Transcriptional profiling of the model organism *A. gossypii*: Comparison of life cycle stages and transcription factor deletions

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel
von

Riccarda Rischatsch

aus Vaz/Obervaz GR

Basel, 2007

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von Prof. Peter Philippsen und Dr. Lars Steinmetz.

Basel, den 16.10.2007

Prof. Hans-Peter Hauri
Dekan

Table of contents

Summary

Experiments described in this PhD thesis used for the first time custom-made oligonucleotide chips to substantially increase our knowledge and understanding of the model organism *A. gossypii.* Four sets of experiments of different scales were performed:

(1) a transcriptome analysis of two strains at different developmental stages
(2) a comparison of the cell wall proteome and transcriptome
(3) a bioinformatics search for conserved promoter elements in target genes of the presumptive mating and filamentation signaling network
(4) a transcriptome analysis in strains lacking the transcriptional repressor AgDig1/2 or the transcriptional activators AgSte12 and AgTec1.

Chapters 1 to 4 and Appendix 1 and 2 describe in detail the steps of duplicate target preparations and data quality checks (chapter 1) followed by the presentation and discussion of transcriptional profiles at eight developmental stages: spores and four time points of germination up to bipolar germlings (chapter 2), two stages of hyphal high speed growth, one in liquid medium, the other on agar medium (chapter 3 and appendix 1), and finally sporulation at limiting nutrient conditions (chapter 4). For these experiments mRNA was isolated from a double auxotrophic derivative (Agleu2Δthr4Δ called ΔlΔt) of the sequenced *A. gossypii* strain. Transcription profiles were also determined for germination and high speed hyphae of strain FDAG, a novel natural isolate (Appendix 2). Key results are:

(1) very high expression of cell wall genes and genes of unknown function
(2) very different expression profiles of a novel cell wall gene family with similarity to the single copy *S. cerevisiae* CWP1 gene
(3) sporulation-specific expression of a novel gene family with similarity to the single gene copy *S. cerevisiae* HSP26 gene
(4) a high percentage (10/20) of homologs to *S. cerevisiae* twin genes among the most abundant transcripts in spores
(5) lack of ribosomal protein transcripts among the top 100 most abundant mRNAs in spores followed by coordinated up-regulation in young germlings
(6) an apparently relaxed glucose catabolite repression, because high speed hyphae on agar employ both pathways, glycolysis and gluconeogenesis including a highly active glyoxylate cycle
(7) a 150-fold up-regulation of an *A. gossypii* specific gene (NOHBY712) in high speed hyphae on agar medium
(8) an up-regulation of a phosphatidyl-inositol-P phosphatase (INP54) in fast speed hyphae indicating substantial changes in membrane composition
(9) one third unknown genes among the 15 most down-regulated genes in high-speed hyphae
(10) two up-regulated *A. gossypii* specific genes of unknown function at sporulation
(11) one highly expressed histone 3 gene during sporulation (also highly abundant mRNA in spores) pointing to substantial gene silencing during germination.

Chapter 5 describes results from a collaboration with Frans M. Klis and Piet de Groot (University of Amsterdam) to characterize the cell wall proteome of *A. gossypii.* Two highly expressed but so far unknown genes were identified as cell wall genes and a complete correlation between 14 cell wall proteins and high expression was established except for AgCCW12, which based

on its predicted amino acid sequence, could not yield a tryptic peptide for detection in mass spectrometry.

Chapter 6 concerns the comparison of genes controlling the mating and filamentation pathways in *S. cerevisiae* with orthologous genes in *A. gossypii*. To shed light on the functional conservation of the two pathways including transcriptional regulation, a bioinformatics analysis was carried out. This analysis looked into conservation of transcription factor binding sites in promoters of target genes of the two pathways. To experimentally investigate regulatory network conservation, the two transcription factors AgTec1 and AgSte12 and their repressor AgDig1/2 were deleted. The key results are:

(1) Components of the mating and filamentation pathway are highly conserved as concluded from sequence comparison between *A. gossypii* and *S. cerevisiae.* However the transcription factors and the repressor seem to have evolved faster during the 100 million years since the separation of the *A. gossypii* and the *S. cerevisiae* lineage.
(2) The Ste12 binding sites were conserved in 13 out of 18 promoters, which suggest that AgSte12 regulates a similar set of genes as ScSte12.
(3) The Tec1 binding sites were conserved in only 7 out of 20 promoters which might suggest that AgTec1 has a (partially) different role and the regulatory network has been rewired during evolution.
(4) Mycelia deleted for AgDIG1/2 showed a reduced maximal growth speed of 121 μm/h, which accounts for 67% of wild type growth speed. In addition the colony surface of Ag*dig1/2*Δ was altered, and no spores were formed in this mutant.
(5) Mycelia deleted for AgTec1 grew invasively after 10 days on full medium plates. The average nuclear distance in Ag*tec1*Δ had nearly doubled compared to wild type. Similar phenotypes were observed in wild type mycelia in response to glucose limitation.
(6) Mycelia deleted for AgSTE12 did not show a changed phenotype under conditions where the mating cascade was not induced by external stimuli.

Chapter 7 describes the differential gene expression in strains lacking the transcriptional repressor AgDig1/2 or the transcriptional activators AgSte12 and AgTec1. The fifteen top up- and down-regulated genes are discussed. To support and/or extent the single gene analysis the activity of predefined groups of genes was scored with t-profiler (Boorsma et al, 2005).

The key results for DIG1/2 are:

(1) AgDIG1/2 has a similar role as ScDIG1/2 which is to repress mating genes and filamentation genes. It also represses a set of genes with *S. cerevisiae* homologs of unknown function.
(2) AgPRY1/PRY2, a gene with homology to the plant PR-1 class of pathogen related proteins, was among the top fifteen up-regulated genes in Agdig1/2Δ and among the top15 down-regulated genes in Agste12Δ. This expression pattern suggest a possible role of this gene in mating.
(3) Glyoxylate cycle genes were significantly down-regulated in Ag*dig1/2*Δ. Either the deletion of AgDIG1/2 down-regulates these genes preventing hyphae to reach fast growth speeds on glucose plates. Or, the deletion of AgDIG1/2 leads to an induction of mating genes thus causing a “cell cycle arrest”.
(4) Meiosis and sporulation genes were down-regulated in Ag*dig1/2*Δ which was consistent with the observed sporulation defect.
(5) Ribosomal protein genes were down-regulated in Ag*dig1/2*Δ.

The key results for TEC1 are:

(1) Thiamine biosynthesis genes and genes involved in utilization of alternative carbon sources were up-regulated in Ag*tec1*Δ.
(2) Ribosome biosynthesis and assembly genes, in particular genes involved in snoRNA binding, processing and maturation of pre-rRNA, were down-regulated in Ag*tec1*Δ. This included the genes AgNOG1, AgNOP7 and AgRLP24 whose *S. cerevisiae* homologs code for proteins that form a complex which is tethered to the nucleolus by nutrient depletion causing cessation of late stages of ribosome biogenesis (Honma et al., 2006).
(3) TOR (target of rapamycin) broadly controls ribosome biogenesis and in *S. cerevisiae* the ScNOG1/NOP7/RLP24 complex. Therefore we hypothesize that components of the TOR pathway are upstream of AgTec1 and that AgTec1 is a transcription factor that contributes to control of ribosome biogenesis via AgNOG1, possibly through the putative AgNog1/Nop7/Rlp24 complex.
(4) In well fed polarly growing germlings, AgTec1 has a dual role as it acts as a repressor on the expression of genes involved in utilization of alternative carbon sources and thiamine biosynthesis and as an activator on the expression of ribosome biosynthesis genes.

The key results for STE12 are:

(1) Even under conditions where the mating cascade was not induced, AgSTE12 played a role in the cell. Genes whose expression was positively affected by deletion of AgSTE12 were mainly genes involved in amino acid biosynthesis.

General introduction

General introduction

General introduction to *Ashbya gossypii*

Ashbya gossypii is a filamentous fungus with a sequenced 8.8 Mb genome that encodes 4718 protein-coding genes (Dietrich *et al.*, 2004). Custom made DNA chips designed by Affymetrix are also available. *A. gossypii* is closely related to the budding yeast *Saccharomyces cerevisiae*. Both fungi belong to the phylum of the ascomycetes. *A. gossypii* was first described in 1926 as a pathogen, transmitted by insects, that causes stigmatomycosis on cotton (*Gossypium hirsutum*) bolls and seeds (Nowell, 1926). It can also infect citrus, tomato and flowering plants (Nowell, 1926; Starmer *et al.*, 1987; Guillermond, 1998) and has been widespread in tropic and sub-tropic regions of Northern and southern America, Africa and South East Asia (Batra, 1973). Although this fungus caused severe economical losses in the first half of the century, insect control permitted to fully eradicate infections (Batra, 1973). *A. gossypii* naturally overproduces riboflavin (vitamin B2) (Wickerham, 1946) and accumulation of this vitamin is responsible for the yellow pigmentation of the mycelium. It has been suggested that riboflavin protects *A. gossypii* spores against UV light (Stahmann *et al.*, 2001). Mutation and screening programs with *A. gossypii* led to the isolation of strains commercially used for riboflavin production.

A. gossypii and *S. cerevisiae* originate from the same ancestor

Ashbya gossypii and *Saccharomyces cerevisiae* originate from the same ancestor. Their lineages separated about 100 million years ago. Comparison of the *A. gossypii* genome to that of *S. cerevisiae* showed that 95% of *A. gossypii* protein-coding genes have a homologue in budding yeast and that 90% are found in synteny (conservation of gene order) in the two genomes (Dietrich *et al.*, 2004). The pattern of synteny is referred to as “ancient synteny” because the two genomes are not strictly co-linear but individual areas of the *A. gossypii* genome correspond to two distinct *S. cerevisiae* chromosomal regions with homologous genes alternatively present on either yeast chromosome. Ancient synteny was observed for 96% of the *A. gossypii* genome and provides strong evidence that the evolution of *S. cerevisiae* included a whole-genome duplication (Dietrich *et al.*, 2004). The evolution of synteny was complicated by the fact that the evolutionary paths of *A. gossypii* and *S. cerevisiae* included not only about 300 translocations and inversions but also the loss of 4000 genes post whole genome duplication. For 90% of all duplicate genes generated by the whole-genome duplication, one copy was lost. The remaining duplicated genes in *S. cerevisiae* are referred to as twin genes. The high quality sequence annotation for the two organisms allowed the identification of 261 *A. gossypii* genes which do not have a homolog in *S. cerevisiae*. These “species-specific” genes are referred to as NOHBYs, which stands for No homolog in baker’s yeast.

The *A. gossypii* life cycle

A. gossypii grows filamentously, producing extremely elongated cells which are called hyphae. Such growth behavior leads to a dense net of hyphae on solid substrates referred to as mycelium. In contrast to most filamentous fungi, *A. gossypii* displays a very simple life cycle, depicted in Figure 0-1 A-H. The *A. gossypii* mycelium covers at 30°C an

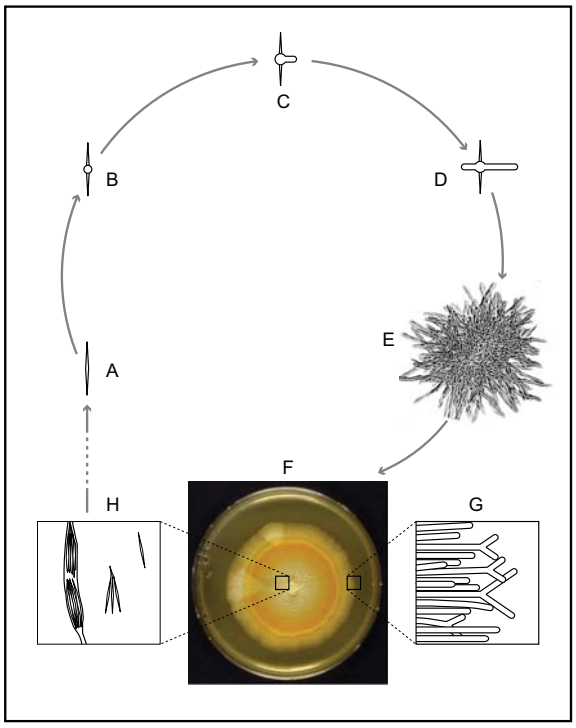


Figure 0-1 Life cycle of *A. gossypii.* The figure summarizes the life cycle of *A. gossypii* from germination to sporulation (Ashby and Nowell, 1926; Knechtle et al., 2006, reviewed by Philippsen et al., 2005 and Wendland and Walther, 2005). Figure was modified from M. Köhli, PhD thesis, 2007.

8 cm-Petri-dish in about 7 days. *A. gossypii* produces needle-shaped spores that are 13 to 16 µm long and 2 µm wide (A). They can germinate when conditions are favorable, for instance in the presence of a rich growth medium. During germination the central region of the spore enlarges in an isotropic manner and forms a germ bubble (B). After an initial isotropic growth phase germlings switch to polar growth. First one hypha emerges from the germ bubble, forming a unipolar germling (C). This event is followed by a second event in which a second hypha emerges at the opposite site giving rise to a bipolar germling (D). Formation of lateral branches (not shown) is typical for young and advanced mycelium (E). Upon maturation, growth speed is dramatically increased. Germ tubes emerging from the germ bubble initially grow with a tip speed of 6-10 µm/h (C and D). An advanced mycelium reaches a growth speed of 80 µm/h (E). The maximal growth speed of 200 µm/h is reached in *A. gossypii* hyphae growing on plates after approximately three days. The tips of these mature hyphae have the ability to split into two equally fast growing hyphae (G).
The *A. gossypii* hyphae are multinucleated and compartmented by open septa which in young parts appear as rings that allow transfer of nuclei and in older parts may appear as closed discs (Kaufmann, 2007). Compartments typically contain around eight nuclei. Sporulation occurs spontaneously on solid medium and is characterized by the drastic enlargement of several segments in close proximity to the tip (H). Spores are formed within these enlarged regions that can be described as sporangia. Spores are arranged in bundles (Nowell, 1926). At maturation, sporangia lyse to release the needle-shaped spores (A).

Goal of this PhD thesis

The main goal of was to achieve high quality gene expression data covering all major life cycle stages of the filamentous fungus *A. gossypii*. Specific goals concerned the development of individual transription profiles a) from spores to bipolar germlings (time frame 9 hours) b) from bipolar germlings to advanced mycelia and to very fast growing hyphae (time frame 9 h and 4 days, respectively) c) from very fast growing hyphae to sporulating mycelium (time frame 2 days). The final goal of this life cycle study was to correlate transcription profiles of genes or gene groups with the biology of different developmental stages.

A second major goal of this PhD thesis concerned the comparison of genes controlling the mating and filamentation pathways in *S. cerevisiae* with orthologous genes in *A. gossypii*. Even though the *A. gossypii* genome encodes homologous genes of all components of these two pathways, there are major differences in the modes of growth between both organisms, and the potential for mating is unknown

in *A. gossypii*. To shed light on functional conservation the two pathways including transcriptional regulation, a bioinformatics analysis was first carried out. This analysis looked into transcription factor binding sites in promoters of target genes of the two pathways. To experimentally investigate regulatory network conservation, the two transcription factors that act at the end of the pathways and their potential repressor were deleted. Subsequent comparisons of mRNA levels of the deletion mutants with wild type allowed the identification of genes and gene groups which are regulated by the transcription factors/repressor.

The expression data was generated using a laboratory strain. Additionally, several life cycle stages were also tested in a strain that was recently isolated from a cotton stainer in Florida (Dietrich, personal communication) allowing the investigation of intra-species differences in gene expression.

Chapter 1: Technical chapter

Chapter 1: Technical chapter

1.1 Introduction

The simple life cycle and the small genome size (8.8 Mb) make the filamentous fungus *A. gossypii* an ideal model organism. An abundance of tools for genetic, molecular genetics and biochemical analysis are available. The genome has been fully sequenced (Dietrich *et al.*, 2004) and encodes 4718 protein coding genes. The complete genome sequence and the annotation provide the basis to move from the genomic to the transcriptomic era of the filamentous fungus *A. gossypii*. The use of microarrays is the method of choice to monitor the transcriptional profile at distinct time points in the *A. gossypii* life cycle and to detect changes in gene expression due to transcription factor deletion. Recently, several genome scale transcription studies in over 20 species of filamentous fungi encompassing a wide variety of research areas have been reviewed (Breakspear *et al.*, 2007). Most of these studies use “traditional“ cDNA chips that have either been constructed using PCR products from expressed sequence tag (EST) libraries or include all predicted open reading frames (ORFs). cDNA chips are simple and cost effective but in order to produce reliable results experienced laboratory staff is required during the chip manufacturing and hybridisation process. State-of-the-art oligonucleotide chips from Affymetrix, considered to be the most reliable platform, have been only used in studies with two filamentous fungi, Aspergillus niger (MacKenzie *et al.*, 2005) and Fusarium graminearum (Guldener *et al.*, 2006). In collaboration with Syngenta and Affymetrix a custom-made oligonucleotide array for *A. gossypii* was designed that covers 94% of the 4718 genes and 190 intergenic or antisense regions. This chapter aims at elucidating the details of the array and describes the experimental procedures from sample preparation to gene chip hybridisation and scanning. Much effort has gone into quality control of the arrays and normalization strategies which will be discussed in the last part of this chapter.

1.2 General introduction to DNA Microarray technology

Basic concept

Microarray technology represents a powerful functional genomics technology, which permits the expression profiling of thousands of genes in parallel (Schena, 1996).
Microarrays works by exploiting the ability of a given mRNA molecule to bind specifically (=hybridize) to the DNA template it originated from. The principle to analyze nucleotides with the help of other labeled nucleotides using hybridisation is an old concept and was introduced by Edwin Southern (E. Southern, 1975).
DNA Microarrays are small, solid supports (such as glass, silicon or nylon membrane) onto which the sequences from thousands of different genes (=probes) are immobilized at distinct locations. The probes are hybridized with labeled nucleotides, named targets. Total RNA that was isolated from biological samples of interest serves as template for the synthesis of the single stranded targets. Hybridisation takes place during a distinct period of time and target molecules that do not bind to one of the probes are washed away. The amount of hybridized target molecule, measured by fluorescence intensity, is proportional to the amount of isolated mRNA. It is important to note that Microarrays do not measure absolute, but relative abundances.

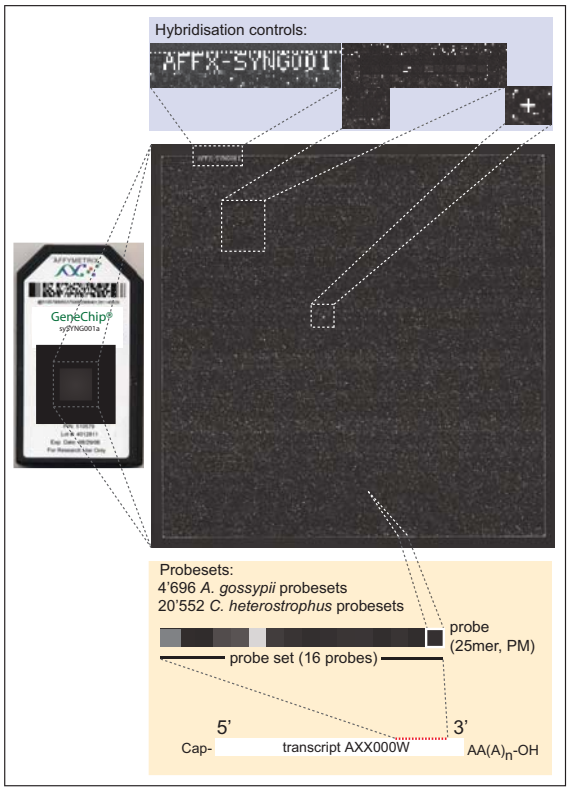


Figure 1-1 *A. gossypii* sySYNG001a chip design. The hybridization controls include from left to right: the name of the chip, an alternating pattern of intensities at the chip border, a region with spiked-in controls in staggered concentrations and a cross in the center. The chip contains 4'696 *A. gossypii* and 20'552 *C. heterostrophus* probe sets. Each probe set consists of 16 perfect match 25mers that match the 3' end of the gene.

Types of DNA Microarrays

Different types of DNA arrays are designed for mRNA profiling. These differ by the type of probes that they carry: cDNA probes (0.1 to 5 kb long) that are conventionally synthesized and immobilized on the chip, or oligonucleotides (20 – 80mers) that are synthesized directly *in situ*. As oligonucleotide microarrays were used for this investigation, this technology will be explained here in more detail. The high density oligonucleotide microarrays from Affymetrix typically include nowadays 11 probes (or probe pairs) per transcript while older formats included up to 20 probes (or probe pairs) per transcript. Chip designs that include for each perfect match (pm) probe a probe that carries a mismatch (mm) in the middle of the sequence are also available on the market. The more recent tiling arrays are designed to contain the entire DNA sequence without prior consultation of existing gene annotation. They enable the discovery of novel transcribed sequences, e.g. regulatory RNA genes, through the unbiased interrogation of genomic loci.

Affymetrix array manufacturing

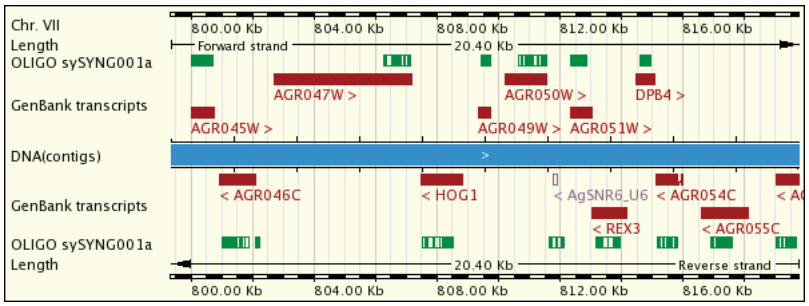
Manufacturing of Affymetrix arrays happens via a photolithographic process. The 25mer oligonucleotides are synthesized *in situ*. The manufacturing process begins by coating a quartz wafer with a light-sensitive chemical compound that prevents coupling between the wafer and the first nucleotide of the DNA probe being created. Lithographic masks are used to either transmit or block light onto specific locations of the wafer surface. The surface is then flooded with a solution containing either adenine, thymine, cytosine, or guanine. Coupling occurs only in those regions on the glass that have been deprotected through illumination.

The sySYNG001a Affymetrix chip design

The Affymetrix oligonuceotide arrays used in this study were custom made (GeneChip® CustomExpress® Array Program, Affymetrix, Santa Clara, USA). The dimension of the chip is 640 x 640 features with 4'696 A. gossypii and 20'552 C. heterostrophus probe sets. Each *A. gossypii* probe set consists of 16 antisense probes, each 25 nucleotides long, matching a region within 600 bases from the 3' end of the open reading frame. Figure 1-1 summarizes the fundamentals of the chip design. Physical genomic positions of probes (figure 1-2) are viewed in the Ashbya genome database AGD 3.0 (Gattiker *et al.*, 2007). The 4696 *A. gossypii* probe sets represent 4452 genes of which 130 do not have a homolog in baker's yeast (NOHBYs). These

are thought to be potential Ashbya-specific genes when compared with *S. cerevisiae*. When compared with other fungi 40 NOHBYs are Ashbya-specific. 190 probe sets are intergenic or sense probe sets that will not be extensively discussed in this work (table 1-1). Affymetrix uses for each probe set an identifier of the pattern 'Ag00XXXX_at'. Probe sets that have to be used with caution because they are unspecific follow the pattern 'Ag00XXX_*_at'. 31 genes, some of which have particularly long open reading frames (e.g. SPA2, DYN1), are represented by more than one probe set (table 1-2). The 5' to 3' ratios of intensity values of these genes are assessed to check the quality of the ds cDNA and cRNA synthesis steps and RNA integrity.

A



B

Feature Information

Location	Length	Name(s)	Mismatches
VII:809428-809452(1)	25	sySYNG001a:Ag004168_at:145:388;	0
VII:809441-809465(1)	25	sySYNG001a:Ag004168_at:157:215;	0
VII:809467-809491(1)	25	sySYNG001a:Ag004168_at:517:126;	0
VII:809481-809505(1)	25	sySYNG001a:Ag004168_at:568:110;	0
VII:809484-809508(1)	25	sySYNG001a:Ag004168_at:328:345;	0
VII:809519-809543(1)	25	sySYNG001a:Ag004168_at:201:266;	0
VII:809547-809571(1)	25	sySYNG001a:Ag004168_at:607:84;	0
VII:809556-809580(1)	25	sySYNG001a:Ag004168_at:115:502;	0
VII:809564-809588(1)	25	sySYNG001a:Ag004168_at:32:115;	0
VII:809573-809597(1)	25	sySYNG001a:Ag004168_at:441:185;	0
VII:809575-809599(1)	25	sySYNG001a:Ag004168_at:505:246;	0
VII:809624-809648(1)	25	sySYNG001a:Ag004168_at:142:175;	0
VII:809626-809650(1)	25	sySYNG001a:Ag004168_at:81:100;	0
VII:809681-809705(1)	25	sySYNG001a:Ag004168_at:418:538;	0
VII:809693-809717(1)	25	sySYNG001a:Ag004168_at:294:91;	0
VII:809712-809736(1)	25	sySYNG001a:Ag004168_at:414:626;	0

Figure 1-2 (A) Physical genomic positions of the probe sets (green boxes) of the sySYNG001a Affymetrix chip are viewed in AGD3.0 (Gattiker et al., 2007). The probe sets, containing 16 probes each, were designed to match a region within 600 bases from 3' end of the open reading frames (red boxes). (B) Genomic coordinates and strand, length, name and mismatch number of oligonucleotide probes belonging to a probe set (e.g. Ag004168_at) can be queried in AGD3.0. The probe set shown here matches the cell wall protein gene AGR049W (CCW12). Due to the shortness of the gene probes are partially overlapping, which is not the standard design.

Table 1-1 Overview of probe sets and genes represented on the *A. gossypii* sySYNG001a Affymetrix chip.

PROBESETS	
total *A. gossypii* antisense probesets (*_at)	4696
probe sets matching *A. gossypii* genes	4499
sense probe sets and probe sets matching intergenic regions	190
probe sets matching other genomes	7
probe sets matching nuclear chromosomes	4686
probe sets matching mitochondrial chromosome	9
specific probe sets	4653
unspecific probe sets (_*_at)	43
not all probes of probe set are unique (_r_at)	22
fewer than the required 16 unique probes (_i_at)	13
all probes exactly match multiple transcripts (_s_at)	6
probe set recognizes more than one gene (_f_at)	2

GENES (total 4718 annotated genes)	
total *A. gossypii* genes represented on chip	4452
genes represented by one probe set	4419
genes represented by two probe sets	29
genes represented by three probe sets	2
A. gossypii genes with homologs in *S.cerevisiae*	4322
A. gossypii specific genes (NOHBYs)	130

Affy ID	Systematic name	Common name(s)		Relative probe set location
Ag004455_at	AAR023C	NOHBY106		3 prime
Ag004456_r_at	AAR023C	NOHBY106		5 prime
Ag000295_at	AAR069W			3 prime
Ag000067_at	AAR069W			5 prime
Ag004462_at	AAR087C			3 prime
Ag004463_at	AAR087C			5 prime
Ag000033_at	AAR125C	DFM1		3 prime
Ag000330_at	AAR125C	DFM1		5 prime
Ag000488_at	ABL066C	MDM31		3 prime
Ag000423_at	ABL066C	MDM31		5 prime
Ag000503_at	ABR111C			3 prime
Ag000504_at	ABR111C			5 prime
Ag004485_at	ABR185W	PRD1		3 prime
Ag004484_at	ABR185W	PRD1		5 prime
Ag000371_at	ABR231W			3 prime
Ag000424_at	ABR231W			5 prime
Ag001179_at	ACR006C			3 prime
Ag001180_at	ACR006C			5 prime
Ag000918_at	ACR023W	KAR9		3 prime
Ag000805_at	ACR023W	KAR9		5 prime
Ag004507_at	ACR077W			3 prime
Ag004506_at	ACR077W			5 prime
Ag001213_at	ACR258W	DYN1		3 prime
Ag000870_at	ACR258W	DYN1		5 prime
Ag001795_at	ADL022C	SPA2	SPH1	3 prime
Ag001845_at	ADL022C	SPA2	SPH1	middle
Ag001803_at	ADL022C	SPA2	SPH1	5 prime
Ag001972_at	ADL377W	PTC4		3 prime
Ag001917_at	ADL377W	PTC4		5 prime
Ag003707_f_at	ADL397C	HSP26		3 prime
Ag001934_f_at	ADL397C	HSP26		5 prime
Ag001791_at	ADR198C	GIS4		3 prime
Ag001706_at	ADR198C	GIS4		5 prime
Ag001718_at	ADR200C	HKR1		3 prime
Ag001717_at	ADR200C	HKR1		5 prime
Ag004578_at	AEL023C	MUC1		3 prime
Ag002225_at	AEL023C	MUC1		5 prime
Ag002653_at	AER029C	ATG19		3 prime
Ag002301_at	AER029C	ATG19		5 prime
Ag003079_r_at	AFL087C	RCS1	AFT2	3 prime
Ag002859_at	AFL087C	RCS1	AFT2	5 prime
Ag003118_at	AFL092C	FLO5		3 prime
Ag002944_at	AFL092C	FLO5		5 prime
Ag003039_at	AFR130W	SSP1		3 prime
Ag003013_at	AFR130W	SSP1		5 prime
Ag003428_at	AFR644C	RCY1		3 prime
Ag002930_at	AFR644C	RCY1		5 prime
Ag003165_at	AFR739C			3 prime
Ag002928_at	AFR739C			5 prime
Ag004141_at	AGL022W	BLM3		3 prime
Ag004143_at	AGL022W	BLM3		5 prime
Ag003747_at	AGL058C	NOHBY706		3 prime
Ag003799_at	AGL058C	NOHBY706		5 prime
Ag004073_at	AGL209W			3 prime
Ag004071_at	AGL209W			5 prime
Ag004696_at	AGL300C	FLO8		3 prime
Ag003742_at	AGL300C	FLO8		middle
Ag004695_at	AGL300C	FLO8		5 prime
Ag003832_at	AGL306C	BUD4		3 prime
Ag003967_r_at	AGL306C	BUD4		5 prime
Ag004157_at	AGR197C	RPS6A	RPS6B	3 prime
Ag004659_at	AGR197C	RPS6A	RPS6B	5 prime
Ag004135_at	AGR313W			3 prime
Ag003829_at	AGR313W			5 prime

Table 1-2 31 *A. gossypii* genes are represented by more than one probe set on the sySYNG001a Affymetrix chip. Please note that the locations are given relative to each other: out of the two (or three) probe sets per gene, the 3 prime probe set is per definition the one that lies closest to the 3 prime end of the gene (this does not imply that it is positioned in absolute terms at the 3 prime end of the gene). An analog rule applies for the middle and 5 prime probes.

1.3 Sample preparation and collection

Sampling of life cycle stages for transcriptional profiling

The early life cycle stages of germinating fungus *A. gossypii* spores (0 h, 2 h, 5 h, 7 h, and 9 h) and the advanced mycelium stage (18 h) were assayed in liquid full medium (AFM). To monitor transcriptional changes occurring later in the life cycle, mycelia were grown on plates as radially expanding large colonies. This enabled us to separate the zone of very fast hyphal elongation showing repeated tip splitting at the border of the colony (OZ = Outer Zone) from the two to three days older zone of sporulation in the center of the colony (IZ = Inner Zone). Expression of a deletion mutant (Ag*dig1/2*Δ) with a sporulation defect was also assayed on plates. The other two transcription factor deletions (Ag*ste12*Δ, Ag*tec1*Δ) were assayed when germlings had developed one or two hyphae (9 h liquid medium). An overview of the sampling method is given in figure 1-3, and table 1-3 at the end of this chapter summarizes experimental conditions and protocols that were used for the 34 *A. gossypii* arrays.

Strains

The *A. gossypii* laboratory strain Agleu2Δthr4Δ (derivative of *A. gossypii* wild type strain #ATCC10895; figure 1-4 A) was used throughout this study. Additionally, several life cycle stages were also tested in a strain that was isolated in 2005 from a cotton stainer in Florida (Fred Dietrich, personal communication; figure 1-4 B). Cotton stainers are insects that feed on cotton seed fibers and transmit *A. gossypii* spores from cotton plant to cotton plant.

Preparation of a highly purified spore suspension

High amounts of pure spores, free from vegetative mycelial contamination, were required as starting material for all experiments described in this thesis. We used a protocol that allows for the preparation of highly purified spore suspensions (Figure 1-4 C) without enzymatic digestion (S.Brachat, PhD thesis). Spores were collected making use of their hydrophobic properties. For this purpose 5 x 10^4 spores were grown for 8 days on Ashbya full medium (AFM) agar plates. The AFM composition was as follows: 10g/l Bacto Peptone (Pancreatic Digest of Casein, Difco), 10g/l Yeast Extract (Micro Granulated, Formedium, Norwich, England) 1g/l Myo-inositol (Merck), 2% D(+)-Glucose-Monohydrate (Merck). 15g/l agar was added to solidify the medium (Formedium, Norwich, England). Glucose was autoclaved separately and Myo-inositol was filter sterilized and added to the medium prior to use.
In the first step of the spore preparation procedure, a silicon layer was covalently attached to the surface of test tubes with screw caps. Then 250 ml of Sigmacote (Sigma-Aldrich, CAT# SLS-100 ml) was added per glass test tube. The glass tubes were inoculated at room temperature in a slowly turning wheel over night. In the morning the tubes were autoclaved. 15 ml of sterile H20 and the mycelium coming from one plate were added per glass tube. A microscopy slide sterilized with 70% ethanol was used to scratch the mycelia from the agar plates. The tubes were vortexed and shaken vigorously to break the mycelium. To let the released spores stick to the glass wall, the tubes were incubated for 2 hours in a turning wheel. The water was then discarded and the tube was rinsed 3 x with pure water to remove residual pieces of mycelium. 0.1% Triton was added and the tubes were vigorously shaken and vortexed to wash the spores from the glass wall. The spore suspensions were transferred to 15 ml Falcon tubes. The glass tubes were rinsed once again with 5 ml 0.1% Triton to maximize the spore yield. The spores were collected by centrifugation at 3000 rmp for 5 min at room temperature. The supernatant was decanted, the spores were resuspended in 1 ml 0.03% Triton and transferred to 2 ml Eppendorf tubes. Three washing steps with 0.03% Triton were carried out. After the last washing step the Triton was completely sucked off by vacuum pump and the spores were resuspended in 500 ml 30% glycerol per spore prep. All spore preps were checked by light microscopy for the absence of mycelial contamination. Spores were incubated for 1h at room temperature to allow glycerol uptake and then frozen at -80°C.

Sample collection from liquid culture

For experiments in liquid culture, 1.6 x 10^7 spores (collected from 5 agar plates) were thawn for 10 min and inoculated per 10 ml Ashbya full medium (AFM) in 100 ml baffle flasks (Duran, Schott, Germany). For isolation of advanced mycelium spores were incubated for 18 h in 200 ml AFM and 500 ml baffle flasks at 30°C at 150 rpm. The AFM composition was as follows: 10g/l Bacto Peptone (Pancreatic Digest of Casein, Difco), 10g/l Yeast Extract (Micro Granulated, Formedium, Norwich, England) 1g/l Myo-inositol (Merck), 2% D(+)-Glucose-Monohydrate (Merck). Glucose was autoclaved separately and Myo-inositol was filter sterilized and added to the medium prior to use.
The cultures were inoculated at 30°C on a shaker at 150 rpm. Previous to sample collection we checked the developmental stages and the absence of contamination by light microscopy. 500 ml aliquots were removed and fixed with formaldehyde. These samples were stored for later inspection. Sample collection was done via filtration because this method allows collection of all *A. gossypii* life cycle stages

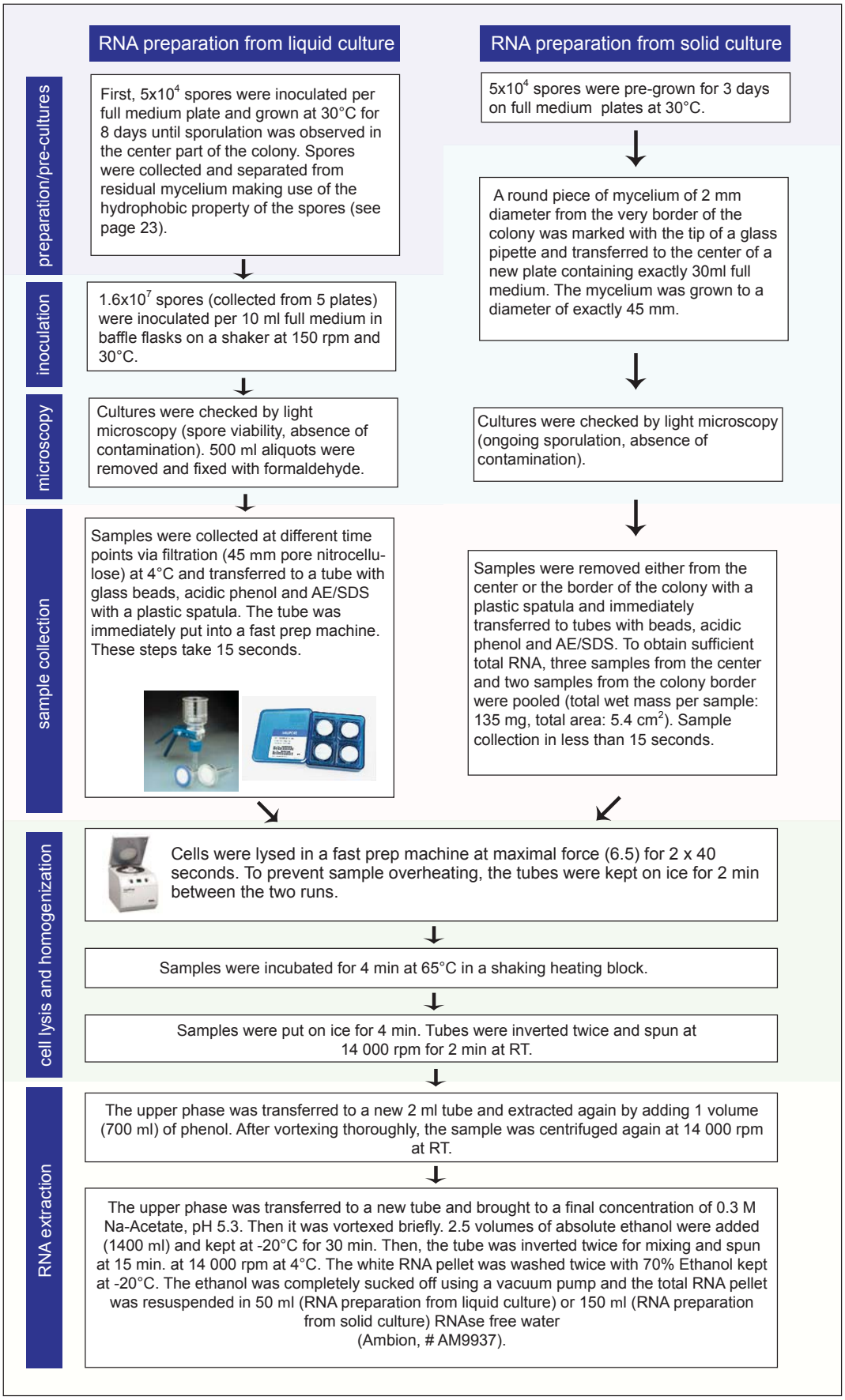


Figure 1-3 Overview of sampling of life cycle stages and RNA isolation for transcriptional profiling

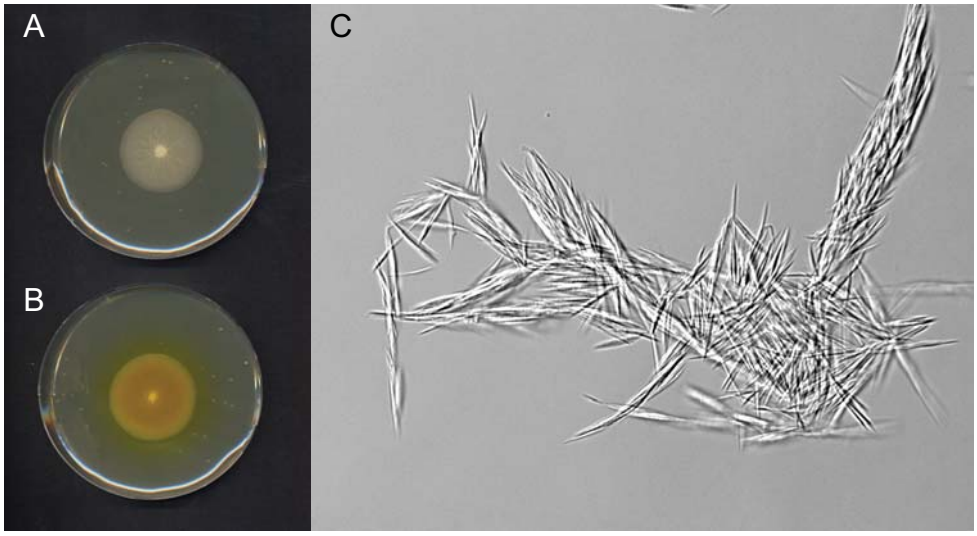


Figure 1-4 (A) Three days old *A. gossypii* laboratory strain ΔIΔt (B) three days old *A. gossypii* natural isolate strain FDAG, the yellow colour comes from the riboflavin (vitamin B2) (C) clean spore preparation devoid of vegetative mycelial contamination

from liquid culture including mycelia which can not be harvested by centrifugation. A 47 mm glass filter holder kit (Millipore, CAT# XX15 047 00) was used for 10 ml cultures and a 90 mm glass filter holder kit (Millipore, CAT# XX1009020) was used for 200 ml cultures. Filters were 45 mm or 90 mm 0.45 μm pore size nitrocellulose (Millipore, CAT# HAWP04700 and CAT# HAWP09000). First, filters were moistened with RNAse free water (DEPC-treated) and then the culture was filtered, the biomass was immediately transferred with a plastic spatula (240 x 15 mm, Semadeni AG, Switzerland) to a tube preloaded with glass beads, ice-cold acidic phenol and AE/SDS (see total RNA preparation) and put into a FastPrep machine for lysis. These steps took 15 seconds.

Sample collection from solid culture

Sampling of late stages in the Ashbya life cycle were done on plates which allows the simultaneous collection of hyphae at the border of the colony that grow at maximal speed (160-180 μm/h) and hyphae in the center of the colony that undergo sporulation. For experiments in solid medium 5 x 10^4 spores were pre-grown for 3 days at 30°C on Ashbya full medium (AFM), solidified with 15 g/l agar (Formedium, Norwich, England). The pre-incubation step was used to make sure that the starting mycelium was growing with the same maximal speed in all conditions. A round piece of mycelium of 2 mm diameter from the very border of the colony was transferred to the center of a new plate that contained 30 ml solid AFM. Several plates were prepated in the same way. The mycelia were grown as previously described to a diameter of 45 mm. To obtain sufficient total RNA two samples from the colony border and three samples from the center of the colony were pooled. The total wet mass per sample was 135 mg collected from 5.4 cm^2. The border of the colony was marked with the bottom part of a 35 x 10 mm Petri dish (Falcon 1000, Becton Dickinson Labware, New Jersey, USA) to alleviate collection of the outermost 2 mm of the colony and prevent cross contamination from other zones of the colony (Figure 1-5 B). The mycelial sample was harvested with a plastic spatula (240 x 15 mm, Semadeni AG, Switzerland). A region from the center of the colony was marked with a custom-made 2 cm diameter polypropylene ring (Figure 1-5 A). The very center of the colony was omitted and a 4 mm wide Ring of sporulating mycelium was removed with a 10 ml plastic loop (Nunc Brand Products). The mycelial samples were immediately transferred to a tube preloaded with glass beads, ice-cold acidic phenol and AE/SDS and put into the FastPrep machine for lysis. These steps took 15 seconds.

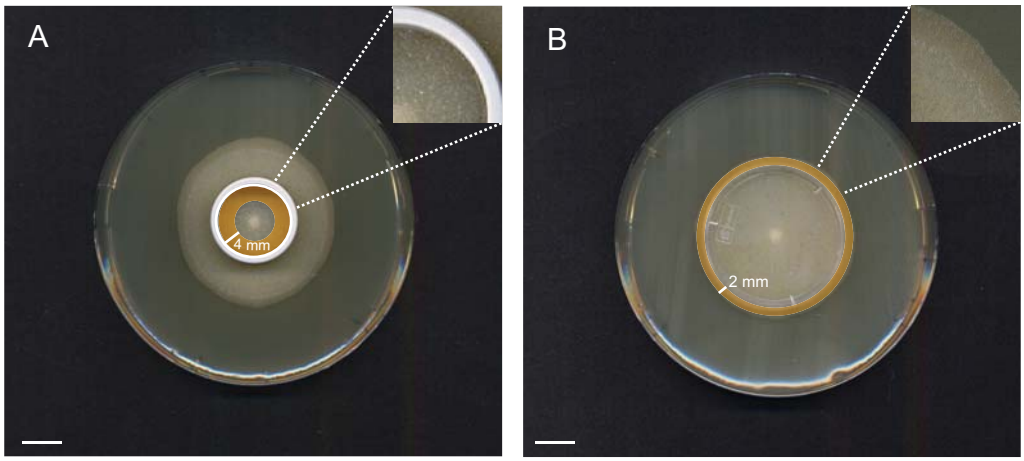


Figure 1-5 Sampling of late stages in the *A. gossypii* life cycle was done from plates because this permits the separation of (A) hyphae in the center of the colony that undergo sporulation from (B) hyphae at the border of the colony that grow at maximal speed (200 µm/h). Annular zones of the mycelia that were used for transcriptional profiling are highlighted in orange. For increased RNA yield fungal material was pooled from two (A) or three (B) plates for one sample. Total wet mass, removed from 5.5 cm^2, was 135 mg. The scale bar represents 10 mm.

Replicates

Replicates are repeated measurements with the same sample that provide a measure of the experimental variation. Replicate measurements are important in the data mining process where they increase the statistical power and reduce noise. There are two types of replicates: technical replicates and biological replicates. Technical variation in an array experiment comes from the quantity and quality of the labeled RNA hybridized as well as differences in reagent concentrations and chip handling. When using the Affymetrix platform technical replicates are obtained by hybridizing the same labeled cRNA to two chips. Due to the high reliability and reproducibility of the Affymetrix platform the necessity for technical replicates has become decreasingly important. Biological variation that is irrelevant to the study may arise from differences in spore preparation, culture conditions, strain background, time etc. In our study two biological replicates per sample were generated, starting from two independent cultures. The germination experiment described in chapter 2 and the experiment with fast growing hyphae in chapter 3 were repeated in an *A. gossypii* natural isolate, so that in total four measurements in two genetic backgrounds per time point or zone are available.

1.4 RNA preparation, cRNA target synthesis and gene chip hybridisation

Total RNA preparation

RNA was isolated using a hot phenol method (Schmitt *et al.*, 1990) with some modifications for *A. gossypii* (P. Demougin, U. Schlecht and R. Rischatsch, unpublished data).
The modified method yields high quality RNA even when the extractions were done under difficult conditions, e.g. extraction from spores. The FastRNA Pro Red Kit (Bio101 Systems, CAT# 6035-050) uses a similar phenol/chloroform extraction protocol except that samples are not heated and that the glass bead size is larger than 0.5 mm. The kit worked well for RNA extraction from advanced mycelium but failed to give high quality RNA from germlings and spores. Due to the limited amount of spore material

we did not try RNA extraction protocols that need pre-grinding of the samples under liquid nitrogen and transfer to columns.
All solutions, tubes and other equipment was RNAse free. Powder-free gloves were worn at all times. All tubes were preloaded with solutions and labeled in advance. To ensure fast handling we proceeded tube by tube from the sample collection step until the incubation step at -20°C for RNA precipitation.
700 ml of ice-cold acidic phenol (AppliChem, CAT# A00470500, water-saturated, stabilized) was loaded into a tube. 0.5 mm glass beads pretreated with concentrated nitric acid (Biospec products, CAT# 11079105, Bartlesville, USA) were added up to half the volume of phenol and the tube was kept on ice. Shortly before starting the experiment 630 ml AE buffer (50 mM NaAc, 10 mM EDTA) were mixed with 70 ml 10% SDS in a separate tube and then carefully added on top of the phenol. *A. gossypii* cells were collected as previously described and transferred with a plastic spatula and loop to the tube in less than 10 seconds. The tube was vigorously shaken three times and then very quickly placed in the Fast-prep® homogenizer (Bio101, Illkirch, France). The settings were 6.5 (max. speed) for 2x 40 s. In between the runs, the tube was cooled on ice for 2 min to prevent overheating and degradation of RNA. The sample was then incubated for 4 min at 65°C in a shaking (1000 rpm) heating block followed by an incubation of 4 min on ice. The tube was inverted twice and centrifuged for 2 min at full speed (>12 000 rpm) at room temperature. The upper phase was transferred to a new 2 ml tube, extracted again by adding 700 ml of phenol and vortexed. The tube was again centrifuged for 2 min at full speed (>12 000 rpm) at room temperature. The upper phase was carefully transferred to a new 1.5 ml tube preloaded with 60 ml of 3M Na-Acetate (pH 5.3). It is very important not to transfer any phenol. The tube was vortexed briefly and 2.5 vol. of absolute ethanol (1400 ml) kept at -20°C were added. The tube was vortexed and kept in the freezer at -20°C for 30 min The tube was inverted twice for mixing and centrifuged for 15 min at full speed (>12 000 rpm) at 4°C. The pellet was washed twice with 70% ethanol kept at -20°C. Remaining ethanol was removed by vacuum pump and the pellet was resuspended in 50 ml (liquid samples) or 150 ml (solid samples) RNAse free water (Ambion, CAT# AM9937). The A260/A280 ratio was measured to assess the level of protein contamination. Total RNA was stored at -80° C. Cleanup of 40 mg total RNA was done on RNeasy columns (Quiagen) including DNAse digestion. The RNA concentration was determined by OD260 and the purity by OD260/280 with the spectrophotometer (NanoDrop, ND-1000, Wilmington, USA). 1 ml of the RNA solution was monitored on a RNA Nano 6000 Labchip in the Agilent 2001 Bioanalyzer (Agilent technologies, USA).

Characterization of RNA quality using the Agilent 2100 Bioanalyzer

RNA quality is critical for gene expression measurements using microarray technology. The Agilent bioanalyzer (Agilent technologies, USA) and the lab chip kit provide a simple, reliable and rapid method to assess the concentration and purity/integrity of the RNA. The lab chip is a miniaturization of RNA capillary electrophoresis. We ran 1 μl of each total RNA solution on a Eukaryote Nano Lab chip. Electropherograms and gels are shown for all RNA extractions used in this investigation in Figure 1-6 A-D. The predominant features of intact total RNA samples are the 18S and 25S ribosomal RNA peaks. A smaller and broader peak at lower migration times which includes the 5.8 and 5s ribosomal RNAs and tRNAs is also present (see ΔlΔt 0 h and 2 h samples). These peaks disappear after total RNA cleanup on columns (see ΔlΔt 5 h to 9 h samples). Good RNA samples are characterized by the absence of smaller well-defined peaks between the two main ribosomal peaks. The baseline between 29 seconds running time and the 18S ribosomal peak is flat. RNA quality is usually determined by quantification of the 25S to 18S ribosomal RNA ratio (Russel, 2001) that is sample dependent and should be close to 2. More recent publications propose other means to assess RNA quality (Copois *et al.*, 2007). RNA degradation (an example is given in figure 1-6 D) produces a shift in the RNA size distribution towards smaller fragments. A decrease in fluorescence signal is observed and discrete bands at lower running times become visible.

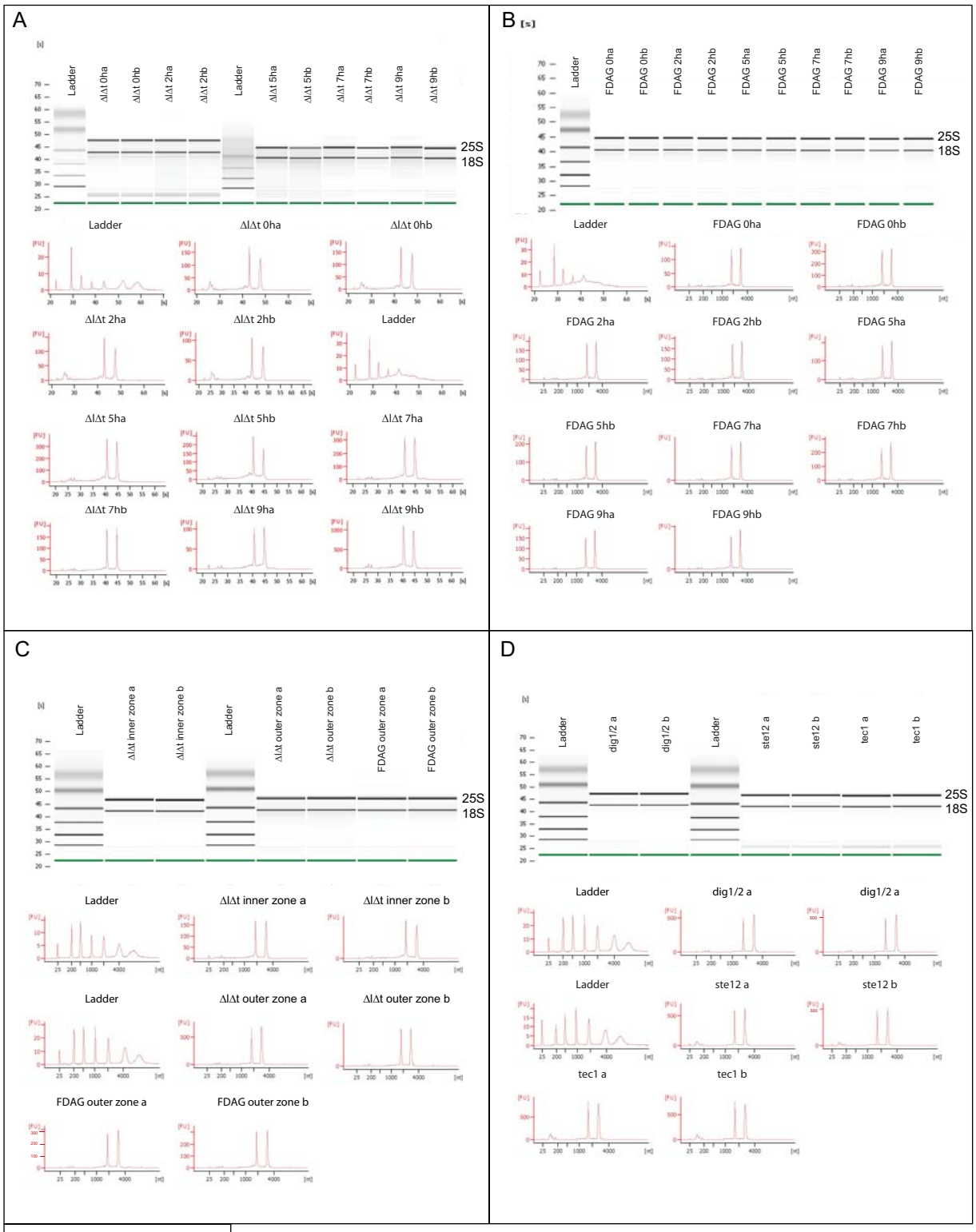
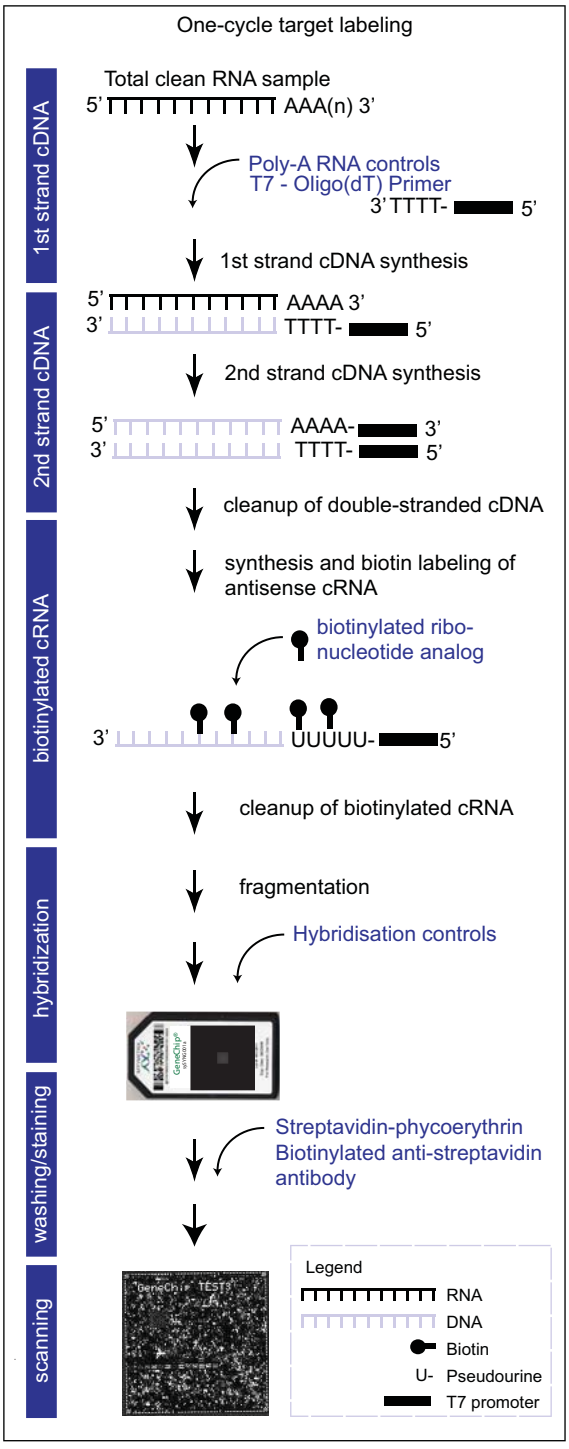


Figure 1-6 RNA quality is critical for gene expression measurements using microarray technology. Electropherograms of total RNA are shown. (A) RNA samples from germinating spores of the lab strain (ΔlΔt) before cleanup (ΔlΔt 0 ha - ΔlΔt 2 hb) and after cleanup (ΔlΔt 5 ha - ΔlΔt 9 hb). (B) RNA samples from germinating spores of a new natural isolate (FDAG) (C) RNA samples isolated from plates (D) RNA isolated from deletion mutants (E) a partially degraded RNA sample of ΔlΔt 9 h is shown for comparison. The sample was not used for downstream experiments.

cRNA Target Synthesis

Biotin labeling of RNA was performed as described in the GeneChip Expression Analysis Technical Manual (Affymetrix, Santa Clara, USA). Double-stranded cDNA was synthesized according to the One-Cycle cDNA Synthesis Kit (Affymetrix, CAT# 900431), starting from 2 µg of total RNA. The material was purified with the Sample Cleanup Module (Affymetrix, CAT# 900371). A graphical overview of the one-cycle target labeling assay is given in figure 1-7.
The purified cDNA was used for an in vitro transcription reaction by using the IVT labeling kit (Affymetrix, CAT# 900449) to synthesize cRNA in the presence of a biotin-conjugated ribonucleotide analog. 75 µg of labeled cRNA from each reaction was purified by using the Sample Cleanup Module, and average size of the cRNA molecules was assessed on RNA Nano 6000 Chips in the 2100 Agilent Bioanalyzer (1-8 B). The cRNA targets were incubated at 94°C for 35 min in the provided fragmentation buffer and the resulting fragments of 50–150 nucleotides were again monitored using the Bioanalyzer (figure 1-8 C). All synthesis reactions were carried out using a PCR machine (T1 Thermocycler; Biometra, Göttingen, Germany) to ensure a high degree of temperature control.

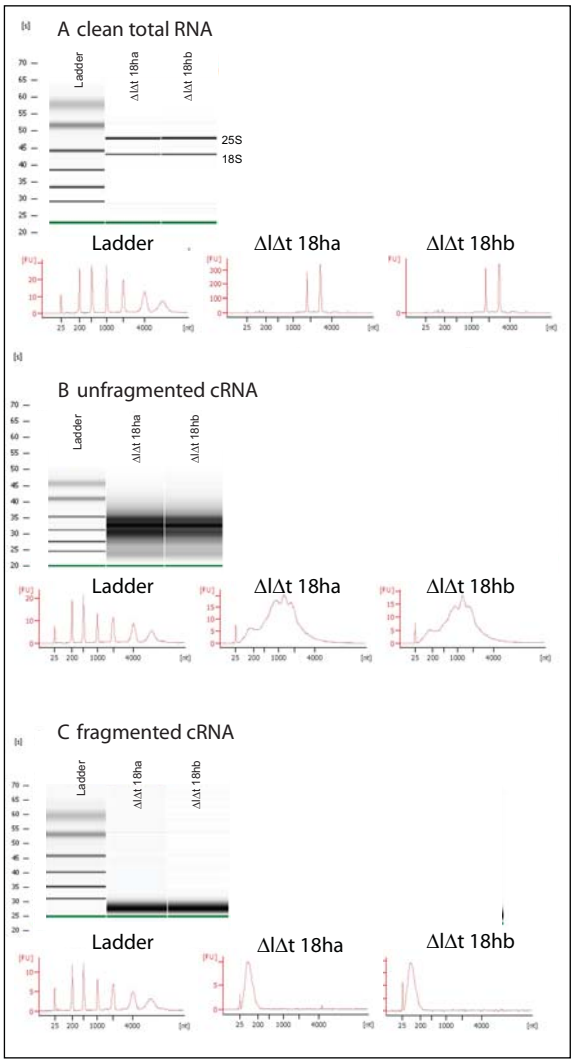


GeneChip Hybridization

The hybridization cocktail (200 µl) containing fragmented biotin-labeled target cRNA at a final concentration of 0.05 µg/µl was

Figure 1-7 Experimental outline of the one-cycle eukaryotic target labeling assay (modified from GeneChip® Expression Analysis Technical Manual, 2007). 2 µg of total clean RNA is first reverse transcribed using a T7-Oligo(dT) Promoter Primer in the first strand cDNA synthesis reaction. Following RNAse H-mediated second-strand cDNA synthesis, the double-stranded cDNA is purified and serves as a template in the subsequent in vitro transcription (IVT) reaction. The IVT reaction is carried out in the presence of T7 RNA Polymerase and a biotinylated nucleotide analog/ribonucleotide mix for complementary RNA (cRNA) amplification and biotin labeling. The biotinylated cRNA targets are then cleaned up, fragmented, and hybridized to GeneChip expression arrays. Note that the steps that lead to the production of cRNA have two advantages. First, the isolated mRNA will be amplified during transcription from cDNA to cRNA which enables the detection of signals that are based on very low numbers of expressed mRNA molecules. Secondly, RNA/DNA hybridization has a higher sensitivity compared to DNA/DNA hybridization.

transferred into Affymetrix sySYNG001a and incubated at 45°C on a rotator in a hybridization oven 640 (Affymetrix) for 16 h and 40 min at 60 rpm. The arrays were washed and stained on a Fluidics Station 450 (Affymetrix) by using the Hybridization Wash and Stain Kit (Affymetrix, CAT# 900720). To increase the signal strength the antibody amplification protocol was used (FS450_0004). The GeneChips were processed with an Affymetrix GeneChip® Scanner 3000 7G (Affymetrix) by using the current default settings. DAT image files of the microarrays were generated using GeneChip Operating Software (GCOS 1.4; Affymetrix).



1.5 Quality controls and normalization

Assessing quality of the hybridization step by visual inspection

Quality control of the hybridisation step included first checking for proper hybridisation of the control regions and second checking for the absence of spatial effects.
A hybridisation control oligo B2 is spiked into each hybridisation cocktail. Hybridisation of B2 is highlighted as alternating pattern of intensities at the border of the array, the array name in the upper left corner and a cross in approximately the center of the chip. BioB, bioC and bioD and cre are additional hybridization controls prepared in staggered concentrations (1.5 pM, 5 pM, 25 pM, and 100 pM final). BioB, bioC and bioD represent genes in the biotin synthesis pathway of *E. coli*. Cre is the recombinase gene from P1 bacteriophage. Hybridisation

Figure 1-8 During the experimental steps that lead from total mRNA to labeled and fragmented cRNA (used for chip hybridization), several electropherograms are run to ensure high quality RNA. The total RNA shown here comes from biological replicates of mycelium grown for 18 h in full medium. The synthesis steps are mRNA -> cDNA -> cRNA. The cDNA is not present in sufficient amounts to be assayed. Successfully amplified cRNA ranges in length from 50 – 4000 nt. Large transcripts often contain secondary structure, which can interfere with hybridization and increase the opportunity for non-specific cross-hybridisation. To obtain a good target for hybridization, cRNA is fragmented to a size of 50 – 200 bases long. Fragmentation helps to improve specificity and raises the feature signal intensity on the array.

control areas on the sySYNG001a chip are shown in figure 1-1.
Due to the, per design, (mostly) random distribution of *C. heterostrophus* and *A. gossypii* probes on the chip, non-random intensity distributions with visible patterns might indicate problems during hybridisation (scratches, particles in hybridisation cocktail etc.). The 'image' function (T. Lumley) in R uses the CEL files (containing pm information) to calculate pseudo-images of the chip that can be saved in pdf or jpeg format. Image plots of all 34 arrays passed the quality control step of visual inspection.

RNA digestion plots

RNA digestion plots are used to assess uniform treatment of samples, RNA integrity and quality of the ds cDNA and cRNA synthesis steps (figure 1-10). For each array and within each probe set, the probes are ordered by their proximity to the 5' end of the gene. The average intensity of the probes classified by this order is shown in the plots. Each array is represented by a line. The slopes and p-values from linear regressions are shown in the lists. Typically

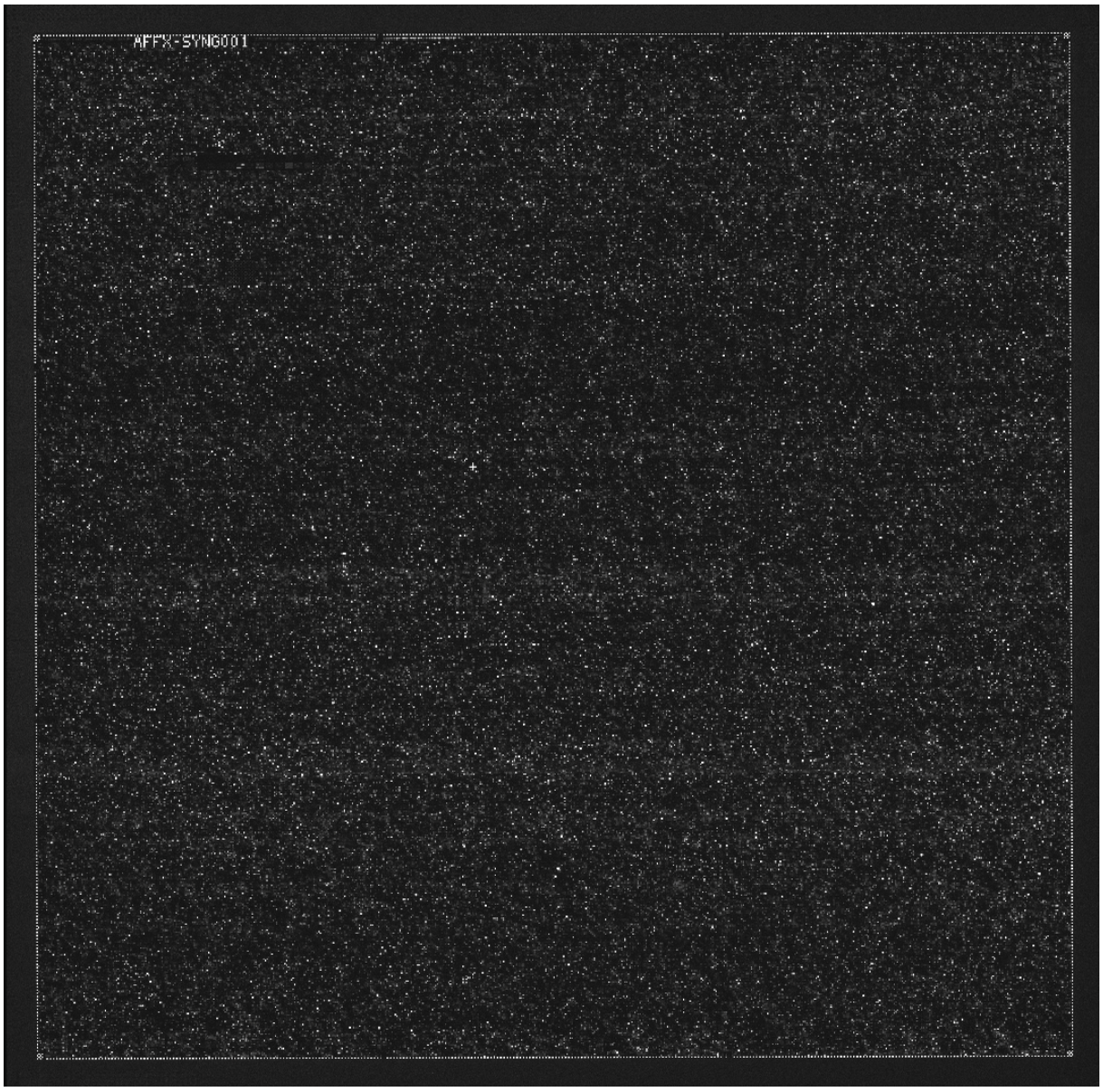
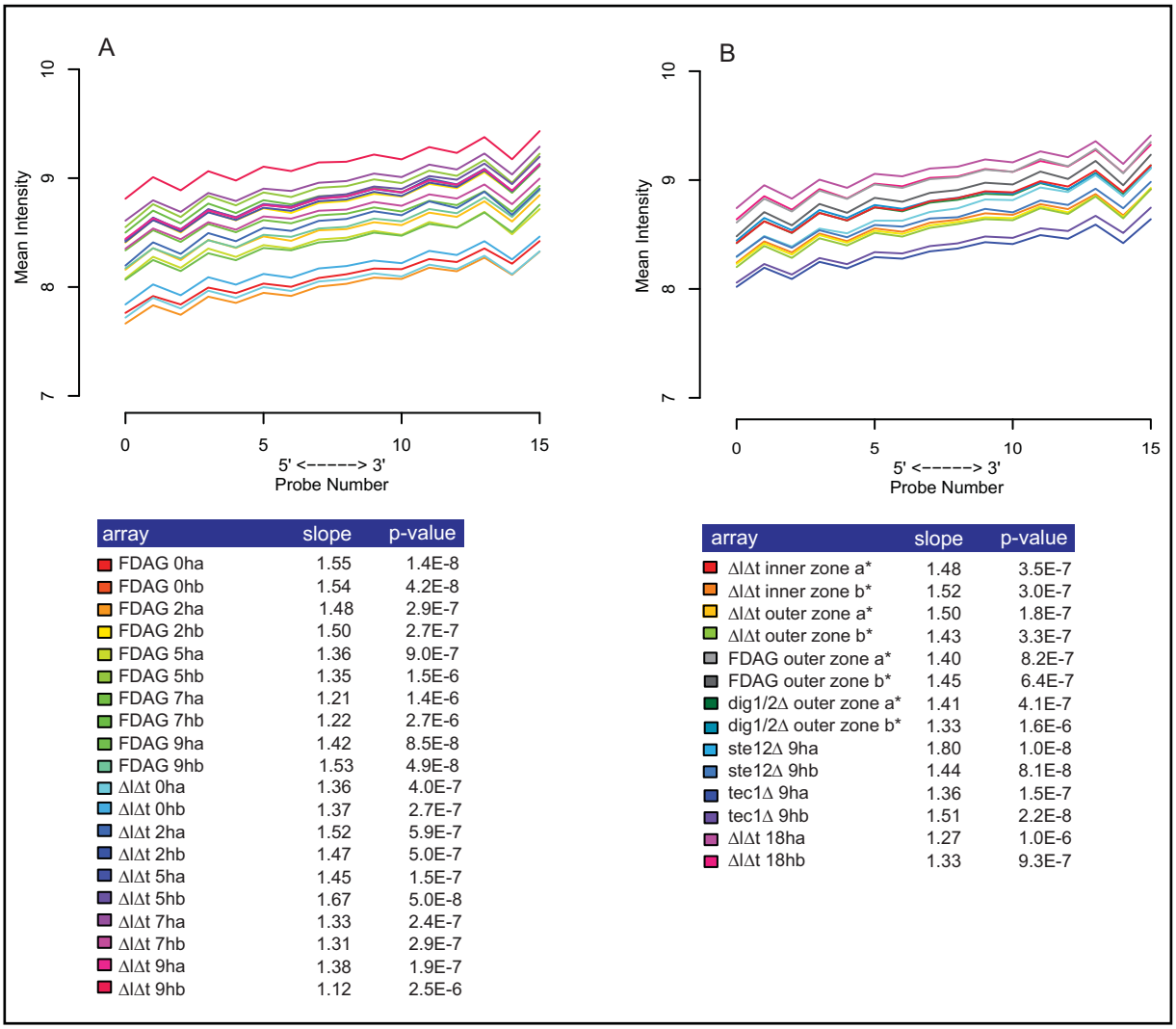


Figure 1-9 Image of the hybridized sySYNG001a array (18 ha experiment)



array	slope	p-value
FDAG 0ha	1.55	1.4E-8
FDAG 0hb	1.54	4.2E-8
FDAG 2ha	1.48	2.9E-7
FDAG 2hb	1.50	2.7E-7
FDAG 5ha	1.36	9.0E-7
FDAG 5hb	1.35	1.5E-6
FDAG 7ha	1.21	1.4E-6
FDAG 7hb	1.22	2.7E-6
FDAG 9ha	1.42	8.5E-8
FDAG 9hb	1.53	4.9E-8
ΔlΔt 0ha	1.36	4.0E-7
ΔlΔt 0hb	1.37	2.7E-7
ΔlΔt 2ha	1.52	5.9E-7
ΔlΔt 2hb	1.47	5.0E-7
ΔlΔt 5ha	1.45	1.5E-7
ΔlΔt 5hb	1.67	5.0E-8
ΔlΔt 7ha	1.33	2.4E-7
ΔlΔt 7hb	1.31	2.9E-7
ΔlΔt 9ha	1.38	1.9E-7
ΔlΔt 9hb	1.12	2.5E-6

array	slope	p-value
ΔlΔt inner zone a*	1.48	3.5E-7
ΔlΔt inner zone b*	1.52	3.0E-7
ΔlΔt outer zone a*	1.50	1.8E-7
ΔlΔt outer zone b*	1.43	3.3E-7
FDAG outer zone a*	1.40	8.2E-7
FDAG outer zone b*	1.45	6.4E-7
dig1/2Δ outer zone a*	1.41	4.1E-7
dig1/2Δ outer zone b*	1.33	1.6E-6
ste12Δ 9ha	1.80	1.0E-8
ste12Δ 9hb	1.44	8.1E-8
tec1Δ 9ha	1.36	1.5E-7
tec1Δ 9hb	1.51	2.2E-8
ΔlΔt 18ha	1.27	1.0E-6
ΔlΔt 18hb	1.33	9.3E-7

Figure 1-10 RNA digestion plots ((A) arrays of samples from germination experiments (B) arrays of samples from plates (*), deletion experiments and advanced mycelium) are used to assess uniform treatment of samples, RNA integrity and quality of the cDNA synthesis step. For each array and within each probe set, the probes are ordered by their proximity to the 5' end of the gene. The average intensities of the probes classified by this order are shown in the plots. Each array is represented by a line. The slopes and p-values from linear regressions are shown in the lists. Typically RNA digestion plots of good quality chips do not show horizontal lines, but a slight trend to lower mean intensities towards the 5' end of the gene. Observed slopes are a characteristic of the type of chip (sySYNG001a) and the laboratory handling. High slopes might (but do not necessarily) indicate poor RNA quality or cDNA synthesis. Due to the fact that comparable data is missing, we focus here on the main purpose of the plots which is to highlight differences in degradation patterns. Parallel lines, as seen in plots A + B, are a good indicator for uniform and reproducible laboratory procedure (plots were done in R (R. Ihaka and R. Gentleman, 1996) using the tool AffyRNAdeg (L. Cope, 2006)).

RNA digestion plots of good quality chips do not show horizontal lines, but a slight trend to lower mean intensities towards the 5' end of the gene. This is because the Affymetrix eukaryotic expression assay has an inherent 3' bias (cDNA synthesis and cRNA synthesis via incorporated T7 promoter start at the 3' end). High slopes might (but do not necessarily) indicate degraded RNA or inefficient transcription of ds cDNA or biotinylated cRNA. Observed slopes are a characteristic of the type of chip (sySYNG001a) and the laboratory handling. Due to the fact that comparable data is missing, we focus here on the main purpose of the plots which is to highlight differences in line patterns. Parallel lines, as seen in plots A + B, are a good indicator for uniform and reproducible laboratory procedure (plots were done in R (R. Ihaka and R. Gentleman) using the tool AffyRNAdeg (L. Cope).

Reproducibility and similarity

Reproducibility is measured by the comparison of biological replicates. A straight-forward way to compare array data is the scatter plot (two arrays) or the pairs

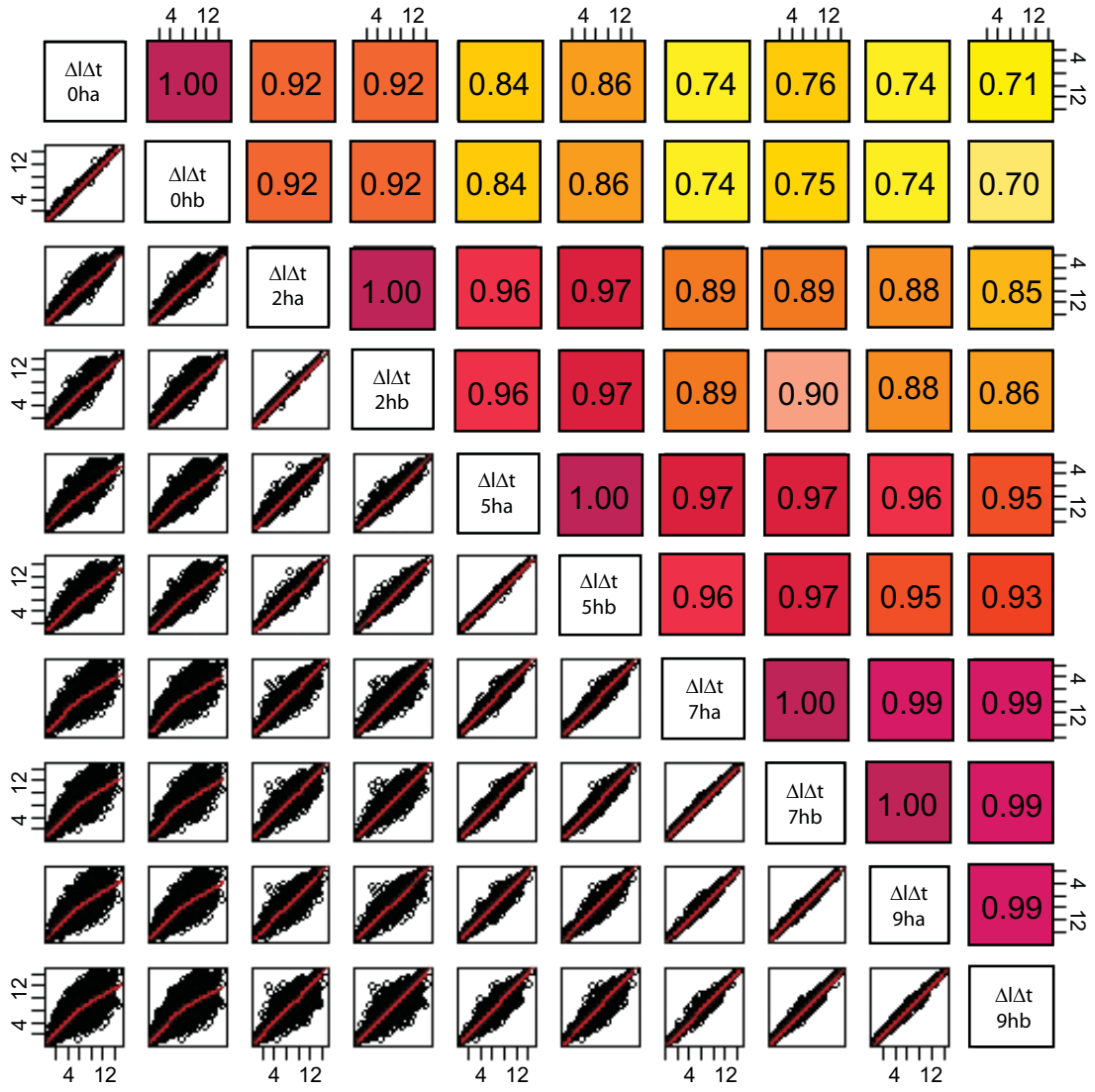


Figure 1-11 A Pairs plot of log2 transformed intensity values. Corresponding correlation coefficients are indicated.

plot (multiple arrays). Two variables are plotted along the axes and dots are drawn according to these coordinates. The dots in the scatter plot fall on a straight line if the two variables are perfectly linearly related. Correlation coefficients serve as similarity measure. Scatter plots of log2 transformed intensity values from our biological replicates showed high reproducibility with correlation coefficients ranging from 0.99 to 1.0 (except for ste12Δ which was 0.98; figure 1-11 F). The data points of replicates are scattered along the diagonal. Pair-wise comparison of the arrays belonging to the germination experiment revealed a steady decrease of the corresponding correlation coefficients with increasing time after spore inoculation independent of the strain that was used. Correlation coefficients of the germination experiment dropped from 1.0 to 0.71 for the laboratory strain ΔlΔt and from 1.0 to 0.67 for the natural isolate FDAG. The data clouds are increasingly spread out and spherical in shape. In biological terms this means that the further the process of germination advances the more the gene expression pattern changes. It increasingly deviates from the pattern observed in spores. This fundamental observation is crucial. When we started the transcriptional profiling project of

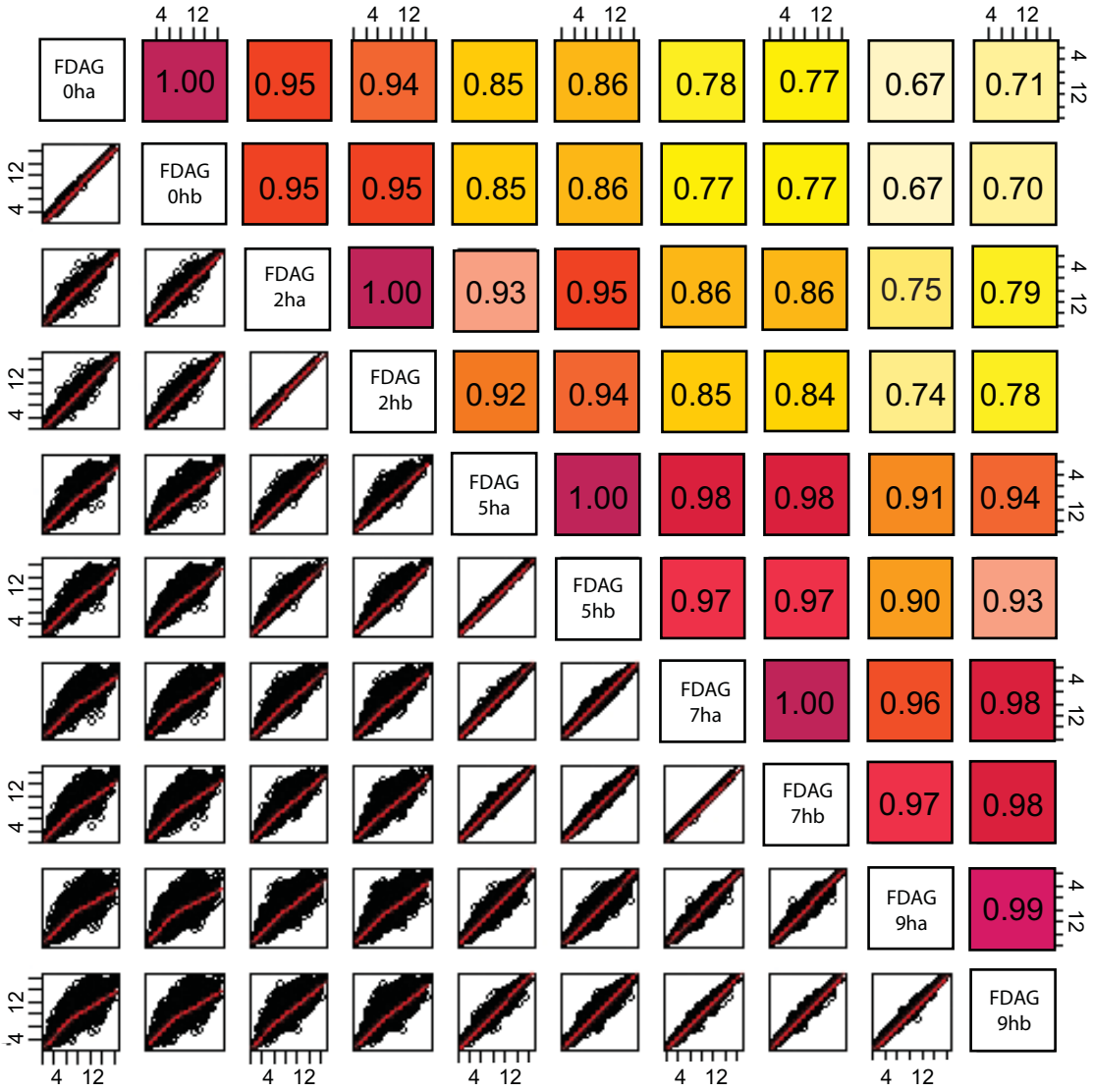


Figure 1-11 B Pairs plot of log2 transformed intensity values. Corresponding correlation coefficients are indicated.

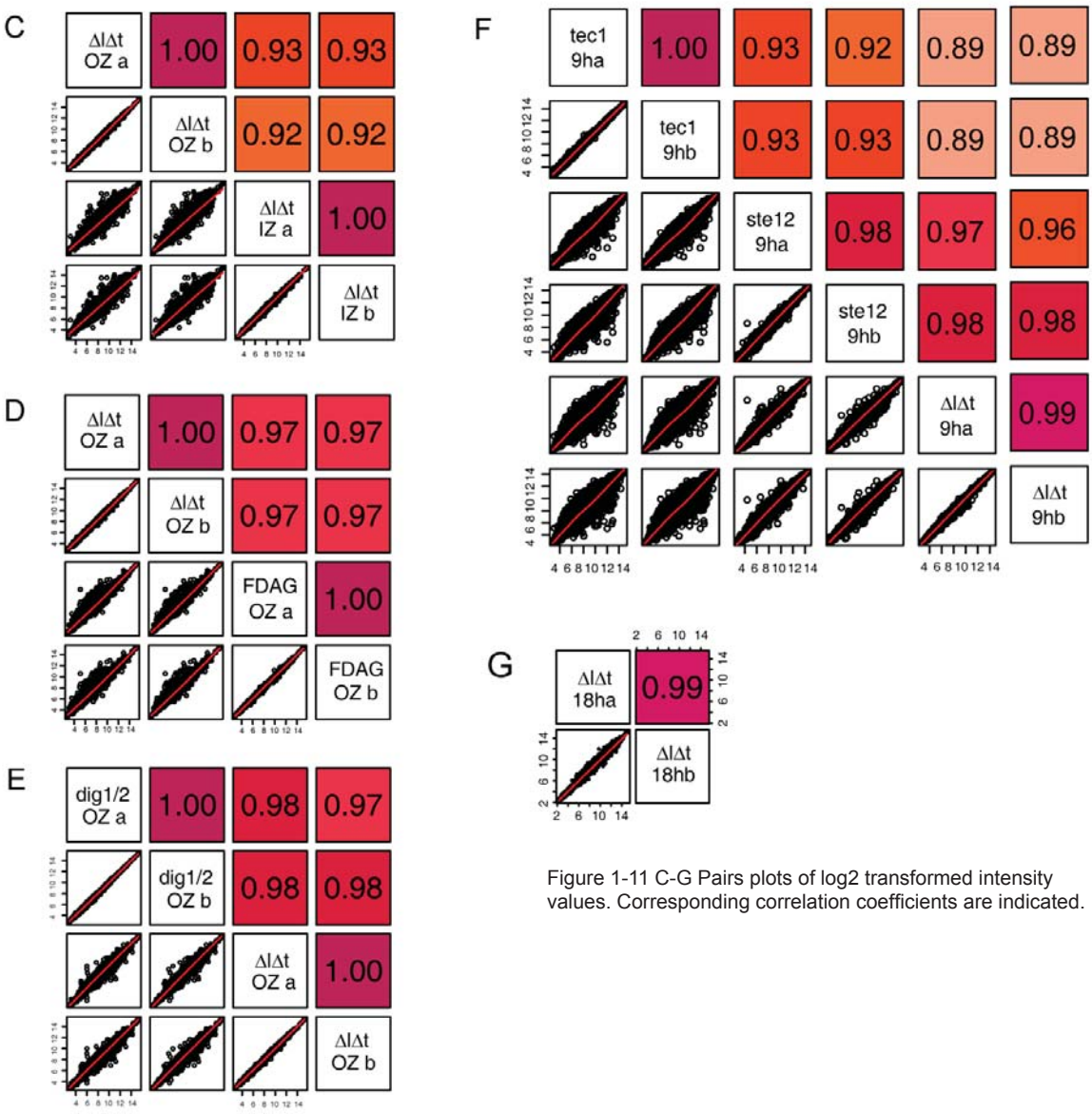


Figure 1-11 C-G Pairs plots of log2 transformed intensity values. Corresponding correlation coefficients are indicated.

A. gossypii life cycle stages it was not clear whether the changes in expression within the first few hours after germination would be adequately captured by the experiment design. A known problem is the mixed cell type population due to asynchronous germination which will be discussed in chapter 2.

Data distribution

Data coming from microarray experiments does in general not follow a normal distribution. The RAW data is positively skewed with just few observations at very high intensity levels and many observations at low intensities. Many statistical tests, like the t-test or ANOVA (analysis of variance), assume that the data is normally distributed. Results coming from statistical tests can be misleading if the data is highly skewed. Log2 transformation is often used to make the data more normally distributed.

General introduction to normalization strategies

Normalization is an essential procedure in the analysis of microarrays to compare data from different one-colour arrays. Measurements on microarrys may be biased by diverse effects such as efficiency of RNA extraction, reverse transcription, label incorporation, hybridisation time, washing, scanning etc. The main aim of normalization is to remove all technical and systematic variation from the data, so that the observed differential expression is (ideally) exclusively due to biological variation of the individual samples. There are three types of internal controls that can be used for normalization.

1. global normalization is based on all genes on the array
2. previously known housekeeping-genes
3. spiked-in control genes

The selection of a set of housekeeping genes, including at least 50 genes, can be a challenging task. The prerequisite of a suitable housekeeping gene is as follows: it should be adequately expressed during all life cycle stages and in all parts of the fungal mycelium and show minimal variability in expression between samples and under the experimental conditions used. It has been shown previously in mammalian expression studies that many housekeeping genes exhibit considerable variability under different experimental conditions (Lee *et al.*, 2002). Furthermore their expression levels are often on the high side, making them unsuitable and unrepresentative of the whole expression intensity range.

When using spiked controls, known (biotin-) labeled cRNA targets are added to the array in defined concentrations and used later for normalization or for quality control of the normalization step. To prevent cross-hybridisation, the spiked-in targets are often from bacterial or viral origin. The disadvantage with spiked-in controls is that they have to be present in sufficient numbers and they have to be included in the Chip design.

Global normalization is based on all genes on the array. The assumption is used that between two conditions the majority of genes do not change in terms of their expression level (Quackenbush, 2002). Many state-of-the art global normalization algorithms exist today for one-color arrays, the most frequently used ones being mas5, vsn, rma and gc-rma. Mas5 and gc-rma consider mismatch (mm) information in addition to perfect match (pm) signals and are therefore not applicable to the pm-only arrays sySYNG001a. RMA (robust multichip average) adjusts the scale of the data by setting not only the median (50% percentile) of differences to 0, but by adjusting all of the various quantiles (Irizarry *et al.*, 2003). Vsn (variance stabilisation normalization) transforms the data to have equal variances for all intensities (Huber *et al.*, 2002). This transformation method adjusts for effects that are often observed after background subtraction (i.e. high-variance for lowly expressed genes). Which normalization method gives the best results is dependent on the experimental design and the type of data. RMA and VSN normalization performed well on our data set and we decided to use RMA as the default normalization method.

Removal of *C. heterostrophus* probes before RMA normalization

The sySYNG001a Chip contains 4696 *A. gossypii* and 20552 *C. heterostrophus* probe sets. Hybridising solely *A. gossypii* targets results in 81% of the probes showing background intensity levels. RMA normalization across all probes causes the median signal intensities to drop to a log2 transformed value of 4, while log2 signals under 6 should be considered as background. To overcome this difficulty we selected exclusively *A. gossypii* probes before normalization. This was done with the R software (The R foundation for statistical computing, 2006) exploiting the fact that all

Affymetrix identifiers of *A. gossypii* probes start with "^Ag". In a first step the CEL files were read into an Affybatch and then, the existing environment was modified with the script shown in table 1-4. Environment modification is based on a package from Laurent Gautier (L. Gautier, 2007). Re-evaluation of the signal intensities post normalization revealed log2 transformed medians of 8.

```
# read .CEL files into an Affybatch a
a <- ReadAffy(filenames=dir()[1:20])

#set new environment for Affybatch object
library(altcdfenvs)
sySYNG001a=make.cdf.env("sySYNG001a.CDF")
ashbycdf=wrapCdfEnvAffy(sySYNG001a,640,640,"sySYNG001a")
ids =geneNames(ashbycdf)
ids.Ag = ids[grep("^Ag",ids)]

agcdf=ashbycdf[ids.Ag]
envagcdf<-as(agcdf,"environment")
a@cdfName="envagcdf"

save(a,agcdf,envagcdf,file="xxx.rda")
```

Table 1-4 Script for environment modification in R

RMA (Robust multichip average) pre-processing for Affymetrix chips

RMA (Irizarry *et al.*, 2003) is an expression measure consisting of three particular pre-processing steps i) background adjustment ii) quantile normalization, and iii) summarization by median polishing (Turkey, 1977). RMA uses only perfect match information. The following section describes the individual steps in more detail.

i) background adjustment

The perfect match intensity is a sum of the background (optical noise and non-specific binding) and the signal ("spot") intensity: $PM_{ijn} = bg_{ijn} + s_{ijn}$. n represents the different genes, i represents different RNA samples, and j represents the probe number (denotes physical position of oligonucleotide in the gene). It is assumed that each array has a common mean background level. The signal s_{ijn} is exponential and the background bg_{ijn} is normally distributed. The values of background and foreground intensities can then be estimated based on these assumptions, and foreground intensities can be adjusted accordingly.

ii) quantile normalization (Bolstad et al., 2003)

Quantile normalization assumes similar intensity value distribution on all arrays. This assumption does not necessarily hold well for genes with high expression levels. Practically, quantile normalization means that the intensity values on all arrays are arranged in ascending order. Arranged values for every probe are then averaged over all the chips. Last, the intensity values for the probes are replaced with their mean on every chip, and the means are arranged in the original order that is based on the arranged intensity values. A practical example with three genes can be found in "DNA Microarray Data Analysis" by (Hovatta, 2005).

iii) summarization

Summarization of probes to a single expression value per gene is done by median polishing (Turkey, 1977). Median polishing is an iterative method, which aims to centralize both column medians and row medians to one. In the case of RMA, the algorithm proceeds as follows (B. Bolstad, PhD thesis): first a matrix is formed for each probe set n such that the probes are in rows and the arrays are in columns. This matrix is augmented with row and column effects. Next, each row is swept by taking the median across columns (ignoring the last column of row effects) subtracting it from each element in that row and adding it to the final column. Then the columns are swept in a similar manner by taking medians across rows, subtracting from each element in those rows and then adding to the bottom row. The procedure continues, iterating row sweeps followed by column sweeps, until the changes become small or zero.

After RMA normalization, the data from each array has about the same median and similar distribution (see data from germination experiment in figure 1-12). The most convenient way to graphically depict the five-way summary of a distribution is

the boxplot (Turkey, 1977) which shows the smallest observation, lower quartile, median, upper quartile, and largest observation. The boxplot also indicates which observations, if any, are considered outliers.
Boxplots of the ΔlΔt and FDAG germination data before and after rma normalization are viewed in figure 1-12.

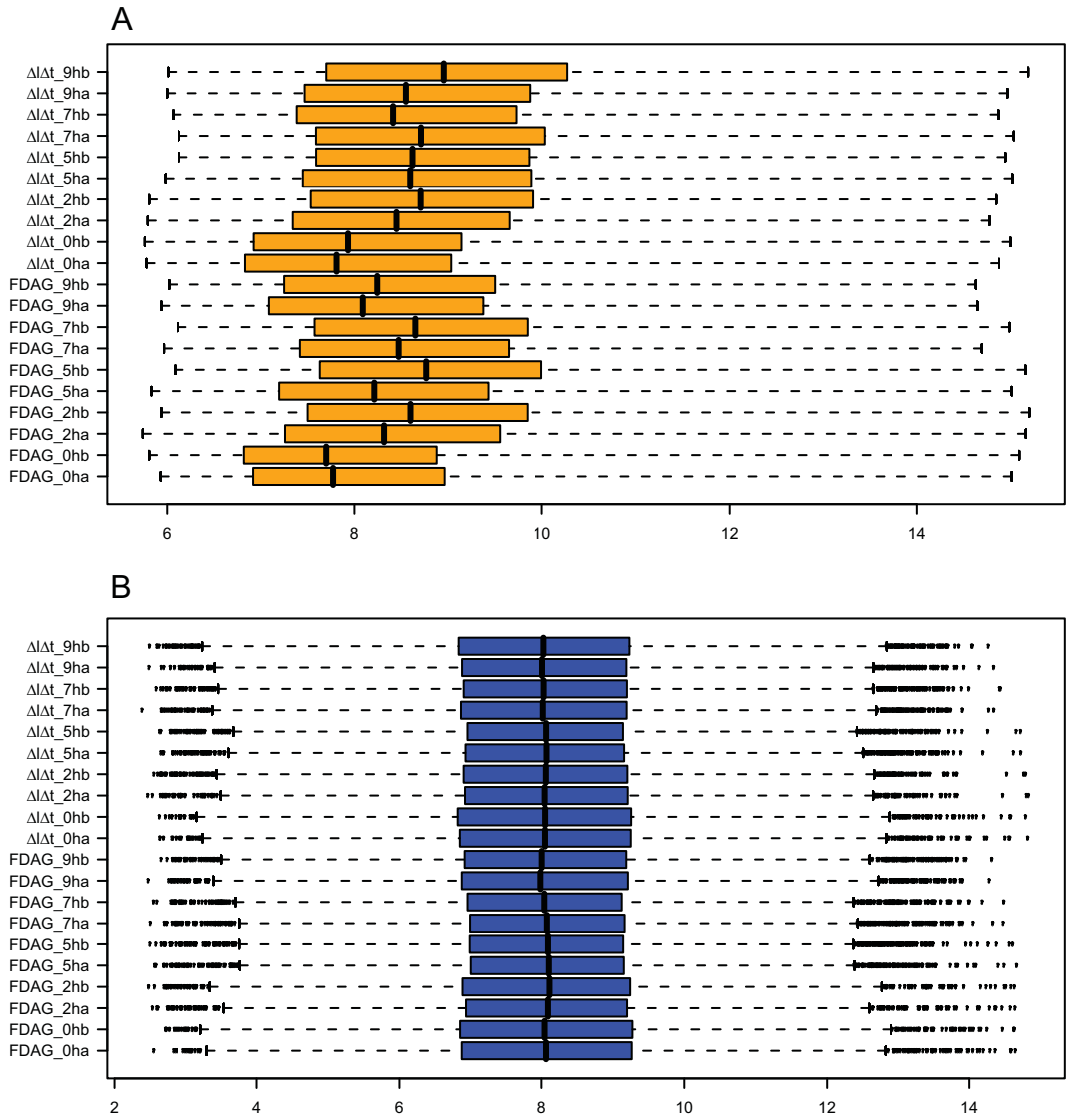


Figure 1-12 Box plots of log2 transformed signal intensity values from the germination experiment before (A) and after RMA normalization (B)

1.6 Data storage

Deposition of expression data in public repositories

There are several good reasons why microarray data should be made available to the public. First, analyzing several experiments together will facilitate the task of finding actual correlations in gene expression. Second, inter-species comparison of expression data will help to identify regulatory networks that have been conserved during evolution. And third, microarray experiments are expensive and adding publicly available information to a study will save money and time. Many journals including nature and science ask for the deposition of microarray data in a public database prior to manuscript publication. However in order to be able to compare data from different researchers and laboratories the data has to be accompanied by supportive annotation. This annotation has to enable an unambiguous interpretation and reproduction of the experiment. A widely accepted standard for the description of microarray experiments is MIAME (Minimum Information About a Microarray Experiment). To facilitate and accelerate the annotation of our MIAME compliant data we used MIMAS. MIMAS is an innovative tool for high density oligonucleotide microarray data management and annotation (Hermida *et al.*, 2006). The local array data management solution MIMAS is on a data node directly connected to the European certified public microarray data repository Array Express located at the EMBL-EBI. Other major public repositories are the Expression omnibus (GEO) and the NCBI at the center for Information Biology Gene Expression (CIBEX) at DDBJ. We submitted 34 Affymetrix CEL files with raw probe-level information to Array Express at EMBL-EBI. The data will be publicly accessible and retrievable after manuscript approval. Normalized linear and log2 transformed data of germinating spores inoculated into full medium for four, five and seven hours was included in the cross-species genome/transcriptome database AGD 3.0 and displayed in the context of the DNA annotation (Gattiker *et al.*, 2007). The remaining data described in this thesis is going to be added to AGD 3.0 in the near future. Table 1-3 summarizes experimental conditions and protocols that were used for the 34 *A. gossypii* arrays.

1.7 Concluding remarks

We present here for the first time a data set comprising transcription data across all major life cycle stages of the filamentous fungus *A. gossypii* and data aiming at studying gene expression changes upon transcription factor deletions. The data set has been generated with 34 state-of-the-art high density oligonucleotide arrays, custom-made for *A. gossypii*. It is close to being complete and comprises 94% of all annotated *A. gossypii* genes and 190 observations in intergenic or antisense regions. Since data quality is crucial to microarray experiments, extensive quality control steps have been introduced throughout the data generation process. Pre-processed data is accessible via our *A. gossypii* genome database (www.agd.unibas.ch) and raw data will be available for download from Array Express at EMBL-EBI.

The experiments were performed with two *A. gossypii* strains, a laboratory strain used for functional genomics (ΔlΔt strain) and a strain recently isolated from an insect named cotton stainer in Florida (F. Dietrich, personal communication). The subsequent chapters will focus on results obtained from the laboratory strain but, whenever useful, results from the natural isolate will be included in discussions. Two full medium growth conditions were used throughout this study. Liquid cultures were used to assay germination and formation of an advanced

Array name	Replicate array	Strain background	Type of experiment	Culture method	Sample collection method	RNA isolation method
ΔlΔt 0ha	ΔlΔt 0hb	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt 0hb	ΔlΔt 0ha	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt 2ha	ΔlΔt 2hb	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt 2hb	ΔlΔt 2ha	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt 5ha	ΔlΔt 5hb	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt 5hb	ΔlΔt 5ha	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt 7ha	ΔlΔt 7hb	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt 7hb	ΔlΔt 7ha	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt 9ha	ΔlΔt 9hb	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt 9hb	ΔlΔt 9ha	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
FDAG 0ha	FDAG 0hb	natural isolate (FDAG)	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
FDAG 0hb	FDAG 0ha	natural isolate (FDAG)	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
FDAG 2ha	FDAG 2hb	natural isolate (FDAG)	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
FDAG 2hb	FDAG 2ha	natural isolate (FDAG)	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
FDAG 5ha	FDAG 5hb	natural isolate (FDAG)	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
FDAG 5hb	FDAG 5ha	natural isolate (FDAG)	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
FDAG 7ha	FDAG 7hb	natural isolate (FDAG)	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
FDAG 7hb	FDAG 7ha	natural isolate (FDAG)	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
FDAG 9ha	FDAG 9hb	natural isolate (FDAG)	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
FDAG 9hb	FDAG 9ha	natural isolate (FDAG)	time course	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt OZa	ΔlΔt OZ b	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	30ml AFM agar plates	harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt OZ b	ΔlΔt OZ a	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	30ml AFM agar plates	harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt IZ a	ΔlΔt IZ b	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	30ml AFM agar plates	harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt IZ b	ΔlΔt IZ a	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	30ml AFM agar plates	harvesting with spatula	hot phenol (Schmitt et al., 1990)
FDAG OZ a	FDAG OZ b	natural isolate (FDAG)	time course	30ml AFM agar plates	harvesting with spatula	hot phenol (Schmitt et al., 1990)
FDAG OZ b	FDAG OZ a	natural isolate (FDAG)	time course	30ml AFM agar plates	harvesting with spatula	hot phenol (Schmitt et al., 1990)
dig1/2Δ OZ a	dig1/2Δ OZ b	ATCC10895; Ag*leu2*Δ*thr4*Δ	knock out	30ml AFM agar plates	harvesting with spatula	hot phenol (Schmitt et al., 1990)
dig1/2Δ OZ b	dig1/2Δ OZ a	ATCC10895; Ag*leu2*Δ*thr4*Δ	knock out	30ml AFM agar plates	harvesting with spatula	hot phenol (Schmitt et al., 1990)
ste12Δ 9ha	ste12D 9hb	ATCC10895; Ag*leu2*Δ*thr4*Δ	knock out	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
ste12Δ 9hb	ste12D 9ha	ATCC10895; Ag*leu2*Δ*thr4*Δ	knock out	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
tec1Δ 9ha	tec1Δ 9hb	ATCC10895; Ag*leu2*Δ*thr4*Δ	knock out	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
tec1Δ 9hb	tec1Δ 9ha	ATCC10895; Ag*leu2*Δ*thr4*Δ	knock out	10ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt 18ha	ΔlΔt 18hb	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	200ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)
ΔlΔt 18hb	ΔlΔt 18ha	ATCC10895; Ag*leu2*Δ*thr4*Δ	time course	200ml liquid AFM	filtration, harvesting with spatula	hot phenol (Schmitt et al., 1990)

Table 1-3 Summary of the 34 arrays that were hybridized and analyzed in this study

mycelium which represent early and mature stages in the *A. gossypii* life cycle. Material from mycelial colonies growing on solid medium was used to monitor developmental differentiation occurring later in the life cycle, tip splitting in fast growing hyphae at the edge of colonies and spore formation in the inner zone of the colony (figure 1-5). Expression of a deletion mutant (*dig1/2*Δ) with a sporulation defect was also assayed on plates. We have chosen elaborated and adapted protocols for sample collection and RNA isolation for the filamentous fungus *A. gossypii*. These protocols are applicable to all life cycle stages and are similar to protocols used in *S. cerevisiae*, which facilitates possible future inter-species comparisons of transcriptional profiles.

Chapter 2: A spore awakes: Development of the transcriptome from spores to bipolar germlings

Chapter 2: A spore awakes: Development of the transcriptome from spores to bipolar germlings

2.1 Introduction

A. gossypii is a filamentous fungus with sequenced genome and custom made DNA chips. Because of its simple life cycle it is perfectly suitable for transcriptional profiling studies across life cycle stages aiming at elucidating differential gene expression. *A. gossypii* spores are easily cleanable in sufficient amounts with a simple, albeit labor intensive, method that makes use of the high spore hydrophobicity and which prevents use of externally added lytic enzyme. Spores were attached to silicone coated glass tubes and detached by washing with TritonX-100 after removal of the hydrophilic vegetative mycelium by washing steps with water as described in more detail in chapter 1 of this thesis (Brachat, PhD thesis, 2003).
Spore germination is not a synchronous process in filamentous fungi and also not for *A. gossypii*. This becomes obvious if the different stages of germination are counted at selected time points post inoculation of spores in liquid medium (figure 2-2 A). During the first two hours, we did not observe a morphological alteration of the spores, e.g. no trace of germ bubbles. At 5 h 37% had developed a germ bubble in the center of the spores; at 7 h a mixed population was seen with 33% spores, 31% germ bubble spores, 32% germlings with one hyphae emerging from the germ bubble (unipolar germlings) and 4% with already two hyphae growing in opposite directions (bipolar germlings). Finally at 9 h we counted 20% unaltered spores, 22% germ bubble spores, 28% unipolar germlings, and 30% bipolar germlings. Regardless of the culture heterogeneity it is possible to obtain a reliable time-resolution in terms of transcriptional changes per nucleus. This is due to the increase in number of nuclei per cell during germination. *A. gossypii* spores invariably contain a single nucleus, germ

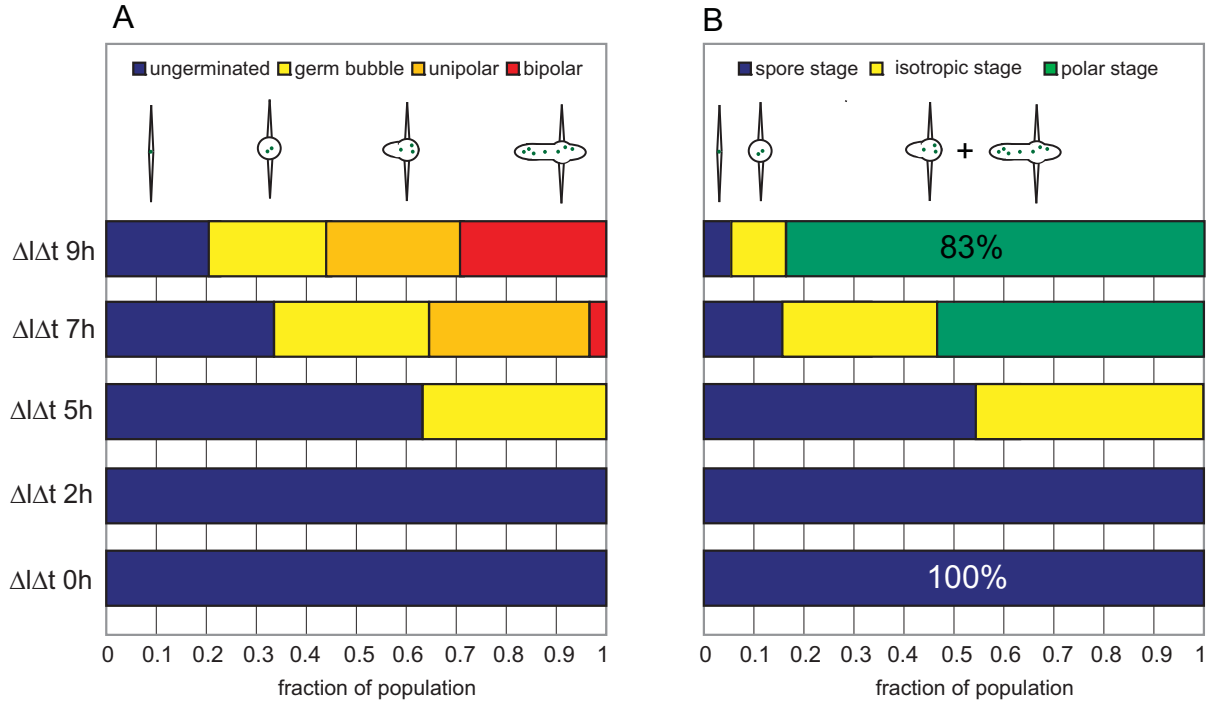


Figure 2-2 Cell type composition (A) and fraction of nuclei belonging to the three growth phases (B) in relation to inoculation time

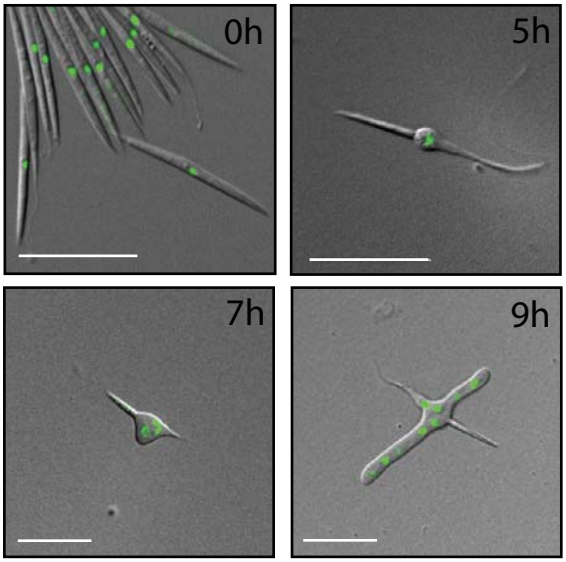


Figure 2-1 Nuclear content of spores and germlings at successive time points after inoculation. The nuclei were in vivo labeled with H4-GFP (pictures kindly provided by Claudia Birrer). Scale bars: 20 µm.

bubble spores one or two, unipolar germlings three or five depending on the length of the hypha and bipolar germlings contain six to nine nuclei. Examples are shown in figure 2-1 using Histone4-GFP labeled nuclei. In addition of counting for each time point the four distinguishable types of germlings, we counted the fraction of nuclei being in a particular growth stage (figure 2-2 B). Based on the distribution shown in figure 2-2, at 5 h 54% of the nuclei are still in morphologically unaltered spores and 46% represent nuclei after the first mitosis. At 7 h there are 18% spore nuclei, 28% germ bubble nuclei and 54% polar-growth-stage nuclei. Finally at 9 h there are only 5% spore nuclei, 12% germ bubble nuclei and 83% polar-growth-stage nuclei. Thus, data at 0 h represent fully the spore transcriptome and data at 9 h reliably represent the transcriptome of nuclei in polar growing germlings with 20% background of less developed nuclei.

RNA was extracted from two independent spore and germling preparations collected at indicated time points. As described in the general chapter 1, cRNA targets were prepared, hybridized to the probes on the chip, and the fluorescence intensities were determined.

The first step in the data analysis was hierarchical clustering. Visualization of global changes in gene expression in a time-dependant manner is preferentially done by such a hierarchical clustering of the data into a condition tree (figure 2-3 A). The conditions are on the horizontal axis and the genes are on the vertical axis in increasing Affymetrix ID order. Replicate experiments cluster together at the highest resolution level, before, at the next level neighboring time points cluster together indicating reproducible transcription changes between the selected time points. Genes were colored according to their expression level compared to their average expression across time, red indicating over-expression and blue under-expression. We also employed the method of scatter plots to visualize global changes in gene expression during germination. In this approach two variables are plotted along the axes and dots are drawn according to these coordinates. As seen in figure 2-3 B, the data points of replicates were scattered along the diagonal. Pair-wise comparison of data from all arrays from 0 h to 9 h revealed data clouds that were increasingly spread out and spherical in shape. In biological terms this meant that the further the process of germination advanced the more the gene expression pattern deviated from the one originally observed in spores. This observation was crucial. When we started the transcriptional profiling project of *A. gossypii* life cycle stages it was not clear whether the changes in expression within the first few hours after germination would be adequately captured by the experiment design. Corresponding correlation coefficients for each pair-wise comparison dropped from 1 for biological replicates to 0.70 for the comparison of the array 0 hb with 9 ha. A very similar decrease in correlation coefficients was observed for the natural isolate strain FDAG which will be discussed in appendix 2 but it is already mentioned here to support the scatter plot analysis. The coefficient dropped from 1 to 0.67 (chapter 1, figure 1-11 B). High inter-time point correlations were observed for

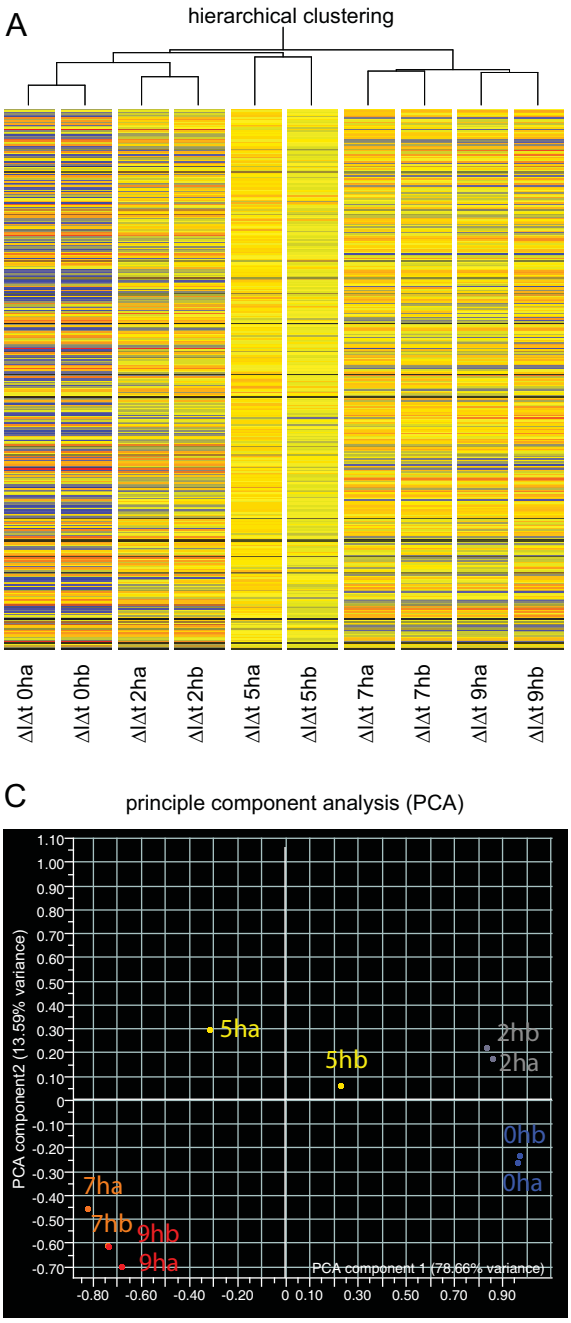


Figure 2-3 Visualization of gene expression changes during germination. (A) Hierarchical clustering of the ten experiments. The samples were displayed horizontally and the genes vertically in increasing Affymetrix ID order. Note that replicate experiments clustered together at the highest resolution level, before, at the next level neighboring time points clustered together. Genes were colored according to their expression level compared to their average expression across time, red indicating over-expression and blue under-expression. (B) Pairs plot of the experiments showed the highest correlation coefficients for replicate measurements (dark pink fields). An overall tendency to lower correlation coefficients (orange and yellow fields) with increasing inoculation time was observed. (C) PCA visualizes inter-experiment variances. Smallest variances were observed for replicate measurements. Individual time points were well separated which is a prerequisite for subsequent differential gene expression analysis.

7 h and 9 h. This was confirmed by the condition tree (figure 2-3 A) and by principal component analysis (PCA) of samples (figure 2-3 C) and indicates that there is not much variance between the two time points compared to earlier time points where most drastic changes in gene expression happen. These two types of analyses clearly show strong global gene expression changes during germination of *A. gossypii* spores. In order to evaluate these data with respect to biological implications we will proceed in two steps. We will start by looking at the 20 most abundant mRNAs found in spores and in germlings with one or two hyphae (9 h). Then we will present and discuss a clustering analysis of the 0 h to 9 h time point data obtained with the laboratory strain.

2.2 Results and Discussion

Abundant mRNA species in spores

Spores contained considerable amounts of mRNA. Some of the mRNAs might have been newly synthesized in hydrated spores for instance during the spore preparation

process while the majority very likely reflects the mRNA population at the time of spore wall closure. A ranking of the 20 most abundant mRNA species in spores revealed a surprising picture (figure 2-4 A and table 2-1 A). The proteins encoded by the genes listed in decreasing relative transcript levels can be assigned to the following eight groups as inferred from the function of the syntenic homologs in *S. cerevisiae*.

1. Cell wall (CCW12, SED1/SPI1, CWP1d, NOHBY621, TDH3, UTH1/NCA3, CWP1b)
2. Protein assembly and aggregation (HSP26, RNQ1, LSB1/PIN3, STF2)
3. Chromatin (HHT1)
4. C2/C3 metabolism (TDH3, ENO1/2, ADH2)
5. Protein degradation (UBI4)
6. Redox homeostasis and protection against oxidative damage (TRX1/2, AHP1)
7. Cortical patches/Endocytosis (LSB1/PIN3, PIL1)
8. Glycolipid complexes (NCE102)

For the fifth highest expressed gene ADL036W-A a syntenic homolog YLR307C-X exists in *S. cerevisiae* (Brachat et al., 2003) but a function is not known. Interestingly none of the ribosomal protein genes ranked among the top 20 genes and not even in the top100 genes. In the following we will briefly discuss these eight groups. Further work of the *A. gossypii* cell wall proteome which was started in collaboration with Piet de Groot and Frans M. Klis (University of Amsterdam) will be described in a extra chapter (chapter 5).

A first glance of highly expressed cell wall genes

AgCCW12 codes for a mannoprotein with a predicted GPI-anchor site. *S. cerevisiae* carries two syntenic homologs with highly conserved sequences, ScCCW12 and YDR134C, which are so-called twins because they originate from the yeast whole genome duplication. **AgSED1/SPI1** codes for a 505 aa cell wall protein with predictable sites for glycosylation and GPI-anchor. Again *S. cerevisiae* carries twin homologs ScSED1

Highly expressed at 0h (Top 20)

0h rank	2h	5h	7h	9h rank	Affy ID	Systematic Name	Common Name		Exp 0h	Exp 9h
1				1	Ag004168_at	AGR049W	CCW12 (*)		27268	25150
2				2	Ag003796_at	AGR138W	SED1 (*)	SPI1(*)	19809	22464
3				1529	Ag000439_at	ABR028C	CWP1d(*)		18698	531
4				910	Ag000012_at	AAL128C	HSP26		15061	917
5				980	Ag004550_i_at	ADL036W-A			14611	866
6				7	Ag002925_at	AFR065W	NOHBY621(*)		12087	17702
7				70	Ag002394_at	AER013W	HHT1		11892	10199
8				3	Ag002696_at	AER031C	TDH3(*)		11559	20893
9				967	Ag003162_at	AFR701C	RNQ1		11391	877
10				245	Ag000970_at	ACL062C	UBI4		11071	3239
11				14	Ag002371_at	AER294C	ENO2	ENO1	10717	15868
12				758	Ag001674_at	ADR322W	UTH1 (*)	NCA3	10381	1094
13				82	Ag001224_at	ACL131W	TRX1	TRX2	10367	9575
14				199	Ag001360_at	ADL154C	AHP1		10263	4201
15				953	Ag000323_at	AAR084W	ADH2		9384	888
16				1509	Ag000455_at	ABR026C	CWP1b(*)		9345	539
17				438	Ag003457_at	AFR320W	LSB1	PIN3	9301	1796
18				346	Ag003617_at	AFR505C	STF2		9252	2380
19				535	Ag002072_at	AEL329W	PIL1		9087	1485
20				466	Ag003189_at	AFR312W	NCE102		8463	1679

Figure 2-4 A Ranking of the 20 most abundant mRNA species at 0 h

(*) cell wall protein genes (inferred from homology to *S. cerevisiae*)

(*) newly identified *A. gossypii* specific cell wall protein gene

color code:
in comparison to average of each gene:
overexpressed underexpressed

and ScSPI1. **NOHBY621** is now known to code for a cell wall protein with a typical PIR repeat (see chapter 5). **AgCWP1b** and **AgCWP1d** originate from a rare quadruplication event starting with a syntenic homolog to ScCWP1. A total of 7 members of this unique *A. gossypii* gene family have been annotated (Dietrich *et al.*, 2004) and four will be discussed in chapter 5. **AgUTH1/ NCA3** encode a cell wall protein of the SUN family. Again, two syntenic homologs (twin) are present in the *S. cerevisiae* genome, ScUTH1 and ScNCA3. **AgTDH3** encodes the glycolytic enzyme glycerin-aldehyd-phosphat-dehydrogenase, the *S. cerevisiae* homolog of which is repeatedly found in yeast cell wall preparations. Therefore AgTDH3 is counted as cell wall gene.

Highly expressed protein assembly/ aggregation genes

A. gossypii carries five genes presently annotated as **homologs of ScHSP26** (YBR072W) which codes for a small heat shock protein with chaperone activity. *A. gossypii* lacks a syntenic homolog but the five non-syntenic homologs code for conserved proteins with highest similarity to ScHSP26. The **AgHSP26 homolog AAL128C** is exceptionally highly expressed. **AgRNQ1** may promote ordered aggregation of Q/N rich proteins as demonstrated for ScRNQ1, its syntenic yeast homolog, which encodes a prion (Sandheimer and Lindquist, 2000). **AgLSB1/PIN3** is again a gene with two syntenic homologs in *S. cerevisiae* ScLSB1 and ScPIN3. The functional annotation of both genes differs considerably and it is difficult to extrapolate the function of the *A. gossypii* syntenic homolog. ScLSB1 is known to bind to ScLAS17 and to be important for assembly of actin patches. ScPIN3 induces the appearance of PIN^+ prions when overexpressed. One important difference is the SH3 domain of ScLSB1 which is lacking in ScPIN3. AgLSB1/ PIN3 carries the SH3 domain. **AgSTF2/ YLR327C** is the *A. gossypii* homolog of the twin genes ScSTF2 and ScYLR327C both of which have ill-defined function. ScSTF2 is described as stabilizing factor for

Highly expressed at 0h (Top 20)

				SGD Description of first *S. cerevisiae* homolog
1	AGR049W	CCW12		Cell wall protein, mutants are defective in mating and agglutination, expression is downregulated by alpha-factor
2	AGR138W	SED1	SPI1	Major stress-induced structural GPI-cell wall glycoprotein in stationary-phase cells, associates with translating ribosomes, possible role in mitochondrial genome maintenance; ORF contains two distinct variable minisatellites
3	ABR028C	CWP1		Cell wall mannoprotein, linked to a beta-1,3- and beta-1,6-glucan heteropolymer through a phosphodiester bond; involved in cell wall organization
4	AAL128C	HSP26		Small heat shock protein with chaperone activity that is regulated by a heat induced transition from an inactive oligomeric (24-mer) complex to an active dimer; induced by heat, upon entry into stationary phase, and during sporulation
5	ADL036W-A	YLR307C-X		Putative protein of unknown function
6	AFR065W	NOHBY621		no description
7	AER013W	HHT1		One of two identical histone H3 proteins (see also HHT2); core histone required for chromatin assembly, involved in heterochromatin-mediated telomeric and HM silencing; regulated by acetylation, methylation, and mitotic phosphorylation
8	AER031C	TDH3		Glyceraldehyde-3-phosphate dehydrogenase, isozyme 3, involved in glycolysis and gluconeogenesis; tetramer that catalyzes the reaction of glyceraldehyde-3-phosphate to 1,3 bis-phosphoglycerate; detected in the cytoplasm and cell-wall
9	AFR701C	RNQ1		[PIN(+)] prion, an infectious protein conformation that is generally an ordered protein aggregate
10	ACL062C	UBI4		Ubiquitin, becomes conjugated to proteins, marking them for selective degradation via the ubiquitin-26S proteasome system; essential for the cellular stress response
11	AER294C	ENO2	ENO1	Enolase II, a phosphopyruvate hydratase that catalyzes the conversion of 2-phosphoglycerate to phosphoenolpyruvate during glycolysis and the reverse reaction during gluconeogenesis; expression is induced in response to glucose
12	ADR322W	UTH1	NCA3	Mitochondrial outer membrane and cell wall localized SUN family member required for mitochondrial autophagy; involved in the oxidative stress response, life span during starvation, mitochondrial biogenesis, and cell death
13	ACL131W	TRX1	TRX2	Cytoplasmic thioredoxin isoenzyme of the thioredoxin system which protects cells against both oxidative and reductive stress, forms LMA1 complex with Pbi2p, acts as a cofactor for Tsa1p, required for ER-Golgi transport and vacuole inheritance
14	ADL154C	AHP1		Thiol-specific peroxiredoxin, reduces hydroperoxides to protect against oxidative damage
15	AAR084W	ADH2		Glucose-repressible alcohol dehydrogenase II, catalyzes the conversion of ethanol to acetaldehyde; involved in the production of certain carboxylate esters; regulated by ADR1
16	ABR026C	CWP1		Cell wall mannoprotein, linked to a beta-1,3- and beta-1,6-glucan heteropolymer through a phosphodiester bond; involved in cell wall organization
17	AFR320W	LSB1	PIN3	Protein containing an N-terminal SH3 domain; binds Las17p, which is a homolog of human Wiskott-Aldrich Syndrome protein involved in actin patch assembly and actin polymerization
18	AFR505C	STF2		Protein involved in regulation of the mitochondrial F1F0-ATP synthase; Stf1p and Stf2p act as stabilizing factors that enhance inhibitory action of the Inh1p protein
19	AEL329W	PIL1		Primary component of eisosomes, which are large immobile patch structures at the cell cortex associated with endocytosis, along with Lsp1p and Sur7p; null mutants show activation of Pkc1p/Ypk1p stress resistance pathways
20	AFR312W	NCE102		Protein of unknown function; contains transmembrane domains; involved in secretion of proteins that lack classical secretory signal sequences; component of the detergent-insoluble glycolipid-enriched complexes (DIGs)

Table 2-1 A SGD description of the top20 highly expressed genes in a germinating culture at 0 h.

regulation of mitochondrial ATP synthase and ScYLR327C was recently found among 77 novel proteins associated with ribosomes (Fleischer *et al.*, 2006). Again the function of the *A. gossypii* homolog is difficult to predict.

One highly expressed histone

AgHHT1 is one of the two *A. gossypii* genes encoding histone H3. The massive overexpression of the AgHHT1 genes is intriguing especially because transcripts of all other histone genes are much less abundant in *A. gossypii* spores. It should be noted that the Affymetrix array lacks a probe set for AgHHF1, one of the two genes for histone H4. The "missing" gene is expressed from the same bidirectional promoter as the highly expressed histone H3 gene. We hypothesize that the "missing" H4 gene is also highly expressed in spores potentially indicating massive gene silencing during sporulation with the help of theH3/H4 histone pair.

Exceptionally highly expressed genes of the C2/C3 metabolism

AgTDH3 coding for Glyceraldehyde-P-dehydrogenase was already mentioned as a presumptive cell wall component despite the clearly predictable function in glycolysis and also in gluconeogenesis. Under certain conditions, both processes seem to operate in parallel in later developmental stages as will be described in the chapter "Induction of the glyoxylate shunt." **AgENO1/2** is the syntenic homolog of the twin genes ScENO1 and ScENO2 both of which catalyze a common step in glycolysis and gluconeogenesis, the conversion of 2-Phosphoglycerate to Phosphoenolpyruvate and vice versa. **AgADH2** is the syntenic homolog of ScADH2 coding for a alcoholdehydrogenase. The transcription is induced when ethanol is used as carbon source and repressed by glucose. The gene coding for the production

Highly expressed at 9h (Top 20)

0h rank	2h	5h	7h	9h rank	Affy ID	Systematic Name	Common Name		Exp 0h	Exp 9h
1				1	Ag004168_at	AGR049W	CCW12(*)		27268	25150
2				2	Ag003796_at	AGR138W	SED1(*)	SPI1(*)	19809	22464
8				3	Ag002696_at	AER031C	TDH3(*)		11559	20893
341				4	Ag002059_at	AEL151C	RPS22A		1374	20157
538				5	Ag000758_at	ABR033C	RPS25A	RPS25B	969	19970
138				6	Ag001088_at	ACR255C	RPS21B	RPS21A	2590	18269
6				7	Ag002925_at	AFR065W	NOHBY621(*)		12087	17702
708				8	Ag001165_at	ACL140C	RPS0A	RPSB	793	17685
108				9	Ag003351_at	AFR356C	ANB1	HYP2	3078	17626
154				10	Ag001066_at	ACR065C	RPS27B	RPS27A	2426	17350
169				11	Ag002186_at	AEL152W	RPS14B	RPS14A	2221	17285
384				12	Ag000262_at	AAR177W	RPS2		1249	16657
433				13	Ag003594_at	AFR199C	ASC1		1156	16447
11				14	Ag002371_at	AER294C	ENO2	ENO1	10717	15868
464				15	Ag003332_at	AFR214C	RPL15B	RPL15A	1109	15565
557				16	Ag001397_at	ADL157C	RPL18A	RPL18B	946	15425
481				17	Ag003644_at	AFR330C	RPL20B	RPL2A	1074	15371
405				18	Ag000575_at	ABR041C	RPS20		1207	15286
212				19	Ag001020_at	ACR104C	RPL5		1921	14745
293				20	Ag003170_at	AFL163C	RPL36B	RPL36A	1500	14509

Figure 2-4 B Ranking of the 20 most abundant mRNA species at 9 h
(*) cell wall protein genes (inferred from homology to *S. cerevisiae*)
(*) newly identified *A. gossypii* specific cell wall protein gene

color code:
in comparison to average of each gene:
overexpressed underexpressed

of ethanol from acetaldehyde during yeast glycolysis is ScADH1 the name of which is misleading because it acts as acetaldehyde reductase considering the glycolytic flux. *A. gossypii* carries a syntenic homolog called AgADH1 (see chapter 3).

Protein degradation in spores

AgUBI4 codes for a ubiquitin 5-mer as concluded from the same size of the ScUBI4 gene. The high level of mRNA indicates that many proteins in *A. gossypii* spores are potentially marked for degradation. The targets may be degraded during sporulation in an early stage of germination.

Redox homeostasis and protection against oxidative damage

The generation of reactive oxygen species like superoxid anion, hydroxyperoxid and alkyl-hydroperoxids can not be prevented in aerobic metabolism. Most cells have several protection systems, e.g. peroxidase, which use either thioredoxins, 100 aa proteins carrying two cysteins, or peroxiredoxins, small proteins carrying either two cysteins or one cystein. The SH-groups of cysteins can oxidize to S-S, -SOH or –SOOH groups either as response to active oxygen species or to control the redox equilibrium together with the tripeptide glutathione (γGlu-Cys-Gly). Oxidized thioredoxins are reduced by redutases and NADPH as cofactor. Oxidized peroxiredoxins are reduced by glutathione-dependent reductases. It is not surprising to find genes encoding components of these systems in spores. **AgTRX1/2** codes for thioredoxin and is the syntenic homolog of the highly conserved twin ScTRX1 and ScTRX2, and **AgAHP1** encodes an alkyl-hydroperoxid reductase as concluded from the known function of the ScAHP1. It should be mentioned that **AgTSA1/2** encoding a thiol-specific antioxidant protein is highly expressed in *A. gossypii* spores.

Highly expressed at 9h (Top 20)

				SGD Description of first *S. cerevisiae* homolog
1	AGR049W	CCW12		Cell wall protein, mutants are defective in mating and agglutination, expression is downregulated by alpha-factor
2	AGR138W	SED1	SPI1	Major stress-induced structural GPI-cell wall glycoprotein in stationary-phase cells, associates with translating ribosomes, possible role in mitochondrial genome maintenance; ORF contains two distinct variable minisatellites
3	AER031C	TDH3		Glyceraldehyde-3-phosphate dehydrogenase, isozyme 3, involved in glycolysis and gluconeogenesis; tetramer that catalyzes the reaction of glyceraldehyde-3-phosphate to 1,3 bis-phosphoglycerate; detected in the cytoplasm and cell-wall
4	AEL151C	RPS22A		Protein component of the small (40S) ribosomal subunit; nearly identical to Rps22Bp and has similarity to E. coli S8 and rat S15a ribosomal proteins
5	ABR033C	RPS25A	RPS25B	Protein component of the small (40S) ribosomal subunit; nearly identical to Rps25Bp and has similarity to rat S25 ribosomal protein
6	ACR255C	RPS21B	RPS21A	Protein component of the small (40S) ribosomal subunit; nearly identical to Rps21Ap and has similarity to rat S21 ribosomal protein
7	AFR065W	NOHBY621		no description
8	ACL140C	RPS0A	RPSB	Protein component of the small (40S) ribosomal subunit, nearly identical to Rps0Bp; required for maturation of 18S rRNA along with Rps0Bp; deletion of either RPS0 gene reduces growth rate, deletion of both genes is lethal
9	AFR356C	ANB1	HYP2	Translation initiation factor eIF-5A, promotes formation of the first peptide bond; similar to and functionally redundant with Hyp2p; undergoes an essential hypusination modification; expressed under anaerobic conditions
10	ACR065C	RPS27B	RPS27A	Protein component of the small (40S) ribosomal subunit; nearly identical to Rps27Ap and has similarity to rat S27 ribosomal protein
11	AEL152W	RPS14B	RPS14A	Ribosomal protein 59 of the small subunit, required for ribosome assembly and 20S pre-rRNA processing; mutations confer cryptopleurine resistance; nearly identical to Rps14Ap and similar to E. coli S11 and rat S14 ribosomal proteins
12	AAR177W	RPS2		Protein component of the small (40S) subunit, essential for control of translational accuracy; has similarity to E. coli S5 and rat S2 ribosomal proteins
13	AFR199C	ASC1		G-beta protein for Gpa2p; involved in translation regulation; required for repression of Gcn4p activity in the absence of amino-acid starvation; core component of the ribosome; ortholog of mammalian RACK1
14	AER294C	ENO2	ENO1	Enolase II, a phosphopyruvate hydratase that catalyzes the conversion of 2-phosphoglycerate to phosphoenolpyruvate during glycolysis and the reverse reaction during gluconeogenesis; expression is induced in response to glucose
15	AFR214C	RPL15B	RPL15A	Protein component of the large (60S) ribosomal subunit, nearly identical to Rpl15Ap and has similarity to rat L15 ribosomal protein; binds to 5.8 S rRNA
16	ADL157C	RPL18A	RPL18B	Protein component of the large (60S) ribosomal subunit, identical to Rpl18Bp and has similarity to rat L18 ribosomal protein; intron of RPL18A pre-mRNA forms stem-loop structures that are a target for Rnt1p cleavage leading to degradation
17	AFR330C	RPL20B	RPL2A	Protein component of the large (60S) ribosomal subunit, nearly identical to Rpl20Ap and has similarity to rat L18a ribosomal protein
18	ABR041C	RPS20		Protein component of the small (40S) ribosomal subunit; overproduction suppresses mutations affecting RNA polymerase III-dependent transcription; has similarity to E. coli S10 and rat S20 ribosomal proteins
19	ACR104C	RPL5		Protein component of the large (60S) ribosomal subunit with similarity to E. coli L18 and rat L5 ribosomal proteins; binds 5S rRNA and is required for 60S subunit assembly
20	AFL163C	RPL36B	RPL36A	Protein component of the large (60S) ribosomal subunit, nearly identical to Rpl36Ap and has similarity to rat L36 ribosomal protein; binds to 5.8 S rRNA

Table 2-1 B SGD description of the top20 highly expressed genes in a germinating culture at 9 h

High expression of cortical patch genes in spores

AgLSB1/PIN3 was already mentioned as a gene coding for a actin patch assembly protein with SH3 domain. Due to the prion-like function of the twin gene ScPIN3 in yeast this *A. gossypii* gene was viewed as an assembly/aggregation factor. ScLSB1 is known to interact with ScLAS17 an essential protein for actin patch initialization in endocytosis. Therefore we view AgLSB1/PIN3 also as endocytosis component without really understanding why its mRNA is so abundant in *A. gossypii* spores. **AgPIL1** is the syntenic homolog of ScPIL1, a highly expressed component of eisosomes, originally discovered in the group of J. A. Cooper as novel cortical patch structure which also affects sphingolipid metabolism (Young *et al.*, 2002) It is highly expressed in yeast and therefore the similarly high expression level in *A. gossypii* is not surprising.

Glycolipid-enriched complexes in spores

Research in the area of Yeast lipid rafts has led to the isolation of detergent-insoluble glycolipid-enriched complexes. Among other ScNCE102 (formerly called ScNCE2) was identified as one of the components (Bagrat et al., 2000). This protein has been characterized previously in a screen to identify candidate proteins allowing secretion without classical secretory signal sequences. Hence the name NCE for non-classical export (Cleves *et al.*, 1996). *S. cerevisiae* carries a twin gene, ScYGR131W. The function of the encoded protein is not known. Sequence comparison shows conservation of four transmembrane domains and the signal peptide. It is surprising to find mRNA of the *A. gossypii* syntenic homolog **AgNCE102/YGR131W** among the most abundant mRNAs in spores. One may speculate that the strange needle-shaped form of *A. gossypii* spores is partly controlled by lipid compositions and/or non-classical protein export.

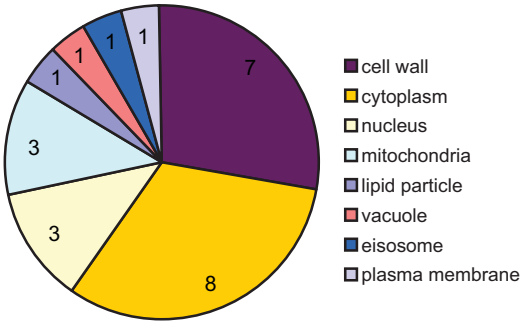
The most abundant mRNA species in spores and germlings differ strongly

Figure 2-5 shows significantly enriched GO categories within the top 20 most highly expressed genes. The categories “GO:5618 cell wall” and “GO:9277 cell wall (sensu Fungi)” were still significantly enriched after 9 h incubation time with 4 out of 20 genes (p = 1.15E-3 and p = 1.15E-3).
A visual representation of the changing gene expression in spores and germlings is given in figure 2-4 B. Genes that are down-regulated within the first nine hours change their color in the heat map from red to blue while up-regulated genes change from blue to yellow. Figure 2-4 B shows clearly that all fifteen ribosomal protein genes are coordinately up-regulated while the spores wake up. Figure 2-4 A shows that some of the cell wall protein genes remain expressed at high level (CCW12, SED1/SPI1, NOHBY621, TDH3) while transcripts of other cell wall proteins are found in decreasing amounts (CWP1d, UTH1/NCA3, CWP1b). Dramatic changes in expression intensity levels as seen for CWP1d from 18698 to 531 suggest active mRNA degradation. Additional genes that were found at a constant high level were the histone 3 gene HHT1 and the genes involved in central carbon metabolism TDH3 and ENO1/2. The thioredoxin gene TRX1/2 and the alkyl hydroxyperoxid reductase gene AHP1 as well as genes LSB1/PIN3, STF2, PIL1 and NCE102 were down-regulated.

Gene clustering of microarray expression profiles during spore germination

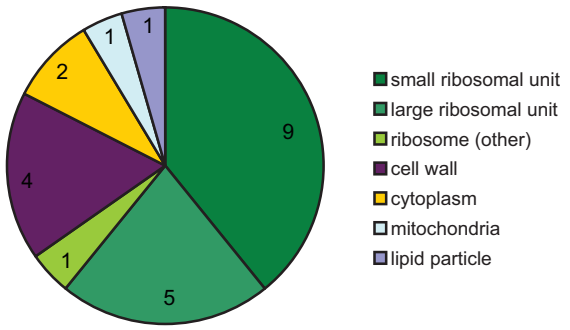
The aim of clustering is to group together

A
GO “cellular component” affiliation of top 20 highly expressed genes at 0h



Gene Ontology "Cellular component"	% of genes in category	p-value
GO:30312: external encapsulating structure	29.41	8.13E-07
GO:5618: cell wall	29.41	8.13E-07
GO:9277: cell wall (sensu Fungi)	29.41	8.13E-07
GO:15: phosphopyruvate hydratase complex	5.882	4.26E-03
GO:788: nuclear nucleosome	5.882	4.18E-02
GO:786: nucleosome	5.882	4.18E-02

B
GO “cellular component” affiliation of top 20 highly expressed genes at 9h



Gene Ontology "Cellular component"	% of genes in category	p-value
GO:5830: cytosolic ribosome (sensu Eukaryota)	68.42	1.91E-17
GO:5840: ribosome	78.95	5.28E-17
GO:5829: cytosol	73.68	3.71E-14
GO:30529: ribonucleoprotein complex	78.95	3.98E-13
GO:5843: cytosolic small ribosomal subunit (sensu Eukaryota)	42.11	1.35E-12
GO:16283: eukaryotic 48S initiation complex	42.11	1.35E-12
GO:16282: eukaryotic 43S preinitiation complex	42.11	1.14E-11
GO:15935: small ribosomal subunit	42.11	3.44E-10
GO:43232: intracellular non-membrane-bound organelle	78.95	1.16E-08
GO:43228: non-membrane-bound organelle	78.95	1.16E-08
GO:43234: protein complex	89.47	1.58E-07
GO:5842: cytosolic large ribosomal subunit (sensu Eukaryota)	26.32	1.88E-06
GO:15934: large ribosomal subunit	26.32	4.48E-05
GO:30312: external encapsulating structure	15.79	1.15E-03
GO:5618: cell wall	15.79	1.15E-03
GO:9277: cell wall (sensu Fungi)	15.79	1.15E-03
GO:15: phosphopyruvate hydratase complex	5.263	4.76E-03
GO:5737: cytoplasm	89.47	2.44E-02
GO:5732: small nucleolar ribonucleoprotein complex	10.53	2.84E-02

Figure 2-5 The most abundant mRNA species in spores and germlings differ strongly. Affiliation of the top 20 most highly expressed genes at 0 h and at 9 h to the Gene Ontology (GO) „cellular component“. The number of genes is given in the pie chart. Please note that one gene can belong to more than one functional category. (A) Spores at 0h are significantly enriched for mRNAs encoding cell wall proteins (p = 8.13E-07). None of the ribosomal protein genes is among the top 20. (B) After 9h in full medium the cell wall protein genes are still significantly enriched (p = 1.15E-3) but the majority of the genes among the top 20 encodes small (p = 3.44E-10) or large (p = 4.48 E-10) ribosomal subunits.

similarly expressed genes and then try to correlate the observations to biology. Clustering organizes the data into relatively homogenous groups with similar patterns irrespective of absolute average expression strengths. To improve cluster quality we applied cleaning steps before clustering. These are summarized in figure 2-6. First, 244 unspecific inter-ORF- or sense probe sets were removed from the data set. Then, 260 probe sets with intensity values below 60 (assumed backround noise) in all ten samples were removed, considered “not expressed” and grouped into a separate cluster. Last, 1511 genes with constant expression levels within the first nine hours in full medium were assigned to a “constant expression” cluster. They were identified by selecting genes that passed a 1-way ANOVA (Analysis of variance) with $p \leq 0.05$ and changed at least 2-fold when the maximum value was compared to the minimum value in the time course experiment. The remaining 2681 genes were clustered based on their expression profiles across the five time points. The k-means clustering algorithm with pearson correlation as similarity measure was used starting from five initial clusters. The five clusters that were obtained were grouped into two up-regulated, one down-regulated and two up- and down-regulated clusters. Individual clusters were named by virtue of the time points of their maximal and minimal average expression, “**Min0-up**”, “**Max9-up**”, “**Max0-down**”, “**Max2-up-down**”, “**Min2-down-up**”, “**constant expression**” and “**not expressed**”. The “**Max-0-down**” cluster

was manually sub-clustered into a cluster containing the genes which were strongly down-regulated within the first two hours (top10 "**strongly down-regulated**" genes). To obtain each of the eight profiles shown in figure 2-7, the sum of each gene's expression value was standardized to 1. Next, time course values for all genes in each cluster were averaged, and the average value for each cluster was scaled to 1. To search for biological functions or processes that were over- or under-represented in a given gene cluster, we used Gene Ontologies (GO) and p-values coming from Fisher's exact tests. Gene Ontology is a collaborative effort to assign controlled vocabulary to gene products in a species independent manner.

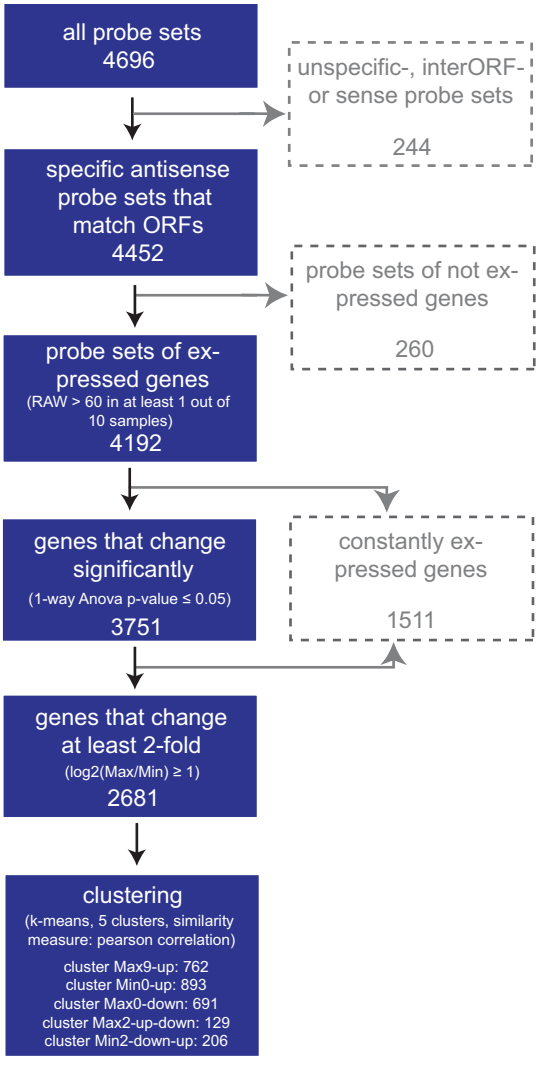


Figure 2-6 Data cleaning steps before clustering

The categories are hierarchically ordered into sub-categories with increasing level of specificity. A single gene can belong to more than one functional GO category. Lists with genes belonging to each cluster and lists with significantly enriched biological process categories for each cluster are available as supplementary material 1 on the compact disc accompanying this thesis.
Genes predicted to be in a particular functional category were not distributed evenly among the eight clusters. We will discuss for each cluster the top two or three enriched categories.

For example genes involved in "GO:30476:spore wall assembly" (p = 3.36E-12), in "GO:7126:meiosis" (p = 2.26E-11) and "GO:30435:sporulation" (p = 2.73E-9) were significantly over-represented in the "not expressed" cluster. 21 out of total 92 genes classified into the GO category "sporulation" fell into this cluster. However the category "GO:30435:sporulation" was also significantly enriched (p = 2.37E-4) in the "**Max0-down**" with 25 out of total 92 sporulation genes. A later section of this chapter will be exclusively dedicated to the expression profiles of *A. gossypii* homologs of genes which are known be sporulation-specific in *S. cerevisiae*.

Genes that function in transcription (GO:6350), protein modification (GO:6464) and chromatin modification (GO:16569) were significantly over-represented in the "constant expression" cluster (p = 8.72E-07, p = 7.95E-6 and p = 2.11E-5).

The "**Min-0-up**" cluster was enriched for the categories "protein biosynthesis" (GO:6412) and "ribosome biogenesis and assembly" (GO:42254) with 179 out of 328 and 125 out of 232 genes (p = 1.05E-30 and p = 2.77E-28). In addition several categories with functions related to ribosomal RNA, for instance rRNA processing, rRNA metabolism, rRNA modification, rRNA export from the nucleus and rRNA transport were found to be enriched in this cluster.

It is known form *N. crassa* that a large fraction of macromolecular biosynthesis in germinating conidia is devoted to ribosomal RNA and the protein synthetic machinery (Sachs *et al.*, 1991). The data on increase in relative mRNA levels of genes involved in ribosomal protein functions during spore germination in *A. gossypii* are consistent with results from similar studies in the filamentous fungus *N. crassa*, where they found that 49 out of 56 genes functioning in ribosome biogenesis were members of a cluster with maximal expression between 1 and 4 hours after spore inoculation in glucose-rich medium (Kasuga *et al.*, 2005). Kasuga et al. reported also that the average expression profile of the cluster into which most ribosome genes fell increased between 0 and 4 hours and then decreased from 4 to 16 hours. In *A. gossypii* we also measured a decrease in the expression of the highly co-regulated ribosomal protein genes. As presented later (Chapter 4 and Appendix 1), we measured a modest decrease in the average ribosomal gene expression from 9 h to 18 h (log2(18 h/9 h)=0.42) and almost a two-fold decrease from 9 h to the outer zone of a fast growing colony (log2(OZ/9 h)=0.96). A four-fold decrease in average ribosomal gene expression was observed in the comparison of the inner zone of a sporulating colony with 9 h (log2(IZ/9 h)=2.00).

The "**Min0-up**" cluster is characterized by a very steep increase in gene expression within the first two hours. The second up-regulated cluster, "**Max9-up**", is characterized by a constant increase from 0 h to 7 h. Genes predicted to be involved in oxidative phosphorylation (GO:6119) and in generation of precursor metabolites and energy (GO:6091) were over-represented in this cluster (p = 1.79E-12 and p = 1.61E-8). Furthermore the functional categories cellular respiration (GO:45333; p = 4.28E-8) and aerobic respiration (GO:9060; p = 3.06E-7) were found to be enriched in the "Max9-up" cluster. mRNAs belonging to these categories are increasingly synthesized, even in spores between 0 h and 2 h where no morphological changes are microscopically visible. The average expression of the cluster changes at least 2-fold from 0 h to 2 h and at least 7-fold from 0 h to 5 h. It would be interesting to measure the oxygen consumption of a germinating culture and correlate it with the expression profile. We would expect to find increasing oxygen consumption during spore germination.

The summed expression values of genes belonging to the "**Max0-down**" cluster decrease with increasing inoculation time. The two significantly enriched categories with the lowest p-values were amino acid biosynthesis (GO:8652; p = 2.74E-8) and amine biosynthesis (GO:9309; p = 1.66E-07). Transcripts from 32 out of total 88 genes involved in amino acid biosynthesis were down-regulated during germination. A comparison with *N. crassa* shows that conidial extracts contained free pools of amino acids and transcripts from 25 genes predicted to be involved in amino acid biosynthesis were detected to peak as early as 1h after inoculation and then transcript numbers dropped dramatically between 1h and 2 h and modestly between 2 h and 12 h.

The "**Max0-down**" was, as previously mentioned, also enriched for transcripts from sporulation genes (GO:30435; p = 2.73E-4). The categories "cellular response to nutrient levels" (GO:31669; p = 3.30E-3) and "cellular response to starvation" (GO:9267; p = 3.3E-3) were over-represented with 5 out of total 9 genes, each.

The "**Max0-down**" cluster was manually sub-clustered into "**Max0-strong-down**", a cluster containing 10 genes that were found by filtering based on fold change and statistical significance (volcano plot). We identified the genes DAL82, PRD1(AGR406C), FBP1, DOG1/2, LEU4, ICL1, MEP1/2, CIT1/2, MDH2 and AER090W to be at least 4-fold down-regulated from 0 h to 2 h with t-test p-values smaller than 0.05. All genes, except for CIT1/2 and AER090W which were up-regulated after

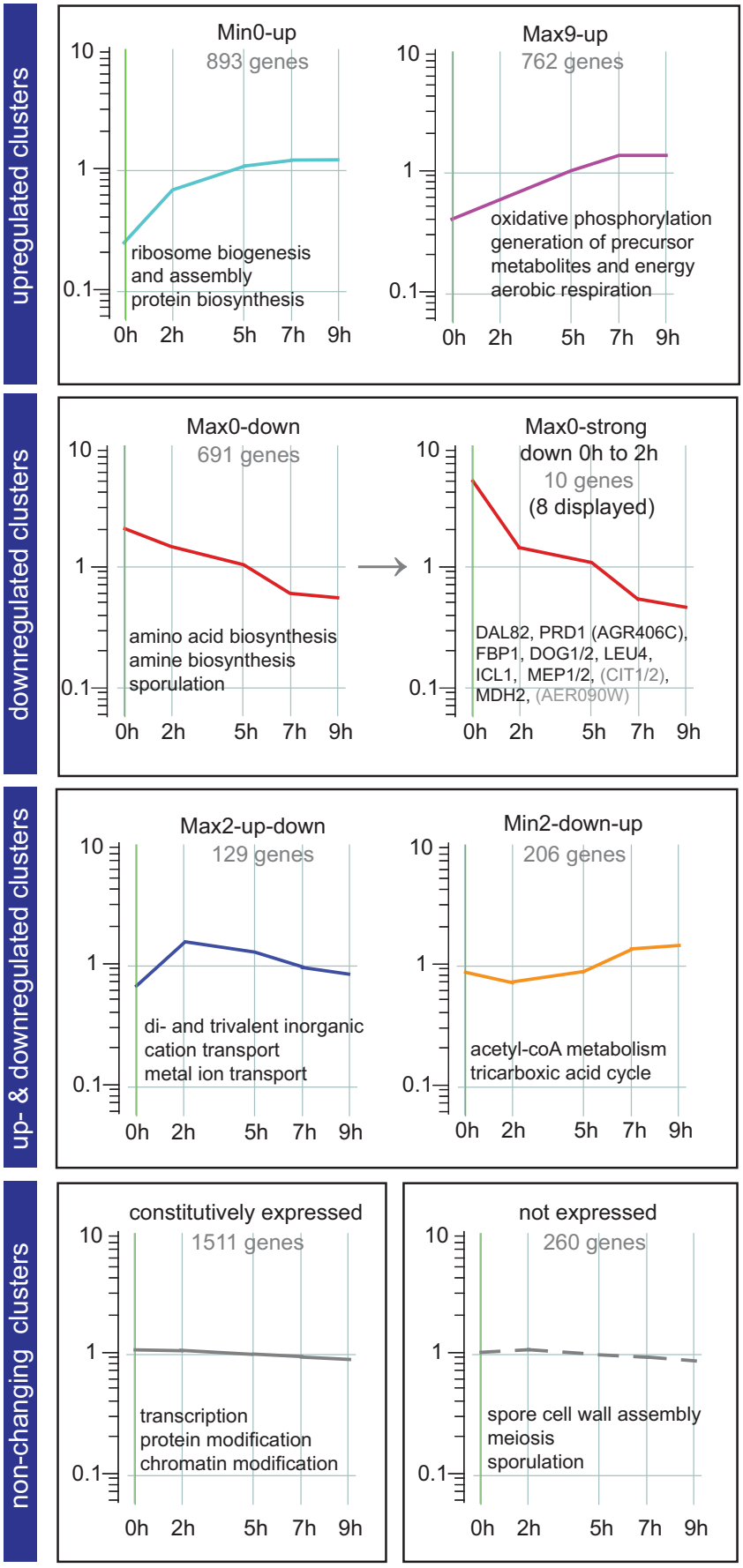


Figure 2-7 Clustering of 4'452 *A. gossypii* genes based on the expression profiles across five time points. The average of each cluster scaled to 1 and the top two or three significantly enriched GO categories are shown.

2 hours, followed a strict down-regulated pattern during germination. DAL82, DOG1/2 and ICL1 showed a strong down-regulation pattern and were not expressed any more after 9 h incubation. If, during germination, these genes were not transcribed any more, there would still be an intensity signal left after 9h. This is because 20% of the spores had not germinated after 9 h incubation time. Assuming a spore viability of 100% and the same RNA extraction efficiency for all cell types, we would expect to still see at 9 h 1/5 of the initial signal measured at 0h. We suggest that the DAL82, DOG1/2 and ICL1 mRNAs are actively degraded after spore inoculation. Protein degradation is very likely during germination because high mRNA levels of AgUBI4 (encoding ubiquitin) have been found in spores.

The remaining two clusters were mixed up- and down-regulated clusters. The 129 genes in the "**Max2-up-down**" cluster are strongly up-regulated from 0 h to 2 h and then down-regulated between 2 h and 9 h. Categories with the lowest p-value in this cluster were di- and trivalent inorganic cation transport (GO:15674; p = 2.15E-5) and metal ion transport (GO:30001; p = 3.65E-5).

The "**Min2-down-up**" contained 206 genes and was enriched for the categories "acetyl-CoA catabolism" (GO:46356; p = 1.10E-12) and "tricarboxylic acid cycle" (GO:6099, P = 1.10E-12) with 11 genes out of total 15 genes, each.

2.3 Material and Methods

Gene expression data

Whole-genome expression data of *A. gossypii* spores or mycelium coming from the laboratory strain AgΔleu2Δthr4 (#ATCC10895) or the natural isolate strain FDAG were obtained as described in chapter 1 of this publication.

Germination counts

Germination of a culture of *A. gossypii* spores is not a synchronous process. The following cell types coexist in a culture 9 h after inoculation: ungerminated spores, spores with a germ bubble, unipolar and bipolar germlings. The cell type composition was counted in relation to the time point post inoculation. The total number of cell types counted was n = 200 per array. Replicate samples were inspected individually and the average composition was calculated per time point. 500 ml samples for investigation of cell type composition were collected shortly before RNA isolation and fixed with 3% formaldehyde for 1h on a shaker at room temperature. The cells were then washed twice with 1 x PBS and stored at -4°C until needed. Trials to reliably count the cell types in suspension under the microscope failed because the hydrophobic spores were sticking tightly together. We developed a protocol that allows better separation of the cells. Cells were transferred to 0.03% Triton and plated with a glass Drigalsky on plates filled with a 2 mm thin 1.8% agar layer. Microscopy was done with an Axioplan 2 microscope (Carl Zeiss, Germany) equipped with a 20 x objective. Pictures were evaluated with the MetaMorph imaging software.

Hierarchical clustering (Condition tree)

Hierarchical clustering is often used to organize the data based on conditions. It can be used as a quality control step when two or more conditions with at least two replicates are tested. Replicates are expected to cluster together first before clustering together with other samples. Hierarchical clustering works by calculating a distance matrix containing all pair wise

distances. The algorithm iteratively joins the two closest clusters starting from single clusters. After each step a new distance matrix is recalculated until the algorithm stops. Distances can be calculated as follows: single linkage (nearest neighbor in cluster is used to calculate distance), complete linkage (furthest neighbor in cluster is used to calculate distance) or average linkage (average distance from any member of the cluster is used to calculate the distance). Hierarchical clustering was done in GeneSpring GX (Agilent) using pearson correlation and the average linkage method. The tree in figure 2-3 A was clustered according to the samples. The genes are displayed horizontally in increasing Affymetrix ID order.

Pairs plots

Combined scatter plots were made as described in chapter 1.

Principle component analysis (PCA)

PCA is a technique for simplifying a data set, by reducing multiple dimension data sets to lower dimensions for analysis such that the greatest variance by any projection of the data comes to lie on the first coordinate (called first principle component), the second greatest variance on the second coordinate and so on. PCA analysis on the laboratory strain (ΔlΔt) captured 78.66% of the variance by the first dimension and 13.59% by the second dimension; analysis on the natural isolate (FDAG) captured 63.53% on the first dimension and 18.58% of the variance. In the 2D-plot biological replicates were close together and clearly separated from non-replicate measurements, except for 7 h and 9 h where the replicate measurements lay close together but showed little inter-time point variance.

Data selection for clustering analysis

The germination data set coming from the laboratory strain ΔlΔt can be described as a matrix of 10 experiments (2 x 5 time points) and 4696 probe sets encoding 4452 genes. The aim of clustering is to group together similarly expressed genes and then try to correlate the observations to biology. Clustering organizes the data into relatively homogenous groups with similar patterns irrespective of expression levels. This means for instance that a gene that changes its expression level from 10 to 60 might be grouped into an up-regulated cluster even though it is not expressed at any time. To improve cluster quality we applied cleaning steps before clustering. These are summarized in figure 2-6. First, 244 unspecific inter-ORF- or sense probe sets were removed from the data set. Then, 260 probe sets with intensity values below 60 in all ten samples were removed. Corresponding genes were considered “not expressed” and grouped into a separate cluster. Last, 1511 genes with constant expression levels within the first nine hours in full medium were assigned to a “constant expression” cluster. They were identified by selecting genes that passed a 1-way ANOVA (Analysis of variance) with $p \leq 0.05$ and changed at least 2-fold when the maximum value was compared to the minimum value in the time course experiment. The remaining 2681 genes were pre-selected for clustering.

Statistical testing

Statistical test are used to address the question whether the differences between two or more observations can be explained by chance alone. T-test and ANOVA (Analysis of Variance) compare the means of the groups. ANOVA (Kerr *et al.*, 2000) is the generalization of the t-test for more than two groups. ANOVA compares variances in

order to test for significance between the means. The variance is calculated in two phases. Fist the sums of squares (error) is calculated which tells us how much individual observations deviate from the group mean (e.g. the mean of the replicates). Then the sums of squares (time) between the groups are calculated which tells us how much the individual observations deviate from the overall mean. SS (error) and SS (time) are divided by their degrees of freedom which results in MS (error) and MS (time). A F-statistic is calculated by dividing MS (time) by MS (error). A p-value can be read from the F-distribution table of critical values. If the calculated p-value is lower than 0.05 the null hypothesis that the difference between the means is by chance alone is rejected. An excellent publication that describes ANOVA as a statistical method to detect differentially expressed genes is "Using ANOVA for gene selection from microarry studies of the nervous system" (Pavlidis, 2003).

Data clustering

Clustering reduces the dimensionality of the data and groups together similarly expressed genes. 2681 genes from the germination experiment conducted with the laboratory strain ΔlΔt were used as input for clustering. The K-means algorithm implemented into Genespring GX (Agilent) has been used as clustering method. K-means is an iterative program without guarantee that the absolute minimum of the objective function is reached. Therefore we tested starting from five additional random clusters and checked for comparable results. The similarity measure and number of clusters was set to Pearson correlation and five, respectively. The five clusters that were obtained were grouped into two up-regulated, one down-regulated and two up- and down-regulated clusters. The ten genes with the highest log2(0 h/2 h) ratio were put into a separate sub-cluster "Max0-strongly down". Individual clusters were named by virtue of the time points of their maximal and minimal average expression.

Volcano plot

The volcano plot serves to easily identify differentially expressed genes based on fold-change and confidence. The log2 ratio is plotted versus the statistical significance given as -log10[(p-value)].

Gene Ontology

Gene Ontology is a collaborative effort to assign controlled vocabulary to gene products in a species independent manner. Standardized terms are used to describe three top categories: molecular function, biological process and cellular component. Gene Ontologies are used analytically to search for biological functions or processes that are over- or underrepresented in a given set of genes. P-values coming from a Fisher's exact test are calculated to answer the question how likely it is to find x members out of y genes in a certain category within the z first genes in the list out of w total genes in the microarray. Since hundreds or thousands of tests are performed the p-values need to be corrected for multiple testing with the False Discovery Rate (FDR) method (Benjamini, 1995). This is because the more analyses are performed the more the results will meet the standard p-value cutoff of 0.05 just by chance. We used Gene Ontologies to add a layer of biological information to the gene lists generated in the clustering step. Then we extracted from all categories that were significantly enriched in a given cluster (p-value ≤ 0.05) the two or three categories with the lowest p-value (and in some cases additionally a couple of interesting categories) including their member genes for discussion.

Chapter 3: Hyphae at super speed

Chapter 3: Hyphae at super speed

3.1 Introduction

Germ tubes emerging from the germ bubble initially grow with a tip speed of 6-10 µm/h. This speed increases during early stages of development (Knechtle et al., 2003). An advanced mycelium at 18 h after inoculation of spores in liquid full medium reaches hyphal growth speeds of up to 80 µm/h. The maximal growth speed of 180-200 µm/h is reached in *A. gossypii* hyphae growing on plates after approximately three days. These mature hyphae have the ability to split at their tips into two equally fast growing hyphae (see figure 3-1). We isolated these very fast growing hyphae from the outer zone (OZ) of two mycelial colonies 103 h after incubation as described in chapter 1. Total RNA of both preparations was extracted and processed as described earlier to obtain duplicates of

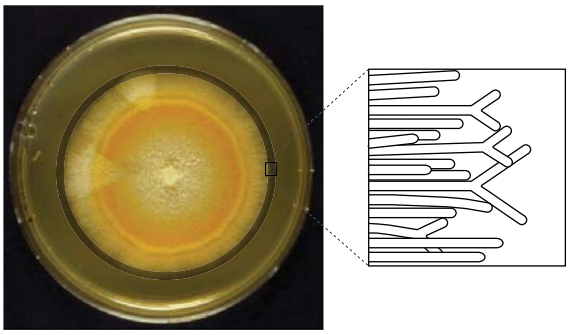


Figure 3-1 Tip splitting of mature fast hyphae in the outer zone (OZ) of the mycelial colony (OZ represented as black ring; Figure modified from M. Köhli, PhD thesis)

transcriptional profiles.
The comparison to the transcriptional profile to fast growing advanced mycelium (18 h) reveales several unexpected differences. The fifteen most up- and down-regulated genes in response to very fast hyphal growth are shown in table 3-1 and in figure 3-2. First, the function of yeast homologs of these genes was used to define, if possible, groups of genes potentially belonging to specific pathways or to similar functions. The fifteen most up- and downregulated genes between very fast growing hyphae (OZ, 180 µm/h) and slow growing hyphae (9 h germlings, 10 µm/h) were also included in this chapter (table 3-2 and figure 3-3). Possible biological implications are presented in a discussion section at the end of this chapter.

3.2 Results

Genes up-regulated in very fast hyphae (180 µm/h) compared to fast advanced hyphae (80 µm/h)

Glyoxylate cycle genes and genes for growth on nonfermentable carbon sources. Six of the top 15 up-regulated genes in response to very fast hyphal growth. This group consists of **AgICL1**, **AgMLS1**, **AgFBP1**, **AgACS1, AgPDH1 and AgYAT2** encodes enzymes for the glyoxylate cycle, gluconeogenesis and use of C-2 carbon sources (see figure 3-4). On this account we investigated the expression of additional genes needed for the pathway from C-2 to glucose as shown in table 3-3. All listed genes were highly up-regulated with the following log2(ratios):

AgADH2 (1.26), AgALD5 (3.07),
AgACS1(4.62), AgCIT3(1.87),
AgICL1(7.13), AgMLS1(6.22),
AgYAT1(3.83), AgMDH2(3.55),
AgPCK1(3.98) and AgFBP1(4.66).

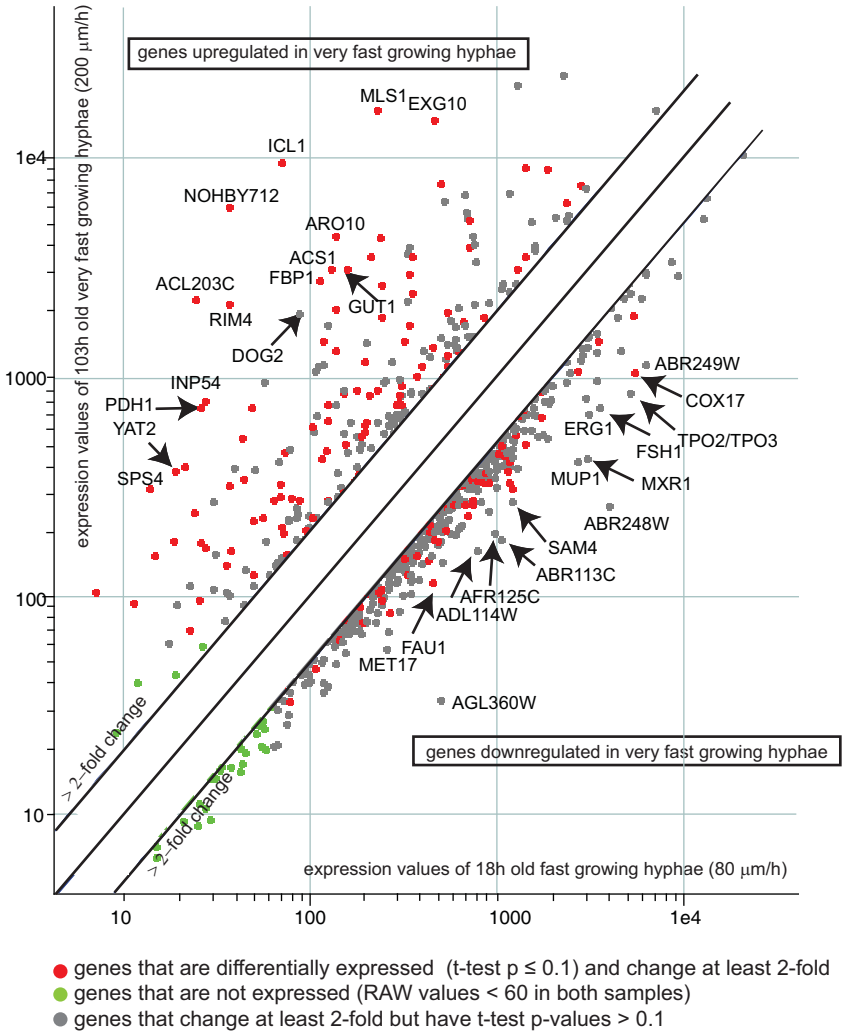
Interestingly the citrate synthase gene AgCIT1/CIT2 (single gene in *A. gossypii*, twin genes in *S. cerevisiae*) was down-regulated with a log2 ratio of (-0.63) suggesting that AgCIT3 encodes the isozyme mainly participating in the glyoxylate cycle. This is in contrast to *S. cerevisiae* where ScCIT2 encodes the glyoxylate-cycle-specific enzyme located in the peroxisomes.

ScCIT1 and ScCIT3 were found in the mitochondria.

Like citrate synthase, malate dehydrogenase (MDH) is also encoded by more than one gene in *S. cerevisiae* and *A. gossypii*: MDH1, MDH2 and MDH3. ScMdh1 is the mitochondrial, ScMdh2 the cytosolic and ScMdh3 the peroxisomal enzyme. In *S. cerevisiae* the cytosolic isozyme of ScMdh2 preferentially functions in the glyoxylate cycle. In *A. gossypii*, the three orthologous genes, showed distinct expression profiles with the following log2 ratios: AgMDH2(3.55), AgMDH3(0.99) and AgMDH1(-0.90). Similarly to *S. cerevisiae* MDH2 was strongest up-regulated followed by AgMDH3. AgMDH1 was down-regulated in very fast growing hyphae.

ScICL1, the *S. cerevisiae* homolog of **AgICL1**, encodes the isocitrate lyase enzyme I which catalyzes the formation of succinate and glyoxylate from isocitrate, the key reaction of the glyoxylate cycle. ScICL1 is glucose repressed indicating at least a partial conservation of regulation.

ScMLS1, the homolog of **AgMLS1**, codes for a malate synthase. This enzyme is involved in the utilization of nonfermentable carbon sources and localizes to the peroxisomes when cells are grown on oleic acid medium. The expression of ScMLS1 is subject to



● genes that are differentially expressed (t-test p ≤ 0.1) and change at least 2-fold
● genes that are not expressed (RAW values < 60 in both samples)
● genes that change at least 2-fold but have t-test p-values > 0.1

Figure 3-2 Scatter plot showing the top 15 up- and down-regulated genes

Table 3-3 Glyoxylate cycle genes and genes needed for growth on ethanol.

Systematic name	Common name	9h	18h	OZ (FAST)	9h	18h	OZ (FAST)	log2 (OZ/18h)
AAR004C	CIT1/CIT2	+++	+++	+++	3128	6981	4497	-0.63
AGR002W	CIT3	-	-	~	62	29	108	1.87
AAR084W	ADH2	++	+++	+++	1013	6955	16680	1.26
ACL044W	ALD5	++	++	+++	1065	673	5663	3.07
ACR268C	MLS1(*)/DAL7	+	+	+++	184	223	16561	6.22
ADL066C	ICL1(*)	-	-	+++	46	69	9609	7.13
ADR152C	MDH1	+++	+++	+++	4602	5029	2700	-0.90
ADL164C	MDH2	+	+	+++	143	334	3920	3.55
ADR252W	MDH3	++	++	++	966	894	1774	0.99
ADR408W	ACS1	~	~	+++	128	126	3104	4.62
AER224W	YAT1	~	~	+++	97	121	1730	3.83
ABR172W	YAT2	-	-	+	19	18	375	4.36
AFR292W	PCK1	+	+	+++	235	491	7725	3.98
AFR593C	FBP1	~	~	+++	119	110	2788	4.66

- < 100
~ < 150
+ < 500
++ < 2000
+++ > 2000

Asterisks denote glyoxylate cycle-specific genes. Numbers represent linear intensity signals. These linear intensity signals were grouped into expression strength classes (legend see above). The log2 intensities are given for the comparison of “very fast growing hyphae in the outer zone (OZ) of the mycelial colony” to “fast growing hyphae of an advanced mycelium (18h)”.

carbon catabolite repression.

The *S. cerevisiae* homolog of **AgFBP1** enodes the Fructose-1,6-bisphosphatase which is one of the two key enzymes in gluconeogenesis. The production of glucose is the principal function of the glyoxylate cycle.

ScACS1, the homolog of **AgACS1**, is expressed on nonfermentable carbon sources and encodes an acetyl-CoA synthetase isoform.

ScPDH1, the homolog of **AgPDH1**, codes for a mitochondrial protein that participates in respiration. ScPDH1 was reported to be induced during the diauxic shift, the changed central carbon metabolism for growth on ethanol after consumption of glucose.

The *S. cerevisiae* homolog of **AgYAT2** encodes an cytoplasmic carnitine acetyltransferase with ill defined function. Its close homolog, ScYAT1 encodes an ethanol inducible carnitine acetyltransferase involved in the transport of activated acyl groups from the cytoplasm into the mitochondrial matrix. Cytoplasmic acetyl-CoA can either be used as building block for lipid synthesis or is shuttled into the mitochondria. ScYAT1 was shown to be highly up-regulated together with glyoxylate cycle genes in *S. cerevisiae* (Lorenz *et al.*, 2001). We found that AgYAT1, the homolog of ScYAT1, was also highly induced with log2 ratio of 3.83.

Genes with no homologs in *S. cerevisiae* (NOHBYs). The gene that changed its expression most in very fast growing hyphae compared to fast growing hyphae was AGL207W (NOHBY712, No homolog in baker’s yeast). It was not expressed in fast growing hyphae at 18 h, but was very highly expressed in very fast growing hyphae at 103 h. NOHBY712 also showed the strongest up-regulation when very fast growing hyphae at 103 h were compared with slow growing hyphae at 9 h. NOHBY712 encodes a 278 aa long protein that has homologs in *K. waltii* and *K. lactis*. The function of NOHBY712 remains to be determined.

Sporulation genes. ScRIM4, homolog of **AgRIM4**, encodes a putative RNA-binding protein that is required for the expression of early and middle sporulation genes.

The second sporulation related gene that was up-regulated in very fast growing

TABLE 3-1: TOP DIFFERENTIALLY EXPRESSED GENES IN VERY FAST (200μm/h) COMPARED TO FAST (80μm/h) growing HYPHAE

	SYSTEMATIC NAME	COMMON NAME	EXP fast (18h)	EXP v.fast (103h)	LOG2 (ratio)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGICAL PROCESS	CELLULAR COMPONENT
Genes expressed at higher levels in fast growing hyphae									
UPREGULATED	AGL207W	NOHBY712	35.88	6015.50	7.39	no description	-	-	-
	ADL066C	ICL1	68.82	9608.50	7.13	Isocitrate lyase, catalyzes the formation of succinate and glyoxylate from isocitrate, a key reaction of the glyoxylate cycle; expression of ICL1 is induced by growth on ethanol and repressed by growth on glucose	glyoxylate cycle	isocitrate lyase activity	-
	ACL203C		23.92	2248.50	6.55	Putative protein with similarity to the allantoate permease (Dal5p) subfamily of the major facilitator superfamily; mRNA expression is elevated by sulfur limitation; YIL166C is a non-essential gene	transporter activity	transport	membrane
	ACR268C	MLS1/DAL7	222.50	16560.50	6.22	Malate synthase, enzyme of the glyoxylate cycle, involved in utilization of non-fermentable carbon sources; expression is subject to carbon catabolite repression; localizes in peroxisomes during growth in oleic acid medium	glyoxylate cycle	malate synthase activity	cytoplasm
	AGL038C	RIM4	35.91	2144.00	5.90	Putative RNA-binding protein required for the expression of early and middle sporulation genes	meiosis	RNA binding	cytoplasm
	ACR211W	ARO10	135.05	4441.50	5.04	Phenylpyruvate decarboxylase, catalyzes decarboxylation of phenylpyruvate to phenylacetaldehyde, which is the first specific step in the Ehrlich pathway	leucine catabolic process	pyruvate decarboxylase activity	cytoplasm
	AAR146W	EXG1/SPR1	453.75	14922.50	5.04	Major exo-1,3-beta-glucanase of the cell wall, involved in cell wall beta-glucan assembly; exists as three differentially glycosylated isoenzymes	cell wall organization and biogenesis	glucan 1,3-beta-glucosidase activity	chitin- and beta-glucan-containing cell wall
	AEL090C	INP54	26.67	774.25	4.86	Phosphatidylinositol 4,5-bisphosphate 5-phosphatase with a role in secretion, localizes to the endoplasmic reticulum via the C-terminal tail; lacks the Sac1 domain and proline-rich region found in the other 3 INP proteins	exocytosis	phosphoinositide 5-phosphatase activity	endoplasmic reticulum
	AGR003W	PDH1	25.19	723.55	4.84	Mitochondrial protein that participates in respiration, induced by diauxic shift; homologous to E. coli PrpD, may take part in the conversion of 2-methylcitrate to 2-methylisocitrate	propionate metabolic process	-	cytoplasm
	AFR593C	FBP1	109.90	2788.00	4.66	Fructose-1,6-bisphosphatase, key regulatory enzyme in the gluconeogenesis pathway, required for glucose metabolism	gluconeogenesis	fructose-bisphosphatase activity	cytosol
	ADR408W	ACS1	125.90	3104.00	4.62	Acetyl-coA synthetase isoform which, along with Acs2p, is the nuclear source of acetyl-coA for histone acetlyation; expressed during growth on nonfermentable carbon sources and under aerobic conditions	histone acetylation	acetate-CoA ligase activity	mitochondrion
	ABR127C	SPS4	13.30	310.05	4.54	Protein whose expression is induced during sporulation; not required for sporulation; heterologous expression in E. coli induces the SOS response that senses DNA damage	sporulation (sensu Fungi)	-	-
	AAL123W	DOG2/DOG1	85.41	1957.00	4.52	2-deoxyglucose-6-phosphate phosphatase, member of a family of low molecular weight phosphatases, similar to Dog1p, induced by oxidative and osmotic stress, confers 2-deoxyglucose resistance when overexpressed	response to stress	2-deoxyglucose-6-phosphatase activity	cytoplasm
	ABR172W	YAT2	18.33	375.10	4.36	Carnitine acetyltransferase; has similarity to Yat1p, which is a carnitine acetyltransferase associated with the mitochondrial outer membrane	alcohol metabolic process	carnitine O-acetyltransferase activity	cytoplasm
	ACL069C	GUT1	153.40	3119.00	4.35	Glycerol kinase, converts glycerol to glycerol-3-phosphate; glucose repression of expression is mediated by Adr1p and Ino2p-Ino4p; derepression of expression on non-fermentable carbon sources is mediated by Opi1p and Rsf1p	glycerol metabolic process	glycerol kinase activity	cytoplasm

	SYSTEMATIC NAME	COMMON NAME	EXP fast (18h)	EXP v. fast (103h)	LOG2 (ratio)	DESCRIPTION OF S. CEREVISIAE HOMOLOG	MOLECULAR FUNCTION	BIOLOGICAL PROCESS	CELLULAR COMPONENT
Genes expressed at lower levels in fast growing hyphae									
DOWNREGULATED	ABR248W		3993.00	258.90	-3.95	Putative cytoplasmic protein of unknown function	-	cytoplasm	cytoplasm
	AGL360W		494.65	33.25	-3.89	Putative protein of unknown function with similarity to oxidoreductases; HOG1 and SKO1-dependent mRNA expression is induced after osmotic shock; GFP-fusion protein localizes to the cytoplasm and is induced by the DNA-damaging agent MMS	response to toxin	cytoplasm	cytoplasm
	AFR239W	MXR1	3000.00	422.55	-2.83	Peptide methionine sulfoxide reductase, reverses the oxidation of methionine residues; involved in oxidative damage repair, providing resistance to oxidative stress and regulation of lifespan	protein-methionine-S-oxide reductase activity	cytoplasm	cytoplasm
	ABL003C	MUP1	2652.00	408.45	-2.70	High affinity methionine permease, integral membrane protein with 13 putative membrane-spanning regions; also involved in cysteine uptake	L-methionine secondary active transmembrane transporter activity	integral to plasma membrane	integral to plasma membrane
	AFR322C	TPO2/TPO3	5097.00	842.30	-2.60	Polyamine transport protein specific for spermine; localizes to the plasma membrane; transcription of TPO2 is regulated by Haa1p; member of the major facilitator superfamily	spermine transmembrane transporter activity	plasma membrane	plasma membrane
	ABR113C		1032.80	180.50	-2.52	High-affinity cysteine-specific transporter with similarity to the Dal5p family of transporters; green fluorescent protein (GFP)-fusion protein localizes to the endoplasmic reticulum; YCT1 is not an essential gene	-	endoplasmic reticulum	endoplasmic reticulum
	ABR249W		6167.50	1134.50	-2.44	Putative cytoplasmic protein of unknown function	-	cytoplasm	cytoplasm
	AAL161W	COX17	5453.50	1051.50	-2.37	Copper metallochaperone that transfers copper to Sco1p and Cox11p for eventual delivery to cytochrome c oxidase	thioredoxin peroxidase activity	mitochondrion	mitochondrion
	AFR125C		957.60	193.05	-2.31	Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to mitochondria; YMR252C is not an essential gene	-	mitochondrion	mitochondrion
	ACR076C	FSH1	3542.00	728.10	-2.28	Serine hydrolase that localizes to both the nucleus and cytoplasm; sequence is similar to Fsh2p and Fsh3p	serine hydrolase activity	cytoplasm	cytoplasm
	ADL114C		767.55	162.10	-2.24	Protein of unknown function; interacts with Sup45p (eRF1), Sup35p (eRF3) and Pab1p; has a role in translation termination efficiency, mRNA poly(A) tail length and mRNA stability	mRNA catabolic process, ribosome biogenesis and assembly, translational readthrough	nucleus	nucleus
	AAL141C	ERG1	3054.00	677.05	-2.17	Squalene epoxidase, catalyzes the epoxidation of squalene to 2,3-oxidosqualene; plays an essential role in the ergosterol-biosynthesis pathway and is the specific target of the antifungal drug terbinafine	squalene monooxygenase activity	endoplasmic reticulum	endoplasmic reticulum
	ABL125W	SAM4	1203.25	271.05	-2.15	S-adenosylmethionine-homocysteine methyltransferase, functions along with Mht1p in the conversion of S-adenosylmethionine (AdoMet) to methionine to control the methionine/AdoMet ratio	homocysteine S-methyltransferase activity	cytoplasm	cytoplasm
	ADL031W	MET17	250.50	57.34	-2.13	O-acetyl homoserine-O-acetyl serine sulfhydrylase, required for sulfur amino acid synthesis	cysteine synthase activity	cytoplasm	cytoplasm
	ADL385C	FAU1	444.95	114.85	-1.95	5,10-methenyltetrahydrofolate synthetase, involved in folic acid biosynthesis	5-formyltetrahydrofolate cyclo-ligase activity	mitochondrion	mitochondrion

hyphae was **AgSPS4**. ScSPS4 codes for a protein whose expression is induced during sporulation but is not required for sporulation.

The reason for up-regulation of these genes in very fast growing hyphae is unclear. It is conceivable that the induction of genes for consumption of ethanol (which accumulates in the first phase of mycelial growth on glucose) signals the beginning of a starvation phase thus inducing sporulation genes.

Cell wall: glucan metabolism. The t-profiler analysis revealed that the GO categories "cell wall" and "glucan metabolism" were significantly up-regulated in very fast growing hyphae (t = 5.81, E = 8.7E-06; t = 4.17, E = 4.10E-02). **AgEXG1/SPR1** (a single gene with twin homologs in *S. cerevisiae*) was among the top 15 up-regulated genes. ScEXG1 codes for the major exo-1,3-beta-glucanase of the cell wall. The enzyme is involved in cell wall beta-glucan assembly and has a possible function in cell-wall expansion during vegetative growth. ScSPR1 is a sporulation-specific exo-1,3-beta-glucanase that contributes to ascospore thermoresistance. It is probably involved in the process of spore formation and in the morphogenesis of ascospore walls. In *A. gossypii* the expression of **AgEXG1/SPR1** was highest in the outer zone of very fast growing hyphae, in the inner zone of sporulation and in spores. This suggests a sporulation-related function of the AgExg1/Spr1 protein.

Other genes. ScINP54, the homolog of **AgINP54**, codes for a phosphatidylinositol 4,5-bisphosphate 5-phosphatase with a role in secretion. It localizes to the endoplasmic reticulum via the C-terminal tail. The high up-regulation of AgINP54 indicates changes in the phosphatidylinositol composition of membranes. **AgARO10** potentially encodes a phenylpyruvate decarboxylase involved in the leucine catabolic process. **AgDOG1/AgDOG2** is also a single gene with twin homologs in the *S. cerevisiae* genome. ScDOG1 codes for a 2-deoxyglucose-6-phosphate phosphatase which is a member of a family of low molecular weight phosphatases. ScDOG2 codes, like ScDOG1, for a 2-deoxyglucose-6-phosphate phosphatase. The cellular function in *S. cerevisiae* is not clear. ScGUT1, the *S. cerevisiae* homolog of **AgGUT1**, codes for a glycerol kinase that converts glycerol to glycerol-3-phosphate. The expression of ScGUT1 is glucose repressed and derepressed on nonfermentable carbon sources similarly as shown here for the *A. gossypii* homolog.

Genes down-regulated in very fast hyphae (180 μm/h) compared to fast hyphae (80 μm/h)

Genes with *S. cerevisiae* homologs of unknown function. Of the top 15 down-regulated genes, **ABR248W** (YIR035C and YIR036C), **AGL360W** (YML131W), **ABR249W** (YIR035C and YIR036C), **AFR125C** (YMR252C) and **ADL114C** (YER049C) encode proteins without a predictable function because the *S. cerevisiae* homologs are uncharacterized genes. ABR248C and ABR249W are members of a tandem gene quadruplication in *A. gossypii* and their homologs YIR035C and YIR036C are members of a tandem gene duplication in *S. cerevisiae*.

Methionine metabolism related genes. **AgMXR1** potentially encodes enzyme that reverses the oxidation of methionine residues. It is involved in oxidative damage repair and in the regulation of lifespan. **AgMUP1** potentially encodes a high affinity methionine permease. **AgMET17** potentially encodes a o-acetyl homoserine-o-acetyl serine sulfhydrylase that is required for sulfur amino acid synthesis. **AgSAM4** potentially codes for a S-adenosylmethionine-homocysteine methyltransferase that functions along with Mht1p in the conversion of S-adenosylmethionine (AdoMet) to

methionine to control the methionine/ AdoMet ratio.

Transporters. AgTPO2/TPO3 is a single gene in *A. gossypii* with twin homologs in *S. cerevisiae*. ScTPO2 encodes a polyamine transport protein specific for spermine that localizes to the plasma membrane. ScTPO3 has a similar function. The down-regulation indicates a much decreasing need for spermine uptake from the medium in very fast growing hyphae. It could also be caused by the growth of a solid medium with less than 50% of the hyphal surface direct in contact with the nutrients. **ABR113C** (homolog of YLL055W) potentially encodes a high-affinity cysteine-specific transporter with similarity to the Dal5p family of transporters. Its down-regulation may be caused by similar reasons as proposed for the spermine transporter.

Other genes. ScCOX17, homolog of **AgCOX17**, codes for a copper metallochaperone that transfers copper for delivery to cytochrome c oxidase. **AgFSH1** potentially codes for a serine hydrolase that localizes to both the nucleus and cytoplasm. ScERG1, the *S. cerevisiae* homolog of **AgERG1**, codes for a squalene epoxidase that catalyzes the epoxidation of squalene to 2,3-oxidosqualene, an essential step in the ergosterol-biosynthesis pathway. **AgFAU1** is potentially involved in folic acid biosynthesis. An explanation for the down-regulation of these genes is presently not possible.

Differential gene expression in very fast hyphae (180 μm/h) compared to slow hyphae (10 μm/h)

There was a good correlation between the up-regulated genes that were identified in the comparison of “very fast growing hyphae (OZ) with fast growing hyphae (18 h)” and “very fast growing hyphae (OZ) with slow growing hyphae (9 h)”.

Seven of the fifteen genes (NOHBY712, AgICL1, ADL392W, AgMLS1, AgRIM4, AgDOG1/DOG2 and AgPDH1) were identified in both comparisons. This can be seen in a comparison of the upper panel in table 3-1 with the upper panel in table 3-2.

The remaining eight top up-regulated genes were: NOHBY301 (ACL091C), AgDSE4, AgCDA1/2, AgCWP1b, AgPCK1 and AgGPH1. These genes, AgGPH1, belong to one of the up-regulated gene groups known from the comparison of very fast hyphae (180 μm/h) to fast hyphae (10 μm/h).

AgPCK1 falls into the group of “use of C-carbon sources”. The yeat homolog ScPCK1 encodes the phosphoenolpyruvate carboxykinase which is the key enzyme in gluconeogenesis. This enzyme catalyzes an early step in gluconeogenesis. Glucose represses ScPCK1 transcription and accelerates mRNA degradation which is regulated by Mcm1p and Cat8p.

AgNOHBY301 is a gene of unknown function and belongs to the group of “genes with no homologs in *S. cerevisiae*”.

AgCDA1/2, **AgCWP1b** and **AgDSE4** are members of the group “cell wall”. ScCDA1 encodes a chitin deacetylase, which together with ScCDA2, is involved in the biosynthesis of the ascospore wall component chitosan. Chitosan is required for proper rigidity of the ascospore wall.

The *S. cerevisiae* homolog of **AgCWP1b** encodes a cell wall mannoprotein that is linked to a beta-1,3- and beta-1,6-glucan heteropolymer through a phosphodiester bond. ScCWP1 is involved in cell wall organization. The AgCWP1b is member of a gene familiy of seven genes in *A. gossypii*. The expression patterns of four of these genes will be discussed in chapter 5.

We had already identified a potential glucanase, AgEXG1/SPR1. Now we find an additional up-regulated potential glucanase: **AgDSE4.** Its yeast homolog ScDSE4 is a daughter cell-specific secreted protein with similarity to glucanases. It degrades cell wall from the bud side causing the bud to separate from the mother cell. It is difficult to infer the

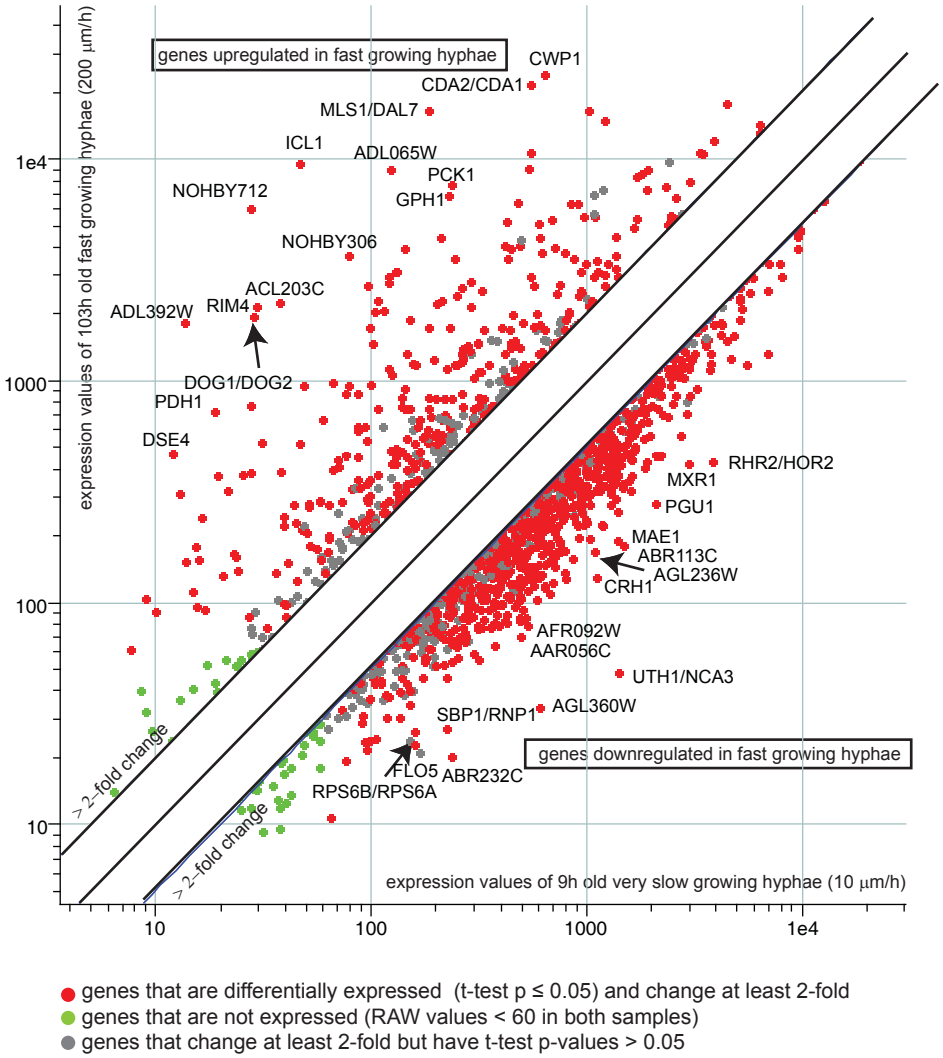


Figure 3-3 Scatter plot showing the top 15 up- and down-regulated genes

function of the *A. gossypii* homolog from *S. cerevisiae* because cell separation does not occur in *A. gossypii*, with the potential exception during spore formation. Indeed, AgDSE4 was not expressed during the first 18 hours of incubation but was expressed in the outer zone (OZ) and expressed at very high level in the inner zone of sporulating mycelium (IZ). Therefore, AgDSE4 is can be considered as a member of both classes, “cell wall: glucan metabolism” and “sporulation genes”.

The only gene that could not be assigned to one of the mentioned gene groups was **AgGPH1**. ScGPH1 encodes a non-essential glycogen phosphorylase required for the mobilization of glycogen. The expression of ScGPH1 is regulated by stress-response elements and by the HOG MAP kinase pathway. Therefore, the up-regulation of the *A. gossypii* homolog may indicate a presently unknown stress when hyphae reach maximal speed on solid medium.

TABLE 3-2: TOP DIFFERENTIALLY EXPRESSED GENES IN VERY FAST (200 μm/h) COMPARED TO SLOW (10μm/h) GROWING HYPHAE

UPREGULATED

SYSTEMATIC NAME	COMMON NAME	EXP slow (9h)	EXP v. fast (103h)	LOG2 (RATIO)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGIGAL PROCESS	CELLULAR COMPONENT
Genes expressed at higher levels in fast growing hyphae								
AGL207W	NOHBY712	27.18	6015.50	7.79	no description	-	-	-
ADL066C	ICL1	45.88	9608.50	7.71	Isocitrate lyase, catalyzes the formation of succinate and glyoxylate from isocitrate, a key reaction of the glyoxylate cycle; expression of ICL1 is induced by growth on ethanol and repressed by growth on glucose	glyoxylate cycle	isocitrate lyase activity	-
ADL392W		13.29	1811.50	7.09	Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the vacuole, while HA-tagged protein is found in the soluble fraction, suggesting cytoplasmic localization	-	-	vacuole, cell cycle-correlated morphology
ACR268C	MLS1/DAL7	183.85	16560.50	6.49	Malate synthase, enzyme of the glyoxylate cycle, involved in utilization of non-fermentable carbon sources; expression is subject to carbon catabolite repression; localizes in peroxisomes during growth in oleic acid medium	glyoxylate cycle	malate synthase activity	cytoplasm
ADL065W		121.10	8950.50	6.21	Putative protein of unknown function; expression induced under carbon limitation and repressed under high glucose			
AGL038C	RIM4	29.20	2144.00	6.20	Putative RNA-binding protein required for the expression of early and middle sporulation genes	meiosis	RNA binding	cytoplasm
AAL123W	DOG2/DOG1	28.13	1957.00	6.12	2-deoxyglucose-6-phosphate phosphatase, member of a family of low molecular weight phosphatases, similar to Dog1p, induced by oxidative and osmotic stress, confers 2-deoxyglucose resistance when overexpressed	response to stress	2-deoxyglucose-6-phosphatase activity	cytoplasm
ACL203C		36.98	2248.50	5.93	Putative protein with similarity to the allantoate permease (Dal5p) subfamily of the major facilitator superfamily; mRNA expression is elevated by sulfur limitation; YIL166C is a non-essential gene	-	-	cytoplasm
ACL091C	NOHBY306	77.21	3661.50	5.57	no description	-	-	-
AGL208C	DSE4	11.76	470.70	5.32	Daughter cell-specific secreted protein with similarity to glucanases, degrades cell wall from the daughter side causing daughter to separate from mother	cytokinesis, completion of separation	glucan 1,3-beta-glucosidase activity	chitin- and beta-glucan-containing cell wall
ADL036C	CDA2/CDA1	545.45	21797.50	5.32	Chitin deacetylase, together with Cda1p involved in the biosynthesis ascospore wall component, chitosan; required for proper rigidity of the ascospore wall	spore wall assembly (sensu Fungi)	chitin deacetylase activity	chitosan layer of spore wall
AGR003W	PDH1	18.66	723.55	5.28	Mitochondrial protein that participates in respiration, induced by diauxic shift; homologous to E. coli PrpD, may take part in the conversion of 2-methylcitrate to 2-methylisocitrate	propionate metabolic process	-	cytoplasm
ABR026C	CWP1b	638.95	23836.00	5.22	Cell wall mannoprotein, linked to a beta-1,3- and beta-1,6-glucan heteropolymer through a phosphodiester bond; involved in cell wall organization	cell wall organization and biogenesis	structural constituent of cell wall	chitin- and beta-glucan-containing cell wall
AFR292W	PCK1	235.45	7725.00	5.04	Phosphoenolpyruvate carboxykinase, key enzyme in gluconeogenesis, catalyzes early reaction in carbohydrate biosynthesis, glucose represses transcription and accelerates mRNA degradation, regulated by Mcm1p and Cat8p, located in the cytosol	gluconeogenesis	phosphoenolpyruvate carboxykinase (ATP) activity	cytosol
							glycogen catabolic process	cytoplasm
AFR547W	GPH1	226.05	6807	4.91	Non-essential glycogen phosphorylase required for the mobilization of glycogen, activity is regulated by cyclic AMP-mediated phosphorylation, expression is regulated by stress-response elements and by the HOG MAP kinase pathway	glycogen phosphorylase activity		

DOWNREGULATED

SYSTEMATIC NAME	COMMON NAME	EXP slow (9h)	EXP v. fast (103h)	LOG2 (RATIO)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGICAL PROCESS	CELLULAR COMPONENT
Genes expressed at lower levels in fast growing hyphae								
ADR322W	UTH1/NCA3	1419.50	48.30	-4.88	Mitochondrial outer membrane and cell wall localized SUN family member required for mitochondrial autophagy; involved in the oxidative stress response, life span during starvation, mitochondrial biogenesis, and cell death	mitochondrion organization and biogenesis	-	chitin- and beta-glucan-containing cell wall
AGL360W		600.70	33.25	-4.18	Putative protein of unknown function with similarity to oxidoreductases; HOG1 and SKO1-dependent mRNA expression is induced after osmotic shock; GFP-fusion protein localizes to the cytoplasm and is induced by the DNA-damaging agent MMS	response to toxin	-	cytoplasm
ABR232C		234.35	20.14	-3.54	Protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to endosomes; YLR073C is not an esssential gene	-	-	endosome
ADL071C	RHR2/HOR2	3835.50	435.25	-3.14	Constitutively expressed isoform of DL-glycerol-3-phosphatase; involved in glycerol biosynthesis, induced in response to both anaerobic and, along with the Hor2p/Gpp2p isoform, osmotic stress	response to drug	glycerol-1-phosphatase activity	cytoplasm
AGR318C	CRH1	1096.90	129.20	-3.09	Cell wall protein that functions in the transfer of chitin to beta(1-6)glucan, putative chitin transglycosidase	cell wall organization and biogenesis	-	chitin- and beta-glucan-containing cell wall
ACL071C	SBP1/RNP1	221.30	26.84	-3.04	Nucleolar single-strand nucleic acid binding protein; associates with small nuclear RNAs	35S primary transcript processing	snoRNA binding	nucleolus
ABR113C		1487.00	180.50	-3.04	High-affinity cysteine-specific transporter with similarity to the Dal5p family of transporters; green fluorescent protein (GFP)-fusion protein localizes to the endoplasmic reticulum; YCT1 is not an essential gene	-	ion transmembrane transporter activity	endoplasmic reticulum
AFL092C	FLO5	166.71	20.85	-3.00	Lectin-like protein involved in flocculation, cell wall protein that binds to mannose chains on the surface of other cells, confers floc-forming ability that is chymotrypsin resistant but heat labile; similar to Flo1p (1, 2 and see	flocculation via cell wall protein-carbohydrate interaction	mannose binding	chitin- and beta-glucan-containing cell wall
ADR327W	PGU1	2093.00	281.25	-2.90	Endo-polygalacturonase, pectolytic enzyme that hydrolyzes the alpha-1,4-glycosidic bonds in the rhamnogalacturonan chains in pectins	pseudohyphal growth	polygalacturonase activity	extracellular region
AGL068W	MAE1	1393.50	191.35	-2.86	Mitochondrial malic enzyme, catalyzes the oxidative decarboxylation of malate to pyruvate, which is a key intermediate in sugar metabolism and a precursor for synthesis of several amino acids	pyruvate metabolic process	malic enzyme activity	mitochondrion
AFR239W	MXR1	2974.00	422.55	-2.82	Peptide methionine sulfoxide reductase, reverses the oxidation of methionine residues; involved in oxidative damage repair, providing resistance to oxidative stress and regulation of lifespan	response to oxidative stress	protein-methionine-S-oxide reductase activity	cytoplasm
AAR056C		488.60	70.05	-2.80	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; green fluorescent protein (GFP)-fusion protein localizes to the nucleus	-	-	nucleus
AGR197C	RPS6B/RPS6A	156.95	22.98	-2.77	Protein component of the small (40S) ribosomal subunit; identical to Rps6Ap and has similarity to rat S6 ribosomal protein	translation	structural constituent of	cytoplasm
AFR032W		524.15	78.76	-2.73	no description	-	-	cytoplasm
AGL236W		1078.50	168.80	-2.68	Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the nucleolus	-	-	nucleolus

3.3 Discussion and Conclusions

The glyoxylate cycle

The glyoxylate cycle, also called glyoxylate shunt, encodes enzymes, necessary for the use of two-carbon compounds as carbon sources (see figure 3-4). It is found in fungi, plants and bacteria, but not in mammals. The glyoxylate cycle has been shown to be an important virulence factor in human and plant pathogenic fungi because it allows a life in a glucose-deficient environment (Lorenz *et al.*, 2001; Idnurm *et al.*, 2002).

The glyoxylate cycle involves two critical steps catalysed by the enzymes isocitrate lyase (ICL) and malat synthase (MLS). These steps bypass the two decarboxylation steps of the TCA cycle. Isocitrate lyase hydrolyses isocitrate (C6) to succinate (C4) and glyoxylate (C2) and subsequent condensation of glyoxylate and acetyl-CoA (C2) by malate synthase produces malate (C4). Malate is further oxidized to oxaloacetate (C4) which is used in three pathways. 1. The condensation of oxalacetate with acetyl-CoA generates citrate (C6) which feeds the glyoxylate cycle. 2. Oxaloacetate is the “precursor” for aspartate, methionine, threonine and isoleucin. 3. Oxaloacetate is used for gluconeogenesis because it can be converted to phosphoenol-pyruvate by PCK1. Glucose produced by gluconeogenesis is necessary, for instance for the pentose-phosphate cycle, the glycosylation of cell wall proteins, and the synthesis of cell wall glucans.
Malate, oxaloacetate, citrate and isocitrate are intermediates in the glyoxylate and TCA cycle; and the glyoxylate cycle replenishes the TCA cycle. Thus, the glyoxylate cycle is used for growth on gluconeogenic carbon sources, such as ethanol or acetate, and it is usually activated under conditions of nutrient limitation. Ethanol is oxidized to acetaldehyde by the enzymatic function of the alcohol dehydrogenase II (Adh2) and further transformed into acetate by the aldehyde dehydrogenase V (Ald5). Acetate is converted to acetyl-CoA by the glucose-repressible acetyl-CoA synthetase I (Acs1), which is essential for the entry of ethanol/ acetate into the tricarboxylic acid cycle.

Up-regulation of glyoxylate cycle genes and associated genes

The two key genes AgICL1 and AgMLS1 and genes encoding “down-stream and up-stream” enzymes were highly up-regulated in very fast growing hyphae (figure 3-4 and table 3-3). The main function of the glyoxylate cyle is its efficient support in using C2-compounds for the production of glucose via gluconeogensis. Under laboratory conditions *A. gossypii* grows very efficiently on glucose although the natural carbon source of *A. gossypii* is neutral fats in cotton seeds. This means that the *A. gossypii* metabolism is primed for growth on C2-compounds because the fatty acids released by lipase are first converted to acetyl-CoA. Glucose, a C6-compound, was the main carbon source added to the Ashbya full medium. But then, where do the C-2 compounds come from? Even though ethanol (C2-compound) production as the main endproduct in glycolysis is fairly common among certain yeast species, filamentous fungi are not well known for their fermentative abilities. Besides the fermentation of glucose to ethanol, many fungi produce a variety of organic acids, such as acetic acid, fumaric acid, malic acid and lactic acid (Skory *et al.*, 1997). In addition to genes that function directly in the glyoxylate cycle, there were many *A. gossypii* genes up-regulated that are potentially ethanol inducible, for instance AgADH2, AgPDH1, AgYAT1 and AgYAT2. We therefore propose that *A. gossypii* releases ethanol into the medium during the early stages of growth and then re-consumes the ethanol at later stages.

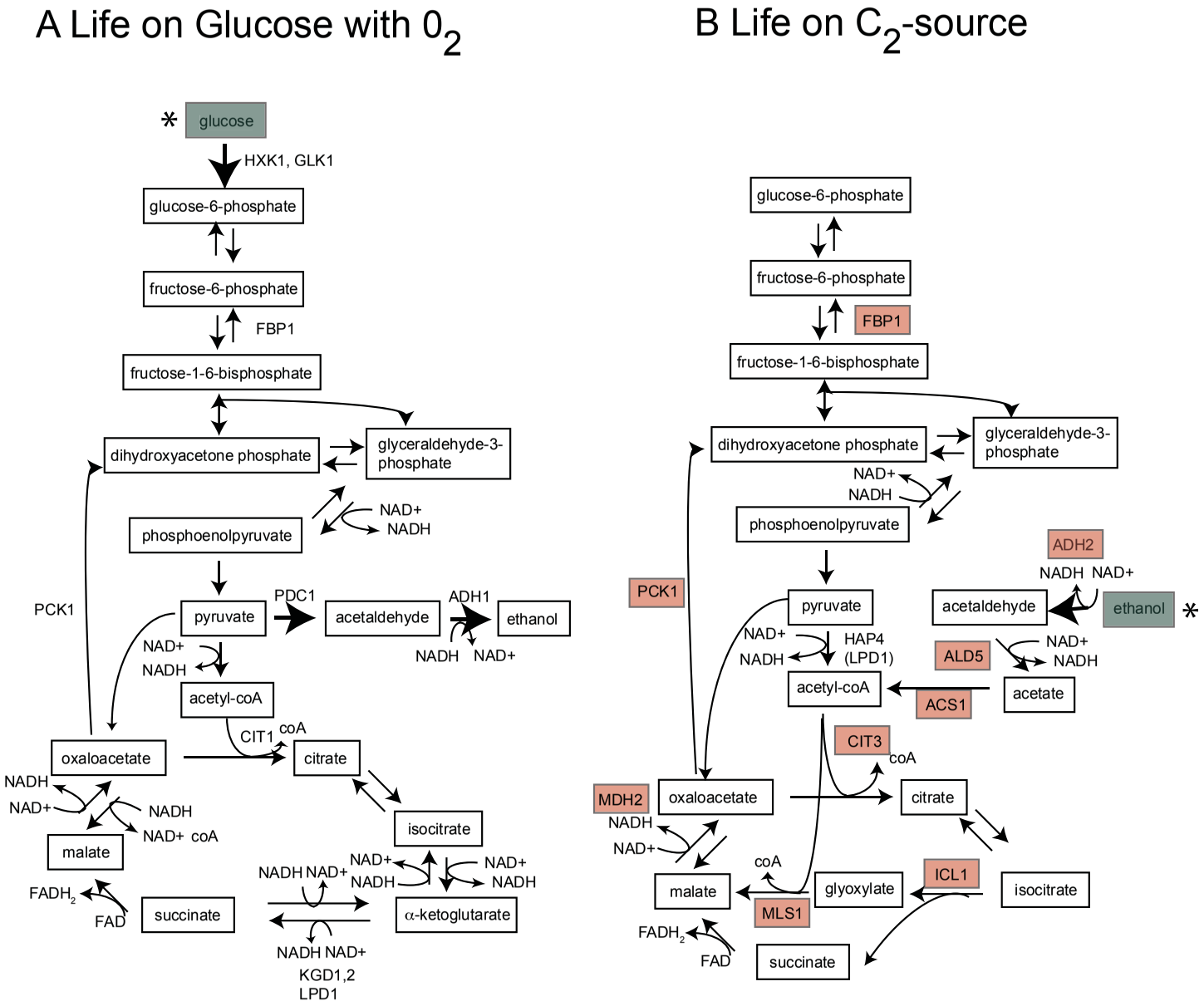


Figure 3-4 Glycolysis (A) and glyoxylate cycle/shunt (B). Genes that were up-regulated in the outer zone (OZ) of the mycelium, where hyphae grow with maximal speed, are indicated in red. This figure was modified from (Klein et al., 1998).

The glyoxylate cycle and gluconeogenesis is induced "too" early in *A. gossypii*

In *S. cerevisiae* the transcription of a number of genes is repressed in the presence of rapidly fermentable sugars, i.e. glucose, fructose and mannose. In the presence of glucose, the transcription of genes essential for catabolism of slowly fermentable sugars is repressed, as well as genes with gluconeogenetic, mitochondrial, glyoxylate cycle and respiratory functions. (Gancedo, 1992; Ronne, 1995). Since glucose is the most repressing sugar, the term "glucose catabolite repression" is often employed. Glucose repression in *S. cerevisiae* is mediated by ScMig1 and other repressors such as ScMig2. Control mechanisms other than glucose repression, for example increased mRNA degradation, decreased translational efficiency, or carbon catabolite inactivation are also operative (Klein *et al.*, 1998).
Due to the unidirectional glycolytic reactions of phosphofructokinase and pyruvate kinase, the functions of ScFBP1 (fructose-1,6-bisphosphatase) and ScPCK1 (phosphoenolpyruvate carboxykinase) are essential for gluconeogenesis. Both mRNAs are exceptionally sensitive towards glucose and are rapidly degraded (Mercado *et al.*, 1994; Yin *et al.*, 1996). In *A. gossypii* AgFBP1 and AgPCK1 mRNAs were found even in the presence of glucose in the medium (table 3-3). The mRNAs found at 9 h might as well be remnants from the sporulation process packaged into spores. But in advanced mycelium (18 h) and in the very fast growing hyphae (OZ) the measured mRNA levels are certainly due to transcription of AgFBP1 and AgPCK1. Based on measurements with a glucose meter (Accu-Chek, Roche), we found that at least 50-75% of the initial glucose was still present in the medium after 18 h. Furthermore, the total dry biomass after 18 h was measured, which was 0.29 g. The 0.6 g glucose, 0.3 g yeast extract and 0.3 g bacto peptone are most probably sufficient to produce this biomass without running into glucose limiting conditions.

In addition to the key gluconeogenetic genes, genes required for ethanol conversion to acetyl-CoA (AgADH2, AgALD5, AgACS2), were also expressed in the presence of glucose. Transcripts of the alcohol dehydrogenase AgADH2 were detected very early in the life cycle. High amounts of transcripts were found at 9 h and 18 h and very high amounts were found in the outer zone of fast growing hyphae. Surprisingly, just one of the two key genes (AgMLS1, AgICL1) of the glyoxylate cycle was already up-regulated in advanced mycelium (18 h), whereas both were highly expressed in fast growing hyphae (OZ). Even though five of the seven genes shown in figure 3-4 were expressed at 18h, it is difficult to predict whether the glyoxylate cycle is indeed actively used in this growth/nutrient stage.

None the less, it becomes clear that glucose repression seems to have an inferior role in the filamentous fungus *A. gossypii*. Genes involved in the glyoxylate cycle and gluconeogenesis are expressed very early in the life cycle. Hence, we speculate that regarding carbon source utilization, *A. gossypii* lives a mosaic-like life. It grows with 0_2 on glucose, but at the same time is also able to use ethanol via the glyoxylate cycle. It is important to note that the *A. gossypii* genome encodes a MIG1 homolog but not a MIG2 homolog which may partly explain the relaxed catabolite repression.

Several genes that function in the cell wall and early during sporulation are up-regulated in very fast growing hyphae

AgCDA1/2, AgCWP1b, AgDSE4, AgCWP1b, AgEXG1/SPR1 were highly up-regulated in fast growing hyphae and were assigned to the functional group "cell wall". Surprisingly, all of these genes show sporulation-related expression patterns. In addition, the early

sporulation genes ScRIM4 and AgSPS4 were found to be highly up-regulated in fast growing hyphae. Although no morphological changes are visible in the periphery of a mycelial colony, the hyphae seem to prepare for sporulation. The expression of sporulation-specific genes was investigated in detail in chapter 4.

3.4 Materials and Methods

Gene expression data

Gene expression data was obtained as previously described in the technical chapter 1 of this thesis.

T-profiler analysis

T-profiler is a tool that uses the t-test to score changes in the average activity of predefined groups of genes (Boorsma *et al.*, 2005). Currently, t-profiler supports gene expression data from *S. cerevisiae* and *C. albicans*. Even though *A. gossypii* is not (yet) supported, we were able to use t-profiler. This was done by uploading a text file with the name of the first *S. cerevisiae* homolog next to the *A. gossypii* log2 ratios. Log2 ratios of the 130 Ashbya-specific genes were excluded from the analysis. To prevent the identification of significantly up- or down-regulated gene groups that are not expressed, we excluded genes with intensity values < 60 in both the deletion mutant and wild type from the text file. Microarray data was uploaded on the web

TABLE 3-5: Changes in activity of predefined groups in very fast (200mμ/h) compared to fast (100μm/h) growing hyphae

	CATEGORY	ASPECT	t-VALUE	E-VALUE	MEAN	ORFs
UPREGULATED	carbohydrate metabolism	P	7.81	8.00E-12	0.53	135
	aldehyde metabolism	P	6.26	5.30E-07	1.94	8
	vacuole	C	6.21	7.40E-07	0.41	131
	cell differentiation	P	6.17	9.50E-07	0.61	56
	O-acyltransferase activity	F	5.03	6.80E-04	1.41	9
	cell wall	C	4.80	2.20E-03	0.63	32
	vitamin B2 and derivative metabolism	P	4.29	2.50E-02	1.33	8
DOWN-REGULATED	intracellular	C	-9.05	< 1.0E-15	-0.21	3394
	nucleus	C	-4.60	5.90E-03	-0.11	1387
	biosynthesis	P	-5.14	3.80E-04	-0.13	667
	alcohol catabolism	P	-4.97	9.30E-04	-0.91	20

TABLE 3-6: Changes in activity of predefined groups in very fast (200mμ/h) compared to slow (10μm/h) growing hyphae

	CATEGORY	ASPECT	t-VALUE	E-VALUE	MEAN	ORFs
UPREGULATED	carbohydrate metabolism	P	9.17	<1.0E-15	0.94	134
	peroxisomal matrix	C	6.87	8.90E-09	2.19	13
	cell differentiation	P	5.86	6.40E-06	0.93	55
	vacuole	C	5.30	1.60E-04	0.59	130
DOWNREGULATED	nucleus	C	-18.26	<1.0E-15	-0.61	1417
	ribosome biogenesis and assembly	P	-13.93	<1.0E-15	-0.99	197
	protein biosynthesis	P	-13.26	<1.0E-15	-0.61	357
	mitochondrion	C	-9.70	<1.0E-15	-0.26	554
	nucleobase, nucleoside, nucleotide and nucleic acid metabolism	P	-7.48	1.00E-10	-0.07	1045
	nucleolus	C	-6.98	4.10E-09	-0.37	173
	nucleotidyltransferase activity	F	-5.06	5.80E-04	-0.49	64
	intracellular	C	-4.67	4.20E-03	0.10	3421
	S-adenosylmethionine-dependent methyltransferase activity	F	-4.54	7.80E-03	-0.55	55
	chaperone activity	F	-4.43	1.30E-02	-0.59	46

at http://www.t-profiler.org. The t-profiler output tables give the categories which are up- or down-regulated with a p-value < 0.05; the aspect (P = biological process, F = molecular function, C = cellular component) and the t-value. Gene groups with Bonferroni corrected E-values < 0.05 are though to be significantly regulated. “Mean” is the mean log2 expression ratio and “ORFs” are the number of ORFs in each category.

Chapter 4: The *A. gossypii* response to decreasing nutrient supply: sporulation

Chapter 4: The *A. gossypii* response to decreasing nutrient supply: sporulation

4.1 Introduction

A. gossypii is capable of producing needle-shaped haploid spores in specialized compartments. Sporulation has been observed in liquid culture with limiting nutrients and on plates after several days of growth. It is assumed that spore formation is induced by nutrient deprivation. On solid medium an *A. gossypii* colony extends radially with very fast growing hyphae at the periphery. Hyphae in the center of the discoidal colony are two to three days older and undergo sporulation in this area of decreased nutrient supply. To unveil the gene expression changes which underlie the process of spore formation/nutrient limitation in *A. gossypii*; mRNA was isolated from sporulating mycelium in the inner zone (IZ) of the colony as described in chapter 1. The transcriptional profile obtained with this mRNA was compared to the transcriptional profile of very fast growing hyphae at the periphery of the same colony, hereafter called outer zone (OZ).

4.2 Results

The transcription data was evaluated in two approaches. The first approach was an unbiased approach that aimed at identifying the thirty most differentially expressed genes in the inner zone of the fungal colony compared to the transcription profile of fast growing hyphae in the outer zone. This unbiased approach included a t-profiler analysis in which the activity of predefined groups of genes was scored (table 4-2). The second data evaluation approach was based on a compilation of meiosis- and sporulation-specific *S. cerevisiae* genes and aimed at identifying the transcription patterns of the *A. gossypii* orthologs.

The fifteen top up-regulated genes (table 4-1 and figure 4-1)

Sporulation-related genes.
The *A. gossypii* genome encodes a CWP1 related cell wall protein gene family with distinct expression patterns (see chapter 5). AgCWP1d (ABR028C) is one of four syntenic homolog of ScCWP1 (the gene family contains additional non-syntenic homologs). The expression of **AgCWP1d** is clearly related to the life cycle stage. It is not (or only marginally) expressed in very fast growing hyphae but 27-fold up-regulated in the sporulation zone (see also chapter 5, figure 5-1). The sporulation-specific function of AgCWP1d remains to be determined.

AgHSP26 is one of five homologous genes encoding a small heat shock protein, most likely with chaperone activity. Interestingly there is only one homolog in *S. cerevisiae*. AgHSP26 mRNA was the fourth most abundant mRNA in spores (see chapter 1).

The *S. cerevisiae* homolog of **AgFMP45**, ScFMP45, encodes an integral membrane protein found in mitochondria and eisosomes, cortex located mainly immobile patches (Walther et al., 2006). ScFMP45 is assumed to be associated with endocytosis, sporulation and the distribution/concentration of sphingolipids in membranes.

The *S. cerevisiae* homologs of **AgDIT1**, AgDIT2 and AgDTR1 are required for spore cell wall maturation. ScDIT1 and ScDIT2 are enzymes involved in the production of a soluble LL-dityrosine containing precursor of the spore wall. The transcripts of ScDIT1 are known to accumulate at the

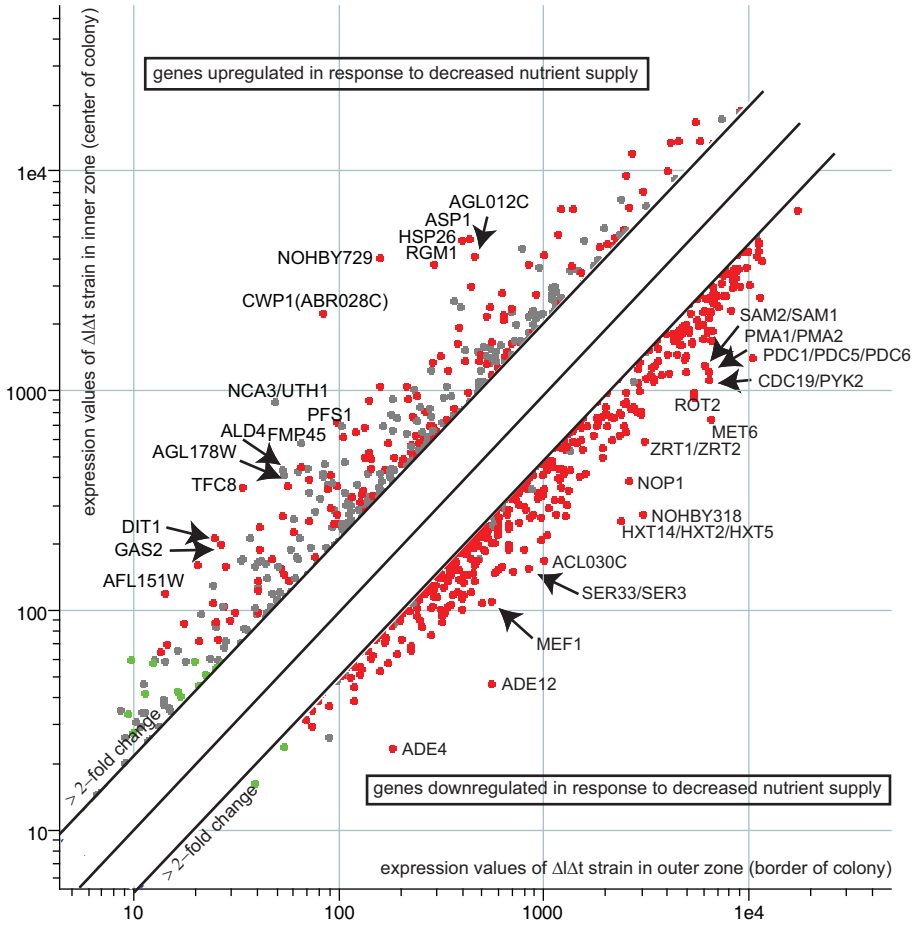


Figure 4-1 Scatter plot showing the top 15 up- and down-regulated genes

time of prospore enclosure. Dityrosine is a component of the yeast ascospore wall that is essential for the resistance of the spores to adverse environmental conditions (Briza *et al.*, 1994). The *A. gossypii* spore was also shown to contain dityrosin (Prillinger *et al.*, 1997). ScDTR1 is a transporter of the major facilitator superfamily that facilitates the translocation of bisformyl dityrosine through the prospore membrane.

The *S. cerevisiae* homolog of **AgGAS2**, codes for a 1,3-beta-glucanosyltransferase which is involved together with Gas4p in spore wall assembly. ScGAS2 has similarity to Gas1p. AgGAS4 was also up-regulated with a log2 ratio of 2.36.

ScPFS1, the homolog of **AgPFS1**, is a sporulation enzyme required for prospore membrane formation at spindle poles. It ensures the functionality of all four spindle pole bodies during meiosis II in a sporulating yeast cell, but it is not required for meiotic recombination or meiotic chromosome segregation.

As expected, the t-profiler analysis (table 4-2) revealed that sporulation genes were significantly up-regulated (t = 6.34, E = 3.2E-07).

Genes with no homologs in *S. cerevisiae* (NOHBYs). Surprisingly, two NOHBYs were among the fifteen top up-regulated genes in response to sporulation. NOHBY449 (ADL400W) has no homologs in other fungal species. Neither has NOHBY611 which potentially encodes a 181 aa long protein of unknown function. NOHBY 611 was not expressed in fast growing hyphae and weakly expressed in sporulating mycelium.

Genes involved in impaired cell growth and cell death. AgUTH1/NCA3 (a single gene in *A. gossypii* but twin genes in *S. cerevisiae*) potentially encodes a cell wall protein of the SUN family. The AgUTH1/NCA3 transcript was also found at very high levels in spores (see chapter 2).

ScRGM1, the homolog of **AgRGM1**, codes for a putative transcriptional repressor with proline-rich zinc fingers. Overproduction of ScRGM1 impairs cell growth.

Other genes. AgASP1 potentially encodes a cytosolic L-asparaginase which is involved in asparagine catabolism.

AgALD4 potentially codes for a mitochondrial aldehyde dehydrogenase that is required for growth on ethanol and conversion of acetaldehyde to acetate. The activity of the yeast homolog ScALD4 is K+ dependent and utilizes either NADP+ or NAD+ as coenzymes. Its expression is glucose repressed.

ScTFC8, the homolog of **AgTFC8**, encodes one of six subunits of RNA polymerase III transcription initiation factor complex (TFIIIC).

Genes involved in meiosis. None of the yeast homologs of the fifteen top up-regulated genes was potentially involved in meiosis, except the homolog for AgPFS1. However in the t-profiler analysis the category “meiosis” was significantly up-regulated (t = 4.4, E = 3.3E-04; see table 4-2).

The fifteen top down-regulated genes (table 4-1 and figure 4-1)

Amino acid biosynthesis.
AgMET6 potentially codes for a methionine synthase which is involved in amino acid biosynthesis. The *S. cerevisiae* homolog of **AgADE4** encodes a phosphoribosylpyrophosphate amidotransferase which catalyzes first step of the ‘de novo’ purine nucleotide biosynthetic pathway.
AgSER33/SER3 is a single gene in *A. gossypii* and a twin gene in *S. cerevisiae*. ScSER33 codes for a 3-phosphoglycerate dehydrogenase which catalyzes the first step in serine and glycine biosynthesis. ScSer33p is an isozyme of Ser3p.

Other genes. The *S. cerevisiae* homolog of **AgADE12** codes for an adenylosuccinate synthase which catalyzes the first step in the synthesis of adenosine monophosphate during purine nucleotide biosynthesis. Interestingly, AgHXT14/HXT5, a gene which potentially codes for a hexose transporter was strongly down-regulated. ScHXT5 is induced by a decrease in growth rate. ScHXT14 codes for a protein of unknown function with similarity to hexose transporter family members. Its expression is induced by low levels and is repressed by high levels of glucose.

AgPDC1/PDC5 (one gene in *A. gossypii*, twin genes in yeast) potentially codes for the most abundant of three pyruvate decarboxylase isozymes. ScPdc1p is the key enzyme in alcoholic fermentation and decarboxylate pyruvate to acetaldehyde. ScPDC5 codes for a minor isoform of the pyruvate decarboxylase.

AgNOP1 potentially codes for a nucleolar protein of the small subunit processome complex which is required for processing of pre-18S rRNA.

The *S. cerevisiae* homolog of **AgROT2** codes for the glucosidase II catalytic subunit required for normal cell wall synthesis.

AgPMA1/PMA2 (one gene in *A. gossypii*, twin genes in yeast) potentially codes for a plasma membrane H+-ATPase that pumps protons out of the cell. The homolog ScPMA1 is the major regulator of cytoplasmic pH and plasma membrane potential. ScPma2 is an

TABLE 4-1: TOP DIFFERENTIALLY EXPRESSED GENES IN RESPONSE TO DECREASING NUTRIENT SUPPLY/SPORULATION

UPREGULATED

SYSTEMATIC NAME	COMMON NAME	EXP outer zone	EXP inner zone	LOG2 (RATIO)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGIGAL PROCESS	CELLULAR COMPONENT
ABR028C	CWP1d	82.65	2259.00	4.77	Cell wall mannoprotein, linked to a beta-1,3- and beta-1,6-glucan heteropolymer through a phosphodiester bond; involved in cell wall organization	cell wall organization and biogenesis	structural constituent of cell wall	chitin- and beta-glucan-containing cell wall*
ADL400W	NOHBY449	155.65	4055.00	4.70	no description	-	-	-
ADR322W	UTH1/NCA3	48.30	894.55	4.21	Mitochondrial outer membrane and cell wall localized SUN family member required for mitochondrial autophagy; involved in the oxidative stress response, life span during starvation, mitochondrial biogenesis, and cell death	mitochondrion organization and biogenesis	-	chitin- and beta-glucan-containing cell wall
AFL136W	RGM1	287.65	3796.00	3.72	Putative transcriptional repressor with proline-rich zinc fingers; overproduction impairs cell growth	negative regulation of transcription from RNA polymerase II promoter	RNA polymerase II transcription factor activity	nucleus
AAL128C	HSP26	394.45	4884.00	3.63	Small heat shock protein with chaperone activity that is regulated by a heat induced transition from an inactive oligomeric (24-mer) complex to an active dimer; induced by heat, upon entry into stationary phase, and during sporulation	response to stress	unfolded protein binding	cytoplasm
AGR357W	ASP1	427.25	4914.00	3.52	Cytosolic L-asparaginase, involved in asparagine catabolism	asparagine catabolic process	asparaginase activity	cytosol
AGL009C	TFC8	33.45	364.30	3.45	One of six subunits of RNA polymerase III transcription initiation factor complex (TFIIIC); part of TFIIIC TauB domain that binds BoxB promoter sites of tRNA and other genes; linker between TauB and TauA domains; human homolog is TFIIIC-90	transcription initiation from RNA polymerase III promoter	RNA polymerase III transcription factor activity	transcription factor TFIIIC complex
AGL012C		454.15	4099.50	3.17	no description	-	-	cytoplasm
ABL156C	FMP45	64.86	582.35	3.17	Integral membrane protein localized to mitochondria (untagged protein) and eisosomes, immobile patches at the cortex associated with endocytosis; sporulation and sphingolipid content are altered in mutants; has homologs SUR7 and YNL194C	cell wall organization and biogenesis	molecular_function	mitochondrion
AFR401W	DIT1	24.52	214.85	3.13	Sporulation-specific enzyme required for spore wall maturation, involved in the production of a soluble LL-dityrosine-containing precursor of the spore wall; transcripts accumulate at the time of prospore enclosure	spore wall assembly (sensu Fungi)	catalytic activity	-
AFL151W	NOHBY611	13.99	120.10	3.10	no description	-	-	-
ADR417W	ALD4	51.93	438.80	3.08	Mitochondrial aldehyde dehydrogenase, required for growth on ethanol and conversion of acetaldehyde to acetate; activity is K+ dependent; utilizes NADP+ or NAD+ equally as coenzymes; expression is glucose repressed	ethanol metabolic process	aldehyde dehydrogenase (NAD) activity	mitochondrion
AGL178W		52.99	418.90	2.98	TyB Gag-Pol protein; proteolytically processed to make the Gag, RT, PR, and IN proteins that are required for retrotransposition	-	-	-
ACL182C	GAS2	26.19	199.40	2.93	1,3-beta-glucanosyltransferase, involved with Gas4p in spore wall assembly; has similarity to Gas1p	spore wall assembly (sensu Fungi)	1,3-beta-glucanosyltransferase	cytoplasm
AFR132C	PFS1	96.32	720.35	2.90	Sporulation protein required for prospore membrane formation at selected spindle poles, ensures functionality of all four spindle pole bodies of a cell during meiosis II; not required for meiotic recombination or meiotic chromosome segregation	-	-	-

DOWNREGULATED

SYSTEMATIC NAME	COMMON NAME	EXP outer zone	EXP inner zone	LOG2 (RATIO)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGIGAL PROCESS	CELLULAR COMPONENT
ABL186W	ADE12	549.30	46.13	-4.88	Adenylosuccinate synthase, catalyzes the first step in synthesis of adenosine monophosphate from inosine 5'monophosphate during purine nucleotide biosynthesis; exhibits binding to single-stranded autonomously replicating (ARS) core sequence	telomere maintenance	DNA replication origin binding	cytoplasm
ACR027C		2986.00	273.05	-4.18	no description	-	-	-
AFR602W	HXT14/HXT5/ HXT2	2359.50	256.15	-3.54	Protein with similarity to hexose transporter family members, expression is induced in low glucose and repressed in high glucose; the authentic, non-tagged protein is detected in highly purified mitochondria in high-throughput studies	hexose transport	galactose transmembrane transporter activity	mitochondrion
ABR212C	MET6	6431.00	741.50	-3.14	Cobalamin-independent methionine synthase, involved in amino acid biosynthesis; requires a minimum of two glutamates on the methyltetrahydrofolate substrate, similar to bacterial metE homologs	methionine biosynthetic process	5-methyltetrahydropteroyl-triglutamate-homocysteine S-methyltransferase activity	cytoplasm
AGL334W	ADE4	179.10	23.53	-3.09	Phosphoribosylpyrophosphate amidotransferase (PRPPAT; amidophosphoribosyltransferase), catalyzes first step of the 'de novo' purine nucleotide biosynthetic pathway	de novo' IMP biosynthetic process	amidophosphoribosyltransferase activity	cytoplasm
ACL134C	PDC1/PDC5	10348.00	1415.50	-3.04	Pyruvate decarboxylase isozyme, key enzyme in alcoholic fermentation, decarboxylates pyruvate to acetaldehyde; subject to glucose-, ethanol-, and autoregulation; involved in amino acid catabolism	pyruvate metabolic process	pyruvate decarboxylase activity	cytoplasm
AER214C	NOP1	2554.00	391.05	-3.04	Nucleolar protein, component of the small subunit processome complex, which is required for processing of pre-18S rRNA; has similarity to mammalian fibrillarin	ribosome biogenesis and assembly	methyltransferase activity	nucleolus
ACL030C		980.80	168.65	-3.00	no description	-	-	cytoplasm
AAR173C	ROT2	5358.50	940.45	-2.90	Glucosidase II catalytic subunit required for normal cell wall synthesis; mutations in rot2 suppress tor2 mutations, and are synthetically lethal with rot1 mutations	chitin- and beta-glucan-containing cell wall biogenesis	alpha-glucosidase activity	mitochondrion
AGL085C	PMA1/PMA2	6366.50	1135.50	-2.86	Plasma membrane H+-ATPase, pumps protons out of the cell; major regulator of cytoplasmic pH and plasma membrane potential; part of the P2 subgroup of cation-transporting ATPases	regulation of pH	hydrogen-exporting ATPase activity, phosphorylative	mitochondrion
ADR368W	CDC19/PYK2	5309.00	983.85	-2.82	Pyruvate kinase, functions as a homotetramer in glycolysis to convert phosphoenolpyruvate to pyruvate, the input for aerobic (TCA cycle) or anaerobic (glucose fermentation) respiration	glycolysis	pyruvate kinase activity	cytosol
ACL032C	SER33/SER3	831.30	155.05	-2.80	3-phosphoglycerate dehydrogenase, catalyzes the first step in serine and glycine biosynthesis; isozyme of Ser33p	serine family amino acid biosynthetic process	phosphoglycerate dehydrogenase activity	cytoplasm
AFR692C	SAM2/SAM1	6337.00	1221.50	-2.77	S-adenosylmethionine synthetase, catalyzes transfer of the adenosyl group of ATP to the sulfur atom of methionine; one of two differentially regulated isozymes (Sam1p and Sam2p)	methionine metabolic process	methionine adenosyltransferase activity	-
AGL143C	ZRT2/ZRT1	3037.50	592.20	-2.73	Low-affinity zinc transporter of the plasma membrane; transcription is induced under low-zinc conditions by the Zap1p transcription factor	low-affinity zinc ion transport	low-affinity zinc ion transmembrane transporter activity	plasma membrane
ABR227C	MEF1	547.75	109.30	-2.68	Mitochondrial elongation factor involved in translational elongation	translational elongation	translation elongation factor activity	mitochondrion

TABLE 4-2: Changes in activity of predefined groups in response to decreasing nutrient supply

	CATEGORY	ASPECT	t-VALUE	E-VALUE	MEAN	ORFs
UPREGULATED	cellular component unknown	C	7.18	9.70E-10	0.37	363
	sporulation (sensu Fungi)	P	6.34	3.20E-07	1.19	19
	meiosis	P	5.17	3.30E-04	0.50	84
	cell wall	C	4.40	1.50E-02	0.63	35
DOWNREGULATED	ribosome	C	-19.28	< 1.0E-15	-1.03	192
	mitochondrion	C	-10.94	< 1.0E-15	-0.27	548
	metabolism	P	-9.49	< 1.0E-15	-0.04	2284
	cytoplasm	C	-9.25	< 1.0E-15	0.01	2441
	intracellular	C	-9.19	< 1.0E-15	0.04	3391
	ribosome biogenesis	P	-8.42	< 1.0E-15	-0.41	161
	biosynthesis	P	-8.17	4.60E-13	-0.12	664
	purine nucleotide metabolism	P	-5.82	8.20E-06	-0.66	35
	RNA modification	P	-5.76	1.20E-05	-0.43	68
	endoplasmic reticulum	C	-5.60	3.00E-05	-0.15	291
	mitochondrial electron transport chain	C	-5.46	6.60E-05	-0.85	19
	DNA-directed RNA polymerase activity	F	-4.52	8.60E-03	-0.51	31

isoform of ScPma1.

AgCDC19/PYK2 is one gene in *A. gossypii* with twin genes in *S. cerevisiae*. ScCDC19 encodes a pyruvate kinase that functions in glycolysis to convert phosphoenolpyruvate to pyruvate, the input for aerobic (TCA cycle) or anaerobic (glucose fermentation) respiration. ScPYK2 codes for a pyruvate kinase is probably modulated by phosphorylation.

AgSAM1/SAM2 potentially codes for a S-adenosylmethionine synthetase that catalyzes the transfer of the adenosyl group of ATP to the sulfur atom of methionine. The twin homologs ScSAM1 and ScSAM2 of

AgZRT1/ZRT2 code for zinc transporters of the plasma membrane. **AgMEF1** potentially encodes a mitochondrial elongation factor involved in translational elongation.

Transcriptional profiling of syntenic homologs of meiosis/sporulation-specific yeast genes

Yeast homologs of many *A. gossypii* genes that were differentially expressed in the first analysis and genes in predefined t-profiler gene groups are not only expressed during meiosis/sporulation but also in vegetatively growing cells. In her PhD thesis, Sophie Brachat used the data of (Primig *et al.*, 2000) and compiled one list of sporulation-specific genes and a second list of meiosis-specific genes. Starting with these two lists which are shown in tables 4-3 and 4-4, the individual transcription profiles of the ortholog pairs were compared. The transcription profiles of the *A. gossypii* orthologs can be classified into five groups:

1. high in spores, low in growing mycelium, increased in spore forming mycelium
2. low in spores and growing mycelium and increasing in spore formation
3. high in spores, low in growing mycelium and still low in sporulating mycelium
4. highest during mycelial growth
5. not detectable

The expression profiles of the last two groups are inconsistent with potential roles in spore formation/meiosis.

4.3 Discussion and Conclusions

Expression of genes presumably involved in sporulation

Top up-regulated genes associated with cell wall formation and membrane functions

Not surprisingly, homologs of genes that are involved in sporulation of the budding yeast *S. cerevisiae* are also up-regulated during sporulation in *A. gossypii*. Five of the fifteen genes which were most up-regulated during sporulation were sporulation genes. Interestingly, these genes have mainly functions related to the (spore) cell wall and the (spore) membrane.

Expression profiles of genes that are consistent with a role in sporulation

The expression patterns of genes which have sporulation-specific homologs in *S. cerevisiae* are shown in table 4-3. Out of 23 expressed genes, 16 showed clearly sporulation-related gene expression with three distinct patterns. Eight genes followed the expression pattern 1 ("high in spores, low in growing mycelium, increased in spore forming mycelium"). Eight other genes followed the pattern 2 ("low in spores and growing mycelium and increasing in spore formation") and two genes followed the pattern 3 ("high in spores, low in growing mycelium and still low in sporulating mycelium").
Patterns 1 and 3 suggest the presence of sporulation-specific mRNAs in spores. The presence of these transcripts can be explained by a pool of mRNA that is packaged into the prospore before enclosure. Alternatively, although less likely, these transcripts might have been newly synthesized in hydrated spores and have functions in spore survival and maintenance of germination capacity. After inoculation of spores in full medium we observed a decrease of these mRNAs. This resulted in almost all cases in 18 h signals below the cutoff value of 60, chosen as background signal. The slow decline in the signal for AgSPO1 or AgADY4 can be explained by the following dilution model. After onset of germination, the genes are not transcribed and the mRNAs present in the spores are not actively degraded. Because 20% of the spores had not germinated after 9 h incubation time (see chapter 2 figure 2-2), we would expect to still see 1/5 of the initial signal at 9 h assuming a spore viability of 100% and the same RNA extraction efficiency for all developmental stages. Rapid declines as seen for AgDIT1 and AgDTR1 within the first two hours can only be explained by fast mRNA degradation. In *S. cerevisiae* germination has been previously described to lead to very rapid degradation of mRNAs. For instance, degradation with a half life time of 15 min has been observed for ScSPS100, a gene required for spore cell wall maturation (Brengues *et al.*, 2002).

Expression profiles of genes that are not consistent with a role in sporulation

Five out of 23 expressed genes, followed the expression pattern 4 ("highest during mycelial growth") and the transcripts of two genes belong to pattern 5 ("expression not detectable"). These profiles are inconsistent with a specific role in spore formation. However, genes belonging to pattern 4 might still have a function in spore formation in *A. gossypii*, but have acquired an additional function, for instance in mitosis. Both, changes in sporulation-specific gene composition and/or alterations in their regulation might attribute to differences in spore morphology. The ascospore of *S. cerevisiae* is smooth and spherical whereas the *A. gossypii* spore is needle shaped.

Table 4-3 *A. gossypii* genes with **sporulation-specific** homologs in *S. cerevisiae*

Systematic name	Common name	ΔIΔt 0h	ΔIΔt 2h	ΔIΔt 5h	ΔIΔt 7h	ΔIΔt 9h	ΔIΔt 18h	ΔIΔt OZ	ΔIΔt IZ	YPD description of *S. cerevisiae* homolog
pattern of gene expression: 										
AAL027W	SPO1	350	325	262	136	110	82	131	216	Transcriptional regulator involved in sporulation, has strong similarity to phospholipase B enzymes
AAR146W	SPR1	7852	5208	4604	1784	956	351	12201	11236	Sporulation-specific exo-beta-1,3-glucanase, involved in late sporulation
ACR227W	SHC1	178	199	120	51	39	18	90	289	Protein involved in cell wall chitin synthesis or deposition
AFR401W	DIT1	858	353	305	93	44	20	34	174	First enzyme in the pathway for biosynthesis of dityrosine in the outer layer of the spore wall; expressed late (10-16 hr) in sporulation
AGL038C	RIM4	118	54	35	23	22	19	1343	2313	Protein required for sporulation and formation of meiotic spindle, has two RNA recognition motif (RRM) domains
AGL290C	SPR6	288	483	335	184	125	46	110	284	Sporulation-specific protein
AGR174W	ADY4	92	112	87	52	54	48	56	117	Protein with possible role in sporulation
AGR195W	DTR1	265	81	62	41	42	41	202	894	dityrosine transporter, essential for spore wall synthesis, facilitates the translocation of bisformyl dityrosine through the prospore membrane
pattern of gene expression: 										
ABR127C	SPS4	21	12	11	10	10	7	212	964	Protein expressed in mid-late (8-14 hr) sporulation, possible cell wall component
ACR135C	SSP2	26	18	17	9	8	6	83	391	Sporulation-specific protein
ADL225W	SGA1	67	97	62	48	47	55	201	939	Glucoamylase (glucan-alpha-1,4-glucosidase), sporulation-specific
ADL315C	SMK1	50	45	51	45	46	50	241	863	Sporulation-specific protein kinase of the MAP family required for completion of sporulation
AER135W	SMA2	57	93	112	163	145	87	141	223	Protein with possible role in sporulation
AFR571W	SPR3	10	10	10	9	9	9	19	71	Sporulation-specific septin
AFR724C	SPS1	20	42	37	20	17	11	78	290	Serine/threonine protein kinase involved in middle/late stage of meiosis
AGR098W	ISC10	107	129	102	88	91	78	188	632	Meiosis-specific protein required for spore formation
pattern of gene expression: 										
ADL168C	MCK1	1371	995	829	841	872	760	738	828	Member of the GSK3 subfamily of protein kinases, positive regulator of meiosis and spore formation
AER068C	BDF1	1292	973	641	505	435	342	382	573	Protein required for sporulation, contains two bromodomain motifs and interacts with histones H3 and H4
expression profiles not consistent with the role in spore formation:										
AAL167W	SPO75	14	20	24	17	20	65	25	48	Protein with possible role in in spore formation
ABL193C	ADY3	23	73	76	89	81	56	57	71	Protein involved in accumulation of dyads, has a region of low similarity to a region of C. albicans Cht3p (chitinase 3), which is involved in cell wall biosynthesis
ADL140C	SPO12	157	201	546	628	614	215	308	251	Sporulation protein required for chromosome division in meiosis I
AEL334W	SPO7	104	113	117	117	110	138	131	122	Protein required for meiosis
AGL235C	CSM1	282	118	131	144	149	196	198	225	Protein possibly involved in sporulation
transcripts not detected:										
AFR400C	DIT2	7	5	6	6	6	5	6	6	Cytochrome P450 56, second enzyme in the pathway for biosynthesis of dityrosine in the outer layer of the spore wall
AGR105C	SPO71	14	20	16	14	12	9	18	29	Protein involved in spore wall formation
genes not represented on Affymetrix array or genes with no homologs (NH) in *A. gossypii*:										
ACL179C	SPO77	ND	ND	ND	ND	ND	ND	ND	ND	Protein with possible role in sporulation
ADL029W	SPO74	ND	ND	ND	ND	ND	ND	ND	ND	Protein with similarity to phosphoribulokinase precursor (phosphopentokinase), involved in sporulation
AGR274C	MPC54	ND	ND	ND	ND	ND	ND	ND	ND	Component of the meiotic plaque required for formation of the prospore membrane, coiled-coil protein induce early in sporulation
NH	SMA1	ND	ND	ND	ND	ND	ND	ND	ND	Protein with possible role in sporulation
NH	SPO16	ND	ND	ND	ND	ND	ND	ND	ND	Early meiotic protein required for efficient spore formation
NH	SPO20	ND	ND	ND	ND	ND	ND	ND	ND	Protein required for sporulation and growth of prospore membrane, interacts with Dbf2p
NH	SPO22	ND	ND	ND	ND	ND	ND	ND	ND	Protein required for meiosis and sporulation, has cecropin family signature
NH	SPS100	ND	ND	ND	ND	ND	ND	ND	ND	Sporulation specific protein involved in spore wall formation, first appears at 12h of sporulation

Table 4-3 Expression levels of *A. gossypii* genes with **sporulation-specific** homologs in *S. cerevisiae*. Normalized linear intensity levels of the laboratory strain (ΔIΔt) in dependance of time post inoculation or mycelial zone are shown. OZ = outer zone of fast hyphal growth, IZ = inner zone of sporulation, ND = No data available, NH = No homolog in *A. gossypii*. This gene list was modified from (S. Brachat, unpublished data) and (Primig et al, 2000).

Table 4-4 *A. gossypii* genes with **meiosis-specific** homologs in *S. cerevisiae*

Systematic name	Common name	ΔlΔt 0h	ΔlΔt 2h	ΔlΔt 5h	ΔlΔt 7h	ΔlΔt 9h	ΔlΔt 18h	ΔlΔt OZ	ΔlΔt IZ	YPD Description of homolog
pattern of gene expression: 										
MEK1	ADR379C	56	75	56	33	28	21	57	65	Serine/threonine protein kinase required for meiotic recombination
SCM4	AEL283C	1483	742	460	471	555	1057	1444	1740	Protein that suppresses temperature-sensitive allele of CDC4 when overexpressed
pattern of gene expression: 										
ACR218W	RIM15	275	257	296	274	242	224	352	544	Serine/threonine protein kinase, positive regulator of IME2 expression and of sporulation
ADL135C	AMA1 (SPO70)	102	129	125	117	117	84	167	295	Protein required for proper meiosis and sporulation, associates with and regulates the activity of anaphase promoting complex/cyclosome (APC/C), member of Cdc20 family of proteins
ADR253W	RIM11	799	470	468	562	584	633	1033	1415	Member of the GSK3 subfamily of protein kinases, required for induction of IME2 by Ime1p
AEL292W	REC8 (SPO69)	128	141	154	228	261	396	384	178	Protein required for meiosis and sporulation
AFR076W	IME2	23	29	28	25	28	29	170	227	Serine/threonine protein kinase and positive regulator of early, middle, and late meiotic genes
AFR474W	HOP2	9	9	8	9	9	9	77	148	Meiosis-specific gene required for the pairing of homologous chromosomes
AGL038C	RIM4	118	54	35	23	22	19	1343	2313	Protein required for sporulation and formation of meiotic spindle, has two RNA recognition motif (RRM) domains
AGR101C	DMC1	26	24	27	28	26	48	56	78	Meiosis-specific recombination protein structurally related to S. cerevisiae Rad51p, Rad55p, Rad57p, and to E. coli RecA
AGR260W	RED1	38	43	40	33	35	35	74	97	Meiosis-specific protein associated with lateral elements of the synaptonemal complex, involved in homologous chromosome synapsis and chiasmata formation
AGR347W	NDT80	21	20	30	22	20	15	28	54	Meiosis-specific protein required for exit from pachytene
pattern of gene expression: 										
ABR082W	HOF1	154	193	143	135	105	123	92	65	Protein involved in cytokinesis, has an SH3 domain
ADL048C	SAE2	89	100	88	61	55	32	32	39	Protein involved in meiotic recombination
ADL168C	MCK1	1371	995	829	841	872	760	738	828	Member of the GSK3 subfamily of protein kinases, positive regulator of meiosis and spore formation
ADR141W	RIM9	88	119	94	96	93	66	72	63	Sporulation protein that stimulates expression of IME1
expression profiles not consistent with the role in presumptive meiosis:										
ABL103W	MUM2	115	304	238	206	201	221	243	313	Protein required for premeiotic DNA synthesis and sporulation
ADL093W	IRR1	185	385	558	577	525	402	332	317	Component of cohesin complex, required for sister chromatid cohesion during DNA replication
AEL007W	MCD1	40	247	393	481	446	314	159	55	Cohesin, protein required for mitotic chromatid cohesion and chromosome condensation
AER439W	SKI8	172	216	246	279	321	257	223	207	Protein involved in meiotic recombination and in protection from double-stranded RNA (dsRNA) viruses, member of the WD (WD-40) repeat family
AFR190C	RIM101	363	318	366	367	315	308	469	337	Transcription factor involved in induction of IME1, IME2, DIT1, and DIT2 transcription, has three C2H2-type zinc fingers
AGL023W	SMC1	27	69	87	93	83	57	58	68	Coiled-coil protein of the SMC family involved in chromosome condensation and segregation
AGL076W	PDS5	44	204	242	298	258	164	151	105	(putative) involved in sister chromosome cohesion during mitosis
AGL099C	UME6	109	82	111	140	150	124	116	119	Global transcriptional regulator containing a zinc binuclear cluster domain involved in pathway specific repression or induction
AGR229W	SLK19	272	285	275	237	229	132	183	261	Protein involved with Kar3p in control of spindle dynamics
AGR236W	SMC2	31	51	60	73	62	46	45	77	Subunit of condensin protein complex required for chromosome condensation and segregation, coiled-coil protein of the SMC family
transcripts not detected:										
AAL085C	ZIP1	7	7	6	6	6	5	8	9	Structural protein of the synaptonemal complex central element
AAL182W	SMC3	21	48	58	50	46	31	30	33	Cohesin, coiled-coil protein of the SMC family involved in chromosome condensation and segregation
ADR293C-A	SAE3	45	35	25	22	22	22	26	37	Protein involved in meiotic recombination
AGL133W	SCC2	17	41	49	57	55	45	42	51	Cohesin, protein required for mitotic sister chromatid cohesion
AGL304W	MAM1	25	27	20	17	15	14	21	39	Monopolin, protein involved in chromosome attachment to the meiotic spindle, induced in sporulation
AGR017W	REC102	7	8	7	8	8	8	10	15	Meiotic recombination protein, required for early meiotic events and for formation of the synaptonemal complex but not required for mitosis
genes not represented on Affymetrix array or genes with no homologs (NH) in *A. gossypii*:										
AFR407C	SPO13	ND	ND	ND	ND	ND	ND	ND	ND	Meiosis-specific protein required for sister centromere cohesion during meiosis I chromosome division
AGL308W	IME1	ND	ND	ND	ND	ND	ND	ND	ND	Transcription factor required for sporulation, positive regulator of IME2 and many sporulation genes
NH	RME1	ND	ND	ND	ND	ND	ND	ND	ND	Zinc-finger transcription factor that represses meiosis in non-a/alpha cells
NH	REC104	ND	ND	ND	ND	ND	ND	ND	ND	Meiosis-specific protein, mutants are defective in synaptonemal complex formation and have reduced recombination
NH	ZIP2	ND	ND	ND	ND	ND	ND	ND	ND	Protein involved in meiotic recombination, chromosome synapsis, and synaptonemal complex formation

Table 4-4 Expression levels of *A. gossypii* genes with **meiosis-specific** homologs in *S. cerevisiae*. Normalized linear intensity levels of the laboratory strain (ΔlΔt) in dependance of time post inoculation or mycelial zone are shown. OZ = outer zone of fast hyphal growth, IZ = inner zone of sporulation, ND = No data available, NH = No homolog in A. gossypii. This gene list was modified from (S. Brachat, unpublished data) and (Primig et al., 2000).

Expression of potential meiosis genes

Is sporulation in *A. gossypii* a sexual process?

The *A. gossypii* life cycle may include a sexual phase with mating and meiosis. This assumption is justified because its genome contains homologs of nearly all genes from *S. cerevisiae* known to be involved in the sexual cycle. This includes genes involved in mating (see chapter 6 and 7) as well as in meiotic transcriptional control, meiostic recombination, synaptonemal complex formation, sister chromatid cohesion, control of meiotic chromosome segregation and ascospore formation. (S. Brachat, PhD thesis, 2003; Dietrich *et al*., 2004). The needle shaped *A. gossypii* spores carrying one haploid nucleus are visible as packages of four spores in sporulating mycelia. Haploid ascospores are formed even though diploid nuclei have not been identified in *A. gossypii* hyphae. It is conceivable that *A. gossypii* can produce meiosporangia. However, there is no experimental evidence that sporulation seen in the center of mycelial colonies is a sexual process. If the sporulation process in *A. gossypii* involved meiotic divisions, we would expect to find the expression of genes which are homologous to *S. cerevisiae* meiosis genes derepressed during sporulation.

Homologs to meiosis-specific genes are up-regulated during sporulation in *A. gossypii*

Since there were no genes with exclusively meiotic functions among the fifteen most up-regulated genes during sporulation (table 4-1), t-profiler was used to score the activity of the meiosis genes as a group. The category “meiosis” was significantly up-regulated (t = 4.4, E = 3.3E-04) during sporulation. It is known that several genes share meiotic and mitotic functions in *S. cerevisiae*. Therefore, we focused on homologs of yeast genes which are meiosis-specific. The *A. gossypii* homologs were grouped according to their expression patterns in table 4-4. Out of 26 expressed genes, 16 showed a clear sporulation-dependent expression pattern. This is very similar to the 16 out of 23 presumable sporulation genes that followed a sporulation-specific pattern.

The hypothesis that sporulation is a sexual process is, at least to some extent, supported by the gene expression data. However the analysis shows at the same time that the expression pattern of some genes is inconsistent with a specific role in meiosis.

Chapter 5: The *A. gossypii* cell wall genes: Transcription levels and identification of cell wall proteins

Chapter 5: The *A. gossypii* cell wall genes: transcription levels and identification of cell wall proteins

5.1 Introduction

In chapter 2, genes coding for cell wall proteins were found to be among the 20 most highly expressed genes in spores. This observation raises the questions a) how the transcription profiles of cell wall genes change across the life cycle and b) whether the high cell wall gene expression is reflected in high cell wall protein levels. The latter question is important because proteins and not mRNAs drive functional processes in cells. If proteins were indeed abundant in the *A. gossypii* cell wall, it should be possible to detect them by mass spectrometry. Therefore a collaboration was started with Piet de Groot and Frans M. Klis (University of Amsterdam) that aimed at identifying the *A. gossypii* cell wall proteome and chitin content.
The cell wall proteome analysis was carried out with advanced mycelia grown for 18 hours in liquid full medium after spore inoculation. Advanced mycelia were used first, because of the analogy of this stage to a logarithmic growing *S. cerevisiae* culture; and second, because of the requirement for high amounts of biological material (5 to 10 g wet mass) which can hardly be provided in the form of spores or germlings.

5.2 Results

The mass spectrometry analysis of such mycelia resulted in the identification of fifteen *A. gossypii* cell wall proteins out of the 59 *in silico* predicted cell wall proteins (Piet de Groot, personal communication) which are listed in table 5-1. Cell wall proteins mediate important functions of the cell wall, for instance adhesiveness and protection against desiccation (Klis et al., 2007). They are processed through the secretory pathway where they are O- and also N- glycosylated. The majority receive a GPI-anchor at the ω amino acid close to the carboxy terminus and are linked through highly branched beta-1,6-glucan chains to beta-1,3-glucan. In addition, some proteins, the so-called proteins with internal repeats (PIR proteins), are highly linked to the beta-1,3-glucan network probably through an ester linkage (Klis *et al.*, 2007). The first part of this results section looks into the expression of selected cell wall protein genes and into the expression of chitin synthases. Chitin is in addition to the cell wall proteins an important structural component of the cell walls of many fungi. The second part of this results section will discuss two highly expressed genes with previously unknown functions that are, thanks to the mass spectrometry analysis, now also known to code for cell wall proteins (Piet de Groot, personal communication). The third part will include a section about the correlation of the *A. gossypii* cell wall transcriptome and the cell wall proteome.

Expression of selected cell wall protein genes

Expression patterns of genes that were confirmed by mass spectometry to code for cell wall proteins

Fourteen of the fifteen *A. gossypii* genes that were identified to code for cell wall proteins by mass spectrometry are represented on the *A. gossypii* sySYNG001a array. The expression patterns of these genes across the life cycle are shown in figure 5-1.
AgCDA1/2 (a single gene with twin homologs in *S. cerevisiae*) and **NOHBY702** (AGL034C) displayed a clear sporulation-related expression pattern. ScCDA1 codes

	Affymetrix ID	Systematic name	Common name(s)		0h ΔlΔt	2h ΔlΔt	5h ΔlΔt	7h ΔlΔt	9h ΔlΔt	18h ΔlΔt	ΔlΔt OZ	ΔlΔt IZ
Top 23	Ag003796_at	AGR138W	SED1	SPI1	19809	26637	32596	25822	22464	31153	37778	43580
	Ag004168_at	AGR049W	CCW12		27268	27274	29856	26082	25150	30924	30420	37713
	Ag002925_at	AFR065W	NOHBY621		12087	8885	17602	17192	17702	21709	25029	18353
	Ag000558_at	ABL063C	ECM33	PST1	7928	5289	9053	9020	9128	12482	11424	10695
	Ag003772_at	AGL354C	SCW4	SCW10	7354	4623	6508	5605	5410	7908	8702	9582
	Ag003735_at	AGL343C	BGL2		6936	5186	6181	4980	5073	7858	8647	12368
	Ag002964_at	AFR711C	FIG2		686	1166	8074	10496	9126	7723	11231	2688
	Ag003990_at	AGL352W	GAS1b		1982	3521	5781	5404	5393	7514	5768	4016
	Ag002641_at	AER372C	SUN4	SIM1	2644	4445	8300	8320	8044	7474	9770	7224
	Ag004477_at	ABR025C	CWP1a		5867	4021	5031	4366	4762	6569	6425	11250
	Ag002764_at	AEL111C	PIR3		4072	5004	4655	3384	3149	5976	5235	8956
	Ag003905_at	AGL321W	SOD1		6520	3776	4605	4185	3870	5786	4322	9223
	Ag002632_at	AER058W	CCW14		4937	3724	4913	3111	2869	4541	5842	8621
	Ag002011_at	ADL398C	CWP21		1586	455	817	1205	1716	3043	3453	5082
	Ag003734_at	AGL034C	NOHBY702		549	701	2854	3336	3047	2961	9766	18456
	Ag000455_at	ABR026C	CWP1b		9345	3606	2239	720	539	2218	23836	32721
	Ag000078_at	AAR076C			166	206	383	736	870	1756	868	275
	Ag001862_at	ADL021W	GAS5		260	472	1382	1590	1429	1575	1529	919
	Ag004532_at	ADL190C	KRE1		7797	5326	4828	1731	1287	1364	1362	2281
	Ag001083_r_at	ACR079W	PMP3		3030	1130	768	432	525	1346	2421	5224
	Ag001841_at	ADL036C	CDA1	CDA2	2189	1241	1054	504	433	1270	21798	30170
	Ag001266_at	ADR056W	TOS1		1138	1388	2312	1984	1591	1154	1589	1777
	Ag003519_at	AFL175C	ROT1		791	738	818	833	854	1128	793	708
	Ag000796_at	ABR060W	DCW1		800	828	1106	818	851	759	673	803
	Ag003517_at	AFR530W	DFG5		1147	1268	1541	918	785	746	864	2261
	Ag001080_at	ACR150W	YPS7		1816	1575	1678	802	816	586	597	1010
	Ag003118_at	AFL092C	FLO5a		353	3216	2429	1215	846	322	127	136
	Ag002051_at	AEL110W	PIR1	HSP150	154	158	149	133	127	297	766	4466
	Ag004379_at	AGR318C	CRH1		898	681	1385	1394	964	240	129	352
	Ag002149_at	AEL220C	EXG2		167	159	179	159	151	187	308	450
	Ag001104_at	ACR144W	PEP4b		234	136	76	73	70	173	535	1560
	Ag001733_at	ADR238C	NOHBY432		1267	495	183	107	92	115	1460	407
	Ag002561_at	AEL130C			64	66	103	83	76	106	138	260
	Ag004464_at	AAR101W	NOHBY112		164	238	451	213	134	95	134	288
	Ag000977_at	ACL200W	DFG50a		690	290	126	62	59	91	181	248
	Ag003777_at	AGL208C	DSE4		63	29	23	14	14	75	471	1507
	Ag002106_at	AER347W	GAS3		64	44	86	128	85	74	224	772
	Ag000439_at	ABR028C	CWP1d		18698	10399	5973	1491	531	73	83	2259
	Ag003880_at	AGL351W	GAS1a		50	45	30	20	19	72	70	446
	Ag001674_at	ADR322W	UTH1	NCA3	10381	8232	7569	2184	1094	70	48	895
	Ag000844_at	ACL182C	GAS2		9	11	7	7	8	48	26	199
	Ag001385_at	ADL175W	GAS4		11	11	9	10	9	42	17	88
	Ag002115_at	AEL312C			283	977	1159	331	193	34	24	109
	Ag003235_at	AFL095W	FLO5b		64	54	47	36	27	33	28	39
	Ag001545_at	ADR078C	CRR1		23	23	28	22	19	26	39	127
	Ag004230_at	AGR327C	SAG1		14	18	22	19	17	21	14	24
	Ag002805_at	AFR723C	SPS22	SPS2	10	8	11	9	10	18	180	927
	Ag000379_at	ABR064W	NOHBY214		32	33	16	9	8	7	14	59
	not on chip	ABL123C	PEP4a		ND	ND	ND	ND	ND	ND	ND	ND
	not on chip	ABR027C	CWPc		ND	ND	ND	ND	ND	ND	ND	ND
	not on chip	ACL201W	DFG50b		ND	ND	ND	ND	ND	ND	ND	ND
	not on chip	ACL202W	DFG50c		ND	ND	ND	ND	ND	ND	ND	ND
	not on chip	ACR272C	CWP20a		ND	ND	ND	ND	ND	ND	ND	ND
	not on chip	ACR273W	CWP20b		ND	ND	ND	ND	ND	ND	ND	ND
	not on chip	ADR019C	TIR1	TIR3	ND	ND	ND	ND	ND	ND	ND	ND
	not on chip	ADR246C	NOHBY434		ND	ND	ND	ND	ND	ND	ND	ND
	not on chip	AFR215W			ND	ND	ND	ND	ND	ND	ND	ND
	not on chip	AFR473W	NOHBY644		ND	ND	ND	ND	ND	ND	ND	ND
	not on chip	AEL103W	CIS3	PIR1	ND	ND	ND	ND	ND	ND	ND	ND

legend:

identified MS
GPI-anchored
other CWPs
newly identified cell wall proteins

Table 5-1 The *A. gossypii* genome encodes 59 predicted cell wall protein genes, many of which are highly expressed.
The list was sorted in decreasing expression level order according to the 18 h time point, the reference time point for mass spectrometry. Affymetrix IDs of genes that are predicted to encode GPI-anchored proteins are in blue, IDs of other predicted cell wall protein genes are in green and genes with cell wall proteins identified by mass spectrometry are in red. Mass spectrometry is known to be biased against identifying abundant proteins. 13 of the top 23 highly expressed genes encode proteins that were found in the *A. gossypii* cell wall by mass spectrometry indicating that there is a good correlation between mRNA levels and protein abundance in the case of highly expressed cell wall proteins. ND = No data available.

for a chitin deacetylase, which is together with Cda2p involved in the biosynthesis of the ascospore wall component chitosan. ScCDA1 and ScCDA2 are required for proper rigidity of the ascospore wall. A similar function of AgCDA1/2 would explain the observed pattern of gene expression. A possible function of NOHBY702 will be discussed later in this chapter.

AgCCW12 was expressed during vegetative growth, but also showed a clear up-regulation in sporulating mycelium. ScCCW12 encodes a GPI-anchored mannoprotein. Compared to wild type, the mutant ccw12Δ grows more slowly, is highly sensitive to Calcofluor white and contains 2.5 times more cell wall chitin (Ragni et al., 2007). AgCCW12 may play a role in agglutination and mating.

AgPIR3, **AgCWP1a**, **AgSCW4/10** and **AgSED1/SPI1** were slightly upregulated during sporulation. The PIR proteins are a family of covalently-linked cell wall proteins that contain internal repeats. ScPIR3 codes for a O-glycosylated cell wall protein that is required for cell wall stability. ScCWP1 is a cell wall mannoprotein involved in cell wall organization. AgSCW4/10 is a single gene in *A. gossypii* with twin homologs in *S. cerevisiae*. ScSCW10 and ScSCW4 are genes which code for proteins with similarity to glucanases. AgSED1/SPI1 codes for a 505 aa cell wall protein with predictable sites for glycosylation and GPI-anchor. AgSED1/SPI1 was the strongest expressed gene in advanced mycelium. *S. cerevisiae* carries twin homologs ScSED1 and ScSPI1. ScSED1 encodes a major stress-induced structural GPI-cell wall glycoprotein in stationary-phase cells which associates with translating ribosomes. ScSPI1 encodes a GPI-anchored, serine/threonine rich cell wall protein of unknown function.

The *A. gossypii* genome encodes two non-syntenic homologs of the ScFLO5. The expression of **AgFLO5a** (AFL092C) peaks very early in the life cycle. ScFLO5 codes for a lectin-like protein involved in flocculation (aggregation of cells) by binding to mannose chains on the surface of other cells. Based on its expression pattern AgFLO5a might play a role early in the development, maybe during germination.

The four genes **AgGAS5, AgGAS1b, AgFIG2** and **AgTOS1** displayed highest expression during vegetative growth. We might speculate that the fast constitutive filamentous growth of *A. gossypii* requires constant and high expression of these proteins that ensure integrity of the cell wall. The GAS multigene family in *S. cerevisiae* consists of five genes, ScGAS1-ScGAS5. The latter one codes (like ScGAS1) for a glycosylphosphatidylinositol (GPI) anchored beta-1,3-glucanosyltransferases required for cell wall assembly. ScFIG2, the homolog of AgFIG2, codes for a cell wall adhesin that is expressed specifically during mating and may be involved in maintenance of cell wall integrity during mating. ScTOS1 codes for a covalently-bound cell wall protein of unknown function.

Two genes, **NOHBY621** and **AgECM33/ PST1**, did not follow a specific pattern. NOHBY621 codes for a novel *A. gossypii* cell wall protein that will be discussed later in this chapter.
AgECM33/PST1 is a single gene in *A. gossypii* with twin homologs in *S. cerevisiae*. ScECM33 encodes a GPI-anchored protein of unknown function that has a possible role in apical bud growth. ScPST1 codes for a cell wall protein that contains a putative GPI-attachment site. The expression of ScPST1 is up-regulated by activation of the cell integrity pathway and by cell wall damage via disruption of ScFKS1. It is difficult to infer a role for AgECM33/PST1 based on its homologs which have rather diverse functions.

The *A. gossypii* genome encodes a CWP1 gene family with distinct expression patterns

The *A. gossypii* genome encodes seven homologs of the *S. cerevisiae* cell wall protein gene CWP1 (figure 5-2). Four of these homlogs, ABR025C, ABR026C, ABR027C and ABR028C, are syntenic homologs that origin from a tandem gene quadruplication. They were named CWP1a, CWP1b, CWP1c and CWP1d. The three non-syntenic homologs are ACR272C (CWP20a), ACR273W (CWP20b) and ADL398C (CWP21). Repeated sequences are frequently found in cell wall proteins. Repeated sequences found in ABR027C and ABR028C provide the explanation for the significant differences in length between the open reading frames (ORFs) which are 3396 bp and 1161 bp, respectively. The *S. cerevisiae* CWP1 protein contains a PIR repeat (proteins with internal repeats) which was not found in the *A. gossypii* homolog ABR025C. ABR025C was the only member of the CWP gene family in *A. gossypii* for which a GPI-anchor was predicted *in silico*

Figure 5-1 Expression patterns of 14 genes which encode proteins that were found in the *A. gossypii* cell wall.

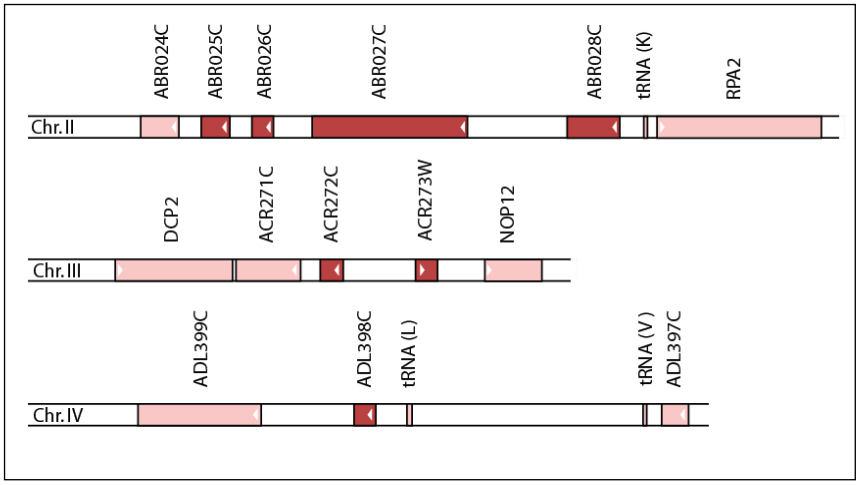
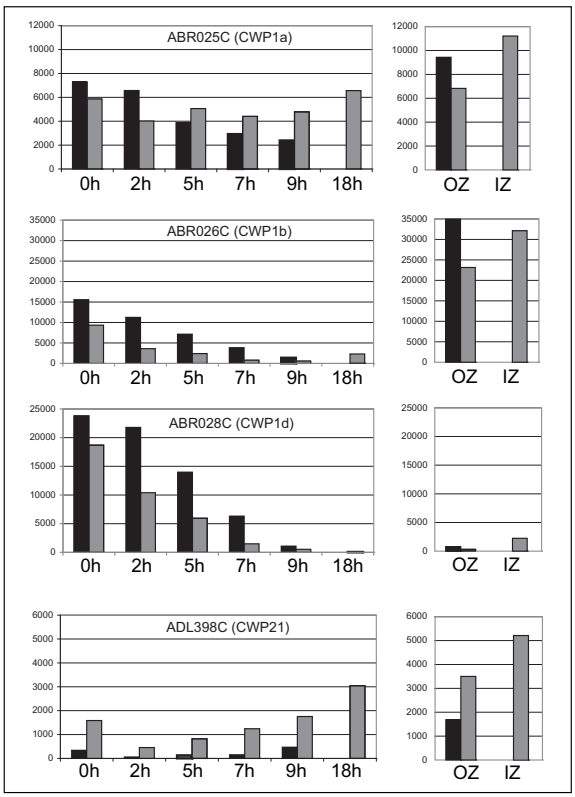


Figure 5-2 *A. gossypii* has seven homologs of the cell wall protein gene CWP1 (dark red boxes). On Chromosome II we have identified a syntenic tandem gene quadruplication of the CWP1 gene (map kindly provided by Mark Finlayson). Four of the CWP1 homologs (ABR025C, ABR026C, ABR028C and ADL398C) are represented as probesets on the sySYNG001a Gene Chip®.

and the only member of the family that was identified as a cell wall protein by mass spectroscopy (Piet de Groot, personal communication).

In *S. cerevisiae* a few specific antisera against individual Cwps have been described. Upon probing immunoblots with anti-Cwp1, we obtained specific signals for Cwp1, indicating that at least one *A. gossypii* homolog is present in the cell wall (Piet de Groot, personal communication).



Four of the members of the CWP family were represented on the Affymetrix array. Interestingly the expression patterns of ABR025C (CWP1a), ABR026C (CWP1b), ABR028C (CWP1d) and ADL398C (CWP21) differed remarkably (figure 5-3). Very high levels of both, ABR026C and ABR028C mRNAs, were found within the first few hours (in decreasing relative amounts). While the ABR026C transcript was found at very high levels in the outer zone of fast growing hyphe (OZ) and the inner zone of sporulation (IZ), high expression of ABR028C transcript was clearly related to the life cycle stage of the spore. ABR025C was nearly constantly expressed and expression levels of ADL398C increased (except for 0 h time point) with increasing inoculation time. A possible sporulation-

Figure 5-3 The barplots show expression levels of four (of the seven) *A. gossypii* homologs of the cell wall protein gene CWP1 in FDAG (black) and ΔlΔt (gray) strains at time points after inoculation. Expression levels are also shown for the tip region of fast growing mycelium (OZ = outer zone) and in a region with decreased nutrient supply (IZ = inner zone) where starting sporulation was observed. Interestingly the expression patterns of the four CWP1 copies differed remarkably. Very high levels of both, ABR026C and ABR028C mRNAs, were found within the first few hours (in decreasing relative amounts). While the ABR0026C transcript was found at very high levels in the outer- and inner zones (OZ and IZ), high expression of ABR028C transcript was clearly related to the lifecycle stage of the spore (ΔlΔt signal intensities found at 18 h and in the OZ were at background level). ABR025C was constitutively expressed in all life cycle stages and expression levels of ADL398C increased (except for 0h time point) with increasing inoculation time.

specific role of ABR026C is supported by a comparison with the natural isolate strain FDAG (see chapter 8) that follows the same expression pattern but at a higher absolute level. The higher expression level in FDAG makes sense if ABR026C was a sporulation-specific gene. A t-profiler analysis (Appendix 2) has shown that the sporulation gene group is significantly up-regulated in the natural isolate FDAG compared to the laboratory strain in the outer zone of fast hyphal growth (OZ).

The *S. cerevisiae* CWP1 is known to be a target gene of the filamentation pathway regulated by the transcription factor Tec1 (see chapters 6 and 7). Nitrogen starvation in haploid cells and glucose starvation in diploid cells leads to an activation of the filamentation pathway which results in invasive growth or filamentous growth, respectively. In *S. cerevisiae* Tec1 is repressed by Dig1. Upon pathway induction Dig1 is phosphorylated and released from the transcription factor Tec1 which is now free to mediate transcription from downstream promoters, for example from the CWP1 promoter. In *A. gossypii* ADL398C (CWP21) was the only of the four CWP genes on the array that was more than two fold up-regulated upon AgDIG1/2 deletion. ABR028C (CWP1d) was four fold down-regulated upon Ag*tec1* deletion and upon Ag*ste12* deletion.

Expression of chitin synthases

The polysaccharide chitin is an important structural component of the cell walls of many fungi. Chitin synthesis in *S. cerevisiae* is governed by three chitin synthases, ScCHS1, ScCHS2 and ScCHS3. ScCHS3 activity also depends on products of other genes. Chitin synthases in other fungi have been grouped into three classes, class I–III, according to their sequence similarity to the yeast chitin synthases. The rest are defined as chitin synthase class IV and V and are highly similar to ScCHS3. Chitin synthase class V genes have only been identified in filamentous fungi and their other functions are unknown. Class IV genes, which include ScCHS3 are involved in the synthesis of most chitin in the yeast cells (Valdivieso, 1999). Chitin synthases of class I are involved in repair functions at the end of cytokinesis; Chitin synthases of class II are responsible for the synthesis of the primary septum that separates mother and daughter cells; and Chitin synthases of class III are responsible for the formation of the ring (bud scar) where most of the cell wall chitin is located in the yeast *S. cerevisiae*.

Interestingly, the chitin synthase 3 (CHS3) which is responsible for the majority of the cell wall chitin was not expressed in *A. gossypii* during vegetative growth and was transcriptionally restricted to the sporulation process. AgCHS7 probably involved in chitin biosynthesis by export of Chs3p from the ER in *S. cerevisiae* based on the function of the yeast homolog, followed a similar expression pattern as AgCHS3 with two-fold up-regulation during sporulation. In addition, ACR227W, the yeast homolog of which encodes the activator of AgChs3p, showed sporulation-specific expression in *A. gossypii* with 5-fold up-regulation in the inner zone of sporulation compared to the outer zone of fast hyphal growth. ACR227W has two twin homologs in yeast. One, ScSHC1, has a similar expression pattern as ACR227W and is a sporulation-specific activator of ScChs3p whereas the other homolog, ScSKT5, activates ScChs3p during vegetative growth. These data suggest that, unlike the chitin synthase 3 of *S. cerevisiae*, the role of chitin synthase 3 in *A. gossypii* is limited to the sporulation process where it might be involved in the synthesis of spore wall chitosan. Sporulation-specific gene expression of AgCHS3, AgSHC1 and AgCHS7 is highlighted in gray in table 5-2 A.

All *S. cerevisiae* chitin synthases reside in membranes, either in vesicles or in the plasma membrane. Export of ScChs3 from the Golgi and transport to the plasma membrane is mediated via a group of four related,

A

Affymetrix ID	Systematic name(s)	Common name(s)	ΔlΔt 0h	ΔlΔt 2h	ΔlΔt 5h	ΔlΔt 7h	ΔlΔt 9h	ΔlΔt 9h	ΔlΔt 18h	ΔlΔt OZ	ΔlΔt IZ
Ag000519_at	ABL153W	CHS1	6	6	11	11	10	11	9	8	13
Ag002389_at	AEL190W	CHS2	60	193	574	861	759	793	1400	491	109
Ag002224_at	AEL189W	CHS3	25	19	9	8	7	8	28	10	59
Ag000828_at	ACR227W	CHS4 SHC1/SKT5	121	130	100	47	36	34	22	70	394
Ag003804_at	AGL051C	CHS5	210	180	238	313	375	349	387	409	332
Ag003494_at	AFR088W	CHS6/FMP50	97	87	105	172	217	234	631	451	192
Ag003188_at	AFR033C	CHS7	36	31	47	50	46	58	80	70	145
Ag003513_at	AFR528W	BCH1/BUD7	121	210	386	499	530	592	607	585	550

B

Common name(s)	SGD description
CHS1	Chitin synthase I, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for repairing the chitin septum during cytokinesis; transcription activated by mating factor
CHS2	Chitin synthase II, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for the synthesis of chitin in the primary septum during cytokinesis
CHS3	Chitin synthase III, catalyzes the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for synthesis of the majority of cell wall chitin, the chitin ring during bud emergence, and spore wall chitosan
CHS4 SHC1/SKT5	SHC1: Sporulation-specific activator of Chs3p (chitin synthase III), required for the synthesis of the chitosan layer of ascospores; has similarity to Skt5p, which activates Chs3p during vegetative growth; transcriptionally induced at alkaline pH; SKT5: Activator of Chs3p (chitin synthase III), recruits Chs3p to the bud neck via interaction with Bni4p; has similarity to Shc1p, which activates Chs3p during sporulation
CHS5	Protein involved in export from the Golgi to plasma membrane; involved in chitin biosynthesis through its role in Chs3p localization; interacts with Arf1p, Bch1p, Fmp50p, Bud7p, and Chs6p
CHS6/FMP50	Member of the Chs5p-Arf1p-binding proteins (ChAPs), a group of 4 related, interacting proteins (Bch1p, Fmp50p, Bud7p, Chs6p) that mediate export of specific cargo proteins, including chitin synthase Chs3p, from the Golgi to plasma membrane
CHS7	Protein of unknown function, involved in chitin biosynthesis by regulating Chs3p export from the ER
BCH1/BUD7	BCH1: Member of the Chs5p-Arf1p-binding proteins (ChAPs), a group of 4 related, interacting proteins (Bch1p, Fmp50p, Bud7p, Chs6p) that mediate export of specific cargo proteins, including chitin synthase Chs3p; may interact with ribosomes. BUD7: Member of the Chs5p-Arf1p-binding proteins (ChAPs), a group of 4 related, interacting proteins (Bch1p, Fmp50p, Bud7p, Chs6p) that mediate export of specific cargo proteins, including chitin synthase Chs3p, from the Golgi to plasma membrane

Table 5-2
A) Expression of *A. gossypii* genes CHS1, CHS2 and CHS3 that encode chitin synthases I-III and expression of genes involved in regulation and transport of chitin synthases. Sporulation-specific expression is indicated in gray (B) SGD description of *S. cerevisiae* homologs

interacting proteins (ScBch1, ScFmp50, ScBud7, ScChs6). We have already mentioned that the ScChs3p activators ScSHC1 and ScSKT5 are twin genes, meaning that we have only one homolog of this gene in *A. gossypii*. Interestingly, four members of the export complex are also twin pairs (ScCHS6 and ScFMP50; and ScBCH1 and ScBUD7). Thus, *A. gossypii* seems to have a much simpler system to regulate and export AgChs3p.

The chitin synthase 1, AgChs1p, was not expressed under the tested conditions. Based on the function of the yeast homolog we may speculate that AgChs1p has a function during mating or/and may be required for repair of the chitin septum. In *A. gossypii* hyphae multiple nuclei reside in the same cytoplasm and divide asynchronously (Gladfelter *et al.*, 2006). If only few nuclei, for instance the ones next to septa, expressed AgCHS1 we might not be able to detect these transcripts on the microarray. This is because, even for a decently transcribed gene, the fraction of the transcripts coming from that gene would be very low in comparison to all other transcripts.

The chitin synthase 2, AgChs2p, was expressed during all life cycle stages with a 23-fold increase in expression in advanced mycelia (18 h) compared to ungerminated spores (0 h). Mycelia at later life cycle stages growing on plates (OZ and IZ) showed reduced expression of AgCHS2. Judging

from expression data we can conclude that, unlike in *S. cerevisiae*, most of the chitin in *A. gossypii* is synthesized by AgChs2p.
Measurements of the overall composition of the *A. gossypii* cell wall revealed that chitin represents only 0.2% of the wall mass. The amount of chitin is lower as compared to *S. cerevisiae*, measured as 1.4% in the same assay condition (P. de Groot, unpublished data). This can be explained by the fact that most of the chitin in budding yeast is localized in the septa and the chitin ring and only about 10% of the chitin is present in the lateral walls (Molano et al., 1980). In *A. gossypii* we have only one chitin synthase expressed in advanced mycelium (18 h) whereas all three chitin synthases are expressed in a logarithmic growing culture of *S. cerevisiae*. This could at least partially explain the differences in chitin content. On the other hand post-translational regulation should not be disregarded as it has been reported to be important for chitin synthases, particularly for ScChs1p and ScChs2p.

AFR065W and AGL034C, two genes with previously unknown functions, encode cell wall proteins

AFR065W (NOHBY621) is a novel Ashbya-specific cell wall protein gene

There are 261 genes annotated in the *A. gossypii* genome that do not have homologs in the closely related budding yeast *S. cerevisiae*. AFR065W and AGL034C are two of these genes that were named NOHBYs (no homolog in baker's yeast). The NOHBYs have gained continuous interest from the very beginning in our laboratory. The initial idea that the 5% Ashbya-specific genes do account for the differences in the mode of growth between *S. cerevisiae* and *A. gossypii* is most likely wrong. None of the over 20 NOHBY deletions has turned *A. gossypii* into a budding fungus. Only two deletions showed slower growth, the others were indistinguishable from wild type. But what are the functions of the NOHBYs then, if not to contribute to sustained polar growth? We consider 83% of the 130 NOHBYs present on the array to be expressed which supports their annotation as authentic genes (definition of not expressed: signal intensity values < 60 in all ΔlΔt experiments from 0 h to 9 h). Furthermore, NOHBYs have homologs in species other than *S. cerevisiae* and domains can be predicted *in silico*. Nevertheless the assignment of a function to those genes needs conclusive experiments. As already mentioned, only two NOHBY knock outs resulted in the little informative phenotype "slow growth". When we discovered that one NOHBY was as highly expressed as cell wall protein genes, we started our collaboration about the cell wall proteome. To our pleasant surprise, NOHBY621 (AFR065W) was identified as cell wall protein by mass spectrometry. The 184 aa long protein had been predicted *in silico* to contain one PIR repeat (proteins with internal repeats) and it was extractable from the cell wall fraction with mild alkali (P. de Groot, unpublished data). In order to make predictions about possible functions of the newly discovered cell wall protein gene, BLAST searches were performed against the non-redundant database at NCBI. Manual inspection of the hits revealed no significant hits in fungal species such as *C. albicans*, *C. glabrata, D. hansenii, K. lactis, K. marxianus, K. waltii, N. crassa, P. Pastoris, S. cerevisiae, S. pombe, Y. lipolytica* (Sylvia Voegeli, personal communication). We conclude that AFR065W (NOHBY621) is a highly expressed cell wall protein gene that has not been described previously and is specific to the filamentous fungus *A. gossypii*. Future studies will hopefully shed light on the function of this novel cell wall protein.

AGL034C (NOHBY702) is a cell-wall protein gene with homology to superoxide dismutases

Hydroxyl and superoxide radicals react with cellular components, resulting in oxidation of proteins and nucleic acids as well as lipid peroxidation. Superoxide dismutatases play an important role in cell detoxification by converting a damaging superoxide radical, O_2-, into a less damaging hydrogen peroxide and an oxygen molecule (dioxygen; Bermingham-McDonogh et al., 1988; Steinmann et al., 1980). The *S. cerevisiae* genome encodes two superoxide dismutases, a CuZnSOD (SOD1) and a MnSOD (SOD2), the first with cytosolic and the second with mitochondrial localization. Syntenic homologs, SOD1 (AGL321W) and SOD2 (ADR160W), are present in *A. gossypii*. Proteome analysis revealed an additional cell-wall related CuZnSOD in *A. gossypii*, NOHBY702 (AGL034C), that shares homology with a CuZnSOD from *N. crassa* (best reciprocal BLAST hit) and *C. albicans* (P. de Groot, personal communication). The fact that the third *A. gossypii* superoxide dismutase has homologs in pathogens might be a hint to the importance of this gene in infection. *N. crassa* is a saprophyte (growing on decayed plant material) and not a real pathogen, nevertheless the genome sequence has revealed many proteins and systems with functions that have previously been found to be important in plant pathogens, for example, proteins such as cytochrome P450 and efflux systems involved in detoxification of plant antifungal compounds (Galagan *et al.*, 2003).
Aerobic eukaryotic pathogens can encounter superoxide radicals from several internal or external sources. An important internal source is the mitochondrial respiratory chain (Boveris *et al.*, 1978; Casteilla *et al.*, 2001; Lenaz, 2001). External sources are superoxide radicals coming from oxidative bursts produced by the host to cope with pathogenic infections. Another external source in the natural environment of *A. gossypii* is UV-light. Two possible functions of cell-wall related superoxide dismutases might be to protect *A. gossypii* against attacks of the cotton plant during and after infection or against cell damage induced by UV-light. *A. gossypii* SOD1 (AGL321W) and NOHBY702 (AGL034C) were predicted *in silico* to be GPI-anchored cell wall proteins and the presence of NOHBY702 (AGL034C) in the *A. gossypii* cell wall was confirmed by mass spec analysis of the cell wall fraction (Piet de Groot, personal communication). SOD1 (AGL321W) was nearly constantly expressed at very high level. The highest level measured in the inner zone of sporulation mycelium (IZ). The newly identified cell-wall protein gene NOHBY702 (AGL034C) was 5.5-fold up-regulated within the first 18 hours of inoculation and 34-fold up-regulated in the inner zone of sporulating mycelium (IZ).
Possible roles of superoxide dismutases during the sporulation process have not been intensively discussed in the literature so far. In *S. cerevisiae*, it has been reported that the double deletion of the CuZn-SOD (SOD1) and the Mn-SOD (SOD2) leads, in addition to hypersensitivity to oxygen toxiticity, to a sporulation defect. Single deletion of the SOD1 gene in *N. crassa* did not have any effect on sporulation (Narasipura *et al.*, 2003). We found two research articles about slime molds, one reporting elevated superoxide dismutase-specific activity during starvation-induced sporulation in the myxomycete Didymium iridis (Lott *et al.*, 1981) and the other reporting expression of SODD, a Cu/Zn superoxide dismutase in Dictyostelium discoideum homologous to *S. cerevisiae* SOD1, at elevated levels at late developmental stages. SODD gene expression was not detected in the sporeless mutant, indicating that sodD is a spore cell-specific gene (Akaza *et al.*, 2002).

Correlation of cell wall transcriptome and proteome

The recent availability of platform technologies for high throughput proteome analysis has led to the emergence of integrated messenger RNA and protein expression data (Hack, 2004). Two early studies (Futcher *et al.*, 1999; Gygi *et al.*, 1999) have quantitatively examined the yeast proteins and mRNAs by gathering information from two-dimensional gel electrophoresis (2DE) and from serial analysis of gene expression (SAGE) or oligonucleotide arrays, respectively. These studies arrive at contradictory conclusions about the goodness of the correlation. Today it is generally accepted that mRNA is not a direct indicator of protein levels. The amount of protein is dependent on many factors such as its half-life and its cellular location, or the protein may have to be modified or interact in complexes to become active. Methods for the global analysis of protein expression are just emerging and traditional methods to study the proteome such as 2DE followed by mass spectrometric analysis are biased against large, small or highly charged proteins, less soluble proteins and less abundant proteins (Humphrey-Smith, 1997). Nevertheless both early studies come to the conclusion that the correlation between mRNA and protein levels is good for abundant proteins. In addition an experiment looking at the protein turnover revealed that it was insignificant for abundant proteins (Futcher *et al.*, 1999).

As many of the predicted *A. gossypii* cell wall protein genes are highly expressed (table 5-1) we would expect to find high protein levels for most of them and hence be able to detect them by mass spec. Let us first have a look at “highly” expressed genes. In the experiment with 18 h old mycelium, 80 percent of the data lies within a signal intensity value range from 0.1 to 901, and 90 percent of the data lies within a range from 0.1 to 1842. Of the 48 *in silico* predicted *A. gossypii* cell wall protein genes present on the array, 23 are above the 80th percentile and 16 are above the 90th percentile. There was a total of 15 proteins identified as cell wall proteins by mass spectrometry and 14 of them are represented on the array. The expression profiles are shown in figure 5-1. 13 of the 14 identified proteins are found among the top 23 on mRNA level (table 5-1). The six proteins that were among the top 23 in mRNA level but were not found by mass spec are Bgl2, Sun4, Cwp21 (ADL398C), Cwp1b (ABR026C), Sod1, Ccw12, AAR076C, Kre1, Pmp3 and Rot1. Sun4 is probably secreted into the medium, the Cwp1 homologs do not have predicted GPI-anchors and Ccw12 is a very small highly mannosylated protein that is not found in the mass spec due to lack of suitable unglycosylated tryptic peptides (P. de Groot, personal communication). The protein that was detected by mass spectrometry but was not among the top 23 on mRNA level was Flo5a (AFL092C). The reason for that might be that the signal intensity of 322 is an underestimation. We found that the probe set targeting Flo5a, Ag003188_at, was designed over a short stretch of sequence positioned in the middle of the 2808 bps long gene rather than at the 3’ end. Since the Affymetrix eukaryotic expression assay has an inherent 3’ bias, the signal intensity obtained from a 3’ probe set might be an estimated 3-4 times higher.

5.3 Concluding remarks

There is an excellent overlap between the most highly expressed cell wall protein genes and the cell wall proteins that were identified by mass spectrometry. Since mass spectrometry is biased towards identifying abundant proteins, we conclude that the observed high expression of cell wall protein genes is reflected in high protein abundance, which is an important observation because proteins and not mRNAs drive functional processes in cells.

The fungal cell wall is an essential organelle that fulfills a wide variety of functions. Possession of a cell wall makes it possible to

withstand turgor pressures, maintain stable osmotic conditions, maintain the cell shape, allow morphogenesis, and protect the cell against physical damage (Klis *et al.*, 2007). The cell wall is also directly in contact with the environment, and interaction with the outside world may determine the successful survival of the cell. The budding yeast *S. cerevisiae* and the filamentous fungus *A. gossypii* have different modes of growth and are found in different habitats in nature. These differences are reflected in the cell wall proteome. For instance, the *A. gossypii* cell wall possesses two proteins that have no homologs in *S. cerevisiae*, one of which shares homology to superoxide dismutases. Furthermore the chitin content differs remarkably between the two fungi (Piet de Groot, personal communication) which might at least partially be due to differences in expression of the three chitin synthases.

The question why cell wall protein genes are so highly expressed in *A. gossypii* remains to be answered. The growth of the cell wall must accompany that of the cell that it encloses. In particular new cell wall must be generated, for example at the emergence of a new bud in *S. cerevisiae*; its synthesis continues throughout the growth of the bud. Growth in *A. gossypii* is a continuous process that is limited to the hyphal tip region. The maximal surface expansion rate of the cell wall is 40 μm^2/min in fast (3 μm/min) elongating hyphae (Knechtle *et al.*, 2006) and (H. Helfer, unpublished data). This constitutes a 40 x higher surface expansion rate compared to the bud growth phase in *S. cerevisiae* explains the very high expression levels of cell wall proteins and their putative assembly enzymes. In *S. cerevisiae* it is known that cell wall construction and usage of specific cell wall macromolecules are temporally and spatially controlled and closely coordinated with cell cycle progression (Caro *et al.*, 1998; Rodriguez-Pena *et al.*, 2002; Cabib *et al.*, 2005; Sumita *et al.*, 2005; Smits *et al.*, 2006). This is reflected by cell cycle-controlled expression of many cell wall proteins and putative assembly enzymes (Ram *et al.*, 2006; Smits *et al.*, 2006). In *A. gossypii* controls similar to cell-cycle are difficult to imagine because of the lack of cell separation. Whereas transcription of several cell wall genes seems to be constitutive, others are for sure “life cycle” regulated.

Chapter 6: Conservation of the pheromone and filamentation pathways in the filamentous fungus *A. gossypii*

Chapter 6: Conservation of the pheromone and filamentation signaling pathways in the filamentous fungus *A. gossypii*

6.1 Introduction

In order to understand the molecular basis of phenotypic diversity and evolution, characterization of interspecies differences in gene regulation is crucial. *A. gossypii* is a filamentous ascomycete that carries a genome which resembles that of *S. cerevisiae* prior to its whole genome duplication. Despite the high degree of conservation of gene order and gene content, *A. gossypii* strictly grows as a fungal mycelium with constantly elongating and branching hyphae.

In the yeast *S. cerevisiae* key regulators determine the cell fate and control developmental pathways. The transcription factor Ste12 is required in haploids for mating and in diploids for pseudohyphal development in cooperation with the transcription factor Tec1. Mating-specific genes and the pseudohyphal pathway are regulated in *S. cerevisiae* by the repressors Dig1 and Dig2. A comparison of the mating and filamentation mitogen-activated protein kinase (MAPK) pathways and regulatory networks in *A. gossypii* and *S. cerevisiae* is of particular interest. As we will show in this chapter, all major components of the pathways are conserved. Nevertheless, there are differences in the modes of growth, and the potential for mating has not been experimentally established in *A. gossypii* due to the lack of the opposite mating type to the Mat a-like genotype determinded by sequencing of the original and only *A. gossypii* isolate used in the laboratory.

As a first step, a bioinformatics analysis was carried out. This analysis looked into the conservation of transcription factor binding sites in target genes of the mating and the filamentation pathway. Then, AgDIG1/2, AgTEC1 and AgSTE12 were deleted and phenotypes were characterized under conditions where the MAPK cascades were not induced by external stimuli.

6.2 Results

The *A. gossypii* genome encodes homologous genes of all components of the *S. cerevisiae* mating and filamentation pathway

Mitogen-activated protein (MAP) kinases cascades mediate signal transduction from the cell surface to the nucleus in response to a wide variety of signals. In *S. cerevisiae* mating signals (pheromone) and filamentation signals (nitrogen or glucose depletion) lead to the activation of ScSte12, a transcription factor that binds to pheromone responsive elements (PREs) of mating genes. Filamentation genes are induced by binding of active ScTec1/ScSte12 complexes to ScTec1 binding sites (Chou *et al.*, 2006) upstream of filamentation genes. Under conditions where the pathways are not induced, ScSte12 is repressed by either ScDig1 and ScDig2 or ScDig1, depending on the ScSte12 complex (table 6-1). Figure 6-1 shows a diagram of the MAP kinase pathways that control the mating and the pheromone response in *S. cerevisiae*. Many components are shared between the two pathways (e.g. ScSte11, ScSte7, ScDig1, ScDig2 and ScSte12, depicted in blue) whereas other components are specific to the mating (ScSte5, ScFus3; depicted in black) or the filamentation pathway (ScKss1, ScTec1; depicted in gray). Crosstalk between the two pathways

is prevented under mating conditions by ScFus3 mediated phosphorylation and subsequent degradation of Tec1 (Bao *et al.*, 2004). We were wondering whether one does find homologs of these two networks in *A. gossypii*, an exclusively filamentously growing fungus.

The *A. gossypii* genome encodes homologs of all major components that are required in *S. cerevisiae* to respond to mating pheromone and filamentation signals. This includes not only the p21-activated kinase Ste20, the MEK kinase Ste11, the MAP/ERK kinase Ste7, and the mitogen-activated protein (MAP) kinases Fus3 and Kss1 but also the pheromone receptors (Ste2, Ste3), the G-protein subunits (Ste18, Ste4, Gpa1), the scaffold protein Ste5, the repressor Dig1/2, the transcription factors Ste12 and Tec1, and most target genes. MAP kinases are a family of serine threonine kinases, whose function have been highly conserved during evolution across all kingdoms and phyla.

The overall protein identity of the MAP kinases AgFus3 and ScFus3 was 69% and that of AgKss1 and ScKss1 was 66% (table 6-2). A pairwise comparison of all homologous components in the pathways revealed that the transcription factors and the repressor were the least conserved proteins. DIG1 and DIG2 are twin genes in *S. cerevisiae*. This means that *A. gossypii* carries one syntenic homolog for each of these twin genes, that was named AgDIG1/2. Interestingly, the protein sizes differ remarkably. ScDig1 is 422 aa, ScDig2 is 323 and AgDig1/2 is 227 aa in size. Big size differences were

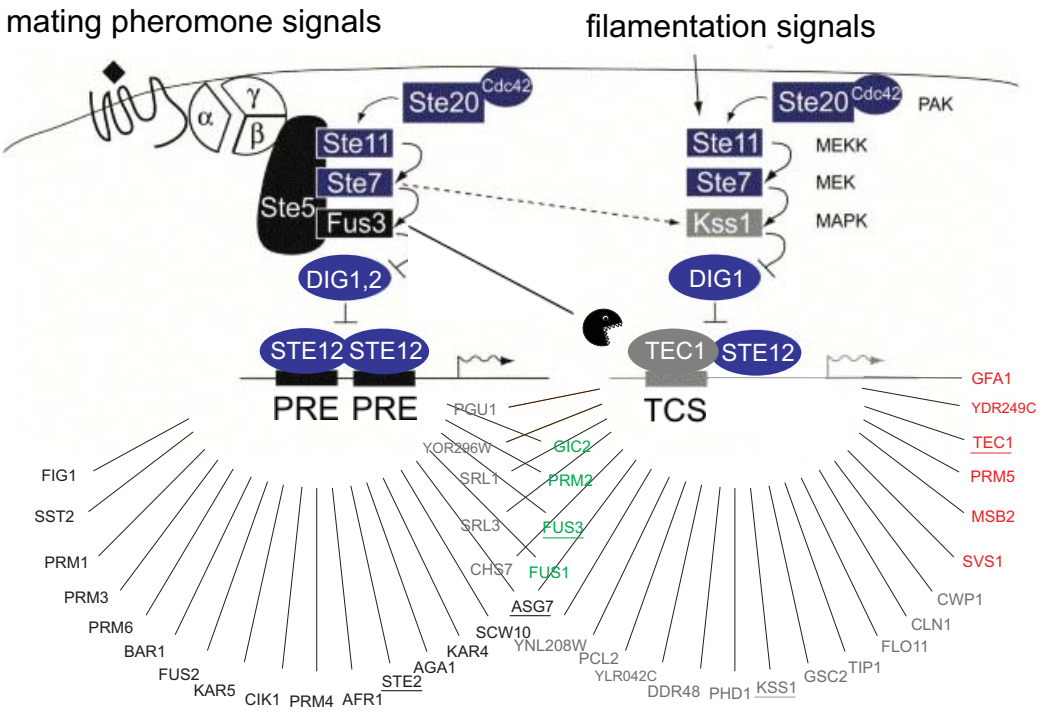


Figure 6-1 Diagram of the MAP kinase pathways that control the mating and the pheromone response and filamentous growth in *S. cerevisiae*. Mating-specific components are depicted in black, filamentous-specific components in gray and shared components in blue. Ste12p is involved in both transcriptional programs and regulates genes via PRE (Pheromone responsive element; TGAAACR). Tec1p is a filamentation-specific component that regulates target genes via TCS elements (TEC1 binding site; CATTCY) and is degraded in the mating process which prevents crosstalk between the two pathways. Dig1p and Dig2p bind and inactivate Ste12p complexes. Target genes of the two pathways, depicted in semi-circular arrangements, were identified because of their at least two-fold up-regulation in *dig1Δdig2Δ* versus wt microarray data (Hughes et al., 2003; Liu et al., 2006). Target genes of the mating pathway are depicted in black and target genes of the filamentation pathway in gray. Green genes are thought to be regulated by the mating pathway but have also TCS binding sites in their promoters and red genes are thought to be regulated by the filamentation pathway, but have also PRE binding sites in their promoters. The figure was modified from (Bao et al., 2004).

protein	*S.c.* function	% identity
Components of the MAP kinase pathway		
Ste20	p21-activated kinase	47%
Ste11	MEK kinase	57%
Ste7	MAP/ERK kinase	47%
Fus3	MAP kinase	69%
Kss1	MAP kinase	66%
Ste5	scaffold protein	30%
G-protein subunits and pheromone receptors		
Gpa1	α-subunit of G-protein	69%
Ste4	β-subunit of G-protein	53%
Ste18	γ-subunit of G-protein	43%
Ste2	α-factor receptor	38%
Ste3	a-factor receptor	46%
Transcription factors and their repressors		
Dig1/2	repressor	19/25%
Tec1	transcription factor	20%
Ste12	transcription factor	36%

Table 6-2 Pairwise comparison of homologous *A. gossypii* and *S. cerevisiae* proteins reveals that the transcription factors are the least conserved components of the mating and filamentation pathway.

also observed for ScTec1 (486aa) and AgTec1 (791aa). The N-terminal part with the highly conserved DNA-binding domain TEA at positions 127-192 (*S. cerevisae*) and 130-195 (*A. gossypii*) and the central part of Tec1 with the phosphothreonine at positions 273 (*S. cerevisiae*) and 274 (*A. gossypii*) were highly conserved. The observed size difference is mainly due to a 305 aa extension at the C-terminus of AgTec1 that was found to be glutamine-rich and histidine-rich. The function of these stretches in AgTec1 is not known but they might be involved in transcriptional activity. Recently, a strong positive correlation between glutamine-rich activators and strong transcriptional activation has been reported (Titz *et al.*, 2006). Polyglutamine stretches are of medical importance in the Huntington disease which is caused by a dominant polyglutamine expansion within the N-terminus of the huntingtin protein that changes its transcriptional activity. The result is oxidative stress, energetic insufficiency and striatal degeneration. The ScTec1 lacks the glutamine-rich region but multiple sequence alignments of Tec1 from several fungi revealed a shorter conserved polyglutamine stretch in the C-terminus of *S. kluyveri*.

The protein sizes of ScSte12 and AgSte12 were 655 aa and 688 aa, respectively. Changes in protein length could account for the low protein identities in Dig1/2 and Tec1, but not in Ste12. We were wondering whether homologous *S. cerevisiae* and *A. gossypii* transcription factor pairs were less conserved on a global scale compared to all possible homologous *S. cerevisiae* and *A. gossypii* protein pairs. For this investigation we used 200 annotated transcription factors from SGD (figure 6-3). The result of the investigation was that transcription factors have indeed evolved faster during the 100 million years of evolution since divergence of the *A. gossypii* from the *S. cerevisiae* lineage. Low overall conservation and changed putative activation domain in Tec1p notwithstanding, the DNA-binding sites of Tec1 and Ste12 were highly conserved, not only in comparison with *S. cerevisiae* but also in comparison with other fungal species. The 69 amino acid long AgTec1 DNA-binding site, for instance, shared 74% identity and 86% similarity with the ScTec1 DNA-binding site, and 62% identiy and 86% similarity in a multiple sequence alignment with *S. mikatae, S. cerevisiae,* S*. bayanus, S. kudriavzevii, S. kluyveri and S. castelli*.

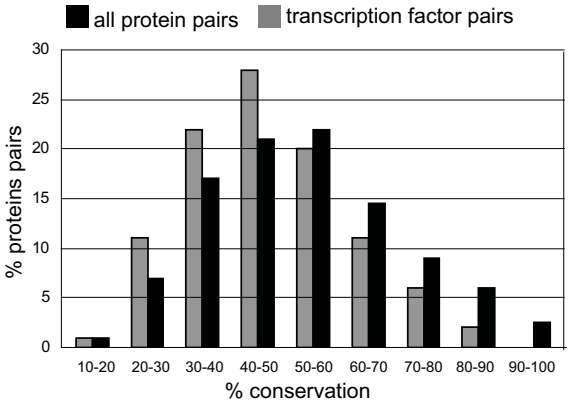


Figure 6-3 Protein pairs annotated as transcription factors are less conserved than other protein pairs. A density plot of protein identities of 200 *A. gossypii* and *S. cerevisiae* transcription factor pairs in the context of all possible protein pairs is shown.

Identification of *A. gossypii* homologs of ScSte12 and ScTec1 target genes

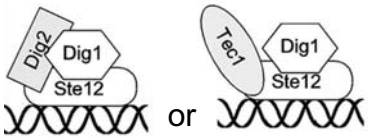
The first step in the analysis was to identify target genes of the two pathways in *S. cerevisiae*. This data was kindly provided by H. Liu and S. Chou prior to publication and has been published in the meantime (Chou *et al.*, 2006). Liu *et al*. looked for genes that were at least two fold up-regulated in Sc*dig1*Δ*dig2*Δ microarrays (Hughes *et al.*, 2000) versus wild type and contained conserved Ste12 PRE (TGAAACR) and/or Tec1 TCS (CATTCY) sites in four closely related Saccharomyces species. The 1000-bp upstream sequences of each gene from all four species were searched. Ty genes were excluded from the list. In addition FLO11, CLN1, CWP1, and TEC1 were added as they were shown in other studies to be regulated by the filamentation pathway in a Ste12/Tec1 dependent manner (Madhani et al., 1999; Zeitlinger et al., 2003). The identified *S. cerevisiae* target genes are shown in figure 6-1. Genes depicted in green are thought to belong to the mating pathway but contained in addition to the PRE site(s) at least one TCS site. The genes in red are target genes of the filamentation pathway but their 5' upstream regions contain in addition to the TCS site(s) at least one PRE site. We identified *A. gossypii* homologs for all genes in figure 6-1, except for PRM4, AGA1 YNL208W and YLR042C.

Target gene groups of the pheromone and the filamentation pathway displayed differential levels of promoter element conservation

In order to elucidate the potential for mating and filamentation regulatory network conservation, we analyzed the *A. gossypii* 5' upstream regions (entire inter-ORFs) for TCS and PRE sites. The results are reported in table 6-1 in context with the data from *S. cerevisiae* and in figure 6-4. Most *S. cerevisiae* promoters belonging to the mating gene group have PRE elements (Ste12p binding sites) but lack TCS elements (Tec1p binding sites). We found that 72% (13/18) of the promoters which had homologs in *A. gossypii* had also at least one PRE element. The enrichment of PRE elements in the mating gene group suggests conservation of the mating regulatory network. All *S. cerevisiae* promoters from the filamentation gene group have at least one TCS element and most promoters lack PRE elements. Interestingly only 35% (7/20) of the homologous *A. gossypii* promoters have TCS elements. The original PRE and TCS elements were randomly permuted to obtain a TCS control motif (TAARCGA) and a PRE control motif (CYTACT). Not surprisingly the TCS control motif was conserved more frequently than the PRE control motif. The TCS control motif was shorter (6mer) than the PRE control motif (7mer) and its GC content (33-50%) reflected better the average GC content of the *A. gossypii* inter-ORFs (48.8%). This observation strengthens the high level of conservation that was found for PRE elements.

The low conservation level of TCS elements in the filamentation group allows for the following possibilites. AgTec1 binds the same set of genes as ScTec but the transcription factor binding sites are not recognized any more by our search algorithms. This is because of co-evolution of the binding sites and the DNA binding site in AgTec1. However this option is not very likely because the Tec1 DNA binding site was highly conserved. Furthermore we should be able to find at least one changed conserved binding site in the upstream regions. We searched for any 5-8 nucleotide long conserved motifs 10 to 500 bps upstream of each gene in the Tec1 target gene set and did not find any significantly enriched binding site. The sequence CGGTAT was observed upstream of 10 out of the 20 genes. Although the sequence occurred so often, the conservation was not significant because the false positive probability was 0.208 (after multiple testing correction). In

PRE promoters

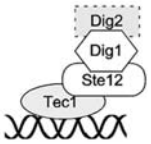


S.c. promoters with PRE elements				Promoters of *A.g.* homologs with PRE and TCS elements			
S.c. genes	PRE	TCS	pathway	*A.g.* homologs	PRE	TCS	presence of at least one PRE in *A.g.*
FIG1	1	0	mating	ACL175W	1	0	yes
SST2	2	0	mating	ABR176C	2	1	yes
PRM1	1	0	mating	AFR685C	2	(1)	yes
PRM3	1	0	mating	NOASH			
PRM6	1	0	mating	AFR378W	2	0	yes
BAR1	2	0	mating	AGR240W	0	0	
FUS2	2	0	mating	AGR240W	0	0	
KAR5	1	0	mating	ABL013C	1	0	yes
CIK	1	0	mating	AFL170C	0	0	
PRM4	1	0	mating	ACR146W	1	0	yes
AFR1	1	0	mating	AGR154C	1	0	yes
STE2	1	0	mating	AFR522C	0	0	
AGA1	3	0	mating	NOASH			
KAR4	2	0	mating	AFR736C	2	1	yes
SCW10	2	0	mating	AGL354C	1	0	yes
ASG7	2	0	mating	AEL129C	1	0	yes

S.c. promoters with PRE and TCS elements				Promoters of *A.g.* homologs with PRE and TCS elements			
S.c. genes	PRE	TCS	pathway	*A.g.* homologs	PRE	TCS	presence of at least one PRE in *A.g.*
FUS1	3	1	mating	AFR699C	3	2	yes
FUS3	2	1	mating	AFR019W	2	0	yes
PRM2	1	1	mating	ADL106C	0	0	
GIC2	1	1	?	AAL047C	1	0	yes

Table 6-1 PRE (TGAAACR) and TCS (CATTCY) elements in *S. cerevisiae* and *A. gossypii* promoters of genes involved in mating and filamentation. Protein complexes that bind and regulate gene expression in *S. cerevisiae* are shown (Liu et al., 2006). Most *S. cerevisiae* promoters belonging to the mating gene group have PRE elements (Ste12p binding sites) but lack TCS elements (TEC1p binding sites) and most promoters belonging to the filamentation group have TCS elements but lack PRE elements. Homologous *A. gossypii* inter-ORF regions were searched for conservation of at least one conserved PRE (PRE promoters) or TCS site (TCS promoters).

TCS promoters



S.c. promoters with TCS elements				Promoters of *A.g.* homologs with PRE and TCS elements			
S.c. genes	PRE	TCS	pathway	*A.g.* homologs	PRE	TCS	presence of at least one TCS in *A.g.*
PGU1	0	2	filamentation	ADR327W	0	0	
YOR296W	0	1	filamentation	ABL016C	0	0	
SRL1	0	1	filamentation	AAL092C	0	0	
SRL3	0	2	filamentation	ADR233W	0	0	
CHS7	0	1	filamentation	AFR033C	0	0	
YNL208W	0	2	?	NOASH			
PCL2	0	3	?	AFR159C	0	0	
YLR042C	0	2	filamentation	NOASH			
YMR173W	0	1	filamentation	AFL125W	0	0	
PHD1	0	1	filamentation	ABR055C	0	0	
KSS1	0	1	filamentation	ACL191C	1	2	yes
GSC2	0	1	?	AAR053W	0	0	
TIP1	0	1	?	ADR019C	0	1	yes
FLO11	0	2	filamentation	AEL023C	0	0	
CLN1	0	2	filamentation	AFL174C	0	4	yes
CWP1	0	1	filamentation	ABR026C	0	0	

S.c. promoters with PRE and TCS elements				promoters of *A.g.* homologs with PRE and TCS elements			
S.c. genes	PRE	TCS	pathway	*A.g.* homologs	PRE	TCS	presence of at least one TCS in *A.g.*
SVS1	3	2	filamentation	AAL092C	1	3	yes
MSB2	2	3	filamentation	AGR019C	1	2	yes
PRM5	1	2	filamentation	ADL251C	0	1	yes
TEC1	4	3	filamentation	AER177W	0	0	
YDR249C	1	3	?	AGL243W	0	1	yes
GFA1	1	2	filamentation	ABL036C	0	0	

the literature the sequence CGGTAT has not been described as a transcription factor binding site so far.
A second possibility is that AgTec1 regulates the same set of genes as ScTec1 but not by direct binding of the target genes. The third possibility is that AgTec1 has a (partially) different role and the regulation network has been rewired during evolution. These possibilities can be addressed by investigation of differentially expressed genes upon AgTEC1 deletion in microarray experiments.

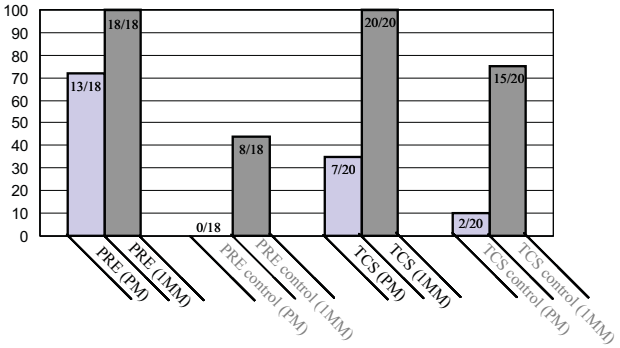


Figure 6-4 Conservation of (at least one) PRE and TCS elements in 18 mating- and 20 filamentation pathway target genes of *S. cerevisiae* and their homologs in *A. gossypii.* 72% (13/18) of PRE elements and 35% (7/20) of TCS elements are conserved in *A. gossypii*. Searches were repeated allowing 1 mismatch (mm) anywhere in the consensus sequence and with random control motifs (TCS control = TAARCGA; PRE control = CYTACT).

Deletion of the transcription factors TEC1, STE12 and the repressor DIG1/2 in *A. gossypii*

Deletions of entire open reading frames were made via the one-step gene targeting method (Wendland et al., 2000) using pairs of oligonucleotides for each locus, which contained at least 45-bp homology upstream and downstream of the ORFs (table 6-2). The primary transformation produces heterokaryon cells, which have a mixture of wild-type and transformed nuclei; thus, even mutations in essential genes produce viable transformants. To evaluate phenotypes, the heterokaryon was sporulated to produce uninucleated spores, which were isolated by micromanipulation and then allowed to germinated under selective conditions.
We compared the radial growth speed on plates of the three deletion mutants from day 3 to day 7. Figure 6-5 shows homokaryotic mycelia of ΔlΔt, *dig1/2*Δ, *tec1*Δ and *ste12*Δ on full medium plates. Wild type mycelium reached the maximum growth speed of 180 µm/h after three days and covered a standard petri dish in approximately seven days. We measured a reduced growth speed of 121 µm/h in *dig1/2*Δ. This accounts for 67% of wild type growth speed. In addition the colony surface of *dig1/2*Δ was uneven and wrinkled, and we realized that no spores formed in this mutant (figures 6-5 B and 7-1 B).
Cells deleted for STE12 grew with normal growth speed of 177 µm/h and showed no phenotypic changes under all conditions tested. However, mating conditions could not be tested with our mating type a laboratory strain because an opposite mating type has not been isolated from nature to this day. In *S. cerevisiae* artificial induction of the mating cascade can been done in several ways, for instance by deletion of the α–subunit of the G-protein Gpa1 (Miyajima *et al.*, 1987),

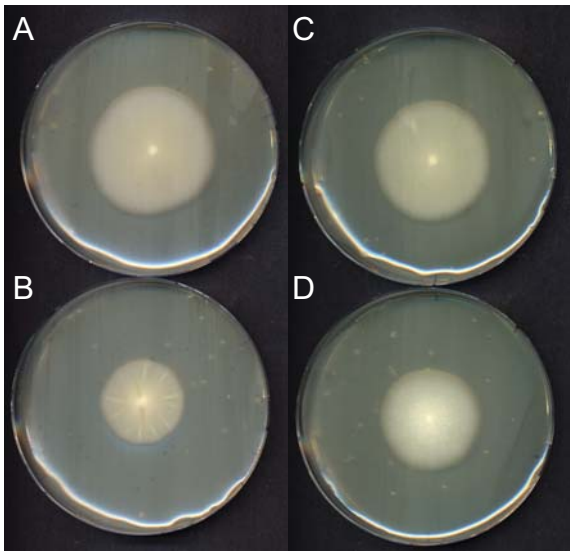


Figure 6-5
Mycelia at 103 h of
(A) ΔlΔt (wild type)
(B) *dig1/2*Δ
(C) *ste12*Δ
(D) *tec1*Δ

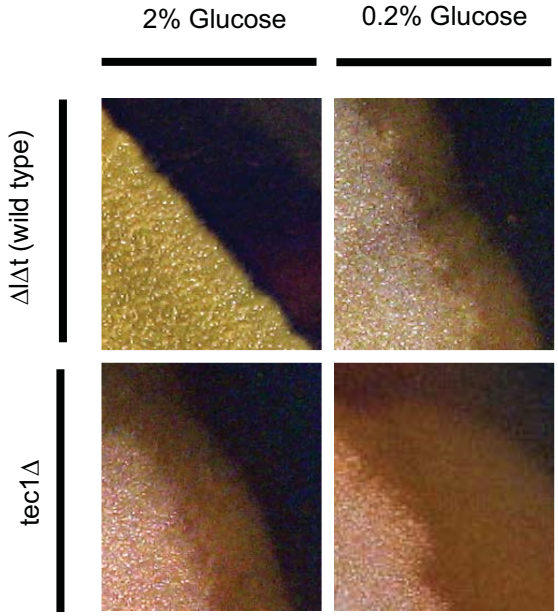


Figure 6-6 Invasive growth at the border of the colony with complete submergence of hyphae into the agar is observed in Ag*tec1*Δ cells after 10 days on full medium plates (bottom left). Invasive growth was observed in wild type, but so far only under glucose limiting conditions (top right).

by deleting the Ste12p repressor Dig1/2 (Bardwell *et al.*, 1989), by a constitutively active Ste11 mutant (Harris *et al.*, 2001) or simply by addition of mating pheromone. In *A. gossypii* trials to artificially induce the mating cascade are problematic because after all it is not known with certainty which morphological or molecular changes are to be expected in *A. gossypii* upon mating pathway induction in the absence of a mating partner.

Cells deleted for TEC1 grew with nearly normal speed (163 μm/h) but started after ten days to grow invasively at the border of the colony with hyphae completely submerging into the agar and disappearing from the agar surface. Invasive growth had not been observed in wild type under normal conditions but we found that a ten-fold reduction in glucose concentration (0.2% glucose) had the same phenotypic effect on wild type as the deletion of TEC1 (figure 6-6). Interestingly, the deletions of ScTEC1 and AgTEC1 have opposite effects: in *S. cerevisiae* TEC1 is required for invasive growth upon nitrogen starvation whereas in *A. gossypii* the deletion of TEC1 leads to invasive growth.

To further characterize the phenotypes, the actin cytoskeleton, the septa and the nuclei were stained and examined under the fluorescence microscope. The cytoskeleton and the septa were normal in all deletion mutants. The nuclear distribution was normal in Ag*dig1/2*Δ and Ag*ste12*Δ, but Ag*tec1*Δ showed an increased nuclear distance of 6.3 μm which was almost 50% longer than the distance measured for wild type under the same conditions (figure 6-7).

6.3 Concluding remarks

Components of the mating and filamentation MAPK pathways were highly conserved as concluded from sequence comparison between *A. gossypii* and *S. cerevisiae*. However the transcription factors at the end of the cascades, Ste12, Tec1 and the repressor Dig1/2, seem to have evolved faster during the 100 million years since the separation of the *A. gossypii* and the *S. cerevisiae* lineage. An *in silico* comparison of AgTec1 and ScTec1 revealed a changed potential activation domain in the AgTec1 C-terminus that contained glutamine- and histidine-rich sequences. We were wondering whether AgSte12/AgTec1 regulate the same set of genes as ScSte12/ScTec1. In a first step to elucidate regulatory network conservation, a bioinformatics analysis was carried out. The analysis looked into conservation of transcription factor binding sites in *S. cerevisiae* target genes of the mating and the filamentation pathway and their *A. gossypii* homologs. We found that Ste12 binding sites were highly conserved whereas binding sites for Tec1 were poorly conserved. This suggests that either AgTec1 regulates the same set of genes as ScTec1 by binding to another, yet unknown transcription factor binding site or by indirect regulation. Or, AgTec1 has a (partially) different role and the regulatory network has been rewired during evolution.

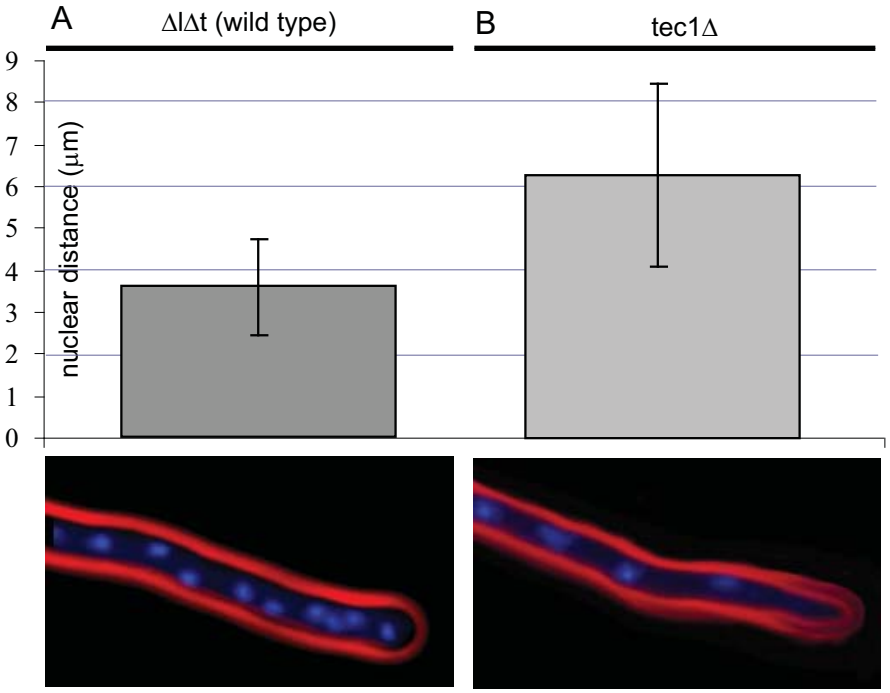


Figure 6-7 The average nuclear distance in Ag*tec1Δ* had nearly doubled to 6.3 µm (B) compared to the wild type nuclear distance of 3.6 µm (A). Mycelia were grown under identical conditions for 18 h. Fluorescent and phase contrast pictures were overlayed to show the nuclei stained with Hoechst (blue) and the outlines of the hyphae (red).

The AgSTE12, AgTEC1 and AgDIG1/2 genes were disrupted and the deletion phenotypes were characterized under conditions where the mitogen-activated protein (MAP) kinase cascades were not induced by external stimuli. *A. gossypii* cells that lack one of the two transcription factors or the repressor have phenotypes which remarkably differ from the phenotypes of the homologous deletions in *S. cerevisiae*. Mycelial colonies of Ag*dig1/2*Δ showed an altered surface, grew 30% slower compared to wild type and were defective in sporulation. Cells deleted for AgTec1 showed no changed phenotype when the mating cascade was not induced. Hyphae of Agtec1Δ showed increased nuclear distances and grew invasively on agar plates after ten days. Invasive growth was observed in wild type *A. gossypii* under glucose limiting conditions (0.2% glucose). Increased nuclear distances have been observed in *A. gossypii* in response to nutrient deprivation. The response, which involved phosphorylation of Agcdc28Y18F, was reported to be independent of the type of nutrient that was restricted and was observed when either carbon (AFM with 0.1% glucose) or nitrogen (ASD-Asn or with one quarter amino acid concentration) was the limiting resource (Helfer and Gladfelter, 2006).

To test the hypothesis that nutrient availability, in particular glucose availability, is signaled via a pathway that involves AgTec1 as a transcription factor, we will look in the next chapter into a genome-wide microarray experiment for genes that are regulated by Tec1. This chapter will also contain results from Ag*dig1/2*Δ and Ag*ste12*Δ mRNA samples analyzed by whole-genome microarray hybridization.

6.4 Materials and Methods

Searches for transcription factor binding sites

Sequences of 38 *A. gossypii* inter-ORF regions belonging to genes that were homologous to *S. cerevisiae* target genes of the mating and filamentation pathway were derived from the sequence information deposited in the AGD 3.0 database and scanned for potential transcription factor binding sites. The consensus sequences

PRE (pheromone responsive element; TGAAACR) and TCS (Tec1p binding site; CATTCY) that were used to search the *A. gossypii* inter-ORF regions for potential transcription factor binding sites were known transcription factor binding sites from the yeast *S. cerevisiae* (Kellis et al., 2003). Searches were done with the latest version of the biological sequence analysis package EMBOSS (P. Rice et al., 2000) that was running on a remote UNIX machine. The EMBOSS command-line program fuzznuc (Bleasby, 2000; Williams, 2001) was used in custom-made vi scripts to search for conserved transcription factor binding sites in the 38 inter-ORF sequences. Fuzznuc uses PROSITE style patterns to search nucleotide sequences. Patterns are specifications of a (typically short) length of sequence to be found. They can specify a search for an exact sequence or they can allow various ambiguities, matches to variable lengths of sequence and repeated subsections of the sequence. In the first round we searched both strands for perfect matches to the consensus sequences and in the second round, we allowed one mismatch anywhere in the consensus sequence. Test sequences with the same base composition as the PRE and TCS elements were generated randomly. GeneSpring GX (Agilent) was used to search for conserved potential regulatory sequences upstream of a given set of genes. The search criteria were to find an enriched ($p<0.05$) 5-8 nucleotide sequence within 5-500 bases upstream of each gene. The statistics was done relative to the sequence upstream of other genes (and not relative to the whole genome sequence which would be less accurate). A local nucleotide density correction that calculates the exact frequencies for nucleotides in *A. gossypii* inter-ORF regions is used instead of assuming ¼ for A, T, G, C.

Gene deletions

A. gossypii deletion mutants were made via the one-step PCR-based gene targeting method (Wendland et al., 2000) using pairs of primers for each locus, which contained at least 45-bp homology upstream and downstream of the ORFs (table 6-2).

Primers for construction of deletion cassette	
DIG1/2_N*S1	GTATATATAGTGGCGCGGGGGCGCGGGCGGCGGCGGGCGGCACAGCGatgggtaccactcttgacgac
DIG1/2_N*S2	TAGCACATATTAAACAGTCCTTGTATACACTGGCGGGCGACTGGCTAggggcagggcatgctcatgtag
TEC1_N*S1	GCTGTGAGCAACGCAGCAAAAGAGCGGGATTTGGCGAAAGCGCAGCGatgggtaccactcttgacgac
TEC1_N*S2	GCTCAGTATTTTCTTCTTTAAACAGGATATATTATGCCCACTTCATGTCAggggcagggcatgctcatgtag
STE12_N*S1	GCGGGCGTAGCACAGTGCGAGCAGGGGTTGCGGAGGAGGCCGACAAGatgggtaccactcttgacgac
STE12_N*S2	ATGAGCACTCGTGCCTTGTTCTGCCCGCCTCATTTTCCCGGTATTTTTCAggggcagggcatgctcatgta
Specific verification primers	
DIG1/2_V1*NAT1	ACGGCCATGTTGTCGCACAG
DIG1/2_V4*NAT1	GCATCGGCGTCAGAATCAGC
DIG1/2_I	CGCGGGACCAGATAGACAAG
TEC1_V1*NAT1	GAACTCTCCGACGCAGACCT
TEC1_V4*NAT1	TGAATCAGGGCGGGAGCCAA
TEC1_I	CTTCAGCGCCGACAGAAACG
STE12_V1*NAT1	GGTCTACGACGCCTATATGG
STE12_V4*NAT1	GACGTTCTTCCACCCAAACC
STE12_I	GCGCCGTGGCAAGAAAGAAC
General verification primers	
V2*NAT1	CTTCAGCGCCGACAGAAACG
V3*NAT1	GGTCTACGACGCCTATATGG

Table 6-2 Primers used for gene deletion and verification. The upper case sequences in the primers for construction of deletion cassettes are homologous to regions upstream or downstream of the *A. gossypii* ORFs. The lower case sequences bind to the plasmid pAG25 (clonat) used for PCR amplification of the deletion cassettes.

A. gossypii laboratory strain ΔlΔt cells were transformed with PCR products amplified off the pAG25 template (containing the clonat resistance cassette) with the respective N*S1/N*S2 oligonucleotide pairs. The plasmid pAG25 had been linearized with BamHI/EcoRI and NcoI/EcoRI before amplification of the PCR products to improve the PCR amplification. The primary transformants are heterokaryotic cells because they carry a mixture of wild-type and transformed nuclei; thus, even deletions of essential genes produce viable primary transformants. Transformants were selected on Ashbya full medium (AFM) plates containing 50µg/ml clonat (Werner Bio Agents). Heterokaryotic transformants were verified with two primer pairs detecting the deletion cassette (gene name_V1*NAT1/V2*NAT1 and V3*NAT1/gene name_V4*NAT1) and one primer pair detecting the wild type gene (gene nameV4*NAT1/I). Three independent transformants were characterized for each mutant. To evaluate phenotypes, the heterokaryon was sporulated to produce uninucleated spores, which were then germinated under selective conditions. The resulting homokaryotic mycelia were again verified to contain the deletion cassette and lack the wild type gene. The primers were the same as the ones used for heterokaryon. Homokaryotic spore preparations were made as described previously and frozen at -80°C until use after a 1 h incubatation time in 30% glycerol at room temperature. Due to the severe sporulation defect of the *dig1/2*Δ mutant, spore preparations containing exclusively transformed spores can not be obtained. Spore preparations obtained from heterokaryotic mycelium and homokaryotic mycelium itself were frozen at -80°C.

Radial growth speed

5 x 10^4 spores, prepared from homokaryotic mycelium, were germinated and pre-grown for 3 days on selective Ashbya full medium (AFM) agar plates at 30°C. The sporulation deficient strain *dig1/2*Δ was pre-grown starting from homokaryotic mycelium. For each deletion mutant, a round piece of mycelium of 2 mm diameter from the very border of the colony was transferred to the center of a new non-selective plate. The pre-incubation step was used to make sure that the starting mycelium was growing with maximal speed. Radial colony speed was determined on AFM plates by measuring every 24 hours the average diameter of the colony at four positions. The growth speed that was reported for each deletion mutant is the average for three independent transformants. All deletion strains and the wild type reached a constant growth speed after 3 days, hence we compared growth speeds between day 3 and day 7.

Culture conditions for microscopy

3.2 x 10^6 spores were inoculated in 500 ml baffle flasks (Duran, Schott, Germany) in 200 ml Ashbya full medium on a shaker at 150 rpm and 30°C for 18 hours. The AFM composition was as follows: 10g/l Bacto Peptone (Pancreatic Digest of Casein, Difco), 10g/l Yeast Extract (Micro Granulated, Formedium, Norwich, England) 1.0 g/l Myo-inositol (Merck), 2% D(+)-Glucose-Monohydrate (Merck). Glucose was autoclaved separately and Myo-inositol was filter sterilized and added to the medium prior to use.

Microscopy

Mycelia were collected in 15 ml Falcon tubes by 5 min centrifugation in a Heraeus Minifuge at 3000 rpm.
Nuclei were stained with Hoechst dye. The cells were fixed for 1 h in 3.7% formaldehyde in Ashbya full medium at 30°C, washed once with 1x PBS, and resuspended in a 5 µg/ml solution of Hoechst dye. The mycelia were incubated with dye for 30 min at room

temperature and washed three times in 1x PBS (phosphate buffered saline).
The actin cytoskeleton (pictures not shown) was visualized by Alexa Fluor Phalloidin 488 as described in (Knechtle *et al.*, 2003) with modifications to improve the yield as described in (Helfer, PhD thesis, 2006). The cells were incubated for 20 min in 4% p-formaldehyde, washed twice with PBS and resuspended in 100 ml mounting medium (Vectashield, Vector Laboratories, Burlingame, CA) and then mounted on slides.
Chitin rings (pictures not shown) were visualized with Calcofluor white. We added 20 µl of a 1 mg/ml Calcofluor solution to 100 ml of liquid culture. After 5 min incubation time at room temperature, the mycelia were washed twice with H_20 and mounted on glass slides for subsequent microscopic inspection (Hanspeter Helfer, personal communication).
The microscopy setup was the same as previously described in (Hoepfner *et al.*, 2000; Knechtle *et al.*, 2003; Gladfelter *et al.*, 2006). The filter sets (Carl Zeiss) were specific for each fluorophore: #02 for Hoechst and Calcofluor; #10 for Alexa 488 and #15 for rhodamine 568. 20-50 planes with distances between 0.3 and 0.5 µm in the Z-axis were taken using the MetaMorph 4.1.7 imaging software (Universal Imaging Corp). In order to flatten the acquired z-stacks into a single plane, we used the stack arithmetic “maximize” command. Outlines of the cells were obtained by doing phase-contrast Z-series. The stacks were passed through the Metamorph filter “detect edges: Laplace 2” and flattened into a single plane using the stack arithmetic “sum” command. Fluorescent and phase-contrast images were scaled and overlayed using “color align”. Images were resized and cropped in Adobe Photoshop CSII and Adobe Illustrator 10.0.

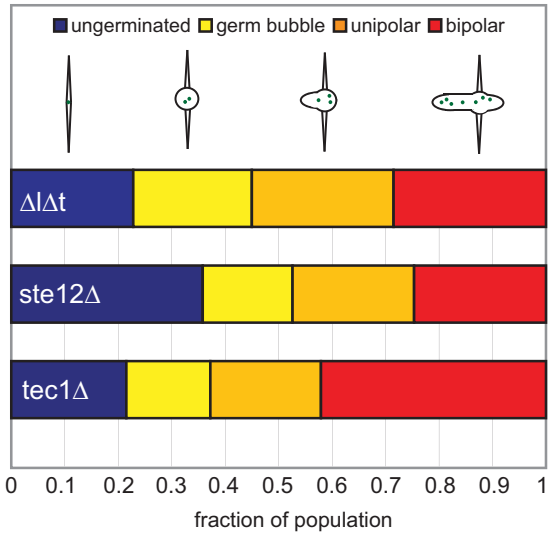
Chapter 7: Transcription in hyphae lacking Dig1/2, Tec1 and Ste12

Chapter 7: Transcription in hyphae lacking Dig1/2, Tec1 and Ste12

7.1 Introduction

There are several developmental stages in the *A. gossypii* life cycle that could be chosen for comparison of the deletion mutants with wild type on microarrays. We decided to collect samples from the Ag*tec1*Δ and the Ag*ste12*Δ deletion mutants nine hours after spore inoculation in full liquid medium (AFM). This stage is characterized by young germlings with over 80% of all nuclei being in a polar growth stage (see chapter 2 figure 2-2 B). In addition the seven to eight nuclei present in a small bipolar germling very likely have the same pattern of gene expression, which is probably not the case at advanced mycelial growth stages. After 9 hours the developmental stages in germinating cultures of Ag*tec1*Δ, Ag*ste12*Δ and ΔlΔt were similar as presented in figure 7-1. The small differences in cell type composition were comparable to differences observed for replicate samples (not shown). Since the correlation coefficients for replicate microarrays were close to 1.0 (Chapter 1 figures 1-11 E and F) we can be sure that the small differences in developmental stages of Ag*ste12*Δ, Ag*tec1*Δ and ΔlΔt do not to provoke variations in the data sets.
The *dig1/2*Δ mutant could not be assayed in a similar manner because the homokaryon hyphae, which carry only transformed nuclei, did not sporulate. Primary transformants are always heterokaryotic and able to produce a mixture of transformed and non-transformed spores. Germinating a mixture of transformed and non-transformed spores under selective conditions would not have been a good option knowing that ungerminated wild type spores contribute to the pool of total mRNAs. Furthermore, it would have been difficult to guarantee reproducibility of the experiment because the fraction of transformed and non-transformed spores in the mixture can not be controlled. For this reason we grew homokaryotic *dig1/2*Δ mycelium on plates to

A Agste12Δ and Agtec1Δ



B Agdig1/2Δ

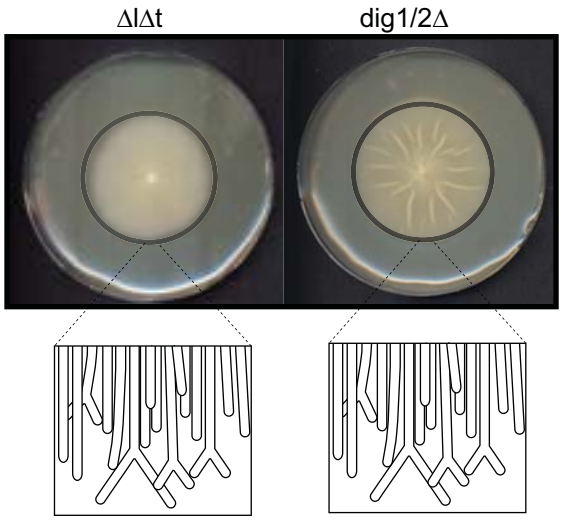


Figure 7-1 Comparison of developmental stages used for transcription profiling. (A) Ag*ste12*Δ and Ag*tec1*Δ hyphae were assayed in liquid medium (B) Ag*dig1/2*Δ hyphae were assayed on solid medium. Hyphae from the outer zone (OZ, black rings) of two mycelial colonies were pooled per preparation.

a diameter of 45 mm (see figure 7-1) and compared the outermost 2 mm of the colony with mycelium from the ΔlΔt strain obtained in the same way as decribed in chapter 4.

7.2 Results

Under non-limiting nutrient conditions we found that of the three deletions, Ag*tec1*Δ had the strongest effect on global gene expression with 221 up-regulated and 1349 down-regulated genes (>2-fold change). This is also reflected in the lowest correlation coefficient (0.89) for comparison of the Ag*tec1*Δ arrays with the ΔlΔt arrays (chapter 1 figure 1-11 F). Corresponding correlation coefficients for Ag*dig*1/2Δ and Ag*ste12*Δ were 0.98 and 0.98, respectively. Surprisingly, the strong phenotype notwithstanding, the Ag*dig*1/2Δ deletion resulted in differential expression of few genes. 172 genes were up-regulated and 47 genes were down-regulated. Deletion of Ag*ste12*Δ resulted in up-regulation of 247 genes and down-regulation of 31 genes.

One of the key challenges in the analysis of gene expression data is how to relate the expression level of individual genes to the underlying transcriptional programs and cellular state. We systematically analyzed the data from the six arrays: the top 15 up-regulated and the 15 top down-regulated genes in each deletion are shown in tables 7-1 to 7-3. The expression levels (EXP), the log2 ratios, the description of the first yeast homolog and the Gene Ontologies “Molecular function”, “Biological process” and “Cellular component” are also shown. In the discussions, the functional descriptions of *A. gossypii* genes were always inferred from the SGD descriptions of the first *S. cerevisiae* homologs, if not otherwise stated. The 30 most differentially expressed genes in Ag*dig1/*2Δ, Ag*tec1*Δ and Ag*ste12*Δ will be discussed group-wise and in detail. A summary of the up- and down-regulated gene groups is given in table 7-4.
The 30 most differentially expressed genes are displayed again in scatter plots where the expression values from the mutant were plotted against the values from wild type (figures 7-2 to 7-4). Genes that were not expressed are shown as green dots (raw value < 60 in deletion mutant and wild type). Genes that were differentially expressed are shown as red dots. These genes changed at least 2-fold in response to gene deletion and passed a t-test with a p-value ≤ 0.05 (unless otherwise stated). Genome-wide expression data for all deletions and wild type is available in electronic version as supplementary material 2.

To support and/or extend the single gene analysis we scored the activity of predefined groups of genes with t-profiler (Boorsma *et al.*, 2005). T-profiler is a tool that uses the t-test to score changes in the average activity of predefined groups of genes (GO categories) without the need to choose parameters. Whenever useful, we will include results from the t-profiler analysis. T-profiler tables are available at full length as supplemental material 3.

Differential gene expression in cells lacking Dig1/2

Up-regulated genes in Dig1/2Δ

Genes homologous to mating genes.
The *S. cerevisiae* homolog of **AgFUS3** codes for a mitogen-activated protein kinase involved in mating response. It is activated by phosphorylation by Ste7p and its substrates include Ste12p, Far1p, Bni1p, Sst2p. ScFUS3 inhibits invasive growth during mating by phosphorylating ScTec1p which promotes its degradation.
The other FUS gene that was induced upon DIG1/2 deletion was **AgFUS1.** The yeast homolog ScFUS1 encodes a membrane

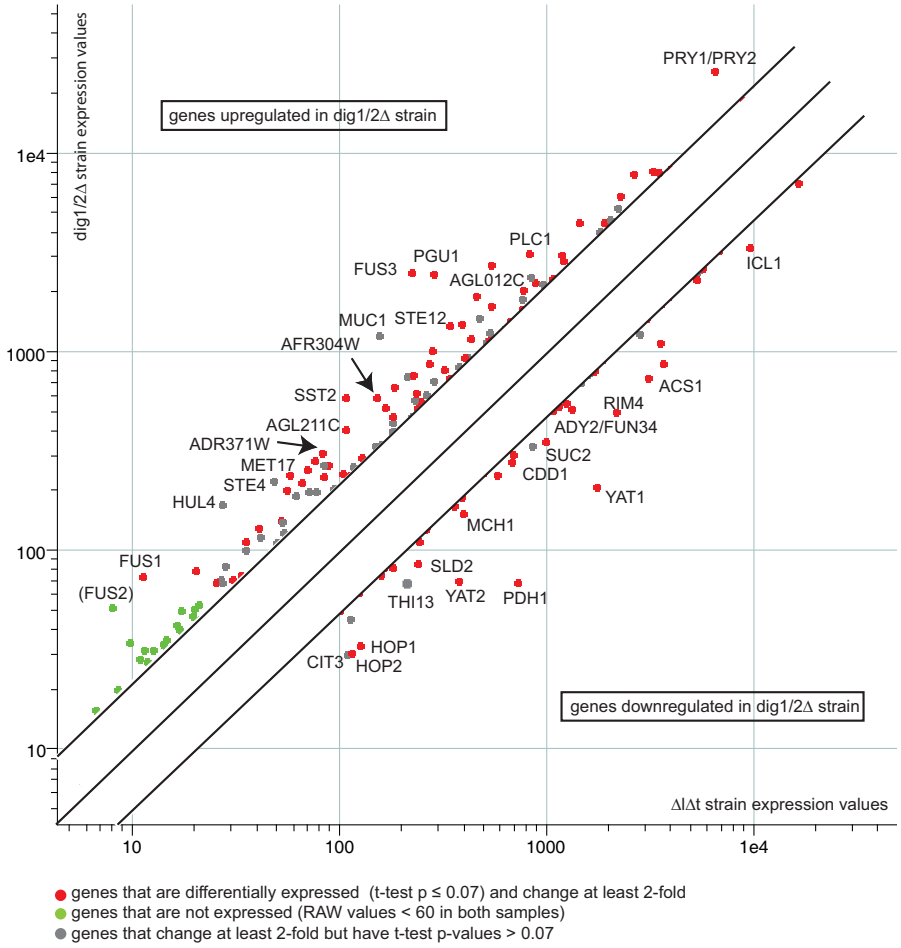


Figure 7-2 Scatter plot showing the top 15 up- and down-regulated genes

protein that localizes to the shmoo tip. It is required for cell fusion. The expression of ScFUS1 is regulated by mating pheromone. ScFUS1 is proposed to coordinate signaling, fusion, and polarization events that are required for fusion. It is thought to be a Cdc28p substrate.

The *S. cerevisiae* homolog of **AgSST2**, ScSST2, codes for a GTPase-activating protein for Gpa1p. It regulates desensitization to alpha factor pheromone and is also required to prevent receptor-independent signaling of the mating pathway. It is a member of the RGS (regulator of G-protein signaling) family. The **AgSTE4** homolog, ScSTE4, encodes the G protein beta subunit. It forms a dimer with Ste18p to activate the mating signaling pathway and forms a heterotrimer with Gpa1p and Ste18p to dampen signalling. It may recruit Rho1p to the polarized growth site during mating. We found that AgSTE18 and AgGPA1 were also up-regulated, albeit with small log2 ratios of 0.4 for AgSTE18 and 0.3 for AgGPA1.

The *S. cerevisiae* homolog of **AgSTE12** encodes a transcription factor that is activated by a MAP kinase signaling cascade. It can activate both, genes involved in mating or pseudohyphal/invasive growth pathways. It has been shown to cooperate with the ScTec1p transcription factor to regulate genes specific for invasive growth.

Genes homologous to filamentation genes. The strongest up-regulated gene upon DIG1/2 deletion was **AgMUC1**, also called AgFLO11. ScMUC1 encodes a GPI-anchored cell surface glycoprotein that is required for diploid pseudohyphal formation and haploid invasive growth. ScMUC1 is

transcriptionally regulated by the MAPK pathway (via Ste12p and Tec1p) and the cAMP pathway (via Flo8p).

The *S. cerevisiae* homolog of **AgPGU1** codes for a endo-polygalacturonase which is a pectolytic enzyme that hydrolyzes the alpha-1,4-glycosidic bonds in the rhamnogalacturonan chains in pectins. AgPGU1 is among the top 15 highly up-regulated genes in response to AgDIG1/2 deletion. It was found to be among the top15 down-regulated in Ag*tec1*Δ and Ag*ste12*Δ deletion strains. Strong down-regulation of AgPGU1 mRNA that yields a close-to-background signal intensity on the array, as seen in Ag*tec1*Δ, does not have an effect on cell viability or growth speed.

The homolog of **AgPLC1**, ScPLC1 codes for a phosphoinositide-specific phospholipase C that hydrolyzes phosphatidylinositol 4,5-biphosphate (PIP2) to generate inositol 1,4,5-triphosphate (IP3) and 1,2-diacylglycerol (DAG). It is involved in kinetochore function and pseudohyphal differentiation.

Genes of unknown function. Four genes with proteins of unknown function were found amongst the fifteen most highly up-regulated.
The *S. cerevisiae* homologs of **AGL012C** (YPL014W) and **ADR371W** (YOR342C and YAL037W) localizes to the nucleus and the cytoplasm as determined by GFP-fusions.

The *S. cerevisiae* homolog of **AFR304W,** YGR125W, encodes a putative protein of unknown function. The deletion mutant has decreased rapamycin resistance and the green fluorescent protein (GFP)-fusion protein localizes to the vacuole.

AgPRY1/PRY2, the homolog of YJL079C, YJL078C and YKR013W, encodes a protein of unknown function with similarity to the plant PR-1 class of pathogen related proteins. Interestingly, PRY1/2 was amongst the fifteen most strongly down-regulated genes in the STE12 deletion.

Down-regulated genes in Dig1/2Δ

Glyoxylate shunt and ethanol inducible genes.
The most strongly down-regulated gene was **AgPDH1**, the yeast homolog of which encodes a mitochondrial protein that partipates in repiration and is induced by diauxic shift.

The second and third most strongly up-regulated genes were **AgYAT1** and **AgYAT2**, two genes with homology to carnitine acetyltransferases that are ethanol inducible. ScYAT1 in involved in the transport of activated acyl groups from the cytoplasm into the mitochondrial matrix.

There were additional genes among the fifteen most strongly down-regulated genes that allow a life on C2-sources, like ethanol: **AgCIT3** (citrate synthase), **AgACS1** (acetyl-CoA synthetase) and **AgICL1** (isocytrate lyase).
ScCIT3 encodes a dual specificity mitochondrial citrate and methylcitrate synthase that catalyzes the condensation of acetyl-CoA and oxaloacetate to form citrate and that of propionyl-CoA and oxaloacetate to form 2-methylocitrate.
The *S. cerevisiae* homolog of **AgACS1** encodes a acetyl-CoA synthetase which, along with ScACS2, is the nuclear source of acetyl-CoA for histone acetylation. ScACS1 is expressed during growth on nonfermentable carbon sources and under aerobic conditions.
The *S. cerevisiae* homolog of **AgICL1**, ScICL1, is induced by growth on ethanol and encodes the isocitrate lyase which catalyzes the formation of succinate and glyoxylate from citrate, the key reaction in the glyoxylate cycle. The glyoxlate cycle including the genes that are required for a life on C2-sources is depicted in figure 3-4 B.
We found that all genes in the glyoxlate cycle were down-regulated upon DIG1/2 deletion with the following log2 ratios: **AgADH2**

TABLE 7-1. TOP DIFFERENTIALLY EXPRESSED GENES IN CELLS LACKING DIG1/2

UPREGULATED

Genes expressed at higher levels in cells lacking DIG1/2

SYSTEMATIC NAME	COMMON NAME	EXP $_{\Delta/\Delta t}$	EXP $_{dig1/2\Delta}$	LOG2 (RATIO)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGIGAL PROCESS	CELLULAR COMPONENT
AEL023C	MUC1(=FLO11)	103.01	1366.50	3.73	GPI-anchored cell surface glycoprotein required for diploid pseudohyphal formation and haploid invasive growth, transcriptionally regulated by the MAPK pathway (via Ste12p and Tec1p) and the cAMP pathway (via Flo8p)	pseudohyphal growth	-	plasma membrane
AFR019W	FUS3	219.95	2526.50	3.52	Mitogen-activated protein kinase involved in mating pheromone response; activated by phosphorylation by Ste7p; provides specificity during the mating vs. filamentous growth response by phosphorylating transcriptional and cytoplasmic targets	protein amino acid phosphorylation	MAP kinase activity	cytoplasm
ADR327W	PGU1	281.25	2447.50	3.12	Endo-polygalacturonase, pectolytic enzyme that hydrolyzes the alpha-1,4-glycosidic bonds in the rhamnogalacturonan chains in pectins	pseudohyphal growth	polygalacturonase activity	extracellular region
AFR699C	FUS1	11.04	74.13	2.75	Membrane protein localized to the shmoo tip, required for cell fusion; expression regulated by mating pheromone; proposed to coordinate signaling, fusion, and polarization events required for fusion; potential Cdc28p substrate	conjugation with cellular fusion	-	plasma membrane
AEL068W	HUL4	26.64	170.40	2.68	Protein with similarity to hect domain E3 ubiquitin-protein ligases, not essential for viability	protein monoubiquitination	ubiquitin-protein ligase activity	-
ABR176C	SST2	106.35	591.65	2.48	GTPase-activating protein for Gpa1p, regulates desensitization to alpha factor pheromone; also required to prevent receptor-independent signaling of the mating pathway; member of the RGS (regulator of G-protein signaling) family	signal transduction	GTPase activator activity	plasma membrane
ACR097W	STE4	47.52	223.65	2.23	G protein beta subunit, forms a dimer with Ste18p to activate the mating signaling pathway, forms a heterotrimer with Gpa1p and Ste18p to dampen signaling; may recruit Rho1p to the polarized growth site during mating; contains WD40 repeats	invasive growth (sensu Saccharomyces)	signal transducer activity	plasma membrane
AGL012C		454.15	1913.50	2.07	Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the cytoplasm and to the nucleus			
ADL031W	MET17	57.34	241.05	2.07	O-acetyl homoserine-O-acetyl serine sulfhydrylase, required for sulfur amino acid synthesis	response to drug	cysteine synthase activity	cytoplasm
ADR304W	STE12	336.90	1368.00	2.02	Transcription factor that is activated by a MAP kinase signaling cascade, activates genes involved in mating or pseudohyphal/invasive growth pathways; cooperates with Tec1p transcription factor to regulate genes specific for invasive growth	pseudohyphal growth	transcription factor activity	nucleus
AAL179W	PRY1/PRY2	6472.5	25793.00	1.99	Protein of unknown function, has similarity to Pry2p and Pry3p and to the plant PR-1 class of pathogen related proteins	-	-	endoplasmic reticulum
AFR304W		150.6	589.40	1.97	Putative protein of unknown function; deletion mutant has decreased rapamycin resistance but normal wormannin resistance; green fluorescent protein (GFP)-fusion protein localizes to the vacuole	-	-	vacuole, cell cycle-correlated morphology
AGL211C		105.325	407.15	1.95	Zinc finger protein; putative transcription factor that may interact with proteins involved in histone acetylation or deacetylation; may be involved in altering acetylation on histone lysines	-	-	cytoplasm
ACR024W	PLC1	824.55	3127.00	1.92	Phosphoinositide-specific phospholipase C, hydrolyzes phosphatidylinositol 4,5-biphosphate (PIP2) to generate inositol 1,4,5-triphosphate (IP3) and 1,2-diacylglycerol (DAG); involved in kinetochore function and pseudohyphal differentiation	pseudohyphal growth	phosphoinositide phospholipase C activity	nucleus
ADR371W		81.45	307.05	1.91	Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the cytoplasm and the nucleus	-	-	cytoplasm

DOWNREGULATED

Genes expressed at lower levels in cells lacking DIG1/2

SYSTEMATIC NAME	COMMON NAME	EXP $_{\Delta/\Delta t}$	EXP $_{dig1/2\Delta}$	LOG2 (RATIO)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGIGAL PROCESS	CELLULAR COMPONENT
AGR003W	PDH1	723.55	68.42	-3.40	Mitochondrial protein that participates in respiration, induced by diauxic shift; homologous to E. coli PrpD, may take part in the conversion of 2-methylcitrate to 2-methylisocitrate	propionate metabolic process	-	cytoplasm
AER224W	YAT1	1729.50	207.05	-3.06	Outer mitochondrial carnitine acetyltransferase, minor ethanol-inducible enzyme involved in transport of activated acyl groups from the cytoplasm into the mitochondrial matrix	alcohol metabolic process	carnitine O-acetyltransferase activity	mitochondrion
ABR172W	YAT2	375.10	69.66	-2.43	Carnitine acetyltransferase; has similarity to Yat1p, which is a carnitine acetyltransferase associated with the mitochondrial outer membrane	alcohol metabolic process	carnitine O-acetyltransferase activity	cytoplasm
AGL038C	RIM4	2144.00	501.50	-2.10	Putative RNA-binding protein required for the expression of early and middle sporulation genes	meiosis	RNA binding	cytoplasm
ADR408W	ACS1	3104.00	736.55	-2.08	Acetyl-coA synthetase isoform which, along with Acs2p, is the nuclear source of acetyl-coA for histone acetlyation; expressed during growth on nonfermentable carbon sources and under aerobic conditions	histone acetylation	acetate-CoA ligase activity	mitochondrion
ACL029W	HOP1	125.20	32.96	-1.93	Meiosis-specific DNA binding protein that displays Red1p dependent localization to the unsynapsed axial-lateral elements of the synaptonemal complex; required for homologous chromosome synapsis and chiasma formation	synapsis	four-way junction DNA binding	condensed nuclear chromosome
AFR474W	HOP2	112.55	30.67	-1.88	Meiosis-specific protein that localizes to chromosomes, preventing synapsis between nonhomologous chromosomes and ensuring synapsis between homologs; complexes with Mnd1p to promote homolog pairing and meiotic double-strand break repair	meiotic recombination	protein binding	condensed nuclear chromosome
AGR002W	CIT3	107.70	30.13	-1.84	Dual specificity mitochondrial citrate and methylcitrate synthase; catalyzes the condensation of acetyl-CoA and oxaloacetate to form citrate and that of propionyl-CoA and oxaloacetate to form 2-methylcitrate	tricarboxylic acid cycle	citrate (Si)-synthase activity	mitochondrion
ADL066C	ICL1	9608.50	3352.50	-1.52	Isocitrate lyase, catalyzes the formation of succinate and glyoxylate from isocitrate, a key reaction of the glyoxylate cycle; expression of ICL1 is induced by growth on ethanol and repressed by growth on glucose	glyoxylate cycle	isocitrate lyase activity	-
AER451W	THI13	201.90	72.91	-1.47	Protein involved in synthesis of the thiamine precursor hydroxymethylpyrimidine (HMP); member of a subtelomeric gene family including THI5, THI11, THI12, and THI13	thiamin biosynthetic process	-	-
ABL045W	SLD2	238.35	86.15	-1.47	Protein required for DNA replication, phosphorylated in S phase by S-phase cyclin-dependent kinases (Cdks), phosphorylation is essential for DNA replication and for complex formation with Dpb11p; potential Cdc28p substrate	DNA strand elongation during DNA replication	-	replication fork
AFR529W	SUC2	976.40	356.25	-1.45	Invertase, sucrose hydrolyzing enzyme; a secreted, glycosylated form is regulated by glucose repression, and an intracellular, nonglycosylated enzyme is produced constitutively	sucrose catabolic process	beta-fructofuranosidase activity	cytoplasm
AER192W	ADY2/FUN34	1330.5	511.65	-1.38	Acetate transporter required for normal sporulation	meiosis	ammonium transmembrane transporter activity	mitochondrion
AFR595W	MCH1	389.90	154.20	-1.34	Protein with similarity to mammalian monocarboxylate permeases, which are involved in transport of monocarboxylic acids across the plasma membrane; mutant is not deficient in monocarboxylate transport	transport	transporter activity	membrane
AGL123W	CDD1	846.35	336.70	-1.33	Cytidine deaminase; catalyzes the modification of cytidine to uridine in vitro but native RNA substrates have not been identified, localizes to both the nucleus and cytoplasm	pyrimidine salvage	cytidine deaminase activity	cytoplasm

(-0.48; alcohol dehydrogenase 2), **AgACS1** (-2.1; acetyl-CoAsynthetase), **AgCIT1**(-0.24; citrate synthase), **AgICL1**(-1.5; isocytrate lyase), **AgMLS1**(-1.2; malate synthase), **AgPCK1**(-0.7; phosphoenolpyruvate carboxykinase) and **AgFBP1**(-1.2; fructose-1,6-bisphosphatase).

Ribosomal protein genes. The t-profiler analysis revealed that the categories "large ribosomal subunit" (t = -8.9, E = 1.0E-15) and "small ribosomal subunit" (t = -5.72, E = 1.4E-05) were significantly down-regulated in Ag*dig1/2*Δ cells. This makes sense. In a growing cell, a large portion of total transcription and also a large portion total cellular energy are dedicated to ribosome biogenesis. Under slow growth conditions, e.g. in response to nutrient limitation, ribosomal proteins need to be down-regulated.

Sporulation and meiosis genes. A group of four genes strongly down-regulated genes, AgRIM4, AgHOP1, AgHOP2 and AgADY2/FUN34, is involved in meiosis and sporulation.
This is consistent with the observed sporulation-defect of Ag*dig1/2*Δ. Here, we compared the expression profiles of not-yet-sporulating hyphae growing at the border of a expanding mycelial colony. We would expect to find even more down-regulated sporulation genes in a comparison of the expression profiles from hyphae in the center of the colony.

The homolog of **AgRIM4**, ScRIM4, codes for a putative mRNA binding protein required for the expression of early and middle sporulation genes.
AgADY2/FUN34 is a twin gene in *S. cerevisiae*. ScADY2 codes for an acetate transporter that is required for sporulation. ScFUN34 encodes a putative transmembrane protein that is involved in the export of ammonia.
Other genes. ScTHI13 encodes a protein involved in the synthesis of a thiamin precursor. **AgTHI13** was strongly up-regulated in AgTEC1 deletion.

The *S. cerevisiae* homolog of **AgSLD2** encodes a protein that is essential for DNA replication.

The homolog of **AgSUC2**, ScSUC2, codes for an invertase which is a sucrose hydrolyzing enzyme. There are two forms of Suc2 in *S. cerevisiae*: the first form is glycosylated, secreted and glucose-repressed and the second form is intracellular, non-glycosylated and produced constitutively.

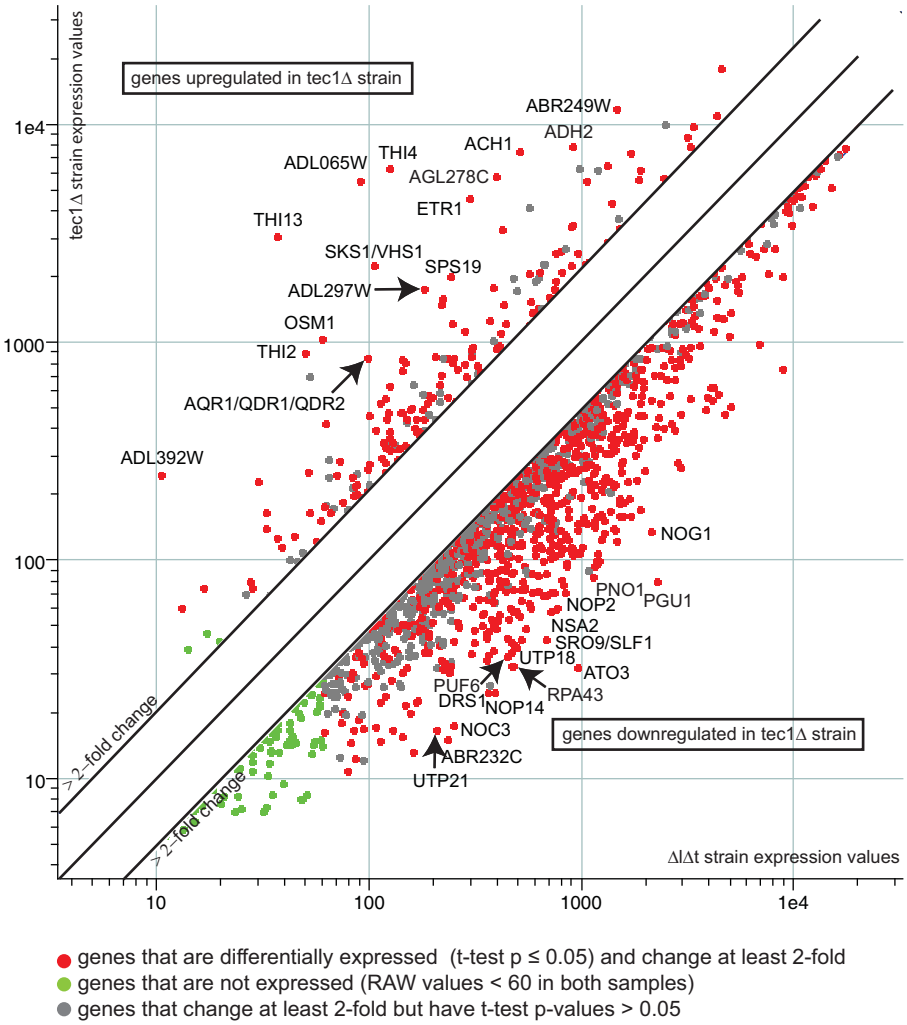
ScMCH1, homolog of **AgMCH1**, encodes a protein with similarity to mammalian monocarboxlate permeases which are involved in the transport of monocarboxylic acids across the plasma membrane. ScCDD1 encodes a cytidin deaminase which catalyzes the modification of citidine to uridine.

Differential gene expression in cells lacking Tec1

Up-regulated genes in Ag*tec1*Δ

Thiamin biosynthesis genes. Most genes involved in thiamin (vitamin B1) biosynthesis were not expressed in wild type cells in full medium. **AgTHI13**, **AgTHI4** and **AgTHI2** were 85-fold, 62-fold and 18-fold up-regulated in cells lacking AgTEC1. AgTHI13 and AgTHI4 were the two top up-regulated genes upon Tec1 deletion. We also found AgTHI6, AgTHI7 and AgTHI20/THI21 expressed at higher levels in Ag*tec1*Δ hyphae. AgTHI1 and AgTHI10 were not represented on the array. The t-profiler analysis supported the findings from the single gene analysis. The GO category "thiamin biosynthesis" was significantly up-regulated (t-value 6.06, E-value 1.9E-6).

Interactions between Tec1 and thiamin biosynthesis genes are poorly documented in



● genes that are differentially expressed (t-test p ≤ 0.05) and change at least 2-fold
● genes that are not expressed (RAW values < 60 in both samples)
● genes that change at least 2-fold but have t-test p-values > 0.05

Figure 7-3 Scatter plot showing the top 15 up- and down-regulated genes

the literature. ScTec1 has not been reported to be a transcription factor that directly binds and regulates thiamin biosynthesis genes.

We searched *in silico* for enrichment of the ScTec1p binding site (CATTCY) upstream of six *A. gossypii* thiamine biosynthesis genes. Only two genes, AgTHI6 and AgTHI13, had a potential Tec1p binding site, which was statistically not a significant overrepresentation. Even though a direct interaction has not been documented, there are a couple of publications that observed an interaction between thiamin or thiamin gene expression and glucose levels in the medium (see next section).
ScTHI4 was shown to be differentially expressed in yeast when glucose concentration was increased (Kaeberlein *et al.*, 2002). In S. pombe the intracellular pool of thiamin was decreased when cells were starved for glucose (Schweingruber *et al.*, 1991). And, in microarray experiments, ScTHI13 was up-regulated during the diauxic shift (DeRisi *et al.*, 1997).

Genes involved in utilization of alternative carbon sources. The second strongest up-regulated gene was **ADL065W**, a gene homologous to the twin *S. cerevisiae* genes YIL057C and YER067W whose functions are not known.

The expression of the *S. cerevisiae* gene YIL057C is induced under carbon limitation. Among the top 15 up-regulated genes was

a second gene which is glucose repressed in yeast. **AgADH2** encodes the alcohol dehydrogenase II enzyme that catalyzes the oxidation of ethanol to acetaldehyde. We also found that the potential activator of AgADH2, AgADR1, was transcriptionally induced in *A. gossypii* upon TEC1 deletion. Ethanol, a C2-compound, accumulates during fermentation when yeast is growing on glucose. As glucose levels decrease ScADH2 becomes derepressed allowing yeasts to consume the ethanol they have produced. Hence, ethanol is an alternative carbon source. Amongst the top 15 up-regulated genes we found two other genes involved in the utilization of alternative carbon sources. The first one, **AgACH1**, shares homology with ScACH1 which encodes an acetyl-CoA hydrolase that is primarily localized to mitochondria and is required for acetate utilization. And the second one, **AgSPS19**, is homologous to ScSPS19 which encodes a peroxisomal 2.4-dienoyl-CoA reductase. This is an auxiliary enzyme of fatty acid beta-oxidation, that is transcriptionally induced in the presence of oleate.

The observation that glucose repressed genes and genes involved in the utilization of alternative carbon sources were up-regulated lead us to an investigation of the glucose transporters. Genes involved in glucose uptake are up-regulated in *S. cerevisiae* during glucose starvation (Wu *et al.*, 2004). The presumptive low affinity glucose transporter AgHXT6/HXT7, whose *S. cerevisiae* twin genes ScHXT6 and ScHXT7 are highly up-regulated under glucose starvation (Wu *et al.*, 2004), was not expressed in *A. gossypii*. But this transporter was not even expressed under conditions where glucose limitation was very likely, for instance in sporulating *A. gossypii* mycelium.
The glucose transporters ScHXK1 and ScHXT2 are twin genes in *S. cerevisiae*. They are involved in glucose uptake and in the first irreversible step in glucose metabolism, the conversion of glucose to glucose-6-phophate. ScHXK1 is repressed when glucose is limiting and ScHXT2 is induced by low levels of glucose. The *A. gossypii* homolog AgHXT1/2 was 3-fold up-regulated upon Tec1 deletion.
A third gene which is involved in glucose uptake in yeast, the glucokinase ScGLK1, has been shown to be rapidly de-repressed on nonfermentable carbon sources. We found that the expression AgGLK1 gene was 9-fold up-regulated in absence of Tec1.

Genes with unknown function. The *S. cerevisiae* homologs of **ADL392W** (YHR202W), AGL278C (YBL095W), ADL297W (YBR085C-A) and ABR249W (YIR035C and YIR036C) have unknown functions. YHR202W and YBR085C-A were found in the cytoplasm and YBL095W C in the mitochondria.

Other genes. Two homologs for **AgVHS/SKS1** were found in *S. cerevisiae* (twin). ScVHS1 codes for a cytoplasmic serine/threonine protein kinase that has a role in G1/S phase progression. ScSKS1 encodes also a putative serine/threonine protein kinase which is involved in the adaptation to low concentrations of glucose independent of the SNF3 regulated pathway.

ScOSM1, the homolog of **AgOSM1**, codes for a fumarate reductase, that catalyzes the reduction of fumarate to succinate. It is required for the reoxidation of intracellular NADH under anaerobic conditions. Mutations in the ScOSM1 gene cause osmotic sensitivity.

AgETR1 potentially codes for a 2-enoyl thioester reductase which is a member of the medium chain dehydrogenase/reductase family. The enzyme is localized in the mitochondria, where it has a probable role in fatty acid synthesis. The expression level of AgETR1 was also highly up-regulated in Ag*ste12*Δ.

The *S. cerevisiae* homolog of **AgAQR1**, ScAQR1, encodes a plasma membrane

TABLE 7-2. TOP DIFFERENTIALLY EXPRESSED GENES IN CELLS LACKING TEC1

	SYSTEMATIC NAME	COMMON NAME	EXP Δ1Δt	EXP tec1Δ	LOG2 (RATIO)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGIGAL PROCESS	CELLULAR COMPONENT
Genes expressed at higher levels in cells lacking TEC1									
UPREGULATED	AER451W	THI13	38.47	3258.00	6.40	Protein involved in synthesis of the thiamine precursor hydroxymethylpyrimidine (HMP); member of a subtelomeric gene family including THI5, THI11, THI12, and THI13	thiamin biosynthetic process	-	-
	ADL065W		89.54	5519.00	5.95	Putative protein of unknown function; expression induced under carbon limitation and repressed under high glucose	-	-	cytoplasm
	AAL137W	THI4	124.05	6282.00	5.66	thiazole synthase, catalyzes formation of the thiazole moiety of thiamin pyrophosphate; required for thiamine biosynthesis and for mitochondrial genome stability	mitochondrial genome maintenance	-	mitochondrion
	ADL392W		10.34	243.90	4.56	Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the vacuole, while HA-tagged protein is found in the soluble fraction, suggesting cytoplasmic localization	-	-	vacuole, cell cycle-correlated morphology
	ADR163W	VHS1/SKS1	103.25	2239.00	4.44	Cytoplasmic serine/threonine protein kinase; identified as a high-copy suppressor of the synthetic lethality of a sis2 sit4 double mutant, suggesting a role in G1/S phase progression; homolog of Sks1p	protein amino acid phosphorylation*	protein kinase activity	cytoplasm
	AFL033W	THI2	49.63	889.90	4.16	Zinc finger protein of the Zn(II)2Cys6 type, probable transcriptional activator of thiamine biosynthetic genes	positive regulation of transcription from RNA polymerase II promoter	transcriptional activator activity	nucleus
	AFR367W	OSM1	59.05	1030.65	4.13	Fumarate reductase, catalyzes the reduction of fumarate to succinate, required for the reoxidation of intracellular NADH under anaerobic conditions; mutations cause osmotic sensitivity	metabolic process	fumarate reductase (NADH) activity	mitochondrion
	AEL081W	ETR1	293.30	4572.50	3.96	2-enoyl thioester reductase, member of the medium chain dehydrogenase/reductase family; localized to in mitochondria, where it has a probable role in fatty acid synthesis	aerobic respiration	enoyl-[acyl-carrier-protein] reductase activity	mitochondrion
	AFR020W	ACH1	504.15	7513.50	3.90	Acetyl-coA hydrolase, primarily localized to mitochondria; required for acetate utilization and for diploid pseudohyphal growth	acetate metabolic process	acetyl-CoA hydrolase activity	mitochondrion
	AGL278C		391.20	5737.00	3.87	Putative protein of unknown function; the authentic, non-tagged protein is detected in highly purified mitochondria in high-throughput studies	-	-	mitochondrion
	ADL297W		179.60	1757.00	3.29	Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the cytoplasm and to the nucleus	-	-	cytoplasm
	AAR084W	ADH2	906.30	7865.00	3.12	Glucose-repressible alcohol dehydrogenase II, catalyzes the conversion of ethanol to acetaldehyde; involved in the production of certain carboxylate esters; regulated by ADR1	NADH oxidation	alcohol dehydrogenase activity	cytoplasm
	ADL258W	AQR1	97.93	846.95	3.11	Multidrug transporter required for resistance to quinidine, barban, cisplatin, and bleomycin; may have a role in potassium uptake; member of the major facilitator superfamily of transporters conferring multiple drug resistance (MFS-MDR)	drug transport	monocarboxylic acid transmembrane transporter activity	plasma membrane
	ABL163W	SPS19	238.15	1995.00	3.07	Peroxisomal 2,4-dienoyl-CoA reductase, auxiliary enzyme of fatty acid beta-oxidation; homodimeric enzyme required for growth and sporulation on petroselineate medium; expression induced during late sporulation and in the presence of oleate	sporulation (sensu Fungi)	2,4-dienoyl-CoA reductase (NADPH) activity	peroxisomal matrix
	ABR249W		1462.00	11668.50	3.00	Putative cytoplasmic protein of unknown function			

	SYSTEMATIC NAME	COMMON NAME	EXP Δ1Δt	EXP tec1Δ	LOG2 (RATIO)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGIGAL PROCESS	CELLULAR COMPONENT
Genes expressed at lower levels in cells lacking TEC1									
DOWNREGULATED	AFR144W	ATO3	945.40	32.05	-4.88	Plasma membrane protein, regulation pattern suggests a possible role in export of ammonia from the cell; member of the TC 9.B.33 YaaH family of putative transporters	nitrogen utilization	ammonium transmembrane transporter activity	plasma membrane
	ADR327W	PGU1	2275.00	80.21	-4.83	Endo-polygalacturonase, pectolytic enzyme that hydrolyzes the alpha-1,4-glycosidic bonds in the rhamnogalacturonan chains in pectins	pseudohyphal growth	polygalacturonase activity	extracellular region
	AGR194W	NOG1	2125.00	135.55	-3.97	Putative GTPase that associates with free 60S ribosomal subunits in the nucleolus and is required for 60S ribosomal subunit biogenesis; constituent of 66S pre-ribosomal particles; member of the ODN family of nucleolar G-proteins	ribosome biogenesis and assembly	GTP binding	nucleolus
	AFR713W	SRO9/SLF1	672.15	42.98	-3.97	RNA binding protein that associates with polysomes; proposed to be involved in regulating mRNA translation; involved in the copper-dependent mineralization of copper sulfide complexes on cell surface in cells cultured in copper salts	translation	RNA binding	polysome
	ADR072C	NOP14	385.10	24.74	-3.96	Nucleolar protein, forms a complex with Noc4p that mediates maturation and nuclear export of 40S ribosomal subunits; also present in the small subunit processome complex, which is required for processing of pre-18S rRNA	ribosome biogenesis and assembly	snoRNA binding	nucleolus
	ABR232C		231.55	15.14	-3.94	Protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to endosomes; YLR073C is not an esssential gene	-	-	endosome
	AEL153W	UTP18	474.55	32.87	-3.85	Possible U3 snoRNP protein involved in maturation of pre-18S rRNA, based on computational analysis of large-scale protein-protein interaction data	ribosome biogenesis and assembly	snoRNA binding	nucleolus
	AAL164C	DRS1	358.55	24.88	-3.85	Nucleolar DEAD-box protein required for ribosome assembly and function, including synthesis of 60S ribosomal subunits; constituent of 66S pre-ribosomal particles	ribosome biogenesis and assembly	ATP-dependent RNA helicase activity	nucleolus
	ADR377W	RPA43	468.55	32.90	-3.83	DNA dependent RNA polymerase I subunit A43	ribosome biogenesis and	DNA-directed RNA polymerase	DNA-directed RNA
	ACR032C	NOC3	245.80	17.57	-3.81	Protein that forms a nuclear complex with Noc2p that binds to 66S ribosomal precursors to mediate their intranuclear transport; also binds to chromatin to promote the association of	ribosome biogenesis and	protein binding	nucleus
	AFR390C	PNO1	1121.00	84.49	-3.73	Essential nucleolar protein required for pre-18S rRNA processing, interacts with Dim1p, an 18S rRNA dimethyltransferase, and also with Nob1p, which is involved in proteasome biogenesis; contains a KH domain	ribosome biogenesis and assembly*	unfolded protein binding	nucleus
	ADL254W	NOP2	745.40	58.97	-3.66	Probable RNA m(5)C methyltransferase, essential for processing and maturation of 27S pre-rRNA and large ribosomal subunit biogenesis; localized to the nucleolus; constituent of 66S pre-ribosomal particles	ribosome biogenesis and assembly	S-adenosylmethionine-dependent methyltransferase activity*	nucleolus
	ACL050W	NSA2	756.60	60.66	-3.64	Protein constituent of 66S pre-ribosomal particles, contributes to processing of the 27S pre-rRNA	ribosome biogenesis and assembly	-	nucleus
	AFR342C	PUF6	445.15	35.95	-3.63	Pumilio-homology domain protein that binds ASH1 mRNA at PUF consensus sequences in the 3' UTR and represses its translation, resulting in proper asymmetric localization of ASH1 mRNA	ribosome biogenesis and assembly	mRNA binding	nucleus
	AER081C	UTP21	205.00	16.69	-3.62	Possible U3 snoRNP protein involved in maturation of pre-18S rRNA, based on computational analysis of large-scale protein-protein interaction data	ribosome biogenesis and assembly	snoRNA binding	nucleus

transporter of the major facilitator superfamily that confers resistance to short-chain monocarboxylic acids and quinidine.

Down-regulated genes in Ag*tec*1Δ

Nitrogen utilization and pseudohyphal growth. AgATO3 was the gene whose expression was strongest down-regulated. The *S. cerevisiae* homolog of this gene is a plasma membrane protein with a possible role in export of ammonia from the cell.

The second most strongly down-regulated gene, **AgPGU1**, presumably encodes an endo-polygalacturonase that hydrolyzes the alpha-1,4-glycosidic bonds in the rhamnogalacturonan chains in pectins based on the function of the yeast homolog ScPGU1. ScPGU1 may also have a function in the process of pseudohyphal growth.

Ribosome biosynthesis and assembly. Eleven of the top 15 down-regulated genes, were involved in ribosome biosynthesis and assembly. The significant down-regulation of the whole group of genes involved in "ribosome biogenesis and assembly" was confirmed by the t-profiler analysis (t = -23.39, p = 1.0E-15). Nine of the previously mentioned eleven genes showed extremely low expression intensity values and were most probably not expressed in the strain deleted for Tec1.

Ribosome biosynthesis and assembly: snoRNA binding. *S. cerevisiae* homologs of the three genes **AgNOP14**, **AgUTP18** and **AgUTP21** were all shown to be involved in snoRNA binding. ScNOP14 is a nucleolar protein that forms a complex with ScNoc4p. ScNoc4p mediates maturation and nuclear export of 40S ribosomal subunits. ScNOP14 is also present in the small subunit processome complex, which is required for processing of pre-18S rRNA. The *S. cerevisiae* homolog of **AgUTP18**, ScUTP18, encodes a U3 snoRNP protein involved in maturation of pre-18S rRNA. The second gene with a homolog in *S. cerevisiae* that encodes a U3 snoRNP protein is **AgUTP21**.

Ribosome biosynthesis and assembly: processing/maturation of the pre-18S or the 27 pre-rRNA.
The *S. cerevisiae* homologs of a group of six genes, AgNOP14, AgUTP18, AgPNO1, AgNOP2, AgNSA2 and AgUTP21, is involved in either processing/maturation of the pre-18S rRNA or the 27 pre-rRNA. **AgNOP14** was already described as a putative snoRNA binding gene. ScNOP14 is also present in the small subunit processome complex, which is required for processing of pre-18S rRNA.
The putative functions of **AgUTP18** and **AgUTP21** were already described in the previous section.
The *S. cerevisiae* homolog of **AgPNO1** is an essential KH domain-containing nucleolar protein required for pre-18S rRNA processing. It interacts with ScDim1p, an 18S rRNA dimethyltransferase, and also with ScNob1p, which is involved in proteasome biogenesis.
ScNOP2, homolog of **AgNOP2**, is a probable RNA m(5)C methyltransferase, essential for processing and maturation of 27S pre-rRNA and large ribosomal subunit biogenesis. It localized to the nucleolus and is a constituent of 66S pre-ribosomal particles.
The *S. cerevisiae* homolog of **AgNSA1** is a protein constituent of 66S pre-ribosomal particles that contributes to processing of the 27S pre-rRNA (reviewed in (Fromont-Racine *et al.*, 2003)).

Ribosome biosynthesis and assembly: other functions. The *S. cerevisiae* homolog of **AgNOG1**, ScNOG1, is a GTPase that associates with free 60S ribosomal subunits in the nucleolus and is required for 60S subunit biogenesis. In *S. cerevisiae* ScNOG1 forms a complex that includes 60S ribosomal proteins and pre-ribosomal proteins ScNop7 and ScRlp24. The ScNog1 complex shuttles

between the nucleolus and the nucleoplasm for ribosome biogenesis, but it is tethered to the nucleolus by both nutrient depletion and AgTOR (target of rapamycin) inactivation, causing cessation of the late stages of ribosome biogenesis (Honma *et al.*, 2006).
We found that the expression of all components of the AgNog1 complex was strongly down-regulated upon AgTec1 deletion. AgNOG1 mRNA was 16-fold, AgNOP7 mRNA was 12-fold and AgRLP24 mRNA was 6.6-fold down-regulated.
The yeast homolog of **AgSRO9/SLF1** codes for a RNA binding protein that associates with polysomes and has been proposed to be involved in regulating mRNA translation.
ABR232C (homolog of YLR073C) is a gene of unknown function that is not expressed in cells deleted for Tec1.
As inferred from its *S. cerevisiae* homolog, **AgDRS1** encodes a nucleolar DEAD-box protein required for ribosome assembly and function. It is involved in the synthesis of the 60S ribosomal subunits constituent of the 66S pre-ribosomal particles.
The homolog of **AgRPA43**, ScRPA43, encodes the A43 subunit of the DNA dependent RNA polymerase I. ATP-dependent RNA helicase activity has been proposed to be the biological function of ScRPA43.
ScNOC3 is the homolog of **AgNOC3** and binds to the 66S ribosomal precursors to mediate their intranuclear transport.
The *S. cerevisiae* homolog of **AgPUF6**, ScPUF6, is a pumilio-homology domain protein that binds in *S. cerevisiae* ASH1 mRNA and represses its translation resulting in proper asymmetric localization of the ASH1 mRNA. ScASH1 encodes a zinc-finger inhibitor of HO endonuclease gene expression that is required for mating type switching. The possible function of the AgASH1 homolog is difficult to extrapolate as the *A. gossypii* genome does not encode a HO endonuclease.
Ribosomal proteins. Since the majority of the genes that were down-regulated were involved in ribosome biogenesis and assembly, particularly in rRNA maturation/ processing, we were wondering whether the ribosomal protein genes themselves would also be down-regulated. And indeed, the ribosomal protein genes were down-regulated in hyphae lacking AgTec1 (log2ratios between 0.5 and -2.5).

7.2.3 Differential gene expression in cells lacking STE12

Up-regulated genes in Ags*te12*Δ

Amino acid biosynthesis. The GO category "amino acid biosynthesis" was significantly up-regulated in the t-profiler analysis (t-value = 14, E = 1.0E-15).

Arginine biosynthesis. The *S. cerevisiae* homolog of **AgARG3** codes for a Ornithine carbamoyltransferase (carbamoylphosphate:L-ornithine carbamoyltransferase) which catalyzes the sixth step in the biosynthesis of the arginine precursor ornithine. **AgARG4** potentially encodes an argininosuccinate lyase which catalyzes the final step in the arginine biosynthesis pathway.
The *S. cerevisiae* homolog of **AgECM40** codes for a mitochondrial ornithine acetyltransferase that catalyzes the fifth step in arginine biosynthesis. It also possesses acetylglutamate synthase activity, this means it regenerates acetylglutamate while forming ornithine.

Isoleucine biosynthesis. The *S. cerevisiae* homolog of **AgILV1** encodes a threonine deaminase which catalyzes the first step in isoleucine biosynthesis. Its expression is under general amino acid control. The ILV1 locus exhibits highly positioned nucleosomes whose organization is independent of the known ScILV1 regulation.

Lysin biosynthesis. AgLYS20/LYS21 potentially enodes a homocitrate synthase

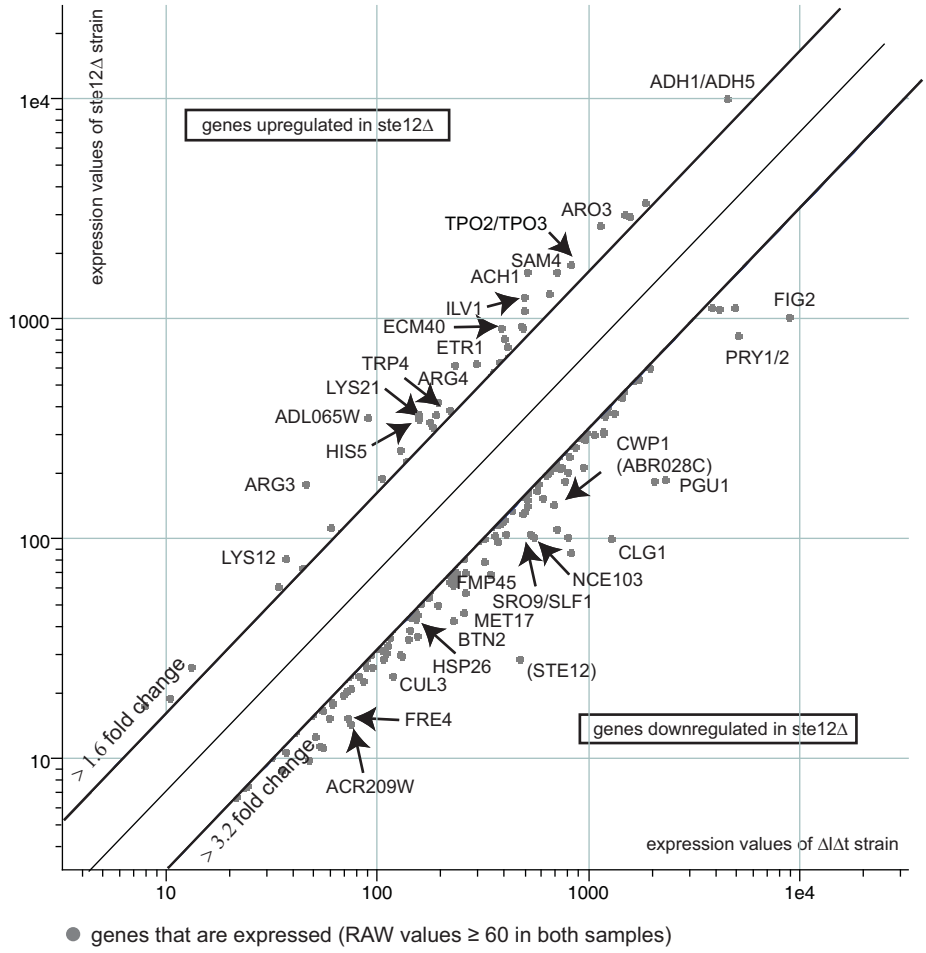


Figure 7-4 Scatter plot showing the top 15 up- and down-regulated genes

isozyme that catalyzes the condensation of acetyl-CoA and alpha-ketoglutarate to form homocitrate, which is the first step in the lysine biosynthesis pathway. ScLys20 is highly similar to the other isozyme, ScLys21p. The *S. cerevisiae* homolog of **AgLYS12** codes for a homo-isocitrate dehydrogenase, an NAD-linked mitochondrial enzyme required for the fourth step in the biosynthesis of lysine, in which homo-isocitrate is oxidatively decarboxylated to alpha-ketoadipate.

Aromatic amino acid biosynthesis. AgARO3 belongs to the aromatic amino acid biosynthesis familiy. ScARO3 codes for 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) synthase which catalyzes the first step in aromatic amino acid biosynthesis. It is feedback-inhibited by phenylalanine or high concentration of tyrosine or tryptophan.

Methionine biosynthesis. The *S. cerevisiae* homolog of **AgSAM4** encodes a S-adenosylmethionine-homocysteine methyltransferase. It functions along with ScMht1p in the conversion of S-adenosylmethionine (AdoMet) to methionine to control the methionine/AdoMet ratio.

Histidin biosynthesis. AgHIS5 potentially codes for a histidinol-phosphate aminotransferase which catalyzes the seventh step in histidine biosynthesis. It responds to general control of amino acid biosynthesis.

Tryptophan biosynthesis.
The *S. cerevisiae* homolog of **AgTRP4** codes

TABLE 7-3. TOP DIFFERENTIALLY EXPRESSED GENES IN CELLS LACKING STE12

	SYSTEMATIC NAME	COMMON NAME	EXP ΔlΔt	EXP ste12Δ	LOG2 (RATIO)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGIGAL PROCESS	CELLULAR COMPONENT
UPREGULATED	ADL065W		89.54	711.70	2.99	Putative protein of unknown function; expression induced under carbon limitation and repressed under high glucose	-	-	cytoplasm
	AFR098W	ARG3	45.40	354.02	2.96	Ornithine carbamoyltransferase (carbamoylphosphate:L-ornithine carbamoyltransferase), catalyzes the sixth step in the biosynthesis of the arginine precursor ornithine	arginine biosynthetic process	ornithine carbamoyltransferase activity	cytosol
	AFR020W	ACH1	504.15	3261.00	2.69	Acetyl-coA hydrolase, primarily localized to mitochondria; required for acetate utilization and for diploid pseudohyphal growth	acetate metabolic process	acetyl-CoA hydrolase activity	mitochondrion
	AEL019W	ARG4	230.40	1237.80	2.43	Argininosuccinate lyase, catalyzes the final step in the arginine biosynthesis pathway	arginine biosynthetic process	argininosuccinate lyase activity	cytosol
	ACR232C	ILV1	493.15	2525.00	2.36	Threonine deaminase, catalyzes the first step in isoleucine biosynthesis; expression is under general amino acid control; ILV1 locus exhibits highly positioned nucleosomes whose organization is independent of known ILV1 regulation	branched chain family amino acid biosynthetic process	L-threonine ammonia-lyase activity	mitochondrion
	ADR107W	LYS21/LYS20	155.05	741.10	2.26	Homocitrate synthase isozyme, catalyzes the condensation of acetyl-CoA and alpha-ketoglutarate to form homocitrate, which is the first step in the lysine biosynthesis pathway; highly similar to the other isozyme, Lys20p	lysine biosynthetic process via aminoadipic acid	homocitrate synthase activity	nucleus
	ADR138C	ECM40	380.25	1813.20	2.25	Mitochondrial ornithine acetyltransferase, catalyzes the fifth step in arginine biosynthesis; also possesses acetylglutamate synthase activity, regenerates acetylglutamate while forming ornithine	cell wall organization and biogenesis	amino-acid N-acetyltransferase activity	mitochondrion
	ABL102C	ARO3	1119.80	5338.00	2.25	3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) synthase, catalyzes the first step in aromatic amino acid biosynthesis and is feedback-inhibited by phenylalanine or high concentration of tyrosine or tryptophan	aromatic amino acid family biosynthetic process	3-deoxy-7-phosphoheptulonate synthase activity	cytoplasm
	ABL125W	SAM4	695.75	3282.00	2.24	S-adenosylmethionine-homocysteine methyltransferase, functions along with Mht1p in the conversion of S-adenosylmethionine (AdoMet) to methionine to control the methionine/AdoMet ratio	methionine biosynthetic process from S-adenosylmethionine*	homocysteine S-methyltransferase activity	cytoplasm
	ADL214C	LYS12	36.15	163.01	2.17	Homo-isocitrate dehydrogenase, an NAD-linked mitochondrial enzyme required for the fourth step in the biosynthesis of lysine, in which homo-isocitrate is oxidatively decarboxylated to alpha-ketoadipate	lysine biosynthetic process	homoisocitrate dehydrogenase activity	mitochondrion
	AER032W	ADH1/ADH5	4506.00	20254.00	2.17	Alcohol dehydrogenase, fermentative isozyme active as homo- or heterotetramers; required for the reduction of acetaldehyde to ethanol, the last step in the glycolytic pathway	NADH oxidation	alcohol dehydrogenase activity	cytosol
	ADL249W	HIS5	156.40	700.80	2.16	Histidinol-phosphate aminotransferase, catalyzes the seventh step in histidine biosynthesis; responsive to general control of amino acid biosynthesis; mutations cause histidine auxotrophy and sensitivity to Cu, Co, and Ni salts	histidine biosynthetic process	histidinol-phosphate transaminase activity	intracellular
	AER239W	TRP4	190.35	839.20	2.14	Anthranilate phosphoribosyl transferase of the tryptophan biosynthetic pathway, catalyzes the phosphoribosylation of anthranilate, subject to the general control system of amino acid biosynthesis	tryptophan biosynthetic process	anthranilate phosphoribosyltransferase	cytoplasm
	AFR322C	TPO2/TPO3	821.45	3577.00	2.12	Polyamine transport protein specific for spermine; localizes to the plasma membrane; transcription of TPO2 is regulated by Haa1p; member of the major facilitator superfamily	polyamine transport	spermine transmembrane transporter activity	plasma membrane
	AEL081W	ETR1	293.30	1266.30	2.11	2-enoyl thioester reductase, member of the medium chain dehydrogenase/reductase family; localized to in mitochondria, where it has a probable role in fatty acid synthesis	aerobic respiration	enoyl-[acyl-carrier-protein] reductase activity	mitochondrion

	SYSTEMATIC NAME	COMMON NAME	EXP ΔlΔt	EXP ste12Δ	LOG2 (RATIO)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGIGAL PROCESS	CELLULAR COMPONENT
DOWNREGULATED	AER440C	CLG1	1259.00	201.63	-2.64	Cyclin-like protein that interacts with Pho85p; has sequence similarity to G1 cyclins PCL1 and PCL2	cell cycle	cyclin-dependent protein kinase regulator activity	cyclin-dependent protein kinase holoenzyme complex
	ADR327W	PGU1	2275.00	372.20	-2.61	Endo-polygalacturonase, pectolytic enzyme that hydrolyzes the alpha-1,4-glycosidic bonds in the rhamnogalacturonan chains in pectins	pseudohyphal growth	polygalacturonase activity	extracellular region
	ADL397C	HSP26	815.65	173.04	-2.24	Small heat shock protein with chaperone activity that is regulated by a heat induced transition from an inactive oligomeric (24-mer) complex to an active dimer; induced by heat, upon entry into stationary phase, and during sporulation	response to stress	unfolded protein binding	cytoplasm
	AFR711C	FIG2	8831.00	2035.80	-2.12	Cell wall adhesin, expressed specifically during mating; may be involved in maintenance of cell wall integrity during mating	cellular morphogenesis during conjugation with cellular fusion*	-	chitin- and beta-glucan-containing cell wall
	AGR019C	MSB2	706.50	221.30	-1.67	Mucin family member at the head of the Cdc42p- and MAP kinase-dependent filamentous growth signaling pathway; also functions as an osmosensor in parallel to the Sho1p-mediated pathway; potential Cdc28p substrate	establishment of cell polarity (sensu Fungi)	osmosensor activity	integral to plasma membrane
	AAL179W	PRY1/PRY2	5057.50	1682.50	-1.59	Protein of unknown function, has similarity to Pry2p and Pry3p and to the plant PR-1 class of pathogen related proteins	-	-	vacuole, cell cycle-correlated morphology
	ADL031W	MET17	255.40	92.49	-1.47	O-acetyl homoserine-O-acetyl serine sulfhydrylase, required for sulfur amino acid synthesis	response to drug	cysteine synthase activity	cytoplasm
	ADL226C	NCE103	547.15	204.82	-1.42	Carbonic anhydrase; poorly transcribed under aerobic conditions and at an undetectable level under anaerobic conditions; involved in non-classical protein export pathway	response to oxidative stress	carbonate dehydratase activity	cytoplasm
	AFR325W	BTN2	224.40	85.16	-1.40	v-SNARE binding protein that facilitates specific protein retrieval from a late endosome to the Golgi; modulates arginine uptake, possible role in mediating pH	retrograde transport, endosome	protein binding	cytosol
	ACR209W	NOHBY332	73.59	28.56	-1.37	no description	-	-	-
	ABR028C	CWP1	528.40	212.04	-1.32	Cell wall mannoprotein, linked to a beta-1,3- and beta-1,6-glucan heteropolymer through a phosphodiester bond; involved in cell wall organization	cell wall organization and biogenesis	structural constituent of cell wall	chitin- and beta-glucan-containing cell wall*
	AFR498W	CUL3	117.40	47.50	-1.31	Ubiquitin-protein ligase, member of the cullin family with similarity to Cdc53p and human CUL3; required for ubiquitin-dependent degredation of the RNA Polymerase II subunit RPO21	ubiquitin-dependent protein catabolic process	protein binding	cytoplasm
	AGR328C	FRE4	71.55	30.75	-1.22	Ferric reductase, reduces a specific subset of siderophore-bound iron prior to uptake by transporters; expression induced by low iron levels	siderophore-iron transport	ferric-chelate reductase activity	plasma membrane
	AFR713W	SRO9/SLF1	672.15	289.38	-1.22	Cytoplasmic RNA-binding protein that associates with translating ribosomes; involved in heme regulation of Hap1p as a component of the HMC complex, also involved in the organization of actin filaments; contains a La motif	translation	RNA binding	polysome
	ABL156C	FMP45	258.80	114.16	-1.18	Integral membrane protein localized to mitochondria (untagged protein) and eisosomes, immobile patches at the cortex associated with endocytosis; sporulation and sphingolipid content are altered in mutants; has homologs SUR7 and YNL194C	cell wall organization and biogenesis	-	mitochondrion

for an anthranilate phosphoribosyltransferase of the tryptophan biosynthetic pathway that catalyzes the phosphoribosylation of anthranilate. It is subject to the general control system of amino acid biosynthesis.

Genes of unknown function. The gene which was most up-regulated upon STE12 deletion was **ADL065W**. ADL065W is a gene homologous to the *S. cerevisiae* genes YIL057C and YER067W whose functions are not known. The expression of the *S. cerevisiae* gene YIL057C is induced under carbon limitation. ADL065W was also found to be up-regulated in the strain Ag*tec1*Δ.

Other genes. AgACH1 shares homology with ScACH1 which encodes an acetyl-CoA hydrolase that is primarily localized to mitochondria and is required for acetate utilization. AgACH1 mRNA levels were also up-regulated in Ag*tec1*Δ.
AgADH1/5 was up-regulated upon AgSTE12 deletion. ScADH1 encodes an alcohol dehydrogenase, a fermentative isozyme which is active as homo- or heterotetramer;. It is required for the reduction of acetaldehyde to ethanol which is the last step in the glycolytic pathway. ScADH5 encodes a formaldehyde dehydrogenase which detoxifies from formaldehyd.
The *S. cerevisiae* homolog of **AgTPO2/TPO3** encodes a polyamine transport protein specific for spermine. The proein localizes to the plasma membrane. The transcription of ScTPO2 is regulated by Haa1p. ScTPO2 is a member of the major facilitator superfamily. ScTPO3 has a similar function.
The *S. cerevisiae* homolog of **AgETR1** codes for a 2-enoyl thioester reductase which is a member of the medium chain dehydrogenase/reductase family. The enzyme is localized in the mitochondria, where it has a probable role in fatty acid synthesis. The expression level of AgETR1 was also highly up-regulated in Ag*tec1*Δ.

Down-regulated genes in Ag*ste12*Δ

Extracellular/cell wall protein genes.
The *S. cerevisiae* homolog of **AgPGU1** codes for a endo-polygalacturonase which is a pectolytic enzyme that hydrolyzes the alpha-1,4-glycosidic bonds in the rhamnogalacturonan chains in pectins. AgPGU1 is among the top 15 highly up-regulated genes in response to DIG1/2 deletion. Interestingly, it was found to be among the top15 down-regulated genes in Ag*tec1*Δ.
ScFIG2, the homolog of **AgFIG2**, codes for an integral membrane protein required for efficient mating. It may participate in or regulate the low affinity Ca2+ influx system, which affects intracellular signaling and cell-cell fusion during mating.
AgCWP1d (ABR028C) is one of seven members of the *A. gossypii* CWP1 homolog family. Members of this gene family have distinct patterns of gene expression (see chapter 5). The single *S. cerevisiae* homolog encodes a cell wall mannoprotein, linked to a beta-1,3- and beta-1,6-glucan heteropolymer through a phosphodiester bond. It is involved in cell wall organization.

Cellular transport. The *S. cerevisiae* homolog of **AgBTN2** encodes a v-SNARE binding protein that facilitates specific protein retrieval from a late endosome to the Golgi. It modulates arginine uptake and has a possible role in mediating pH homeostasis between the vacuole and plasma membrane.
ScFMP45, the homolog of **AgFMP45**, codes for an integral membrane protein that is localized to the mitochondria and to eisosomes. Eisosomes are immobile patches at the cortex probably associated with endocytosis. ScFMP45mutants show altered sporulation and sphingolipid content.
Response to stress/osmotic changes/drug. AgHSP26 is one of five genes presently annotated as homologs of ScHSP26 (YBR072W) which codes for a

small heat shock protein with chaperone activity.
AgMSB2 has a homolog in *S. cerevisiae*, ScMSB2, that encodes a mucin family member at the head of the Cdc42p- and MAP kinase-dependent filamentous growth signaling pathway. It also functions as an osmosensor in parallel to the Sho1p-mediated pathway and is a potential Cdc28p substrate. The *S. cerevisiae* homolog of
AgMET17 codes for a O-acetyl homoserine-O-acetyl serine sulfhydrylase that is required for sulfur amino acid synthesis. According to its molecular function GO classification, it is involved in the response to drugs.
AgNCE103 has a homolog in *S. cerevisiae* that encodes a carbonic anhydrase. The gene is poorly transcribed under aerobic conditions and at an undetectable level under anaerobic conditions. It is involved in non-classical protein export pathway. High levels of mRNA of another gene involved in non-classical export, **AgNCE102**, was found in spores. It is not clear why the non-classical export pathways are expressed at higher levels in *A. gossypii* compared to *S. cerevisiae*. It might be related to fast filamentous growth that requires efficient secretion of proteins and other molecules.

Other genes. The gene which was strongest down-reglated upon Ste12 deletion was **AgCLG1**. ScCLG1 encodes a cyclin-like protein that interacts with Pho85p and has sequence similarity to the G1 cyclins PCL1 and PCL2.
AgPRY1/PRY2, the homolog of YJL079C, YJL078C and YKR013W, encodes a protein of unknown function with similarity to the plant PR-1 class of pathogen related proteins. Interestingly, AgPRY1/2 was amongst the fifteen most up-regulated genes in hyphae deleted for AgDIG1/2.
The yeast homolog of **AgSRO9/SLF1** codes for a RNA binding protein that associates with polysomes and has been proposed to be involved in regulating mRNA translation. AgSRO9/SLF1 was also highly down-regulated in Ag*tec1*Δ.
NOHBY332 (ACR209W) has no homolog in *S. cerevisiae* and was found to be poorly (if at all) expressed in wild type and was clearly not expressed in Ag*ste12*Δ.
The *S. cerevisiae* homolog of **AgCUL3** codes for a ubiquitin-protein ligase that is a member of the cullin family with similarity to ScCdc53p and human CUL3. It is required for ubiquitin-dependent degredation of the RNA Polymerase II subunit RPO21. AgCUL3 was not (or very poorly) expressed in Ag*ste12*Δ.
The same holds for the expression of **AgFRE4**, a homolog of ScFRE4. ScFRE4 codes for a ferric reductase that reduces a specific subset of siderophore-bound iron prior to uptake by transporters. Its expression is induced by low iron levels.

7.3 Discussion and Conclusions

The role of AgDig1/2

Genes that were expressed at considerable higher levels upon AgDIG1/2 deletion were mating genes and filamentation genes. Among the top 15 up-regulated genes, we found five genes (AgFUS1, AgFUS3, AgSST2, AgSTE4 and AgSTE12) the yeast homologs of which are involved in mating and three genes (AgMUC1, AgPGU1, AgPLC1) the yeast homologs of which are involved in filamentation.
This means that AgDIG1/2 has a similar role as ScDIG1/2 which is to repress mating genes and filamentation genes. Among the top 15 up-regulated genes, there were two genes, AgPGU1 and PRY1/PRY2, which were among the top 15 down-regulated genes in other experiments (AgPGU1 in Ag*tec1*Δ and in Ag*ste12*Δ and AgPRY1/PRY2 in Ag*ste12*Δ). In addition STE12 was up-regulated in Ag*dig1/2*Δ. This could mean that AgDig1/2p acts as a repressor on AgTEC1 and AgSTE12. AgSte12p might, if

DIG1/2
Up-regulated gene groups
Genes homologous to *S. cerevisiae* mating genes
Genes homologous to *S. cerevisiae* filamentation genes
Genes of unknown function
Down-regulated gene groups
Glyoxylate cycle genes
Ethanol inducible genes
Sporulation and meiosis genes
Ribosomal protein genes
TEC1
Up-regulated gene groups
Thiamin biosynthesis genes
Genes involved in utilization of alternative carbon sources
Genes of unknown function
Down-regulated gene groups
Ribosome biosynthesis and assembly (snoRNA binding, processing and maturation of pre-rRNA)
Ribosomal protein genes
Nitrogen utilization and pseudohyphal growth
STE12
Up-regulated gene groups
Amino acid biosynthesis (ARG, ILE, LYS, ARO, MET, HIS, TRP)
Down-regulated gene groups
Extracellular/cell wall protein genes
Cellular transport
Response to stress/osmotic changes/drugs

Table 7-4 Functional groups to which the 15 most up- or down-regulated genes belong

not repressed by AgDig1/2p, act in a positive feed-back loop as a transcription activator on its own promoter. The function of the *S. cerevisiae* homolog of AgPRY1/PRY2, a gene with homology to the plant PR-1 class of pathogen related proteins, is not known. The expression pattern of AgPRY1/PRY2 suggests a possible function in mating in *A. gossypii*.

The fifteen genes that were expressed at considerable lower levels upon DIG1/2 deletion were glyoxylate cycle genes (AgCIT3, AgICL1) and ethanol inducible genes (AgACS1, AgYAT1, AgYAT2, AgPDH1). For an overview of the glyoxylate cycle please refer to chapter 3 figure 3-4. The glyoxylate cycle involves two critical steps that catalyzed by the enzymes isocitrate lyase and malate synthase, which bypass the two decarboxylation steps of the TCA cycle. The glyoxylate cycle is required for growth on gluconeogenetic carbon sources, such as acetate or ethanol.

The down-regulated glyoxlate cycle genes might cause the slow growth of *dig1/2*Δ or the slow growth effects their down-regulation. In other words, either the deletion of Ag*dig1/2* down-regulates the glyoxylate cycle genes preventing hyphae to reach fast growth speeds on glucose plates. Or, the deletion of Ag*dig1/2* leads to the induction of mating genes thus causing a “cell cycle arrest”. As a consequence energy requirements are reduced which is reflected in the down-regulated glyoxylate shunt. The t-profiler analysis revealed that in addition, the ribosomal protein genes were significantly down-regulated. This makes sense. In a growing cell, a large portion of total transcription and also a large portion of total cellular energy are dedicated to ribosome biogenesis. Under slow growth conditions, e.g. in response to nutrient limitation, ribosomal proteins need to be down-regulated.

A group of four genes being potentially involved in sporulation (AgRIM4, AgADY2/

FUN34) and meiosis (AgHOP1, AgHOP2) were found to be highly down-regulated upon AgDIG1/2 deletion. This is consistent with the observed sporulation-defect of Ag*dig1/2*Δ. Here, we compared the expression profiles of not-yet-sporulating hyphae growing at the border of the mycelial colony. We would expect to find even more down-regulated sporulation genes in a comparison of the expression profiles from hyphae in the center of the colony.
It is interesting to find meiosis and sporulation genes down-regulated at the same time. This could be a hint that meiosis is linked to spore formation in *A. gossypii* which is interesting because it is still not known whether *A. gossypii* produces gametangia or meiosporangia, which, if known, would tell whether spore formation is a sexual or asexual process in *A. gossypii*.

The role of AgTec1

Genes whose expression was markedly increased by deletion of AgTec1 included thiamine biosynthesis genes, genes involved in the utilization of alternative carbon sources (acetate, ethanol, oleate) and a group of genes with unknown functions.
Genes which are involved in the utilization of alternative carbon sources are, at least in *S. cerevisiae*, repressed when sufficient glucose is present in the medium (glucose-repression). In Ag*tec1*Δ we found these genes even under glucose rich conditions highly up-regulated, for instance AgADH2 (alcohol dehydrogenase II), a gene which catalyzes the conversion of ethanol to acetaldehyde. Furthermore, AgACH1, a gene potentially involved in acetate utilization and AgSPS19, a gene potentially induced in the presence of oleate, were highly up-regulated upon AgTec1 deletion. In addition we found a 3-fold up-regulation of the glucose transporter AgHXT1/2 and a 9-fold up-regulation of the glukokinase AgGLK1, the yeast homolog of which is de-repressed on nonfermentable carbon sources.

These data suggest that deletion of AgTEC1 mimics glucose limitation. This is in agreement with the previous observation that Ag*tec1*Δ hyphae have increased nuclear distances which is normally observed in wild type *A. gossypii* hyphae starved for glucose (see chapter 6). Also the invasive growth phenotype after 10 days on full medium plates is similar to the one observed in wild type mycelia grown under limiting glucose conditions (0.2% glucose).

Interactions between Tec1p and thiamine biosynthesis genes is not or only poorly documented in the literature, but an interaction between medium glucose levels and expression of thiamin biosynthesis genes has been observed in *S. cerevisiae* (Kaeberlein et al, 2002; DeRisi et al., 1997). In order to elucidate whether the transcription factor AgTec1 directly binds thiamin biosynthesis genes, *in silico* searches were done for conserved AgTec1 binding sites 10-500 bps upstream of thiamin biosynthesis open reading frames. No conserved Tec1p or alternative binding sites were found, which suggests that the regulation must be indirect.

Genes whose expression was negatively affected by deletion of AgTec1 were predominately genes involved in ribosome biosynthesis and assembly. Many of the most strongly down-regulated genes were involved in snoRNA binding or in the processing and maturation of pre-rRNAs. In addition the ribosomal protein genes themselves were also significantly down-regulated as could be shown in the t-profiler analysis.
Ribosomal protein genes are regulated by the nutrient-sensitive, conserved TOR (target of rapamycin) signaling pathway (Hall et al., 1999). As a central controller of cell growth, TOR regulates several growth-related processes including ribosome biogenesis via the RAS-cAMP-PKA pathway (Schmelzle *et al.*, 2004). In *S. cerevisiae*, RNA Pol I- and RNA Pol III-dependent transcription and 35S rRNA processing are strongly reduced upon

TOR inhibition by rapamycin treatment. Moreover, rapamycin treatment leads to a rapid and pronounced downregulation of RNA Pol II-dependent RP genes (Zaragoza *et al.*, 1998; Hardwick *et al.*, 1999; Powers *et al.*, 1999; Shamji *et al.*, 2000; Schmelzle *et al.*, 2004). Thus, TOR broadly controls ribosome biogenesis. Nutrient limitation, for instance glucose limitation, is signaled via TOR and leads to down-regulation of the expression of ribosomal protein genes.

In *S. cerevisiae* NOG1 forms a complex that includes 60S ribosomal proteins and pre-ribosomal proteins Nop7 and Rlp24. The Nog1 complex shuttles between the nucleolus and the nucleoplasm for ribosome biogenesis, but it is tethered to the nucleolus by both nutrient depletion and TOR inactivation, causing cessation of the late stages of ribosome biogenesis (Honma et al., 2006). We found that the expression of all components of the Nog1 complex was strongly down-regulated upon Tec1 deletion. AgNOG1 mRNA was 16-fold, AgNOP7 mRNA was 12-fold and AgRLP24 mRNA was 6.6-fold down-regulated.

Therefore we hypothesize that components of the TOR signaling pathway are upstream of AgTec1. Under glucose non-limiting conditions, AgTec1 negatively regulates the expression of downstream genes that are involved in the utilization of alternative carbon sources and thiamin biosynthesis.
Furthermore we hypothesize that AgTec1 is a transcription factor that contributes to control of ribosome biogenesis via AgNOG1, possibly through the putative AgNog1/Nop7/Rlp24 complex. In well fed polarly growing germlings, AgTec1 has a dual role as it acts as a repressor on the expression of genes involved in utilization of alternative carbon sources and thiamin biosynthesis and as an activator on the expression of ribosome biosynthesis genes.

The role of AgSTE12

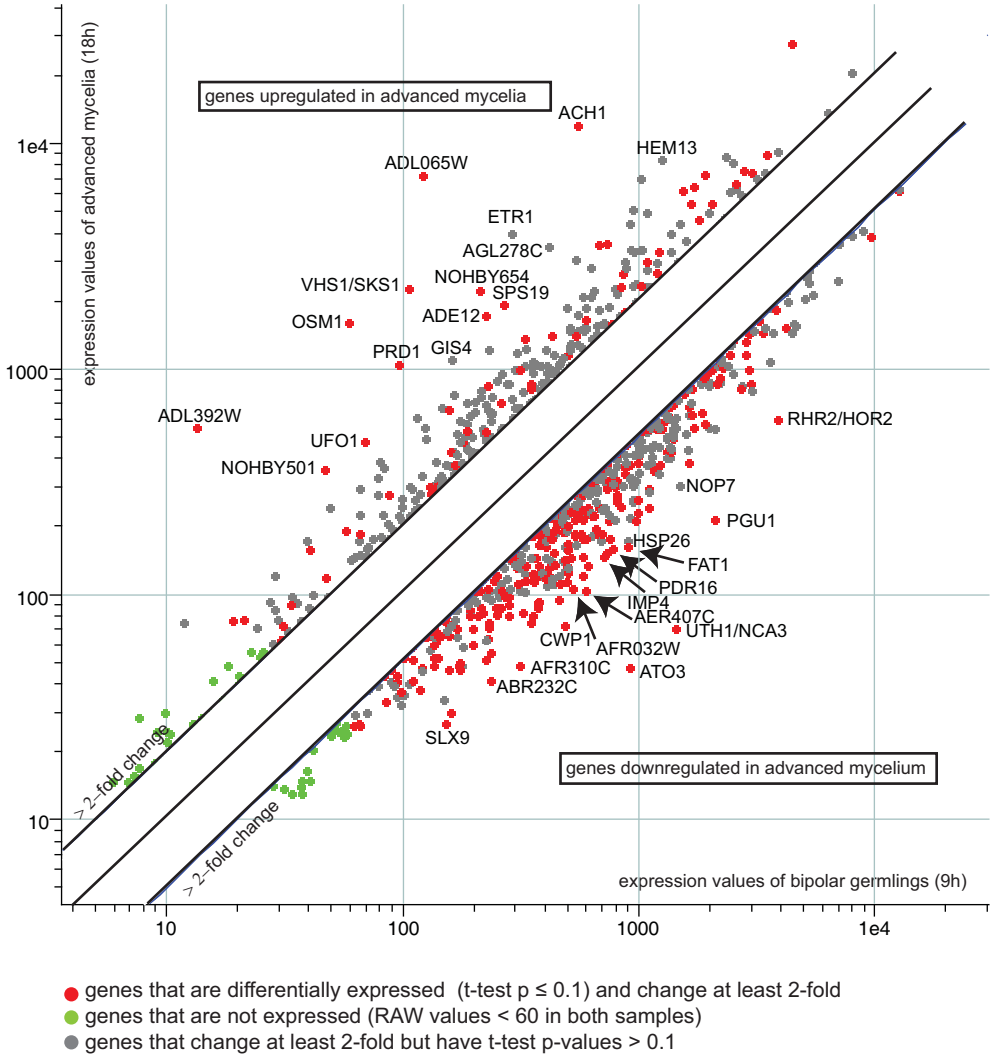
Genes whose expression was positively affected by deletion of AgSTE12 were mainly genes involved in amino acid biosynthesis (ARG, ILE, LYS, ARO, MET, HIS, TRP). Under conditions where the mating cascade is not induced, AgSTE12 might have a function in repressing the expression of amino acid biosynthesis genes.
Three genes that were among the top 15 up-regulated were also found to be among the 15 top up-regulated in Ag*tec1*Δ: ADL065W, AgACH1 and AgETR1. This could mean that AgSTE12 and AgTEC1 act together some times, maybe as a complex on certain genes. Interestingly, we found MSB2 as the fifth most strongly down-regulated gene in Ag*ste12*Δ. ScMSB2 encodes a mucin family member at the head of the Cdc42p- and MAP kinase-dependent filamentous growth signaling pathway. Msb2, is a FG-pathway-specific factor that promotes differential activation of the MAPK for the FG pathway, Kss1. Msb2 is localized to polarized sites on the cell surface and interacts with Cdc42 and with the osmosensor for the high osmolarity glycerol response (HOG) pathway (Cullen *et al.*, 2004). Why the deletion of STE12 would down-regulate MSB2 expression is not know.

In this chapter the role of AgTec1 was investigated under conditions where the mating cascade was not induces. It was surprising to find that even under such conditions. AgSTE12 plays a role in the cell. In order to fully understand the role of AgSTE12 experiments with induced mating cascades are necessary. The deletion of DIG1/2 resulted in the transcriptional activation of many mating genes. A possibility would be to compare the transcriptional profiles of AgDIG1/2 with the double deletion AgDIG1/2Δ STE12Δ.

Appendix

Appendix 1: Development from a bipolar germling to an advanced mycelium

The transcription profile of advanced mycelium (18 h) was already described in chapter 3 where we compared it to the profile of very fast growing hyphae. The transcription profile of advanced mycelium was also used in chapter 5 where the *A. gossypii* cell wall transcriptome was compared to the proteome. Here, we use this transcription profile for a comparison with the transcription profile of the bipolar germling (9 h; see chapter 2). The data was evaluated in the same manner as previously described and is widely self-explanatory. Table A1-1 lists the top fifteen up-and down-regulated genes, figure A1-1 shows the expression levels these genes in a scatter plot and table A1-2 lists the significantly up- and down-regulated gene groups.



- genes that are differentially expressed (t-test p ≤ 0.1) and change at least 2-fold
- genes that are not expressed (RAW values < 60 in both samples)
- genes that change at least 2-fold but have t-test p-values > 0.1

Figure A1-1 Scatter plot showing the top 15 up- and down-regulated genes

TABLE A1-1: TOP DIFFERENTIALLY EXPRESSED GENES IN ADVANCED MYCELIA (18h) COMPARED TO A GERMINATING CULTURE (9h)

	SYSTEMATIC NAME	COMMON NAME	EXP 9h	EXP 18h	LOG2 (ratio)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGICAL PROCESS	CELLULAR COMPONENT
	Genes expressed at higher levels in advanced mycelia								
UPREGULATED	ADL065W		121.10	7183.00	5.89	Putative protein of unknown function; expression induced under carbon limitation and repressed under high glucose			
	ADL392W		13.29	550.90	5.37	Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the vacuole, while HA-tagged protein is found in the soluble fraction, suggesting cytoplasmic localization	-	-	-
	AFR367W	OSM1	58.67	1603.00	4.77	Fumarate reductase, catalyzes the reduction of fumarate to succinate, required for the reoxidation of intracellular NADH under anaerobic conditions; mutations cause osmotic sensitivity	metabolic process	fumarate reductase (NADH) activity	mitochondrion
	AFR020W	ACH1	546.35	12011.50	4.46	Acetyl-coA hydrolase, primarily localized to mitochondria; required for acetate utilization and for diploid pseudohyphal growth	acetate metabolic process	acetyl-CoA hydrolase activity	mitochondrion
	ADR163W	VHS1/SKS1	105.10	2259.50	4.43	Cytoplasmic serine/threonine protein kinase; identified as a high-copy suppressor of the synthetic lethality of a sis2 sit4 double mutant, suggesting a role in G1/S phase progression; homolog of Sks1p			
	AEL081W	ETR1	287.60	4001.00	3.80	2-enoyl thioester reductase, member of the medium chain dehydrogenase/reductase family; localized to in mitochondria, where it has a probable role in fatty acid synthesis	aerobic respiration	enoyl-[acyl-carrier-protein] reductase activity	mitochondrion
	AGR406C	PRD1	95.88	1053.20	3.46	Zinc metalloendopeptidase, found in the cytoplasm and intermembrane space of mitochondria; with Cym1p, involved in degradation of mitochondrial proteins and of presequence peptides cleaved from imported proteins	proteolysis	metalloendopeptidase activity	cytoplasm
	AFR589C	NOHBY654	210.65	2223.00	3.40	no description	-	-	-
	AGL278C		413.15	3480.00	3.07	Putative protein of unknown function; the authentic, non-tagged protein is detected in highly purified mitochondria in high-throughput studies	-	-	-
	AEL035W	NOHBY501	46.31	358.30	2.95	no description	-	-	-
	ABL186W	ADE12	223.05	1718.00	2.95	Adenylosuccinate synthase, catalyzes the first step in synthesis of adenosine monophosphate from inosine 5'monophosphate during purine nucleotide biosynthesis; exhibits binding to single-stranded autonomously replicating (ARS) core sequence	telomere maintenance	DNA replication origin binding	cytoplasm
	ABL163W	SPS19	263.55	1921.50	2.87	Peroxisomal 2,4-dienoyl-CoA reductase, auxiliary enzyme of fatty acid beta-oxidation; homodimeric enzyme required for growth and sporulation on petroselineate medium; expression induced during late sporulation and in the presence of oleate	sporulation (sensu Fungi)	2,4-dienoyl-CoA reductase (NADPH) activity	peroxisomal matrix
	ADR198C	GIS4	160.45	1107.20	2.79	CAAX box containing protein of unknown function, proposed to be involved in the RAS/cAMP signaling pathway	cell ion homeostasis	-	plasma membrane
	AFR438W	UFO1	68.54	470.75	2.78	F-box receptor protein, subunit of the Skp1-Cdc53-F-box receptor (SCF) E3 ubiquitin ligase complex; binds to phosphorylated Ho endonuclease, allowing its ubiquitylation by SCF and subsequent degradation	ubiquitin-dependent protein catab	protein binding	cytoplasm
	AGR030C	HEM13	1237.00	8492.50	2.78	Coproporphyrinogen III oxidase, an oxygen requiring enzyme that catalyzes the sixth step in the heme biosynthetic pathway; localizes to the mitochondrial inner membrane; transcription is repressed by oxygen and heme (via Rox1p and Hap1p)	heme biosynthetic process	coproporphyrinogen oxidase activity	mitochondrial inner membrane

	SYSTEMATIC NAME	COMMON NAME	EXP 9h	EXP 18h	LOG2 (ratio)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGICAL PROCESS	CELLULAR COMPONENT
	Genes expressed at lower levels in advanced mycelia								
DOWNREGULATED	ADR322W	UTH1/NCA3	1419.50	69.71	-4.35	Protein of the SUN family (Sim1p, Uth1p, Nca3p, Sun4p) that may participate in DNA replication, promoter contains SCB regulation box at -300 bp indicating that expression may be cell cycle-regulated	mitochondrion organization and biogenesis	-	chitin- and beta-glucan-containing cell wall
	AFR144W	ATO3	898.20	47.62	-4.24	Plasma membrane protein, regulation pattern suggests a possible role in export of ammonia from the cell; member of the TC 9.B.33 YaaH family of putative transporters	nitrogen utilization	ammonium transmembrane transporter activity	mitochondrion
	ADR327W	PGU1	2093.00	214.80	-3.28	Endo-polygalacturonase, pectolytic enzyme that hydrolyzes the alpha-1,4-glycosidic bonds in the rhamnogalacturonan chains in pectins	pseudohyphal growth	polygalacturonase activity	extracellular region
	ABR028C	CWP1	485.40	72.68	-2.74	Cell wall mannoprotein, linked to a beta-1,3- and beta-1,6-glucan heteropolymer through a phosphodiester bond; involved in cell wall organization	cell wall organization and biogenesis	structural constituent of cell wall	chitin- and beta-glucan-containing cell wall
	AFR310C		308.50	47.82	-2.69	Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the cytoplasm; specifically phosphorylated in vitro by mammalian diphosphoinositol pentakisphosphate (IP7)	-	-	cytoplasm
	ADL071C	RHR2/HOR2	3835.50	600.80	-2.67	Constitutively expressed isoform of DL-glycerol-3-phosphatase; involved in glycerol biosynthesis, induced in response to both anaerobic and, along with the Hor2p/Gpp2p isoform, osmotic stress	response to drug	glycerol-1-phosphatase activity	cytoplasm
	AER407C		593.25	103.78	-2.52	Predicted protein shares weak similarity with uridine kinases and with phosphoribokinases; null exhibits no apparent phenotype	-	-	cytoplasm
	ABR232C		234.35	41.16	-2.51	Protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to endosomes; YLR073C is not an esssential gene	-	-	endosome
	AEL319C	SLX9	150.25	26.50	-2.50	Protein required for pre-rRNA processing; associated with the 90S pre-ribosome and 43S small ribosomal subunit precursor; interacts with U3 snoRNA; deletion mutant has synthetic fitness defect with an sgs1 deletion mutant	rRNA processing	snoRNA binding	nucleus
	ACL174W	FAT1	893.20	161.55	-2.47	Fatty acid transporter and very long-chain fatty acyl-CoA synthetase, may form a complex with Faa1p or Faa4p that imports and activates exogenous fatty acids	very-long-chain fatty acid metabolic process	long-chain-fatty-acid-CoA ligase activity	plasma membrane
	AAL128C	HSP26	889.30	174.60	-2.35	Small heat shock protein with chaperone activity that is regulated by a heat induced transition from an inactive oligomeric (24-mer) complex to an active dimer; induced by heat, upon entry into stationary phase, and during sporulation	response to stress	unfolded protein binding	cytoplasm
	AFR032W		524.15	106.25	-2.30	Putative protein of unknown function, identified by homology	-	-	cytoplasm
	AAR094W	NOP7	1469.00	302.15	-2.28	Nucleolar protein involved in rRNA processing and 60S ribosomal subunit biogenesis; constituent of several different pre-ribosomal particles; required for exit from G₀ and the initiation of cell proliferation	ribosome biogenesis and assembly	-	nucleus
	AER430W	IMP4	708.40	146.15	-2.28	Component of the SSU processome, which is required for pre-18S rRNA processing; interacts with Mpp10p; member of a superfamily of proteins that contain a sigma(70)-like motif and associate with RNAs	ribosome biogenesis and assembly	rRNA primary transcript binding	small nucleolar ribonucleoprotein complex
	ABL198C	PDR16	777.50	161.10	-2.27	Phosphatidylinositol transfer protein (PITP), downregulates Plb1p-mediated turnover of phosphatidylcholine, found in the cytosol and microsomes, homologous to Pdr16p, deletion affects phospholipid composition	response to drug	phosphatidylinositol transporter activity	cytoplasm

TABLE A1-2: Changes in activity of predefined groups in response to development into a advanced mycelium

	CATEGORY	ASPECT	t-VALUE	E-VALUE	MEAN	ORFs
UPREGULATED	cellular component unknown	C	6.82	1.30E-08	0.27	346
	oxidoreductase activity, acting on the CH-CH group of donors	F	6.72	2.50E-08	1.61	12
	alcohol metabolism	P	5.22	2.50E-04	0.49	102
	peroxisome	C	5.10	4.70E-04	0.66	45
	purine base metabolism	P	5.03	6.80E-04	1.04	13
	endoplasmic reticulum	C	5.01	7.60E-04	0.20	293
	nucleolar ribonuclease P complex	C	4.98	8.80E-04	1.37	8
	vacuole	C	4.97	9.30E-04	0.41	131
DOWNREGULATED	nucleolus	C	-22.76	< 1.0E-15	-1.28	173
	nucleobase, nucleoside, nucleotide and nucleic acid metabolism	P	-14.18	< 1.0E-15	-0.25	1042
	ribosome	C	-13.33	< 1.0E-15	-0.57	192
	nucleus	C	-8.65	< 1.0E-15	-0.01	1412
	nucleolus	C	-22.76	< 1.0E-15	-1.28	173
	nucleobase, nucleoside, nucleotide and nucleic acid metabolism	P	-14.18	< 1.0E-15	-0.25	1042
	ribosome	C	-13.33	< 1.0E-15	-0.57	192
	nucleus	C	-8.65	< 1.0E-15	-0.01	1412
	RNA metabolism	P	-7.73	1.50E-11	-0.15	362
	mitochondrion	C	-6.28	4.70E-07	-0.04	553
	intracellular	C	-4.98	8.80E-04	0.08	3419

Appendix 2: Comparison of a novel isolate with the strain used for functional genomics

The germination experiment described in chapter 2 and the experiment with fast growing hyphae described in chapter 3 were repeated with a novel natural isolate of *A. gossypii*, named FDAG, that was isolated in 2005 from a cotton stainer (insect) in Florida (Fred Dietrich, personal communication). RNA was extracted from two independent spore, germling or fast hyphae preparations, collected at previously indicated time points (0 h, 2 h, 5 h, 7 h, 9 h and 103 h). As described in the general chapter 1, cRNA targets were prepared, hybridized to the probes on the same chip that was designed according to the sequence of the original *A. gossypii* strain (#ATCC10895), and the fluorescence intensities were determined. The data from a comparison of very fast growing natural isolate hyphae with laboratory strain hyphae was evaluated and will be presented here. Table A2-1 contains the 30 top differentially expressed genes (15 up- and 15 down-regulated). Figure A2-1 shows these genes in a scatter plot, and in table A2-2, significant up- and down-regulated groups of genes

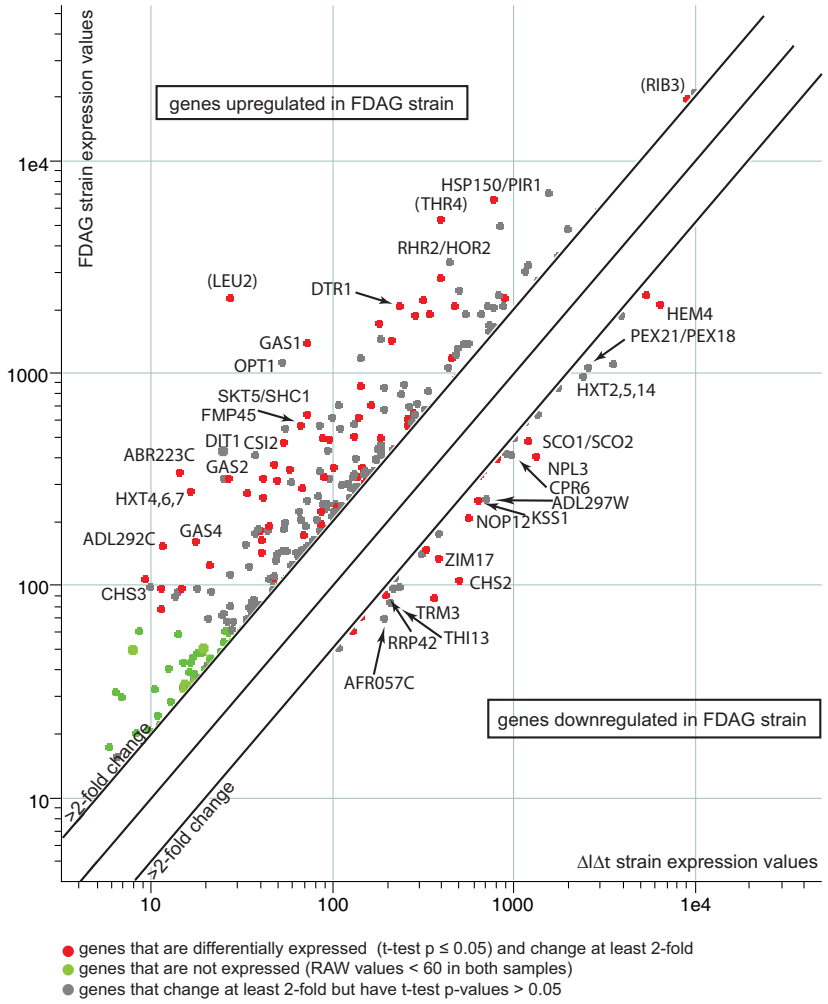


Figure A2-1 Scatter plot showing the top 15 up- and down-regulated genes

are given. The tables and the figure are self-explanatory. Nonetheless we will here point out some peculiarities of the natural isolate strain.

Comparison of gene expression in fast growing natural isolate strain hyphae with laboratory strain hyphae

Morphogenesis of the natural isolate is very similar to the sequenced *A. gossypii* laboratory strain. Growth speeds on Ashbya full medium (AFM) plates are comparable (see figure 1-4), but the natural isolate strain seems to sense and signal earlier a starvation condition, i.e. nutrient decline. Spore formation in the laboratory strain does not occur in fast growing hyphae in the outer zone (OZ) of the mycelial colony. However, in the natural isolate spore formation already strarted just behind this zone (Sylvia Voegeli, personal communication). This difference is also apparent when the transcription profile of very fast growing hyphae from the natural isolate is compared to that of the laboratory strain. Sporulation genes were already significantly up-regulated (t = 12.00, E = 1.0E-15). Four cell wall protein genes (**AgGAS1, AgGAS2, AgGAS4, AgPIR1/HSP15**) were among the fifteen top up-regulated genes many of which have *S. cerevisiae* homologs with sporulation-related functions. Interestingly, **AgCHS3**, the gene that potentially encodes chitin synthase III was expressed and **AgCHS2**, the gene that potentially encodes chitin synthase II, was down-regulated. **AgSKT5/SHC1**, a gene that potentially encodes the activator of AgChs3, and **AgCSI2**, a gene that potentially encodes a structural component of the chitin synthase III complex were also highly up-regulated. These results are consistent with a sporulation-specific role of AgCHS3 (see chapter 5). Interestingly, the t-profiler algorithm did not identify meiosis genes as significantly up-regulated, even though homologs of *S. cerevisiae* meiotic genes were up-regulated during sporulation in the laboratory strain.

A spore awakes II: Development from spores to bipolar germlings (natural isolate strain)

As already shown in chapter 2, spore germination is not a synchronous process in the filamentous fungi and also not for *A. gossypii*. When we counted the different stages of germination at selected time points post inoculation of natural isolate spores in liquid medium, we observed differences with respect to the laboratory

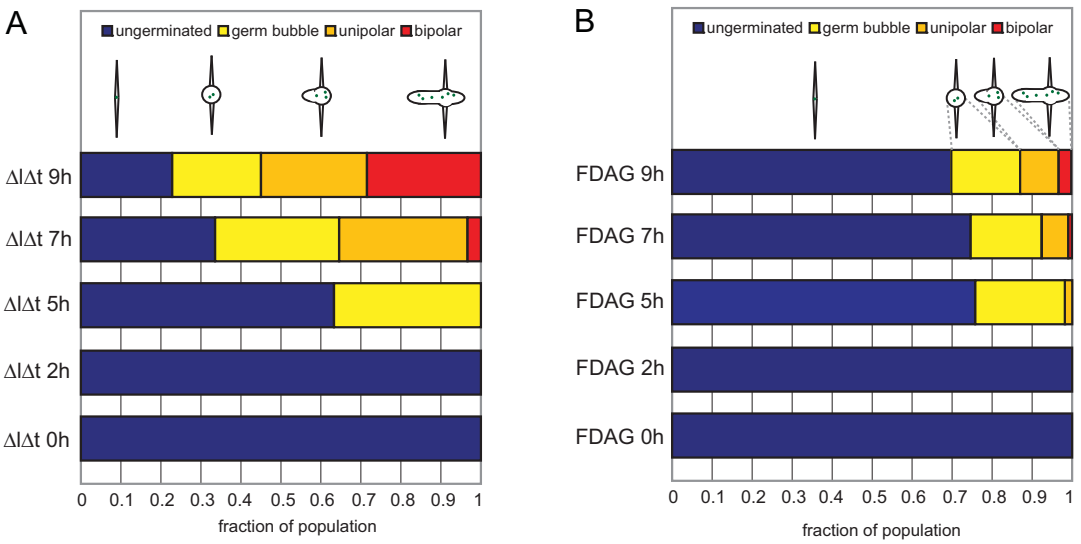


Figure A2-2 Comparison of developmental stages (A) in the laboratory strain ΔIΔt and (B) the natural isolate FDAG

strain. 70% of the spores had no trace of spore bubble 9 h post inoculation compared to 20% observed in the laboratory strain experiment (figure A2-2). Presently it is not clear whether this is due to reduced spore viability or delayed germination. Despite these differences in developmental stages, we observed in general comparable gene expression patterns, albeit the absolute expression levels differed sometimes. Higher expression levels in the sequenced laboratory strain might indicate base pair changes in the natural isolate. On the other hand, the Affymetrix array design with 16 probes per gene is very robust against outliers. Base pair changes in one or two probes very likely do not influence the overall signal intensity. Increased expression levels in the natural isolate though, can only be explained by increased gene expression. Here is an example: The natural isolate produces increased amounts of riboflavin which is apparent from its strong yellow colour (chapter 1 figure 1-4 B). The *S. cerevisiae* homolog of AgRIB3, ScRIB3 encodes a 3,4-dihydroxy-2-butanone-4-phosphate synthase which is required for riboflavin biosynthesis from ribulose-5-phosphate. ScRib3 also has an unrelated function in mitochondrial respiration. AgRIB3 was up-regulated from a signal intensity of 8900 to 19'842, which constitutes a 2.2-fold change. Other genes involved in riboflavin biosynthesis were also moderately up-regulated.

Based on comparable global gene expression, we confirmed earlier species assignment based on rDNA sequences (Fred Dietrich, personal communication) that the new natural isolate is indeed an *A. gossypii* strain. The genome of the natural isolate is presently sequenced allowing the investigation of intra-species differences in gene sequence, gene content, gene order and gene expression.

TABLE A2-2: Changes in activity of predefined groups in a new natural isolate (FDAG) compared with the strain used for functional genomics ($\Delta I \Delta t$)

	CATEGORY	ASPECT	t-VALUE	E-VALUE	MEAN	ORFs
UP-REGULATED	sporulation (sensu Fungi)	P	11.49	< 1.0E-15	1.57	19
	cellular component unknown	C	11.41	< 1.0E-15	0.33	364
	cell wall organization and biogenesis	P	5.82	8.20E-06	0.33	97
	plasma membrane	C	4.66	4.40E-03	0.21	122
DOWN-REGULATED	cell	C	-11.09	< 1.0E-15	0.02	3631
	intracellular	C	-10.57	< 1.0E-15	0.09	3388
	ribosome biogenesis and assembly	P	-7.14	1.30E-09	-0.26	188
	metabolism	P	-5.97	3.30E-06	-0.03	2282
	cellular component	C	-5.29	1.70E-04	0.02	4013
	ribosome	C	-4.26	2.80E-02	-0.16	192

TABLE A2-1: TOP DIFFERENTIALLY EXPRESSED GENES IN A NATURAL ISOLATE (FDAG) COMPARED TO THE STRAIN USED FOR FUNCTIONAL GENOMICS (ΔIΔt)

	SYSTEMATIC NAME	COMMON NAME	EXP FDAG	EXP ΔIΔt	LOG2 (ratio)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGICAL PROCESS	CELLULAR COMPONENT
Genes expressed at higher levels in the natural isolate									
UPREGULATED	AGL351W	GAS1	1366.50	69.54	4.30	Beta-1,3-glucanosyltransferase, required for cell wall assembly; localizes to the cell surface via a glycosylphosphatidylinositol (GPI) anchor	cell wall organization and biogenesis	1,3-beta-glucanosyltransferase activity	mitochondrion
	ABR156W	OPT1	943.35	51.28	4.20	Proton-coupled oligopeptide transporter of the plasma membrane; also transports glutathione and phytochelatin; member of the OPT family	sulfur metabolic process	oligopeptide transporter activity	endoplasmic reticulum
	AFL207C	HXT4/HXT6/ HXT7	292.25	16.16	4.18	High-affinity glucose transporters of the major facilitator superfamily	hexose transport	glucose transmembrane transporter activity	plasma membrane
	AFR401W	DIT1	398.25	24.52	4.02	Sporulation-specific enzyme required for spore wall maturation, involved in the production of a soluble LL-dityrosine-containing precursor of the spore wall; transcripts accumulate at the time of prospore enclosure	spore wall assembly (sensu Fungi)	catalytic activity	-
	ADL292C		159.00	11.20	3.83	Putative protein of unknown function with similarity to acyl-carrier-protein reductases	-	-	-
	ACL182C	GAS2	323.95	26.19	3.63	1,3-beta-glucanosyltransferase, involved with Gas4p in spore wall assembly; has similarity to Gas1p	spore wall assembly (sensu Fungi)	1,3-beta-glucanosyltransferase activity	cytoplasm
	ABR223C		98.09	8.92	3.46	Putative protein of unknown function	-	-	-
	AGR195W	DTR1	2188.50	228.35	3.26	Multidrug resistance dityrosine transporter of the major facilitator superfamily, essential for spore wall synthesis, facilitates the translocation of bisformyl dityrosine through the prospore membrane	spore wall assembly (sensu Fungi)	multidrug transporter activity	prospore membrane
	ACR227W	SKT5/SHC1	665.65	70.43	3.24	Activator of Chs3p (chitin synthase III), recruits Chs3p to the bud neck via interaction with Bni4p; has similarity to Shc1p, which activates Chs3p during sporulation	response to osmotic stress	enzyme activator activity	cellular bud neck
	AEL110W	PIR1/HSP15	7165.50	766.20	3.23	O-glycosylated protein required for cell wall stability; attached to the cell wall via beta-1,3-glucan	cell wall organization and biogenesis	structural constituent of cell wall	chitin- and beta-glucan-containing cell wall
	AEL189W	CHS3	88.66	9.56	3.21	Chitin synthase III, catalyzes the transfer of N-acetylglucosamine (GlcNAc) to chitin; required for synthesis of the majority of cell wall chitin, the chitin ring during bud emergence, and spore wall chitosan	spore wall assembly (sensu Fungi)	chitin synthase activity	cytoplasm
	ADL071C	RHR2/HOR2	4025.50	435.25	3.21	Constitutively expressed isoform of DL-glycerol-3-phosphatase; involved in glycerol biosynthesis, induced in response to both anaerobic and, along with the Hor2p/Gpp2p isoform, osmotic stress	response to drug	glycerol-1-phosphatase activity	cytoplasm
	ABL156C	FMP45	591.00	64.86	3.19	Integral membrane protein localized to mitochondria (untagged protein) and eisosomes, immobile patches at the cortex associated with endocytosis; sporulation and sphingolipid content are altered in mutants; has homologs SUR7 and YNL194C	cell wall organization and biogenesis	-	mitochondrion
	ACL009C	CSI2	415.75	46.25	3.17	structural component of the chitin synthase 3 complex	-	-	vacuole, cell cycle-correlated morphology
	ADL175W	GAS4	150.90	17.06	3.15	1,3-beta-glucanosyltransferase, involved with Gas2p in spore wall assembly; has similarity to Gas1p; localizes to the cell wall	spore wall assembly (sensu Fungi)	1,3-beta-glucanosyltransferase activity	chitin- and beta-glucan-containing cell wall

	SYSTEMATIC NAME	COMMON NAME	EXP FDAG	EXP ΔIΔt	LOG2 (ratio)	DESCRIPTION OF *S. CEREVISIAE* HOMOLOG	MOLECULAR FUNCTION	BIOLOGICAL PROCESS	CELLULAR COMPONENT
Genes expressed at lower levels in the natural isolate									
DOWNREGULATED	AEL190W	CHS2	102.71	491.40	-2.26	Chitin synthase II, requires activation from zymogenic form in order to catalyze the transfer of N-acetylglucosamine (GlcNAc) to chitin	cytokinesis	chitin synthase activity	cellular bud neck
	ADL294C	TRM3	82.73	356.70	-2.11	2'-O-ribose methyltransferase, catalyzes the ribose methylation of the guanosine nucleotide at position 18 of tRNAs	tRNA methylation	tRNA (guanosine) methyltransferase activity	cytoplasm
	ADR183C	NPL3	403.50	1293.50	-1.68	RNA-binding protein that carries poly(A)+ mRNA from the nucleus into the cytoplasm	pseudohyphal growth	mRNA binding	cytoplasm
	ADL173C	ZIM17	131.35	380.50	-1.53	Heat shock protein with a zinc finger motif; essential for protein import into mitochondria; may act with Pam18p to facilitate recognition and folding of imported proteins by Ssc1p (mtHSP70) in the mitochondrial matrix	protein folding	protein binding	mitochondrion
	AER351W	HEM4	2178.00	6286.50	-1.53	Uroporphyrinogen III synthase, catalyzes the conversion of hydroxymethylbilane to uroporphyrinogen III, the fourth step in the heme biosynthetic pathway	heme biosynthetic process	uroporphyrinogen-III synthase activity	-
	ADL297W		244.85	685.15	-1.48	Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the cytoplasm and to the nucleus	-	-	cytoplasm
	ACR274W	NOP12	207.80	557.90	-1.42	Nucleolar protein, required for pre-25S rRNA processing; contains an RNA recognition motif (RRM) and has similarity to Nop13p, Nsr1p, and putative orthologs in Drosophila and S. Pombe	rRNA metabolic process	RNA binding	nucleolus
	AFR602W	HXT14/HXT5	885.10	2359.50	-1.41	Protein with similarity to hexose transporter family members, expression is induced in low glucose and repressed in high glucose; the authentic, non-tagged protein is detected in highly purified mitochondria in high-throughput studies	hexose transport	galactose transmembrane transporter activity	mitochondrion
	AFR057C		70.95	188.15	-1.41	Putative protein of unknown function; cytoplasmically localized protein proposed to function in phospholipid binding; YBL107C is an essential gene	-	phospholipid binding	cytoplasm
	ACL191C	KSS1	253.50	623.50	-1.30	Mitogen-activated protein kinase (MAPK) involved in signal transduction pathways that control filamentous growth and pheromone response	invasive growth (sensu Saccharomyces)	MAP kinase activity	nucleus
	AEL191C	SCO1/SCO2	485.05	1183.00	-1.29	Copper-binding protein of the mitochondrial inner membrane, required for cytochrome c oxidase activity and respiration	protein complex assembly	thioredoxin peroxidase activity	mitochondrion
	ADL151W	PEX21/PEX18	1050.50	2542.50	-1.28	Peroxin required for targeting of peroxisomal matrix proteins containing PTS2; interacts with Pex7p; partially redundant with Pex18p	protein import into peroxisome matrix	protein binding	cytosol
	ADR087C	CPR6	397.20	940.45	-1.24	Peptidyl-prolyl cis-trans isomerase (cyclophilin), catalyzes the cis-trans isomerization of peptide bonds N-terminal to proline residues; binds to Hsp82p and contributes to chaperone activity	protein folding	unfolded protein binding	cytoplasm
	ADL295W	RRP42	90.99	210.10	-1.21	Protein involved in rRNA processing; component of the exosome 3->5 exonuclease complex with Rrp4p, Rrp41p, Rrp43p and Dis3p	35S primary transcript processing	3'-5'-exoribonuclease activity	nuclear exosome (RNase complex)
	AER451W	THI13	88.77	201.90	-1.19	Protein involved in synthesis of the thiamine precursor hydroxymethylpyrimidine (HMP); member of a subtelomeric gene family including THI5, THI11, THI12, and THI13	thiamin biosynthetic process	-	-

Appendix 3: Transcription data assists in improving gene annotation

NOHBY702 was identified as a new *A. gossypii* cell-wall protein that shares homology with superoxide dismutases (chapter 5). The story about NOHBY702 (AGL034C) is an example how transcription data can assist in improving gene annotation. The probe set originally designed to target NOHBY702 (AGL034C), Ag004677_at, showed no expression at 18 h (figure A3-1 A, next page). But for the adjacent probe set, Ag003734_at, very high signal intensities were measured. The two probe sets followed a very similar expression pattern through time but at different absolute levels. This discrepancy led to careful re-investigation of the annotation. The final decision was that the open reading frame (ORF) of AGL034C needed to be elongated at the 3' end. The consequence was that an open reading frame annotated on the opposite strand, AGL035W, was deleted (Sylvia Voegeli, personal communication). Figure A3-1 B shows another example of annotation correction based on expression data. A previously overlooked homology of NOHBY412 to AHP1 was detected because the gene showed similar expression levels and pattern of expression as the thioredoxin TRX1/TRX2.

There are 190 probe sets that, according to the final *A. gossypii* genome annotation, do not match a gene. This is either due to changes in the *A. gossypii* genome annotation since 2001, when the *A. gossypii* Affymetrix array was designed or due to the use of different open reading frame (ORF) prediction algorithms used by Affymetrix and our laboratory. For most of the 190 probe sets no intensity signals are detected on the array. However, 78 probe sets have at least once a signal intensity value above 100 in any of the 16 laboratory strain arrays. 13 of these 78 probe sets are according to the Affymetrix nomenclature unspecific because not all probe sets are unique (*_r_at), fewer than the required 16 unique probes are used (*_i_at) or all probes exactly match multiple transcripts (*_s_at). Of the remaining 65 probe sets, 12 probe sets have intensity values above 100 in all 16 arrays and 29 probe sets have in at least half of the arrays intensity values above 100. A careful bioinformatics and experimental re-investigating of the sequence at loci where the following probe sets match might further improve the high quality annotation: Ag000023_at, Ag001509_at, Ag002119_at, Ag002292_at, Ag003706_at, Ag003828_at, Ag004518_at, Ag004552_at, Ag004610_at, Ag004632_at, Ag004675_at, Ag004683_at.

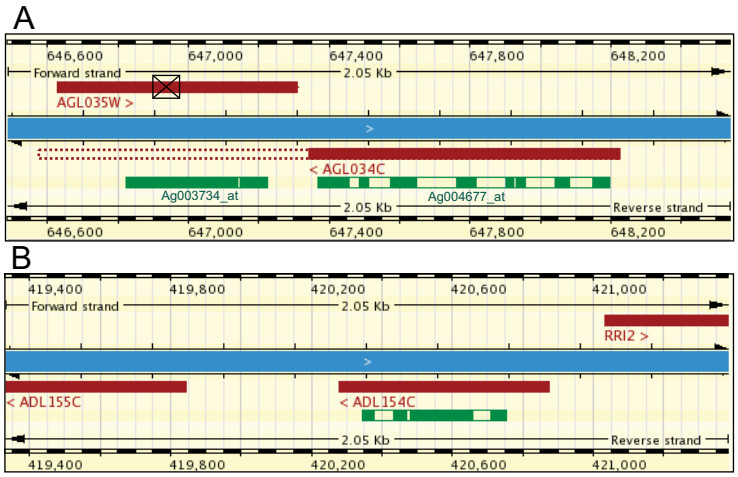


Figure A3-1Transcription data assists in improving gene annotation.

(A) AGL034C (NOHBY702) has recently been identified as an abundant cell wall protein in 18 h old mycelium by mass spec (Piet de Groot, personal communication). The probe set originally designed to target AGL034C (Ag004677_at) showed no expression at 18 h. But the adjacent probe set Ag003734_at gave very high signal intensities. This discrepancy led to careful re-investigation of the annotation. The open reading frame of AGL034C was elongated (red dashed line) and the previously annotated AGL035W was removed (Sylvia Voegeli, personal communication). (B) ADL154C was previously annotated as a gene with no homolog in S. cerevisiae (NOHBY412). The annotation of ADL154C was re-investigated because this gene was found to be among the 20 most highly expressed genes in A. gossypii in spores at 0 h. We found that ADL154C encodes the peroxiredoxin AHP1 (Peter Philippsen, personal communication).

References

References

Akaza, Y., A. Tsuji, et al. (2002). «Analysis of the gene encoding copper/zinc superoxide dismutase homolog in Dictyostelium discoideum.» Biol Pharm Bull **25**(12): 1528-32.

Bao, M. Z., M. A. Schwartz, et al. (2004). "Pheromone-dependent destruction of the Tec1 transcription factor is required for MAP kinase signaling specificity in yeast." Cell **119**(7): 991-1000.

Bardwell, J. C., P. Regnier, et al. (1989). "Autoregulation of RNase III operon by mRNA processing." Embo J **8**(11): 3401-7.

Batra (1973). "Nematosporaceae (hemiascomycetidae): taxonomy, pathogenicity, distribution and vector relations." USDA Technol Bull **1469**(1-71).

Benjamini (1995). "Controlling the false discovery rate: a practical and powerful approach to multiple testing. ." J. Roy. Statist. Soc. Ser. **B 57**: 289-300.

Bleasby (2000). "The fuzznuc program."

Bolstad, B. M., R. A. Irizarry, et al. (2003). "A comparison of normalization methods for high density oligonucleotide array data based on variance and bias." Bioinformatics **19**(2): 185-93.

Boorsma, A., B. C. Foat, et al. (2005). "T-profiler: scoring the activity of predefined groups of genes using gene expression data." Nucleic Acids Res **33**(Web Server issue): W592-5.

Boveris, A., A. O. Stoppani, et al. (1978). "Superoxide anion production and trypanocidal action of naphthoquinones on Trypanosoma cruzi." Comp Biochem Physiol C **61 C**(2): 327-9.

Breakspear, A. and M. Momany (2007). "The first fifty microarray studies in filamentous fungi." Microbiology **153**(Pt 1): 7-15.

Brengues, M., L. Pintard, et al. (2002). "mRNA decay is rapidly induced after spore germination of Saccharomyces cerevisiae." J Biol Chem **277**(43): 40505-12.

Briza, P., M. Eckerstorfer, et al. (1994). "The sporulation-specific enzymes encoded by the DIT1 and DIT2 genes catalyze a two-step reaction leading to a soluble LL-dityrosine-containing precursor of the yeast spore wall." Proc Natl Acad Sci U S A **91**(10): 4524-8.

Cabib, E. and A. Duran (2005). "Synthase III-dependent chitin is bound to different acceptors depending on location on the cell wall of budding yeast." J Biol Chem **280**(10): 9170-9.

Caro, L. H., G. J. Smits, et al. (1998). "Transcription of multiple cell wall protein-encoding genes in Saccharomyces cerevisiae is differentially regulated during the cell cycle." FEMS Microbiol Lett **161**(2): 345-9.

Casteilla, L., M. Rigoulet, et al. (2001). "Mitochondrial ROS metabolism: modulation by uncoupling proteins." IUBMB Life **52**(3-5): 181-8.

Chou, S., S. Lane, et al. (2006). "Regulation of mating and filamentation genes by two distinct Ste12 complexes in Saccharomyces cerevisiae." Mol Cell Biol **26**(13): 4794-805.

Cleves, A. E., D. N. Cooper, et al. (1996). "A new pathway for protein export in Saccharomyces cerevisiae." J Cell Biol **133**(5): 1017-26.

Copois, V., F. Bibeau, et al. (2007). "Impact of RNA degradation on gene expression profiles: assessment of different methods to reliably determine RNA quality." J Biotechnol **127**(4): 549-59.

Cullen, P. J., W. Sabbagh, Jr., et al. (2004). «A signaling mucin at the head of the Cdc42- and MAPK-dependent filamentous growth pathway in yeast.» Genes Dev **18**(14): 1695-708.

DeRisi, J. L., V. R. Iyer, et al. (1997). «Exploring the metabolic and genetic control of gene expression on a genomic scale.» Science **278**(5338): 680-6.

Dietrich, F. S., S. Voegeli, et al. (2004). "The Ashbya gossypii genome as a tool for mapping the ancient Saccharomyces cerevisiae genome." Science **304**(5668): 304-7.

Fleischer, T. C., C. M. Weaver, et al. (2006). "Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes." Genes Dev **20**(10): 1294-307.

Fromont-Racine, M., B. Senger, et al. (2003). "Ribosome assembly in eukaryotes." Gene **313**: 17-42.

Futcher, B., G. I. Latter, et al. (1999). "A sampling of the yeast proteome." Mol Cell Biol **19**(11): 7357-68.

Galagan, J. E., S. E. Calvo, et al. (2003). "The genome sequence of the filamentous fungus Neurospora crassa." Nature **422**(6934): 859-68.

Gancedo, J. M. (1992). «Carbon catabolite repression in yeast.» Eur J Biochem **206**(2): 297-313.

Gattiker, A., C. Niederhauser-Wiederkehr, et al. (2007). «The GermOnline cross-species systems browser provides comprehensive information on genes and gene products relevant for sexual reproduction.» Nucleic Acids Res **35**(Database issue): D457-62.

Gladfelter, A. S., A. K. Hungerbuehler, et al. (2006). "Asynchronous nuclear division cycles in multinucleated cells." J Cell Biol **172**(3): 347-62.

Guillermond, M. A. (1998). «Recherches sur quelques ascomycetes inférieurs isolés de la stigmatomycose des graines de cotonnier.» Revue Génerale de Botanique **40**: 328-342, 327-414, 475-485, 555-574, 606-624, 690-704.

Guldener, U., K. Y. Seong, et al. (2006). «Development of a Fusarium graminearum Affymetrix GeneChip for profiling fungal gene expression in vitro and in planta.» Fungal Genet Biol **43**(5): 316-25.

Gygi, S. P., Y. Rochon, et al. (1999). «Correlation between protein and mRNA abundance in yeast.» Mol Cell Biol **19**(3): 1720-30.

Hack (2004). "ntegrated transcriptome and proteome data: The challenges ahead." Briefings in Functional Genomics and Proteomics **3**(3): 212-219.

Hardwick, J. S., F. G. Kuruvilla, et al. (1999). "Rapamycin-modulated transcription defines the subset of nutrient-sensitive signaling pathways directly controlled by the Tor proteins." Proc Natl Acad Sci U S A **96**(26): 14866-70.

Harris, K., R. E. Lamson, et al. (2001). "Role of scaffolds in MAP kinase pathway specificity revealed by custom design of pathway-dedicated signaling proteins." Curr Biol **11**(23): 1815-24.

Hermida, L., O. Schaad, et al. (2006). "MIMAS: an innovative tool for network-based high density oligonucleotide microarray data management and annotation." BMC Bioinformatics **7**: 190.

Hoepfner, D., A. Brachat, et al. (2000). "Time-lapse video microscopy analysis reveals astral microtubule detachment in the yeast spindle pole mutant cnm67." Mol Biol Cell **11**(4): 1197-211.

Honma, Y., A. Kitamura, et al. (2006). “TOR regulates late steps of ribosome maturation in the nucleoplasm via Nog1 in response to nutrients.” Embo J **25**(16): 3832-42.

Hovatta (2005). “DNA Microarray Data Analysis.”

Huber, W., A. von Heydebreck, et al. (2002). “Variance stabilization applied to microarray data calibration and to the quantification of differential expression.” Bioinformatics **18 Suppl 1**: S96-104.

Hughes, T. R., C. J. Roberts, et al. (2000). “Widespread aneuploidy revealed by DNA microarray expression profiling.” Nat Genet **25**(3): 333-7.

Humphrey-Smith, I., Cordwell, S.J. and Blackstock, W. (1997). “Proteome research: complementarity and limitations with respect to the RNA and DNA worlds.” Electrophoresis **18**: 1217–1242.

Idnurm, A. and B. J. Howlett (2002). “Isocitrate lyase is essential for pathogenicity of the fungus Leptosphaeria maculans to canola (Brassica napus).” Eukaryot Cell **1**(5): 719-24.

Irizarry, R. A., B. Hobbs, et al. (2003). “Exploration, normalization, and summaries of high density oligonucleotide array probe level data.” Biostatistics **4**(2): 249-64.

Kaeberlein, M., A. A. Andalis, et al. (2002). “High osmolarity extends life span in Saccharomyces cerevisiae by a mechanism related to calorie restriction.” Mol Cell Biol **22**(22): 8056-66.

Kasuga, T., J. P. Townsend, et al. (2005). “Long-oligomer microarray profiling in Neurospora crassa reveals the transcriptional program underlying biochemical and physiological events of conidial germination.” Nucleic Acids Res **33**(20): 6469-85.

Kaufmann (2007). “Polarized growth and septation in the filamentous ascomycete Ashbya gossypii analyzed by life cell imaging.” PhD thesis.

Kerr, M. K., M. Martin, et al. (2000). “Analysis of variance for gene expression microarray data.” J Comput Biol **7**(6): 819-37.

Klein, C. J., L. Olsson, et al. (1998). “Glucose control in Saccharomyces cerevisiae: the role of Mig1 in metabolic functions.” Microbiology **144 (Pt 1)**: 13-24.

Klis, F. M., M. de Jong, et al. (2007). «Extraction of cell surface-associated proteins from living yeast cells.» Yeast **24**(4): 253-8.

Knechtle, P., F. Dietrich, et al. (2003). “Maximal polar growth potential depends on the polarisome component AgSpa2 in the filamentous fungus Ashbya gossypii.” Mol Biol Cell **14**(10): 4140-54.

Knechtle, P., J. Wendland, et al. (2006). “The SH3/PH domain protein AgBoi1/2 collaborates with the Rho-type GTPase AgRho3 to prevent nonpolar growth at hyphal tips of Ashbya gossypii.” Eukaryot Cell **5**(10): 1635-47.

Lee, P. D., R. Sladek, et al. (2002). “Control genes and variability: absence of ubiquitous reference transcripts in diverse mammalian expression studies.” Genome Res **12**(2): 292-7.

Lenaz, G. (2001). “The mitochondrial production of reactive oxygen species: mechanisms and implications in human pathology.” IUBMB Life **52**(3-5): 159-64.

Lorenz, M. C. and G. R. Fink (2001). “The glyoxylate cycle is required for fungal virulence.” Nature **412**(6842): 83-6.

Lott, T., S. Gorman, et al. (1981). “Superoxide dismutase in Didymium iridis: characterization and changes in activity during senescence and sporulation.” Mech

Ageing Dev **17**(2): 119-30.

MacKenzie, D. A., T. Guillemette, et al. (2005). "UPR-independent dithiothreitol stress-induced genes in Aspergillus niger." Mol Genet Genomics **274**(4): 410-8.

Mercado, J. J., R. Smith, et al. (1994). "The levels of yeast gluconeogenic mRNAs respond to environmental factors." Eur J Biochem **224**(2): 473-81.

Miyajima, I., M. Nakafuku, et al. (1987). "GPA1, a haploid-specific essential gene, encodes a yeast homolog of mammalian G protein which may be involved in mating factor signal transduction." Cell **50**(7): 1011-9.

Narasipura, S. D., J. G. Ault, et al. (2003). "Characterization of Cu,Zn superoxide dismutase (SOD1) gene knock-out mutant of Cryptococcus neoformans var. gattii: role in biology and virulence." Mol Microbiol **47**(6): 1681-94.

Nowell, A. a. (1926). Annals of Botany **XL**.

Pavlidis, P. (2003). "Using ANOVA for gene selection from microarray studies of the nervous system." Methods **31**(4): 282-9.

Powers, T. and P. Walter (1999). "Regulation of ribosome biogenesis by the rapamycin-sensitive TOR-signaling pathway in Saccharomyces cerevisiae." Mol Biol Cell **10**(4): 987-1000.

Prillinger, H., W. Schweigkofler, et al. (1997). "Phytopathogenic filamentous (Ashbya, Eremothecium) and dimorphic fungi (Holleya, Nematospora) with needle-shaped ascospores as new members within the Saccharomycetaceae." Yeast **13**(10): 945-60.

Primig, M., R. M. Williams, et al. (2000). "The core meiotic transcriptome in budding yeasts." Nat Genet **26**(4): 415-23.

Quackenbush, J. (2002). "Microarray data normalization and transformation." Nat Genet **32 Suppl**: 496-501.

Ram, A. F. and F. M. Klis (2006). "Identification of fungal cell wall mutants using susceptibility assays based on Calcofluor white and Congo red." Nat Protoc **1**(5): 2253-6.

Rodriguez-Pena, J. M., C. Rodriguez, et al. (2002). «Mechanisms for targeting of the Saccharomyces cerevisiae GPI-anchored cell wall protein Crh2p to polarised growth sites.» J Cell Sci **115**(Pt 12): 2549-58.

Ronne, H. (1995). "Glucose repression in fungi." Trends Genet **11**(1): 12-7.

Russel, S. a. (2001).

Sachs, M. S. and C. Yanofsky (1991). "Developmental expression of genes involved in conidiation and amino acid biosynthesis in Neurospora crassa." Dev Biol **148**(1): 117-28.

Schena (1996). "DNA Microarrays: A Practical Approach." Bioassays.

Schmelzle, T., T. Beck, et al. (2004). "Activation of the RAS/cyclic AMP pathway suppresses a TOR deficiency in yeast." Mol Cell Biol **24**(1): 338-51.

Schmitt, M. E., T. A. Brown, et al. (1990). "A rapid and simple method for preparation of RNA from Saccharomyces cerevisiae." Nucleic Acids Res **18**(10): 3091-2.

Schweingruber, A. M., J. Dlugonski, et al. (1991). «Thiamine in Schizosaccharomyces pombe: dephosphorylation, intracellular pool, biosynthesis and transport.» Curr Genet **19**(4): 249-54.

Shamji, A. F., F. G. Kuruvilla, et al. (2000). «Partitioning the transcriptional program induced by rapamycin among the effectors of the Tor proteins.» Curr Biol **10**(24): 1574-81.

Skory, C. D. and H. A. Wyckoff (1997). “Fabrication of an autoclavable one-way valve for venting fermentation gases.” Biotechniques **22**(1): 82.

Smits, G. J., L. R. Schenkman, et al. (2006). “Role of cell cycle-regulated expression in the localized incorporation of cell wall proteins in yeast.” Mol Biol Cell **17**(7): 3267-80.

Stahmann, K. P., H. N. Arst, Jr., et al. (2001). “Riboflavin, overproduced during sporulation of Ashbya gossypii, protects its hyaline spores against ultraviolet light.” Environ Microbiol **3**(9): 545-50.

Starmer, W. T., P. F. Ganter, et al. (1987). «The ecological role of killer yeasts in natural communities of yeasts.» Can J Microbiol **33**(9): 783-96.

Sumita, T., T. Yoko-o, et al. (2005). “Comparison of cell wall localization among Pir family proteins and functional dissection of the region required for cell wall binding and bud scar recruitment of Pir1p.” Eukaryot Cell **4**(11): 1872-81.

Titz, B., S. Thomas, et al. (2006). “Transcriptional activators in yeast.” Nucleic Acids Res **34**(3): 955-67.

Turkey (1977). “Exploratory data analysis.” Addison-Wesley, Reading, MA.

Valdivieso (1999). “Chitin synthases in yeast and fungi.” EXS **87**: 55-69.

Wickerham (1946). “The production of riboflavin by Ashbya gossypii.” Arch Biochem **9**(95-98).

Williams, R. M. (2001). “The fuzznuc program.”

Wu, J., N. Zhang, et al. (2004). “Global analysis of nutrient control of gene expression in Saccharomyces cerevisiae during growth and starvation.” Proc Natl Acad Sci U S A **101**(9): 3148-53.

Yin, Z., R. J. Smith, et al. (1996). “Multiple signalling pathways trigger the exquisite sensitivity of yeast gluconeogenic mRNAs to glucose.” Mol Microbiol **20**(4): 751-64.

Young, M. E., T. S. Karpova, et al. (2002). “The Sur7p family defines novel cortical domains in Saccharomyces cerevisiae, affects sphingolipid metabolism, and is involved in sporulation.” Mol Cell Biol **22**(3): 927-34.

Zaragoza, D., A. Ghavidel, et al. (1998). “Rapamycin induces the G0 program of transcriptional repression in yeast by interfering with the TOR signaling pathway.” Mol Cell Biol **18**(8): 4463-70.

List of abbreviations

List of abbreviations

AFM Ashbya full medium
ATP Adenosine triphosphate
NER nucleotide excision repair
BLAST basic local alignment search tool
cRNA complementary ribonucleic acid
DNA deoxyribonucleic acid
GO GeneOntology
IVT in vitro transcription
kb kilobases
min minutes
NCBI National Center for Biotechnology Information
MM mismatch
wt wild-type
ml milliliter
mRNA messenger ribonucleic acid
ORF open reading frame
pH pondus hydrogenii
PM perfect match
RMA Robust Multi-array Analysis from Bioconductor
rpm revolutions per minute
rRNA ribosomal ribonucleic acid
SAPE streptavidine phycoerythrine conjugate
sec second
tRNA transfer ribonucleic acid
µg microgram
µl microliter
wt wild-type

Microarray glossary

Microarray glossary

Array, chip, probe array, microarray: A small (a few centimeters on each side) glass or other solid surface on which thousands of immobilized oligonucleotide probes have been synthesized or robotically deposited in a predetermined array, so that automated recording of fluorescence from each of the positions may be easily determined. Affymetrix GeneChip® probe arrays are synthesized *in situ* on a glass surface.

Affybatch: An Affybatch is an object in which the raw data from an Affymetrix gene chip experiment as well as accompanying experimental annotation can be stored.

Background: Unwanted intensity observed on an array. Sources of background include fluorescence of the glass, stray DNA, and many other sources. The background is dependent on the sequence of the oligonucleotide.

CDF file: The CDF file describes the layout of the chip. The probe set information in the CEL file by itself is not particularly useful as there is no indication in the file as to which probe set a probe belongs. This information is stored in the CDF library file associated with a chip type.

CEL file: The information about each probe on the chip is extracted from the image data by the Affymetrix image analysis software. The information is stored in the CEL file. The CEL file contains information about where each probe is on the chip and also the intensity values for the probe.

CHP file: The CHP file contains the results of the experiment. These include the average signal measures for each probe set as determined by the Affymetrix software and information about which probe sets are called as present, absent or marginal and the p-values for these calls.

DAT file: The raw image data from chip scanner is saved in the DAT file.

Environment: In the statistical program R, the environment associates information to the CEL files, e.g. which CDF file has to be used.

Feature: A distinct site on the array that has millions of copies of the same 25-mer representing a particular sequence.

False positive: When an entity is detected and it is not present. For microarrays, there are two types often discussed. First, when an absent transcript is reflected as a significantly present output. Second, when a differential change is falsely indicated when there is no change in an expression level.

False negative: When an entity is not detected, and it is present. For microarrays, there are two types of false negative rates. First, when a transcript is indicated as being absent when it is expressed in the sample. Second, when a (specified) differential change is falsely indicated as unchanged.

R: Statistical environment and programming language (The R foundation for statistical computing, 2006)

RMA (Robust multichip average) pre-processing for Affymetrix chips

Robustness: production of reasonably accurate results in the presence of outliers.

Sensitivity: The ability to detect an entity when it is present. The ability to detect a (specified) transcript as being present in a sample and the ability to detect a differential change between two conditions when there is a change.

Signal: A summary value for the observed

intensities in a probe set reflecting a common transcript.

Specificity: The ability to exclude an entity when it is not there. For microarrays there are two types of specificity. When an absent transcript is correctly indicated as absent and when an unchanged expression level is correctly indicated as unchanged.

Probe: A labeled, single-stranded DNA or RNA molecule of specific base sequence which is used to detect the complementary base sequence by hybridization. At Affymetrix, probe refers to unlabeled oligonucleotides synthesized on a GeneChip® probe array.

Probe set: A group of probes that together detect transcript(s).A probe set includes a series of probes and represents a transcript. The current Affymetrix design uses 11 probes (or probe pairs) per gene. Earlier designs (e.g. sySYNG001a) use 16 probes per gene.

Target: refers to the material that hybridizes to the GeneChip® array. Target is a biotin-labeled product prepared from a RNA sample.

Target sequence: The relatively small (≦600 nucleotide) sequence segment from the 3’ end of the consensus or the exemplar sequence; used to design the probes for any given probe set.

Acknowledgements

Acknowledgements

First of all I would like to thank Prof. Peter Philippsen for giving me the opportunity to work in his laboratory, for providing me with a great working environment and for sharing his scientific expertise. I appreciated his constant support, his patience, his generosity and the tons of nice chocolates and pastries I received during my PhD. I also enjoyed the freedom that I had in finding a PhD thesis project and in developing and contributing own ideas. Furthermore I greatly appreciated the opportunity to participate in stimulating meetings and courses in Basel and abroad.

I am very thankful to Dr. Lars Steinmetz for the supporting interest, for critical reading of this thesis and for being a member of my thesis board.

I would like to thank Prof. Frans M. Klis from the University of Amsterdam for stimulating discussions about the *A. gossypii* cell wall, for sending me interesting papers and for introducing me to t-profiler, a useful tool for microarray data analysis.

Piet de Groot from the University of Amsterdam contributed the cell wall proteome data and identified the new *A. gossypii* cell wall proteins. Working with Piet has always been a pleasure.

I am grateful to Prof. Mike Primig for contributing his scientific expertise in the beginning of this study. He has been a driving force in realizing this project.

Many thanks to Fred Dietrich who created the sequence database for the Affymetrix chip. The sySYNG001a chip was designed by Affymetrix in collaboration with Syngenta.

Many thanks to Haoping Liu who kindy shared with us her promoter data prior to publication.

I am very thankful to Philippe Demougin from the former Prof. Primig laboratory for being a reliable and competent college, for his great sense of humor, for teaching me RNA and chip handling and for allowing me to work in the microarray facility at any time, day and night.

Lise Bernard-Granger and Isabelle Tanneau were two talented and motivated summer students in our laboratory and helped me with the transcription factor deletions.

Many thanks to Michael Köhli for giving me a crash course in InDesign.

My warm thanks to all former and current members of the Philippsen lab for the good discussions, the great support, the many “thai” lunches and the good times we shared:
Claudia Birrer, Dr. Kamila Boudier , Dr. Sophie Brachat, Mark Finlayson, Dr. Amy Gladfelter, Virginie Galati, Dr. Sandrine Grava, Dr. Hanspeter Helfer, Dr. Dominic Hoepfner, Dr. Katrin Hungerbühler, Dr. Kumiko Masai, Dr. Michael Köhli, Dr. Andreas Kaufmann, Dr. Philippe Knechtle, Dr. Philippe Laissue, Lars Molzahn, Dr. Florian Schärer, Dr. Hans-Peter Schmitz, Sylvia Voegeli and Jeannine Winkler. I would also like to thank the technical staff, particularly Brigitte Berglas, and our secretary Bettina Hersberger.

I am indebted to my mother who gave me the strength to achieve this work and supported me bravely as a single mom over all these years of study. Without her love and support, I would not have been able to accomplish all what I have today. I want to thank Daniel for his love, patience and comprehension.

Curriculum vitae

Curriculum vitae

(omitted for reasons of data protection)

Erklärung

Erklärung

Ich erkläre, dass ich diese Dissertation “Transcripional profiling of the model fungus *A. gossypii:* Comparison of life cycle stages and transcription factor deletions” nur mit der darin angegebenen Hilfe verfasst und bei keiner anderen Fakultät eingereicht habe.

Basel, Oktober 2007

Riccarda Rischatsch