Ch16 have been determined previously (Chikashige et al., 2007), for analysis of gene expression in partial aneuploid strains using Agilent microarrays, it is necessary to know if this mini-chromosome still shows the same break point on these arrays. Also, we wanted to estimate the breakpoint of IsoCh16 on the left arm of chromosome III and genes contained in it. To estimate copy numbers of genes present in Ch16 and IsoCh16 with breakpoints, comparative genome hybridisation (CGH) was performed. Genomic DNA samples from both aneuploid strains were differentially hybridised on DNA microarrays that contained probes for coding and noncoding regions of the fission yeast genome.

First, to determine the copy number of genes contained in both mini-chromosomes, hybridization signals obtained for aneuploid strains were divided by those from the wild type haploid strain to get the hybridisation signal ratio by using normalisation method mentioned in section 2.8.4. CGH analysis indicated an increase in average ratio by 2.05-fold (Stdev 0.18) in the peri-centromeric region of chromosome III in the Ch16 containing strain, suggesting the presence of 2 copies of genes in this area (figure 3.2 A). Similarly, for aneuploid 2 ~3.29-fold (Stdev 0.3) increases in hybridisation ratio was observed in genes in the left arm of chromosome III (figure 3.2 B). This finding suggests that there are 3 copies of some genes in the left arm of chromosome III in aneuploid 2.

Next, we determined the break points of both Ch16 and IsoCh16 from CGH data. The right end of the Ch16 was mapped to SPCC11E10.02c located on the right arm of chromosome III (from position 1454816 bp to 1455958 bp) and the left end mapped to SPNCRNA.06 on the left arm of chromosome III (from position 968168 bp to 968477 bp).

For IsoCh16, both the right and left ends were mapped to SPNCRNA.06 (968168 bp to 968477 bp) on the left arm of chromosome III. This feature reflects the duplication of the left arm of Ch16 as mentioned previously. These areas were selected by hand to create gene lists. We found that Ch16 contains 167 genes while IsoCh16 contains 40 genes (Appendix IV).



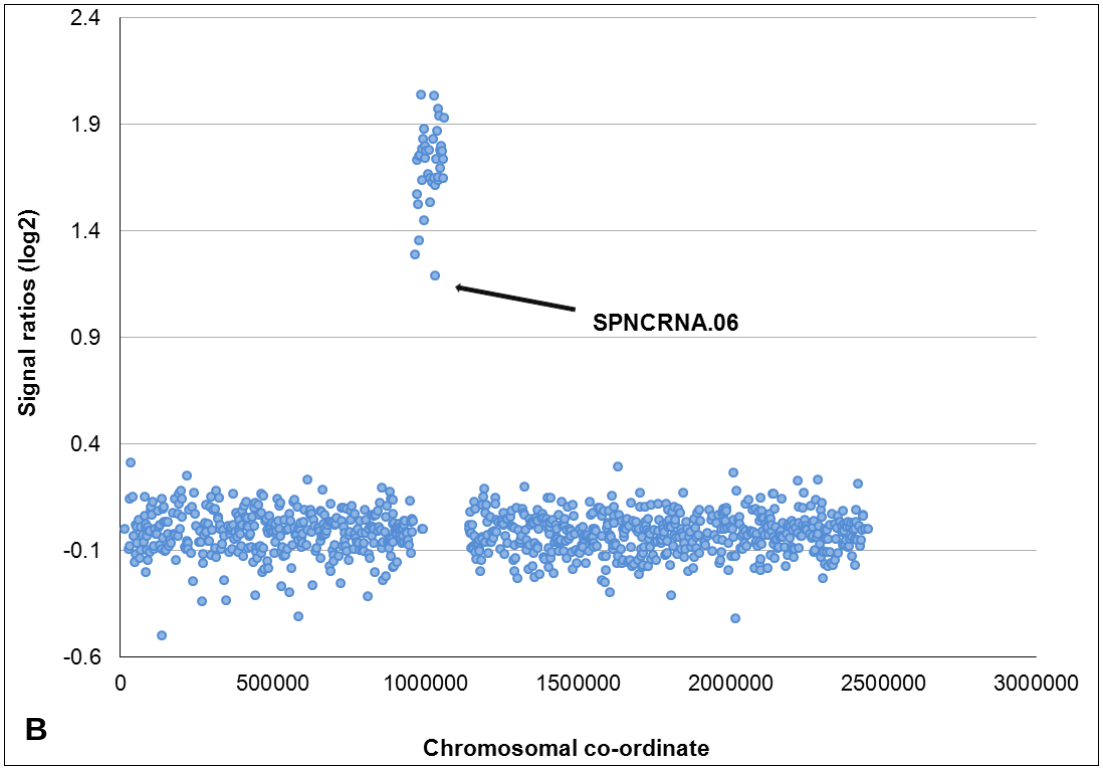


Figure 3.2: Graphs showing log2 of CGH ratios for chromosome III in aneuploid 1 (A) and aneuploid 2 (B). Expression ratios were obtained by normalising expression values obtained for aneuploid 2 against those obtained for

the haploid control. The X-axis represents the chromosomal coordinates, while Y-axis represents the log2 of CGH ratios. Regions on the chromosome III showing the log2 values of CGH ratios equal to or greater than 1 correspond to the regions contained in Ch16 and IsoCh16, respectively. Genes written next to arrows represent the breakpoints on chromosome III where Ch16 and IsoCh16 were mapped.

Comparison with the previous study showed that the right end of the minichromosome maps to the same position on chromosome III as reported (Chikashige et al., 2007), but the left in previous study maps to the region between SPCPB16A4.06c and SPCC1742.01. This difference might be a result of higher resolution and the presence of non-coding RNA probes on the Agilent microarrays used in our study.

3.3 Gene expression increases with an increase in copy number of genes.

To examine the effects of aneuploidy on fission yeast gene expression, labelled cDNA from both aneuploid strains was differentially hybridized to microarrays. Expression values obtained for each aneuploid strain were normalised against those obtained for a wild type control by using normalisation methods mentioned in material and methods, section 2.8.4. The average of expression ratios was found to be increased by 1.97-fold (Stdev 0.5) for disomic areas corresponding to the presence of 2 copies of genes in this area (figure 3.3). Similarly, for IsoCh16, we observed an increase in average of gene expression ratio by 3.0-fold (Stdev 0.6) for the trisomic area (figure 3.4). For monosomic areas on chromosome I, II and III, the expression ratio was measured around 1 except for few genes, which may reflect noise or compensatory changes caused by aneuploidy. These finding suggests that an increase in copy number of genes results in a corresponding increase in the expression of those genes.

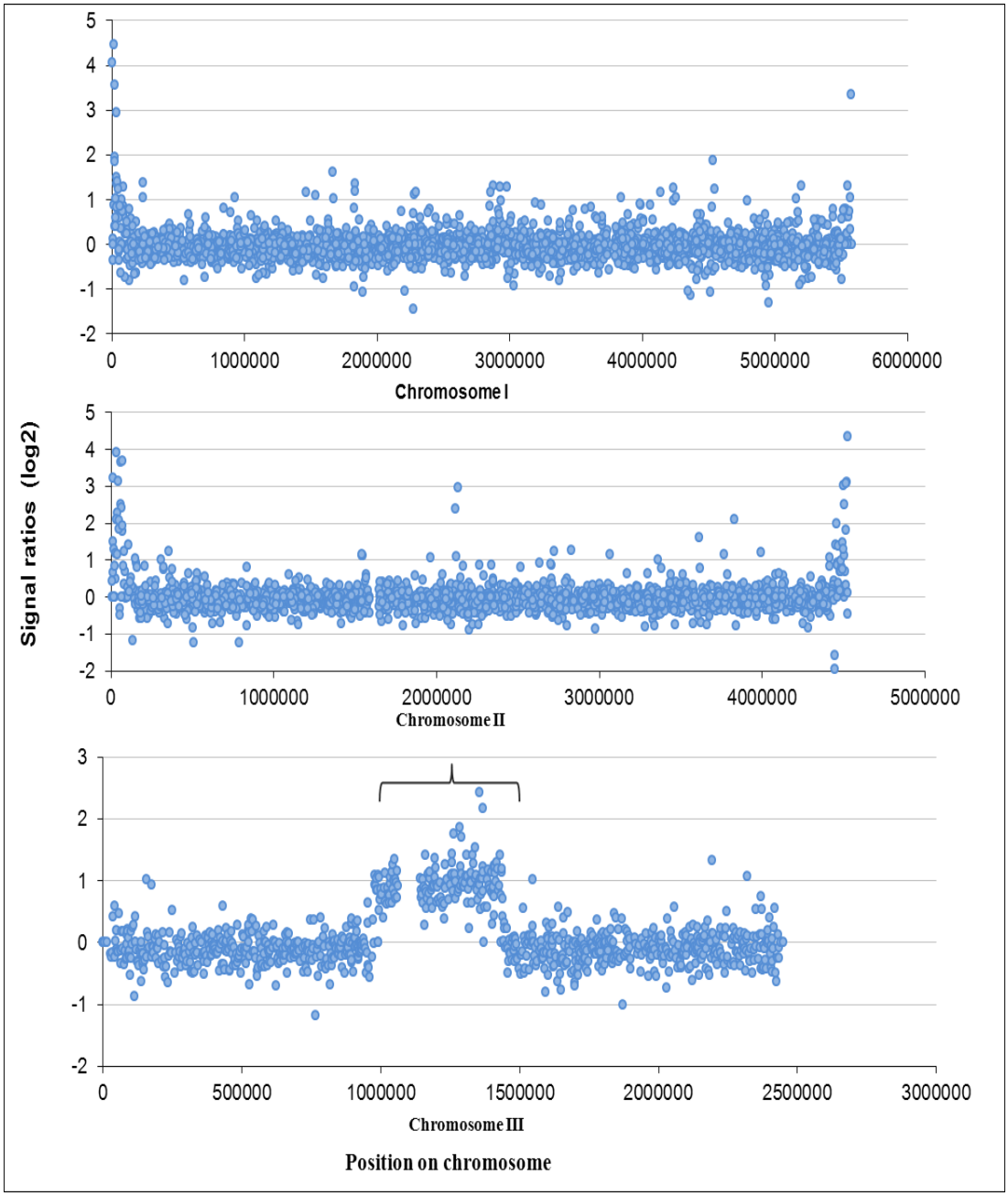


Figure 3.3: Graph showing expression analysis for all three chromosomes of aneuploid 1. Expression ratios were obtained by normalising expression values obtained for aneuploid 1 with those obtained for the haploid control strain. The X axis represents the chromosomal coordinates, while the Y axis represents the $\log_2$ of expression ratios. The region of chromosome III under parenthesis represents the duplication with $\log_2$ of expression ratios greater than 1.

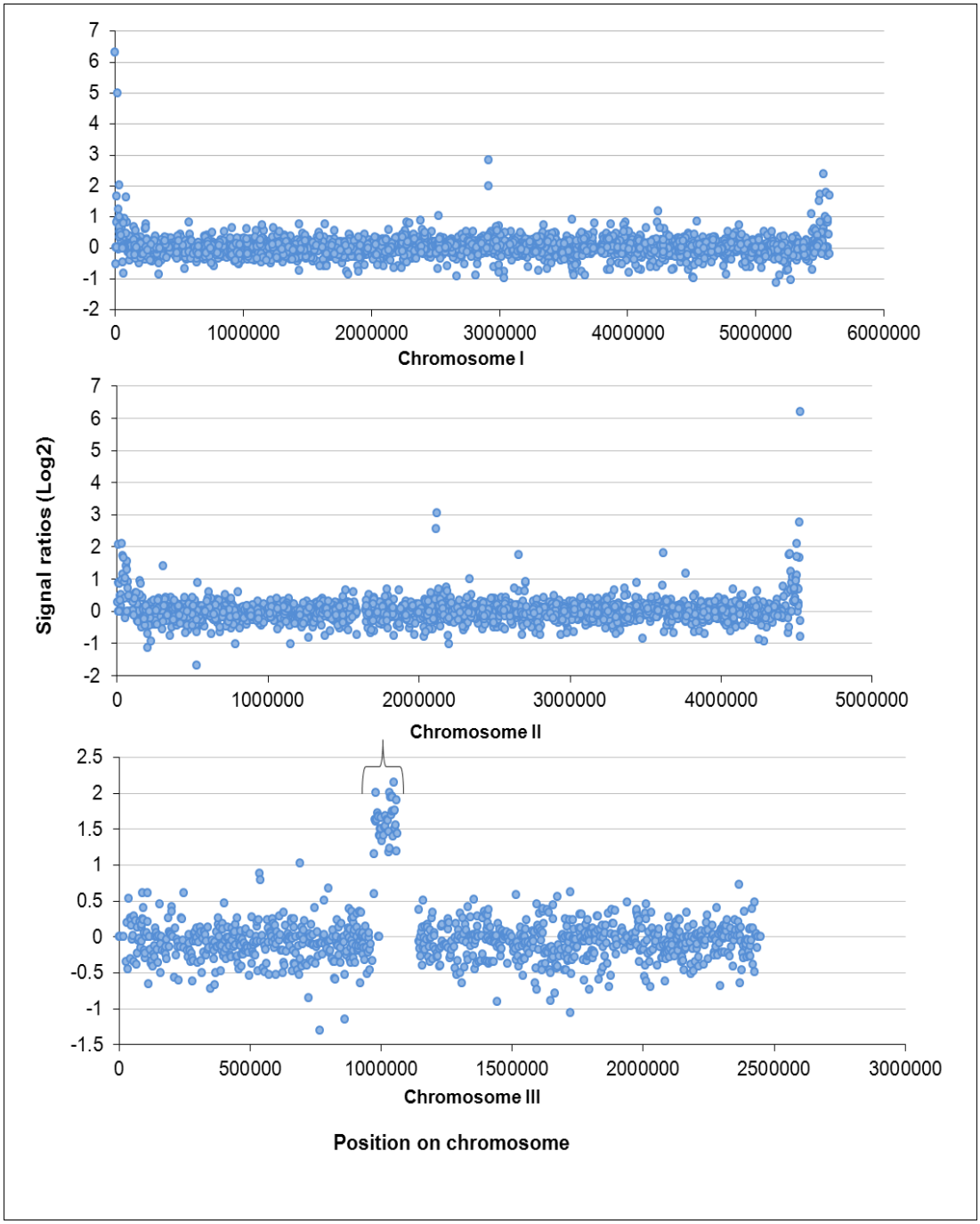


Figure 3.4: Expression analysis for all three chromosomes of aneuploid 2. Expression ratios were obtained by normalising expression values obtained for aneuploid 2 with those obtained for haploid control**.** The X axis represents the chromosomal coordinates, while the Y axis represents the expression ratios on log2 scale. Area of chromosomes III under parenthesis showing expression ratios of ~3 represent the triplicated portion present on IsoCh16.

3.4 Aneuploidy causes differential regulation of genes in monosomic region

This microarray-based expression study indicated the presence of differentially regulated genes in monosomic regions of both aneuploid strains (figure 3.3 and 3.4), which is in agreement with the previous study (Chikashige et al., 2007). We used Gene list analyser, present as a web based resource on Bähler group's web site (www.bahlerlab.info), to find any functional enrichments for these differentially regulated genes based on Gene Ontology (GO), and/or, their location on the chromosomes. As we could not find any enrichment based on GO term, therefore, for better and easy understanding of these gene lists, we have manually categorised these genes into groups based on similar biological functions obtained from the PomBase model organism database: www.pombase.org (Appendix V, table 1 A).

3.4.1 Aneuploid 1

In the aneuploid 1 strain, a total of 92 genes present in monosomic regions showed an increase of ≥2 fold in expression ratios when compared to the haploid control (Appendix V, table 1A). The majority of these genes (88 genes) were present on chromosomes I and II. Genes present on chromosome III include *mde7,* coding for an RNA binding protein that is specifically up-regulated during meiosis, *rec7* which encodes a meiotic recombination protein involved in segregation of homologous chromosomes, *psi1* which encodes a protein of unknown function but predicted to be involved in initiation of translation, and *tf2-13* which is a retro transposable element. Many of the up-regulated genes (65 genes) were found to be located near telomeres of chromosomes I and II. Eleven of the up-regulated genes are *S. pombe* specific and of unknown molecular functions. Four of these genes code for protein products required during meiosis such as; *mde2,* required for meiotic chromosome segregation; *mde7,* coding for a meiotic protein;, *matPc* and *mat3-Mc,* both coding for proteins required for sexual conjugation.

Some other well characterised genes showing up-regulation in aneuploid 1 include *atf1,* encoding a transcription factor required for regulation of various stress responses and meiotic recombination, *rec7,* coding for a protein involved in meiotic recombination and segregation of homologous chromosomes, *tlh1* and *tlh2,* coding for proteins involved in telomere maintenance via recombination, *psi1,* involved in translation initiation, *aes1,* involved in chromatin silencing, and *ppr1* involved in regulation of rRNA and mRNA stability.

Few genes in aneuploid 1 were also found with ≥2 fold decreased in expression ratios compared to their haploid control. A total of 14 such genes were present in this aneuploid, and 6 of them encode proteins that are present on plasma membrane (Appendix V, table 1B). Other genes showing down-regulation include *arg3* which encodes a protein involved in ornithine biosynthesis in urea cycle, *mei2* which encodes an RNA binding protein involved in meiosis, and *cid11* which codes for a protein product of unknown molecular function but predicted to be RNA binding.

3.4.2 Aneuploid 2

In aneuploid 2, 51 genes present in monosomic region showed an increase of ≥2-fold in expression ratio compared to the haploid control (Appendix V, table 2A). All of these genes were found to be located on chromosomes I and II, and 42 of these genes were found in the vicinity of telomeres. Some well characterised genes showing up-regulation include *str3,* coding for a protein involved in transmembrane transport and cellular homeostasis of iron, *fio1,* involved in import of ferrous ions and assimilation of iron by reduction and transport, *hsp16,* coding for protein involved in mRNA export from nucleus in response to heat shock. Seven of the up-regulated genes code for *S. pombe* specific proteins that have not been characterised.

Similarly, 12 genes showing down-regulation were also found in monosomic areas of this aneuploid (Appendix V, table 2B). Among other down-regulated genes of various functions in aneuploid 2, *fhn1* encodes a protein that is involved in plasma membrane organisation and protein secretion.

3.4.3 Comparison between aneuploid 1 and 2

To see whether differentially regulated genes present in monosomic regions of both aneuploids are similar or not, we used the Venn diagram feature of GeneSpring to see the overlap between genes up-regulated or down-regulated (figure 3.5). Comparison of up-regulated genes indicates the presence of 33 genes that are regulated in both aneuploids. These genes are involved in a variety of functions (Table 3.1) such as metabolism, meiosis, and transport of various metabolites across membranes. Three of these genes (*tlh1, tlh2* and SPAC212.06c) are involved in recombination and telomere maintenance. Fourteen genes code for proteins of unknown molecular functions, and among them 7 are *S. pombe* specific.

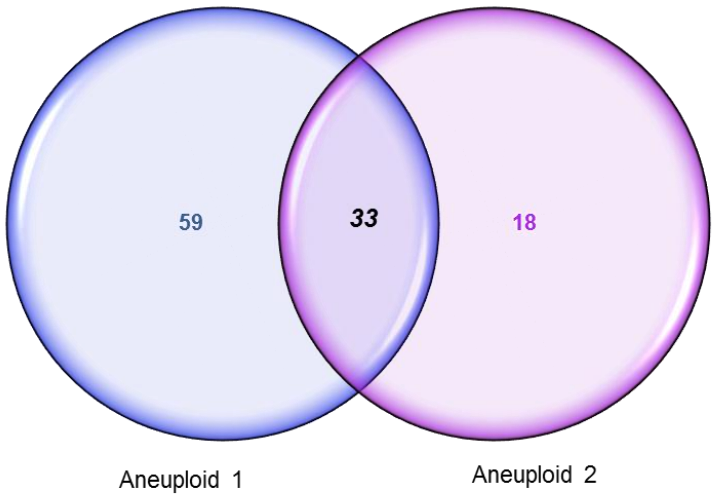


Figure 3.5: Venn diagram showing the overlap of over expressed genes between aneuploid 1 and 2. Genes showing ≥2-fold over expression in monosomic regions in each aneuploid, when compared to the haploid control, are shown in individual colours representing aneuploid 1 and 2, while genes that overlap between two aneuploids are written as italic, bold in black colour.

Table 3.1: Up-regulated genes in monosomic regions of aneuploid 1 and 2

Systematic ID	Common name	Description
Metabolic genes		
SPBPB2B2.01		amino acid permease (predicted)
SPBPB21E7.01c	*eno102*	enolase (predicted)
SPBPB21E7.04c		human COMT ortholog 2
SPAC186.08c		L-lactate dehydrogenase (predicted)
SPBC23G7.10c		NADH-dependent flavin oxidoreductase (predicted)
SPBPB21E7.02c		phosphoglycerate mutase family
Meiosis		
SPBC1711.02	*mat3-Mc*	mating-type m-specific polypeptide
SPBC31F10.08	*mde2*	Mde2 protein
Unknown functions		
SPBPB2B2.17c		*S. pombe* specific 5Tm protein family
SPBC1348.03		*S. pombe* specific 5Tm protein family
SPAC977.02		*S. pombe* specific 5Tm protein family
SPBPB2B2.19c		*S. pombe* specific 5Tm protein family
SPAC212.04c		*S. pombe* specific DUF999 family protein 1
SPAC212.12		*S. pombe* specific GPI anchored protein family
SPAC750.07c		*S. pombe* specific GPI anchored protein family 1
SPBPB2B2.18		sequence orphan
SPBCPT2R1.06c		pseudogene
SPBPB21E7.06		pseudogene
SPBCPT2R1.09c		pseudogene
SPNCRNA.288		unknown function
SPAC212.08c		GPI anchored protein (predicted)
SPBPB2B2.15		conserved fungal family
Recombination		
SPAC212.11	*tlh1*	RecQ type DNA helicase
SPBCPT2R1.08c	*tlh2*	RecQ type DNA helicase
SPAC212.06c		DNA helicase in rearranged telomeric region, truncated
Surface transporters		
SPBPB8B6.02c		urea transporter (predicted)
SPBC1348.11		membrane transporter, pseudogene
SPAC1F7.08	*fio1*	iron transport multicopper oxidase
SPBC1348.14c	*ght7*	hexose transporter (predicted)
Various functions		

SPBC1348.13		similar to fragment of cox1 intron protein
SPNCRNA.445	snoR61	small nucleolar RNA
SPBC1773.12		transcription factor (predicted)
SPBP4G3.03		PI31 proteasome regulator related

Comparison of down-regulated genes revealed only two genes in common between aneuploid 1 and 2. One of them is SPBC530.02 that codes for a protein of unknown molecular function but is predicted to be present in the plasma membrane, with a probable role in transmembrane transport. The other gene is *gnr1*, which encodes a protein involved in negative regulation of pheromone dependent signal transduction required for conjugation.

3.5 Transcription factors in the disomic and trisomic areas

Differential regulation of the genes in monosomic areas suggests the presence of genes encoding transcription factor that could be involved in the transcriptional regulation of these genes. To check whether any such genes are present in the mini-chromosomes, we manually analysed genes contained in the two mini chromosomes for their functions (Appendix IV). Interestingly, we found a total of 5 such genes that are either transcription factors or a part of the regulatory networks that can drive the expression of other genes. Two of these genes were common between two aneuploids, while 3 were observed only in aneuploid 1.

Two genes that are contained in both mini-chromosomes include *rxt3* which encodes a transcriptional regulatory protein that is a part of the histone deacetylase complex Rpd3S (Shevchenko et al., 2008) and is essential for cell survival (Kim et al., 2010); and *taf8* which encodes a transcription factor TFIID complex subunit 8 with a probable involvement in RNA polymerase II dependent transcription initiation, based on its orthologs in human and *S. cerevisiae* (www.pombase.org). Deletion mutants of *taf8* are inviable (Kim et al., 2010).

Three genes present only on minichromosome Ch16 include *pof3* which codes for an F box protein that is involved in the regulation of the G2/M transition of the mitotic cell cycle and in telomere maintenance (Katayama et al., 2002), and *rsc7* which

codes for an RSC complex subunit that is probably involved in regulation of transcription from RNA polymerase II promoters and is essential for cell survival (Kim et al., 2010). Its ortholog in *S. cerevisiae* is involved in importing nuclear proteins and maintenance of proper telomere length (Bossie et al., 1992; Askree et al., 2004). The third gene is *srk1*, encoding a MAPK activated protein kinase, involved in the regulation of mitosis and stress activated signalling pathway (López-Avilés et al., 2005).

The presence in both mini-chromosomes of these genes functioning in transcriptional control suggests their involvement in the transcriptional regulation of genes located in the monosomic regions.

3.6 Comparison of differentially regulated genes with previous study

In a previous study, Chikashige et al (2007) found 79 genes that showed up-regulation and 18 genes that showed down-regulation in the monosomic region of Ch16 containing aneuploid (Chikashige et al., 2007). Out of these 79 genes, 28 were present near telomeres of chromosomes I and II. In our study, a total of 92 genes showing up-regulation were found in the monosomic region, and 65 of these genes were found near telomeres of chromosomes I and II.

A Venn diagram comparing both studies showed an overlap of 29 genes, 18 of which are located near telomeres of chromosomes I and II (Table 3.2). 63 genes that are up-regulated only in this study include 18 genes that have previously been described as stress responsive genes (Chen et al., 2003), 6 genes of retro-transposable elements, 3 genes coding for transcription factors, while other genes code for proteins of various functions (Appendix V, table 1A).

On the other hand, when the comparison was carried out for down-regulated genes, we found only 2 genes that overlap between the two gene lists. One of them, SPBC8E4.01c, encodes a protein that is probably involved in transport of inorganic phosphate, while the other, *pho1*, codes for acid phosphatase involved in cellular response to phosphate starvation. Comparison between the two studies suggests that

most of the differentially regulated genes in monosomic areas might be regulated by some other factors such as growth condition, or media, or it could reflect differences in the genotypes of the two strains.

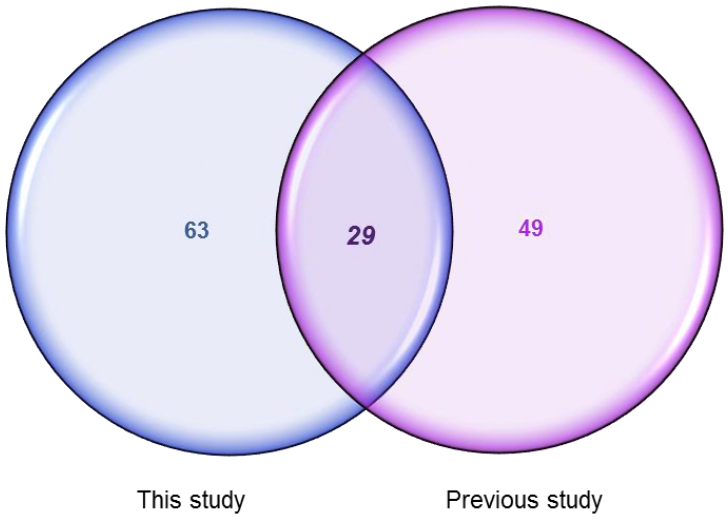


Figure 3.6: Venn diagram showing overlap of over expressed genes between two studies. Genes showing ≥2-fold over expression in monosomic regions in aneuploid 1 in this study and previous study, when compared to a haploid control, are shown in individual colours, while genes that overlap between two aneuploids are written as italic, bold in maroon colour.

Table 3.2: Up-regulated genes in monosomic region of aneuploid 1 showing overlap between the two studies

Systematic ID	Common name	Description
SPACUNK4.16c		alpha,alpha-trehalose-phosphate synthase (predicted)
SPAC637.03		conserved fungal protein
SPBC19C7.04c		conserved fungal protein
SPBC30D10.14		dienelactone hydrolase family (predicted)
SPAC212.06c		DNA helicase in rearranged telomeric region, truncated
SPCC830.07c	*psi1*	DNAJ domain protein, involved in translation initiation Psi1
SPAC212.08c		GPI anchored protein (predicted)
SPBC16D10.08c		heat shock protein Hsp104 (predicted)
SPAC13G7.02c	*ssa1*	heat shock protein Ssa1 (predicted)
SPBC1348.14c	*ght7*	hexose transporter Ght7 (predicted)
SPBC3B9.01		Hsp70 nucleotide exchange factor (predicted)
SPBC1348.05		membrane transporter (predicted)
SPBC1348.11		membrane transporter, pseudogene
SPBC23G7.10c		NADH-dependent flavin oxidoreductase (predicted)
SPAC212.05c		pseudogene
SPBC1348.03		*S. pombe* specific 5Tm protein family
SPAC977.02		*S. pombe* specific 5Tm protein family
SPBPB2B2.19c		*S. pombe* specific 5Tm protein family
SPAC750.05c		*S. pombe* specific 5Tm protein family
SPAC977.01		*S. pombe* specific 5Tm protein family
SPAC212.04c		*S. pombe* specific DUF999 family protein 1
SPAC977.06		*S. pombe* specific DUF999 family protein 3
SPBC1348.07		*S. pombe* specific DUF999 protein family 6
SPAC750.07c		*S. pombe* specific GPI anchored protein family 1
SPBC21C3.19		SBDS family protein Rtc3 (predicted)
SPBC11C11.06c		sequence orphan
SPBC83.17		transcriptional coactivator, multiprotein bridging factor Mbf1 (predicted)
SPAC977.04		truncated C terminal region of membrane transporter
SPBC1348.12		zinc finger protein

3.7 Conclusion

In this chapter, we used two partial aneuploids of fission yeast, named aneuploid 1 and aneuploid 2, to study the effect of aneuploidy on both local and global gene expression, by using Agilent expression microarrays.

We observed that an increase in copy number of a few genes can result in a corresponding increase in expression of those genes. This primary imbalance in gene expression indirectly disturbed the gene expression signature throughout the genome by increasing or decreasing the expression of other genes present in the monosomic regions.

A total of 92 genes were up-regulated in monosomic regions of aneuploid 1. Among them 88 were found on chromosomes I and II. Among these, 65 genes were found to be located near telomeres of chromosomes I and II. Fourteen genes in this aneuploid showed down-regulation compared to their haploid control. Similarly, for aneuploid 2, 51 genes were up-regulated and 12 showed down-regulation. Many of these differentially regulated genes are located in the vicinity of telomeres of chromosomes I and II.

We also observed the presence of genes encoding transcription factors in the mini-chromosomes that might be involved in the differential regulation of genes located in the monosomic areas.

Another aim of conducting this study was to see how relative changes in gene dosage are reflected on these microarrays, and how normalisation methods will detect these changes when analysing the data. We observed that relative changes in gene dosage can be detected easily and quantitatively by using these arrays. Also, normalisation methods used to analyse the data was efficient to observe these changes in copy number as well as expression levels of genes present in disomic and trisomic areas.

Chapter 4

4 Impact of polyploidy on gene expression

This chapter describes experiments designed to observe the impact of polyploidy on fission yeast gene regulation. Two different diploid strains (sporulating and non-sporulating) and one tetraploid strain (non-sporulating) were analyzed. The relationship between ploidy and gene expression level discussed here is based on the experimental evidence provided by microarray analysis derived from 7 independent repeat, including 6 biological and 1 technical repeat.

4.1 Introduction

The idea of this study developed from a previous study in *S. cerevisiae*, where the role of polyploidy in regulation of gene expression was explored. In this microarray-based study, (Galitski et al, 1999) identified 17 genes that showed ploidy dependent expression. Among these ploidy-dependent genes, 10 were induced, while 7 were repressed as a function of ploidy. To distinguish gene expression effects based on ploidy from those based on different mating types; Galitski et al. constructed 3 isogenic sets of strains, with each set differing only in their ploidy. They established the role of ploidy in regulation of gene expression by comparing the pattern of gene expression in strains which share the same mating type but differ in ploidy. Although it was the first study to clearly establish a role of ploidy on gene expression, they did not show any functional relationship between ploidy and gene expression. This could be due to the small number of ploidy dependent genes that did not show enrichment for any particular functional category. Therefore, to understand the outcome of the change in ploidy in fission yeast, the study presented here was designed. Our major goal was to understand the molecular basis of the change in global transcription associated with genome duplication. We also tried to find out whether any functional relationship between ploidy and gene expression exists in fission yeast.

One challenge to start this project was to generate the series of strains that differ in their ploidies. The cell cycle of *S. cerevisiae* alternates between the haplophase and the diplophase, with the diplophase being dominant. Due to this characteristic, the generation and maintenance of polyploid populations are much easier than in *S. pombe* which much prefers to grow as haploid cells. Polyploids up to hexaploids

have been reported in *S. cerevisiae*, that are obtained by simply crossing strains in the laboratory (Mortimer, 1958).

4.2 Polyploidy in fission yeast

Unlike *S. cerevisiae*, fission yeast has a haplontic life cycle under normal growth conditions (Molnar and Sipiczki, 1993). However, sporulating diploid and tetraploid cells can be generated by crossing haploid cells of opposite mating type, either by nitrogen starvation or by protoplast fusion. Under nitrogen starvation, cells of opposite mating type (h^+/h^-) fuse together by the process of sexual conjugation, resulting in a transient zygote. Soon after the completion of nuclear fusion, this diploid zygote undergoes meiosis and forms an ascus containing four spores. If nitrogen levels are increased suddenly after a short period, this zygote can return to the mitotic cell cycle instead of undergoing meiosis, hence giving rise to a diploid population of fission yeast (Molnar and Sipiczki, 1993). Diploid cells generated in this way have been reported to be less stable, as they stop growth and can undergo meiosis as soon as nitrogen levels fall in the growth media ((Egel, 1989). Diploid and tetraploid cells in *S. pombe* can also be generated through the use of a protoplast fusion protocol. The polyploid arising from fusion of protoplasts have also been shown to be unstable as they accumulate cells of lower ploidies during exponential growth (Molnar and Sipiczki, 1993), probably due to problems in mitotic chromosome segregation.

Although, few marker systems have been used previously by (Molnar and Sipiczki, 1993) for selections of diploid and tetraploid cells in fission yeast, we needed a marker system that has a minimum or no effect on the physiology of fission yeast cells. For this purpose, we took advantage of the *ade6* mutation system in fission yeast cells to generate polyploid cells. This marker system not only offers easy selection of polyploids but the resulting cells also show a normal phenotype. Mutants of *ade6* are easy to select as haploid cells have a single mutation at the *ade6* locus (*ade6-M210* or *ade6-M216*). These cells cannot synthesize adenine and appear as pink colonies on plates containing media with low adenine (YEA). Colonies of *ade6-M210* appear as dark pink, while colonies of *ade6-M216* are light pink, due to accumulation of a precursor metabolite. When these haploid cells harbouring

different mating types (h^+ or h^-) are streaked on plates with media containing low nitrogen (MEA), they fuse together via conjugation. In the resulting diploids, these mutations complement each other and appear as white colonies on YEA plates. These diploid cells show normal growth in the absence of adenine. These cells were maintained on YES plates to suppress any sporulation. We also used homothallic haploid cells that had insertion of different antibiotic markers at the *ade6* locus (*kan*^{r}, *Hyg*^{r} or *Nat*^{r}). These haploid cells were mated on MEA plates to generate diploid cells (*ade6::KanMx6 h90/ade6::NatMx6 h90* or *ade6::KanMx6 h90/ade6::hygMx6 h90*) that can be selected on YES plates containing a combination of two respective antibiotics. These haploid and diploid cells were provided by Dr. Francesc Xavier Marsellach Castellví. The idea behind these diploids was to construct tetraploid cells that could be selected on plates containing a combination of three antibiotics.

Although we were successful in generating diploid cells using the above mentioned procedures, they were found to be sporulating. To supress sporulation, these diploid cells were grown in YES media. These diploid cells were crossed further with each other on MEA to generate tetraploid cells that could be selected either on YEA or on YES plates containing a combination of three antibiotics. Tetraploids generated in this way were highly unstable, therefore could not be used for microarray based analysis of gene expression.

Our goal was to generate diploid and then tetraploid fission yeast cells that are not only stable but also of the same mating type, to minimize the effect of mating type differences on gene expression. It has been shown earlier that diploid cells sharing the same genotype at the mating type locus are stable. In some strains of fission yeast, diploid cells arise spontaneously, possibly as a result of endoreplication. We used this characteristic to select for homozygous diploid cells of the same mating type. For this purpose, haploid cells (*ade6-M210 h*$^-$ or *ade6-M216 h*$^-$) were grown to exponential phase and were then plated on Phloxin B containing rich media (YEP) to obtain single colonies. Diploid cells arising as a result of endoreplication appear as dark red colonies on these plates. As these diploids cannot synthesize adenine, they were selected and further maintained on YES plates. Diploids arising in this way were non-sporulating and hence stable. They were stored at -80°C for further use (while sporulating diploids were constructed fresh for every experiment). These

diploids were then used to generate non-sporulating tetraploid cells by the process of protoplast fusion. Tetraploid cells were selected on YEA plates, as white colonies, being prototrophic for adenine.

Using these approaches, we were able to generate 4 different types of diploid strains: 2 were sporulating, while 2 were non-sporulating. We also generated one type of tetraploid strain that originated from protoplast fusion of non sporulating diploids strains. For simplicity, haploids, diploids and tetraploids used in this chapter will be named as follows:

Haploids

Haploid A (*ade6 M210 h^-)*

Haploid B (*ade6 M216 h^-)*

Haploid C *(ade6 M210 h^+)*

Haploid D (*ade6::Kan Mx6 h^{90}*)

Diploids

Non sporulating

Diploid A (*ade6 M210h^-/ade6 M210 h^-*)

Diploid B (*ade6 M216 h^-/ade6 M216h^-*)

Sporulating

Diploid C (*ade6 M210 h^+/ade6 M216 h^-*)

Diploid D (*ade6::KanMx6 h^{90}/ade6::Nat Mx6 h^{90}*)

Tetraploid

Tetraploid (*ade6 M210 h^- ade6 M210 h^-/ade6 M216 h^- ade6 M216 h^-)*

As we did not succeed in generating stable tetraploids from sporulating diploids, results generated by using such unstable tetraploids are provided in separate appendixes. All results presented in this chapter (except figure 4.1) are obtained from non-sporulating polyploids.

4.3 Cellular changes with altered ploidy

The DNA content of polyploid strains was confirmed by FACS analysis (figure 4.1). Haploid cells from all four backgrounds showed 2C DNA content corresponding to their G2 phase, while DNA content of diploid and tetraploid cells was evident as 4C and 8C peaks. FACS analysis showed a clear single peak for haploids and diploids corresponding to their DNA content, while for tetraploids a broader peak was observed, ranging from 4C to 8C. This finding suggested the presence within the tetraploid cell culture of some cells with lower ploidies and aneuploidies, ranging from diploids to tetraploids. This interpretation was confirmed by measurements of cell size (figure 4.2 and 4.3).

To analyse the effect of tetraploidy on regulation of gene expression, tetraploid cultures were selected after performing FACS analysis. Only those cultures that had more than 65% of cells showing 8C DNA content were used for microarray-based studies.

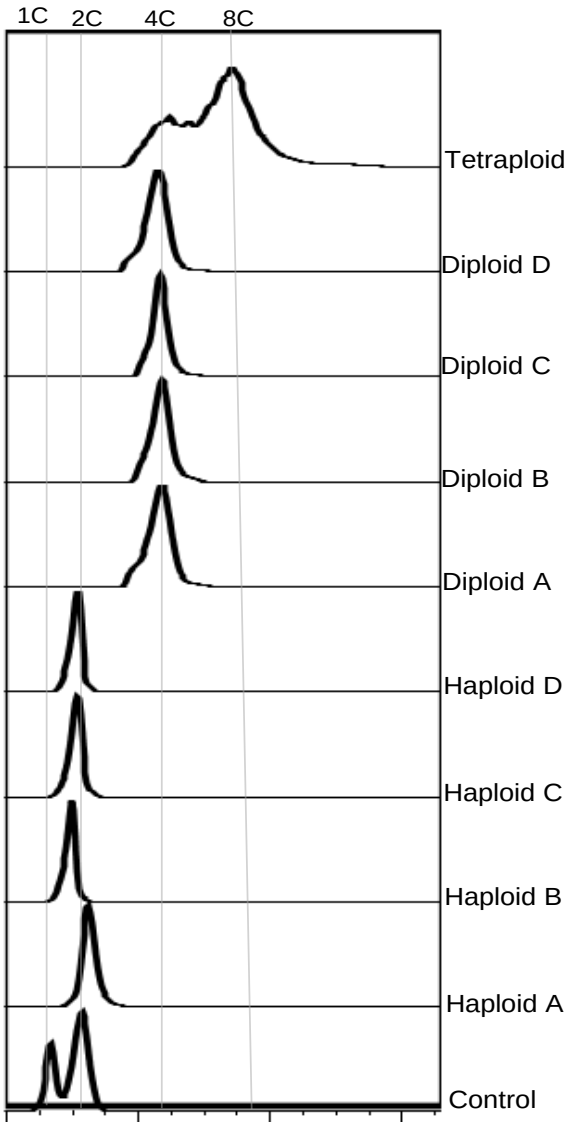


Figure 4.1: FACS profile showing ploidy of cells. Cells stained with Sytox Green were processed in a CyAn ADP flow cytometer. The resulting data was analysed by using FlowJo 7.6. For the control, wild type cells were starved of nitrogen for 3 hours to partially arrest the cell cycle such that some cells show 1C while others show 2C DNA content, corresponding to G1 and G2 phase of the cell cycle, respectively.

4.3.1 Cell size shows linear relationship with ploidy

By microscopy, diploid and tetraploid cells showed a different morphology from haploid cells. They appear larger than haploid cells (figure 4.2A-C). Apart from cell size, all diploid strains showed a normal phenotype, but in tetraploid populations some cells with irregular shapes were observed. For example, many cells with double septa were also present in tetraploid strains that could give rise to unequal divisions (figure 4.2). The presence of double septa and unequal division could be a cause of heterogeneity in our tetraploid population, both in terms of size and ploidy.

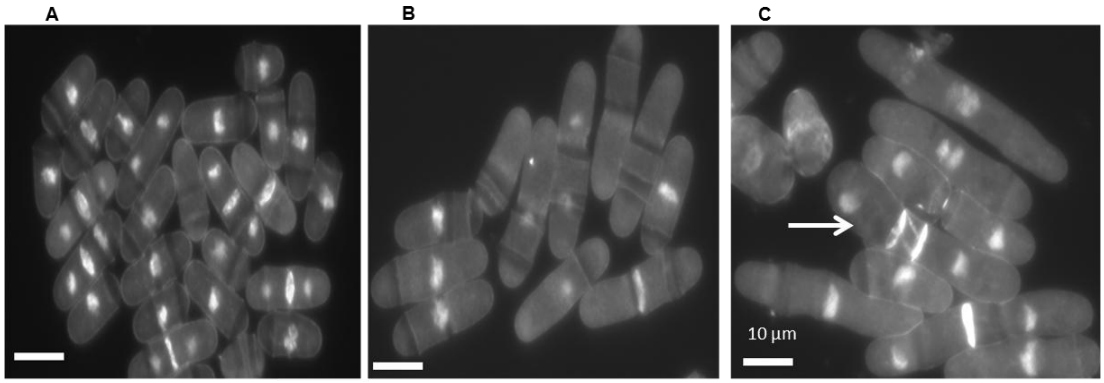


Figure 4.2: Micrographs of haploid (A), diploid (B), and tetraploid (C) cells. Cells were stained with Calcofluor (50 µg/mL; Polysciences Inc.) to visualize the septa and cell wall, and with DAPI (10 µg/mL) to stain nuclei. Images were recorded with DAPI filter using a Hamamatsu digital camera C4742-95 fitted to a Zeiss Axioskop microscope with plan-Apochromat X 63 1.25 oil objective. Bars = 10 µm. Arrow shows a tetraploid cell with double septa.

To further understand the relationship between ploidy and cell size, lengths of 170 septated cells were measured (figure 4.3) each for haploid A, haploid B, diploid A, diploid B and tetraploid. The mean size of haploids was found to be 13.8 µm (Stdev $\pm$ 1.3) and 13.6 µm (Stdev $\pm$ 1.2) for haploids A and B, respectively, 22.3 µm (Stdev $\pm$ 2.5) and 23.04 µm (Stdev $\pm$ 2.2) for diploids A and B, respectively, and 34 µm (Stdev $\pm$ 7.9) for the tetraploid strain. These measurements show a linear relationship between ploidy and cell lengths (figure 4.3). Cell lengths for tetraploid strain

represent of a heterogeneous population of cells, suggesting the presence of cells with varying ploidy, ranging from tetraploids to diploids, which is in complete agreement with the FACS analysis.

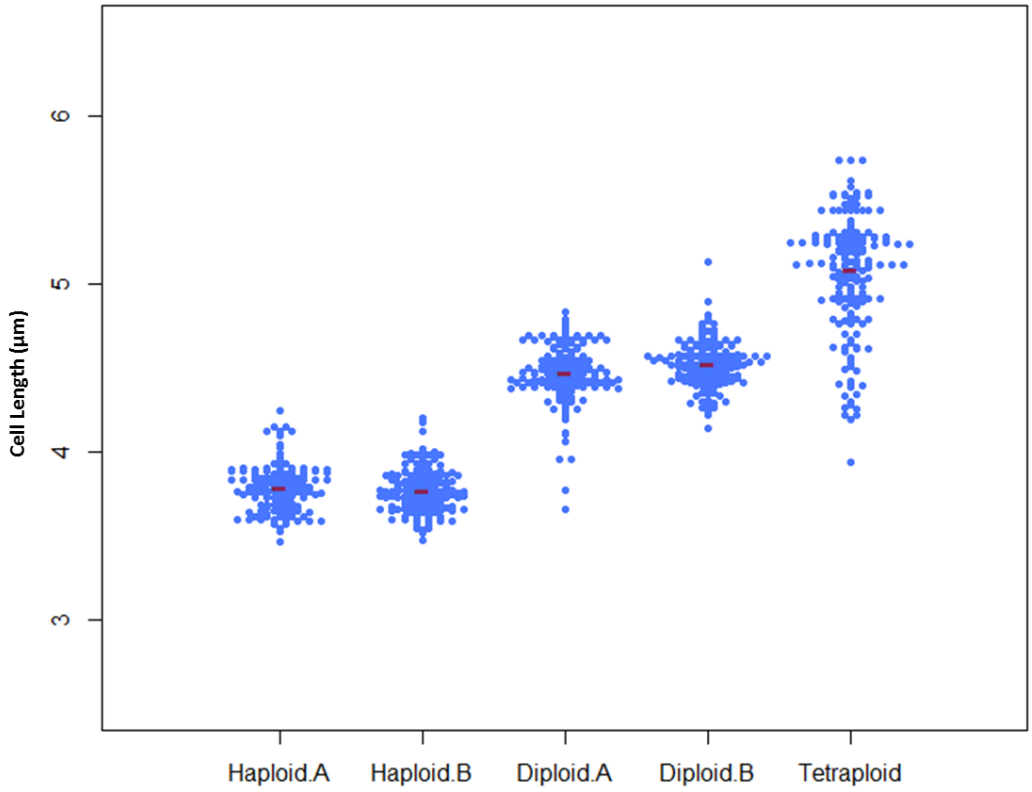


Figure 4.3: Dot plot showing the relationship between ploidy and cell size. Pictures were obtained as described for figure 4.2. Measurements were made using the objectj plugin of ImageJ (NIH), for about 170 septated cells in each group. The X-axis represents the 5 strains analyzed as indicated, while the Y-axis represents cell lengths in log scale. Each dot represents the length for an individual cell, whereas the red bar represents the mean cell length for each strain.

4.4 Gene expression increases with increasing ploidy

Next, we asked, whether a change in ploidy has any effect on cellular processes, such as regulation of gene expression, or whether it simply creates larger cells with more DNA. To examine the cellular control over transcription due to increasing ploidy, we first determined the amount of RNA for each strain as a function of optical density (OD). RNA was extracted from 25 ml of cell culture for each sample at OD 0.5, and the amount of RNA was determined as mentioned in section 2.7.3. The amount of RNA extracted was similar for all these strains (figure 4.4, Table 4.1).

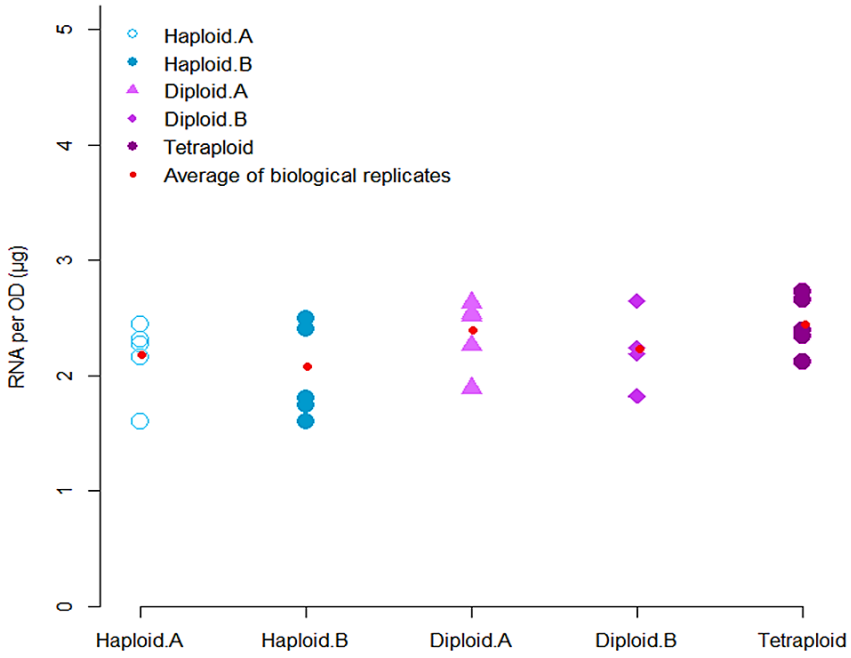


Figure 4.4: Amount of RNA for haploid, diploid and tetraploid cells. 25 ml of cell culture at OD 0.5 was processed for RNA extraction. The RNA pellet obtained was dissolved in 50 µl of DEPC treated water. The amount of RNA/µl was measured at 260 nm using a NanoDrop instrument. Values obtained for 5 independent biological repeats are shown here, where each dot represents an individual value obtained for each repeat. The X-axis represents the strains used for RNA extraction as indicated, while the Y-axis represents the amounts of RNA obtained in µg for same OD cell culture.

We then determined the total amount of RNA per cell for each of the haploid, diploid and tetraploid strains. For this purpose, 25 ml of cells at the same OD were processed for RNA extraction, and the number of cells was determined for each strain as mentioned in secetion 2.6.2. The total amount of RNA extracted was divided by the total number of cells present in 25 ml of a culture. For both diploids, we observed an almost 2-fold (1.85x) increase in the amount of RNA compared to their corresponding haploids. While for tetraploids, instead of an expected 4-fold increase in the amount of RNA, we observed an increase of 3.03-fold as compared to haploids. This lesser amount of cellular RNA in tetraploids could be explained by either the presence of cells of lower ploidies, or the saturation of the RNA extraction system. The amount of RNA per cell clearly established a linear relationship between an increase in ploidy and cellular RNA content (figure 4.5, Table 4.1), which probably reflects increased transcription due to increased copy numbers for each gene.

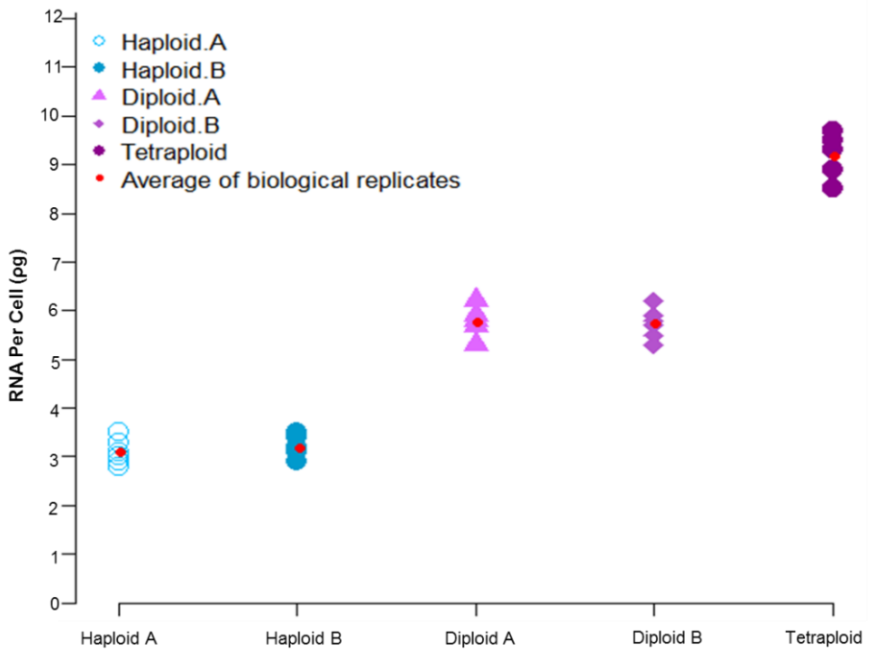


Figure 4.5: Amount of RNA per cell for haploid, diploid and tetraploid cells. 25 ml of cell culture at OD 0.5 was processed for RNA extraction. The RNA pellet obtained was dissolved in 50 µl of DEPC treated water. Number of cells was determined as mentioned in section 2.6.2 by using a Beckman coulter. Amount of

RNA/cell was determined by dividing total amount of RNA in 50 µl by total number of cells used for RNA extraction. Values obtained for 5 independent biological repeats are shown here, where each dot represents an individual value obtained for each repeat. X-axis represents experimental samples used for RNA extraction, while Y-axis represents amount of RNA/cell.

Table 4.1: Amount of RNA in haploid, diploid and tetraploids strains.

	Haploid A	**Haploid B**	**Diploid A**	**Diploid B**	**Tetraploid**
Amount of RNA/OD (µg)	2.18 ± 0.30	2.07 ± 0.40	2.39 ± 0.28	2.22 ± 0.37	2.42 ± 0.23
Amount of RNA/Cell (ρg)	3.1 ± 0.26	3.16 ± 0.25	5.76 ± 0.29	5.73 ± 0.31	9.18± 0.48

4.5 Microarray based analysis of gene expression

The increase in total RNA/cell in polyploids might have some effect on the pattern of gene expression at a global level. To test this idea, we performed microarray-based gene expression analysis for diploids and tetraploids. We performed 6 independent biological repeats for non sporulating polyploids and 2 repeats for sporulating diploids. The ploidy for all strains was confirmed by FACS analysis before proceeding with hybridisation of labelled cDNA to microarrays (figure 4.6).

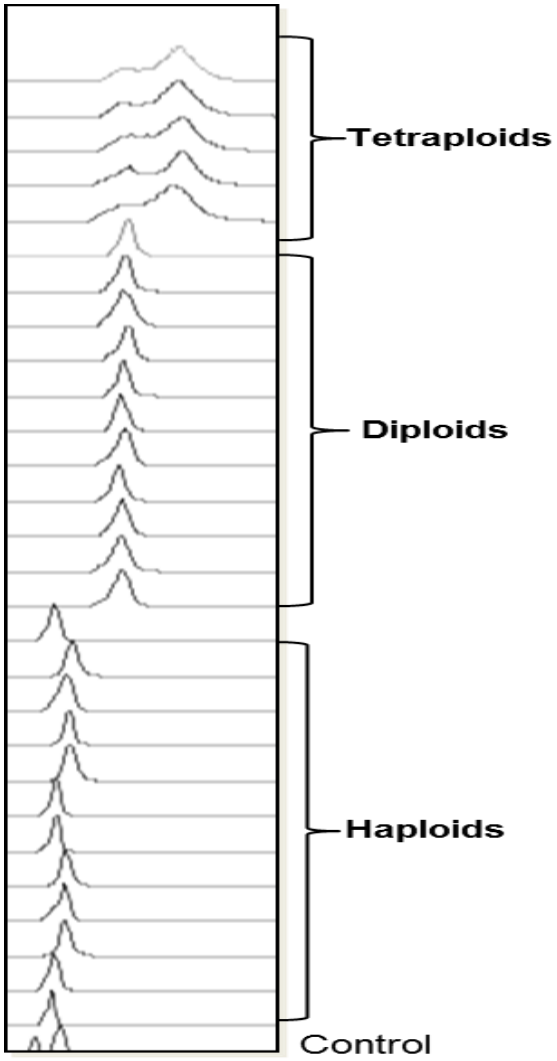


Figure 4.6: FACS analysis for haploids, diploids and tetraploids used for microarray-based analysis of gene expression. Samples were prepared as mentioned in section 2.6.1.

4.5.1 Ploidy does not affects the ratio of total RNA to mRNA

As mentioned earlier (section 4.4), an increase in DNA content results in a proportional increase in the amount of total RNA in polyploid cells. Before proceeding to find out differentially regulated genes, it was important to establish whether it is the mRNA or total RNA that are increased in polyploids. The mRNA constitutes less than 5% of total RNA with remaining 95 % of RNA being largely composed of tRNA and rRNA.

To monitor global changes, if any, in mRNA levels for all genes in polyploids, we incorporated spikes (see section 2.8.1). Spikes are exogenous control RNA molecules, which are added in known quantities in all samples before proceeding for labelling of cDNA as mentioned in section 2.8.2.2. We introduced 9 spikes (controls 1-9), at a concentration that ranged from 1 ρmol/mg to 1000 ρmol/mg, to cover a whole range of mRNA expression levels. The Microarray used had oligonucleotide probes for each spike, spotted several times to generate solid data. As these spikes are not expressed in fission yeast, their expression pattern remains the same amongst haploid and polyploid samples. Their expression ratios are expected to lie along the diagonal when plotted for haploids on the X-axis and polyploids on the Y-axis. If there is a global increase in mRNA levels of all genes, such that it affects the ratio between total RNA to mRNA, then the expression values for genes will lie away from the diagonal (figure 4.7 panels B and C). When we plotted the expression values for haploids and polyploids obtained for all genes along with those of the spikes, they always lied along the diagonal (figure 4.8, shown only for a tetraploid and haploid A). Except for a few genes, the expression values for all genes in polyploids were found to lie along the diagonal, indicating that there was a global increase of all RNA (Figure 4.8a).

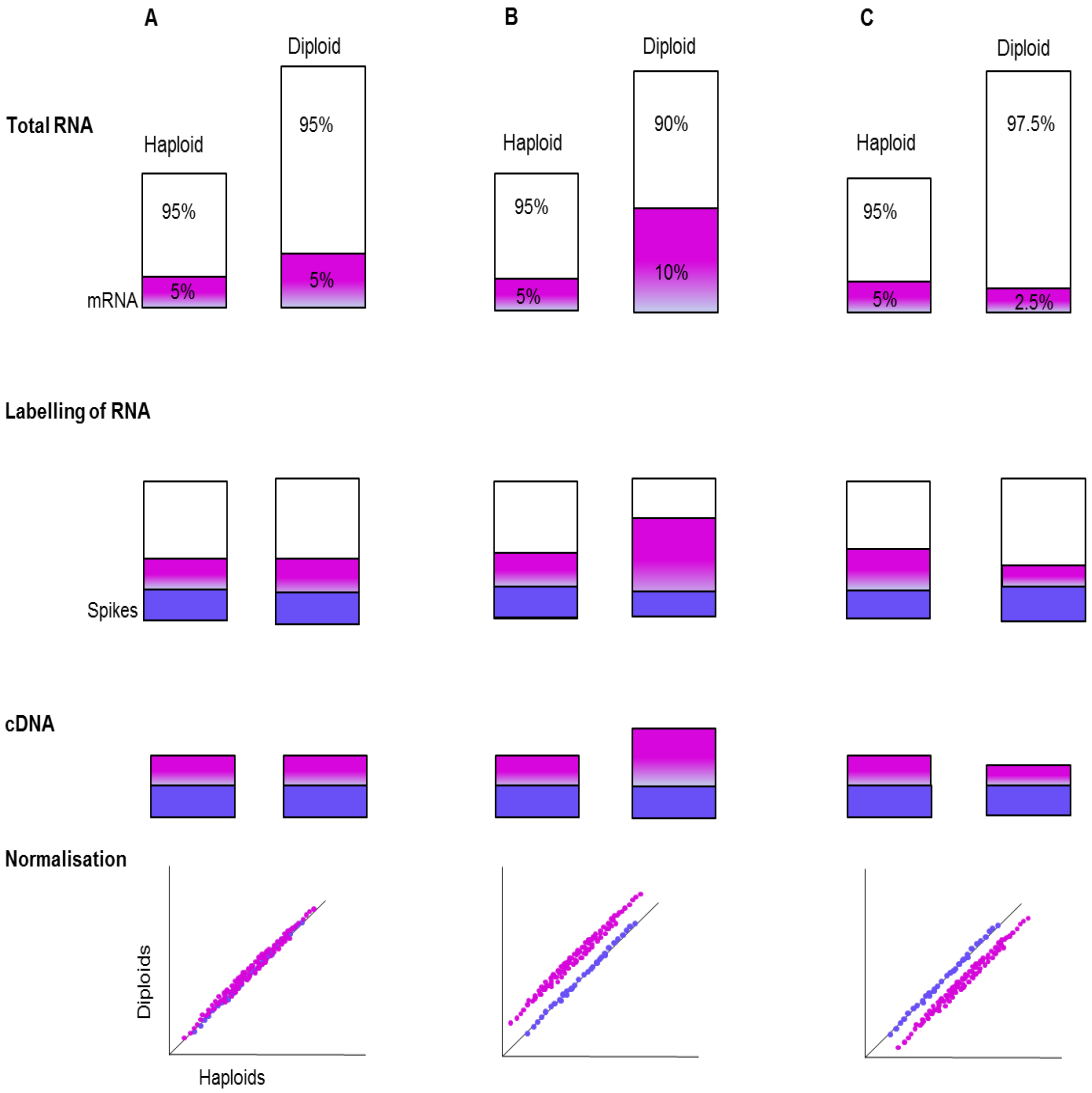


Figure 4.7: schematic representation of global changes in mRNA. Total RNA/cell in diploids is double compared to haploids. Spikes are introduced during labelling of RNA. If the increase in RNA is such that it does not affect the ratio of total RNA to mRNA in diploid cells (A), then expression values for spikes along with genes will lie along the diagonal, when plotted for haploid on x-axis and diploid on Y-axis. If the ratio between total RNA to mRNA is disturbed in diploids such that mRNA is either increased or decreased (B and C), then expression values for spikes will lie along the diagonal but that of genes for diploids will lie above or below the diagonal.

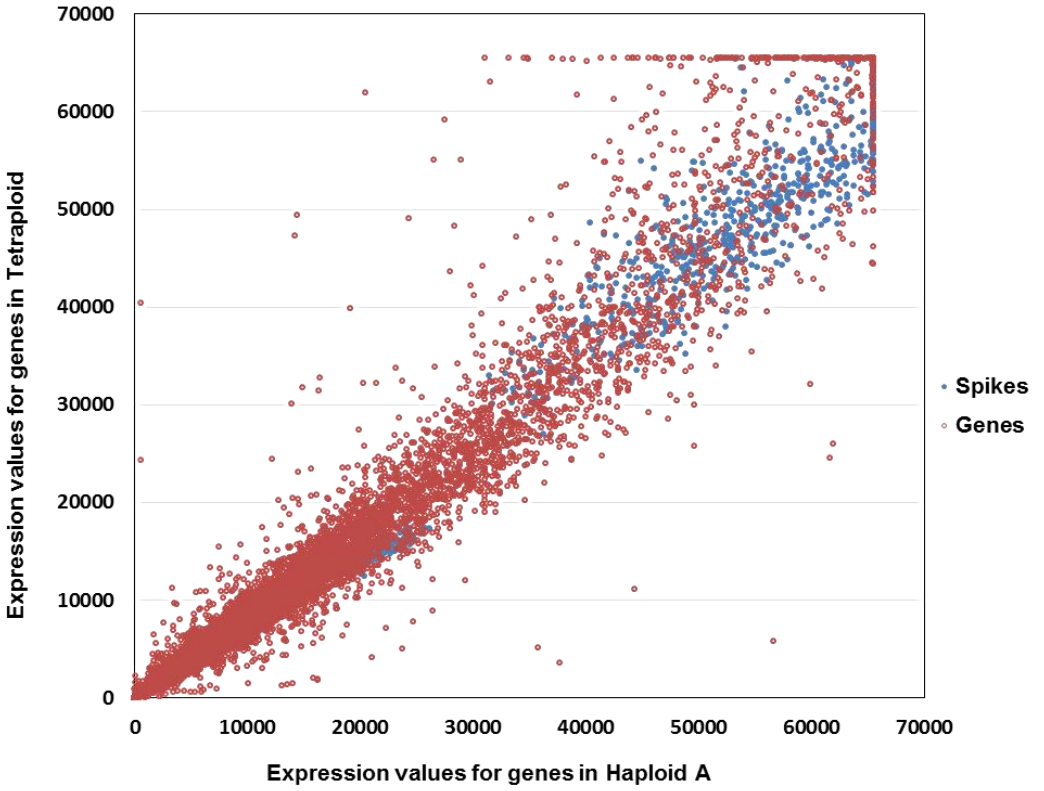


Figure 4.8: Graph showing distribution of expression values for spikes and all genes. Expression values obtained after analysing in GenePix Pro 6.0 (section 2.8.4.1) both for spikes and all genes are plotted for haploid A on the X- axis and tetraploid on the Y-axis.

When signal intensities for all genes obtained in both channels (alexa-647/alexa-555) were normalised with that of spikes using GeneSpring software, they showed an even distribution around 1, showing no large global changes in expression profiles between polyploid and haploid cells (figure 4.9).

Thus the amount of RNA/cell and normalisation of expression values of all genes with spikes indicates that although there is a linear increase in total RNA in diploids and tetraploids, the ratio of total RNA to mRNA has not been affected.

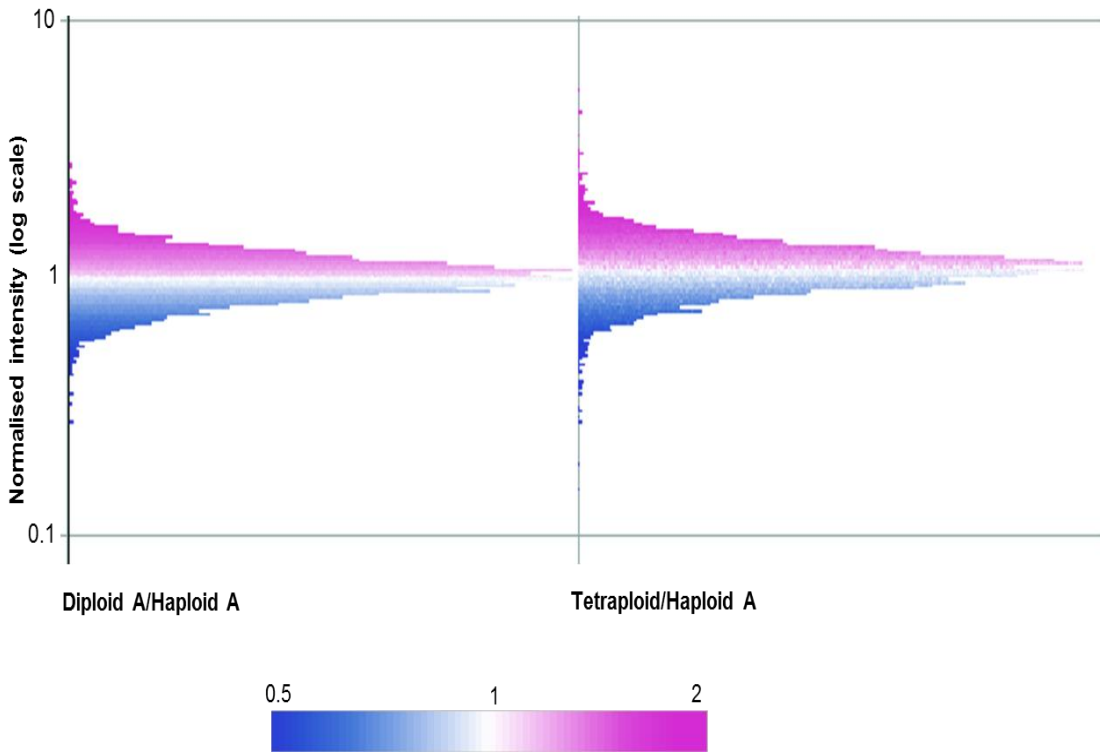


Figure 4.9: Graph showing the distribution of normalised signal intensities for all genes in diploid A and tetraploid cells. Signal intensities for all genes obtained in both channels (alexa-647/alexa-555) normalised with that of spikes in GeneSpring are shown on the Y-axis, while the X-axis represents the samples as indicated. Colour bar shows the range of expression values.

4.6 Selection of differentially regulated genes

To select the differentially regulated genes in polyploids, data obtained from GenePix Pro 6 was further normalised as described in sections 2.8.4 and 2.8.5. For both diploid strains, their corresponding haploid strains were used as a control to normalise their expression values. As tetraploids were constructed by protoplast fusion of diploid A and diploid B, both haploid A and haploid B were used separately as their controls. Six independent biological and one technical repeat each of diploid A/haploid A, diploid B/ haploid B, tetraploid/haploid A and

tetraploid/haploid B were used to perform microarray based gene expression analysis, hence adding up to a total of 28 repeats.

4.6.1 Diploid A

Significance analysis of microarray (SAM) was performed for all 7 repeats of diploid A to select a gene list with a FDR equal to zero. All genes passing this criterion were further analysed in GeneSpring. A cut-off based approach was applied to find genes that were expressed above or below 1.2 fold in at least 2 of 3 of repeats (5 out of 7 repeats). A total of 71 genes were obtained, among them 47 genes were found to be up-regulated, while 25 genes were down-regulated regulated relative to the corresponding haploid strain (Table 4.2 A and B).

These differentially regulated genes were classified into different groups on the basis of their biological functions. Functions of genes were obtained from the *S. pombe* model organism database PomBase (www.pombase.org). This grouping on the basis of biological functions is not rigorous and was adapted to facilitate the understanding of results.

Table 4.2 A: Genes showing 1.2 fold up-regulation in Diploid A

Systematic Name	Common Name	Gene Description
Metabolic genes involved in oxidation reduction		
SPAC222.11	*hem13*	coproporphyrinogen III oxidase (predicted)
SPCC794.12c	*mae2*	malate dehydrogenase (oxaloacetate decarboxylating)
SPAC977.14c		aldo/keto reductase, unknown biological role
SPAC5H10.05c		FAD binding oxidoreductase (predicted)
SPBC8E4.04		aldo/keto reductase involved in pentose catabolism (predicted)
SPBC16A3.02c		mitochondrial peptidase (predicted)
SPBC215.11c		aldo/keto reductase, unknown biological role
SPCC663.06c		short chain dehydrogenase (predicted)
SPAPB24D3.08c		NADP-dependent oxidoreductase (predicted)
SPBC1773.06c		alcohol dehydrogenase (predicted)

SPAC9E9.11	*plr1*	pyridoxal reductase Plr1
SPAC2E1P3.01		dehydrogenase (predicted)
SPAC19G12.09		NADH/NADPH dependent indole-3-acetaldehyde reductase
SPAC6G10.03c		mitochondrial cardiolipin-specific phopholipase (predicted)
SPCC663.08c		short chain dehydrogenase
SPBC2A9.02		NAD dependent dehydratase family protein
SPCC18B5.02c		cinnamoyl-CoA reductase pseudogene
SPBC1289.16c	*cao2*	copper amine oxidase-like protein
Other metabolic genes		
SPAC977.16c	*dak2*	dihydroxyacetone kinase
SPBPB2B2.05		peptidase family protein with hydrolase activity
SPCC965.07c	*gst2*	glutathione S-transferase
SPBC16E9.16c	*lsd90*	Involved in long-chain fatty acid biosynthetic process
SPBPB21E7.04c		human COMT ortholog 2
Cell surface transporters		
SPBC354.13	*rga6*	Rho-type GTPase activating protein (predicted), present on cell surface
SPCC18B5.01c	*bfr1*	brefeldin A efflux transporter, coupled to transmembrane movement of substances
SPCC663.03	*pmd1*	leptomycin efflux transporter present on cell surface
SPAC977.17		MIP water channel(predicted), present on cell surface
Various functions		
SPAC3C7.14c	*obr1*	ubiquitinated histone-like protein,involved in chromatin silencing at silent mating-type cassette
SPAC17A2.13c	*rad25*	14-3-3 protein
SPBC3E7.02c	*hsp16*	heat shock protein
SPBC337.08c	*ubi4*	protein ubiquitination required for meiosis
SPAC16C9.05	*cph1*	Clr6 histone deacetylase associated PHD protein-1
SPCC1223.13	*cbf12*	CBF1/Su(H)/LAG-1 family transcription factor
Unknown function		
SPCC63.13		DNAJ domain protein
SPBC1E8.05		conserved fungal protein
SPBC1271.05c		zf-AN1 type zinc finger protein
SPBC26H8.11c		conserved fungal protein
SPAC186.05c		human TMEM165 homolog
SPAC750.01		hypothetical protein
SPBPB2B2.19c		*S. pombe* specific 5Tm protein family

SPAC750.05c		*S. pombe* specific 5Tm protein family
SPAC977.01		*S. pombe* specific 5Tm protein family
SPAC977.15		dienelactone hydrolase family
SPBC1348.01		*S. pombe* specific DUF999 protein family 5
SPAC212.01c		*S. pombe* specific DUF999 family protein 2
SPAPB24D3.07c		sequence orphan
SPBC17D11.03c		conserved fungal protein

4.6.1.1 Biological functions of up-regulated genes in Diploid A

We then looked into known functions of the differentially regulated genes. Among the 47 up-regulated genes, 18 encoded proteins involved in oxidation reduction reactions. The majority of up-regulated genes were found to code for proteins involved in metabolic processes of cell. Among these, protein products of 18 genes have roles in oxidation and reduction processes. This group of genes includes *mae2* (malate dehydrogenase catalysing the oxidative decarboxylation of malate to pyruvate and carbon dioxide in the presence of $NADP^+$ and divalent cations), *plr1* (pyridoxal reductase catalysing reduction of pyridoxal in the presence of NADPH to pyridoxine and $NADP^+$), *cao2* (member of copper amine oxidase family, but with no apparent role in amine oxidase activity), *hem13* (coproporphyrinogen III oxidase inferred to be involved in heme biosynthetic process). Other members of this group are not experimentally characterised yet, and their roles have been inferred on the basis of sequence homology with other organisms.

Genes that are involved in other metabolic processes include *dak2* (dihydroxy acetone kinase required for metabolism of glycerol), *gst2* (with glutathione transferase activity involved in stress regulation), *lsd90* (required for metabolism of very long chain fatty acid containing phospholipid), and SPAPB21E7.04c that has catechol O-methyltransferase activity (role inferred based on the sequence similarity with human COMT gene).

A few up-regulated genes encode proteins involved in the transport of various substances and located on the cell surface. Examples are *rga6* (probable role in regulation of Rho GTPase activity), *bfr1* (required for multidrug resistance of cells)

and *pmd1*, a structural and functional counterpart of mammalian P glycoprotein which confers multidrug resistance to fission yeast cells.

Up-regulated genes with other functions are *rad25* (involved in check point, meiosis and negative regulation of Byr2, when overexpressed, it reduces mating and sporulation efficiency in homothallic wild type cells) and *cbf12* (involved in cell adhesion and coordination of cell and nuclear division).

14 genes were found to encode proteins whose functions are not yet characterised, with 5 of these genes being present only in *S. pombe.*

Table 4.2 B: Genes showing 1.2 fold down-regulation in Diploid A

Systematic Name	**Common Name**	**Gene Description**
Membrane transporter		
SPBPB10D8.04c		membrane transporter (predicted)
SPBC4B4.08	*ght2*	hexose transporter
SPBC359.01		amino acid permease (predicted)
SPAC139.02c	*oac1*	mitochondrial anion transporter (predicted)
SPBC359.03c	*aat1*	amino acid transporter (predicted)
SPBC359.05	*abc3*	ABC transporter, unknown specificity
SPAC23D3.12		inorganic phosphate transporter (predicted)
SPAC29B12.10c	*pgt1*	glutathione transporter
Amino acid metabolism		
SPBC21C3.08c	*car2*	ornithine transaminase
SPBC359.02	*alr2*	alanine racemase (predicted)
SPBC1271.07c		N-acetyltransferase (predicted)
SPAC11D3.02c		ELLA family acetyltransferase (predicted)
SPCC790.03		rhomboid family protease
Other metabolic genes		
SPAC8F11.10c	*pvg1*	pyruvyltransferase
SPAC4G9.12		gluconokinase
SPAC9E9.09c		aldehyde dehydrogenase (predicted)
SPBP4H10.15		aconitate hydratase/mitochondrial ribosomal protein subunit L49 (predicted)
SPCC330.03c		NADPH-hemoprotein reductase (predicted)
Various functions		
SPAC56F8.14c	*mug115*	sequence orphan
SPCC1884.01		sequence orphan
SPBC1347.11	*sro1*	stress responsive orphan 1
SPAC2F7.11	*nrd1*	RNA-binding protein

SPBC23G7.15c	*rpp202*	60S acidic ribosomal protein P2B subunit
SPBC23G7.12c	*rpt6*	19S proteasome regulatory subunit (predicted)
SPSNORNA.44	*snR92*	small nucleolar RNA

4.6.1.2 Biological functions of down-regulated genes in diploid A

Among 25 genes that showed down regulation in diploid A, two subgroups of genes were evident, named as, membrane transporters and metabolic genes. Eight of such genes encode for proteins involved in transportation of various substrates. This subgroup of genes includes *ght2* (involved in transportation of D-glucose), *abc3* (member of ATP binding cassette (ABC) protein superfamily, involved in transport of various substrates), *oac1* (involved in transport of oxaloacetate mainly inside mitochondria), *aat1* (amino acid transmembrane transporter) and *pgt1* (involved in glutathione import into cell). Exact functions of three of the membrane transporters SPBPB10D8.04c, SPBC359.01 and SPAC23D3.12, are still uncharacterised.

Among 11 metabolic genes, 5 encode protein products involved in amino acid metabolism. This sub group includes genes like *car2* (involved in catabolism of arginine), *alr2* (alanine racemase activity). Functions of other genes in this sub group such as, SPBC1271.07c (protein acetylation) and SPCC790.03 (serine-type endopeptidase activity) still need to be experimentally characterised.

Genes that are sub grouped as other metabolic genes include *pvg1* (pyruvylytransferase required for biosynthesis of N-linked galactosmannans that is a component of fungal type cell wall) and SPAC4G9.12 (involved in D-gluconate metabolism).

Among other genes that are down regulated in diploid A, *sro1* (mitochondrial gene involved in regulation of cellular response under stress conditions) and *nrd1* (negative regulator of cellular differentiation and conjugation) are worth mentioning.

4.6.2 Diploid B

In diploid B, a total of 77 genes passed the cut off of 1.2 fold. Among these, 45 genes were found to be up-regulated, while 32 were down-regulated relative to the corresponding haploid strain.

Table 4.3 A Genes showing 1.2 fold up-regulation in Diploid B

Systematic Name	Common Name	Gene Function
Metabolic genes involved in oxidation reduction		
SPAC977.14c		aldo/keto reductase, unknown biological role
SPBC215.11c		aldo/keto reductase, unknown biological role
SPAC5H10.05c		FAD binding oxidoreductase (predicted)
SPCC663.06c		short chain dehydrogenase (predicted)
SPAC9E9.11	*plr1*	pyridoxal reductase Plr1
SPCC794.12c	*mae2*	malic enzyme, malate dehydrogenase (oxaloacetate decarboxylating)
SPAPB24D3.08c		NADP-dependent oxidoreductase (predicted)
SPBC8E4.04		aldo/keto reductase involved in pentose catabolism (predicted)
SPAC19G12.09		NADH/NADPH dependent indole-3-acetaldehyde reductase
SPCC18B5.02c		cinnamoyl-CoA reductase pseudogene
SPCC663.08c		short chain dehydrogenase
SPBC1289.16c	*cao2*	copper amine oxidase-like protein
SPAC513.07		flavonol reductase/cinnamoyl-CoA reductase family
Other metabolic genes		
SPBPB2B2.13	*gal1*	galactokinase
SPCC965.07c	*gst2*	glutathione S-transferase
SPBC337.08c	*ubi4*	protein ubiquitination required for meiosis
SPAC977.16c	*dak2*	dihydroxyacetone kinase
SPAC6G10.03c		mitochondrial cardiolipin-specific phopholipase (predicted)
SPBC2A9.02		NAD dependent epimerase/dehydratase family protein
SPBC16A3.02c		mitochondrial peptidase (predicted)
SPAC977.15		dienelactone hydrolase family
SPBC800.11		inosine-uridine preferring nucleoside hydrolase (predicted)
Cell surface		
SPAC513.03	*mfm2*	M-factor precursor

SPAC1705.03c	*ecm33*	cell wall protein
SPBC1105.05	*exg1*	glucan 1,6-beta-glucosidase
SPBC29B5.02c	*isp4*	OPT oligopeptide transporter family
Various functions		
SPAC1565.04c	*ste4*	adaptor protein
SPAC27D7.03c	*mei2*	RNA-binding protein involved in meiosis
SPCC338.12	*pbi2*	proteinase B inhibitor (predicted)
SPBC12D12.02c	*cdm1*	DNA polymerase delta subunit
SPAC9E9.13	*wos2*	p23 homolog, predicted co-chaperone
SPAPYUG7.03c	*mid2*	medial ring protein
Unknown function		
SPAC750.01		hypothetical protein
SPBPB2B2.19c		*S. pombe* specific 5Tm protein family
SPAC186.05c		human TMEM165 homolog
SPAPB24D3.07c		sequence orphan
SPBPB21E7.04c		human COMT ortholog 2
SPAC750.05c		*S. pombe* specific 5Tm protein family
SPAC14C4.01c		DUF1770 family protein
SPBC19C7.04c		conserved fungal protein
SPAC977.01		*S. pombe* specific 5Tm protein family
SPAC3C7.14c	*obr1*	ubiquitinated histone-like protein Uhp1
SPCC737.04		*S. pombe* specific UPF0300 family protein 6
SPBC1271.05c		zf-AN1 type zinc finger protein
SPAC167.06c	*mug143*	sequence orphan

4.6.2.1 Biological functions of up-regulated genes in Diploid B

A total of 47 genes were up-regulated in Diploid B, which can be further sub divided in different categories on the basis of functional enrichment. 48% (22 genes) of all up-regulated genes encode proteins involved in various metabolic processes. Among these metabolic genes, 60% (13 genes) code for proteins involved in oxidation-reduction processes of cells, such as *prl1*, *mae2* and *cao2*.

Other genes involved in metabolic processes of cells include *gst2*, *dak2* (functions mentioned earlier), *gal1* (involved in galactose metabolism) and SPBC800.11 (involved in nucleoside metabolism acting on glycosyl bonds).

Genes coding for cell surface proteins include *mfm2* (causes induction of conjugation with cellular fusion during cellular response to nitrogen starvation), *ecm33* (cell wall protein involved in calcium ion homeostasis), *exg1* (cell wall protein involved in 1,6-beta gucosidase activity) and *isp4* (required for tetrapeptide transmembrane transport, and also causes induction of conjugation under nitrogen starvation).

Other genes with known biological functions include *ste4* (adaptor protein involved in pheromone dependent signal transduction required for conjugation by cellular fusion), *mei2* (required for induction of meiosis), *cdm1* (involved in DNA synthesis and DNA repair), *wos2* (involved in regulation of mitotic cell cycle) and *mid2* (acts during cytokinesis to promote cell separation by organizing the septin ring).

About 28% (13 genes) of these induced genes are still uncharacterised. Most of these genes are not conserved and only found in fission yeast.

Table 4.3 B Genes showing 1.2 fold down-regulation in Diploid B

Systematic Name	Common Name	Gene Function
Amino acid metabolism		
SPAC11D3.02c		ELLA family acetyltransferase (predicted)
SPCC965.14c		cytosine deaminase (predicted)
SPBC21C3.08c	*car2*	ornithine transaminase Car2
SPCC790.03		rhomboid family protease
SPAC328.09		mitochondrial 2-oxoadipate and 2-oxoglutarate transporter (predicted)
SPBC19F5.04		aspartate kinase (predicted)
SPCC569.07		aromatic aminotransferase (predicted)
SPBC359.02	*alr2*	alanine racemase Alr2 (predicted)
Other metabolic genes		
SPAC23D3.04c	*gpd2*	glycerol-3-phosphate dehydrogenase Gpd2
SPAC6G10.08	*idp1*	isocitrate dehydrogenase Idp1 (predicted)
SPCC338.11c	*rrg1*	methyltransferase Rrg1 (predicted)
SPBP8B7.15c		ubiquitin-protein ligase E3 RBBP6 family involved in mRNA cleavage (predicted)
SPBC1A4.02c	*leu1*	3-isopropylmalate dehydrogenase Leu1
SPCC330.03c		NADPH-hemoprotein reductase (predicted)
Cell surface/ transporters		
SPAC2E1P3.05c		fungal cellulose binding domain protein
SPAC13G6.10c	*asl1*	cell wall protein Asl1, predicted O-glucosyl hydrolase

SPAC1002.13c	*psu1*	cell wall protein Psu1, beta-glucosidase (predicted)
SPAC139.02c	*oac1*	mitochondrial anion transporter (predicted)
SPBC359.03c	*aat1*	amino acid transporter Aat1 (predicted)
SPCC1529.01		membrane transporter (predicted)
SPBC359.01		amino acid permease (predicted)
SPAPB24D3.02c		amino acid permease, unknown 3 (predicted)
SPBPB10D8.04c		membrane transporter (predicted)
SPBC4B4.08	*ght2*	hexose transporter Ght2
Various functions		
SPCC622.19	*jmj4*	Jmj4 protein (predicted)
SPBC1711.15c		sequence orphan
SPCC1884.01		sequence orphan
SPBC23G7.12c	*rpt6*	19S proteasome regulatory subunit Rpt6 (predicted)
SPAPB1A10.14		F-box protein (predicted)
SPBC23G7.14		sequence orphan
SPBC23G7.15c	*rpp202*	60S acidic ribosomal protein P2B subunit
SPSNORNA.44	*snR92*	small nucleolar RNA snR92

4.6.2.2 Biological functions of down-regulated genes in Diploid B

A total of 36 genes were down-regulated, 14 of which are involved in metabolism of various substrates. Eight genes are involved in amino acid metabolism such as *car2* and *alr2*. The functions of the majority of genes in this group are predicted based on their sequence similarity with genes in other organisms. *gpd2* (involved in glycerol biosynthesis) is also included among the metabolic genes. Other genes included in this group are *idp1*, *rrg1* and *leu1*, along with others, but their functions still need to be experimentally characterised.

Ten genes encode proteins involved in the transport of various metabolites such as *oac1* and *aat1*, or are components of the cell wall such as *asl1*, *psu1* and *ght2*. The functions of majority of these genes are predicted on the basis of the presence of transmembrane domains and sequence homology with transmembrane genes in other organisms.

4.6.3 Comparison between the two diploid strains

To see whether changes in ploidy affect the expression behaviour of similar genes, we compared differentially regulated genes between diploids A and B. For this purpose the Venn diagram feature of GeneSpring was used, to see the overlap between two or three gene lists. Comparing up-regulated genes, 27 genes were found to be identical between the two diploids (figure 4.10, table 4.4 A). The other genes (20 genes in diploid A and 18 genes in diploid B) were only identified as up-regulated in one or the other strain and therefore seem to be strain specific. Among the 27 genes common between the two diploids; 11 encode proteins involved in oxidation reduction processes, 7 encode proteins involved in metabolism of various substrates, and 9 encode proteins of unknown molecular functions. Among 20 genes with induced expression in diploid A only, 7 code for proteins involved in oxidation reduction of various substrates, 4 code for cell surface transporters such as *rga6*, *bfr1*and *pmd1*, while the others are involved in various cellular functions. Genes that showed strain specific overexpression in diploid B include 4 genes coding for various cell surface proteins such as *mfm2*, *ecm33*, *exg1* and *isp4*, while the other genes are involved in various cellular functions, including *ste4*, *mei2*, *pbi2*, *cdm1*, *wos2* and *mid2*.

Similarly, 14 genes with reduced expression were found to be identical between the two diploids (figure 4.10, table 4.4 B). Four of these genes code for proteins involved in amino acid metabolism, while 5 genes code for proteins involved in transport of various metabolites across the membranes, including the well characterised *ght2*. The remaining genes encode for proteins involved in various other activities. The other down-regulated genes (11 genes in diploid A and 17 in diploid B) appear to be strain specific. Genes that are specifically down-regulated in diploid A include 3 genes coding for membrane transporters such as *abc3* and *pgt1*, 4 genes coding for proteins involved in various metabolic processes such as *pvg1* and SPAC4G9.12, while others code for proteins involved in various other processes. Among 17 genes exhibiting reduced expression specifically in diploid B, 4 are involved in amino acid metabolism, 5 code for proteins involved in metabolism of various substrates, including *gpd2*, *idp1*, *rrg1* and *leu1*. Five genes specifically down-regulated in diploid B include cell surface/membrane transporter proteins such

as *asl1* and *psu1*. The remaining down-regulated genes are involved in various functions in these diploids.

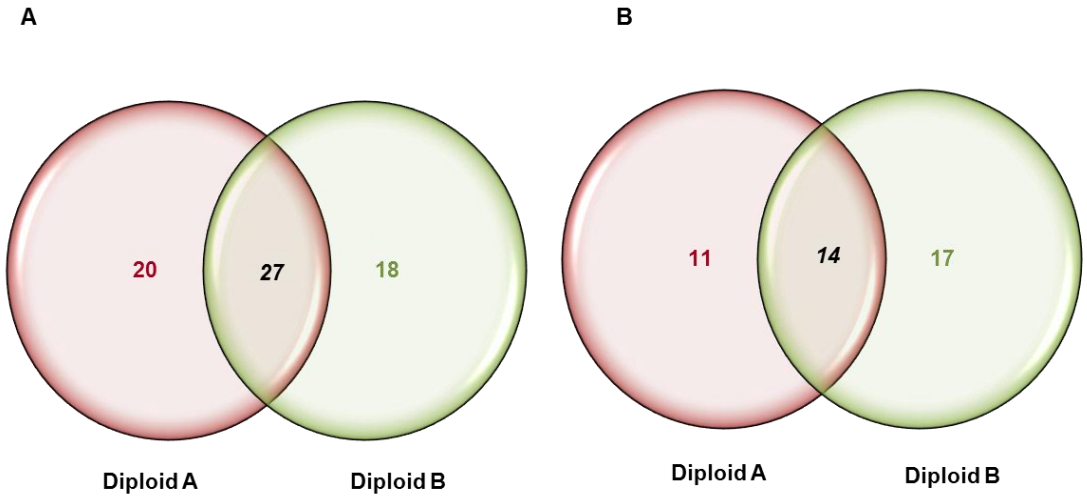


Figure 4.10: Venn diagram showing the overlap of differentially regulated genes in the two diploid strains. The overlap of up-regulated genes (A) and down-regulated genes (B) between Diploid A and Diploid B. Genes that overlap between two diploids are shown as italic and bold in black colour, while genes that are specific to each diploid are shown in individual colours representing each diploid.

Table 4.4 A: Up-regulated genes common between two diploids

Systematic Name	Common Name	Gene Function
Metabolic genes involved in oxidation reduction		
SPAC977.14c		aldo/keto reductase, unknown biological role
SPBC215.11c		aldo/keto reductase, unknown biological role
SPAC5H10.05c		FAD binding oxidoreductase (predicted)
SPCC663.06c		short chain dehydrogenase (predicted)
SPAC9E9.11	*plr1*	pyridoxal reductase Plr1
SPCC794.12c	*mae2*	malic enzyme, malate dehydrogenase (oxaloacetate decarboxylating)
SPAPB24D3.08c		NADP-dependent oxidoreductase (predicted)
SPBC8E4.04		aldo/keto reductase involved in pentose catabolism (predicted)
SPCC18B5.02c		cinnamoyl-CoA reductase pseudogene
SPCC663.08c		short chain dehydrogenase
SPAC19G12.09		NADH/NADPH dependent indole-3-acetaldehyde reductase AKR3C2
SPBC1289.16c	*cao2*	copper amine oxidase-like protein
Other metabolic genes		
SPCC965.07c	*gst2*	glutathione S-transferase
SPBC337.08c	*ubi4*	protein ubiquitination required for meiosis
SPAC977.16c	*dak2*	dihydroxyacetone kinase
SPAC6G10.03c		mitochondrial cardiolipin-specific phopholipase (predicted)
SPBC2A9.02		NAD dependent epimerase/dehydratase family protein
SPBC16A3.02c		mitochondrial peptidase (predicted)
SPAC977.15		dienelactone hydrolase family
Unknown function		
SPAC750.01		hypothetical protein
SPBPB2B2.19c		*S. pombe* specific 5Tm protein family
SPAC186.05c		human TMEM165 homolog
SPAPB24D3.07c		sequence orphan
SPBPB21E7.04c		human COMT ortholog 2
SPAC750.05c		*S. pombe* specific 5Tm protein family
SPAC977.01		*S. pombe* specific 5Tm protein family
SPAC3C7.14c	*obr1*	ubiquitinated histone-like protein Uhp1
SPBC1271.05c		zf-AN1 type zinc finger protein

Table 4.4 B: Down-regulated genes common between two diploids

Systematic Name	Common Name	Gene Function
Amino acid metabolism		
SPAC11D3.02c		ELLA family acetyltransferase (predicted)
SPBC21C3.08c	*car2*	ornithine transaminase Car2
SPCC790.03		rhomboid family protease
SPBC359.02	*alr2*	alanine racemase Alr2 (predicted)
Other metabolic genes		
SPCC330.03c		NADPH-hemoprotein reductase (predicted)
Cell surface/ transporters		
SPAC139.02c	*oac1*	mitochondrial anion transporter (predicted)
SPBC359.03c	*aat1*	amino acid transporter Aat1 (predicted)
SPBC359.01		amino acid permease (predicted)
SPBPB10D8.04c		membrane transporter (predicted)
SPBC4B4.08	*ght2*	hexose transporter Ght2
Various functions		
SPCC1884.01		sequence orphan
SPBC23G7.12c	*rpt6*	19S proteasome regulatory subunit Rpt6 (predicted)
SPBC23G7.15c	*rpp202*	60S acidic ribosomal protein P2B subunit
SPSNORNA.44	*snR92*	small nucleolar RNA snR92

4.6.4 Differentially regulated genes in tetraploids

Significance analysis of microarray was performed for all 14 repeats involving the tetraploid strain to select a conservative gene list with FDR equal to zero. All genes passing this criterion were further analysed in GeneSpring. A total of 95 genes passed the cut off value of 1.2 fold. Among them 81 genes were found up-regulated, while 14 genes were down-regulated (Table 4.5 A and B).

Table 4.5 A Genes showing 1.2 fold up-regulation in tetraploids

Systematic Name	Common Name	Gene Description
Metabolic genes involved in oxidation reduction		
SPAC5H10.05c		FAD binding oxidoreductase (predicted)
SPCC663.06c		short chain dehydrogenase (predicted)
SPAC977.14c		aldo/keto reductase, unknown biological role
SPCC663.08c		short chain dehydrogenase
SPAPB24D3.08c		NADP-dependent oxidoreductase (predicted)
SPAC9E9.11	*plr1*	pyridoxal reductase
SPCC965.06		potassium channel subunit (predicted)
SPBC215.11c		aldo/keto reductase, unknown biological role
SPAC2E1P3.01		dehydrogenase (predicted)
SPAC19G12.09		NADH/NADPH dependent indole-3-acetaldehyde reductase AKR3C2
SPBC16A3.02c		mitochondrial peptidase (predicted)
SPAC3A12.18	*zwf1*	glucose-6-phosphate 1-dehydrogenase (predicted)
SPBC23G7.10c		NADH-dependent flavin oxidoreductase (predicted)
SPAC26F1.14c	*aif1*	apoptosis-inducing factor homolog (predicted)
SPCC18B5.02c		cinnamoyl-CoA reductase pseudogene
Protein metabolism		
SPAC3A11.10c		dipeptidyl peptidase (predicted)
SPCC338.12	*pbi2*	proteinase B inhibitor (predicted)
SPAC1006.01	*psp3*	vacuolar serine protease (predicted)
SPAC4A8.04	*isp6*	vacuolar serine protease
SPAC19G12.10c	*cpy1*	vacuolar carboxypeptidase y
SPAC20G4.03c	*hri1*	eIF2 alpha kinase
Other metabolic genes		
SPCC965.07c	*gst2*	glutathione S-transferase
SPAC6G10.03c		mitochondrial cardiolipin-specific phopholipase (predicted)
SPBPB2B2.13	*gal1*	galactokinase
SPBPB21E7.04c		human COMT ortholog 2
SPBC119.03		human COMT homolog 1
SPBC2A9.02		NAD dependent epimerase/dehydratase family protein
SPACUNK4.16c		alpha,alpha-trehalose-phosphate synthase (predicted)
SPBC16E9.16c	*lsd90*	Involved in long-chain fatty acid biosynthetic process
Cell surface transporters		

SPBC609.04	*caf5*	spermine family transmembrane transporter
SPAC1705.03c	*ecm33*	cell wall protein
SPBC1105.05	*exg1*	glucan 1,6-beta-glucosidase
SPAC11E3.13c	*gas5*	cell wall protein Gas5, 1,3-beta-glucanosyltransferase (predicted)
SPBC725.10		tspO homolog involved in the transport cytoplas/mitochondrial of haem (predicted)
SPCPB1C11.02		amino acid permease (predicted)
SPAC16A10.04	*rho4*	Rho family GTPase
SPAC18B11.04	*ncs1*	related to neuronal calcium sensor Ncs1
Cell cycle regulation		
SPAC3F10.15c	*spo12*	Spo12 family protein
SPBC428.18	*cdt1*	replication licensing factor
SPAC17A2.13c	*rad25*	14-3-3 protein
SPAC20G8.05c	*cdc15*	involve din cytokinesis
SPBC27.04	*uds1*	septation protein
SPAC1565.08	*cdc48*	AAA family ATPase
SPAC8C9.03	*cgs1*	cAMP-dependent protein kinase regulatory subunit
SPBC337.08c	*ubi4*	protein ubiquitination required for meiosis
Various functions		
SPBC1105.14	*rsv2*	transcription factor Rsv2
SPAC3C7.14c	*obr1*	ubiquitinated histone-like protein Uhp1
SPBC1105.12	*hhf3*	histone H4 h4.3
SPAC27D7.09c		But2 family protein
SPBC1348.13		similar to fragment of cox1 intron protein
SPAC19B12.08	*atg4*	Atg8 deconjugator Atg4 (predicted)
SPAC23C4.13	*bet1*	SNARE Bet1 (predicted)
SPBC3E7.02c	*hsp16*	heat shock protein
SPBC16D10.08c		heat shock protein Hsp104 (predicted)
SPAC13G7.02c	*ssa1*	heat shock protein (predicted)
SPAC22H10.13	*zym1*	metallothioneininvolved in zinc and copper ion homeostasis
Unknown function		
SPAC977.01		*S. pombe* specific 5Tm protein family
SPAC186.05c		human TMEM165 homolog
SPBPB2B2.19c		*S. pombe* specific 5Tm protein family
SPAC750.05c		*S. pombe* specific 5Tm protein family
SPAC750.01		hypothetical protein
SPAC167.06c	*mug143*	sequence orphan
SPAPB24D3.07c		sequence orphan
SPBC56F2.06	*mug147*	sequence orphan

SPAC637.03		conserved fungal protein
SPBC19C7.04c		conserved fungal protein
SPCC737.04		*S. pombe* specific UPF0300 family protein 6
SPAC17G6.13	*slt1*	sequence orphan Slt1
SPAC212.01c		*S. pombe* specific DUF999 family protein 2
SPBC1348.01		*S. pombe* specific DUF999 protein family 5
SPAC15E1.02c		DUF1761 family protein
SPBC725.03		conserved fungal protein
SPAC14C4.01c		DUF1770 family protein
SPAC1687.07		conserved fungal protein
SPBC1E8.05		conserved fungal protein
SPAC11D3.13		ThiJ domain protein
SPBC1271.05c		zf-AN1 type zinc finger protein of unknown function

4.6.4.1 Biological functions of up-regulated genes in tetraploids

A total of 81 genes were up-regulated in tetraploids relative to the haploid control strain. These genes can be divided in different groups on the basis of functional similarity. A total of 32 of these genes involved in metabolic processes were further sub-divided into 3 groups. Fifteen of these metabolic genes encode proteins involved in oxidation reduction of various substrates such as *prl1*, while products of 6 genes involved in protein metabolism include *isp6* (vacoular protease involved in cellular protein catabolism), *cpy1* (involved in protein heterodimerisation) and *hri1* (involved in protein phosphorylation and acts as negative regulator of translation initiation in response to osmotic shock). Other well-characterised metabolic genes include *gst2*, *gal1* and *lsd90*. Among these metabolic genes, SPBPB21E7.04c and SPBC119.03, of unknown molecular functions, are worth mentioning. They encode proteins that share homology with human COMT, encoding Catechol-O-methyltransferase. This enzyme catalyses the transfer of a methyl group from S-adenosylmethionine to catecholamines which include neurotransmitters dopamine, norepinephrine and epinephrine. COMT is required for metabolism of catechol drugs used in the treatment of Parkinson's disease, hypertension and asthma (Arts et al., 2013, http://www.ncbi.nlm.nih.gov/gene/1312).

Among the overexpressed genes, 8 encode proteins that are required for cell wall assembly or assist in transport of metabolites across membranes. *Rho4* (required for fungal type cell wall organisation or biogenesis and primary cell septum disassembly), *ncs1* (negatively regulates ascospore formation and also negatively regulates induction of conjugation with cellular fusion), *caf5* (involved in polyamine transport across membrane and plays a role in caffeine resistance) and *gas5* (of unknown molecular function, possibly involved in fungal type cell wall organisation) are also included in this group, along with *ecm33* and *exg1*.

Some genes involved at various steps of the cell cycle are also up-regulated in tetraploids. These genes include *spo12* (involved in progression of mitotic cell cycle), *cdt1* (promotes onset of S phase by facilitating the assembly of an active replication initiation complex), *rad25* (required for mitotic G_2 DNA damage checkpoint and DNA repair), *cdc15* (regulation of mitotic cell cycle and organisation of the actin cytoskeleton to form septum), *uds1* (involved in septation), *cdc48* (involved in segregation of sister chromatids during mitosis) and *cgs1* (negative regulation of meiotic cell cycle, also acts as positive regulator of protein export from the nucleus in response to glucose starvation).

Other genes that showed up-regulation include *rsv2* (that codes for a transcription factor involved in induction of stress related genes during spore formation), *obr1* (coding for a protein involved in chromatin silencing at silent mating-type locus), and *hhf3* (coding for histone).

Twenty-two genes up regulated in tetraploids code for proteins whose functions are not characterised yet, and some of these genes are *S. pombe* specific.

Table 4.5 B Genes showing 1.2 fold down-regulation in tetraploids

Systematic Name	Common Name	Gene Description
Cell surface/membrane transporters		
SPAC2E1P3.05c		fungal cellulose binding domain protein
SPAC1002.13c	*psu1*	cell wall protein, beta-glucosidase (predicted)
SPAC13G6.10c	*asl1*	cell wall protein, predicted O-glucosyl hydrolase
SPBC359.01		amino acid permease (predicted)
Amino acid metabolism		
SPCC790.03		rhomboid family protease
SPBC21C3.08c	*car2*	ornithine transaminase
Various functions		
SPBC23G7.15c	*rpp202*	60S acidic ribosomal protein P2B subunit
SPAPB1A10.14		F-box protein (predicted)
SPBP8B7.15c		ubiquitin-protein involved in mRNA cleavage (predicted)
SPCC622.19	*jmj4*	Jmj4 protein (predicted)
SPBC3D6.03c	*trz2*	mitochondrial 3'-tRNA processing endonuclease
SPCC962.05	*ast1*	asteroid homolog, XP-G family protein of unknown molecular function
SPCC338.11c	*rrg1*	methyltransferase (predicted)
SPCC4E9.02	*cig1*	regulation of mitotic cell cycle

4.6.4.2 Biological functions of down-regulated genes in tetraploids

In tetraploid cells, 14 genes were down-regulated, 4 of which code for proteins required for cell wall assembly and transport across membranes, such as *psu1*, *asl1* and SPAC2E1P3.05c. Other down-regulated genes include *trz2* that codes for endoribonuclease involved in 3'- tRNA processing and *cig1* that is probably involved in regulation of cyclin-dependent protein serine/threonine kinase activity. Other genes encode for proteins which need functional characterisation.

4.6.5 Differentially regulated genes in diploids show similar expression pattern in tetraploid

If the expression pattern of genes (differentially regulated in diploids and tetraploid) would be affected by an increase in ploidy, then these genes should show similar expression patterns among all polyploids. To check for this, we compared gene lists (mentioned earlier in Table 4.2-4.4) by creating Venn diagrams. The Venn diagram for up-regulated genes (figure 4.11 A) showed that 40 genes up-regulated in diploids A and B were also up-regulated in the tetraploid. Twenty-three of these genes, involved in various functions, overlapped in all three polyploids (table 4.6). Interestingly, few genes overlapped between one of the diploid and the tetraploid, but were absent in the other diploid. These include 9 genes overlapping between diploid A and tetraploid but missing in diploid B, and 8 genes overlapping between diploid B and tetraploid but not present in diploid A. Similarly, 8 genes overlapping between the two diploids were missing in the tetraploid, while some genes were found to be specific for each polyploid.

Comparison of down-regulated genes among three polyploids (figure 4.11 B) showed an overlap of 11 genes, only 4 of these genes were common among all (SPBC359.01- probably an amino acid permease; SPCC790.03, a rhomboid family protease; *car2,* an ornithine transaminase; and *rpt6,* a19S proteasome regulatory subunit). 7 genes showed overlap only between diploid B and the tetraploid. Similarly, 10 genes showing overlap between diploid A and B were did not show down regulation in the tetraploid. Other genes were specific for each strain.

For further analysis and follow up study, differentially regulated genes were selected that showed up- or down-regulation in at least 70% of all experimental repeats (20 out of 28). From a total of 50 genes, 45 up-regulated and 5 down-regulated, are included in this list (Table 4.6).

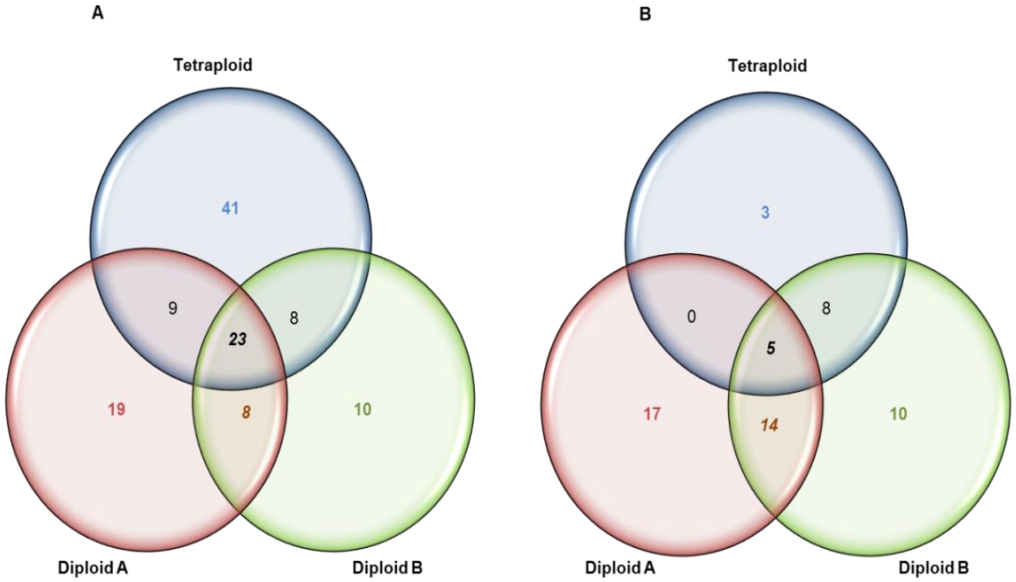


Figure 4.11: Venn diagram showing the overlap of up-regulated genes (A) and down-regulated genes (B) among diploid A, diploid B and tetraploid. Genes that are similar in all three polyploids are shown as italic and bold in black color, while the ones similar between any of diploid and tetraploid are shown in black color. Genes overlapping between two diploids are shown as italic and bold in dark brown color. While genes that are specific to each polyploid are shown in individual colors representing each of polyploid.

Table 4.6: Differentially regulated genes in polyploids

systematic Name	Common Name	Gene Description
Up-regulated Genes		
SPAC750.05c		*S. pombe* specific 5Tm protein family
SPBPB2B2.19c		*S. pombe* specific 5Tm protein family
SPAC186.05c		human TMEM165 homolog
SPAC977.14c		aldo/keto reductase, unknown biological role
SPCC18B5.02c		cinnamoyl-CoA reductase pseudogene
SPCC663.08c		short chain dehydrogenase
SPCC965.07c	*gst2*	glutathione S-transferase
SPAC3C7.14c	*obr1*	ubiquitinated histone-like protein
SPAC750.01		hypothetical protein
SPAPB24D3.08c		NADP-dependent oxidoreductase (predicted)
SPAC5H10.05c		FAD binding oxidoreductase (predicted)
SPAC977.01		*S. pombe* specific 5Tm protein family
SPAC6G10.03c		mitochondrial cardiolipin-specific phopholipase (predicted)
SPBPB21E7.04c		human COMT ortholog 2
SPCC663.06c		short chain dehydrogenase (predicted)
SPBC215.11c		aldo/keto reductase, unknown biological role
SPAPB24D3.07c		sequence orphan
SPBC337.08c	*ubi4*	ubiquitin
SPAC9E9.11	*plr1*	pyridoxal reductase
SPCC338.12	*pbi2*	proteinase B inhibitor (predicted)
SPBC1271.05c		zf-AN1 type zinc finger protein
SPBC3E7.02c	*hsp16*	heat shock protein
SPBC2A9.02		NAD dependent epimerase/dehydratase family protein
SPBC19C7.04c		conserved fungal protein
SPCC737.04		*S. pombe* specific UPF0300 family protein 6
SPAC19G12.09		NADH/NADPH dependent indole-3-acetaldehyde reductase AKR3C2
SPBC725.10		tspO homolog, involved in the transport cytoplas/mitochondrial of haem (predicted)
SPBC16D10.08c		heat shock protein Hsp104 (predicted)
SPBC1E8.05		conserved fungal protein
SPAC977.15		dienelactone hydrolase family
SPBPB2B2.13	*gal1*	galactokinase Gal1
SPAC1705.03c	*ecm33*	cell wall protein
SPAC167.06c	*mug143*	sequence orphan
SPAC17A2.13c	*rad25*	14-3-3 protein Rad25
SPAC2E1P3.01		dehydrogenase (predicted)

SPAC14C4.01c		DUF1770 family protein
SPBC1348.01		*S. pombe* specific DUF999 protein family 5
SPBC16E9.16c	*lsd90*	Lsd90 protein
SPCC965.06		potassium channel subunit (predicted)
SPBC1105.05	*exg1*	glucan 1,6-beta-glucosidase
SPAC22H10.13	*zym1*	metallothionein
SPAC212.01c		*S. pombe* specific DUF999 family protein 2
SPBC16A3.02c		mitochondrial peptidase (predicted)
SPAC2F7.11	*nrd1*	RNA-binding protein
Down-regulated Genes		
SPAC2E1P3.05c		fungal cellulose binding domain protein
SPAPB1A10.14		F-box protein (predicted)
SPBC23G7.15c	*rpp202*	60S acidic ribosomal protein P2B subunit
SPCC965.14c		cytosine deaminase (predicted)
SPCC790.03		rhomboid family protease
SPBC359.01		amino acid permease (predicted)

4.7 Differential regulation of genes is controlled by cell size

During the course of this study, Wu *et al*. reported that in *S. cerevisiae* genes showing differential regulation in polyploids are not regulated by changes in ploidy but by changes in cell size (Wu et al., 2010). They used RNA-seq to identify differentially regulated genes in a tetraploid strain of *S. cerevisiae* and selected genes which show altered expression in the tetraploid as compared to their isogenic haploid strain. A significant number of these genes were found to code for cell surface proteins. The authors suggested that this may be relevant due to the altered surface-to-volume ratio in larger cells of higher ploidy. Next, they examined the expression profiles of these genes in haploid mutants of larger size and found that many of these differentially regulated genes exhibit similar patterns of expression in these haploid mutants as they did in tetraploid cells. These finding indicate a role of cell size, rather than ploidy, in transcriptional regulation.

In this study, we also found a significant number of genes that code for proteins involved in either cell wall assembly or transmembrane transport of various metabolites in both diploid strains and the tetraploid strain. To test whether differentially regulated genes are affected by a change in cell size, we compared their expression in haploid mutants of larger and smaller cell sizes. We selected the haploid mutants Δ*tor1* and *cdc25*-22, which produce enlarged cell size, as well as the mutants Δ*pka1* and *wee1-50,* which produce cells of smaller size. Δ*tor1* produces cells of size 18.4 $\pm$ 1.3 µm and *cdc25-22* of 19.2 $\pm$ 2.3 µm, which is larger than haploids but slightly smaller than diploids used in this study. On the other hand, Δ*pka1* produces cells of size 11.3 $\pm$ 0.71 µm that is slightly smaller than haploids and *wee1-50* of 6.6 $\pm$ 1.1 µm that is almost half the cell size of haploids used in this study.

If differential regulation of genes in polyploids is a result of a change in cell size rather than a change in DNA content, then these genes should exhibit similar expression patterns in haploid mutants of larger size, and the reverse patterns in mutants of smaller size. To test this hypothesis, we analysed the expression profiles of selected genes in haploid mutants (differentially regulated genes common between diploids and tetraploid). For this purpose, microarray expression data for these mutants previously generated in our lab were used. Comparison was done by using

the GeneSpring feature "clustering gene" that groups together the genes showing maximum similarity in gene expression in different conditions. Clustering analysis showed that in Δ*tor1* mutants, most of the differentially regulated genes have a similar expression pattern, while in *cdc25-22* few genes have a similar expression pattern to polyploids.

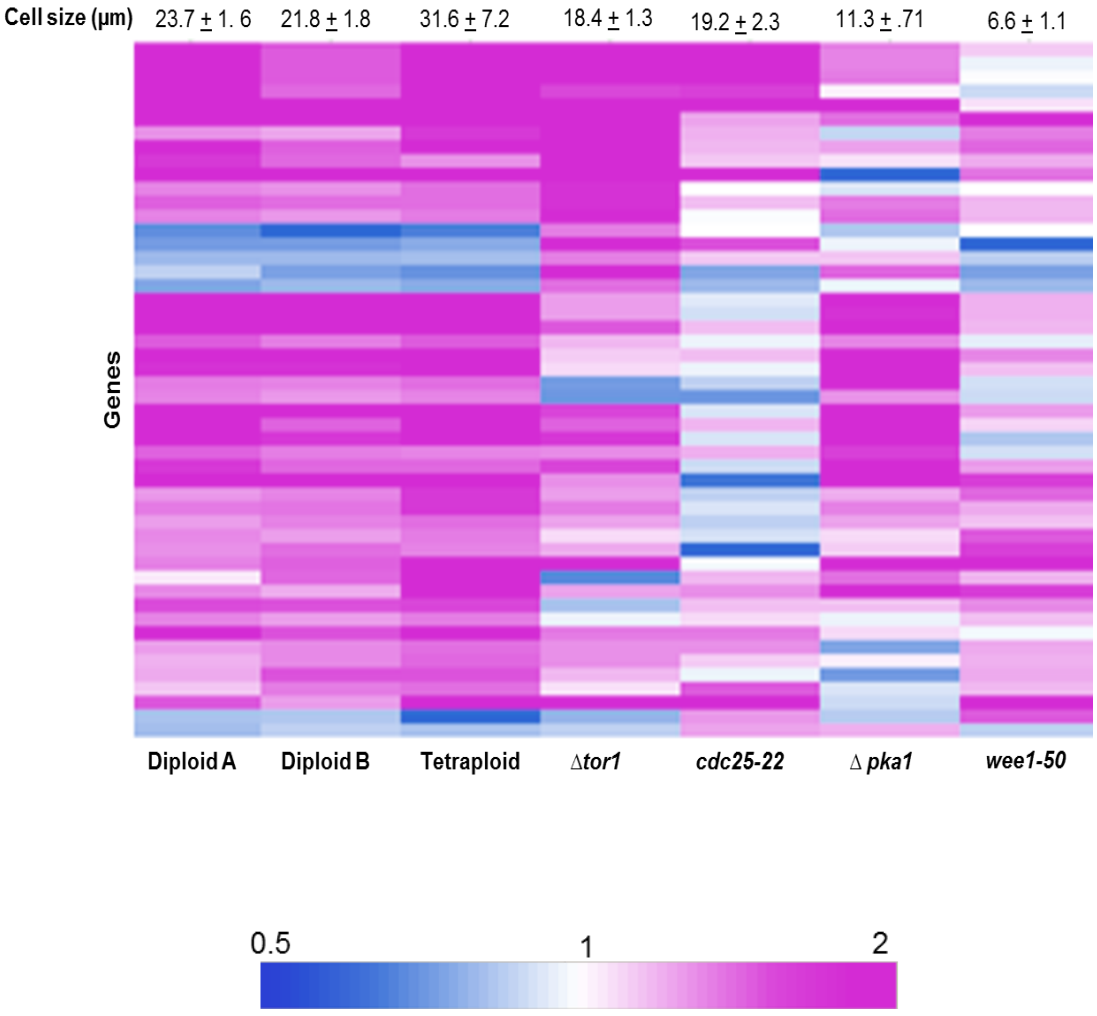


Figure 4.12: Clustering of differentially expressed genes in polyploids. Each column represents an experimental sample, *i.e.,* different polyploids and haploid mutants as indicated, while rows represent genes. The size of each sample is mentioned on top of the column. Pink colour represents up-regulation, and blue represents down-regulation; white: no change in expression (see colour bar)

As *cdc25-22* showed different expression patterns for many genes, we looked for GO enrichment of those particular genes. These genes are mainly involved in metabolic processes and have previously been reported to be involved in stress regulation in fission yeast. When we compared gene expression profiles of differentially regulated genes, but excluding stress responsive genes, most up-regulated genes show similar expression patterns in haploid mutants of larger size (figure 4.13). In mutants of smaller size, like *wee1-50,* some genes showed similar expression patterns, while few genes either did not show any change in expression.

In *wee1-50* mutants, many of these genes either show down-regulation or no change in expression. These genes include SPBPB2B2.19c, SPAC977.01, SPAC212.01c, SPAC5H10.05c and SPBC1E8.05. The first three genes encode proteins of unknown biological function but are probably located in the plasma membrane or in membranous organelles, as they are predicted to contain one or more transmembrane domains. The protein product of SPAC5H10.05c is predicted to be a FAD binding oxidoreductase, but its exact function needs to be explored. Similarly, the function of protein product of SPBC1E8.05 is also unknown but is predicted to be a cell surface protein due to the presence of a Glycosylphosphatidylinositol (GPI) domain. In Δ*pka1*, many genes up-regulated in polyploids are also up-regulated in this mutant

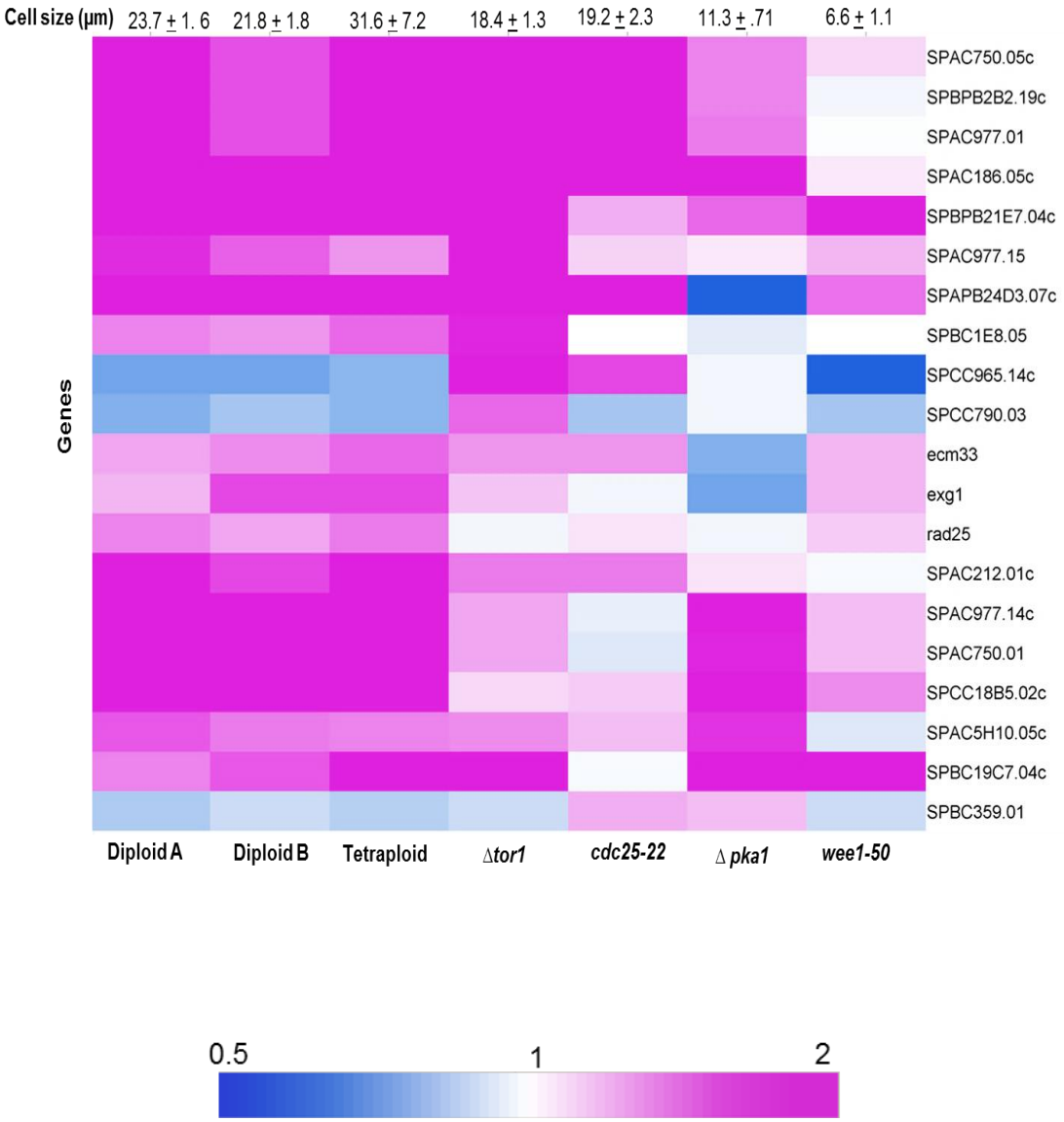


Figure 4.13: Clustering of differentially expressed genes in polyploids, excluding stress responsive genes. Each column represents an experimental sample as indicated, including ploypoids and haploid mutants, with their respective cell sizes shown on top of each column, while rows represent genes. Colour code as in figure 4.12.

4.8 Conservation of ploidy or cell size-regulated gene expression

As fission yeast and budding yeast are only very distantly related, they represent a good complementary system for identification of conserved mechanisms among eukaryotes. To explore whether the function and regulation of the differentially expressed genes identified in our study are conserved between the two yeasts, we compared our genes with those reported for budding yeast. Two different studies in budding yeast have identified genes showing differential regulation by using either microarrays or RNA-Seq (Galitski, 1999; Wu et al., 2010). These two studies identified 17 genes and 65 genes that are differentially regulated in polyploids. For comparison, both budding yeast gene lists were used: Galitski's list containing 17 genes (Galitski, 1999) and Wu's list containing 65 genes (Wu et al., 2010).

Homologous *S. cerevisiae* genes were identified by using a web-integrated resource "eukaryotic Orthology" (YOGY) listed on the Bähler group website (http://www.bahlerlab.info/resources), which searches for orthologous proteins in 11 organisms including *S. cerevisiae* (Penkett et al., 2006). For most genes identified here, no orthologs were detected between the two organisms. No overlap was found for differentially expressed genes described in this study and those for budding yeast. We conclude that the genes that are differentially expressed as a function of ploidy and/or cell size are not necessarily conserved.

4.9 Conclusion:

In this study, we tried to explore the impact of polyploidy on regulation of fission yeast gene expression. We used *ade6* mutants for construction and easy selection of polyploids. To analyze the effect of only cell ploidy but not strain background, we generated a series of polyploids that share the genotype at the mating type locus. For this purpose, we generated 2 different diploids, named as diploid A and diploid B, which originated spontaneously, probably as a result of endoreplication. The tetraploid strain, on the other hand, was constructed by the protoplast fusion method. Using this approach, we successfully generated isogenic series of polyploids consisting of 2 diploids and 1 tetraploid.

Diploids showed normal morphology when compared to haploids, while few tetraploid cells appeared irregular in shape with multiple septa. Diploids showed 4C DNA content corresponding to G2 phase, hence confirming their 2N ploidy state. Tetraploid cells, on the other hand, showed a range of ploidy from 2N to 4N, corresponding to 4C to 8C DNA peaks by FACS analysis, suggesting the presence of cells of lower ploidies amongst these cells. Cell lengths of both diploids and the tetraploid showed a linear increase with ploidy, but the tetraploid, reflecting its unstable nature, showed cell lengths ranging from those expected for tetraploid cells to as small as that of diploid cells. This finding is consistent with the tetraploid strain readily losing chromosomes and reverting back to lower ploidies. The amount of RNA as a function of the same OD could not clearly establish the role of ploidy on cellular transcription. However, we observed a strong linear correlation between ploidy and the amount of RNA, i.e., transcription/cell. Using spike-in controls and normalization approaches, we could show that in polyploids increased transcription does not affects the ratio between total RNA and mRNA.

Very few genes showed differential regulation in diploids and tetraploid when compared to their respective haploid counterparts. When analysed for their GO enrichment, these genes showed enrichment for those encoding proteins involved in oxidation reduction of various metabolites and other metabolic processes. Interestingly, several of these differentially regulated genes were found to encode cell surface proteins that are involved in either cell wall biosynthesis or transport of various metabolites across the membrane. Comparison of differentially regulated

genes with haploid mutants of larger cell sizes, indicate a role of cell size, rather than of ploidy, in the regulation of these genes.

We could not find any overlap between genes showing differential regulation in fission yeast polyploids with those previously described in budding yeast, although in both yeasts cell surface proteins are differentially expressed as a function of cell size. This intriguing finding could be explained by the altered cell surface-to-cell volume ratio that accompanies altered cell size, and the amount of cell surface proteins may therefore need adjusting to altered cell size.

Chapter 5

5 Analysis of genes that are differentially expressed as a function of cell size

This chapter describes the phenotypic effects conferred by the deletion and overexpression of genes that were found to be differentially expressed as a function of cell size in chapter 4 (Table 5.1). Phenotypic effects of these genes were analysed on various aspects of cellular physiology, such as cell growth in response to various stresses, but the main focus is on cell size. All the results mentioned here are derived from at least two independent biological repeats

5.1 Introduction

In a microarray based gene expression analysis of the polyploid strains, we found some genes that showed differential regulation in response to increased cell size (Table 4.6, figure 4.13). Interestingly, these genes are known or predicted to encode cell surface or transmembrane proteins (Table 5.1), which is consistent with the idea that these genes might be regulated in response to altered cell size rather than ploidy. We chose a few of these genes (genes 1 to 6, listed in table 5.1) that were highly expressed in the polyploids as well as in haploid mutants of larger cell size. These genes were further analysed for their effects on cellular physiology, with an emphasis on cell size.

Deletion and overexpression of particular genes represent powerful molecular genetic tools to explore the phenotypic affects and hence the function of a gene. Often, the first indication of gene function comes from the phenotype conferred by its deletion. We used these tools to explore the relationships, if any, between these genes and cell size. The idea was that if these genes were differentially regulated in response to an increase in the cell size, then their deletion and overexpression might in turn result in a decrease or an increase in the cell size, respectively. These genes were deleted and overexpressed by using the one step PCR-based approach (Bähler et al., 1998), as discussed in section 2.4.

Table 5.1: Highly expressed genes in larger cells

No	systematic Name	Common Name	Gene Description
1	SPAC186.05c		human TMEM165 homolog
2	SPBPB2B2.19c		*S. pombe* specific 5Tm protein family
3	SPAC977.01		*S. pombe* specific 5Tm protein family
4	SPAC750.05c		*S. pombe* specific 5Tm protein family
5	SPBC1348.01		*S. pombe* specific DUF999 protein family 5
6	SPAC212.01c		*S. pombe* specific DUF999 family protein 2
7	SPAC1705.03c	*ecm33*	cell wall protein
8	SPBC1105.05	*exg1*	glucan 1,6-beta-glucosidase

Literature search indicates that two of the genes, *exg1* and *ecm33*, play a role in cell wall assembly and biosynthesis (Dueñas-Santero et al., 2010; Takada et al., 2010). The gene *exg1* encodes glycoside hydrolase, located in the periplasmic space, and is involved in the cleavage of β (1, 6)-glucans - a major component of the fission yeast cell wall. Deletion of *exg1* did not show any phenotype when compared to wild type cells, which was attributed to the redundancy of its function due to the presence of two other Exg proteins encoded by *exg2* and *exg3* (Dueñas-Santero et al., 2010). On the other hand, deletion of *ecm33*, encoding a glycosyl-phosphatidylinositol (GPI)-anchored cell surface protein, produced cells of a smaller size (Takada et al., 2010).

Other genes mentioned in table 5.1 (SPAC186.05c, SPBPB2B2.19c, SPAC977.01, SPAC750.05c, SPBC1348.01 and SPAC212.01c) have not been characterised yet. Except for SPAC186.05c, encoding a human TMEM165 homolog, all others are found only in *S. pombe*. These genes are predicted to encode membranous proteins that could be integral to the plasma membrane, due to the presence of transmembrane helix domains in their sequences (www.pombase.org). As *exg1* and *ecm33* have already been characterised, we decided to exclude them in this study, and instead focused on genes 1 to 6 mentioned in Table 5.1.

5.2 Presence of paralogous genes

The *S. pombe* genome database indicated that the genes SPBPB2B2.19c, SPAC977.01, SPAC750.05c, SPBC1348.01 and SPAC212.01c are present in regions close to telomeres that contain duplicated genes (www.pombase.org). If a gene shows high similarity in sequence with other genes (paralogs) then its phenotypic effect could be masked by the presence of redundant proteins encoded by other paralogs. To find whether these genes have any potential paralogs, we carried out a BLAST search of the *S. pombe* genome, by using the BLASTN programme provided at the www.pombedb.org (replaced now by www.pombase.org). This search indicated the presence of 1 other paralog for SPAC186.05c (figure 5.1), 4 other paralogs for SPBPB2B2.19c (including two of the highly regulated genes, SPAC977.01 and SPAC750.05c, table 5.1) (figure 5.2), and 9 paralogs for SPAC212.01c that also includes the SPBC1348.01 (figure 5.3). Sequences of the above mentioned genes and their paralogs were aligned by using the web integrated resource "Clustal Omega" (www.ebi.ac.uk/Tools/msa/clustalo). We also used this programme to create the Percent Identity Matrix (PIM) that gave information about the sequence similarity among different paralogs.

SPAC186.05c and SPAC17G8.08c show 57% sequence similarity. Except for one (SPAC977.02 that shows 48.5% sequence homology), all other paralogs of SPBPB2B2.19c show more than 99% sequence similarity with each other (Appendix VI: table 1a and figure III a). Similarly, for SPAC212.01c all paralogs show more than 80 % sequence homology (Appendix VI: table 2a, figure III b).

The presence of such paralogs could raise the question of whether higher expression of these genes is a result of multiple DNA copies, or of an increase in ploidy and/or cell size. However, the normalisation methods (mentioned earlier in section 2.8.4) and controls used remove any such effects that are caused by an increase in DNA copy number of genes. These methods normalise the expression of all genes expressed two- or four-fold in diploids and tetraploids, respectively, against the haploid control, by assuming that only few genes show differential expression. Therefore, expression of these genes is induced in response to increasing ploidy and/or cell size rather by the presence of paralogs.

Due to the higher similarities among paralogs, it is difficult to assess which of the paralogs has shown the higher expression. We therefore mapped the location of microarray probes to see whether they are unique for each paralogous gene. For SPAC186.05c, we observed that the probe is unique and can clearly distinguish between two paralogs (figure 5.1). Therefore, it is the SPAC186.05c that showed higher expression in our study and not the SPAC17G8.08c.

.

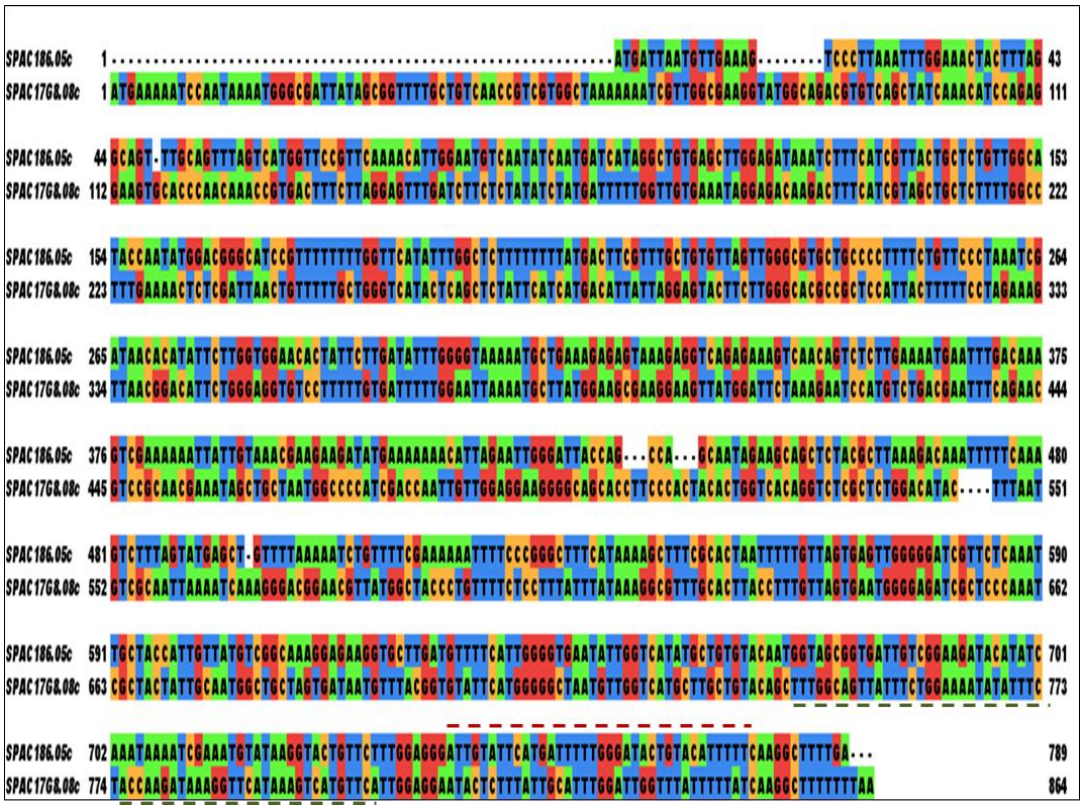


Figure 5.1: Sequence alignment for SPAC186.05c and its paralog. DNA sequences for coding regions of both genes, downloaded from www.pombase.org, were aligned by using the web based programme Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo/). For better visualisation and to colour the aligned sequences on the basis of nucleotides, Jalview software was used (www.jalview.org/). Red and green dashed lines indicate the probe sequences for SPAC186.05c and for SPAC17G8.08c, respectively.

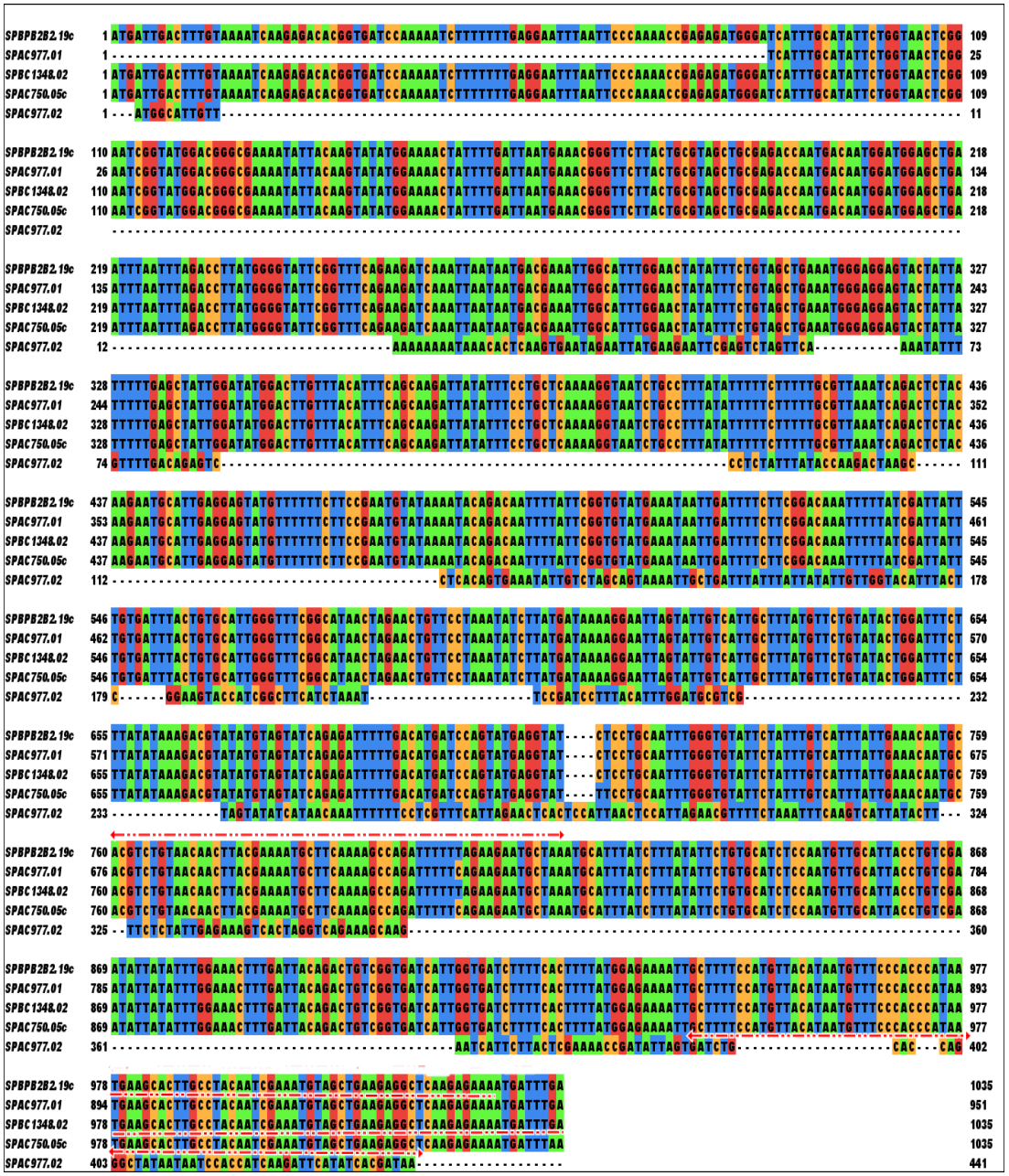


Figure 5.2: Sequence alignment for SPBPB2B2.19c and its paralog. Sequence alignment and further analysis was performed exactly as mentioned for figure 5.1. Red lines above a particular sequence indicate the sequence of a probe for those particular genes on a microarray. Absence of a probe in this figure indicates that similar probes were used for these as for SPBPB2B2.19c.

Figure 5.3: Sequence alignment for SPAC212.01c and its paralog. Sequence alignment and further analysis was performed exactly as mentioned for figure 5.1. Red lines above a particular sequence indicate the sequence of a probe for those particular genes on a microarray. Absence of a probe in this figure indicates that similar probes used for these as for SPAC212.01c.

For paralogous genes that show sequence homology of more than 99%, probe sequences are similar, thus making it impossible to know which of the paralog is highly expressed in this study.

These genes and their paralogs have been given common names in this study (Table 5.2). As our data so far and further down suggest that these genes are up-regulated as a function of increased cell size, they have been named as **U**p regulated in response to **C**ell **L**ength (ucl).

Table: 5.2 Common names given to the genes of interest and their paralogs in this study

Gene	Paralogs	Name given
SPAC186.05c		ucl1-a
	SPAC17G8.08c	ucl1-b
SPBPB2B2.19c		ucl2-a
	SPAC977.01	ucl2-b
	SPBC1348.02	ucl2-c
	SPAC750.05c	ucl2-d
	SPAC977.02	ucl2-e
SPBC1348.01		ucl3-a
	SPAC212.01c	ucl3-b
	SPBCPT2R1.01c	ucl3-c
	SPBCPT2R1.04c	ucl3-d
	SPAC212.04c	ucl3-e
	SPAC750.06c	ucl3-f
	SPAC977.06	ucl3-g
	SPBC1348.07	ucl3-h
	SPBPB2B2.07c	ucl3-i
	SPBPB2B2.14c	ucl3-j

To establish the effects of the genes on cell size, cell growth and stress response, we deleted and overexpressed *ucl1-a, ucl2-a, ucl3-a* and *ucl3-d* in haploid B and diploid B as mentioned in section 2.4. Deletion and overexpression of these genes was confirmed by PCR. The next section describes in detail the phenotypic effects of deletion and overexpression of these mutants on cellular physiology.

5.3 Effects of gene deletions on cell size

5.3.1 Haploid deletion mutants

To test whether the deletion of any of these genes (*ucl1-a, ucl2-a, ucl3-a* and *ucl3-d*) had any effect on cell size, all the deletion mutants along with the haloid B (control) were grown to mid log phase in EMM. As these cells carry an *ade6* mutation, EMM was supplemented with adenine. Cells were then collected, washed and fixed for microscopy as mentioned in section 2.5. Cell lengths were measured for at least 150 septated cells by using the ImageJ plugin object (NIH) as mentioned in section 2.5.3. When plotted, we observed that these deletion mutants show a slight decrease in cell size as compared to their haploid control. Similar results were observed for the independent biological replicates (figure 5.4). The mean cell sizes for the two biological repeats were as follows: for haploid B, 15.45 μm (±1.15 μm) and 15.21 μm (±1.13 μm); for Δ*ucl1-a,* 14.11 μm (±0.93 μm) and 14.34 μm (±0.94 μm); for Δ*ucl2-a* 13.45 μm (±0.94 μm) and 13.51 μm (±0.91 μm); for Δ*ucl3-a,* 13.83 μm (±1.15 μm) and 14.46 μm (±1.34 μm); and for Δ*ucl3-d,* 14.87 μm (±1.02 μm) and 15.05 μm (±1.06 μm).

Cell sizes of two biological replicates of Δ*ucl1-a* were found to be ~91% and ~94% of the size of haploid controls 1 and 2, respectively (where 1 and 2 denotes the two biological replicates). Similarly, biological replicates of Δ*ucl2-a,* Δ*ucl3-a* and Δ*ucl3-d* showed cell sizes that were about 87 and 88%, 89.5 and 95%, and 96 and 98% of the cell sizes of two biological replicates of haploid B, respectively.

To check whether the mean cell size of each deletion mutant differs significantly from that of the control, we applied the Wilcoxon rank test to calculate the p-value for each mutant compared with their respective haploid control. All the mutants showed significant decreases in cell size as compared to their respective haploid control, except for the 2nd biological replicate of Δ*ucl3-a*. In a third repeat, we observed a similar trend as for the first repeat (data not shown here).

These findings suggest that the deletion of either *ucl1-a, ucl2-a, ucl3-a,* and *ucl1-d* result in a decrease in cell size of their respective mutants, as compared to the control.

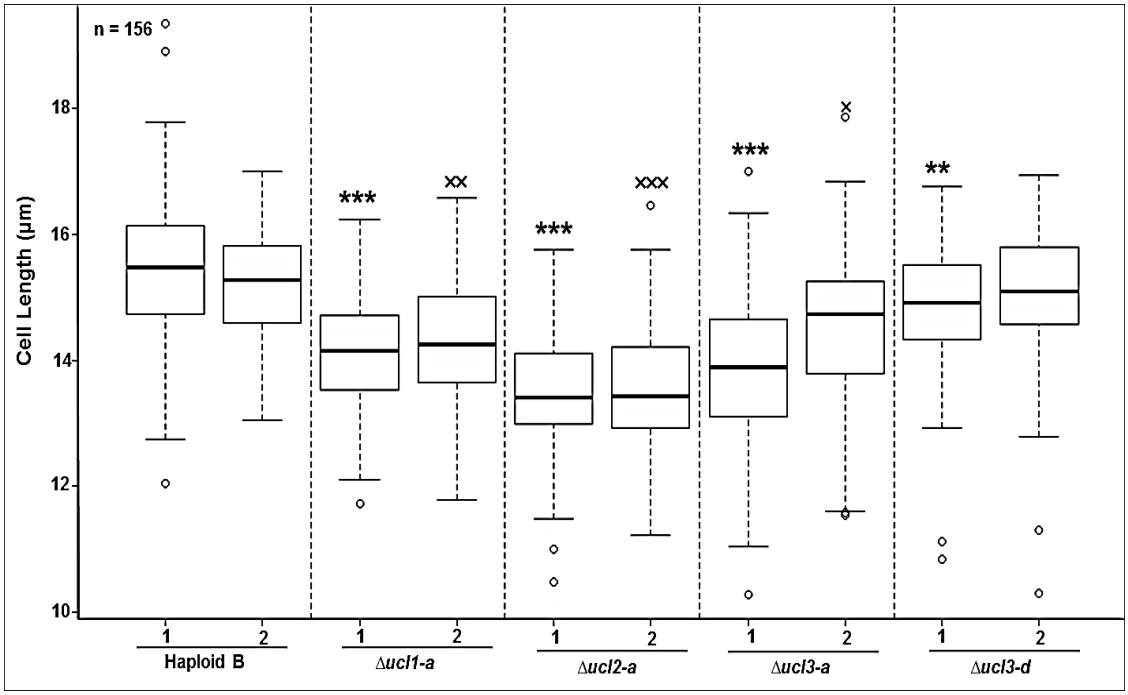


Figure 5.4: Box plots showing the cell sizes for haploid deletion mutants and controls. Pictures of the cells were obtained as discussed in section 2.5. Measurements were made using objectj plugin of ImageJ (NIH), for about 156 septated cells for each sample. The X-axis represents the different mutant strains along with the control (haploid B), where 1 and 2 correspond to the two independent biological repeats, while the Y-axis represents the cell lengths on a linear scale. The horizontal line inside each box represents the mean of the cell lengths, whereas individual dots above or below the whiskers represent the outliers. * and ˣ represent the p -values calculated for the mutants and control of 1st and 2nd biological repeat, respectively. ***, ˣ ˣ ˣ represent the p-values *<2.2e-16, ** p-value = 0.0002, ˣ ˣ p –value <4.8e-11 and ˣ p = 2.6e-06.*

5.3.2 Diploid deletion mutants

To see whether halving the number of the above mentioned genes result in the reduction of cell size in diploids, we deleted one copy of each of these genes in diploid B, thus making these diploids hemizygous for the deleted genes. Cells were grown to mid log phase in EMM supplemented with adenine and were processed for fixation, imaging and cell length measurements as mentioned earlier in section 2.5. When analysed for their cell size, we could not observe any significant decrease in the cell size of these diploid mutants, as compared to the diploid B used as a control (data not shown here). One reason could be that deletion of only one copy of these genes does not affect the cell size in diploid mutants, especially since several paralogous genes could back up their functions.

5.3.3 Double deletions in haploids

As these genes have paralogs in the genome, it is possible that they might act in a dose dependent manner to affect the cell size. To check whether these paralogs have any such effect, we constructed double deletion mutants. We used Δ*ucl1-a,* Δ*ucl2-a,* Δ*ucl3-a,* and Δ*ucl3-d* as background strains to delete another copy of these genes by using a pFa6MX-Nat cassette for selection as mentioned in section 2.4. We successfully constructed several double mutants as follows: Δ*ucl1-a* Δ*ucl2-a*; Δ*ucl2-a* Δ*ucl3-a*; Δ*ucl2-a* Δ*ucl2-b*; *ucl2-a* Δ*ucl3-d*; Δ*ucl3-a* Δ*ucl3-d*; and Δ*ucl3-a* Δ*ucl2-a*. These double mutants were confirmed with the help of PCRs of the deletion junctions.

Cells were grown and processed for measurement of cell length as mentioned previously. All samples were processed at the same time; therefore we show the cell size of the control obtained from only one repeat. The mean cell size of haploid B was 15.63 μm (±1.02 μm), while for Δ*ucl1-a* Δ*ucl2-a* it was found to be 15.09 μm (±1.45 μm) and 15.38 μm (± 1.21 μm) for the two independent biological repeats (figure 5.5). In the first repeat, the double mutant showed a slight decrease in cell size (size of the double mutant was 96% of the size of control), whereas in the second repeat there was no noticeable effect on cell size. Δ*ucl2-a* Δ*ucl3-a* showed cell sizes of 92 and 92.5%, Δ*ucl2-a* Δ*ucl2-a* of 90.3 and 90%, and *cl2-a* Δ*ucl3-d* of

94 and 94.2% of the size of the control for the two repeats, respectively. Similarly, cell sizes for Δ*ucl3-a* Δ*ucl3-d* and Δ*ucl3-a* Δ*ucl2-a* were found to be about 92 to 95% of the size of the haploid B control.

Although all double mutants showed a significant decrease in cell size as compared to the haploid control, we could not observe any gene dosage dependent effect of the double deletions on cell size. Instead these decreases in cell size were similar to the decreases observed for the single deletion mutants.

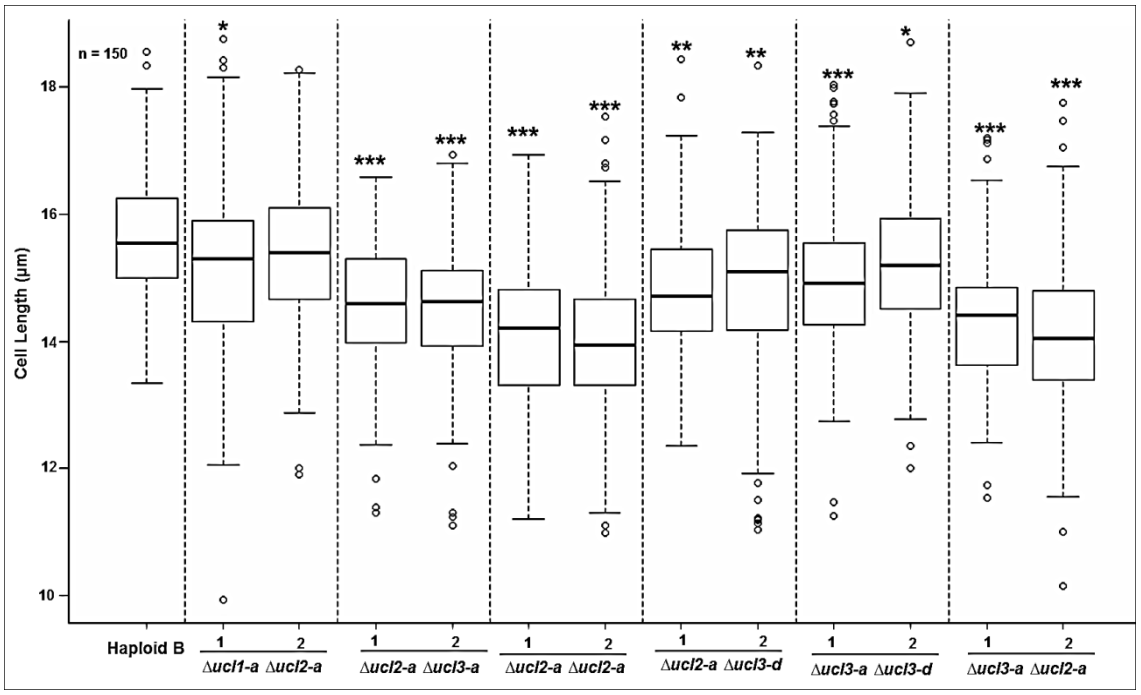


Figure 5.5: Box plots showing the cell size for haploid double mutants and control. Cell lengths were measured for 150 septated cells in each sample. The X-axis represents the different mutant strains along with the control (haploid B), where 1 and 2 correspond to the two independent biological repeats. The Y-axis represents the cell lengths on a linear scale. Horizontal line inside each box represents the mean of the cell lengths, whereas individual dots above or below the whiskers represent the outliers. ***, represent *p-value < 2.2e-12, ** p-value < 4.1e-07,* * p-value ≤ 0.002.

5.3.4 Effects of deletion mutants on ploidy of the cells

ucl1-a, ucl2-a, ucl3-a and *ucl1-d* were highly up-regulated in the polyploid cells. To see whether deletion of these genes itself can cause any effect on the ploidy of cells, we performed FACS analysis for the deletion mutants along with their biological replicates used for cell size measurements. Samples were prepared for the FACS analysis as mentioned in section 2.6.

The FACS profiles did not reveal any change in the ploidy state of the deletion mutants (figure 5.6). Haploid mutants harbouring either a single gene deletion (figure 5.6A) or double gene deletions (figure 5.6B) showed the same DNA content as haploid B. These results show that deletion of these genes does not cause any change in the ploidy state of the deletion mutants.

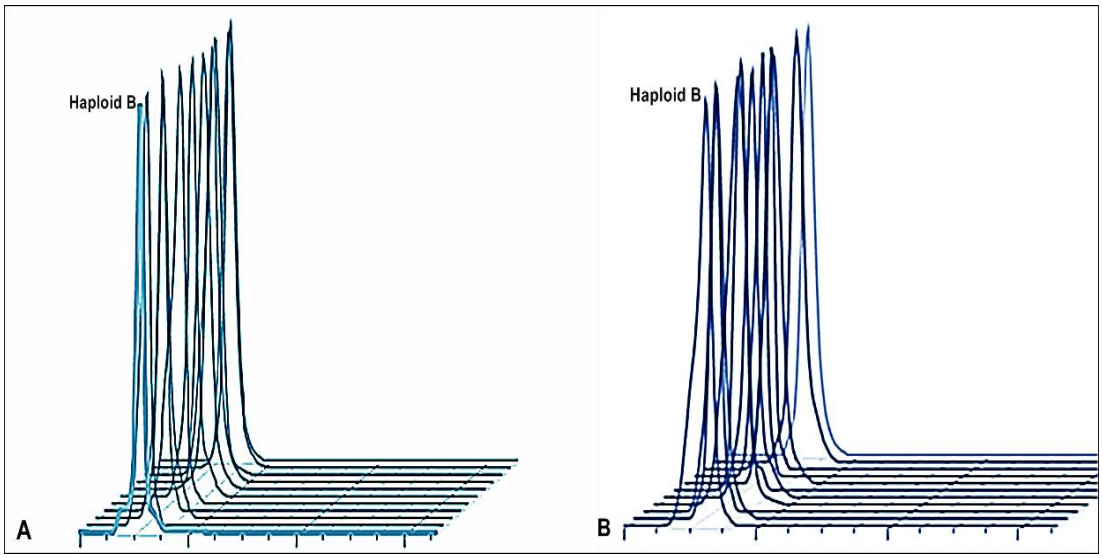


Figure 5.6: FACS profiles for deletion mutants. Samples were prepared for FACS as mentioned in section 2.6. (A) represents the FACS profile for haploid single mutants, while (B) represents the FACS profiles for haploid double mutants.

5.4 Effects of overexpression mutants on cell size

5.4.1 Overexpression in haploids

If *ucl1-a, ucl2-a* and *ucl3-a* have any effect on cell size then their overexpression might result in an increase in the cell size. To assess the effect of their overexpression on cell size, we constructed mutants where the ORFs of these genes were placed under the control of a regulatable promoter, P3nmt1, both in haploid B and diploid B cells. This was achieved by using the one-step PCR based approach (Bähler et al., 1998), as mentioned in section 2.4. Unlike the deletion mutants, construction of overexpression mutants for these genes proved to be difficult. We were able to construct the overexpression mutants *P3nmt1- ucl2-a* and *P3nmt1- ucl3-a* in both the haploid and diploid background, but *P3nmt1-ucl1-a* only in a haploid background. These overexpression mutants were grown in EMM containing adenine (225mg/ L) and thiamine (15 μM) for at least 12 hours, and were then transferred to EMM plus adenine, after washing them three times with EMM to remove the thiamine from the cell cultures. Cells were then collected at different time points and were processed for cell length measurements exactly as mentioned in section 2.5.

In the first repeat, cell lengths were measured for at least 350 cells regardless of whether they are septated or non septated to observe the effect of overexpression of *ucl1-a, ucl2-a* and *ucl3-a* on overall cell population rather only on septated cells. Samples were taken soon after removing the thiamine from the media (T0), after growing them for 17 hours (T1), 22 hours (T2) and 25 hours (T3) in EMM + adenine. When plotted, a slight but significant increase in cell size was observed for *P3nmt1- ucl2-a* and *P3nmt1- ucl3-a* containing haploids, while no change in cell size was observed for *P3nmt1- ucl1-a* containing haploids (figure 5.7 A). For mutants overexpressing *P3nmt1- ucl2-a,* contamination occurred after collecting cells at T2; therefore, data for T3 are not shown for this mutant. For haploid B (control) data is shown here only for T1.

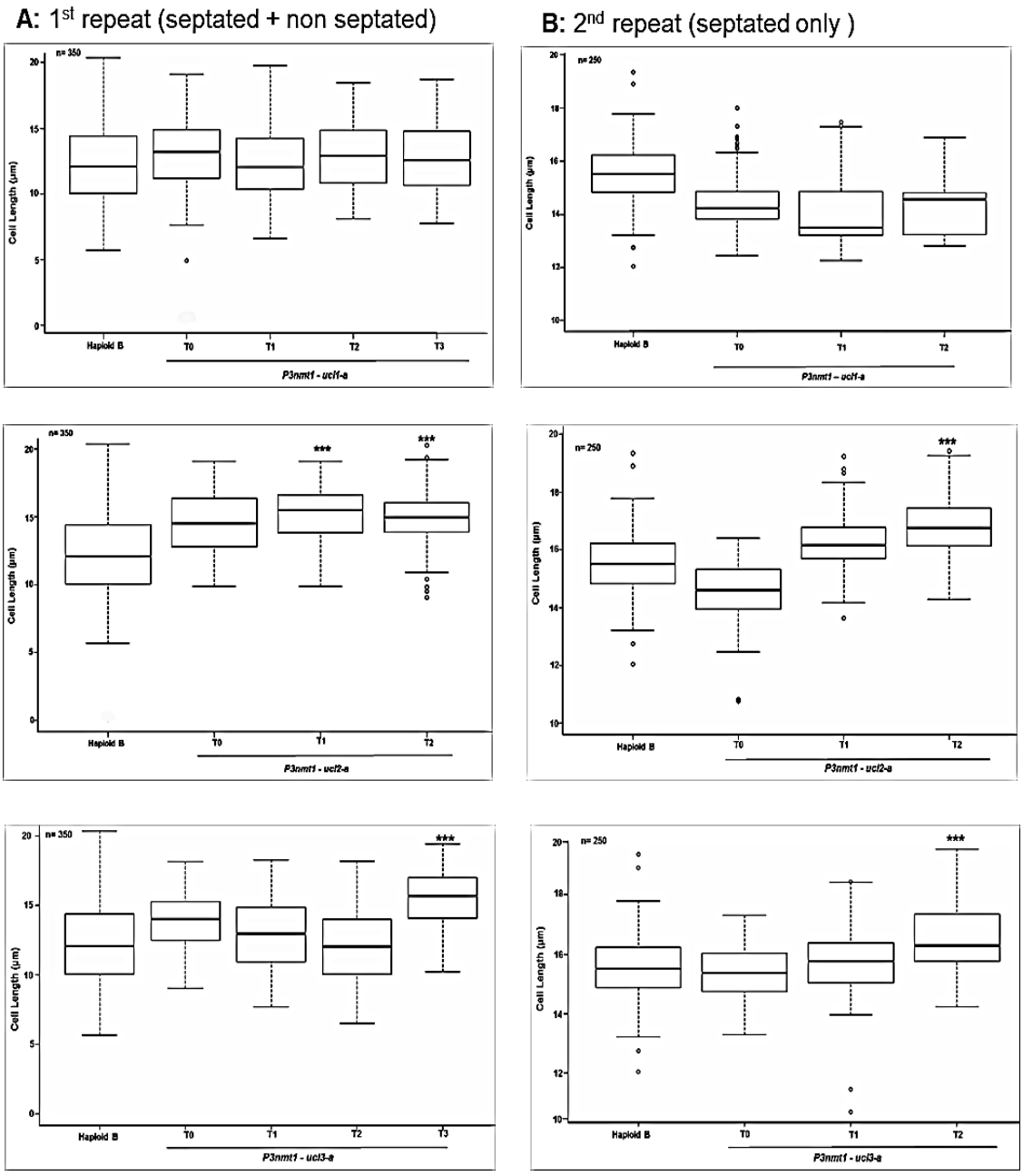


Figure 5.7: Box plots showing the cell sizes for haploid overexpression mutants. (A) Cell lengths were measured for 350 cells (septated and non septated) in each sample. The X-axis represents the different mutant strains along with the control (haploid B), where T0 (soon after removing the thiamine), T1 (17 hrs), T2 (22 hrs) and T3 (25 hrs) correspond to the time points at which samples were taken. The Y-axis represents the cell lengths on a linear scale. **(B)** 2nd repeat of the same experiment, where cell lengths were measured only for 250 septated cells. Here T1 = 25 hrs and T2 = 35 hrs. ***, represent *p-value < 2.2e-11.*

Mean cell length of *P3nmt1- ucl2-a* haploid mutants was found to be 14.67 μm at T0, which was ~20% larger than that of haploid B (12.23 μm). An increase of 3.7% and 2% was observed at T1 and T2, respectively, when compared to T0, while an increase of 25.4% and 22.6%, when compared to the haploid B. The decrease in cell length at T2 as compared to T1 could be due to the presence of a larger number of non septated cells at this time. Similarly, for haploid mutants containing *P3nmt1- ucl3-a*, an increase of 12% in cell length was observed as compared to T0, which is 26.6% larger than haploid B.

In the second repeat, samples were grown in EMM + adenine for about 35 hours, to see whether overexpressing these genes for a longer time results in cells of even bigger sizes. This time, cell lengths for about 250 septated cells were measured. When plotted, we observed a similar trend in cell size increase for *P3nmt1- ucl2-a* and *P3nmt1- ucl3-a* containing haploids, as earlier, while no significant change size occurred for *P3nmt1- ucl1-a* containing haploids (figure 5.7B). Haploid mutants containing overexpressing *ucl2-a* showed an increase of 16% (16.83 μm) in cell length at T2 as compared to the cell length at T0 (14.48 μm), and about 9% increase in cell length at T2 as compared to the cell length of haploid B (15.52 μm). Similarly for *P3nmt1- ucl3-a*, an increase of 8.6% (16.52 μm) in cell length was observed at T2 as compared to the cell length at T0 (15.2 μm), while an increase of 6.5% when compared to the cell length of haploid B. Although this increase in cell length of haploid mutants overexpressing *P3nmt1- ucl2-a* or *P3nmt1- ucl3-a* is slight, it is highly significant.

Overexpression of these genes was confirmed by using the Reverse transcriptase based PCR reaction (qRT-PCR). Samples were collected at all these time points and were processed for RNA extraction, subsequent cDNA synthesis and qRT-PCR were performed as mentioned in section 2.7.5 and 2.7.6. Analysis of the qRT-PCR assay showed overexpression of the *ucl1-a, ucl2-a* and *ucl3-a* in their corresponding mutants (figure 5.8). Interestingly, overexpression of *ucl1-a* was also observed in the mutants, despite the fact that they did not show any increase in cell size (figure 5.8 A).

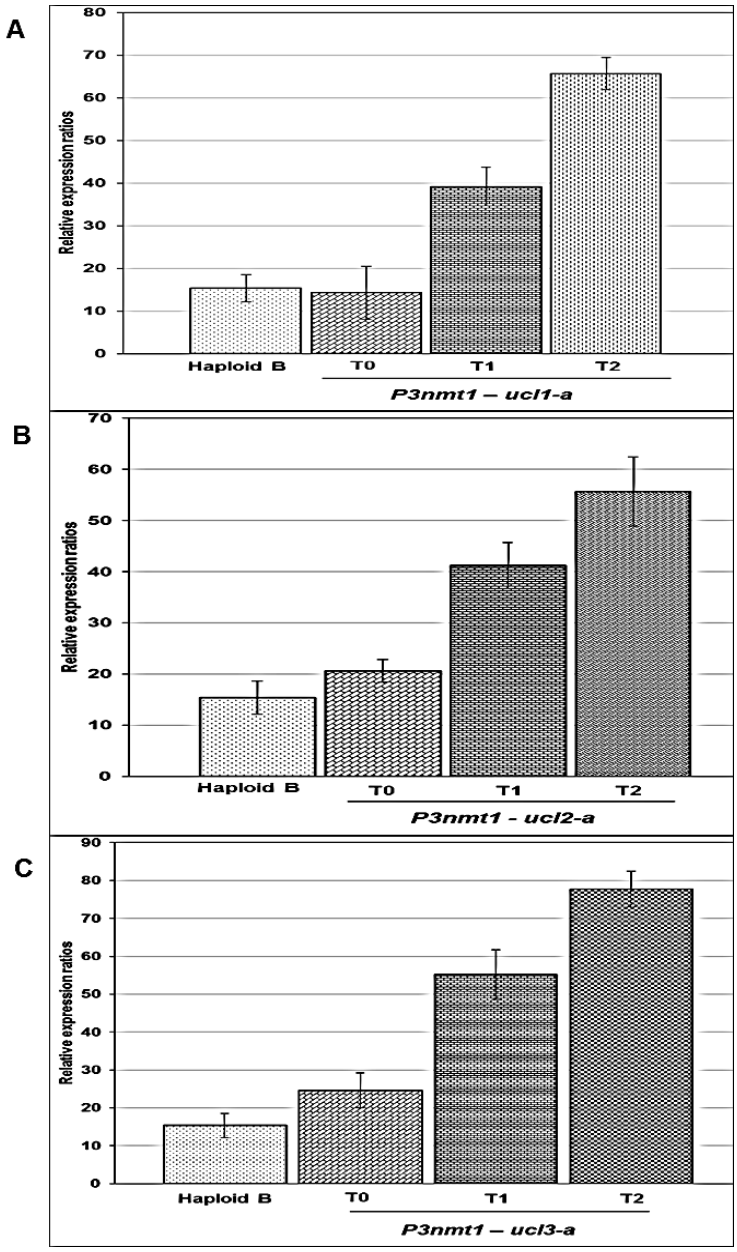


Figure 5.8: Bar plots showing the overexpression of *p3nmt1-ucl1-a* (A), *p3nmt1-ucl2-a* (B) and *p3nmt1-ucl3-a* (C). 20 ml of the cell cultures taken at each time points were processed for RNA extraction and subsequent qRT-PCR. Values obatined were normalised against the values obatined for *aap1* gene that was used as a control. Values represented here are the average of two independent biological repeats, whereas each repeat was processed as triplicate. The X-axis represents the different mutant strains along with the control (haploid B), where T0 (soon after removing thiamine), T1 (17 hrs), T2 (22 hrs) and T3 (25 hrs) correspond to the time points at which samples were taken for RNA processing, while the Y-axis represents relative expression ratios. Error bars represent the standard error.

5.4.2 Overexpression in diploids

Diploid mutants overexpressing either *ucl2-a* or *ucl3-a* were grown in a similar way as mentioned for haploid overexpression mutants. Cells were collected from growing cultures at T0, T1 (17 hrs), T2 (25 hrs) and T3 (35 hrs). Cell lengths were measured for at least 165 septated cells. For both diploid mutants, slight but significant increases in cell size were observed (figure 5.9A).

Cell length for diploids overexpressing *ucl2-a* was found to be 25.64 µm at T3, which is 8 % larger than the cell length at T0 (23.7 µm). When compared with the cell length of diploids overexpressing *ucl2-a* at T3 with that of diploid B (22.92 µm), an increase of ~12% in length was observed for diploid mutants. Similarly, for diploids overexpressing *ucl3-a*, an increase of 15% in cell length was observed at T3 (26.36 µm) as compared to the cell length of diploid B (22.92 µm), but only 6.5% larger than the cell length at T0 (24.76 µm).

To confirm the overexpression of these genes in the diploid mutants, we performed qRT-PCR analysis. Expression values of corresponding genes at each time point were normalised against the expression values obtained for *aap1*, which is used as a control. When plotted, we observed an increase in the expression levels of *ucl2-a* and *ucl3-a*, at time points T2 and T3 (figure 5.9B).

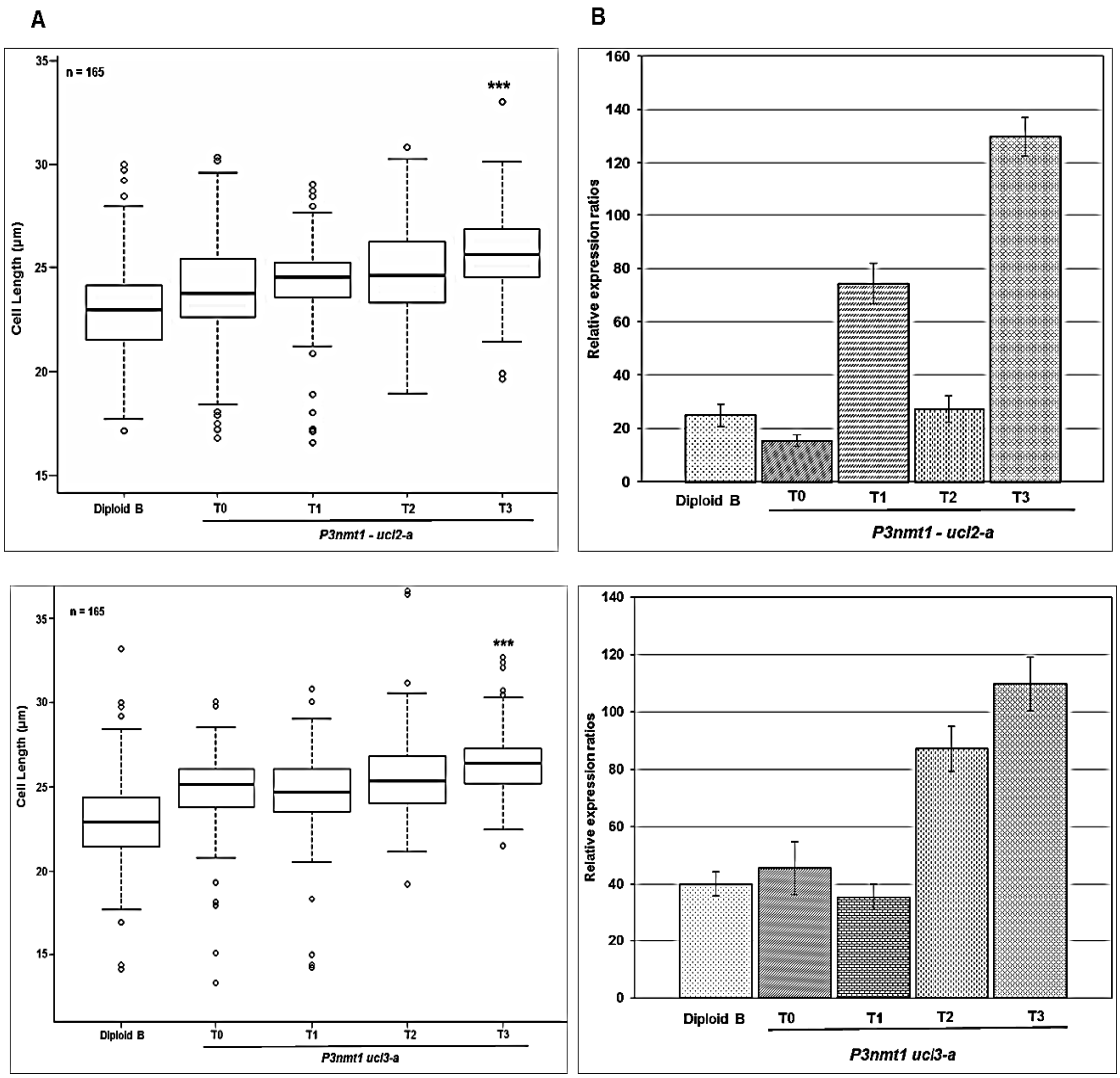


Figure 5.9: Box plots and bar plots showing the cell size and expression ratios in diploid overexpression mutants. (A) Cell lengths were measured for 165 cells septated cells in each sample. The X-axis represents the different mutant strains along with the control (Diploid B), where T0 (soon after removing the thiamine), T1 (17 hrs), T2 (25 hrs) and T3 (35 hrs) correspond to the time points at which samples were taken, while the Y-axis represents the cell lengths on a linear scale. ***, represent *p-value < 9e-11.* (B) qRT-PCR analysis showing the overexpression of *ucl2-a* and *ucl3-a* in these overexpression mutants. The X-axis is the same as for (A), while the Y-axis represents the expression ratios for corresponding time point. Error bars represent the standard error obtained for two independent biological repeats.

5.4.3 Effects of overexpression mutants on ploidy

To test whether the overexpression of Δ*ucl1-a*, Δ*ucl2-a* and Δ*ucl3-a* has any effect on the ploidy state of the mutants, we performed FACS analysis for all the haploid and diploid overexpression mutants at each time point. Samples were prepared for the FACS analysis as mentioned earlier. When analysed, the data obtained in an overlay, no change was observed in the DNA content of the haploid and diploid mutants overexpressing these genes (figure 5.10). The FACS analysis indicates that these genes do not affect the ploidy of the cells.

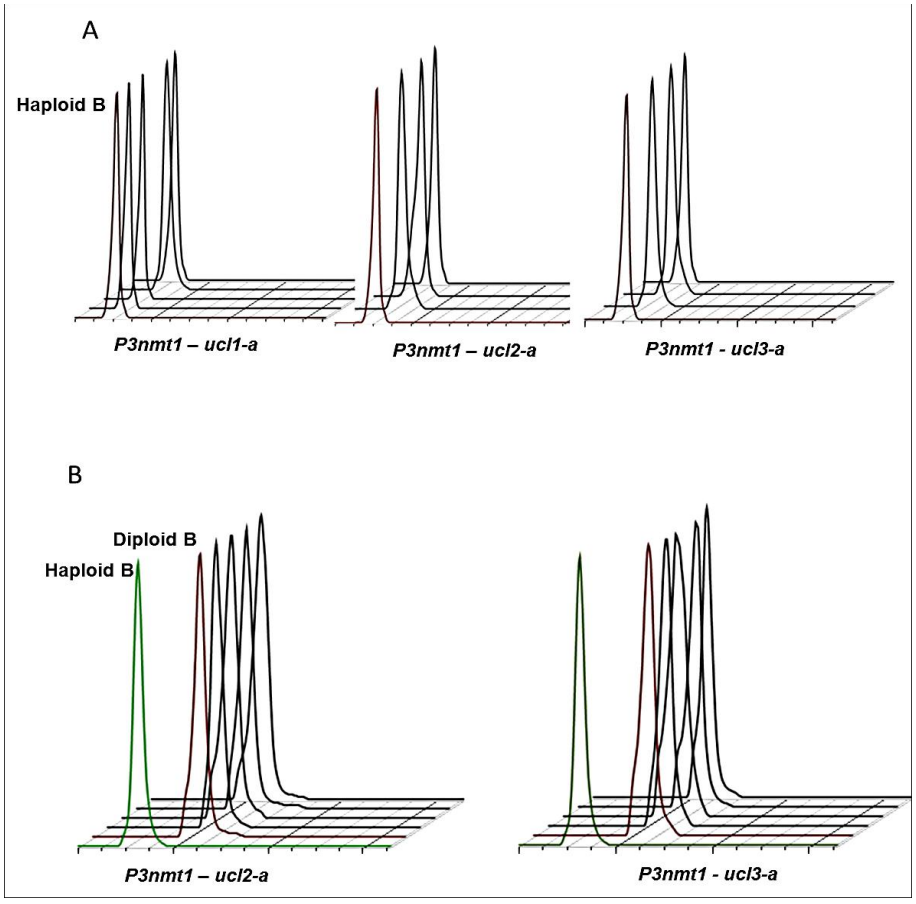


Figure 5.10: FACS analysis for haploid (A) and diploid (B) overexpression mutants. Samples for FACS analysis were collected at the time points mentioned in figure 5.9A from each of the mutant mentioned in (A) and (B) here. For diploid overexpression mutants (B) haploid B was used as a control along with the diploid B.

5.5 Cellular response of deletion mutants to various stresses

To address the role of Δ*ucl1-a,* Δ*ucl2-a,* Δ*ucl3-a* and Δ*ucl3-d* in the cellular response to various stressors, we examined their growth in a high throughput quantitative growth assay, by using the Singer RoToR robot. We assessed the growth of haploid as well as diploid deletion mutants (two independent biological repeats of each) along with different controls, such as wild type *972h*$^{-}$, Δ*pka,* Δ*git3,* Δ*gpa2,* and Δ*rad3.* For the control mutants, stress related phenotypes have already been characterised for various stress causing agents such as Δ*pka1,* which is sensitive to all salts, DNA damaging agents and oxidative stresses (Oowatari et al., 2009; Gupta et al., 2011); Δ*git3* and Δ*gpa2,* which are sensitive to DNA damaging agents and camptothecin (Han et al., 2010); and Δ*rad3,* which is sensitive to UV and hydroxyurea (Shikata et al., 2007). Cellular growth was assessed in the presence of various stressors that can be categorised as follows: Oxidative stressors - 1mM and 2mM H_2O_2; Osmotic stressors – KCl (0.6M), $MgCl_2$ (0.1 and 0.2 M) and $CaCl_2$ (150 mM); cell wall damaging agents – calcofluor (1.2 and 1.4 mg/ml) and SDS (0.005, 0.01 and 0.1 mM); and protein synthesis inhibition – chyclohexamide (10 μg/ml). All these strains were grown in YES to mid log phase in 96 well plates and were replica plated onto the YES agar plates containing different stressors at 96–384–1536 format with the help of the Singer RoTor robot. This format allows all strains to be present in the form of quadruplicates. After incubating them at 32°C for 3 days, all plates were photographed. Images were processed with workspace software that generates the numeric values for the images of the colonies. This data was further analysed by Dr. Martin Převorovský, who used different statistical tests (to take the averages of the quadruplicates and then to normalise the resultant values against wild type control).

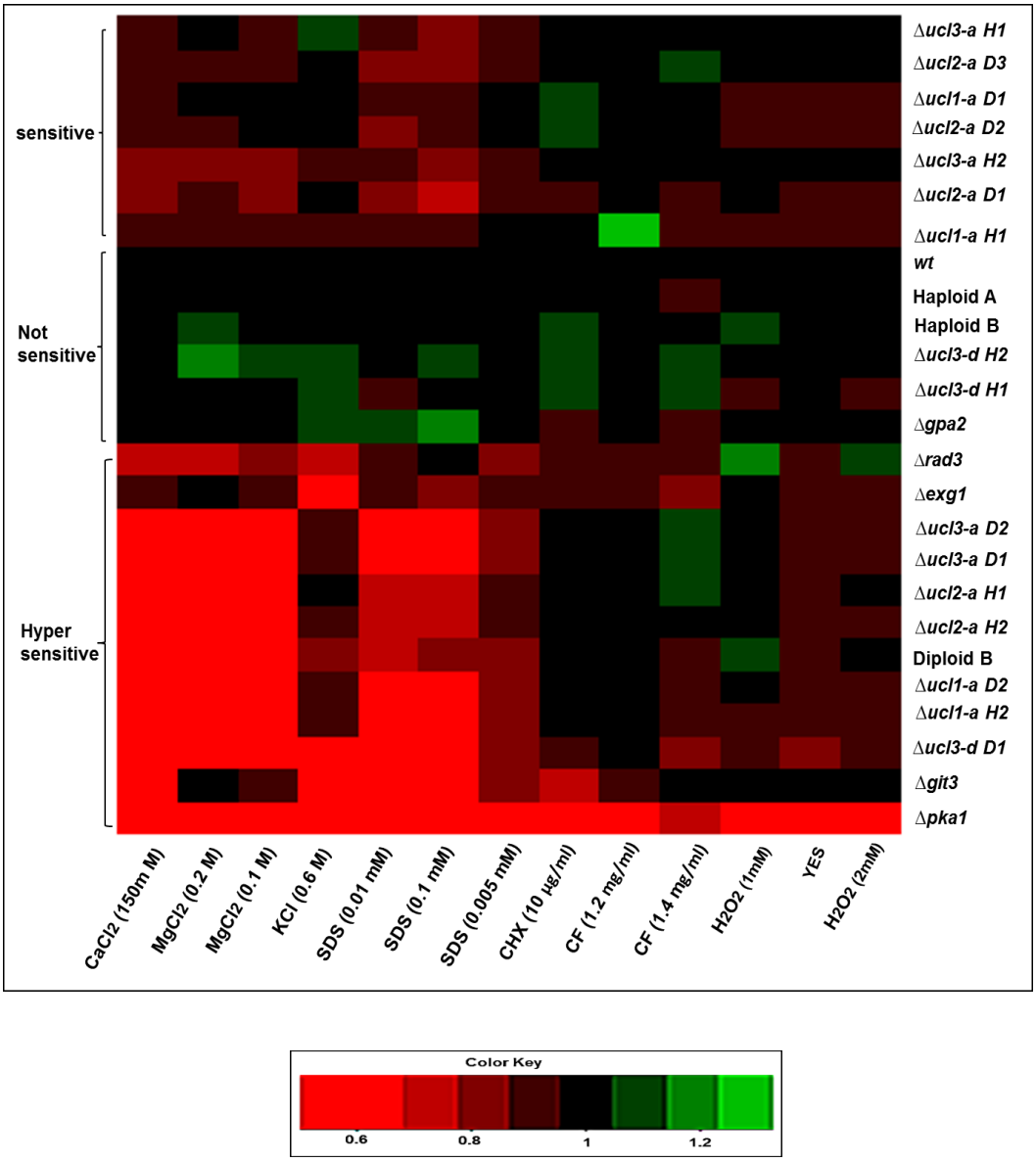


Figure 5.11: Phenotypic clustering of mutants on the basis of their sensitivity to various stresses. Deletion mutants along with the controls were grown to OD_{595} 0.5 in YES and were arrayed on YES+ agar in 1536 format. Images were taken by using BioDoc-It® Imaging System (UVP) and were processed with workspace software to generate numeric values. These values were further processed by applying different statistical tests. The X-axis represents the different stressors used, where H_2O_2 = hydrogen peroxide, CF = calcofluor, CHX =chyclohexamide, SDS = sodium dodecyle sulphate, KCL = Potassium chloride, $MgCl_2$ = Magnesium chloride and $CaCl_2$ = Calcium chloride. The Y- axis represents the different strains used, where H and D denote haploid and diploid strains, respectively, while numbers after H or D represent the biological repeats 1 and 2. The Colour key represents the strength of sensitivity.

On the basis of the resultant growth, strains can be divided into three categories: hyper sensitive, sensitive and not sensitive (figure 5.11). Strains categorised as **not sensitive** include the *wt*, haploid A and haploid B (used as controls), which confirms the reliability of this essay. The stressor's concentrations were chosen to be moderate such that they should not affect the wild type control. Haploid A showed a slight sensitivity in growth in the presence of CF (1.4 mg/ml), whereas, haploid B showed slightly higher growth in the presence of H_2O_2 (1mM) and $MgCl_2$ (0.2 M). Interestingly, the haploid mutants for *ucl3-d* also clustered in this group, which means that deletion of this gene does not affect the cellular growth in the presence of stressors; instead these cells formed bigger colonies in the presence of CF (1.4 mg/ml), CHX and KCL. Δ*gpa2* that was previously shown to be sensitive for DNA damaging agents and camptothecin (Han et al., 2010) also falls in this group, which shows that this mutant's growth is not affected by these stressors.

Sensitive strains include both biological replicates of Δ*ucl3-a H,* Δ*ucl2-a D,* one biological repeat of Δ*ucl1-a* H and D, along with the diploid B. All of these strains showed sensitivity in growth in the presence of SDS, $MgCl_2$ and $CaCl_2$. In addition, Δ*ucl2-a D* also showed sensitivity to the presence of H_2O_2. All other mutants fall into the category of **highly sensitive.** Interestingly, Δ*pka1* proved to be highly sensitive to all the stresses, which is in complete agreement with its previously reported phenotype in response to stress (Oowatari et al., 2009; Gupta et al., 2011). Mutants included in this category include: Δ*ucl2-a H,* Δ*ucl3-a D;* 1 repeat for each of Δ*ucl3-d D;* Δ*ucl1-a H* and *D.* Except for CF (1.2mg/ml), CHX, H_2O_2 (1mM), these mutants were highly sensitive to all other stresses. These mutants also showed smaller colonies on YES plates, revealing that they might be slow growing. These findings suggest that deletion of these genes confers sensitive phenotypes to these mutants in the presence of different stressors

5.6 Conclusion

In this chapter we tried to explore the role of genes that were found to be highly expressed in polyploids as well as in haploid mutants of larger cell size, which indicated that their overexpression might be caused by the larger cell size rather than by an increase in the ploidy. We selected genes for the follow up study that encode or are predicted to encode plasma membrane proteins, due to the presence of transmembrane domains (table 5.1). By using the BLASTN tool, we found that these genes contain orthologs in the genome, where many of them show more than 98 % sequence similarity among each other.

We named these genes and their paralogs Up regulated as a function of Cell Length (ucl). For this study, we deleted and overexpressed *ucl1-a, ucl2-a, ucl3-a* and *ucl3-d* genes in a haploid strain, as well as in diploid mutants. Unlike the deletion mutants, the overexpression mutants posed difficulties in construction. Several rounds of transformations were performed to obtain them, but even then we could not succeed in getting the overexpression mutants for *ucl1-a*, which indicates that overexpression of this gene might have some deleterious effect.

When analysed for their effect on cell size, we observed a slight but significant decrease in size of Δ*ucl2-a,* Δ*ucl3-a* and Δ*ucl3-d.* Accordingly, overexpression of these genes in haploids and diploids showed an increase in cell size. However for Δ*ucl1-a*, we could not observe any such effect on cell size. The FACS analysis showed that these genes do not have any effect on the ploidy state of the cells.

Cellular growth response to various stressors has revealed that many of these mutants are highly sensitive to several stresses, specifically in response to cell wall damaging agents and salt stresses. These findings suggest that these mutants have altered cell wall or membrane composition, consistent with their predicted functions at the cell periphery.

Although these genes have other paralogs in the genome that could mask their phenotypic effects by redundancy of their protein products, their effect on cell size and high sensitivity to various stresses indicate their important functions in the cell.

Chapter 6

6 Discussion

Ploidy – the number of the homologous chromosomes present within the nucleus of a cell, is a fundamental but poorly characterised genetic trait of eukaryotic organisms. Ploidy of an organism can vary as a function of different processes such as evolution, development and disease. All of these processes can result in several rounds of duplication of a genome, thus giving rise to the polyploid repertoire of the cells that has the ability to develop novel roles in a coordinated manner on a whole genome level. Cells with higher ploidies exhibit chromosomal instability and result in aneuploid cells that are a characteristic feature of many cancers. While a plethora of aneuploid and polyploid cell types are now known, how change in ploidy state alters gene expression and cellular physiology remains poorly understood. The availability of whole genome sequence of *S. pombe* and the microarray technology, which allows measuring the transcript levels across the entire genome, has made it possible to identify changes in transcription profiles of a large number of genes. In this study, I have used partial aneuploids and polyploid strains of fission yeast to demonstrate the effect of changing ploidy on regulation of gene expression. Here I discuss the main findings observed in the previous chapters and their implications in a wider context.

6.1 Aneuploidy accompanies large scale changes in gene expression

We used two partial aneuploid strains carrying an extra copy of a smaller portion of chromosome III, named aneuploid 1 and aneuploid 2, to understand the role of aneuploidy in regulating fission yeast gene expression. By using the CGH technique, we not only identified the location of Ch16 and IsoCh16 on chromosome III, but also the genes contained in them (Appendix 1). The right and left end of Ch16 were mapped to SPCC11E10.02c, located on the right arm of chromosome III (from position 1454816 bp to 1455958 bp) and to SPNCRNA.06 on the left arm of chromosome III (from position 968168 bp to 968477 bp), respectively. Mapping of the right arm agrees with a previous study (Chikashige et al., 2007), but the left end of Ch16 in that study was mapped to the region between SPCPB16A4.06c and SPCC1742.01. This difference in mapping at the left end could be attributed to the better resolution and presence of non-coding RNA probes on the Agilent microarrays

we used here. Similarly, for IsoCh16, both the right and left ends were mapped to SPNCRNA.06 (968168 bp to 968477 bp) on the left arm of chromosome III, which reflects the duplication of the left arm of Ch16 (Tinline-Purvis et al., 2009). Ch16 was found to contain 167 genes, while IsoCh16 contained 40 genes (Appendix IV).

Gene expression profiles of aneuploid 1 and aneuploid 2 indicate that the presence of an extra copy of a few genes results in an increase in the corresponding transcript levels, which scales with the DNA copy number of those particular genes. In addition, a relative change in copy number of a few genes is also accompanied by altered expression levels of some other genes present in the monosomic regions. A total of 92 and 51 genes present in the monosomic regions showed an increase of ≥2 fold in expression ratios in aneuploid 1 and aneuploid 2, respectively (Appendix V, table 1A). Overall, 95% and 82% of these genes were present on chromosomes I and II, respectively, in both aneuploids. Of the genes overexpressed in monosomic region of aneuploid 1, 32% have been characterised previously as a stress regulated genes (Chen et al., 2003). Other genes that are up-regulated in monosomic regions of both aneuploids are enriched for metabolic genes, especially the ones that are involved in the oxidation-reduction processes in the cell (Appendix V). In addition to the up-regulated genes, 14 genes in aneuploid 1 and 12 genes in aneuploid 2 exhibited ~2-fold repression in mRNA levels. Many of these differentially regulated genes (more than 70%) were located in the vicinity of telomeres of chromosomes I and II in both aneuploids, which is in agreement with the previous study (Chikashige et al., 2007). Interestingly, an enrichment of differentially regulated genes near telomeres has been observed only in fission yeast aneuploid strains. Other studies analysing the role of aneuploidy in gene expression using budding yeast, plants, mouse embryonic fibroblasts or human cell lines have not reported any such enrichment of genes in telomeric regions (Kahlem et al., 2004; Lyle et al., 2004; Torres et al., 2007, 2008; Huettel et al., 2008; Pavelka et al., 2010; Sheltzer et al., 2012; Stingele et al., 2012). This altered expression of telomeric genes was found to be due to the decreased binding of Swi6 to these regions (Chikashige et al., 2007).

Differential regulation of genes present in the monosomic regions of both aneuploids could be a result of regulatory genes present in disomic and trisomic regions, which in turn leads to changes in controlling target genes in monosomic regions. We found

a total of 5 such potential regulatory genes that are either transcription factors or part of regulatory networks that can drive the expression of other genes. Interestingly, we found *srk1*on Ch16, which encodes a MAPK activated protein kinase, involved in the regulation of mitosis and stress signalling pathway (López-Avilés et al., 2005). The presence of this gene in the disomic region might be involved in the overexpression of stress-regulated genes in aneuploid 1.

In comparison with the previous study (Chikashige et al., 2007), we found a relatively larger set of the up-regulated genes (92 in this study, compared to 79 in the previous study) in the monosomic region of aneuploid 1. Only 29 up-regulated and 2 down-regulated genes were found to be identical between the two studies (table 3.2), while all other genes were specific to either study. Such a small overlap between the two studies suggests that up-regulation of many of the genes contained in the monosomic region might be caused by the differences in the strain background (both studies have used the Ch16 containing aneuploids with different markers for selection), technical issues such as environmental perturbations and/or the different microarray platforms used. Both of these studies, however, have found an enrichment of stress responsive genes. Stress genes are also found to be a signature of aneuploidy in budding yeast and *A. thaliana* (Torres et al., 2007; Huettel et al., 2008; Pavelka et al., 2010).

The presence of an extra chromosome has been shown to be associated with severe growth defects, particularly in the G1 and S phase, in a variety of organisms ranging from yeast to human (Niwa et al., 2006; Torres et al., 2007; Williams et al., 2008; Pavelka et al., 2010). Overexpression of genes contained in the extra chromosome as well as the differential expression of other genes in the monosomic regions in response to the aneuploidy might be the cause of these defects.

Taken together, the findings in this study demonstrate that the presence of an extra chromosome itself is an important factor in causing large-scale imbalances in gene expression levels, which could lead to altered cellular fitness.

6.2 Polyploidy causes differential regulation of few genes

In this study, the effect of polyploidy for the regulation of fission yeast gene expression was also explored. Polyploidy in fission yeast is not well tolerated. Fission yeast is predominantly a haploid organism, but diploid and polyploids can be generated by nutrient starvation and by protoplast fusion. For the selection of polyploids, we needed to use the auxotrophic marker system that could allow us not only to easily select the cells of higher ploidies but also has minimum effects on cellular physiology. Strains auxotrophic for *ura4* or *leu* markers show somewhat compromised growth even in the presence of the proper supplement. For example, *leu-* cells are slow to exit from starvation, whereas, *ura-* cells exhibit some cell wall defects (Nurse Lab manual, http://www-bcf.usc.edu/~forsburg/plasmids.html). We therefore used *ade6* mutants for the generation of polyploid strains, as they have not been shown to confer any physiological phenotype, and also they allow the easy selection of diploid and tetraploid cells. Diploids and tetraploid strains generated by the method of nutrient starvation in homothallic strains or in strains of opposite mating types were too unstable to be used for microarray based gene expression analysis. Thus, we chose *ade6* h^- mutants to construct stable diploid and tetraploid strains that were homozygous at the mating type locus (h^-/h^-). Diploid strains used in this study were stable, while the tetraploid strains accumulated cells of lower ploidies, from 2N to 4N, that were evident from FACS analysis and cell length measurements. The presence of cells of lower ploidies indicates that the cells of fission yeast cannot tolerate the presence of a large number of extra chromosomes; therefore they lose some chromosomes during cell division, probably through aberrant chromosomal segregation. Chromosomal instability in the tetraploid cells has been found to be commonly associated with cancer cells, where tetraploids lose chromosomes and give rise to the aneuploid cells (Lengauer et al., 1997).

Diploids showed normal morphology when compared to haploids, while few tetraploid cells appeared irregular in shape with multiple septa. Cell lengths of both diploids and the tetraploid showed a linear increase with ploidy, which agrees with the cell size observations in response to polyploidy made in other organisms (Cavalier-Smith, 2005; Turner et al., 2012).

6.3 Differential gene expression is result of altered cell size

Our results indicate an increase in transcription per cell in a ploidy-dependent manner. Notably, this increase in transcription per cell does not affect the ratio between total RNA and mRNA. We also identified transcripts (about 1.8%) whose relative level of expression is altered in polyploid fission yeast cells. Many of these transcripts showed enrichment for metabolic genes, specifically the ones involved in oxidation-reduction processes inside the cells. Overexpression of these genes might be a requirement for the larger cells to fulfil their energy needs. Interestingly, about one third of the differentially regulated transcripts encode, or are predicted to encode, the proteins that are mainly localized to the cell wall or plasma membrane. These proteins are involved either in cell wall/membrane assembly or transport of various metabolites across the cell surface. This enrichment of cell surface proteins points towards a role of these genes in larger cells, and cell size might be important for the differential regulation of these genes. Wu et al (2010) have also reported the role of cell size, rather than the ploidy, in regulation of gene expression. They used only tetraploid cells to identify the differentially regulated genes and have established the role of two MAPK pathways regulating these genes in response to the cell size (Wu et al., 2010).

We could not find any overlap between genes showing differential regulation in fission yeast polyploids with those previously described in budding yeast. However, in both yeasts cell surface proteins are differentially expressed as a function of cell size. This effect was also observed for the diploid and tetraploid fission yeast cells, where most differentially regulated cell surface annotated genes are different between two diploids as well as between diploid and tetraploid cells. Thus, only a small number of cell surface protein encoding genes show overlap among all polyploids. The intriguing finding of differentially expressed cell surface proteins could be explained by the altered cell surface-to-cell volume ratio that accompanies altered cell size, and the amount of cell surface proteins may therefore need adjustment to altered cell size.

This differential expression of genes encoding cell surface proteins can be attributed to either the larger cell size of polyploids or to the change in surface area to volume

ratio of these polyploid cells. Surface area to the volume ratio is usually lower for tetraploids compared to diploid and haploid cells. This difference could affect the interactions between cell surface components and their cytoplasmic counterparts. Decreased expression of cell surface transporters in polyploids could be explained by the reduced surface area of these cells. To test the role of cell size, rather than ploidy, in regulation of these genes, we observed their expression in haploid mutants of large and small sizes. Many of the differentially regulated genes in polyploids exhibited similar expression profiles in mutants of larger cell size, thus supporting a role of cell size, rather than ploidy, in regulation of these genes.

To analyse the role of differentially regulated genes on cell size, we selected only those genes that were highly up-regulated in response to polyploidy or cell size in diploids as well as in the tetraploid. Genes that are known to encode cell surface proteins but do not show overlap among all polyploids used in this study were not selected for the follow up study. We selected a list of genes (table 5.2) that showed consistently increased expression in cells of larger size, and that have paralogs in the genome and are not characterised yet. Among these genes, deletion or overexpression of *ucl2-a* and *ucl3-a* showed a slight but significant decrease or increase in cell lengths, respectively. This finding is intriguing as it suggests not only that *ucl2* and *ucl-3* are up-regulated with increasing cell size, but that they seem to be involved themselves in controlling cell size.

The presence of paralogs for these genes leads us to hypothesize that they might act in a dosage dependent manner. To confirm this, we also constructed double mutants for these genes and analysed their effect on the cell size. Although we observed a decrease in cell lengths, it was of similar magnitude as that of single deletion mutants. The presence of multiple copies in the genome could mask the effects conferred by the deletion or overexpression of one or two copies due to redundancy of protein function. However, we not only observed an effect on cell length, but also on growth phenotypes in response to various stressors, specifically in the presence of osmotic and cell wall stressors. Thus, the presence of paralogs is not sufficient for cellular function in the mutants.

This study has provided an insight into the role of ploidy in transcriptional regulation of fission yeast, where aneuploidy affects the transcript levels for a larger number of genes compared to polyploidy. This indicates that changes in gene dosage for a relatively few genes could be more detrimental than global changes in gene dosage, where cells adjust their overall transcriptional response to the altered gene numbers and cell size. This also holds true for cancer cells (largely aneuploids) and megakaryocytes (highly polyploids), where aneuploidy leads to the death of the cells while polyploidy results in differentiation of the platelets at the faster rate.

Although more work needs to be done to establish the role of differentially regulated genes in the regulation of cell size, our data suggest an association between cell size and gene regulation. This shows that cells can regulate their transcription in response to changes in cell size. This finding agrees with a previous study where mRNA synthesis was found to correlate with changes in cell size up to two-fold (Zhurinsky et al., 2010). For a given genome, cells can coordinate their transcriptome and proteome with changing cell size, hence cells can regulate the RNA to DNA and/or Protein to DNA ratios in response to change in cell size (Marguerat and Bähler, 2012). Global tuning of gene expression in response to the larger cell size could compensate for the physiological consequences of polyploidy during tumour progression, differentiation and growth. In eukaryotes, the threshold of cell size determines the growth rate and cell division (Cook and Tyers, 2007). The relationship between cell size and gene expression explains the necessity of maintaining the uniform cell size to maintain transcriptional homeostasis.

6.4 Future work

This work has highlighted the transcriptional response of aneuploidy in fission yeast. As overexpression of a large number of genes near telomeric regions has been observed only in fission yeast partial aneuploids strains it would be interesting to

- Analyse the transcriptional response of other aneuploids in fission yeast that are disomic for chromosomes I and II. This study can tell whether overexpression of telomeric genes is a signature response of fission yeast cells to aneuploidy, or whether it is due to the presence of some regulatory genes for Swi6 on Ch16 and IsoCh16.
- Analyse the phenotypic response of aneuploids and polyploids in response to various stressors/drugs. Our data has indicated an overexpression of stress responsive genes in both aneuploids and polyploids. It would be interesting to determine whether aneuploidy/polyploidy in this organism has conferred any fitness advantage or disadvantage as compared to the wild type haploid cells. The presence of even one or a few beneficial/harmful genes on an extra chromosome in aneuploids can result in a novel phenotype**,** which can help these organisms to adapt to adverse growth conditions.
- Analyse the regulation of the proteome in response to aneuploidy in fission yeast. In budding yeast, one study has shown the role of aneuploidy in protein aggregate formation due to a compromised protein quality control system (Oromendia et al., 2012). As fission and budding yeast are only distantly related, they represent a good complementary system for the identification of conserved mechanisms among eukaryotes. Therefore, it would be interesting to check whether a "poor protein quality control system" is conserved in eukaryotes in response to the aneuploidy, or whether it is only observed in budding yeast.
- Analyse the role of cell surface genes differentially expressed in some diploids and tetraploids but that do not show an overlap among all of them (table 4.2A to table 4.4B). These are clearly annotated as cell surface/plasma membrane protein encoding genes, with most of them being surface transporters. It would be interesting to delete these genes in haploid mutants of larger cell size i.e., *cdc25-22* and overexpress them in mutants of smaller

size such as *wee1-50*. This deletion and overexpression in haploid mutants of larger and smaller size could establish whether they can affect the cell size.

- Analyse the transcriptome of deletion/overexpression of *ucl1-a, ucl2-a* and *ucl3-a* in haploid cells. Even though these genes have paralogs in the genome, the cell size and stress responsive phenotypes indicate their important role in cellular physiology. By microarray based gene expression analysis, we might identify their regulators that contribute to controlling these genes as a function of cell size.
- Protein localisation of *ucl1-a, ucl2-a* and *ucl3-a* to determine their location inside the cells. We have tagged these genes at both the C-terminus and N-terminus with GFP, but could not proceed with the experiments to localize them.

Appendices

Appendix I: Strains used in this study

Strains	**Source**
972 h-	Lab collection
ade6 M210 h-	Lab collection
ade6 M216 h-	Lab collection
ade6 M210 h+	Lab collection
ade6::Kan Mx6 h90	Lab collection
ade6 M210h-/ade6 M210 h-	This study
ade6 M216 h-/ade6 M216 h-	This study
ade6 M210 h+/ade6 M216 h-	This study
ade6::KanMx6 h90/ade6::Nat Mx6 h90	Lab collection
ade6 M210 h- ade6 M210 h-/ade6 M216 h- ade6 M216 h-	This study
Δ*ucl1-a ::KanMx6 ade6 M216 h-*	This study
Δ*ucl2-a ::KanMx6 ade6 M216 h-*	This study
Δ*ucl3-a::KanMx6 ade6 M216 h-*	This study
Δ*ucl3-d::KanMx6 ade6 M216 h-*	This study
Δ*ucl1-a::KanMx6* Δ*ucl2-a::Nat Mx6ade6 M216 h-*	This study
Δ*ucl2-a ::KanMx6* Δ*ucl3-a::Nat Mx6ade6 M216 h-*	This study
Δ*ucl2-a ::KanMx6* Δ*ucl2-a::Nat Mx6ade6 M216 h-*	This study
Δ*ucl2-a ::KanMx6*Δ*ucl3-d::Nat Mx6ade6 M216 h-*	This study
Δ*ucl3-a ::KanMx6*Δ*ucl3-d::Nat Mx6ade6 M216 h-*	This study
Δ*ucl3-a ::KanMx6* Δ*ucl2-a::Nat Mx6ade6 M216 h-*	This study
kanMX6::P3nmt1- ucl2-a ade6 M216 h-	This study
kanMX6::P3nmt1- ucl3-aade6 M216 h-	This study
kanMX6::P3nmt1-ucl1-a ade6 M216 h-	This study
Δ*rad3::KanMX4 h-*	Lab collection
Δ*gpa2::KanMX4 h-*	Lab collection
Δ*git3::KanMX4 h-*	Lab collection
Δ*pka::ura4 ura4-D h-*	Lab collection
Chr16 yps1::arg3, ert1:: Mata G418,ade6-M216, cid2::his3, Chr3 ade6-M210. Leu1-32, ura4-D18, his3-D1, arg3-D4 h^+	Provided by Tim Humphrey (Tinline-Purvis et al., 2009)
Leu1-32, ura4-D18, Chr16 h^+	Provided by Tim Humphrey

Appendix II: Primers used in this study

Primer Name	Primer sequence sequence (5' to 3')
Primers for qRT-PCR	
ucl1-a_ PF1	ACCAATATGGACGGGCATCC
ucl1-a_ PR1	TTCCGACAATCACCGCTACC
ucl1-a_ PF2	TCGTTACTGCTCTGTTGGCA
ucl1-a_ PR2	AGCTGCTTCTATTGCTGGCT
ucl2-a_ PF1	CTCGGAATCGGTATGGACGG
ucl2-a_ PR1	CCGAAACCCAATGCACAGTA
ucl2-a_ PF2	CGTAGCTGCGAGACCAATGA
ucl2-a_ PR2	ACACCCAAATTGCAGGAGAT
ucl3-a_ PF1	TCGACAATCCCGACAATCCC
ucl3-a_ PR1	CAGACGTCTCCTGGTGTCAC
ucl3-a_ PF2	ATTTTCGTTGGCCCTTGCTT
ucl3-a_PR2	TATAGTTGCACCCAGGGCTG
ucl3-d_ PF1	ATTTTCGTTGGCCCTTGCTT
ucl3-d_ PR1	GACTGCACCCAAGACTGCTA
ucl3-d_ PF2	CGATCTCGACAATCCCGACA
ucl3-d_ PR2	GCACCCAGGGCTGCTATAAT
Primers for Nat Cassette	
PLa NAT	CCTCGACATCATCTGCCCA
PLb NAT	CTATTTTAATCAAATGTTAG
PRa NAT	CTCGGCGTCCCCCGGGAC
PRb NAT	GTGGCGGTGACGCGGAAG
Primers for Kan Cassette	
KAN_PL	TTATGCCTCTTCCGACCATC
KAN_PR	ATTCCGACTCGTCCAACATC
KAN_PL2	TGGCCTGTTGAACAAGTCTG
KAN_PR2	CGACAGCAGTATAGCGACCA
KAN_PL3	GCCCGTACATTTAGCCCATA
KAN_PR3	GATGGTCGGAAGAGGCATAA

Primers used in this study

Primer Name	Primer sequence sequence (5' to 3')
Deletion checking primers	
ucl1-a_PF_delch	TTTAAGCTTCGGATTGGATCAT
ucl1-a_PR_delch	AACGAGTAAATGGCATTGCTTT
ucl2-a_PF_delch2	AATGATTATAGCAGCCTTGGGA
ucl2-a_PF_delch	CCTAAAGAAGCCAAAAAGCTCA
ucl2-a_PR_delch	AAATGGGAAAGCAAGTTGAAAA
ucl3-a_PF_delch	AAGGATTACGAGATGCCAACAT
ucl2-a_PR_delch	GACGCTGAATTTTGGGTTAAAG
C- terminus tagging checking primers	
ucl1-a_PF_tagch	GGGATCGTTCTCAAATTGCTAC
ucl2-a_PF_tagch	GTATCTCCTGCAATTTGGGTGT
ucl3-a_PF_tagch	TGGGAATGATTATAGCAGCCTT
ucl3-a_PR_tagch	TAATATTTTCGCCCGTCCATAC
N- terminus tagging checking primers	
ucl1-a_PR_tagch	TATTGGTATGCCAACAGAGCAG
ucl2-a_PR_tagch	TAATATTTTCGCCCGTCCATAC
ucl3-a_PR_tagch	AACGAAAATTACTGCGGGTAGA
ucl3-d_PR_tagch	AACGAAAATTACTGCGGGTAGA

Primers used in this study

Primer Name	Primer sequence sequence (5' to 3')
Gene deletion primers	
ucl1-a_PF_del	TGAAGCTTCAAATAACTGTCGATTAGAGTTATTGCTTATTCATTGTTTTATCTTCCTTTACCCTCTAACGTAAATGAATCCGGATCCCCGGGTTAATTAA
ucl1-a_PR_del	ATAAAACGTAGCAACCGTTGAAAATTATTTCTTCACTTAGCTATAAAATATCAAAGTAATAAAAATTGTAACCATCATCAGAATTCGAGCTCGTTTAAAC
ucl3-a_PF_del	TAAAATATTTTAGTATATTTTCTTTTTTATTTTCCATAAACCACTCTTTTTAATTTTACTTTGGGGAAAAATACTAGTGACGGATCCCCGGGTTAATTAA
ucl3-a_PR_del	CATGATTTGGACAGGACAGTCAATTCGAACATTGTTAAGAATAGTTATATTCTATATGATCATGAAAACGATAAAGAAAAGAATTCGAGCTCGTTTAAAC
ucl2-a_PF_del	ATAAAAGAGAAATTGCTTTGTCTTTTTTTGAAGCAAATGAATGGCAAAATATAATTGGCTGGAATGATCAAGTTGAGGAGCGGATCCCCGGGTTAATTAA
ucl2-a_PR_del	TAATCCATTGATAAACAAATAATAAATAAAATAATCCCGGTAACCCCTTTAAAGGTGAATCAAGTAAACCATGGTTTTTGGAATTCGAGCTCGTTTAAAC
ucl3-d_PF_del	CCATTTTGGTTTTTTTTTTTTAATTTTCTCGTTTATGGCTTAAAATATTTTAGTATATTTTCTTTTTTATTTTCCATAAACGGATCCCCGGGTTAATTAA
ucl3-d_PR_del	ATTAAGTGATAATATCTTAATTAAAACCATAATTTGGTTTGGTTCCATAATTGAAATTCATAACGAACGCACCTTATGTGGAATTCGAGCTCGTTTAAAC
C- terminus tagging primers	
ucl1-a_PF_Ctag	AAATCGAAATGTATAAGGTACTGTTCTTTGGAGGGATTGTATTCATGATTTTTGGGATACTGTACATTTTTCAAGGCTTTCGGATCCCCGGGTTAATTAA
ucl1-a_PR_Ctag	ATAAAACGTAGCAACCGTTGAAAATTATTTCTTCACTTAGCTATAAAATATCAAAGTAATAAAAATTGTAACCATCATCAGAATTCGAGCTCGTTTAAAC
ucl2-a_PF_Ctag	GTTACATAATGTTTCCCACCCATAATGAAGCACTTGCCTACAATCGAAATGTAGCTGAAGAGGCTCAAGAGAAAATGATTCGGATCCCCGGGTTAATTAA
ucl2-a_PR_Ctag	TAATCCATTGATAAACAAATAATAAATAAAATAATCCCGGTAACCCCTTTAAAGGTGAATCAAGTAAACCATGGTTTTTGGAATTCGAGCTCGTTTAAAC
ucl3-a_PR_Ctag	CATGATTTGGACAGGACAGTCAATTCGAACATTGTTAAGAATAGTTATATTCTATATGATCATGAAAACGATAAAGAAAAGAATTCGAGCTCGTTTAAAC
ucl3-a_PF_Ctag	CAGCATTACATTCAGAAACAACAGTTGGTTCTGATATTGAACAAATAGAACTACAAAATATGCCTACTCCTGTGAAAAAACGGATCCCC

	GGGTTAATTAA
ucl3-a_PR_Ctag	GGTCTCGCAGCTACGCAGTAAGAACCCGTTTCATTAATCAAAATAGTTTTCCATATACTTGTAATATTTTCGCCCGTCCAGAATTCGAGCTCGTTTAAAC
Gene overexpression primers	
ucl1-a_PR_OE	CCAATGTTTTGAACGGAACCATGACTAAACTGCAAACTGCCTAAAGTAGTTTCCAAATTTAAGGGACTTTCAACATTAATCATGATTTAACAAAGCGACTATA
ucl2-a_PR_OE	CCCATCTCTCGGTTTTGGGAATTAAATTCCTCAAAAAAAGATTTTTGGATCACCGTGTCTCTTGATTTTACAAAGTCAATCATGATTTAACAAAGCGACTATA
ucl3-a_PR_OE	ACATTACAAACATCTTCCACCATAAATAACTCATTGTAACCAGGAGGTTCAACCTGTTTTTTCAAGCTTTCTGGATTTGACATGATTTAACAAAGCGACTATA
N- terminus GFP tagging primers	
ucl2-a_PF_Ntag+OE	CACTCTTAAGTGCAGTTGGTTCATTGTGGAGGAATCTAACGGTGCCAGAATACATGCTAATTTTTTTGCACAAAGCTTCGGAATTCGAGCTCGTTTAAAC
ucl2-a_PR_Ntag+OE	ATCTCTCGGTTTTGGGAATTAAATTCCTCAAAAAAAGATTTTTGGATCACCGTGTCTCTTGATTTTACAAAGTCAATCATTTTGTATAGTTCATCCATGC
ucl3-a_PR_Ntag+OE	CCACTCTTTTTAATTTTACTTTGGGGAAAAATACTAGTGAATCAGAAGTCATCAAATTACAGTGCTGTCGTTCTATTCCAGAATTCGAGCTCGTTTAAAC
ucl3-a_PR_Ntag+OE	TTACAAACATCTTCCACCATAAATAACTCATTGTAACCAGGAGGTTCAACCTGTTTTTTCAAGCTTTCTGGATTTGACATTTTGTATAGTTCATCCATGC

Appendix III: Concentration of spikes used in the Spikin mix

Spike No	Length (bp)	MW (kD)	RNA ρmol/mg total	RNA ng/mg total	10 reactions	conc. control (µg/µl)	Conc. (ng/µl)	dilution	add to 10x stock (µl)
1	1227	404.95	1110.83	449.83	4498.31	1.12	1120	1	4.02
2	1611	531.68	333.33	177.23	1772.26	1.33	1330	2.5	3.33
3	1182	390.10	111.08	43.33	433.32	1.13	1130	10	3.83
4	1449	478.22	33.33	15.94	159.39	1.13	1130	25	3.53
5	849	280.20	33.33	9.34	93.39	1.20	1200	50	3.89
6	795	262.38	7.39	1.94	19.39	1.22	1220	250	3.97
7	954	314.85	5.54	1.74	17.44	1.19	1190	250	3.66
8	1080	356.44	1.33	0.47	4.74	1.23	1230	1000	3.85
9	633	208.91	0.44	0.09	0.92	1.08	1080	5000	4.26
				10 µl for 1 mg total RNA	100 µl			Total volume RNA (µl)	34.35
								Sterile water (µl)	65.65
								Final volume	100.00

Appendix IV: Genes contained in Ch16 and IsoCh16 along with their CGH and expression ratios

(Genes contained in IsoCh16 are highlighted in blue color)

Systematic ID	Name	Aneuploid 1 CGH	Aneuploid 1 Expression	Aneuploid 2 CGH	Aneuploid 2 Expression	Description
SPCC1795.06	*map2*	2.51	1.80	4.11	3.31	P-factor pheromone Map2
SPCC825.03c	*psy1*	2.30	1.72	4.09	2.26	SNARE Psy1
SPCC1259.05c	*cox9*	2.28	1.83	3.92	3.36	cytochrome c oxidase subunit VIIa (predicted)
SPCC1259.07	*rxt3*	2.39	2.00	3.84	2.65	transcriptional regulatory protein Rxt3
SPCC1259.14c	*meu27*	2.30	1.88	3.81	2.71	*S. pombe* specific UPF0300 family protein 5
SPCC1795.02c	*vcx1*	2.31	2.06	3.67	2.83	CaCA proton/calcium exchanger (predicted)
SPCC1259.03	*rpa12*	2.18	1.57	3.65	3.25	DNA-directed RNA polymerase complex I subunit Rpa12
SPCC825.02		2.22	1.84	3.56	3.07	glucosidase II Gtb1 (predicted)
SPCC1795.03	*gms1*	2.21	1.44	3.55	2.67	UDP-galactose transporter Gms1
SPCC895.03c		2.18	1.72	3.48	2.84	SUA5/yciO/yrdC family protein Sua5 (predicted)
SPCC1259.10	*pgp1*	2.26	2.04	3.47	3.38	mitochondrial metallopeptidase involved in genome maintenance Pgp1
SPCC1795.05c		2.18	1.80	3.44	3.22	uridylate kinase (predicted)
SPCC1259.09c	*pdx1*	2.04	2.53	3.43	4.46	pyruvate dehydrogenase protein x component, Pdx1 (predicted)
SPCC895.07	*alp14*	2.03	1.85	3.43	2.89	TOG ortholog Alp14
SPCC895.05	*for3*	2.10	1.72	3.42	2.52	formin For3
SPCC1259.12c		2.35	1.96	3.42	2.95	Ran GTPase binding protein (predicted)
SPCC1259.11c	*gyp2*	2.36	1.64	3.40	2.86	GTPase activating protein Gyp2 (predicted)
SPCC1795.07		2.06	2.06	3.37	3.10	mitochondrial ribosomal protein subunit S37 (predicted)

SPCC1795.08c	*vid21*	2.16	1.90	3.37	3.04	NuA4 histone acetyltransferase complex subunit Vid21
SPCC895.04c	*ufe1*	2.00	1.83	3.34	3.16	SNARE Ufe1 (predicted)
SPCC1259.15c	*ubc11*	2.04	2.23	3.34	3.75	ubiquitin conjugating enzyme E2-C, Ubc11
SPCC1259.02c		2.06	2.07	3.33	3.85	Endoplasmic Reticulum metallopeptidase 1 (predicted)
SPCC1795.11	*sum3*	2.06	1.00	3.32	1.51	ATP-dependent RNA helicase Sum3
SPCC1259.08		2.04	1.71	3.24	3.38	conserved fungal protein, DUF2457 family
SPCC895.06	*elp2*	2.03	1.32	3.17	2.66	elongator complex subunit Elp2 (predicted)
SPCC1259.06	*taf8*	2.01	2.39	3.15	3.84	transcription factor TFIID complex subunit 8 (predicted)
SPCC1259.13	*chk1*	1.94	1.66	3.13	2.29	Chk1 protein kinase
SPCC895.09c	*ucp12*	1.90	1.58	3.13	2.91	ATP-dependent RNA helicase Ucp12 (predicted)
SPCC825.04c	*naa40*	2.01	1.70	3.13	2.75	histone N-acetyltransferase Naa40 (predicted)
SPCC1795.04c	*pre10*	1.83	2.14	3.11	3.20	20S proteasome component alpha 7, Pre10 (predicted)
SPCC1259.04	*iec3*	2.00	2.20	3.11	3.87	Ino80 complex subunit Iec3
SPCC825.01		2.00	1.54	3.09	3.10	ribosome biogenesis ATPase, Arb family ABCF1-like (predicted)
SPCC825.05c		1.73	1.57	3.06	2.36	splicing coactivator SRRM1 (predicted)
SPCC1795.12c		1.85	1.28	2.97	2.22	sequence orphan
SPCC895.08c		1.73	2.19	2.89	3.20	conserved fungal protein
SPCC1795.10c		1.75	1.91	2.87	3.09	Sed5 Vesicle Protein Svp26 (predicted)
SPCC1795.01c	*mad3*	1.85	1.61	2.73	2.64	mitotic spindle checkpoint protein Mad3
SPCC1795.09	*yps1*	1.62	2.14	2.56	4.02	aspartic protease, yapsin Yps1
SPCC1259.01c	*rps1802*	1.56	1.93	2.28	4.04	40S ribosomal protein S18 (predicted)
SPNCRNA.06	*prl6*	1.53	0.85	2.44	0.79	non-coding RNA, poly(A)-bearing (predicted)
SPCC1322.16	*phb2*	2.38	2.18	1.15	1.07	prohibitin Phb2 (predicted)
SPCC550.05	*nse1*	2.12	2.07	1.14	1.20	Smc5-6 complex non-SMC subunit 1
SPCC550.04c	*gpi2*	2.28	1.88	1.11	0.99	pig-C (predicted)

SPCC645.02	*gep4*	2.44	1.55	1.11	0.91	mitochondrial matrix PGP phosphatase involved in cardiolipin biosynthesis Gep4 (predicted)
SPCC622.21	*wtf12*	1.54	1.24	1.11	0.97	wtf element Wtf12
SPCC622.07		2.07	2.30	1.11	1.30	dubious
SPCC4B3.08	*lsg1*	2.30	2.68	1.09	1.42	Lsk1 complex gamma subunit (predicted)
SPCC61.04c		2.19	1.17	1.09	1.14	Rab GTPase binding (predicted)
SPCC645.03c	*isa1*	2.18	2.40	1.09	1.11	mitochondrial iron-sulfur protein Isa1
SPCC550.08		2.18	1.58	1.08	0.91	N-acetyltransferase (predicted)
SPCC550.01c		2.14	1.87	1.08	1.07	mitochondrial respiratory chain complex assembly protein (predicted)
SPCC1322.07c	*mug150*	2.14	1.87	1.07	1.26	sequence orphan
SPCC1322.14c	*vtc4*	2.25	1.18	1.07	0.75	vacuolar transporter chaperone (VTC) complex subunit (predicted)
SPCC4B3.07	*nro1*	2.19	1.75	1.07	0.94	nuclear pore complex associated protein
SPCC338.02	*mug112*	2.21	1.78	1.06		dubious
SPCC23B6.05c	*ssb3*	2.11	1.83	1.06	0.74	DNA replication factor A subunit Ssb3
SPCC4B3.16	*tip41*	2.10	1.65	1.06	0.95	TIP41-like type 2a phosphatase regulator Tip41
SPCC645.11c	*mug117*	2.15	2.46	1.05	1.19	meiotically upregulated gene Mug117
SPCC1322.15	*rpl3402*	2.02	1.97	1.05	1.13	60S ribosomal protein L34
SPCC550.07		2.40	2.30	1.05	1.27	acetamidase (predicted)
SPCC1281.03c	*emc4*	2.11	2.04	1.05	1.01	ER membrane protein complex subunit 4 (predicted)
SPCC132.01c		2.32	1.97	1.05	0.95	DUF814 family protein
SPCC622.19	*jmj4*	2.33	1.34	1.05	0.94	Jmj4 protein (predicted)
SPCC23B6.04c		2.17	1.76	1.04	0.91	sec14 cytosolic factor family (predicted)
SPCC550.12	*arp6*	1.83	1.96	1.04	0.86	actin-like protein Arp6
SPCC338.07c	*naa15*	2.33	1.72	1.04	0.97	NatA N-acetyltransferase complex regulatory subunit Naa15 (predicted)
SPCC1281.04		2.26	2.26	1.04	1.20	pyridoxal reductase (predicted)

SPCC1281.02c	*spf30*	2.20	1.49	1.04	0.79	splicing factor Spf30 (predicted)
SPCC338.18		2.06	2.90	1.03	1.21	sequence orphan
SPCC23B6.03c	*tel1*	1.96	1.98	1.03	0.96	ATM checkpoint kinase
SPCC1281.07c		2.21	1.76	1.03	0.80	glutathione S-transferase Gst3 (predicted)
SPCC23B6.01c		2.23	2.16	1.03	0.84	oxysterol binding protein (predicted)
SPCC645.05c	*myo2*	2.10	1.85	1.03	0.96	myosin II heavy chain
SPCC622.11		2.21	2.45	1.03	1.10	LMBR1-like membrane protein
SPCC338.06c		2.23	2.04	1.03	0.77	heat shock protein Hsp20 family (predicted)
SPCC338.03c		2.20	2.21	1.03	1.22	dubious
SPCC645.09	*mrpl37*	2.29	2.14	1.02	1.15	mitochondrial ribosomal protein subunit L37 (predicted)
SPCC550.10	*meu8*	2.10	1.87	1.02	1.01	aldehyde dehydrogenase Meu8 (predicted)
SPCC622.05		2.18	1.36	1.02	0.99	dubious
SPCC338.15	*wbp1*	2.21	1.80	1.02	0.97	dolichyl-di-phosphooligosaccharide-protein glycotransferase subunit Wbp1 (predicted)
SPCC645.06c	*rgf3*	2.15	2.02	1.02	0.93	RhoGEF Rgf3
SPCC338.13	*cog4*	2.18	1.96	1.02	1.02	Golgi transport complex subunit Cog4 (predicted)
SPCC645.08c	*snd1*	2.21	1.91	1.02	0.93	RNA-binding protein Snd1
SPCC11E10.02c	*gpi8*	1.66	1.17	1.02	1.08	pig-K
SPCC622.09	*htb1*	2.27	2.34	1.02	0.79	histone H2B Htb1
SPCC338.12	*pbi2*	2.24	5.39	1.02	1.43	proteinase B inhibitor Pbi2 (predicted)
SPCC645.12c		1.98	2.69	1.02	1.08	sequence orphan
SPCC1281.08	*wtf11*	2.09	2.20	1.02	1.24	wtf element Wtf11
SPCC4B3.12	*set9*	2.13	1.84	1.01	1.05	histone lysine methyltransferase Set9
SPCC4B3.18		2.17	2.05	1.01	1.30	phosphopantothenate-cysteine ligase (predicted)
SPCC622.14		2.18	1.88	1.01	1.01	GTPase activating protein (predicted)

SPCC338.16	*pof3*	2.13	1.87	1.01	1.02	F-box protein Pof3
SPCC645.10		2.09	2.16	1.01	0.90	ATP(CTP) tRNA nucleotidyltransferase (predicted)
SPCC11E10.03	*mug1*	1.00	0.71	1.01	0.83	dynactin complex subunit, dynamitin (predicted)
SPCC1322.01	*rpm1*	2.04	3.62	1.01	0.99	3'-5' exonuclease for RNA 3' ss-tail
SPCC61.05		1.99	0.95	1.01	0.93	*S. pombe* specific multicopy membrane protein family 1
SPCC1322.02		1.99	2.34	1.01	1.14	sequence orphan
SPCC622.06c		2.13	2.36	1.01	1.24	dubious
SPCC1322.05c		2.07	1.97	1.00	1.07	leukotriene A-4 hydrolase (predicted)
SPCC4B3.17	*cbp3*	1.97	1.80	1.00	0.96	ubiquinol cytochrome-c reductase assembly protein Cbp3 (predicted)
SPCC1322.09		2.10	2.11	1.00	0.99	conserved fungal protein
SPCC622.02		2.04	1.87	1.00	1.08	dubious
SPCC132.02	*hst2*	2.05	1.57	1.00	0.84	Sir2 family histone deacetylase Hst2
SPCC4E9.02	*cig1*	2.07	1.31	0.99	0.76	cyclin Cig1
SPCC645.04	*nse3*	2.02	1.89	0.99	1.13	Smc5-6 complex non-SMC subunit Nse3
SPCC550.15c		2.01	1.48	0.99	0.89	ribosome biogenesis protein (predicted)
SPCC4B3.15	*mid1*	2.30	1.71	0.99	0.90	medial ring protein Mid1
SPCC4B3.13		2.03	1.56	0.99	0.83	MatE family transporter (predicted)
SPCC622.04		2.14	1.67	0.98	1.15	dubious
SPCC550.09		1.97	1.91	0.98	1.04	peroxin Pex32 (predicted)
SPCC1322.13	*ade6*	2.33	1.83	0.98	1.03	phosphoribosylaminoimidazole carboxylase Ade6
SPCC645.14c	*sti1*	2.21	3.38	0.98	0.82	chaperone activator Sti1 (predicted)
SPCC4B3.09c		2.27	1.82	0.98	0.90	mitochondrial ribosomal protein subunit L12 (predicted)
SPCC622.15c		1.99	2.67	0.97	0.95	sequence orphan
SPCC550.02c	*cwf5*	2.08	1.46	0.97	0.79	RNA-binding protein Cwf5
SPCC4E9.01c	*rec11*	2.03	1.58	0.97	0.76	meiotic cohesin complex subunit Rec11

SPCC550.03c		1.95	1.91	0.97	0.99	Ski complex RNA helicase Ski2 (predicted)
SPCC622.16c	*epe1*	2.01	1.70	0.97	0.79	Jmjc domain chromatin associated protein Epe1
SPCC61.03		1.93	1.24	0.97	0.91	carbohydrate kinase domain family protein
SPCC132.03		2.00	2.68	0.97	1.34	sequence orphan
SPCC1322.12c	*bub1*	2.10	1.56	0.96	0.97	serine/threonine protein kinase Bub1
SPCC4B3.01	*tum1*	2.21	2.20	0.96	1.09	thiosulfate sulfurtransferase, involved in tRNA wobble position thiolation Tum1 (predicted)
SPCC550.13	*dfp1*	2.10	1.56	0.96	0.82	Hsk1-Dfp1 kinase complex regulatory subunit Dfp1
SPCC622.17	*apn1*	1.98	2.20	0.96	0.98	AP endonuclease Apn1
SPCC622.13c	*tti1*	2.06	1.79	0.96	0.79	Tel Two Interacting protein 1
SPCC61.02	*spt3*	1.99	1.02	0.96	0.54	SAGA complex subunit Spt3
SPCC550.14	*vgl1*	1.98	1.65	0.96	0.83	vigilin (predicted)
SPCC550.06c	*hsp10*	1.99	2.57	0.96	0.81	mitochondrial heat shock protein Hsp10 (predicted)
SPCC61.01c	*str2*	2.00	1.63	0.96	1.02	siderophore-iron transporter Str2
SPCC1322.11	*rpl2302*	1.83	2.10	0.96	1.20	60S ribosomal protein L23
SPCC338.08	*ctp1*	2.08	2.04	0.96	0.93	CtIP-related endonuclease
SPCC1322.10		2.16	2.66	0.96	0.64	cell wall protein Pwp1
SPCC23B6.02c		2.10	1.63	0.95	0.88	pre-ribosomal factor (predicted)
SPCC4B3.03c		2.12	1.74	0.95	0.98	mitochondrial morphology protein (predicted)
SPCC4B3.04c	*nte1*	2.16	2.16	0.95	1.11	lysophospholipase (predicted)
SPNCRNA.122		1.98	2.16	0.95	0.79	non-coding RNA (predicted)
SPCC338.17c	*rad21*	2.02	2.12	0.95	0.91	mitotic cohesin complex, non-SMC subunit Rad21 (kleisin)
SPCC4B3.11c		1.90	1.86	0.95	1.10	mitochondrial conserved eukaryotic protein
SPCC338.05c	*mms2*	1.94	2.33	0.94	1.11	ubiquitin conjugating enzyme Mms2
SPCC622.03c		1.94	1.70	0.94	1.12	dubious

SPCC338.04	*cid2*	1.70	4.51	0.94	1.03	caffeine induced death protein Cid2
SPCC4B3.14	*cwf20*	2.05	2.04	0.94	1.01	complexed with Cdc5 protein Cwf20
SPNCRNA.120		2.00	2.09	0.94	1.11	non-coding RNA (predicted)
SPCC645.07	*rgf1*	1.93	1.61	0.94	0.93	RhoGEF for Rho1, Rgf1
SPCC4B3.10c	*ipk1*	1.95	1.21	0.94	0.76	inositol 1,3,4,5,6-pentakisphosphate (IP5) kinase
SPCC132.04c	*gdh2*	1.93	2.42	0.93	0.80	NAD-dependent glutamate dehydrogenase Gdh2 (predicted)
SPCC4B3.02c		1.92	2.06	0.93	1.13	Golgi transport protein Got1 (predicted)
SPCC622.18	*rpl6*	2.03	2.30	0.93	0.98	60S ribosomal protein L6 (predicted)
SPCC11E10.01		1.93	0.97	0.93	0.92	cystathionine beta-lyase (predicted)
SPCC622.10c		2.12	1.97	0.93	0.97	exocyst complex subunit Sec5 (predicted)
SPCC4B3.06c		1.96	1.48	0.93	1.00	NADPH-dependent FMN reductase (predicted)
SPCC338.11c	*rrg1*	1.91	1.92	0.92	1.02	methyltransferase Rrg1 (predicted)
SPCC550.11		1.90	1.62	0.92	0.88	karyopherin (predicted)
SPCC645.13		1.93	1.79	0.92	1.15	transcription elongation regulator (predicted)
SPCC1281.05	*rsc7*	1.91	1.90	0.91	0.84	RSC complex subunit Rsc7
SPCC1322.06	*kap113*	1.93	1.87	0.91	1.02	karyopherin Kap113
SPCC622.08c	*hta1*	1.82	2.19	0.91	0.78	histone H2A alpha
SPCC4B3.05c	*hem12*	1.95	1.65	0.91	0.89	uroporphyrinogen decarboxylase Hem12 (predicted)
SPCC1281.06c		1.93	1.82	0.91	1.28	acyl-coA desaturase (predicted)
SPCC1322.03		1.90	1.88	0.90	0.70	TRP-like ion channel (predicted)
SPCP25A2.03		1.86	1.79	0.90	0.91	THO complex subunit (predicted)
SPCC338.14		1.99	1.61	0.89	0.83	adenosine kinase (predicted)
SPCC622.01c		1.86	1.80	0.88	0.79	dubious
SPCC132.05c		1.82	2.06	0.88	1.01	trichothecene 3-O-acetyltransferase pseudogene

SPCP25A2.02c	*rhp26*	1.87	1.57	0.87	0.82	SNF2 family helicase Rhp26
SPCC1322.04		1.91	3.28	0.87	0.98	UTP-glucose-1-phosphate uridylyltransferase (predicted)
SPCC622.12c	*gdh1*	1.85	2.29	0.86	0.96	NADP-specific glutamate dehydrogenase Gdh1 (predicted)
SPCC1281.01	*ags1*	1.79	1.50	0.86	0.79	alpha-1,4-glucan synthase Ags1
SPCC338.10c	*cox5*	1.91	1.46	0.85	0.73	cytochrome c oxidase subunit V (predicted)
SPCC1322.08	*srk1*	1.95	1.95	0.85	0.71	MAPK-activated protein kinase Srk1

Appendix V: Differentially regulated genes in aneuploid 1 and 2

Table 1 A: Genes showing 2 fold up regulation in aneuploid 1

Systenatic ID	common names	Description
Transcription		
SPBC1773.12		transcription factor (predicted)
SPBPB8B6.04c	*grt1*	transcription factor (predicted)
SPBC29B5.01	*atf1*	transcription factor, Atf-CREB family
SPBC83.17		transcriptional coactivator, multiprotein bridging factor (predicted)
Metabolic genes		
SPBPB21E7.04c		human COMT ortholog 2
SPAC4H3.08		3-hydroxyacyl-CoA dehydrogenase (predicted)
SPAC3F10.18c	*rpl4102*	60S ribosomal protein (predicted)
SPBPB8B6.03		acetamidase (predicted)
SPBC359.02	*alr2*	alanine racemase (predicted)
SPACUNK4.16c		alpha,alpha-trehalose-phosphate synthase (predicted)
SPBPB2B2.01		amino acid permease (predicted)
SPBPB10D8.02c		arylsulfatase (predicted)
SPBC30D10.14		dienelactone hydrolase family (predicted)
SPBPB21E7.01c	*eno102*	enolase (predicted)
SPAC26H5.09c		gfo/idh/mocA family oxidoreductase (predicted)
SPAC1002.19	*urg1*	GTP cyclohydrolase II (predicted)
SPAC4G9.12		gluconokinase
SPAC186.08c		L-lactate dehydrogenase (predicted)
SPBPB2B2.05		peptidase family C26 protein
SPBPB21E7.02c		phosphoglycerate mutase family
SPBC23G7.10c		NADH-dependent flavin oxidoreductase (predicted)
SPAC1002.17c	*urg2*	uracil phosphoribosyltransferase (predicted)
SPBC1683.06c		uridine ribohydrolase (predicted)
SPAC977.08		short chain dehydrogenase (predicted)
Recombination and repair		
SPBC23G7.11	*mag2*	DNA-3-methyladenine glycosidase
SPCC1753.03c	*rec7*	meiotic recombination protein
SPAC212.06c		DNA helicase in rearranged telomeric region, truncated
SPAC212.11	*tlh1*	RecQ type DNA helicase
SPBCPT2R1.08c	*tlh2*	RecQ type DNA helicase
Heat shock		

SPBC16D10.08c		heat shock protein Hsp104 (predicted)
SPAC13G7.02c	*ssa1*	heat shock protein (predicted)
SPBC1348.14c	*ght7*	hexose transporter (predicted)
SPBC3B9.01		Hsp70 nucleotide exchange factor (predicted)
SPBC4F6.17c		mitochondrial heatshock protein Hsp78 (predicted)
Membrane transporters		
SPAC1F7.08	*fio1*	iron transport multicopper oxidase
SPBC1348.05		membrane transporter (predicted)
SPBC1348.11		membrane transporter, pseudogene
SPBPB2B2.16c		MFS family membrane transporter (predicted)
SPAC977.04		truncated C terminal region of membrane transporter
SPBPB8B6.02c		urea transporter (predicted)
Meiosis		
SPBC1711.02	*mat3-Mc*	mating-type m-specific polypeptide mc 1
SPBC31F10.08	*mde2*	Mde2 protein
SPMTR.01	*matPc*	P-specific polypeptide Pc
SPCC320.07c	*mde7*	RNA-binding protein
Transposons		
SPAC19D5.09c	*Tf2-8*	retrotransposable element
SPAC2E1P3.03c	*Tf2-3*	retrotransposable element
SPAC9.04	*Tf2-1*	retrotransposable element
SPBC1289.17	*Tf2-11*	retrotransposable element
SPBC1E8.04	*Tf2-10*	retrotransposable element: pseudo
SPCC1494.11c	*Tf2-13*	retrotransposable element: pseudo
Unknown Biological functions		
SPBC1348.03		*S. pombe* specific 5Tm protein family
SPBPB2B2.17c		*S. pombe* specific 5Tm protein family
SPAC977.02		*S. pombe* specific 5Tm protein family
SPBPB2B2.19c		*S. pombe* specific 5Tm protein family
SPAC750.05c		*S. pombe* specific 5Tm protein family
SPAC977.01		*S. pombe* specific 5Tm protein family
SPAC212.04c		*S. pombe* specific DUF999 family protein 1
SPAC977.06		*S. pombe* specific DUF999 family protein 3
SPBC1348.07		*S. pombe* specific DUF999 protein family 6
SPAC212.12		*S. pombe* specific GPI anchored protein family
SPAC750.07c		*S. pombe* specific GPI anchored protein family 1
SPAC212.08c		GPI anchored protein (predicted)
SPBPB2B2.18		sequence orphan
SPAC1F8.02c		sequence orphan
SPBC1271.08c		sequence orphan
SPBC83.19c		sequence orphan
SPAC1687.23c		sequence orphan

SPBC11C11.06c		sequence orphan
SPBPB2B2.15		conserved fungal family
SPAC977.05c		conserved fungal family
SPAC343.12	*rds1*	conserved fungal protein
SPBC19C7.04c		conserved fungal protein
SPAC637.03		conserved fungal protein
SPAC57A7.05		conserved protein
SPAC806.11		dubious
SPAC26F1.11		dubious
SPAC343.20		dubious
SPBPB21E7.06		pseudogene
SPBCPT2R1.06c		pseudogene
SPBPB21E7.08		pseudogene
SPAC212.05c		pseudogene
SPBCPT2R1.05c		pseudogene
Various functions		
SPBC21C3.19	*rtc3*	SBDS family protein (predicted)
SPBC1348.13		similar to fragment of cox1 intron protein
SPNCRNA.445	*snoR61*	small nucleolar RNA
SPBC1348.12		zinc finger protein
SPMIT.08		mitochondrial ribosomal small subunit (predicted)
SPAC27D7.11c		But2 family protein
SPCC830.07c	*psi1*	DNAJ domain protein, involved in translation initiation
SPBPB21E7.07	*aes1*	enhancer of RNA-mediated gene silencing
SPAC21E11.04	*ppr1*	L-azetidine-2-carboxylic acid acetyltransferase
SPBP4G3.03		PI31 proteasome regulator related

Table 1 B: Genes showing 2 fold down regulation in aneuploid 1

Systenatic ID	**common names**	**Description**
Membrane transporters		
SPAC23D3.12		inorganic phosphate transporter (predicted)
SPBC1683.01		inorganic phosphate transporter (predicted)
SPBC8E4.01c		inorganic phosphate transporter (predicted)
SPBC530.02		membrane transporter (predicted)
SPAC1B3.15c		vitamin H transporter (predicted)
Metabolic genes		
SPBP4G3.02	*pho1*	acid phosphatase
SPAC23D3.05c		alcohol dehydrogenase pseudogene

SPAC4G9.10	*arg3*	ornithine carbamoyltransferase
Various functions		
SPAC27D7.03c	*mei2*	RNA-binding protein involved in meiosis
SPBC1685.06	*cid11*	poly(A) polymerase (predicted)
SPCC1223.13	*cbf12*	CBF1/Su(H)/LAG-1 family transcription factor
SPAPB1A10.14		F-box protein (predicted)
SPCC1020.09	*gnr1*	heterotrimeric G protein beta subunit
SPAC18G6.01c		calchone related protein family

Table 2 A: Genes showing 2 fold up regulation in aneuploid 2

Systematic ID	Common name	Function
Metabolic genes		
SPMIT.01	*cox1*	cytochrome c oxidase 1 (predicted)
SPMIT.04	*cox3*	cytochrome c oxidase 3 (predicted)
SPBPB21E7.01c	*eno102*	enolase (predicted)
SPBPB21E7.04c		human COMT ortholog 2
SPAC186.02c		hydroxyacid dehydrogenase (predicted)
SPAC186.08c		L-lactate dehydrogenase (predicted)
SPBC23G7.10c		NADH-dependent flavin oxidoreductase (predicted)
SPBPB21E7.02c		phosphoglycerate mutase family
SPBC24C6.09c		phosphoketolase family protein (predicted)
Recombination		
SPAC212.06c		DNA helicase in rearranged telomeric region, truncated
SPAC212.11	*tlh1*	RecQ type DNA helicase
SPBCPT2R1.08c	*tlh2*	RecQ type DNA helicase
Meiosis		
SPBC1711.02	*mat3-Mc*	mating-type m-specific polypeptide
SPBC23G7.17c	*mat3-m*	mating-type M-specific polypeptide
SPBC31F10.08	*mde2*	Mde2 protein
SPAC513.03	*mfm2*	M-factor precursor
Unknown functions		
SPBPB2B2.17c		*S. pombe* specific 5Tm protein family
SPBC1348.03		*S. pombe* specific 5Tm protein family
SPAC977.02		*S. pombe* specific 5Tm protein family
SPBPB2B2.19c		*S. pombe* specific 5Tm protein family
SPAC212.04c		*S. pombe* specific DUF999 family protein 1
SPAC212.12		*S. pombe* specific GPI anchored protein family

SPAC750.07c		*S. pombe* specific GPI anchored protein family 1
SPAC513.04		sequence orphan
SPBPB2B2.18		sequence orphan
SPAC23H3.15c		sequence orphan
SPBCPT2R1.06c		pseudogene
SPBPB21E7.06		pseudogene
SPBPB2B2.15		conserved fungal family
SPAC869.09		conserved fungal protein
SPBPB2B2.08		conserved fungal protein
SPBCPT2R1.09c		Obsolete in GeneDB
SPBPB2B2.03c		Obsolete in GeneDB
SPNCRNA.288		Obsolete Synonym
membrane transpoters		
SPBPB8B6.02c		urea transporter (predicted)
SPAC1F8.03c	*str3*	siderophore-iron transporter
SPBC1348.11		membrane transporter, pseudogene
SPBPB2B2.01		amino acid permease (predicted)
SPBC1348.14c	*ght7*	hexose transporter Ght7 (predicted)
SPAC1F7.08	*fio1*	iron transport multicopper oxidase Fio1
SPAC29B12.12		helper of TIM (predicted)
SPAC212.08c		GPI anchored protein (predicted)
Various functions		
SPBC1348.13		similar to fragment of cox1 intron protein
SPNCRNA.445	*snoR61*	small nucleolar RNA s
SPBP4G3.03		PI31 proteasome regulator related
SPBC3E7.02c	*hsp16*	heat shock protein
SPAC869.06c		HHE domain cation binding protein (predicted)
SPBC1773.12		transcription factor (predicted)
Translation		
SPBC16D10.11c	*rps1801*	40S ribosomal protein S18 (predicted)
SPCC16C4.13c	*rpl1201*	60S ribosomal protein L12.1/L12A
SPMIT.08		mitochondrial ribosomal small subunit (predicted)

Table 2B: Genes showing 2 fold down regulation in aneuploid 2

Systematic ID	Common name	Function
Metabolic genes		
SPBC1815.01	*eno1*	enolase (predicted)
SPBC11B10.02c	*his3*	histidinol-phosphate aminotransferase imidazole acetol phosphate transaminase

Unknown functions		
SPBC660.05		conserved fungal protein
SPAC26F1.11		dubious
SPAC27E2.13		dubious
SPCC191.10		sequence orphan
Membrane transporters		
SPBC409.08		spermine family transporter (predicted)
SPBC530.02		membrane transporter (predicted)
Various function		
SPCC63.13		DNAJ domain protein
SPCC1020.09	*gnr1*	heterotrimeric G protein beta subunit
SPAC2H10.01		transcription factor, zf-fungal binuclear cluster type (predicted)
SPBC1685.13	*fhn1*	Fhn1 plasma membrane organization protein

Appendix VI: Percent identity matrix and phylogenetic trees for paralogs

Table 1a: Percent identity matrix created by Clustal 2.1 sowing sequence similarity among different paralogs of SPBPB2B2.19c

	SPBPB2B2.19c	**SPAC977.01**	**SPBC1348.02**	**SPAC750.05c**	**SPAC977.02**
SPBPB2B2.19c	100	99.89	100	99.71	48.28
SPAC977.01	99.89	100	99.89	99.79	48.12
SPBC1348.02	100	99.89	100	99.71	48.28
SPAC750.05c	99.71	99.79	99.71	100	48.51
SPAC977.02	48.28	48.12	48.28	48.51	100

Figure III a: Phylogenetic tree for SPBPB2B2.19c and its 4 other paralog. DNA sequences for coding regions were aligned as mentioned for figure 5.1. Phylogenetic tree was created in Jalview 2.8 software that calculates average distance between sequences at each aligned position.

Table 1: Percent identity matrix created by Clustal Omega

	SPAC212.01c	SPBCPT2R1.01c	SPBCPT2R1.04c	SPBC1348.01	SPAC212.04c	SPAC750.06c	SPAC977.06	SPBC1348.07	SPBPB2B2.07c	SPBPB2B2.14c
SPAC212.01c	100	99.75	99.88	99.75	82.33	99.76	80.44	81.34	83.4	81.34
SPBCPT2R1.01c	99.75	100	99.63	100	83.35	99.51	80.44	81.34	83.26	81.34
SPBCPT2R1.04c	99.88	99.63	100	99.63	82.45	99.64	80.26	81.19	83.26	81.19
SPBC1348.01	99.75	100	99.63	100	83.35	99.51	80.44	81.34	83.26	81.34
SPAC212.04c	82.33	83.35	82.45	83.35	100	82.09	78.14	80.21	71.65	79.91
SPAC750.06c	99.76	99.51	99.64	99.51	82.09	100	80.07	81.04	83.12	81.04
SPAC977.06	80.44	80.44	80.26	80.44	78.14	80.07	100	99.65	89.3	99.65
SPBC1348.07	81.34	81.34	81.19	81.34	80.21	81.04	99.65	100	85.98	99.42
SPBPB2B2.07c	83.4	83.26	83.26	83.26	71.65	83.12	89.3	85.98	100	85.67
SPBPB2B2.14c	81.34	81.34	81.19	81.34	79.91	81.04	99.65	99.42	85.67	100

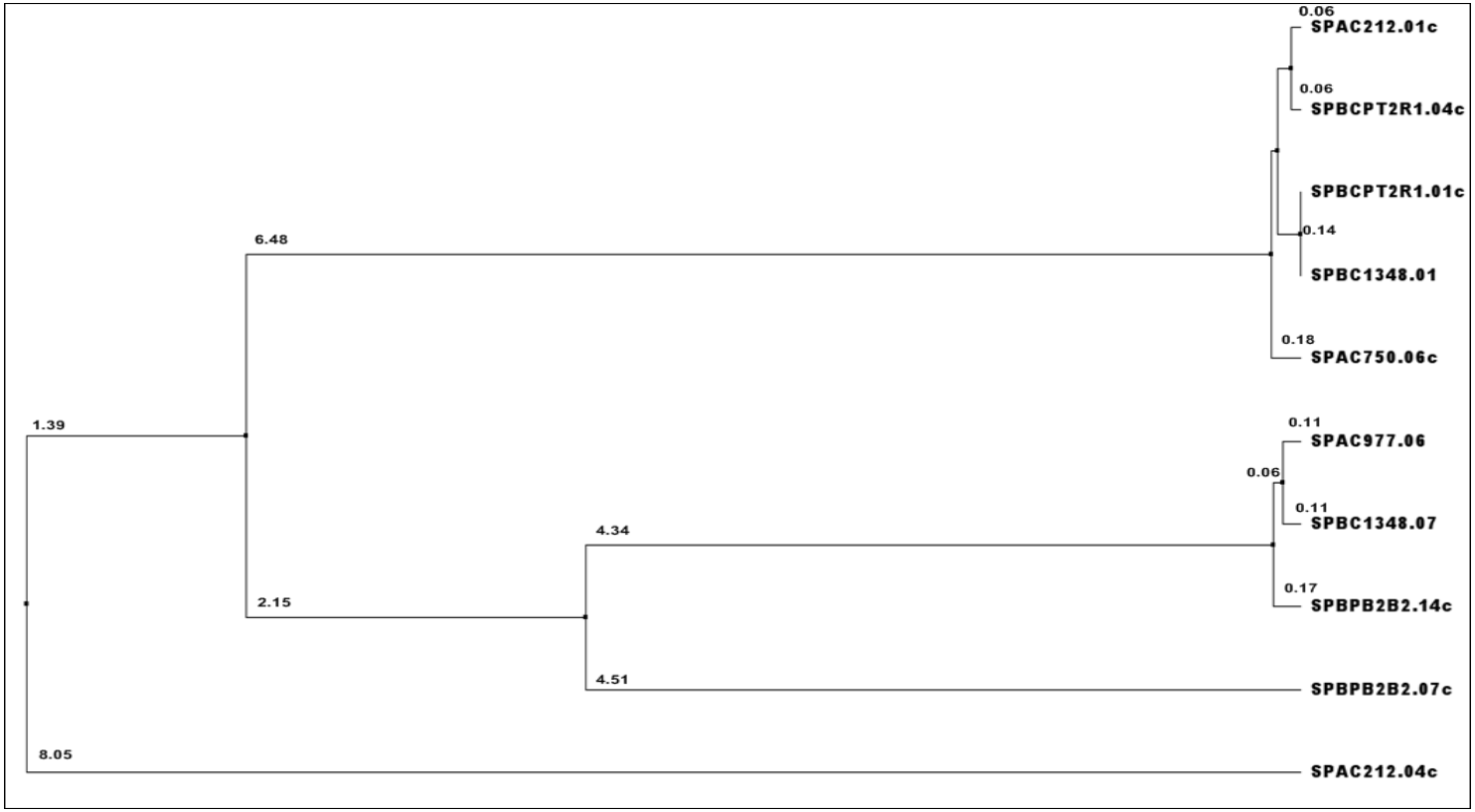


Figure III b: Phylogenetic tree for SPAC212.01c and its 9 other paralog. DNA sequences for coding regions were aligned as mentioned for figure 5.1. Phylogenetic tree was created in Jalview 2.8 software that calculates average distance between sequences at each aligned position.

References

References

Adams, K.L., Percifield, R., and Wendel, J.F. (2004). Organ-specific silencing of duplicated genes in a newly synthesized cotton allotetraploid. Genetics *168*, 2217–2226.

Albertin, W., Balliau, T., Brabant, P., Chèvre, A.-M., Eber, F., Malosse, C., and Thiellement, H. (2006). Numerous and rapid nonstochastic modifications of gene products in newly synthesized *Brassica napus* allotetraploids. Genetics *173*, 1101–1113.

Albertin, W., and Marullo, P. (2012). Polyploidy in fungi: evolution after whole-genome duplication. Proceedings. Biological Sciences / The Royal Society *279*, 2497–2509.

Albertin, W., Marullo, P., Aigle, M., Bourgais, a, Bely, M., Dillmann, C., DE Vienne, D., and Sicard, D. (2009). Evidence for autotetraploidy associated with reproductive isolation in *Saccharomyces cerevisiae*: towards a new domesticated species. Journal of Evolutionary Biology *22*, 2157–2170.

Anatskaya, O. V, and Vinogradov, A.E. (2007). Genome multiplication as adaptation to tissue survival: evidence from gene expression in mammalian heart and liver. Genomics *89*, 70–80.

Andalis, A. a, Storchova, Z., Styles, C., Galitski, T., Pellman, D., and Fink, G.R. (2004). Defects arising from whole-genome duplications in *Saccharomyces cerevisiae*. Genetics *167*, 1109–1121.

Andreassen, P.R., Martineau, S.N., and Margolis, R.L. (1996). Chemical induction of mitotic checkpoint override in mammalian cells results in aneuploidy following a transient tetraploid state. Mutation Research *372*, 181–194.

Arabidopsis, T., and Genome, I. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature *408*, 796–815.

Arts, B., Simons, C.J.P., Drukker, M., and Van Os, J. (2013). Antipsychotic medications and cognitive functioning in bipolar disorder: moderating effects of COMT Val108/158 Met genotype. BMC Psychiatry *13*, 63.

Askree, S.H., Yehuda, T., Smolikov, S., Gurevich, R., Hawk, J., Coker, C., Krauskopf, A., Kupiec, M., and McEachern, M.J. (2004). A genome-wide screen for *Saccharomyces cerevisiae* deletion mutants that affect telomere length. Proceedings of the National Academy of Sciences of the United States of America *101*, 8658–8663.

Aversano, R., Ercolano, M.R., Caruso, I., Fasano, C., Rosellini, D., and Carputo, D. (2012). Molecular tools for exploring polyploid genomes in plants. International Journal of Molecular Sciences *13*, 10316–10335.

Bähler, J., Wu, J.Q., Longtine, M.S., Shah, N.G., McKenzie, A., Steever, a B., Wach, A., Philippsen, P., and Pringle, J.R. (1998). Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. Yeast *14*, 943–951.

Bennett, M.D. (2004). Biological relevance of polyploidy : ecology to genomics Perspectives on polyploidy in plants – ancient and neo. Biological Journal of the Linnean Society *82*, 411–423.

Blanc, G., Hokamp, K., and Wolfe, K.H. (2003). A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. Genome Research *13*, 137–144.

Blanc, G., and Wolfe, K.H. (2004). Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. The Plant Cell *16*, 1679–1691.

Bossie, M. a, DeHoratius, C., Barcelo, G., and Silver, P. (1992). A mutant nuclear protein with similarity to RNA binding proteins interferes with nuclear import in yeast. Molecular Biology of the Cell *3*, 875–893.

Boveri, T. (2008). Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Henry Harris. J Cell Sci *121*, 1–84.

Brown, S. (2008). Miscarriage and its associations. Seminars in Reproductive Medicine *26*, 391–400.

Cavalier-Smith, T. (2005). Economy, speed and size matter: evolutionary forces driving nuclear genome miniaturization and expansion. Annals of Botany *95*, 147–175.

Chen, D., Toone, W.M., Mata, J., Lyne, R., Burns, G., Kivinen, K., Brazma, A., and Jones, N. (2003). Global Transcriptional Responses of Fission Yeast to Environmental Stress. Molecular Biology of the Cell *14*, 214–229.

Chikashige, Y., Tsutsumi, C., Okamasa, K., Yamane, M., Nakayama, J., Niwa, O., Haraguchi, T., and Hiraoka, Y. (2007). Gene expression and distribution of Swi6 in partial aneuploids of the fission yeast *Schizosaccharomyces pombe*. Cell Structure and Function *32*, 149–161.

Comai, L. (2005). The advantages and disadvantages of being polyploid. Nature Reviews. Genetics *6*, 836–846.

Cook, M., and Tyers, M. (2007). Size control goes global. Current Opinion in Biotechnology *18*, 341–350.

Duboule, D., and Dollé, P. (1989). The structural and functional organization of the murine HOX gene family resembles that of *Drosophila* homeotic genes. The EMBO Journal *8*, 1497–1505.

Dueñas-Santero, E., Martín-Cuadrado, A.B., Fontaine, T., Latgé, J.-P., Del Rey, F., and Vázquez de Aldana, C. (2010). Characterization of glycoside hydrolase family 5 proteins in *Schizosaccharomyces pombe*. Eukaryotic Cell *9*, 1650–1660.

Edgar, B. a, and Orr-Weaver, T.L. (2001). Endoreplication cell cycles: more for less. Cell *105*, 297–306.

Egel, R. (1989). Mating-type genes, meiosis, and sporulation. In Molecular Biology of the Fission Yeast., B.F. Nasim, A., Youn,g P., Johnson, ed. (New York: Academic Press), pp. 31–73.

Eiben, B., Bartels, I., Bähr-Porsch, S., Borgmann, S., Gatz, G., Gellert, G., Goebel, R., Hammans, W., Hentemann, M., and Osmers, R. (1990). Cytogenetic analysis of 750 spontaneous abortions with the direct-preparation method of chorionic villi and its implications for studying genetic causes of pregnancy wastage. American Journal of Human Genetics *47*, 656–663.

Ferrezuelo, F., Colomina, N., Palmisano, A., Garí, E., Gallego, C., Csikász-Nagy, A., and Aldea, M. (2012). The critical size is set at a single-cell level by growth rate to attain homeostasis and adaptation. Nature Communications *3*, 1012.

Galitski, T. (1999). Ploidy Regulation of Gene Expression. Science *285*, 251–254.

Ganem, N.J., Godinho, S. a, and Pellman, D. (2009). A mechanism linking extra centrosomes to chromosomal instability. Nature *460*, 278–282.

Gentric, G., Desdouets, C., and Celton-Morizur, S. (2012). Hepatocytes polyploidization and cell cycle control in liver physiopathology. International Journal of Hepatology *2012*, 282430.

Gisselsson, D., Håkanson, U., Stoller, P., Marti, D., Jin, Y., Rosengren, A.H., Stewénius, Y., Kahl, F., and Panagopoulos, I. (2008). When the genome plays dice: circumvention of the spindle assembly checkpoint and near-random chromosome segregation in multipolar cancer cell mitoses. PloS One *3*, e1871.

Gordon, D.J., Resio, B., and Pellman, D. (2012). Causes and consequences of aneuploidy in cancer. Nature Reviews. Genetics *13*, 189–203.

Gorla, G.R., Malhi, H., and Gupta, S. (2001). Polyploidy associated with oxidative injury attenuates proliferative potential of cells. Journal of Cell Science *114*, 2943–2951.

Graham, a, Papalopulu, N., and Krumlauf, R. (1989). The murine and *Drosophila* homeobox gene complexes have common features of organization and expression. Cell *57*, 367–378.

Guo, M., Davis, D., and Birchler, J. (1996). Dosage effects on gene expression in a maize ploidy series. Genetics *142*, 1349–1355.

Gupta, D.R., Paul, S.K., Oowatari, Y., Matsuo, Y., and Kawamukai, M. (2011). Multistep regulation of protein kinase A in its localization, phosphorylation and binding with a regulatory subunit in fission yeast. Current Genetics *57*, 353–365.

Hachet, O., Berthelot-Grosjean, M., Kokkoris, K., Vincenzetti, V., Moosbrugger, J., and Martin, S.G. (2011). A phosphorylation cycle shapes gradients of the DYRK family kinase Pom1 at the plasma membrane. Cell *145*, 1116–1128.

Han, T.X., Xu, X.-Y., Zhang, M.-J., Peng, X., and Du, L.-L. (2010). Global fitness profiling of fission yeast deletion strains by barcode sequencing. Genome Biology *11*, R60.

Hassold, T., Abruzzo, M., Adkins, K., Griffin, D., Merrill, M., Millie, E., Saker, D., Shen, J., and Zaragoza, M. (1996). Human aneuploidy: incidence, origin, and etiology. Environmental and Molecular Mutagenesis *28*, 167–175.

Hassold, T., Chen, N., Funkhouser, J., Jooss, T., Manuel, B., Matsuura, J., Matsuyama, A., Wilson, C., Yamane, J. A, and Jacobs, P. A. (1980). A cytogenetic study of 1000 spontaneous abortions. Annals of Human Genetics *44*, 151–178.

Herget, G.W., Neuburger, M., Plagwitz, R., and Adler, C.P. (1997). DNA content, ploidy level and number of nuclei in the human heart after myocardial infarction. Cardiovascular Research *36*, 45–51.

Huettel, B., Kreil, D.P., Matzke, M., and Matzke, A.J.M. (2008). Effects of aneuploidy on genome structure, expression, and interphase organization in Arabidopsis thaliana. PLoS Genetics *4*, e1000226.

Jiang, C., Wright, R.J., El-Zik, K.M., and Paterson, a H. (1998). Polyploid formation created unique avenues for response to selection in *Gossypium* (cotton). Proceedings of the National Academy of Sciences of the United States of America *95*, 4419–4424.

Johnston, G.C., Pringle, J.R., and Hartwell, L.H. (1977). Coordination of growth with cell division in the yeast *Saccharomyces cerevisiae*. Experimental Cell Research *105*, 79–98.

Jorgensen, P., Nishikawa, J.L., Breitkreutz, B. J., and Tyers, M. (2002). Systematic identification of pathways that couple cell growth and division in yeast. Science (New York, N.Y.) *297*, 395–400.

Jorgensen, P., and Tyers, M. (2004). How cells coordinate growth and division. Current Biology *14*, R1014–27.

Kahlem, P., Sultan, M., Herwig, R., Steinfath, M., Balzereit, D., Eppens, B., Saran, N.G., Pletcher, M.T., South, S.T., Stetten, G., et al. (2004). Transcript level alterations reflect gene dosage effects across multiple tissues in a mouse model of down syndrome. Genome Research *14*, 1258–1267.

Kashkush, K., Feldman, M., and Levy, A. a (2002). Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. Genetics *160*, 1651–1659.

Katayama, S., Kitamura, K., Lehmann, A., Nikaido, O., and Toda, T. (2002). Fission Yeast F-box Protein Pof3 Is Required for Genome Integrity and Telomere Function. Molecular Biology of the Cell *13*, 211–224.

Kim, D.-U., Hayles, J., Kim, D., Wood, V., Park, H.-O., Won, M., Yoo, H. S., Duhig, T., Nam, M., Palmer, G., et al. (2010). Analysis of a genome-wide set of gene deletions in the fission yeast *Schizosaccharomyces pombe*. Nature Biotechnology *28*, 617–623.

Kim, J.A., Jung, Y., Seoh, J.Y., Woo, S.Y., Seo, J.S., and Kim, H.L. (2002). Gene Expression Profile of Megakaryocytes from Human Cord Blood CD34+ Cells Ex Vivo Expanded by Thrombopoietin. 402–416.

Kohli, J. (1987). Genetic nomenclature and gene list of the fission yeast *Schizosaccharomyces pombe*. Current Genetics *11*, 575–589.

Kudryavtsev, B., and Kudryavtseva, M. (1993). Human hepatocyte polyploidization kinetics in the course of life cycle. VVirchows Arch B Cell Pathol Incl Mol Pathol *64*, 387–393.

Lazzerini Denchi, E., Celli, G., and De Lange, T. (2006). Hepatocytes with extensive telomere deprotection and fusion remain viable and regenerate liver mass through endoreduplication. Genes & Development *20*, 2648–2653.

Lengauer, C., Kinzler, K., and Vogelstein, B. (1997). Genetic instability in colorectal cancers. Nature. *386*, 623–627.

Lewis, W.H. (1980). polyploidy: Biological relevance (New York: plenum press).

López-Avilés, S., Grande, M., González, M., Helgesen, A.-L., Alemany, V., Sanchez-Piris, M., Bachs, O., Millar, J.B. a, and Aligue, R. (2005). Inactivation of the Cdc25 phosphatase by the stress-activated Srk1 kinase in fission yeast. Molecular Cell *17*, 49–59.

Lu, B.C. (1964). Polyploidy in the Basidiomycete *Cyathus stercoreus*. American Journal of Botany *51*, 341– 347.

Lyle, R., Gehrig, C., Neergaard-Henrichsen, C., Deutsch, S., and Antonarakis, S.E. (2004). Gene expression from the aneuploid chromosome in a trisomy mouse model of down syndrome. Genome Research *14*, 1268–1274.

Lyne, R., Burns, G., Mata, J., Penkett, C.J., Rustici, G., Chen, D., Langford, C., Vetrie, D., and Bähler, J. (2003). accuracy , reproducibility , and processing of array data. BMC Genomics *4*, 1–15.

Ma, L.-J., Ibrahim, A.S., Skory, C., Grabherr, M.G., Burger, G., Butler, M., Elias, M., Idnurm, A., Lang, B.F., Sone, T., et al. (2009). Genomic analysis of the basal

lineage fungus *Rhizopus oryzae* reveals a whole-genome duplication. PLoS Genetics *5*, e1000549.

Maines, J.Z., Stevens, L.M., Tong, X., and Stein, D. (2004). *Drosophila* dMyc is required for ovary cell growth and endoreplication. Development (Cambridge, England) *131*, 775–786.

Marguerat, S., and Bähler, J. (2012). Coordinating genome expression with cell size. Trends in Genetics *28*, 560–565.

Marshall, W.F., Young, K.D., Swaffer, M., Wood, E., Nurse, P., Kimura, A., Frankel, J., Wallingford, J., Walbot, V., Qu, X., et al. (2012). What determines cell size? BMC Biology *10*, 101.

Martin, S.G., and Berthelot-Grosjean, M. (2009). Polar gradients of the DYRK-family kinase Pom1 couple cell length with the cell cycle. Nature *459*, 852–856.

Masterson, J. (1994). Stomatal Size in Fossil Plants : Evidence for Polyploidy of in Majority of Angiosperms. Science *264*, 421–424.

Minn, a J., Boise, L.H., and Thompson, C.B. (1996). Expression of Bcl-xL and loss of p53 can cooperate to overcome a cell cycle checkpoint induced by mitotic spindle damage. Genes & Development *10*, 2621–2631.

Mitchison, J. (2003). Growth during the cell cycle. Int Rev Cytol *226*, 165–258.

Molnar, M., and Sipiczki, M. (1993). Polyploidy in the haplontic yeast *Schizosaccharomyces pombe*: construction and analysis of strains. Current Genetics *24*, 45–52.

Mortimer, R.K. (1958). Radiobiological and genetic studies on a polyploid series (haploid to hexaploid) of *Saccharomyces cerevisiae*. Radiation Research *9*, 312–326.

Moseley, J.B., Mayeux, A., Paoletti, A., and Nurse, P. (2009). A spatial gradient coordinates cell size and mitotic entry in fission yeast. Nature *459*, 857–860.

Nash, R., Tokiwa, G., Anand, S., Erickson, K., and Futcher, A.B. (1988). The WHI1 + gene of *Saccharomyces cerevisiae* tethers cell division to cell size and is a cyclin homolog. The EMBO Journal *7*, 4335–4346.

Naumov, G.I., Naumova, E.S., Masneuf, I., Aigle, M., Kondratieva, V.I., and Dubourdieu, D. (2000). Natural polyploidization of some cultured yeast *Saccharomyces sensu* stricto: auto- and allotetraploidy. Systematic and Applied Microbiology *23*, 442–449.

Neuber, M., Rehder, H., Zuther, C., Lettau, R., and Schwinger, E. (1993). Polyploidies in abortion material decrease with maternal age. Human Genetics *91*, 563–566.

Nielsen, K., and Yohalem, D.S. (2001). Origin of a polyploid Botrytis pathogen through interspecific hybridization between *Botrytis aclada* and *B. byssoidea*. Mycologia *93*, 1064–1071.

Nigg, E.A. (2002). Reviews and comment from the nature publishing group. Nature Reviews Genetics *3*, 815–815.

Niwa, O., Matsumoto, T., and Yanagida, M. (1986). Construction of a mini-chromosome by deletion and its mitotic and meiotic behaviour in fission yeast. Mol Gen Genet *203*, 397–405.

Niwa, O., Tange, Y., and Kurabayashi, A. (2006). Growth arrest and chromosome instability in aneuploid yeast. Yeast *23*, 937–950.

Nurse, P. (1990). Universal control mechanism regulating onset of M-phase. Nature *344*, 503–508.

Ohno, S. (1970). Evolution by gene duplication. (New York: Springer-Verlag.).

Oowatari, Y., Toma, K., Ozoe, F., and Kawamukai, M. (2009). Identification of *sam4* as a *rad24* Allele in *Schizosaccharomyces pombe*. Bioscience, Biotechnology, and Biochemistry *73*, 1591–1598.

Oromendia, A.B., Dodgson, S.E., and Amon, A. (2012). Aneuploidy causes proteotoxic stress in yeast. Genes & Development *26*, 2696–2708.

Osborn, T.C., Chris Pires, J., Birchler, J. a., Auger, D.L., Jeffery Chen, Z., Lee, H.-S., Comai, L., Madlung, A., Doerge, R.W., Colot, V., et al. (2003). Understanding mechanisms of novel gene expression in polyploids. Trends in Genetics *19*, 141–147.

Otto, S.P., and Whitton, J. (2000). Polyploid incidence and evolution. Annual Review of Genetics *34*, 401–437.

Pavelka, N., Rancati, G., Zhu, J., Bradford, W.D., Saraf, A., Florens, L., Sanderson, B.W., Hattem, G.L., and Li, R. (2010). Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. Nature *468*, 321–325.

Penkett, C.J., Morris, J. a, Wood, V., and Bähler, J. (2006). YOGY: a web-based, integrated database to retrieve protein orthologs and associated Gene Ontology terms. Nucleic Acids Research *34*, 330–334.

Ramsey, J., and Schemske, D.W. (1998). Pathways, Mechanisms, and Rates of Polyploid Formation in Flowering Plants. Annual Review of Ecology and Systematics *29*, 467–501.

Raslova, H., Kauffmann, A., Sekkaï, D., Ripoche, H., Larbret, F., Robert, T., Le Roux, D.T., Kroemer, G., Debili, N., Dessen, P., et al. (2007). Interrelation between polyploidization and megakaryocyte differentiation: a gene profiling approach. Blood *109*, 3225–3234.

Rasmussen, S. a, Wong, L.-Y.C., Yang, Q., May, K.M., and Friedman, J.M. (2003). Population-based analyses of mortality in trisomy 13 and trisomy 18. Pediatrics *111*, 777–784.

Ravid, K., Lu, J., Zimmet, J.M., and Jones, M.R. (2002). Roads to polyploidy: the megakaryocyte example. Journal of Cellular Physiology *190*, 7–20.

Rupes, I. (2002). Checking cell size in yeast. Trends in Genetics *18*, 479–485.

Sanz, N., Díez-Fernández, C., Alvarez, A., and Cascales, M. (1997). Age-dependent modifications in rat hepatocyte antioxidant defense systems. Journal of Hepatoiogy *27*, 525–534.

Schughart, K., Kappen, C., and Ruddle, F.H. (1989). Duplication of large genomic regions during the evolution of vertebrate homeobox genes. Proceedings of the National Academy of Sciences of the United States of America *86*, 7067–7071.

Segal, D.J., and Mccoy, E.E. (1973). Studies on Down's Syndrome in Tissue Culture. J. Cell. Physiol *83*, 85–90.

Sheltzer, J.M., Blank, H.M., Pfau, S.J., Tange, Y., George, B.M., Humpton, T.J., Brito, I.L., Hiraoka, Y., Niwa, O., and Amon, A. (2011). Aneuploidy drives genomic instability in yeast. Science *333*, 1026–1030.

Sheltzer, J.M., Torres, E.M., Dunham, M.J., and Amon, A. (2012). Transcriptional consequences of aneuploidy. PNAS Genetics *109*, 12644–12649.

Shevchenko, A., Roguev, A., Schaft, D., Buchanan, L., Habermann, B., Sakalar, C., Thomas, H., Krogan, N.J., Shevchenko, A., and Stewart, a F. (2008). Chromatin Central: towards the comparative proteome by accurate mapping of the yeast proteomic environment. Genome Biology *9*, R167.

Shikata, M., Ishikawa, F., and Kanoh, J. (2007). Tel2 is required for activation of the Mrc1-mediated replication checkpoint. The Journal of Biological Chemistry *282*, 5346–5355.

Shim, M., Hoover, A., Blake, N., Drachman, J.G., and Anna, J. (2004). Gene expression profile of primary human CD34 CD38 lo cells differentiating along the megakaryocyte lineage. Experimental Hematology *32*, 638–648.

Smith, a V, and Orr-Weaver, T.L. (1991). The regulation of the cell cycle during *Drosophila* embryogenesis: the transition to polyteny. Development (Cambridge, England) *112*, 997–1008.

Stingele, S., Stoehr, G., Peplowska, K., Cox, J., Mann, M., and Storchova, Z. (2012). Global analysis of genome, transcriptome and proteome reveals the response to aneuploidy in human cells. Molecular Systems Biology *8*, 608.

Storchova, Z., and Kuffer, C. (2008). The consequences of tetraploidy and aneuploidy. Journal of Cell Science *121*, 3859–3866.

Storchova, Z., and Pellman, D. (2004). From polyploidy to aneuploidy, genome instability and cancer. Nature Reviews. Molecular Cell Biology *5*, 45–54.

Takada, H., Nishida, A., Domae, M., Kita, A., Yamano, Y., Uchida, A., Ishiwata, S., Fang, Y., Zhou, X., Masuko, T., et al. (2010). The cell surface protein gene ecm33 is a target of the two transcription factors atf1 and mbx1 and negatively regulates pmk1 mapk cell integrity signaling in fission yeast. Molecular Biology of the Cell *21*, 674–685.

Tinline-Purvis, H., Savory, A.P., Cullen, J.K., Davé, A., Moss, J., Bridge, W.L., Marguerat, S., Bähler, J., Ragoussis, J., Mott, R., et al. (2009). Failed gene conversion leads to extensive end processing and chromosomal rearrangements in fission yeast. The EMBO Journal *28*, 3400–3412.

Torres, E.M., Sokolsky, T., Tucker, C.M., Chan, L.Y., Boselli, M., Dunham, M.J., and Amon, A. (2007). Effects of aneuploidy on cellular physiology and cell division in haploid yeast. Science *317*, 916–924.

Torres, E.M., Williams, B.R., and Amon, A. (2008). Aneuploidy: cells losing their balance. Genetics *179*, 737–746.

Turner, J.J., Ewald, J.C., and Skotheim, J.M. (2012). Cell size control in yeast. Current Biology *22*, 350–359.

Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proceedings of the National Academy of Sciences of the United States of America *98*, 5116–5121.

Tyers, M., Tokiwa, G., and Futcher, B. (1993). Comparison of the S*accharomyces cerevisiae* G1 cyclins: Cln3 may be an upstream activator of Cln1, Cln2 and other cyclins. The EMBO Journal *12*, 1955–1968.

Tyers, M., Tokiwa, G., Nash, R., and Futcher, B. (1992). The Cln3-Cdc28 kinase complex of *S. cerevisiae* is regulated by proteolysis and phosphorylation. The EMBO Journal *11*, 1773–1784.

Vignery, a (2000). Osteoclasts and giant cells: macrophage-macrophage fusion mechanism. International Journal of Experimental Pathology *81*, 291–304.

Wang, J., Tian, L., Madlung, A., Lee, H.-S., Chen, M., Lee, J.J., Watson, B., Kagochi, T., Comai, L., and Chen, Z.J. (2004). Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. Genetics *167*, 1961–1973.

Warner, D., Ku, M., and Edward, G. (1987). Photosynthesis, Leaf Anatomy, and Cellular Constituents in the Polyploid C(4) Grass *Panicum virgatum*. Plant Physiology *84*, 461–466.

Weaver, B. a a, and Cleveland, D.W. (2006). Does aneuploidy cause cancer? Current Opinion in Cell Biology *18*, 658–667.

Williams, B.R., Prabhu, V.R., Hunter, K.E., Glazier, C.M., Whittaker, C.A., Housman, D.E., and Amon, A. (2008). Aneuploidy Affects Proliferation and Spontaneous Immortalization in Mammalian Cells. Science *322*, 703–709.

Wolfe, K.H. (2001). Yesterday's polyploids and the mystery of diploidization. Genetics *2*, 333– 341.

Wolfe, K.H., and Shields, D.C. (1997). Molecular evidence for an ancient duplication of the entire yeast genome. Nature *387*, 708–713.

Wu, C.-Y., Rolfe, P.A., Gifford, D.K., and Fink, G.R. (2010). Control of transcription by cell size. PLoS Biology *8*, e1000523.

Yildiz, M. (2013). Plant responses at different ploidy levels. In Current Progress in Biological Research, pp. 364– 385.

Zhurinsky, J., Leonhard, K., Watt, S., Marguerat, S., Bähler, J., and Nurse, P. (2010). A coordinated global control over cellular transcription. Current Biology *20*, 2010–2015.

Zimmet, J., and Ravid, K. (2000). Polyploidy: occurrence in nature, mechanisms, and significance for the megakaryocyte-platelet system. Experimental Hematology *28*, 3–16.

Zybina, T.G., and Zybina, E. V (2005). Cell reproduction and genome multiplication in the proliferative and invasive trophoblast cell populations of mammalian placenta. Cell Biology International *29*, 1071–1083.

Effect of *cbp1* antisense transcription on gene regulation

Effects of *cbp1* antisense transcription on gene regulation

Abstract

Recent large scale studies have uncovered a very large number of non-coding RNAs in eukaryotes that led to an increased interest in exploring their role in the cell. Many studies have suggested that non-coding RNAs are key players in various cellular processes, for example regulation of gene expression, chromatin modification and genome stability. Although pervasive transcription has been shown to be more prevalent than earlier thought, very few studies have established their role in gene expression.

A previous large scale analysis of the fission yeast transcriptome performed in our laboratory uncovered a number of non-coding RNAs transcribed from the antisense strand of protein-coding genes (Wilhelm et al., 2008); one of them is the antisense transcript to *abp1*, also known as *cbp1*. Interestingly, this non-coding antisense transcript gets induced during the heat shock. As *cbp1* is involved in important regulatory roles such as meiotic chromosomal segregation and *Tf2* silencing (Cam et al., 2008), we tried to establish the role of a *cbp1* antisense transcript in gene regulation by expressing it either in *cis* or in *trans*.

Our data indicated that *cbp1* antisense transcript levels increase significantly at 15 minutes of heat shock, along with an increase in *cbp1* sense transcript levels, whereas no effect on Cbp1 levels was detected. We observed a slow growing and temperature resistant phenotype for *cbp1* deletion mutants and tried to rescue these phenotypes by transcribing the *cbp1* sense and antisense in *trans* by putting them individually under the control of an ectopic inducible promoter. Although, our results remained inconclusive due to the phenotype rescue by both sense and antisense transcripts, data suggests that the antisense transcript of *cbp1* might have some role in regulating the *S. pombe* genome.

During the course of PhD studies, I got an opportunity to work on a side project, where I had to analyse the role of antisense transcription in the regulation of *cbp1* gene. This chapter begins with the review of literature providing the basic information about non coding antisense RNA, their prevalence and biological implications across the eukaryotes. Then we discuss the identification of antisense transcript of *cbp1*. Next section contains the materials and methods used to obtain the results mentioned in third section.

Introduction

Tremendous progress has been made in the field of functional genomics since the publication of first draft of human genome sequence which laid the foundation for identification and characterisation of complete human transcriptome in recent years (Venter et al., 2001). Development of new techniques such as tilling arrays and RNA sequencing have totally changed our understanding of genome organisation by revealing that considerably a larger part of the genome is transcribed than previously presumed. For instance, up to 85% of budding yeast, 94% of fission yeast and up to 98% of the human genome is transcribed (Mattick and Makunin, 2005; David et al., 2006; Wilhelm et al., 2008; Derrien et al., 2012), while < 2% transcripts represent protein coding genes (~20,000 genes) in humans (Birney et al., 2007). It shows that the majority of the human genome is actively transcribed into non coding RNAs (ncRNAs). A proportion of ncRNA transcription co-relates with organism complexity (less than 25% for prokaryotes, more than 60% in plants and metazoan and up to 98% in humans) (Mattick, 2004), which means that ncRNAs might be involved in regulating gene expression to add in the complexity and evolution of the species. These ncRNAs include classical "housekeeping" ncRNAs, small ncRNAs, repetitive elements associated ncRNAs and long ncRNAs (lncRNAs). The 'housekeeping' ncRNAs include transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), and small nuclear RNAs (snRNAs) which are required to carry out several cellular processes. The small ncRNAs, for example, microRNAs (miRNAs), small interfering RNA (siRNA) and piwi-interacting RNAs (piRNAs) have been studied extensively in past few years. These small ncRNAs are conserved and can regulate gene expression at transcriptional and post-transcriptional levels by silencing or enhancing the expression of their mRNA targets through complimentary base-

pairing (Ghildiyal and Zamore, 2009). They play an important role in many biological aspects, such as cell fate, tumour progression and development (Stefani and Slack, 2008). Contrary to the short RNAs, lncRNAs, that are >200 nt, have recently appeared as a major class of eukaryotic transcripts. Many recently identified lncRNAs are found to be transcribed by RNA polymerase II, but they do not contain an open reading frame to encode protein products (Mercer et al., 2009). Most lncRNAs are characterised by their nuclear localisation, low expression and low sequence conservation (Mattick, 2011).

They can be derived from intra-, intergenic and repeated regions and are transcribed, either in the sense, or antisense orientation regarding the neighbouring coding sequences. These lncRNA also show variation in their stability, i.e. some are stable and long-lived therefore can be detected in wild type cells, while others are unstable and short-lived as they are quickly degraded by RNA decay method. Although, lncRNAs are found to be involved in regulating transcription, yet for most of them, their function still remains unclear. lncRNAs are found to be lowly expressed (Wilhelm et al., 2008; van Bakel et al., 2010) and less conserved than protein-coding genes. Antisense RNAs constitute one sub group of lncRNA and have gained much attention as important regulators of genes expression in eukaryotes. Here we will discuss this subgroup in detail with a focus on its prevalence and functional importance in yeast.

Antisense RNA

Antisense RNAs are mostly polyadenylated, single stranded transcripts, which are complementary to the protein coding sense strand. Generally, antisense transcripts range from one hundred to several thousand base pairs in length (Beiter et al., 2009; Xu et al., 2009). These can be classified as *cis*- and *trans*-antisense transcripts. The *cis*-antisense transcripts are expressed from the same genomic loci on the opposite DNA strand and are complementary to all or part of its sequence, whereas *trans*-antisense RNAs are transcribed from distinct, non-overlapping loci and show complementarity with a particular region of the sense transcript that they regulate. Since *trans*-antisense RNAs show partial complementarity, they can, therefore, have many more target transcripts, to make complex regulatory networks (Li et al., 2006).

Within a sense/antisense pairs, *cis*-antisense might make extended regions of double stranded RNAs that show perfect complementarity, whereas trans-antisense pairs can result in relatively shorter regions of base pairing, frequently interrupted by the mismatches. Although, definition of sense and antisense is arbitrary, usually, the antisense transcript is assumed to be a regulatory RNA, while sense RNA refers to the protein coding and has a well characterized or more direct role (Yelin et al., 2003; Chen et al., 2004).

Discovery and prevalence of Antisense transcription

The first well characterised antisense RNA was one which regulated the replication of bacterial plasmids. Soon after, antisense RNAs that control expression of endogenous bacterial mRNA were identified (Wagner and Simons, 1994; Storz et al., 2005). Consequently, a better understanding of the role of antisense transcripts in gene regulation in bacteria emerged showing that the *cis*-antisense RNAs in plasmids regulate expression of those genes that are required for replication and stable inheritance of a plasmid (Wagner *et al.* 2002).

Several genome-wide studies illustrated the extent of potential antisense occurrence in different experimental organisms and in human cell lines. In mice, ~ 29% of all transcriptional units (43,553) were detected to overlap with a sense transcripts to the opposite strand (Katayama et al., 2005). In humans, Chen et al. (2004) have reported that ~ 22% of 26,741 transcriptional units form sense–antisense pairs (Chen et al., 2004). Similarly, in a genome-wide tiling array Bertone et al. (2004) found 10,595 novel human transcripts, 1187 of which were antisense to the annotated genes (Bertone et al., 2004). In another genome wide transcriptome study, 2900 to 6400 human genes were found to show antisense transcription, when examined in five different human cell lines (He et al., 2008), while a recent study has reported up to 14,880 lncRNA transcripts, where majority are antisense transcripts (Derrien et al., 2012).

In *Drosophila* genome, 1027 sense–antisense pairs have been identified that constitute ~ 15% of the 13,379 genes (Misra et al., 2002).

Similarly, sense-antisense pairs have also been identified in plants, for example in rice and *Arabidopsis thaliana* ~ 7% of all cDNA clusters contain an antisense partner (Osato et al., 2003, Wang et al., 2005).

Discovery and prevalence of Antisense transcription in Yeast

Fission and budding yeasts are relatively simple model organisms with smaller compact genomes, which permits a detailed study of the transcriptome in different growth conditions, such as meiosis, stress, nutrients deficiency, various carbon sources and different mutant strains (Uhler et al., 2007; Houseley et al., 2008; Wilhelm et al., 2008; Ni et al., 2010; Bitton et al., 2011).

In *S. pombe*, the first antisense RNA identified was the one that overlaps with *spo6*, a meiosis-specific gene (Nakamura et al., 2000). In *spo6* a cell, progression to meiosis is blocked after meiosis I and no mature asci are produced. The *spo6*-antisense RNA overlaps the entire *spo6* gene along with its promoter. Although, this study recognized the presence of an antisense RNA, its potential regulatory role in spo6 expression still needs to be explored.

Tiling array studies of *S. pombe* identified a large number of ncRNAs (>2000) in vegetative cells that were grown in a rich medium (Dutrow et al., 2008; Wilhelm et al., 2008), while 47.4% of all protein coding genes in *S. pombe* have been shown to have antisense transcripts under normal or stressed condition, such as, heat shock (Ni et al., 2010). Among these genes, 5.9% (302 genes) show higher or equal antisense expression than the sense. Different studies have shown that antisense transcripts are enriched for genes whose expression is increased during meiosis (Ni et al., 2010; Bitton et al., 2011; Rhind et al., 2011). In *S. cerevisiae* a large number of antisense transcripts have been identified, which like *S. pombe*, are enriched for meiotic genes (David et al., 2006; Granovskaia et al., 2010), showing that antisense expression is condition specific. When analysed under different growth conditions, such as, heat shock, it was found that the sense-antisense transcripts are individually regulated, which indicates that the expression of antisense transcripts is not simply a result of transcriptional noise (Ni et al., 2010).

The Regulator of meiosis (RME2) is the first antisense RNA in budding yeast with a well-established function. This AS transcript overlaps with the meiotic regulator (IME4), that is vital for the induction of early meiotic genes (Shah and Clancy, 1992; Hongay et al., 2006). Similar to the *spo6* antisense transcript, RME2 spans the coding region and promotor of IME4. Transcription of IME4 is regulated by cell type and RME2 contributes directly to that regulation. In MATa or MATα cells, RME2 is actively transcribed and inhibits IME4 expression.

These examples reveal that antisense RNAs exist in both *S. pombe* and *S. cerevisiae* and may have some regulatory functions. Recent genome-wide studies involving tilling arrays and strand specific RNA sequencing in both, fission and budding yeasts, show that although they are distantly related, antisense transcription is a common feature of both (Dutrow et al., 2008; Wilhelm et al., 2008; Granovskaia et al., 2010; Ni et al., 2010; Bitton et al., 2011; Rhind et al., 2011).

Evolutionary conservation of Antisense transcription in Yeast species

A comparative study of the transcriptome in fission yeast clade, comprising *S. pombe*, *S. octosporus*, *S. cryophilus* and *S. japonicus*, revealed that 328 antisense transcripts (51% of the total) are conserved across two or more species, suggesting their biological importance (Rhind et al., 2011). Across the fission yeast clade ~ 250 genes that show a higher antisense expression than sense, in vegetative cells, are considerably enriched in meiotic genes, consistent with other studies in *S. pombe* and *S. cerevisiae* (Granovskaia et al., 2010; Ni et al., 2010; Bitton et al., 2011). 60% of the specific sense/antisense pairs seen for *S. pombe* meiotic genes are conserved in at least one other species. Such as, the antisense RNA that suppresses the basal transcription of *spo4* gene is conserved in all fission yeasts (Rhind et al., 2011). Similarly, Yassour et al., 2010 has also shown that several antisense transcripts and in some instances, their regulation, are conserved across five different yeast species. Conservation of the antisense transcripts in different yeast species suggests their biological significance in gene regulation.

Mechanisms of Antisense expression:

Sense and antisense transcripts are found to be independently regulated, which shows that antisense transcription is autonomously regulated, by either with their own promoters or through other mechanisms (Ni et al., 2010). Based on their transcriptional initiation sites, antisense transcripts can be divided into three categories: (1) those antisense transcripts that have their own novel transcription start sites, and are expressed as a result of active transcription by RNA Polymerase II at these loci; (2) antisense transcripts whose transcription initiates due to bidirectional promoter of a neighbouring gene. When two genes are in tandem orientation, a bidirectional promoter can derive the transcription of an upstream antisense along with the downstream sense gene. Recent findings strongly support that antisense transcription in budding yeast, mostly results from bidirectional promoters (Xu et al., 2009); for example, pol II dependent cryptic unstable transcripts (CUTs). A similar mechanism is also found to exist in fission yeast, for instance, at the *Guf1/Wis2* locus (Ni et al., 2010). The level of *Guf1* antisense transcripts increases during the heat shock and this increase is synchronized with *Wis2* sense transcription, where *Guf1/Wis2* are oriented in tandem; (3) those antisense transcripts that are created due to transcriptional read-through. In this scenario, two genes are present in convergent orientation. The extended 3′ un-translated region (UTR) of one gene becomes the antisense transcript of other gene, transcribed from the opposite strand. Such as, GRIP and LEA domain- encoding genes are arranged in convergent orientation in *S. pombe* and a read through of the LEA domain encoding gene also gives rise to an antisense transcript for GRIP domain encoding sense transcript (Ni et al., 2010). It is important to consider that these extended 3′ UTRs are unusual, as the average length of 3′ UTRs in *S. pombe* is 170-280 nucleotides, which suggests that their existence is not accidental (Ni et al., 2010).

Post transcriptional regulation of Antisense RNAs:

Transcript levels for ncRNAs including antisense RNAs are maintained at a very low level in cells under normal physiological conditions, as these transcripts are rapidly degraded to keep their levels invisible unless RNA degradation machinery is inactivated. Recent studies involving RNA sequencing of mutants that are defective for RNA processing pathways or chromatin maintenance in yeast have revealed

novel classes of ncRNAs, thus confirming the existence of hidden transcription (Camblong et al., 2009; Xu et al., 2009; Lardenois et al., 2011; van Dijk et al., 2011). For example the transcriptome of *S. cerevisiae rrp6* mutants (defective for the nuclear exosome complex required for the RNA degradation) has revealed a class of ncRNAs referred as Cryptic Unstable Transcripts (CUTs) (Xu et al., 2009). CUTs are transcribed by RNA Pol II, are capped, comparatively small (~ 200 to 500 bp) and with heterogeneous 3' end. Large numbers of these CUTs arises from nucleosome-free regions (NFRs) and are transcribed divergently from the promoter region of annotated genes. Their transcription is controlled by the Nrd1 dependent termination pathway, which initiates an early termination and degradation of these transcripts by targeting them for polyadenylation and degradation by the TRAMP complex and the nuclear exosomes, respectively (Arigo et al., 2006; Thiebaut et al., 2006). Under certain physiological conditions, the unstable nature of CUTs is affected such that they become stable and controls the expression of many other genes. For instance, loss of Rrp6 protein in budding yeast results in the stabilisation of an antisense CUT transcript to the *PHO84* gene, and subsequent repression of *PHO84* transcription. Similarly, during the chronological aging, the Rrp6 protein exhibit a weaker association with the *PHO84* locus and results in raised levels for antisense transcript to *PHO84* (Camblong et al., 2009). Another interesting example for the regulation of antisense ncRNAs is provided by Lardenois and colleagues, where they found a class of ncRNAs specifically up regulated during meiosis, therefore, named as Meiotic Un-annotated Transcripts (MUTs) (Lardenois et al., 2011). These MUTs are normally degraded by Rrp6 during mitosis, but as cells progress towards meiosis, Rrp6 itself is degraded, hence leading to an accumulation of higher transcript levels of MUTs during meiosis and sporulation, where they regulate transcription of other genes, such as antisense transcript *CLN2* regulating the expression of *IME1* (Purnapatre et al., 2002). All these findings suggest that Rrp6 plays an important role in differential regulation of CUTs and MUTs, so that they are expressed only under certain physiological conditions, when they are required to perform their regulatory role in the cell.

Recently, a new class of ncRNA has been described, which consists of antisense transcripts (50% of 1658 transcripts) which are rapidly degraded by the major cytoplasmic 5' - 3' exoribonuclease Xrn1p (Van Dijk et al., 2011). These transcripts

become stable upon mutation of the *xrn1* and are known as Xrn1 Unstable Transcripts (XUTs). XUTs are found to accumulate in a media containing lithium showing their possible role in the cells adaptability to a changing environment.

Antisense transcript levels are also found to be correlated with histone modification that is indicative of an open chromatin state. For example, Churchman and Weissman have shown that knock out of an essential Rpd3S deacetylation complex increases the level of antisense transcription by increasing the H3 and H4 acetylation (Churchman and Weissman, 2011). Antisense transcription is also regulated at post transcriptional level in fission yeast, where a widely conserved histone variant H2A.Z, which localizes at 5' end of genes, facilitates the suppression of antisense transcription. It has been suggested that H2A.Z might interact with RNA Pol II associated exosome complex to degrade antisense transcripts originating from convergent loci under normal growth conditions (Ni et al., 2010).

Such transcriptome studies show that lncRNAs are actively regulated at transcriptional and post-transcriptional levels, either by RNA surveillance or by modification of the chromatin structure.

Antisense transcription in the control of gene expression

The role that lncRNAs play in regulation of gene expression is now widely recognised in many species. It means that control of gene expression cannot be assigned only to those proteins that are required for transcription, post transcriptional modification and translation, but to lncRNAs as well. Antisense transcripts can undergo hybridization with their sense counterpart. Minimum overlap requirement for the sense and antisense transcript is 20 base pairs (Werner and Berdal, 2005). These antisense transcripts can either up-regulate or down-regulate expression of genes in a wide variety of species ranging from prokaryotes to higher eukaryotes and are required in mammals to play a significant role in many physiological and pathological processes, such as alternative splicing, genomic imprinting, translational regulation, X inactivation, mRNA stability, DNA methylation, RNA export and histone modifications. Natural antisense transcripts (NATs) also play a significant role in the regulation of developmental processes, adaptations to various

stresses and response to viral infections (Lavorgna et al., 2004; Lapidot and Pilpel, 2006; Werner and Sayer, 2009).

Antisense transcripts can affect gene expressions either by interacting directly with the sense transcripts on the same locus or through interactions with other targets that might be required for mRNA expression, editing, transport or translation (Li et al., 2006). Regulatory functions of antisense transcripts involving the formation of double-stranded RNA (dsRNAs) are reported to be involved in RNA editing (Bass, 2002), RNA interference (Bass, 2002), RNA masking (Beltran et al., 2008) and transcriptional interference (Prescott and Proudfoot, 2002).

Some Antisense ncRNAs and sense transcripts derived from the same locus might compete for binding to the same transcription factors which shows that antisense and their corresponding sense transcripts could be expressed at same or different timings depending on the pattern of bindings to the transcription factors (Cawley et al., 2004). Antisense ncRNAs have also been found to regulate gene expression at chromatin level by histone modifications (Feinberg and Tycko, 2004).

Antisense transcription in regulation of gene expression in Yeast

The extent of antisense transcription in both, fission and budding yeast, indicate that these ncRNAs play essential tasks by regulating the gene expression under different physiological conditions. In some cases, expression of the antisense transcript supresses the expression of the sense mRNA under non-activating environmental conditions, which renders the gene in an off-state (Xu et al., 2011). In addition, a condition dependent suppression of a gene by an antisense ncRNA has been described in a number of cases. Sense and antisense can also act through different mechanisms such as, in *cis* or in *trans*, while many of them are shown to be regulated by physiological conditions (for example IME4, ZIP2, GAL1-GAL10, PHO84) (Hongay et al., 2006; Houseley et al., 2008; Camblong et al., 2009; Gelfand et al., 2011) in budding yeast.

In fission yeast (Ehrensberger et al., 2013) have reported the regulation of alcohol dehydrogenase1 (adh1) gene, that encodes zinc binding protein and is required for

the conversion of acetaldehyde to ethanol, by an antisense transcript. Under zinc limited conditions this antisense exhibits higher transcript levels and results in the repression of adh1 mRNA levels. Antisense transcript levels also arise in zinc limited cells when expressed in *trans* from a constitutive promoter. Although, the exact mechanism by which antisense transcript regulate expression of *adh1* mRNA remains unclear, there is clear evidence for the reciprocal relation for antisense-sense pair of *adh1* under zinc limited conditions (Ehrensberger et al., 2013).

Previously, it was thought that entry to the meiosis in *S. cerevisiae* is under the control of proteins, such as IME4 and ZIP2. It has been shown now that IME4 and ZIP2 expression and entry into the meiosis in diploids is regulated by the two antisense transcripts RME2 and RME3 respectively. Expression of the antisense is cell type specific as in haploids expression of the antisense RNAs exceeds, whereas in MAT a/α diploids expression of IME4 and ZIP2sense RNA is increased. In these diploids a1/α2 heterodimer supress the transcription of antisense to allow the expression of IME4 and ZIP2 sense expression to enter into meiosis. These transcripts inhibit their sense partners only in a *cis via* transcriptional interference, where increased levels of transcripts from the strand with competent promoter results in the suppression of transcription from opposite strand. Strength of these promoters is cell-type-specific, where RME2 and RME3 antisense transcription determines a cell-type capable of meiotic entry (Hongay et al., 2006; Gelfand et al., 2011).

Although transcriptional interference appears to be the emphasized mechanism of antisense regulation in *S. cerevisiae*, a few other studies have suggested that epigenetic and chromatin modulation mechanisms could also be involved.

Regulation of *GAL1-10* genes cluster is also a result of *cis* regulation by an antisense RNA modification. The GAL cluster is required for galactose metabolism and has three characterised states depending on the availability of carbon source in the media; induced (in the presence of galactose), non-induced (in the presence of raffinose) and repressed (in the presence of glucose). In the repressed state a regulatory protein Reb1 binds to the 3' of *GAL10* ORF and stimulates the transcription of antisense RNA. This antisense RNA results in full suppression of the *GAL* locus by regulating the chromatin remodelling enzymes, e.g., the histone

methyltransferase Set1p and the histone deacetylase complex Rpd3S (Houseley et al., 2008).

Another well studied example of sense-antisense pairing in yeast is provided by PHO84 gene, where the loss of the nuclear exosome component Rrp6 results in the stabilization of two antisense transcripts for PHO84. These antisense RNAs repress PHO84 transcription not only in the *cis* but also in the *trans* manner, by recruiting the histone deacetylases Hda1, which deacetylates the histones at the PHO84 promoter. In addition, PHO84 epression can also be suppressed in *trans* by transcribing the antisense RNA from a plasmid, which does not require Hda1 (Camblong et al., 2009).

Transcriptional regulation by antisense RNA in the *trans* manner is provided by a XUT antisense transcript for *TY1* retrotransposon in *S. cerevisiae*, that starts within the *TY1* retrotransposon and encompasses its 5' LTR. *TY1* antisense RNA is synthesized by RNA Pol II and is polyadenylated and destabilised by the cytoplasmic exosome Xrn1 under normal cellular conditions. On inactivation of Xrn1 these *TY1* antisense RNAs become stable and act in *trans* to repress the transcription and mobility of *TY1* retrotransposons in a mechanism of gene silencing that is dependent on methyltransferase Set1p and histone deacetylase (Berretta et al., 2008).

Another example of regulation of sense RNA by an antisense RNA is provided by *PHO5* that codes for an acid phosphatase and controlled by phosphate availability. Under increased phosphate conditions, four nucleosomes are positioned at the *PHO5* promoter region. Under phosphate starvation, an ncRNA, approximately 2.4-kb long, antisense to *PHO5* is transcribed and plays a role in nucleosome eviction at the promoter. This results in transcriptional activation of the gene (Uhler et al., 2007). Therefore in this case the antisense RNA positively regulates the transcription, whereas in previously mentioned examples, ncRNAs negatively regulate the expression of their respective sense RNAs.

Antisense RNA's implication in disease

An example of sense transcript regulation in disease by antisense transcripts is the cell death-inducing DNA fragmentation factor subunit alpha (DFEA), like effector B

(CIDEB), which is a likely suppressor of tumours. Lavorgna et al. analyzed CIDEB sense and antisense expressed sequence tags (ETS) in different tissues. They noted that, in comparison with normal tissues, CIDEB in cancer exhibits a higher proportion of antisense ESTs (Lavorgna et al., 2004; Li et al., 2006). P15 presents another well-studied tumour suppressor gene in many tumours with raised levels of antisense transcripts. Unlike normal lymphocytes, leukemic lymphocytes have a higher amount of the antisense transcripts of P15 but a lower fraction of P15 sense mRNA (Yu et al., 2009). These findings signify the possibility of a likely correlation between the up regulation of antisense transcripts and the abnormal silence of tumour suppressor genes. Other genes may also be affected by antisense transcripts in a similar way. For example, Natriuretic Peptide Precursor A (NPPA), involved in the regulation of blood pressure is down regulated by its antisense transcripts, resulting in high blood pressure (Annilo et al., 2009).

The β-Secretase-1 (BACE1) enzyme has been found to be associated with Alzeimer`s disease. When cells are subjected to various stresses, expression of BACE1 is up regulated by BACE1 –antisense transcript at both the mRNA and the protein levels through a post-transcriptional concordant mechanism (Faghihi et al., 2008). Likewise, Wrap53, an antisense transcript of p53, increases both the mRNA and protein level of endogenous p53 by targeting the 5′ UTR region (Mahmoudi et al., 2009). Keeping in view their role in different diseases, antisense transcripts are considered as potential drug targets in the therapeutic treatment of various diseases.

Although, much is known about the widespread prevalence and role of antisense transcription in regulation of gene expression, exact mechanisms by which these antisense ncRNAs derive the gene regulation need to be further explored.

Antisense transcripts at fission yeast *cbp1* locus

In a large scale analysis of the fission yeast transcriptome, by using RNA-seq supplemented with tilling array data, our laboratory has uncovered a number of ncRNAs transcribed from an antisense strand of protein-coding genes. These antisense transcripts are detected in actively growing cells, either in rich media or in MM, under different experimental conditions such as meiosis, environmental stress or RNA processing mutants. One of them is antisense transcript to *abp1*, also known

as *cbp1* (Figure 1). Interestingly, this antisense transcript (SPNCRNA. 574) to the *cbp1* gene gets induced during heat shock (Wilhelm et al., 2008). For ease of use, it has been mentioned as *cbp1* antisense ncRNA.

The *cbp1* gene and closely related *cbh1* and *cbh2* are homologues of the human CENP-B protein. The *cbp1* is ~25% identical and ~50% similar to CENP-B. It binds to centromeric DNA with a high affinity and is an abundant nuclear protein (Halverson et al., 1997). Over expression of *cbp1* causes an increase in minichromosome loss and deletion of *cbp1* results in slow growth as well as sporulation defects. It is required for normal mitotic and meiotic chromosomal segregation and is also involved in histone tail modification. Cbp1 binds to the *Tf2* retrotransposons and their remnants and mediates *Tf2* silencing by recruiting the class I histone deacetylase (HDAC) Clr6 and the class II HDAC Clr3 (Cam et al., 2008). CENP-Bs also suppress the expression of many other genes through neighbouring LTRs and enable the HDAC recruitment to the heterochromatic loci and silencing of retrotransposons by incorporating histone deacetylases (Sugiyama et al., 2007).

As *cbp1* is involved in important regulatory roles such as meiotic chromosomal segregation and *Tf2* silencing (Cam et al., 2008), we therefore tried to characterise effects of up-regulation of *cbp1* antisense transcript on gene regulation by expressing it either in *cis* or in *trans*.

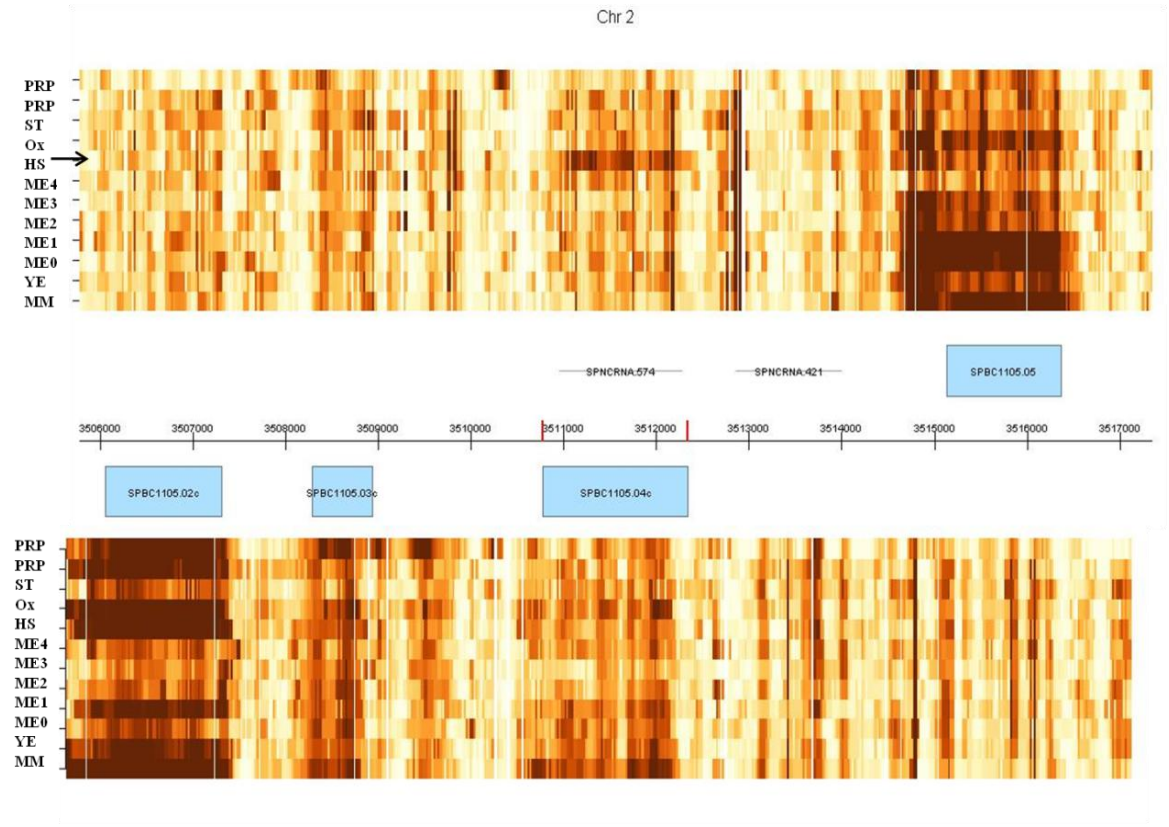


Figure 1: Tiling-chip hybridization signals showing expression of *cbp1* under different experimental conditions. Strength of colour represents the expression levels of transcripts (dark colour corresponds to higher expression, while light colour represents weaker expression), across the genomic region (shown in the centre) for forward (top panel) and reverse (bottom panel) strands. Each row reflects an experimental condition (as labelled on left side of each panel). Rapid proliferation was monitored in rich (YE) and minimal (MM) media (Wilhelm et al., 2008).

Materials and Methods

All strains were maintained in solid or liquid Yeast Extract plus Supplements (YES) or the minimal media (MM). All cultures were grown at 32°C. For long term storage stains were stored at -70 °C.

RNA extraction from cells

50 ml of 0.5 OD cells were harvested by centrifugation for 3 minutes at 3000rpm after which the supernatant was discarded and the cells were stored at -70°C. Cells were then thawed on ice and washed with 1 ml DEPC treated water. 750 µl TES (see appendix) and 750 µl acidic phenol-chloroform (Sigma) were added to each sample. The mixture was vortexed for 10 seconds and incubated at 65°C for 1 hour, vortexing every 10 minutes. Samples were then placed on ice for 1 minute, mixed, and centrifuged (15 minutes at 14000 rpm at 4°C). 700µl upper aqueous phase from each sample were then transferred to Maxtract high density phase lock tubes (Qiagen) together with an equal volume of acidic phenol-chloroform (Sigma) and were centrifuged at 14000rpm for 5 minutes at 4°C. The upper phase from these samples was transferred to phase lock tubes together with 700µl chloroform:isoamyl alcohol (24:1 Sigma) and spun again. RNA was precipitated from the upper phase with 1.5ml 100% ethanol and 50µl 3M NaAc (pH 5.2) overnight at -20°C or for 30 minutes at -70°C. The sample RNA was then collected by centrifugation (15 minutes 14 000 rpm at 4°C) and washed once with 70% ethanol. The sample RNA was dissolved in 200µl DEPC treated water by incubation at 65°C for 1 minute or 10 minutes at room temperature.

To assess RNA integrity 500ng of the sample was loaded on 2% agarose gel together with DNA marker Hyperladder™ IV (Bioline) at 40V in TBE buffer. RNA samples were then purified using the RNeasy Mini Kit (Qiagen) after manufacturer instructions.

Strand-specific reverse transcription and PCR

RNA samples were digested with 2U DNaseI (TURBO DNase™, Ambion) for 25 minutes at 37ºC. Reaction was terminated by incubation with 10 µl inactivation mix (TURBO DNA-free Kit, Ambion) at room temperature for 2 minutes. 1µg of this

RNA was then used to set up the reverse transcription mixture together with 3.33μM of either the forward and reverse primer, 1x First Strand buffer, 0.01M DTT, 40U RNAseOUT (RNAse inhibitor), 1.66mM dNTP. This mixture was incubated at 42ºC for 2 hours. The RNA was then degraded by incubation with 33mM NaOH at 70ºC for 15 minutes, and then 33mM HCl was added to neutralise the sample pH. The RT product was then extracted using MiniElute PCR purification Kit (Qiagen) after manufacturer instructions. For PCR analysis the BIOTAQ™ DNA polymerase system (Bioline) was used after manufacturer instructions. The reaction mixture consist of: 1x NH_4 Reaction buffer, 0.3mM dNTPs, 1.5mM $MgCl_2$, 0.2mM of each primer, 1U BioTaq stock, 25ng cDNA, 6μg/ml actinomycin D. Water was used as a negative control for cDNA whilst genomic DNA acted a positive control. The reaction was run on PCR program (95 ºC for 2 min, 94 ºC for 30 sec, 55 ºC for 30 sec,72 ºC for 30 sec, 72 ºC for 5 min) on PTC-225 thermal cycler (MJ Research), and the product was assessed on 1.5% agarose gel together with DNA marker Hyperladder™ IV (Bioline).

Gel electrophoresis

Samples were run on 1- 2% (w/v) agarose/ ethidium bromide (0.4 μg/ml)/Tris base, boric acid and EDTA (TBE) gel at 100V for 30 minutes in TBE buffer along with DNA marker using EM100 Mini Submarine Gel Unit (Cambridge Electrophoresis Ltd) or Sub cell GT system (Bio-rad). UV visualisation of the DNA was then performed using UVIdoc gel documentation system (Progen Scientific).

Real time -PCR

Reverse transcription was performed on RNA samples using the same method given in strand-specific RT-PCR. Real-time PCR was then performed in duplication in the 7500 Fast Real-Time PCR System (Applied Biosystems). The 20 μl reaction mixtures contained 10 ng cDNA, 0.2 μM of each of the forward and reverse primers, and 10 μl Fast SYBR Green PCR master mix (Applied Biosystems). An external standard curve was also set up using a suitable range of genomic DNA dilutions. The amplification program was as follows: 1 cycle of 95ºC for 20 seconds, 40 cycles of 95ºC for 3 seconds, 60ºC for 20 seconds.

Genomic DNA extraction

50 ml of OD 0.5 cells were harvested by centrifugation at 3000 rpm for 3 minutes. DNA extraction was performed using ZR fungal DNA isolation kit after manufacturer instructions. Quality of DNA was checked by loading 500 ng on 1% agarose gel.

Colony PCR

The BIOTAQ™ DNA Polymerase system (Bioline) was used after manufacturer instructions; each colony were boiled at 100ºC with 1X NH_4 Reaction buffer, 1.5mM $MgCl_2$, 1 µM of each forward and reverse primer for 10 minutes, and then cooled at 4°C for 10 minutes. 0.2 mM dNTP and 2.5U Taq polymerase were then added to the mix before running on the colony PCR program (see appendix). The samples were then analysed on 1% agarose gel along with DNA marker hyperladder I (Bioline).

Protein Extraction

50 ml of 0.5 OD cells were harvested at 3000 rpm for 5 minutes. Washed once in 5 ml of ice-cold stop buffer (150 mM NaCl, 50 mM NaF, 10 mM EDTA, 1 mM NaN3 pH 8), and transfer to 15 ml falcon tube. Spin down at 5000 rpm for 5 minutes. Drained the pellet well and resuspended in 200 µl of RIPA buffer (10 mM sodium phosphate pH 7, 1% Triton X-100, 0.1% SDS, 2 mM EDTA, 150 mM NaCl, 50 mM NaF, 0.1 mM sodium vanadate, 4 µg/ml leupeptin, 1mM PMSF). After adding 1.5 ml of acid washed glass beads (0.5mm diameter, Sigma, G-9268) cells were vortexed for 1 minute and supernatant was collected in a separate eppendorf tube.

Western Blotting

Sample preparation for western blotting

30-40 µg of protein from cell samples was mixed thoroughly with equal volume of 2x loading buffer containing 0.125 M Tris-HCl buffer with pH 6.8, 4% SDS, 20% (v/v) glycerol, 0.2 M dithiothreitol DTT, and 0.02% bromophenol blue with pH adjusted to 6.8. The mixture was heated at 95^{o}C for 5 minutes then placed on ice before loading into the gel.

Gel loading and transfer to the blot

Samples were loaded on the 10% polyacrylamide resolving gel (Sigma, UK) along with protein marker (Biorad, USA) carefully in each well. The upper and lower chambers of the electrophoresis tank were filled with running buffer and the gel was run for approximately 90 minutes at 150 mA using power pac (Biorad). Meanwhile polyvinylidine difluoride (PVDF) (Biorad, USA) was soaked in 100% methanol for 10 minutes and after washing with phosphate buffer saline PBS-0.01% Tween the membrane and the filter paper was soaked with transfer buffer for 15 minutes.

The transfer sandwich was assembled to the semi dry unit (Biorad, USA) in the following order: filter paper/ PVDF membrane/ gel/ filter paper. Air bubbles were removed by rolling a clean test tube over the layer and the unit was set at 180 mA for 120 minutes. It was then blocked with 20 ml of 5% non fat-dry milk in PBS-Tween for 60minutes to block non-specific protein binding sites on the blot. After blocking, the immobilized proteins were probed overnight at 4^{o}C with specific primary rabbit anti-HA (1:500), mouse anti-β-actin (1:10,000, Abcam) followed by washing and incubation with secondary antibody horseradish peroxidase (1:4000). The blot was visualised under Fujifilm LAS-1000 imager.

Cloning

Construct preparation

Expand Long Template PCR system (Roche) was used after manufacturer instruction to prepare *cbp1* construct for cloning. The reaction mixture contained 1 x buffer, 0.5 mM dNTPs, 20 µM of each primer, 3U DNA polymerase mix and 20 ng of template genomic DNA. Purification of PCR product was performed with MiniElute PCR Purification Kit (Qiagen) after manufacturer instruction and assessed on 1 % agarose gel along with DNA marker Hperladder I (Bioline).

Digestion with Restriction enzymes

Construct and plasmid DNA were digested in two separate reactions with *Sal1* and *Sma1* restriction enzymes in a total of 50 µl reaction each. 500 ng of construct and plasmid DNA were incubated at 37 °C with 5 U of *Sma1*, 5 µl of 10 x buffer 3

(NEB) for 3 hours. All samples were then purified using MiniElute PCR Purification Kit (Qiagen) after manufacturer instruction,

Ligation

For ligation 80 ng of construct DNA and 20 ng of plasmid DNA were incubated over night in a total of 20 µl reaction containing 10 U/µl T4 DNA ligase and 2 µl 10 x ligation buffer (Promega).

***E. coli* transformation**

E. coli competent cells were transformed with plasmid (with and without *cbp1*). Transformed *E. coli* cells were selected on ampicillin containing LB plates separately. Single colony from each plate was picked to inoculate the starter culture of 1 ml LB medium containing ampicillin (50µg/ml). After growing them overnight at 37 ˚C with vigorous shaking inoculated 100ml LB medium and grew them at same conditions. Plasmids DNA were extracted and purified by using QIAfilter plasmid Mini Protocol (QIAGEN).

***S. pombe* transformation:**

20 OD of cells were harvested by centrifugation (13000 rpm for 4 minutes) and washed once with distilled water before suspension in 1ml of water. Cells were then washed in 1ml Lithium Acetate - Tris and EDTA (LiAc-TE, see Appendix) before resuspending cells in 55 µl LiAc/TE. After incubation with 10 µl transformed plasmid for 10 minutes at room temperature, cells were further incubated with 260µl of 40% (w/v) polyethylene glycol (PEG 4000)/LiAc-TE for 30-60 minutes at 32°C. 43 µl pre-warmed dimethyl sulfoxide (DMSO, Sigma) were added to the cells, after which the samples were heat shocked at 42°C for 5 minutes. The cells were washed with 1 ml water, resuspended in 500 µl water for plating in duplicate on to EMM plates.

RESULTS

cbp1 antisense transcription

In a tilling array study of *S. pombe* transcriptome many antisense ncRNAs were identified (Wilhelm et al., 2008). Some of these show specific expression patterns under different growth conditions, including rapid proliferation, meiotic differentiation and environmental stress, as well as in RNA processing mutants. Among them, antisense transcript of *cbp1* was of particular interest as it showed higher expression levels in vegetatively growing wild type cells when exposed to heat shock. As *cbp1* is involved in important regulatory roles such as meiotic chromosomal segregation and *Tf2* silencing (Cam et al., 2008), we therefore tried to characterise effects of up-regulation of *cbp1* antisense transcript on gene regulation, when expressed either in *cis* or in *trans*. To characterise the *cbp1* antisense transcripts, we used fission yeast strain JB 747, which has a *HA* tag at 3' end of the *cbp1* gene to characterise the protein levels of Cbp1, and a wild type (*wt)* strain JB 22 as a control.

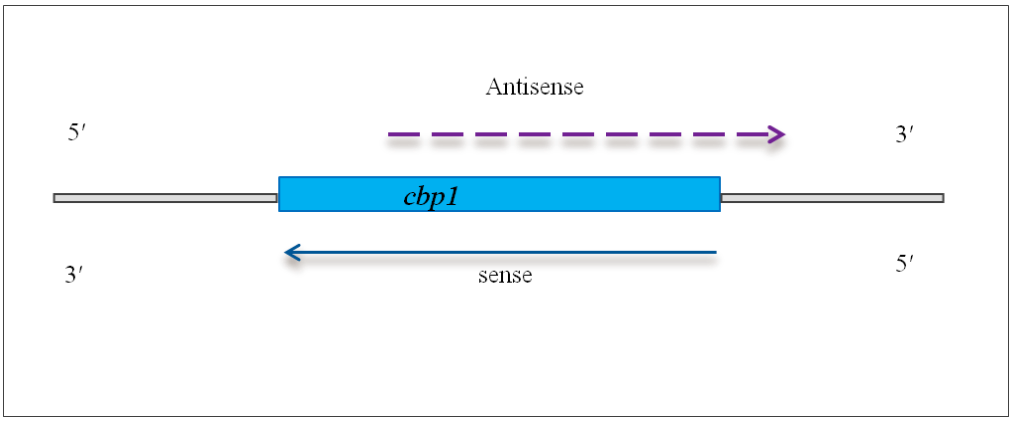


Figure 2: Schematic representation of the sense and antisense transcripts of *cbp1*. This Schematic representation is drawn based on RACE analysis previously done in our lab to map the ends of *cbp1* antisense transcript in (*wt)* strain. Sense transcript is 1569bp long while, antisense transcript is ~ 1350 bp. Antisense ncRNA starts at ~300 bp up-stream of 3' and extends beyond the 5′ of *cbp1* ORF (~ 100bp).

As we used *HA* tagged JB 747 as an experimental strain in this study, there was a possibilty that this insertion of *HA* tag at the 3' end of *cbp1* ORF, might have disrupted the 5' promoter sequences of *cbp1* antisense. This interruption can cause (i) inactivation of antisense transcription (ii) change in a length of antisense transcript. We therefore, first tried to determine, if cbp1 antisense is still actively transcribed after the insertion of *HA* tag, or it has inactivated the promoter sequence for *cbp1* antisense. To test this, we observed *cbp1* antisense transcription in exponentially growing cells, both in *HA* tagged and in *wt* type strains, before and after the heat shock. It was achieved by RNA extraction and subsequent strand specific reverse transcription with a polymerase chain reaction (RT-PCR). In order to see the transcript levels of both the *cbp1* sense and *cbp1* antisense before and after the heat shock, we grew cultures at 25˚C till the mid log phase and collected 50 ml of cells after subjecting them to a heat shock at 39˚C for 15 and 60 minutes.

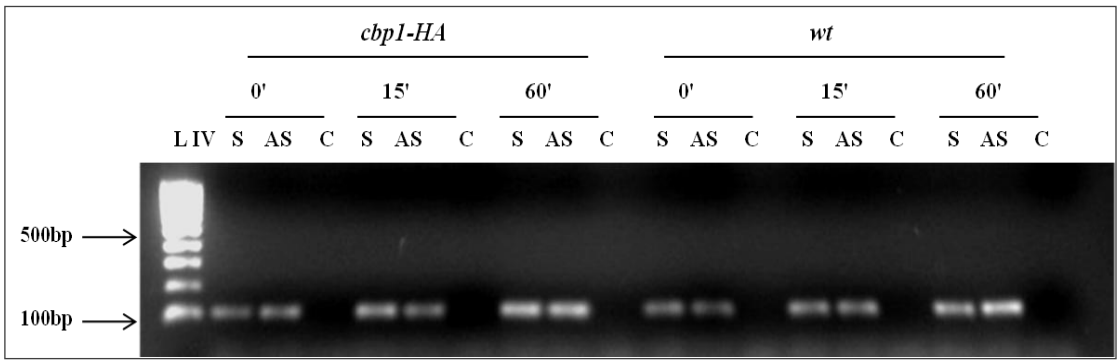


Figure 3: RT-PCR analysis for *cbp1-HA* and *wt* strains. At indicated time points RNA was extracted from 50 ml of exponentially growing cells, after subjecting them to heat shock at 39˚C, to perform RT-PCR. Results clearly show the presence of both the *cbp1* sense (labelled as S) and *cbp1* antisense (labelled as AS) transcripts at alltime points before and after heat shock. The size of both transcripts is 102 bp. Absence of any bands in –RT (C) controls confirms that there is no DNA contamination.

To identify the DNA contamination, if any, during the reverse transcription and PCR procedures, we used some samples without reverse transcriptase and without templates, as negative controls. Amplification products were then resolved on 2% agarose gel, which confirmed the existence of *cbp1* antisense (Figure 3), along with *cbp1* sense before and after the heat shock. This suggests that the promoter sequence for antisense has not been disrupted after the insertion of a HA tag.

To find out, if the boundaries of *cbp1* antisense transcript are still intact after the insertion of the *HA*-tag, we performed the Rapid Amplification of Chromosomal Ends (RACE) for 3' and 5' of sense and antisense transcripts in both strains. As, RACE had been performed previously in our lab to map the location of antisense transcript, we used that information to find, if *cbp1* sense and antisense transcripts are still of the same size.

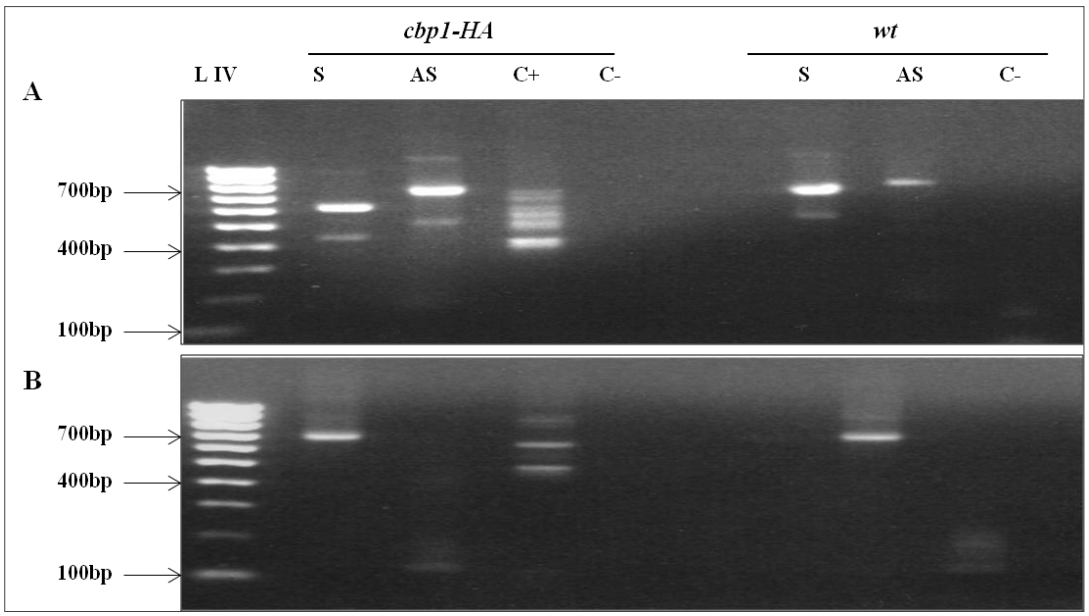


Figure 4: RACE analysis for *cbp1-HA* and wt strain. 3′ RACE analysis (A) and 5′ RACE analysis (B) was performed after synthesizing cDNA for *cbp1-HA* and *wt.* Results clearly show the presence of both the *cbp1* sense (labelled as S) and *cbp1* antisense (labelled as AS) transcripts. Absence of the band from negative control represents the good quality of RACE.

The expected band sizes for 3′ RACE for *cbp1* sense and *cbp1* antisense were 372 bp and 700 bp respectively (Figure 4A), while, we obtained bands of sizes ~ 600-620 bp and ~ 750 bp for sense and antisense respectively. Similarly the expected band sizes for 5' RACE were ~600 bp and 60 bp for sense and antisense respectively and we obtained bands of ~ 700 bp and ~ 120bp (Figure 4B). This observed increase in band size might be due to the presence of RACE adaptors and/or UTRs. These results suggest that the promoter of *cbp1* antisense is intact and HA-tag has not disrupted it.

Effect of *cis* antisense transcription on *cbp1* transcription and translation

As, antisense transcripts can regulate the expression of genes with which they share the locus, in a *cis* manner, by either inducing or repressing their transcription or translation (Hongay et al., 2006; Gelfand et al., 2011), we investigated whether an increase in *cbp1* antisense transcription can change transcriptional and translational response of *cbp1* itself.

Effect of antisense transcription on *cbp1* sense mRNA levels

To assess the influence of antisense expression on sense transcript levels, we tried to observe mRNA levels for both *cbp1* sense and antisense transcripts. We used strand specific quantitative real time PCR (qRT-PCR) to quantify the expression of *cbp1* sense and antisense transcript levels in both strains, before and after heat shock. Here, we used *aap1* mRNA as an internal control for qRT-PCR, as its expression does not change during heat shock in fission yeast.

A qRT-PCR assay showed that both sense and antisense transcripts of *cbp1* were differentially expressed, with induction of antisense transcripts accompanied by an over-expression of sense transcripts or vice versa, both in *HA* tagged and *wt* strains (Figure 5). Levels of antisense transcript were found to be significantly increased after 15 minutes of heat shock, along with a significant increase in *cbp1* sense transcript levels. We further observed that antisense ncRNA levels decreased at 60 minutes of heat shock, but were still significantly higher as compared to the transcript levels before heat shock. Similar, patterns of expression were observed for sense transcript levels after 60 minute of heat shock. P-values were calculated to

check the significance of results by applying student t-test for three independent biological repeats. These findings showed that, when transcribed in *cis*, an increase in antisense transcription is accompanied by an increase in the expression of *cbp1* sense transcripts during heat shock. Interestingly, expression of *cbp1* antisense transcripts was much higher in a *HA*-tagged strain when compared to *wt*, both at 15 and 60 minutes of heat shock. This effect could be due to an interruption of 3' UTR by *HA* tag, which might have an effect in the over expression of the 5' regulatory region of the antisense transcript.

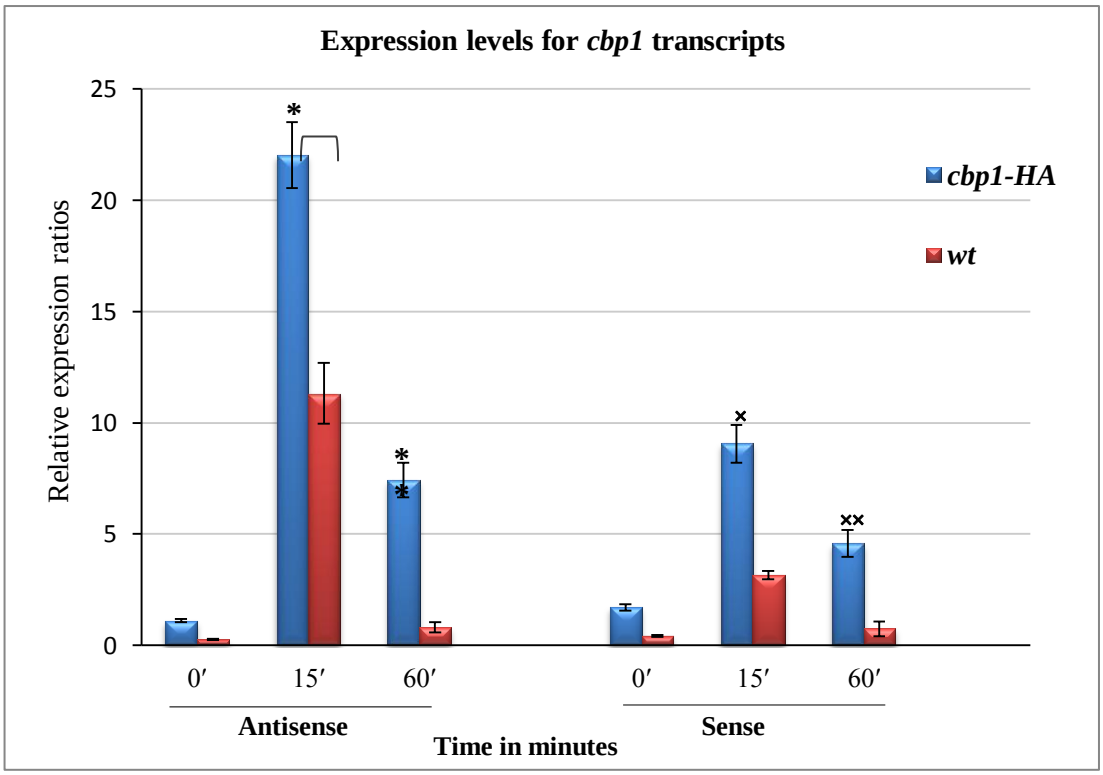


Figure 5: Strand specific qRT-PCR measurements of *cbp1* antisense and sense transcripts before and after heat shock. Expression levels of *cbp1* antisense and sense transcripts, before and after heat shock for 15′ and 60′ at 39˚C are shown here. Expression ratios shown on Y-axis are obtained after normalisation with the expression values for *aap1*. Error bars represent standard deviation among 3 biological replicates. P-value are calculated by using student-t test where, * and ▪ = 0.002, ** = 0.007, × = 0.006, × × = 0.02=, ● = 0.009 and ♦ represents a P- value of 0.01. Each symbol represents the P-value of respective transcript levels before

and after heat shock, except for ♦ that represents P-value for antisense transcripts between two strains at 15′ of heat shock.

Effect of antisense transcription on Cbp1 protein levels

As, an increase in antisense transcription of *cbp1* is accompanied by increased levels of *cbp1* sense mRNA, we next asked whether these increased levels of *cbp1* antisense transcripts effect the translation levels of *cbp1* protein. Both *HA*-tagged and *wt* were grown at 25˚C and exposed 39˚C for 15′ and 60′. 25 ml of the cells were collected before and after the heat shock and were processed for the protein extraction. We detected protein levels before and after heat shock by performing Western blotting.

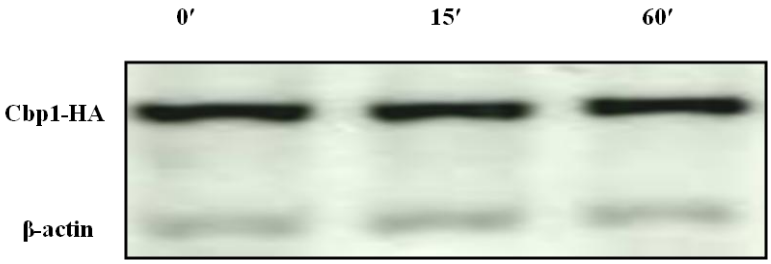


Figure 6: Western blot analysis. Western analysis was done for Cbp1-HA by harvesting 25ml cells for protein extraction. 25ug of protein sample was processed for polyachrylamide gel electrophoresis. Size of Cbp1-HA is 62 KDa. Here β-actin (49 KDa) is used as loading control.

Protein levels were found to be consistent in all samples (Figure 6), as we could not see any considerable difference in *cbp1* protein levels before and after heat shock in western blot analysis. This showed that over expression of *cbp1* antisense or sense transcription does not have any effect at the translational levels of *cbp1* protein during heat shock.

Effect of *cbp1* antisense transcription when transcribed in trans

As there is no obvious effect of antisense transcripts on *cbp1* protein levels during heat shock, when transcribed in *cis,* we therefore, decided to see if it could cause any effect when transcribed in *trans.* To establish their *trans* regulatory role, it was important to observe the phenotypic effects of *cbp1* on *S. pombe.* If the deletion of *cbp1* produces any particular phenotype it could be attributed to, either sense, or antisense transcripts of *cbp1.*Then by transcribing either of them in *trans,* by placing them under the control of an ectopic promoter, we can establish their role in gene regulation and confirm their ability to rescue that particular phenotype.

Phenotypic effect of *cbp1*

To assess the phenotypic effect of cbp1 gene on cell's growth, we used the *cbp1* mutant strain (Δcbp1:: hygB-MX6 leu-32), which carries the deletion of *cbp1* gene. We grew this strain along with *wt* on YES media to observe difference, if any, in growth behaviour between strains. We used two *Δcbp1* strains and two *wt* strains to avoid any chance of getting false results. These strains were grown at 25˚C to OD_{595} 0.5 and performed spot test on YES plates. These plates were considered as control and then incubated at 25˚C, 32˚C and 37˚C to observe any changes in the growth rate at different temperatures.

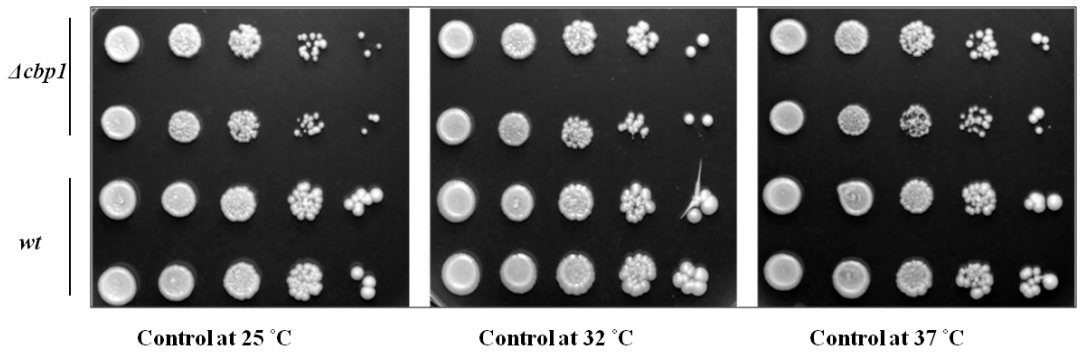


Figure 7: Spot test for the comparison of growth pattern for *Δcbp1* and *wt*. Different dilutions of the exponentially growing cells at 25˚C were spotted on the YES plates and were kept at 25˚C, 32˚C and 37˚C. Colonies from *wt* appeared after 3 days while mutants took 6 days to appear. Size of the colonies is also small for mutants.

We observed slow growth in *Δcbp1* strains at all three temperatures and our results were consistent for both *Δcbp1* strains (Figure 7). Colonies from *wt* cells appeared after 3 days, while *Δcbp1* colonies took 6 days to appear on same plates.

Next, we performed heat shock experiments at 39˚C and 42˚C for *Δcbp1*and *wt* cells to observe the phenotypic effects due to prolong exposure to high temperatures. Cells were collected at different time intervals and different dilutions were spotted on YES plates. Plates were then incubated at 25˚C for 6 days.

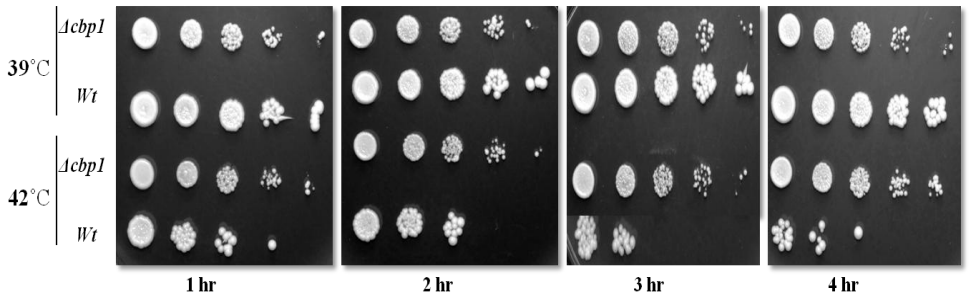


Figure 8: Spot test for the comparison of growth before and after heat shock at 39˚C and 42˚C for *Δcbp1* and *wt*. Both mutant and the *wt* were heat shocked at indicated temperatures for different time intervals after growing them to the OD_{595} 0.5. 5ml of cells were taken out at each time point and different dilutions of the cells

were spotted on YES plates and were incubated at 25˚C. *Wt* colonies appeared after 3 days while mutants took 6 days to appear.

Δcbp1 cells not only showed slow growth but also appeared to be temperature resistant (Figure 8). On the other hand *wt* showed normal growth but they appeared temperature sensitive and could not grow well when exposed to high temperature (42˚C) for 2, 3 and 4 hours.

To further confirm these findings, we analysed the survival rate of both *Δcbp1* and *wt* cells. Minimum of 200 colonies were plated on YES plates, after growing these strains at 25˚C and exposing them to heat shock at 42˚C for, 1, 2, 4 and 15 hours. Plates were incubated for 6 days at 25˚C. Colonies were counted on each plate and % survival was calculated at different time intervals. When analysed, results indicated that *Δcbp1* are long lived as compared to *wt* (Figure 9). The wild type appears to be very sensitive at such a high temperature when exposed for more than 1 hour, which is evident from very few or almost no colonies on the plates. Survival rate for mutant is significantly higher even when exposed to higher temperature for longer time periods such as 4 hours (P= 0.03) or even for 15 hours (P= 0.007). Again colonies from mutants took longer to appear on plates as compared to the *wt,* confirming our initial findings that mutants grow slowly but are heat resistant. These findings clearly indicate the heat resistant but slow growing phenotype of *Δcbp1*.

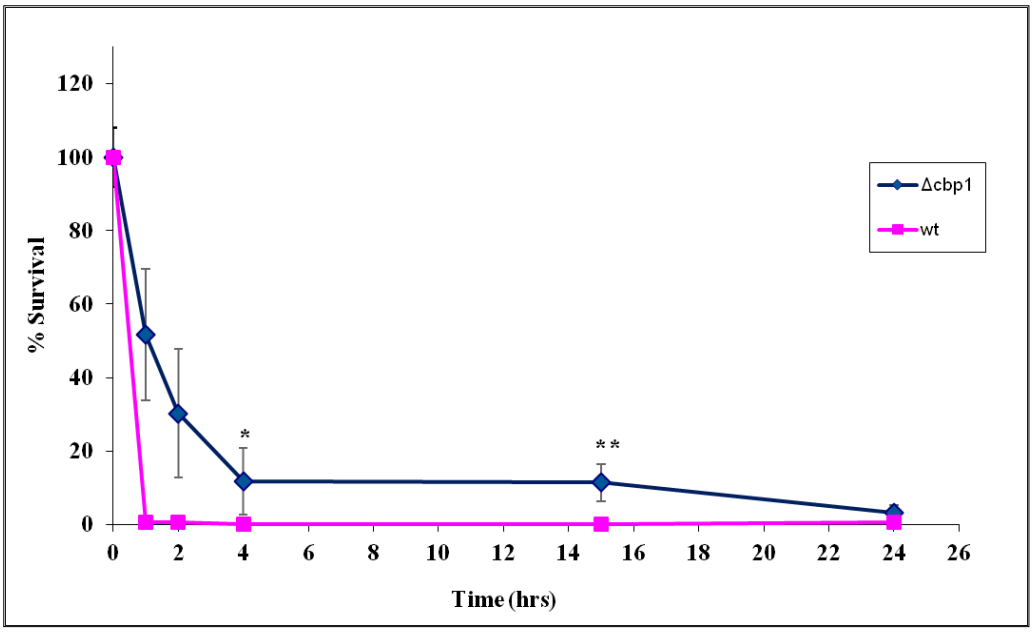


Figure 9: Survival analysis before and after the heat shock 42˚C for *Δcbp1* and *wt*. Both mutant and *wt* were heat shocked at 42˚C for different time intervals after growing them to the OD 0.5. 5ml of cells were taken out at each time point (0 minute, 1 hr, 2 hr, 4 hr and 15 hr after heat shock) and minimum of the 200 cells were plated on YES plates. After incubating at 25˚C for 6 days colonies were counted on each plate to analyse % of survival. Data normalised with time point zero is plotted on Y-axis, whereas, error bars represent standard deviation for 3 repeats. P value is calculated using student t-test, where, ***** represents p=0.03 and ** p=0.007**.**

Phenotype rescue by transcribing either, sense or, antisense transcript of *Δcbp1* in *trans*

Next, we tried to establish the phenotype rescue experiments to see, if it is the *cbp1* antisense, sense, or both, that can rescue the slow growth and temperature resistant phenotype. For this purpose, we tried to transcribe them individually by putting them under the control of an ectopic inducible promoter system.

To achieve this, we used the pURG3x vector system, which uses the promoter of uracil regulatable gene *urg1* (*Purg1)* and *nmt1* terminator system. In this vector,

gene under *Purg1* control can be fully induced and repressed within, 10 minutes of 250mg/ml uracil (Ura) addition and removal, respectively (Watt et al., 2008). To confirm the presence of right inserts in the vector, we setup a double digestion of empty vector and vector with inserts and analysed the products by polyachrylamide gel electrophoresis. These vectors with inserts were also sequenced to further verify the presence of *cbp1* sense and antisense inserts. *Δcbp1* cells were transformed and selected on the EMM without leucine (Leu*)*. Only those *S. pombe* cells were able to grow that were transformed with vector. We also performed colony PCR for the *cbp1* gene to select the transformants having right size of insert.

To assess the influence of *cbp1* sense and antisense on growth of *Δcbp1,* both transformants having either, *cbp1* sense *(Δcbp1pURG3x-cbp1s)* or, antisense *(Δcbp1pURG3x-cbp1as***)** were grown in EMM with and without Ura (+/- Ura). As, a control we used the deletion mutant transformed with an empty vector *(Δcbp1pURG3x).* After growing them at 25˚C, we exposed them to heat shock at 42˚C for 1, 2, 4 and 6 hours. Cells were collected at each of these intervals to spot different dilutions on EMM plates +/- Ura and incubated at 25˚C. We observed slow growth for control on EMM +/- Ura that is consistent with slow growth phenotype of *Δcbp1*, while both transformants showed normal growth (Figure 10).

In another experiment we compared growth behaviour of transformants with wild type. We transformed *wt* with empty vector, as *wt* cells are *leu* minus and cannot grow on EMM media. In this experiment we found similar growth patterns for both transformants and *wt*. Moreover, *wt* with empty vector appeared to be resistant to heat shock. It seems that heat shock sensitive phenotype was media specific.

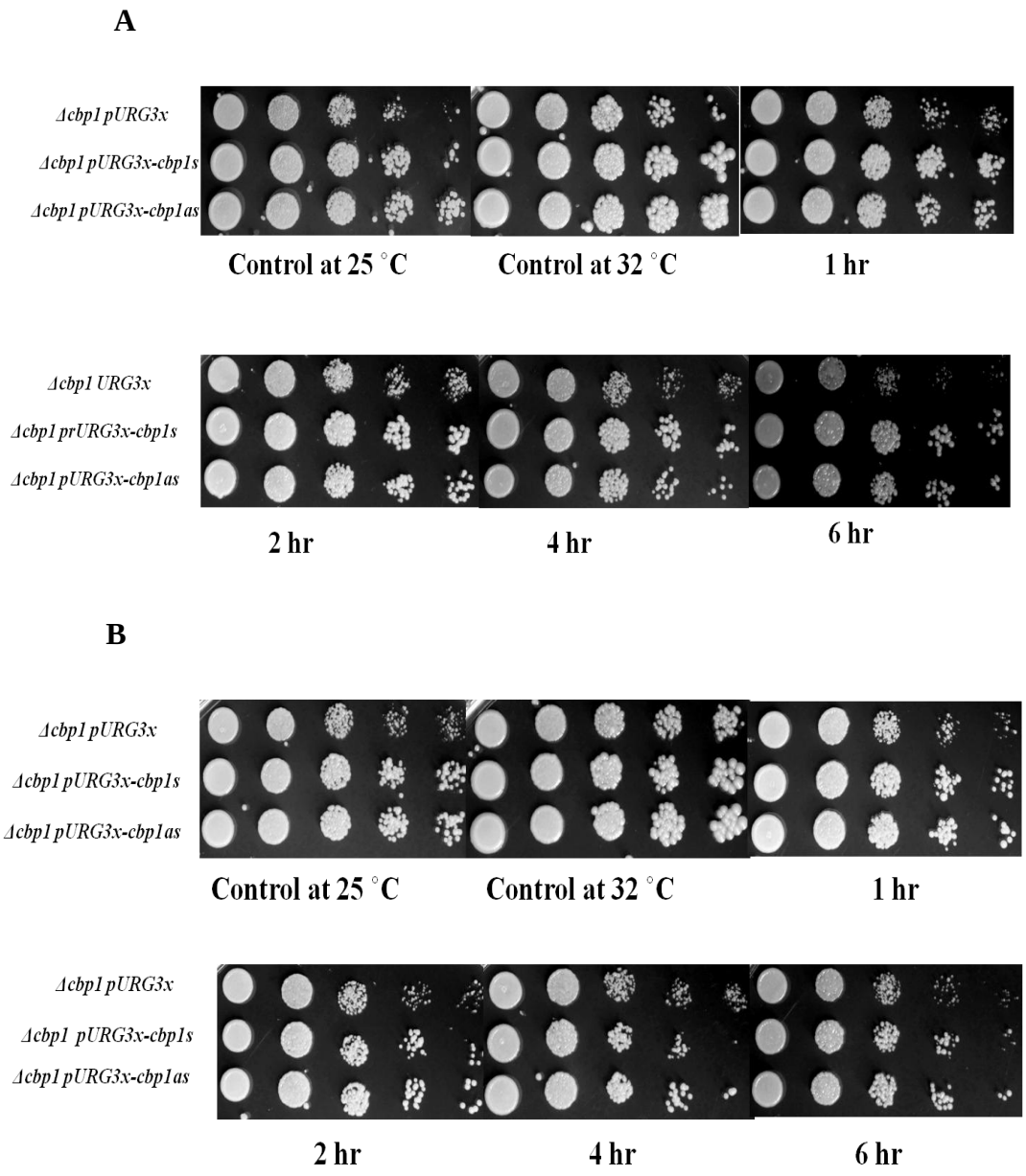


Figure 10: Spot test comparing growth rate of transformants before and after the heat shock at 42°C in EMM (A) and EMM + Ura (B). Each of these transformed *S. pombe* were heat shocked at indicated temperatures for different time intervals after growing them to OD_{595} 0.5. 5ml of cells were taken out at each time point and different dilutions were spotted on EMM +/- Ura (250mg/ml Ura) plates and were incubated at 25°C.

To see, what is happening at transcription level in *Δcbp1pURG3x-cbp1s* and *Δcbp1pURG3x-cbp1as*, we performed RT-PCR for them along with *Δcbp1, Δcbp1 pURG3x* and wild type transformed with empty vector. RNA was extracted from all

these strains after growing to OD_{595} 0.5 in EMM +/- Ura, before heat shock and after exposing them to 42˚C for 60 minutes. When analysed on 1.5 % agarose gel, we found that both sense and antisense transcripts are expressed in both transformants (Figure 11). Interestingly, transcript levels for *cbp1* sense and antisense remain detectable before and after heat shock when grown in the absence or presence of uracil.

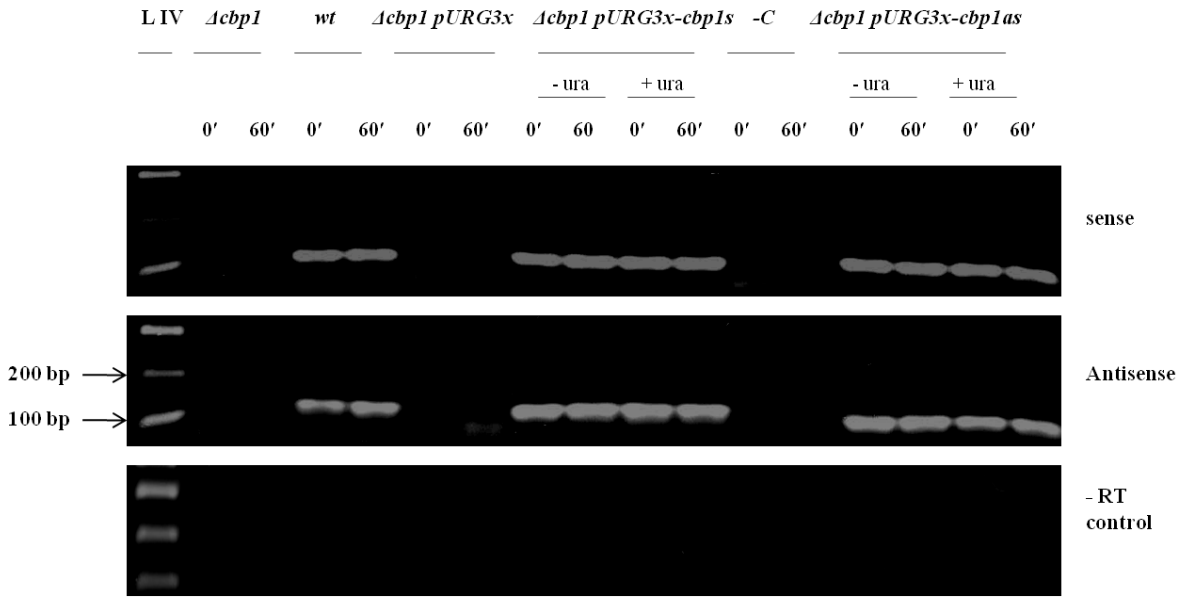


Figure 11: RT-PCR analysis for transformants along with *wt and Δcbp1*. RT-PCR analysis was performed after RNA was extracted from 50ml of exponentially growing cells before and after heat shock at 42˚C at indicated time points. Expected size of both transcripts is 102 bp. Absence of same size bands in –RT, -C (-template) controls and the *Δcbp1 pURG3x* shows good quality of RT-PCR.

Further we tried to quantify the expression levels for *cbp1* sense and antisense transcripts in *Δcbp1pURG3x-cbp1s* and *Δcbp1 pURG3x-cbp1as,* in the presence and absence of uracil in the media. We performed qRT-PCR for these transformants after growing them in EMM +/- Ura, before and after exposing them to heat shock at 42˚C for 60 minutes.

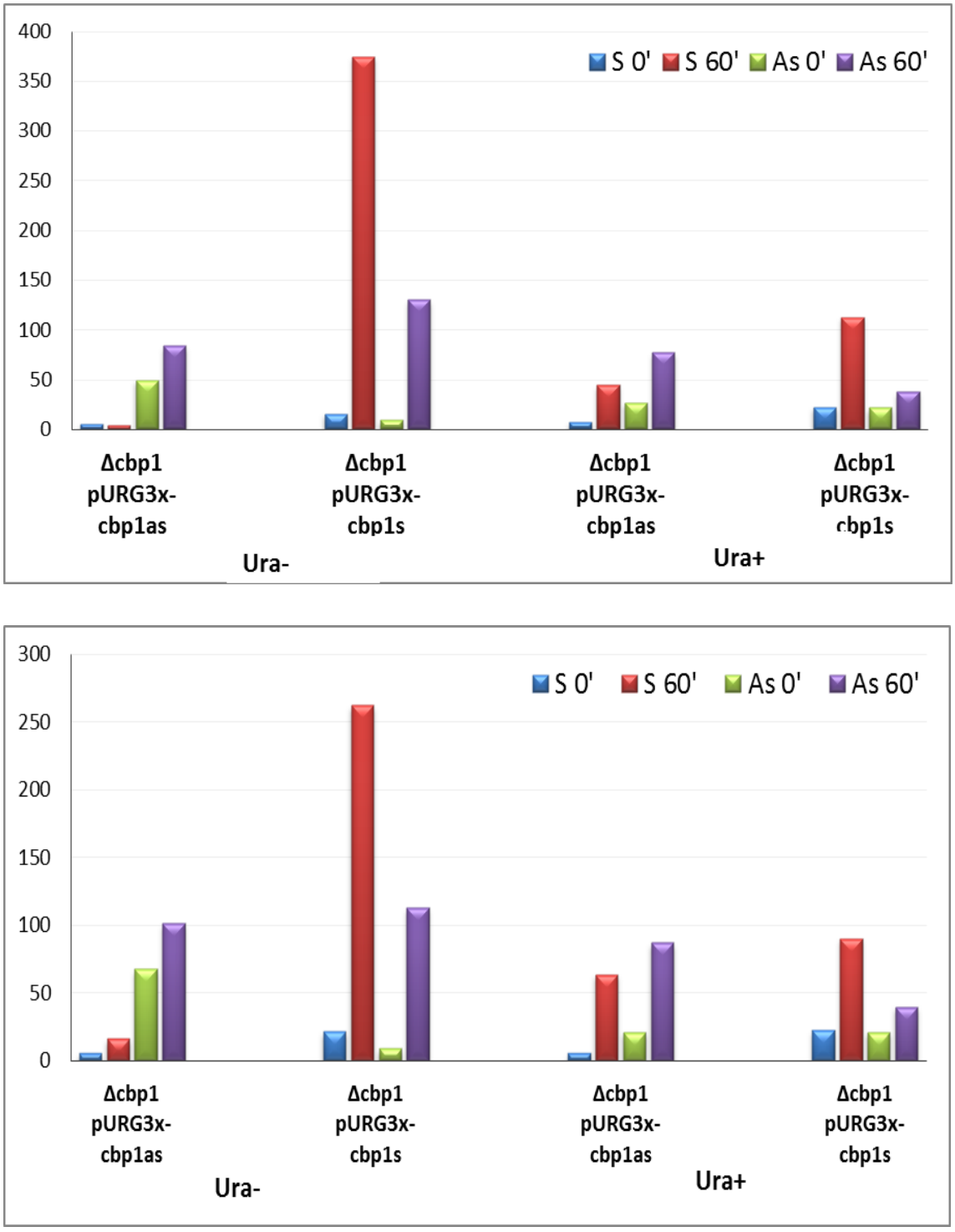


Figure 12: qRT-PCR analysis for transformants with (250 mg/l) or without uracil in the medium. Strand specific qRT-PCR measurements of *cbp1* antisense and sense transcripts to compare their expression levels before and after 60′ of heat shock at 42˚C. Y-axis represents expression ratios for sense and antisense transcripts obtained after normalisation with the expression values for *aap1*. Data is shown from two independent biological repeats.

Sense and Antisense transcripts of *cbp1*are found to be expressed in both *Δcbp1pURG3x-cbp1s* and *Δcbp1pURG3x-cbp1as* after exposure to heat shock at 42˚C (Fig. 12). Antisense transcript levels are significantly increased in both transformants after 60 minutes of heat shock when compared to before heat shock. We observed that expression levels for both sense and antisense ncRNA are detectable even in the absence of uracil. In *Δcbp1pURG3x-cbp1as*, only antisense transcription is dominant before and after heat shock in, while, sense transcripts appear to be very weakly expressed. While, in *Δcbp1pURG3x-cbp1s* expression levels of sense transcript are highly induced after heat shock along with highly significant increase in antisense ncRNA levels. This finding suggests that *pURG3x* is leaky and can give expression even in absence of uracil.

In presence of uracil, in *Δcbp1pURG3x-cbp1as*, levels of antisense ncRNA are significantly increased after heat shock along with an increase in sense mRNA. While, in *Δcbp1 pURG3x-cbp1s,* antisense transcripts show weak but significant expression levels after heat shock.

Although, presence of both sense and antisense transcripts of *cbp1* in *Δcbp1 pURG3x-cbp1s* and *Δcbp1 pURG3x-cbp1as*, makes it ambiguous to conclude the role of antisense transcription in phenotypic rescue, but presence of higher levels of only *cbp1* antisense ncRNA in *Δcbp1 pURG3x-cbp1as* in absence of uracil indicates that this might have a role in the phenotypic rescue, when expressed in *trans*.

Discussion

Antisense transcription is widespread in many eukaryotes including yeast. A recent genome wide analysis of *Schizosaccharomyces pombe* transcriptome has determined the proportion of protein-coding genes with prominent antisense expression (Ni et al., 2010). This study has revealed that approximately 2409 genes (or 47.4% of all *S. pombe* protein-coding genes) have antisense expression under normal and heat shock conditions which is consistent with a recent estimation of antisense expression (20–49%) in human genes (He et al., 2008). There is a chance that transcription of a large number of different antisense ncRNAs allow some of them to develop towards functional molecules, which can then become stable to confer higher fitness to the cell. In addition, these antisense transcripts might provide further opportunity for gene regulation to allow a fine tuned response to environmental conditions.

The important regulatory roles of antisense transcripts have been proposed, based on different studies including a number of large-scale transcription profiling studies. Although, few studies have reported a variety of functional mechanisms involved in antisense ncRNA mediated sense transcript regulation by direct experimental support, the exact mechanism of action for most of this antisense transcription is still unclear. In this study we focused to establish the role of an antisense ncRNA in regulation of *S. pombe* genome. This antisense ncRNA is (~ 1350 bp) transcribed from the opposite strand of *cbp1* sense transcript (Figure 1 & 2), starting at ~ 300 bp upstream of the 3′ end of *cbp1* ORF and extends beyond the transcriptional start site of the gene (~ 100 bp). Its expression is condition specific, as it is expressed only in response to heat shock. Expression of ncRNAs in response to a change in environmental conditions, such as nutrient limitation or stress, indicates their role in helping cells to adapt to these changes. Roles of few such antisense ncRNAs in yeast have been established in response to phosphate limitation (Uhler et al., 2007), carbon source (Houseley et al., 2008), lithium toxicity (Van Dijk et al., 2011) and zinc (Ehrensberger et al., 2013)etc. As, *cbp1* plays an important role in chromosomal segregation and retrotransposon silencing (Cam et al., 2008), therefore, it was interesting to see if its antisense transcript has any role in controlling its expression. We found that antisense transcript of *cbp1* is expressed and becomes considerably higher after exposing fission yeast cells to heat shock at 39˚C for 15′ and 60′.

Antisense ncRNAs that arise from the same locus as their complementary sense transcript can act in *cis,* which limits their function only to the genomic locus of their transcription. Recently, few such ncRNAs have been reported both in fission and budding yeast that can act in *cis* to regulate gene expression. Here, by using Real time-PCR we tried to establish the *cis* effect of antisense ncRNA on *cbp1* regulation. We have shown that antisense transcript levels are significantly higher (P = 0.002) after 15′ of heat shock accompanied by significant increase (P = 0.006) in sense transcript levels for *cbp1*. Antisense transcript levels decrease after 60′ of heat shock, although, still significantly higher (P = 0.007) as compared to the transcript levels before heat shock. This decrease in expression levels might be due to two possibilities, (1) RNA degradations pathways become activated in response to higher antisense ncRNA level and starts degrading them, (2) antisense transcripts have already activated heat shock responsive pathway after 15′ of heat shock. When comparing antisense transcript levels in *cbp1-HA* and *wt,* we observed that *cbp1* antisense transcript levels are significantly higher (P = 0.01) in *HA* tagged strain as compared to *wt*. As, 3' UTR plays an important role in the translation, localization and stability of the mRNA, and that alterations in the length of the 3' UTR due to changes in the position of the physiological termination codon influence the stability and translation of mRNA, therefore, it indicates that insertion of a *HA* tag to the 3' of *cbp1* has disrupted the 3' UTR of *cbp1* such that it has either stabilized, or increased the expression of antisense transcripts.

From the western blot analysis we cannot observe any effect of antisense transcription on protein levels. This indicates that *cbp1* antisense transcript does not have any influence on the protein levels of *cbp1*.

As exposure to heat shock resulted in an over expression of *cbp1* antisense along with its sense partner without affecting protein levels of *cbp1*, it was important to see what, if any, could be the role of *cbp1* antisense transcription . To establish its function in regulating gene expression it was necessary to understand, which of sense, or antisense could rescue the *cbp1* deletion phenotype. To achieve this goal, we decided to express, either sense, or antisense in *trans,* by putting them under the control of an ectopic promoter in *cbp1* deletion background.

Phenotypic assay for *Δcbp1* showed slow growth, as it took 6 days to appear on YES plates. This phenotype is in complete agreement with previous studies [88]. Interestingly, *Δcbp1* appeared to be extremely temperature resistant and showed growth even after prolong exposure to heat shock at 42˚C, whereas, *wt* could not survive this much high temperature (Figure 8 & 9). Temperature resistant phenotype appeared to be media specific i.e. this phenotype was observed only on YES plates. Next, we transformed *Δcbp1* with either, *pURG3x-cbp1s* (containing only sense transcript under the control of *pURG3x* promoter system) or *pURG3x-cbp1as* (antisense transcript under the control of promoter). The idea was to see, which of these transformants *Δcbp1pURG3x-cbp1as,* or *Δcbp1pURG3x-cbp1s,* or both, can rescue the slow growing phenotype, when under the control of an exogenous promoter *Purg1*. As, *pURG3x* vector system contains *leu* marker for selection of transformants, therefore, we were not able to observe this temperature resistant phenotype for transformants, as they were grown on EMM media lacking Leu. Both *Δcbp1pURG3x-cbp1as* and *Δcbp1pURG3x-cbp1s* show normal growth, while *Δcbp1pURG3x* still shows the slow growth (figure 10). Both transformants were able to grow normally after heat shock, even in the absence of Ura in media, that is required for repression of *pURG3x* promoter. RT-PCR analysis confirmed the presence of both *cbp1* sense and antisense transcripts in both of these transformants when grown in EMM +/- Ura (Figure 11). Reverse transcriptase minus (-RT), *Δcbp1* and *Δcbp1pURG3x* controls confirmed the quality of RT-PCR, as we did not get any product of incorrect size in these controls. Presence of *cbp1* sense or antisense transcripts in these transformant in the absence of Ura in the media shows that this promoter system is leaky and can give transcription even when it is repressed.

Expression levels for *cbp1* antisense ncRNA were found to be up regulated in both *pURG3x-cbp1s* and *Δcbp1 pURG3x-cbp1as,* after exposure to heat shock at 42˚C in EMM +/- Ura. Although, transcripts levels for both sense and antisense were detected in both *Δcbp1pURG3x-cbp1s* and *Δcbp1pURG3x-cbp1as,* but the presence of higher transcript levels of only antisense ncRNA in *Δcbp1pURG3x-cbp1as* before and after heat shock (in absence of Ura), suggest its role in phenotypic rescue (Figure 12).

The reason for this phenotypic rescue by both *Δcbp1pURG3x-cbp1s, Δcbp1pURG3x-cbp1as* might be due to the fact that *cbp1* sense and *cbp1* antisense are almost fully overlapping (figure 2). It seems that transcriptional start site of antisense ncRNA lies inside the *cbp1* sense and vice versa, so by putting either of them under the control of an exogenous promoter does not excludes the possibility of other being transcribed.

In a recent RNA-seq based study of fission yeast clade, Rhind and colleagues have reported antisense ncRNA (SPNCRNA.1608) that also shares its locus with *cbp1*.(Rhind et al., 2011) This antisense transcript is predicted to be ~1420 bp with its transcription start site lying ~ 96 bp downstream of transcriptional start site of *cbp1* antisense ncRNA, thus showing 93% overlap. Presence of another ncRNA (SPNCRNA.1608) antisense to *cbp1,* suggests the possibility of gene expression regulation by either or both of these antisense ncRNAs.

Although, role of *cbp1* antisense transcription in regulating gene expression by acting, either in *cis*, or in *trans*, remains unclear and difficult to conclude in this study, the presence of higher levels of *cbp1* antisense ncRNA in *Δcbp1 pURG3x-cbp1as,* in the absence of uracil points towards its possible role in gene regulation. However, further experimental work is required to establish its precise role in mediating gene expression.

Future experiments

To establish the role of *cbp1* antisense RNA, further experiments need to be performed.

- Our attempt to knock down antisense transcripts to express only *cbp1* sense mRNA, when expressed by an ectopic promoter, did not affect antisense ncRNA production. To study its role in regulating gene expression, mutagenesis of the *cbp1* antisense promoter could be performed to get an expression of only *cbp1* sense, when put under the control of an ectopic promoter. Mutation of native promoter of *cbp1* antisense might result in, either no, or very weak expression. This can clearly establish the role, if any, of antisense ncRNA in phenotypic rescue/ gene expression.
- In *S. pombe*, *cbp1* has been shown to silence Long Terminal Repeat (LTR) retrotransposons by recruiting histone deacetylases. In *S. cerevisiae*, Repression of Ty1 transposon mobility and expression has been reported to be mediated by *trans* acting Ty1 antisense ncRNA, by incorporating histone deacetylase and Set1 dependent histone methylation (Berretta et al., 2008). Therefore, it would be interesting to see the effect of over expression of *cbp1* antisense ncRNA on LTR silencing. To achieve this chromatin immuno precipitation (Chip-Chip) experiment need to be performed to find out, if Cbp1 still remains bound to its target, when exposed to heat shock. Also, it would be interesting to establish the effect of over expression of *cbp1* antisense transcript on expression and mobility of LTR retrotransposons by qRT-PCR.
- To establish the role of antisense transcription of *cbp1* in *cis*, or in *trans*, expression levels of longer *cbp1* antisense ncRNA (SPNCRNA.1608) need to be characterised. As, we conducted this study in 2009-2010, Primers we used for RT-PCR and qRT-PCR lie in the overlapping region of two antisense ncRNAs. Primers now can be designed to distinguish their expression levels from each other, which can show, which is involved in changing expression levels of *cbp1*.

These experiments might help to establish the role of both *cbp1* antisense ncRNA either on the gene itself or on any other locus.

References

Annilo, T., Kepp, K., and Laan, M. (2009). Natural antisense transcript of natriuretic peptide precursor A (NPPA): structural organization and modulation of NPPA expression. BMC Molecular Biology *10*, 81.

Arigo, J.T., Carroll, K.L., Ames, J.M., and Corden, J.L. (2006). Regulation of yeast NRD1 expression by premature transcription termination. Molecular Cell *21*, 641–651.

Van Bakel, H., Nislow, C., Blencowe, B.J., and Hughes, T.R. (2010). Most "dark matter" transcripts are associated with known genes. PLoS Biology *8*, e1000371.

Bass, B.L. (2002). RNA Editing by Adenosine Deaminases That Act on RNA. Annu Rev Biochem *71*, 817–846.

Beiter, T., Reich, E., Williams, R.W., and Simon, P. (2009). Antisense transcription: a critical look in both directions. Cellular and Molecular Life Sciences : CMLS *66*, 94–112.

Beltran, M., Puig, I., Peña, C., García, J.M., Álvarez, A.B., Peña, R., Bonilla, F., and Herreros, A.G. De (2008). A natural antisense transcript regulates Zeb2 / Sip1 gene expression during Snail1-induced epithelial – mesenchymal transition. Genes & Development *22*, 756–769.

Berretta, J., Pinskaya, M., and Morillon, A. (2008). A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in *S. cerevisiae*. Genes & Development *22*, 615–626.

Bertone, P., Stolc, V., Royce, T.E., Rozowsky, J.S., Urban, A.E., Zhu, X., Rinn, J.L., Tongprasit, W., Samanta, M., Weissman, S., et al. (2004). Global identification of human transcribed sequences with genome tiling arrays. Science *306*, 2242–2246.

Birney, E., Stamatoyannopoulos, J. a, Dutta, A., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799–816.

Bitton, D. a, Grallert, A., Scutt, P.J., Yates, T., Li, Y., Bradford, J.R., Hey, Y., Pepper, S.D., Hagan, I.M., and Miller, C.J. (2011). Programmed fluctuations in sense/antisense transcript ratios drive sexual differentiation in S. pombe. Molecular Systems Biology *7*, 559.

Cam, H.P., Noma, K., Ebina, H., Levin, H.L., and Grewal, S.I.S. (2008). Host genome surveillance for retrotransposons by transposon-derived proteins. Nature *451*, 431–436.

Camblong, J., Beyrouthy, N., Guffanti, E., Schlaepfer, G., Steinmetz, L.M., and Stutz, F. (2009). Trans-acting antisense RNAs mediate transcriptional gene cosuppression in *S. cerevisiae*. Genes & Development *23*, 1534–1545.

Cawley, S., Bekiranov, S., Ng, H.H., Kapranov, P., Sekinger, E.A., Kampa, D., Piccolboni, A., Sementchenko, V., Cheng, J., Williams, A.J., et al. (2004). Unbiased Mapping of Transcription Factor Binding Sites along Human Chromosomes 21 and 22 Points to Widespread Regulation of Noncoding RNAs 3380 Central Expressway. Cell *116*, 499–509.

Chen, J., Sun, M., Kent, W.J., Huang, X., Xie, H., Wang, W., Zhou, G., Shi, R.Z., and Rowley, J.D. (2004). Over 20% of human transcripts might form sense-antisense pairs. Nucleic Acids Research *32*, 4812–4820.

Churchman, L.S., and Weissman, J.S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature *469*, 368–373.

David, L., Huber, W., Granovskaia, M., Toedling, J., Palm, C.J., Bofkin, L., Jones, T., Davis, R.W., and Steinmetz, L.M. (2006). A high-resolution map of transcription in the yeast genome. Proceedings of the National Academy of Sciences of the United States of America *103*, 5320–5325.

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Research *22*, 1775–1789.

Van Dijk, E.L., Chen, C.L., d'Aubenton-Carafa, Y., Gourvennec, S., Kwapisz, M., Roche, V., Bertrand, C., Silvain, M., Legoix-Né, P., Loeillet, S., et al. (2011). XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. Nature *475*, 114–117.

Dutrow, N., Nix, D. a, Holt, D., Milash, B., Dalley, B., Westbroek, E., Parnell, T.J., and Cairns, B.R. (2008). Dynamic transcriptome of *Schizosaccharomyces pombe* shown by RNA-DNA hybrid mapping. Nature Genetics *40*, 977–986.

Ehrensberger, K.M., Mason, C., Corkins, M.E., Anderson, C., Dutrow, N., Cairns, B.R., Dalley, B., Milash, B., and Bird, A.J. (2013). Zinc-dependent Regulation of the adh1 Antisense Transcript in Fission Yeast. The Journal of Biological Chemistry *288*, 759–769.

Faghihi, M.A., Modarresi, F., Khalil, A.M., Wood, D.E., Sahagan, B.G., Morgan, T.E., Finch, C.E., St, G., Iii, L., and Paul, J. (2008). Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of β-secretase expression. Nat Medicine *14*, 723–730.

Feinberg, A.P., and Tycko, B. (2004). The history of cancer epigenetics. Nature Reviews. Cancer *4*, 143–153.

Gelfand, B., Mead, J., Bruning, A., Apostolopoulos, N., Tadigotla, V., Nagaraj, V., Sengupta, A.M., and Vershon, A.K. (2011). Regulated antisense transcription controls expression of cell-type-specific genes in yeast. Molecular and Cellular Biology *31*, 1701–1709.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. Nature Reviews. Genetics *10*, 94–108.

Granovskaia, M. V, Jensen, L.J., Ritchie, M.E., Toedling, J., Ning, Y., Bork, P., Huber, W., and Steinmetz, L.M. (2010). High-resolution transcription atlas of the mitotic cell cycle in budding yeast. Genome Biology *11*, R24.

Halverson, D., Baum, M., Stryker, J., Carbon, J., and Clarke, L. (1997). A centromere DNA-binding protein from fission yeast affects chromosome segregation and has homology to human CENP-B. The Journal of Cell Biology *136*, 487–500.

He, Y., Vogelstein, B., Velculescu, V.E., Papadopoulos, N., and Kinzler, K.W. (2008). The antisense transcriptomes of human cells. Science *322*, 1855–1857.

Hongay, C.F., Grisafi, P.L., Galitski, T., and Fink, G.R. (2006). Antisense transcription controls cell fate in *Saccharomyces cerevisiae*. Cell *127*, 735–745.

Houseley, J., Rubbi, L., Grunstein, M., Tollervey, D., and Vogelauer, M. (2008). A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. Molecular Cell *32*, 685–695.

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., et al. (2005). Antisense transcription in the mammalian transcriptome. Science *309*, 1564–1566.

Lapidot, M., and Pilpel, Y. (2006). Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. EMBO Reports *7*, 1216–1222.

Lardenois, A., Liu, Y., Walther, T., Chalmel, F., Evrard, B., Granovskaia, M., Chu, A., Davis, R.W., Steinmetz, L.M., and Primig, M. (2011). Execution of the meiotic noncoding RNA expression program and the onset of gametogenesis in yeast require the conserved exosome subunit Rrp6. Proceedings of the National Academy of Sciences of the United States of America *108*, 1058–1063.

Lavorgna, G., Dahary, D., Lehner, B., Sorek, R., Sanderson, C.M., and Casari, G. (2004). In search of antisense. Trends in Biochemical Sciences *29*, 88–94.

Li, Y.-Y., Qin, L., Guo, Z.-M., Liu, L., Xu, H., Hao, P., Su, J., Shi, Y., He, W.-Z., and Li, Y.-X. (2006). In silico discovery of human natural antisense transcripts. BMC Bioinformatics *7*, 18.

Mahmoudi, S., Henriksson, S., Corcoran, M., Méndez-Vidal, C., Wiman, K.G., and Farnebo, M. (2009). Wrap53, a natural p53 antisense transcript required for p53 induction upon DNA damage. Molecular Cell *33*, 462–471.

Mattick, J.S. (2004). Legislative Activity : RNA regulation : a new genetics ? Genetics *5*, 316– 323.

Mattick, J.S. (2011). The central role of RNA in human development and cognition. FEBS Letters *585*, 1600–1616.

Mattick, J.S., and Makunin, I. V (2005). Small regulatory RNAs in mammals. Human Molecular Genetics *14*, R121–32.

Misra, S., Crosby, M. a, Mungall, C.J., Matthews, B.B., Campbell, K.S., Hradecky, P., Huang, Y., Kaminker, J.S., Millburn, G.H., Prochnik, S.E., et al. (2002). Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review. Genome Biology *3*, 0083.1–0083.22.

Nakamura, T., Kishida, M., and Shimoda, C. (2000). The *Schizosaccharomyces pombe spo6* gene encoding a nuclear protein with sequence similarity to budding yeast Dbf4 is required for meiotic second division and sporulation. Genes to Cells *5*, 463–479.

Ni, T., Tu, K., Wang, Z., Song, S., Wu, H., Xie, B., Scott, K.C., Grewal, S.I., Gao, Y., and Zhu, J. (2010). The prevalence and regulation of antisense transcripts in Schizosaccharomyces pombe. PloS One *5*, e15271.

Osato, N., Yamada, H., Satoh, K., Ooka, H., Yamamoto, M., Suzuki, K., Kawai, J., Carninci, P., Ohtomo, Y., Murakami, K., et al. (2003). Antisense transcripts with rice full-length cDNAs. Genome Biology *5*, R5.

Prescott, E.M., and Proudfoot, N.J. (2002). Transcriptional collision between convergent genes in budding yeast. Proceedings of the National Academy of Sciences of the United States of America *99*, 8796–8801.

Purnapatre, K., Piccirillo, S., Schneider, B.L., and Honigberg, S.M. (2002). The CLN3/SWI6/CLN2 pathway and SNF1 act sequentially to regulate meiotic initiation in *Saccharomyces cerevisiae*. Genes to Cells *7*, 675–691.

Rhind, N., Chen, Z., Yassour, M., Thompson, D.A., Brian, J., Robbertse, B., Goldberg, J.M., Aoki, K., Bayne, E.H., Berlin, A.M., et al. (2011). Comparative Functional Genomics of the Fission Yeasts. Science *332*, 930–936.

Shah, J.C., and Clancy, M.J. (1992). IME4, a gene that mediates MAT andnutritional control of meiosis in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *12*, 1078–1086.

Stefani, G., and Slack, F.J. (2008). Small non-coding RNAs in animal development. Nature Reviews. Molecular Cell Biology *9*, 219–230.

Storz, G., Altuvia, S., and Wassarman, K.M. (2005). An abundance of RNA regulators. Annual Review of Biochemistry *74*, 199–217.

Sugiyama, T., Cam, H.P., Sugiyama, R., Noma, K., Zofall, M., Kobayashi, R., and Grewal, S.I.S. (2007). SHREC, an effector complex for heterochromatic transcriptional silencing. Cell *128*, 491–504.

Thiebaut, M., Kisseleva-Romanova, E., Rougemaille, M., Boulay, J., and Libri, D. (2006). Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance. Molecular Cell *23*, 853–864.

Uhler, J.P., Hertel, C., and Svejstrup, J.Q. (2007). A role for noncoding transcription in activation of the yeast PHO5 gene. Proceedings of the National Academy of Sciences of the United States of America *104*, 8011–8016.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C. a, Holt, R. a, et al. (2001). The sequence of the human genome. Science *291*, 1304–1351.

Wagner, E.G.H., and Simons, R.W. (1994). Antisense RNA control in Bacteria, Phages, and Plasmids. Annu. Rev. Microbiol. *48*, 713–742.

Wang, J.-P.Z., Lindsay, B.G., Cui, L., Wall, P.K., Marion, J., Zhang, J., and dePamphilis, C.W. (2005). Gene capture prediction and overlap estimation in EST sequencing from one or multiple libraries. BMC Bioinformatics *6*, 300.

Watt, S., Mata, J., López-Maury, L., Marguerat, S., Burns, G., and Bähler, J. (2008). Urg1: a Uracil-Regulatable Promoter System for Fission Yeast With Short Induction and Repression Times. PloS One *3*, e1428.

Werner, A., and Berdal, A. (2005). Natural antisense transcripts: sound or silence? Physiological Genomics *23*, 125–131.

Werner, A., and Sayer, J. a (2009). Naturally occurring antisense RNA: function and mechanisms of action. Current Opinion in Nephrology and Hypertension *18*, 343–349.

Wilhelm, B.T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., Penkett, C.J., Rogers, J., and Bähler, J. (2008). Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature *453*, 1239–1243.

Xu, Z., Wei, W., Gagneur, J., Clauder-Münster, S., Smolik, M., Huber, W., and Steinmetz, L.M. (2011). Antisense expression increases gene expression variability and locus interdependency. Molecular Systems Biology *7*, 468.

Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Münster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., and Steinmetz, L.M. (2009). Bidirectional promoters generate pervasive transcription in yeast. Nature *457*, 1033–1037.

Yassour, M., Pfiffner, J., Levin, J.Z., Adiconis, X., Gnirke, A., Nusbaum, C., Thompson, D.-A., Friedman, N., and Regev, A. (2010). Strand-specific RNA

sequencing reveals extensive regulated long antisense transcripts that are conserved across yeast species. Genome Biology *11*, R87.

Yelin, R., Dahary, D., Sorek, R., Levanon, E.Y., Goldstein, O., Shoshan, A., Diber, A., Biton, S., Tamir, Y., Khosravi, R., et al. (2003). Widespread occurrence of antisense transcription in the human genome. Nature Biotechnology *21*, 379–386.

Yu, W., Gius, D., Onyango, P., Muldoon-jacobs, K., Karp, J., Feinberg, A.P., and Cui, H. (2009). Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. Nature *451*, 202–206.