SEXUAL DEVELOPMENT AND MEIOTIC SILENCING IN *Neurospora crassa*

A Dissertation

by

ANA VICTORIA SUESCÚN TORRES

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,   Rodolfo Aramayo
Committee Members,   Alan Pepper
                                  Kathryn Ryan
                                  Dorothy Shippen
Head of Department,   Thomas McKnight

August 2014

Major Subject: Microbiology

ABSTRACT

Meiotic silencing refers to the mechanism of silencing genes or chromosomes without a homologous counterpart (unpaired) during meiotic prophase I. Meiotic silencing has been described in several eukaryotes, including humans. Failure to complete meiotic silencing may be detrimental to the organism. There is a necessity for understanding the regulation of the process. *Neurospora* is a powerful model system to study gene silencing phenomena. Numerous genes have been determined to be involved in meiotic silencing in *Neurospora;* however, very little is known about the molecular mechanisms underlying the process. To understand the regulation of meiotic silencing, it is required to combine different approaches such as genetics, proteomics and biochemical analyses.

There is a need for introducing biochemical approaches to the study of meiotic silencing and other processes occurring during sexual development in *N. crassa*. However, protein extraction from sexual tissue is challenging due to the mechanical difficulties associated with disruption of sexual structure. I standardized a strategy that optimizes protein extraction from sexual tissue. Using this strategy, I studied protein-protein interactions among components of the meiotic silencing machinery and determine the proteome of sexual development.

I identified new protein interactions during meiotic silencing in *N. crassa*, and established protein-binding partners for the suppressor of meiotic silencing SMS-5. These interacting partners, PAF400 and Pianissimo represent new molecular components involved in the nuclear initial stage of the meiotic silencing mechanism. Interactions between SMS-5, PAF400, and Pianissimo may represent the connection between chromatin remodeling, DNA repair, signaling transduction pathways and meiotic silencing.

I describe the experiments and data analyses used to develop a comprehensive proteomics data set and a functional catalogue for *N. crassa* sexual development. I used a global proteomics approach and comparative protein functional analysis to investigate the potential molecular differences between two stages of sexual development in filamentous fungi. The data show that secondary metabolites biosynthesis and cellulase activity are required in fruiting body maturation. *N. crassa* functional catalogue of sexual development proteins will serve as a reference tool for further studies related to sexual development not only in *N. crassa*, but also in other filamentous ascomycetes.

DEDICATION

Many special people deserve to be mentioned in this page; without their sacrifices, support and unconditional love this work would never been possible. I dedicate this work to: my family back in Colombia, especially my grand mom Beatriz Rey de Torres, for her kindness; my niece Dana Camila, one of the biggest gifts and motivations in my life; my brother Carlos for believing in me. To my best friend and loving fiancé Cristian; and with all my heart to the memory of my mom, Nubia Marina Torres, who taught me the greatness of love and now she is my angel in heaven.

Tantas personas especiales merecen ser mencionadas en esta página; sin sus sacrificios, apoyo y amor incondicional este trabajo no hubiera sido posible. Dedico este trabajo a: mi familia en Colombia, especialmente a mi abuelita, Beatriz Rey de Torres, por su ternura y generosidad, mi sobrina Dana Camila, uno de los grandes regalos y motivaciones en mi vida; mi hermano, Carlos por creer en mí. A mi mejor amigo y prometido Cristian, mi ángel en la tierra; y con todo mi corazón dedico este trabajo a la memoria de mi mamita, Nubia Marina Torres, quien me enseñó la grandeza del amor y ahora es mi ángel en el cielo.

ACKNOWLEDGEMENTS

NOMENCLATURE

aRNA	Aberrant RNA
DSBs	Double-Strand Breaks
FunCat	Functional Catalogue
HMTase	Histone Methyltransferase
HR	Homologous Recombination
KAAS	KEGG Automatic Annotation Server
KEGG	Kyoto Encyclopedia of Genes and Genomes
MSCI	Meiotic Sex Chormosome Inactivation
MSUD	Meiotic Silencing by Unpaired DNA
MSUC	Meiotic Silencing of Unsynapsed Chromatin
MIPS	Munich Information Center for Protein Sequences
NHEJ	Non-Homologous Recombination End-Joining
NLS	Nuclear Localization Signal
PTGS	Post-transcriptional Gene Silencing
RIP	Repeat Induced Point Mutation
RBP	RNA-binding Protein
SC	Synaptonemal Complex
TORC2	Target of Rapamycin Complex 2
TE	Transposable Elements

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

From a biological perspective, the purpose for life of an organism is to reproduce and pass its genes on to the next generation, assuring the propagation of the species. There are two general mechanisms for reproduction: asexual and sexual. Sexual reproduction introduces genetic variation, which facilitates the adaptation to the constant changes in the environment. It is also a mechanism for purging deleterious mutations from the species. Therefore, sexual reproduction offers evolutionary and selective advantages over asexual reproduction (HORANDL 2009).

Sexual reproduction involves the fusion of two gametes, from different individuals, to form a zygote, through a process known as fertilization. Gametes, or sex cells, contain only one set of chromosomes (haploid cells) while the zygote, which receives one set from each parent, contains two sets of chromosomes (diploid cell). These complementary chromosomes are called homologous chromosomes. Therefore, sexual reproduction in eukaryotes, such as animals, plants and fungi relies on the precise reduction of chromosome number during gamete formation. Haploid gametes are produced by a special type of cell division called meiosis.

MEIOSIS

Meiosis was first described by Edouard van Beneden in 1883. He discovered that germs cells contained half the number of chromosomes as regular cells by performing cytological studies on female worms (HAMOIR 1992). Meiosis is a double-division cycle preceded by only one DNA replication event. Its primary purpose is to generate haploid daughter cells—gametes in animals and plants or spores in fungi—from a diploid progenitor. This reduction in the number of chromosomes is essential for ensuring continuity of the species (PETRONCZKI *et al.* 2003). Errors in meiosis result in the production of aneuploid gametes (*i.e.* gametes with an incorrect number of chromosomes), and abnormal chromosome number can cause defective or inviable progeny (PAGE and HAWLEY 2003).

The transition from diploid to haploid is achieved by two consecutive meiotic divisions. In meiosis I, or reductional division, homologous chromosomes separate. During meiosis II, or equational division, sister chromatids—identical DNA strands product of chromosome replication—separate. Therefore, four daughter haploid cells are formed from a unique parental diploid cell (PETRONCZKI *et al.* 2003).

Another important characteristic of meiosis is the genetic diversity that results from genetic recombination between DNA molecules of homologous chromosomes, called homologous recombination (HR). HR not only provides a potent source of genetic variation, which is crucial for speeding the evolution of

eukaryotic genomes, but also plays an important mechanical role by ensuring proper chromosome segregation at the first meiotic division. This process occurs in a very early stage of meiosis I known as prophase I (COHEN *et al.* 2006).

Prophase I

Prophase of the first meiotic division is the longest phase of meiosis. Many crucial molecular and cellular events, such as homolog pairing, synapsis and meiotic recombination, take place to ensure that homologous chromosomes accurately segregate (BAUDAT *et al.* 2013).

Prophase I is divided into five stages (Figure 1.1): 1) Leptotene, at this stage individual chromosomes, each consisting of two sister chromatids, change from diffuse chromatin to a condense stage, becoming visible strands within the nucleus. In most organisms, a programmed induction of DNA double-strand breaks (DSBs) is observed. These DSBs initiate pairing of homologs and meiotic recombination (SUN *et al.* 1989; MAHADEVAIAH *et al.* 2001). Axial elements of the synaptonemal complex (SC)—a protein complex that connects the homologous chromosomes—are assembled during the transition from leptotene to zygotene (PAGE and HAWLEY 2004). 2) Zygotene, at this stage homologous chromosome pairs line up and begin synapsis, forming bivalents, with the assistance of the central element of the SC (HARPER *et al.* 2004). Telomere clustering occurs at the nuclear envelope, called the bouquet configuration; the bouquet promotes initial homolog interactions and may be involved in pairing (ZICKLER and

3

KLECKNER 1998; YAMAMOTO and HIRAOKA 2001). 3) Pachytene, the longest stage of prophase I. At this stage the homologs are synapsed along their entire length, forming mature bivalents. Non-sister chromatids of the homologous chromosomes may exchange segments over regions of homology by HR, resulting in crossover events (GUILLON and DE MASSY 2002). 4) Diplotene, at this stage elements of SC start to degrade and homologous chromosomes repeal each other, but remain connected through to the final stage of prophase I at the chiasmata, where the recombination event occurs, (PAGE and HAWLEY 2004). 5) Diakinesis. Chromosomes condense further and much of the SC structure is lost at this stage. The sites of crossing over entangle together ensuring that homologs remain paired through to metaphase I (reviewed in (TSAI and MCKEE 2011; BAUDAT *et al.* 2013)). Several studies in different organisms have contributed to understanding the meiotic recombination process and the role it plays in holding homologous chromosomes together through to first meiotic division (KEENEY *et al.* 1997; DERNBURG *et al.* 1998; MCKIM *et al.* 1998; DAVIS and SMITH 2001; MAHADEVAIAH *et al.* 2001; PHADNIS *et al.* 2011; BAUDAT *et al.* 2013). However, less is known about the mechanism by which homologs recognize each other and synapse (YAMAMOTO and HIRAOKA 2001; TSAI and MCKEE 2011).
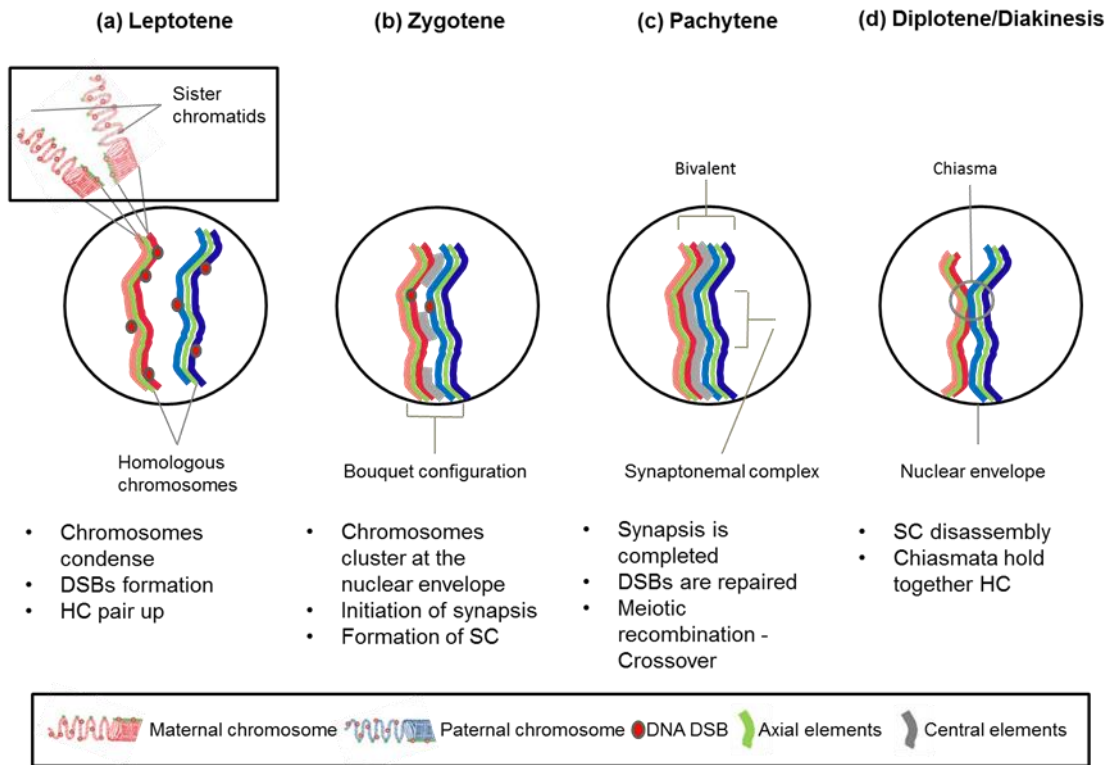
Figure 1.1 An overview of prophase I. During prophase of meiosis I pairing, synapsis and recombination between homologous chromosomes (HC) occur. Interactions between one pair of HC (red and blue) are schematically represented. Sister chromatids, which are products of DNA replication, are shown as different shades of red or blue. (a) During the leptotene stage chromosomes begin to condense (upper box), HC pair up, DNA double-strand breaks (DSBs) form (orange circles) and each pair of sister chromatids begins to assemble axial elements (AE) (green) of the synaptonemal complex (SC). (b) By the zygotene stage chromosomes are completely condensed and cluster at the nuclear envelope through the telomeres forming a bouquet configuration. HC begin synapsis via the central element (gray) to form a SC. (c) The beginning of the pachytene stage is marked by the completion of synapsis and formation of a mature bivalent. DSBs are repaired, with some of the breaks resolving into crossover events. (d) SC is disassembled during diplotene. Chiasmata resulting from the crossover events are observed during diplotene and diakinesis stages and play an important role by holding together HC through to anaphase I, when HC migrate to opposite cellular poles Figure adapted from (Burgoyne et al. 2009).

Recombination Events

Recombination is initiated by the formation of DSBs, followed by DNA damage repair and DSB processing. DSBs are resolved into homologous recombination (crossover) or non-homologous recombination end-joining (NHEJ). The evolutionary conserved topoisomerase, Spo11 in yeast and its homolog in other organisms, is responsible for the formation of meiotic DSBs (KEENEY *et al.* 1997; KEENEY 2001). In most eukaryotes, failure to induce DSBs is associated with meiotic recombination impairment, that may cause sterility (PAGE and HAWLEY 2004). However, in worms and flies, DSBs are not required for synapsis, indicating the existence of more than one mechanism to achieve recombination (DERNBURG *et al.* 1998; MCKIM *et al.* 1998).

Even though DSBs are essential for recombination, they have the potential to be deleterious for the cell. Therefore, they must be repaired before normal cell division proceeds. Initiation of synapsis facilitates DSB repair during zygotene and pachytene stages by allowing ccess to a DNA repair template in the homologous chromosomes (Figure 1.2) (PETRONCZKI *et al.* 2003). The majority of DSBs are resolved into non-homologous recombination. The remaining DSBs are resolved into crossover events forming chiasmata structures that hold homologous chromosomes together after recombination is completed and chromosomes are desynapsed (BORNER *et al.* 2004).
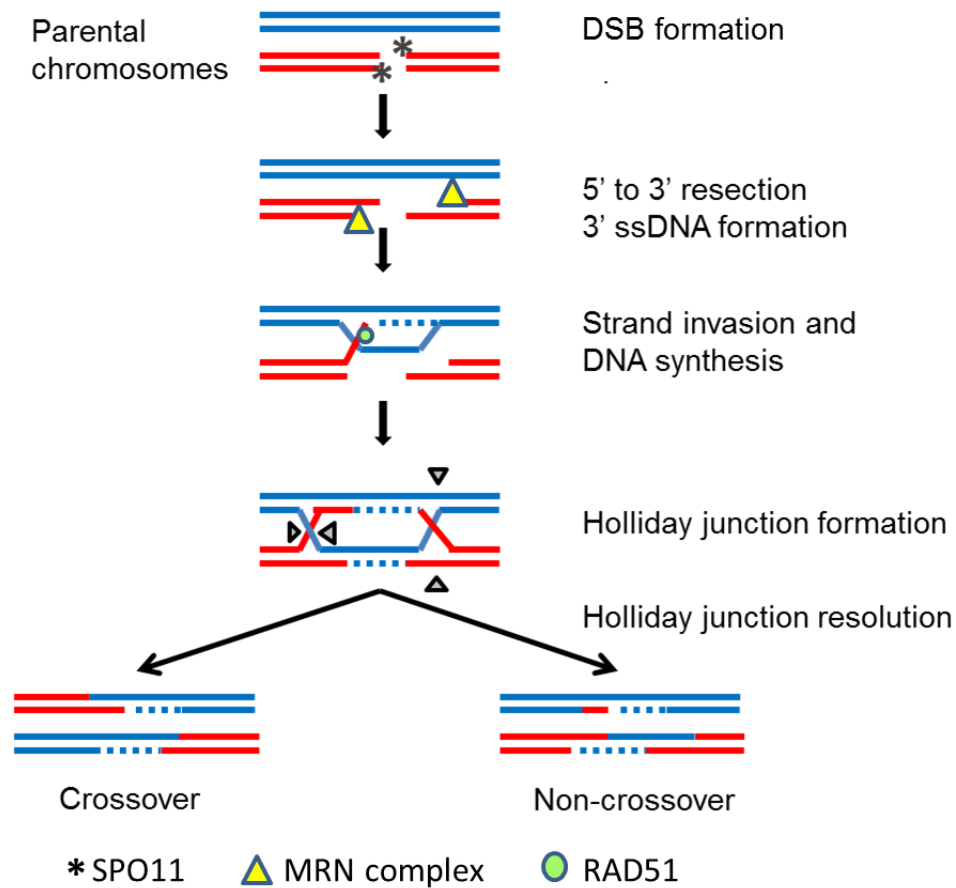
Figure 1.2 A simplified model of recombination events during meiosis. The conserved topoisomerase SPO11 is responsible for the formation of DNA double-strand breaks (DSBs), which are essential for initiation of meiotic recombination. DSBs stimulate phosphorylation of H2AX by ATM. The MRN complex (MRE11, RAD50, and NBS1) is required for DNA repair. After DSBs are formed in one of the homologs, MRN complex is recruited and the resulting breaks are resected from 5 'to 3', creating a free 3' single stranded DNA (ssDNA) end. RAD51 binds the ssDNA end and facilitates invasion of the intact homologous duplex where high sequence identity is found. DNA synthesis starts from the 3' end using the invaded DNA molecule as a template. Holliday junction structures are formed and resolved into crossover or non-crossover events. Figured adapted from (Phadnis et al. 2011)

Like a positive feedback mechanism, correct initial processing of DSBs ensures complete synapsis between homologs. Failure to process the breaks results in partial synapsis or unsynapsed chromosome regions that retain markers of unrepaired DSBs during zygotene, resulting in early meiotic arrest (MARCON and MOENS 2005).

Pachytene Checkpoint

Errors in chromosome synapsis are associated with sterility. This association has been attributed to the presence of surveillance mechanisms or checkpoints that detect problems at the different stages and promote the elimination of defective cells (COHEN *et al.* 2006). Even though sterility is the major consequence of asynapsis, these checkpoints are beneficial because they reduce the frequency of aneuploid gametes which generate chromosomally unbalanced zygotes (BURGOYNE *et al.* 2009).

Asynapsis interferes with the process of repairing DSBs, affecting genome integrity. Therefore, it is crucial that cell division does not proceed in the presence of this DNA damage. In many organisms, including budding yeast (ROEDER and BAILIS 2000), flies (GHABRIAL and SCHUPBACH 1999), worms (BHALLA and DERNBURG 2005) and mammals (ASHLEY *et al.* 2004), if a threshold of unrepaired DSBs is detected after pachytene stage, a pachytene checkpoint is activated and meiotic cells arrest or undergo programmed cell death (HOCHWAGEN and AMON 2006). Studies in the budding yeast *Saccharomyces*

cerevisiae have identified elements of the pachytene checkpoint machinery. Some of these elements are meioitic-specific proteins, and others are factors that play roles in DNA damage signaling in mitotic cells that have been specifically adapted to meiosis (ROEDER and BAILIS 2000). Homologs for most of these proteins have been found, including in humans, and these are also involved in the meiosis checkpoint. This observation underscores the conservation of this mechanism (MACQUEEN and HOCHWAGEN 2011). A number of these factors are recruited to the DSBs at zygotene and are retained into pachytene when chromosome regions fail to synapse. A common denominator of pachytene checkpoint regulation between different organisms is the activation of ATM/ATR. The checkpoint kinases ATM and ATR are key factors for signaling the presence of unrepaired DSBs. Once activated, they phosphorylate a large set of substrates to activate the DNA damage response (MACQUEEN and HOCHWAGEN 2011). After that, the DNA damage response protein BRCA1 and the phosphorylated form of the variant nucleosomal histone H2AX (γH2AX) are responsible for signal amplification. During the transition between zygotene and pachytene stages, BRCA1 and ATR proteins accumulate and spread along the unsynaped chromosomes (BURGOYNE *et al.* 2009).

In addition to the accumulation of unrepaired DSBs and checkpoint activation, unsynapsed chromosomal regions undergo a process of transcriptional repression (BAARENDS *et al.* 2005; TURNER *et al.* 2005). Asynapsis triggers the activation of meiotic silencing mechanisms, which are responsible

9

for the repression of gene expression. Therefore, there is a direct link between asynapsis and meiotic silencing (BURGOYNE *et al.* 2009). It has been proposed that transcriptional inactivation by meiotic silencing may contribute to the arrest of meiotic division at the pachytene stage by affecting genes necessary for meiotic progression (TURNER *et al.* 2005).

MEIOTIC SILENCING MECHANISMS

During meiosis, homologous chromosomes sense each other and pair, both critical steps that assure the correct chromosome alignment for genetic recombination. Meiotic silencing refers to the mechanism of transcriptional or post-transcriptional silencing of genes, genetic regions, or chromosomes without a homologous counterpart (unpaired) during meiotic prophase I.

In higher eukaryotes, meiotic silencing was initially reported as an exclusive phenomenon occurring only in the unpaired sex chromosomes in organisms with heteromorphic sex chromosomes. This phenomenon was called Meiotic Sex Chromosome Inactivation (MSCI) (MCKEE and HANDEL 1993; TURNER 2007). Some of the organisms involved are XO male nematodes (*e.g., Caenorhabditis elegans*) (BEAN *et al.* 2004), XY male mammals (*e.g., Mus musculus)* (HANDEL 2004), and ZW female birds (*e.g., Gallus gallus*) (SCHOENMAKERS *et al.* 2009). Afterwards, it was found that any unpaired chromosome region, even in autosomal chromosomes, is subjected to silencing during meiosis by a phenomenon named Meiotic Silencing of Unsynapsed

Chromatin (MSUC) (SCHIMENTI 2005; TURNER *et al.* 2005). Therefore, it was hypothesized that MSCI evolved as a consequence of MSUC, an evolutionarily conserved pathway and a more general meiotic silencing mechanism (TURNER *et al.* 2006).

MSUC resembles a silencing phenomenon that was originally described in filamentous fungi, in which unpaired chromosome regions—products of ectopic insertions or novel sequences in the genome—are detected by a mechanism known as meiotic *trans*-sensing (ARAMAYO and METZENBERG 1996). Then, the unpaired regions are post-transcriptionally silenced by the process called Meiotic Silencing by Unpaired DNA (MSUD) (SHIU *et al.* 2001). The current hypothesis proposes that MSUD is an ancient genome defense mechanism for silencing foreign sequences (*e.g.,* transposons) that has evolved and adapted to other biological roles like MSUC and MSCI in order to ensure proper chromosome segregation during meiosis (KELLY and ARAMAYO 2007).

MEIOTIC SEX CHROMOSOME INACTIVATION

Meiotic Sex Chromosome Inactivation (MSCI) is a repressive mechanism that transcriptionally silences the heteromorphic sex chromosomes (*e.g.* X and Y). In the case of mammals, sex chromosomes are compartmentalized into a heterochromatic structure called the XY body (MONESI 1965) during the meiotic phase of gametogenesis (HANDEL 2004). Because of the heteromorphic nature of sex chromosomes, chromosome pairing during meiosis occurs either partially

11

or not at all. The X and Y chromosomes of eutherian mammals pair through their small pseudo-autosomal regions (PAR); but in the case of marsupial mammals, sex chromosomes lack significant homology and come together without synapsis (HANDEL 2004; NAMEKAWA *et al.* 2007). Further experiments in mice and nematodes have shown that silencing of sex chromosomes is avoided when either the X or Y chromosome is provided with a homologous partner, forming a synapsed bivalent that is actively transcribed (TURNER *et al.* 2006). This data demonstrated that MSCI is triggered by the lack of pairing.

Originally, MSCI was thought of as a transient silencing process active only during meiosis. However, it has been demonstrated that the repressive stage is maintained in some extent after meiosis, imposing a heritable chromatin imprint on the X chromosome (NAMEKAWA *et al.* 2006).

MSCI is not exclusive to mammals; its presence has been demonstrated in metazoans as diverse as the grasshopper (CABRERO *et al.* 2007), nematode worm (BEAN *et al.* 2004), and birds (SCHOENMAKERS *et al.* 2009). Failure to inactivate partnerless regions of sex chromosomes results in elevated germline apoptosis in both worms and mice (TURNER *et al.* 2006), and may contribute to male infertility in humans (ROYO *et al.* 2010), suggesting that MSCI is required for efficient meiotic progression (CHECCHI and ENGEBRECHT 2011).

Several hypotheses have been proposed to explain the biological function of this silencing mechanism. MSCI could be a meiotic adaptation to prevent deleterious recombination events in non-homologous regions of the X and Y

chromosomes (MCKEE and HANDEL 1993; INAGAKI *et al.* 2010). Also, it has been proposed that MSCI is a genomic defense mechanism against selfish genetic elements (KELLY and ARAMAYO 2007), or the consequence of sexual antagonism, in which genes that enhance reproductive fitness in one sex, reduce it in the other (MEIKLEJOHN and TAO 2010).

Studies of MSCI during mouse spermatogenesis have revealed some of the molecular events that lead to MSCI (Figure 1.3). Chromatin remodeling events, such as replacement of core histones with histone variants (H2A.1) and *de novo* incorporation of histone variants (H3.3) occur in the nucleosomes placed within the XY body. Accumulation of dimethylated and trimethylated histone H3 (H3K9me2/3); and deacetylation of histones H3 and H4 are also hallmarks of MSCI. They all together, physically exclude active RNA polymerase II from sex chromosomes (VAN DER HEIJDEN *et al.* 2007). Additionally, enrichment of phosphorylated histone H2AX (γH2AX) and ubiquitinated histone H2A (UbH2A) have been observed during MSCI in mouse (CELESTE *et al.* 2002). Further work demonstrated that phosphorylation of H2AX by the sensor kinase ATR occurs after the tumor suppressor protein BRCA1 recruits ATR to the XY body (TURNER *et al.* 2004).

Figure 1.3 Representation of MSCI events during spermatogenesis. (A) During leptotene, widespread ATM-dependent phosphorylation of H2AX occurs in response to meiotic DSBs formation. Axial elements (AE) begin to assemble and associate with BRCA1 and ATR proteins. (B) During zygotene, BRCA1, ATR and γH2AX remain as foci on the AEs of chromosomes (sexual or autosomal) that have not yet synapsed and disappear from synapsed chromosome regions. (C) During zygotene-pachytene transition, complete autosomal synapsis is observed and recombination-related γH2AX disappears. BRCA1 and ATR are enriched on the AEs of sexual chromosomes (XY). DNA is arranged in loops (zoom). (D) At early pachytene, ATR dissociates from the AE and re-localizes along DNA loops, where it phosphorylates H2AX, resulting in MSCI and the formation of the sex body. (E) During mid-to-late pachytene, other histone modifications, including H3K9me2 and uH2A ensure the maintenance of MSCI. (F) During the transition from diplotene-to-diakinesis, BRCA1, ATR and γH2AX are lost from the X and Y chromosomes and the sex body migrates to the centre of the nucleus. The other histones modifications are maintained ensuring MSCI proceeds throughout the meiotic divisions and continue in the spermatids. This process is known as post-meiotic sex chromosome repression (PSCR) (Turner *et al.* 2006). Vertical lines show transcriptional activity of the X and Y chromosomes at different stages; high gene expression (green) and repress transcription (red). Figure adapted from (Turner 2007).

XY Transcriptionally active
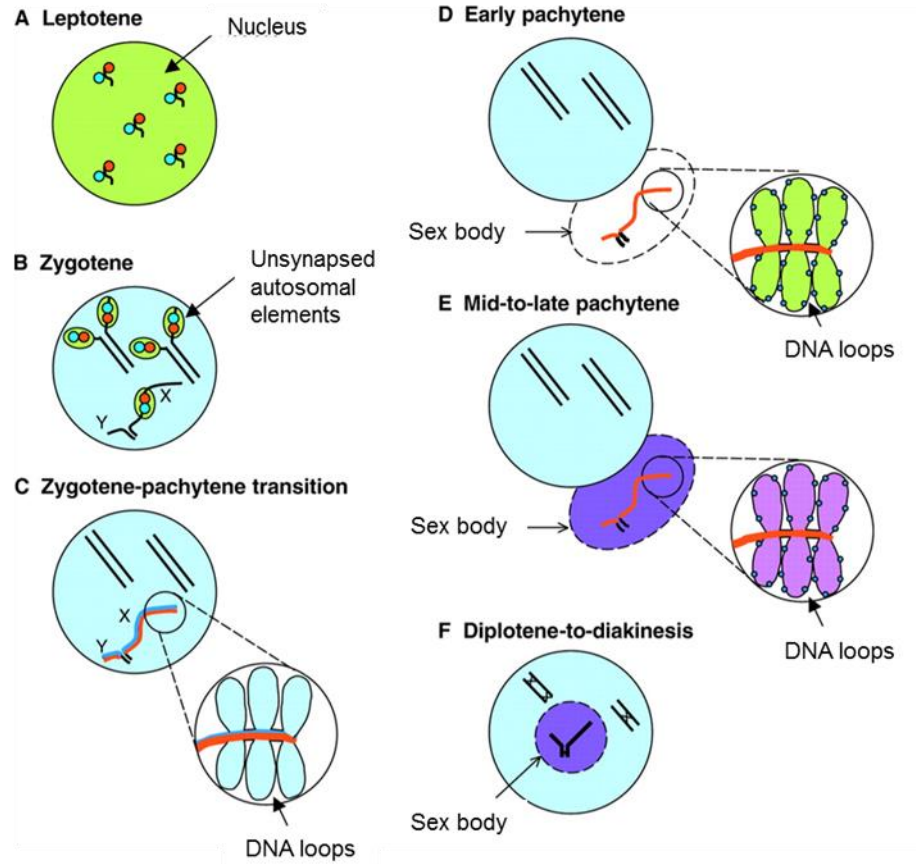
XY Transcriptionally repressed

A Leptotene — Nucleus

B Zygotene — Unsynapsed autosomal elements

C Zygotene-pachytene transition — DNA loops

D Early pachytene — Sex body — DNA loops

E Mid-to-late pachytene — Sex body — DNA loops

F Diplotene-to-diakinesis — Sex body

Axial elements — BRCA1 — ATR — γH2AX — Other modifications: H3K9me2, uH2A — PMSC

15

Although establishment of repressive chromatin structure and transcriptional silencing are common features of MSCI in mammals, worms and birds, each organism has its own molecular machinery to accomplish this feat. Accumulation of UbH2A and γH2AX in chicken is transient. Histone variant H3.3 has not been observed during MSCI in chicken and is depleted from the X chromosome during worm spermatogenesis There is no worm homolog of γH2AX, and the worm homolog of BRCA1 is not required for MSCI (CHECCHI and ENGEBRECHT 2011). In *C. elegans*, the single X chromosome of male meiotic cells is enriched in dimethylation of histone H3 (H3K9me2), a repressive histone mark (BEAN *et al.* 2004). Therefore, MSCI in these organisms appears to be solely dependent upon a condensed chromatin architecture that blocks transcriptional machinery recruitment. Another difference observed among organisms is the timing relative to chromosome synapsis. In eutherian mammals, failure of synapsis during prophase pachytene activates MSCI. However, in marsupial mammals MSCI occurs in early pachytene before co-localization of X and Y (NAMEKAWA and LEE 2009). The same is observed in chickens in which meiotic silencing precedes synapsis of Z and W, suggesting that a homology search mechanism—rather than asynapsis itself— might be the trigger for MSCI (SCHOENMAKERS *et al.* 2009).

MEIOTIC SILENCING OF UNSYNAPSED CHROMATIN

Meiotic Silencing of Unsynapsed Chromatin (MSUC) is the process responsible for the transcriptional silencing of unsynapsed autosomal chromosome region, during meiotic prophase in male and female mammals (SCHIMENTI 2005). In mammals, asynapsis of the X and Y chromosomes, and their silencing through MSCI, is essential for spermatogenesis. However, non-synapsis of autosome regions increases the risk of meiotic segregation errors— aneuploidy—and may be detrimental for meiotic progression causing gametogenic failure and sterility (NAUMOVA *et al.* 2013). Studies of MSUC during male meiosis showed that transcriptional silencing of the unpaired autosomal region precedes and affects silencing of sex chromosomes. It has been proposed that MSUC interferes with the silencing of sex chromosomes, which may be the cause of sterility (HOMOLKA *et al.* 2012).

There are some mechanistic parallels between MSUC and MSCI. As observed in MSCI during male spermatogenesis, BRCA1 protein senses and localizes to unsynapsed chromosome regions, in this case to the autosome, and is responsible for recruiting the kinase ATR followed by the phosphorylation of histone H2AX (TURNER *et al.* 2005). Localization of these three markers and late accumulation of ubiquitinated histone H2A to the unsynapsed region are causally related to transcriptional repression (BAARENDS *et al.* 2005). In the case of unsynapsed autosomes in *C. elegans*, enrichment in H3K9me2 was also observed. This modification is the same repressive histone mark found during

sex chromosome inactivation (BEAN *et al.* 2004). In mammals, MSUC activation depends on the presence of unrepaired meiotic DSBs. Unrepaired DSBs accumulate on the unsynapsed chromosome regions and are then detected by BCRA1 and ATR proteins, key players in the DSB checkpoint (MAHADEVAIAH *et al.* 2008; SCHOENMAKERS *et al.* 2008).

Additional work on mice demonstrated that meiotic silencing also occurred in the single X chromosome of the XO female mouse, where the lone X was covered by BRCA1, ATR and γH2AX, and was transcriptionally quiescent during pachytene-stage of meiotic prophase (TURNER *et al.* 2005). Interestingly, it was demonstrated that in these females, heterologous synapsis and escape from silencing occur. Turner *et al.* showed that homologous synapsis is not required for transcription. Transcriptional repression or chromatin modifications were absent in XO oocytes where the lone X looped back on itself to self-synapse non-homologously (TURNER *et al.* 2005). It is not clear whether the lack of MSUC activation allows heterologous synapsis, or whether heterologous synapsis prevent MSUC activation.

All these observations led to the conclusion that silencing of unsynapsed chromosome regions is a conserved mechanism in mammals. Therefore, it is proposed that MSUC is the general meiotic silencing mechanism and that it evolved into MSCI, a more specialized process, responsible for silencing sex chromosomes during normal male meiosis. It has been hypothesized that MSUC plays a role similar to a meiotic checkpoint and contributes to meiotic arrest

through the silencing of genes that are crucial to meiosis, and, in doing so, may protect against aneuploidy in subsequent generations (TURNER 2007).

This general silencing mechanism resembles a phenomenon described in *Neurospora crassa* in 2001, MSUD, a mechanism that requires components of the RNA interference (RNAi) machinery (SHIU *et al.* 2001). A link with RNA-mediated silencing has also been hypothesized for MSUC/MSCI. Costa and colleagues (COSTA *et al.* 2006) showed that the MAEL protein (mammalian MAELSTROME) was associated not only with XY body, but also with unsynapsed autosomes and interacted with proteins involved in gene silencing, *i.e.* SNF5 (RAYMAN *et al.* 2002) and SIN3B (PAN *et al.* 2005). MAEL is a component of the chromatoid body, a perinuclear germline granule where RNA and RNA processing proteins, including proteins involved in microRNA (miRNA) pathway, accumulate (COSTA *et al.* 2006). Although the exact role of MAEL has not been determined, it is essential for meiotic chromosome synapsis (SOPER *et al.* 2008). Similar to *Drosophila* Maelstrome (FINDLEY *et al.* 2003), MAEL shuttles between the nucleus (unsynapsed chromosomes) and the chromatoid body, probably transporting miRNAs, thus suggesting a link between RNAi and MSUC.

MEIOTIC SILENCING BY UNPAIRED DNA

Meiotic Silencing by Unpaired DNA (MSUD) was originally described in the filamentous fungus *N. crassa* (SHIU *et al.* 2001). MSUD comprises two stages: meiotic *trans*-sensing**,** the process responsible for detecting unpaired

19

DNA (ARAMAYO and METZENBERG 1996) and meiotic silencing, the actual destruction of the transcripts expressed from the unpaired region (SHIU *et al.* 2001). Meiotic silencing is activated during *N. crassa* sexual development after karyogamy, when unpaired DNA is detected. In *Neurospora*, meiosis occurs inside the zygote and is essential for production of sexual ascospores. The zygote is formed after two haploid nuclei from individuals of opposite mating type are fused. During meiosis, unpaired DNA is silenced, but in contrast to MSUC, all DNA sequences homologous to the unpaired region are also silenced, even if those sequences are synapsed. Because detection of odd DNA regions (unpairedness) triggers this silencing mechanism and not only unsynapsis, the distinction between asynapsis and unpairing was highlighted in the name of the phenomenon (SHIU *et al.* 2001). The participation of core components of the RNA-mediated silencing mechanism, RNA interference (RNAi), suggests that gene silencing in MSUD occurs at the post-transcriptional level, after mRNA molecules are synthesized and exported from the nucleus (ARAMAYO and SELKER 2013). Post-transcriptional silencing in MSUD differs from the transcriptional silencing that occurs in MSUC.

Understanding meiotic *trans*-sensing and meiotic silencing is not only important from the molecular perspective of genome defense mechanisms, but also from an evolutionary point of view. It was proposed that MSUD works as a meiosis-specific defense mechanism that protects the genome from infectious genetic parasites (*e.g.* transposable elements (TE)). Therefore, MSUD assures

20

that the integrity of the genome is maintained during sexual reproduction by preventing foreign DNA from being expressed (KELLY 2006; SHIU *et al.* 2006; KELLY and ARAMAYO 2007). Another interesting hypothesis is that meiotic silencing could play a major role in reproductive isolation and speciation. Evolving organisms could accumulate mutations in regions that are not required for activation of meiotic silencing or required for normal meiosis progression, without affecting the interbreeding ability. On the contrary, if genes whose products are essential for meiosis are rearranged, that would activate meiotic silencing and the evolving organism becomes reproductively isolated (LEE *et al.* 2004). According to this idea, interspecific crosses within the genus *Neurospora* become more fertile if the *N. crassa* parent carries a mutation in one of the suppressors of meiotic silencing (SHIU *et al.* 2001) (S. Gajjar and R. Aramayo, unpublished data).

Neurospora Biology

Neurospora crassa is the best understood filamentous fungus and is one of the first eukaryotic model systems. Important features make this fungus attractive for a variety of genetic, biochemical, subcellular and developmental studies. *N. crassa* grows fast and propagates easily on defined growth media, it has a moderate complexity compared with other eukaryotes, a complete genome sequence is available, and the fungal community has access to a

knockout mutant collection (DAVIS and PERKINS 2002; GALAGAN *et al.* 2003; BORKOVICH *et al.* 2004; COLOT *et al.* 2006).

Depending on environmental conditions, *N. crassa* undergoes asexual or sexual development. Under conditions that favor asexual or vegetative phase, a haploid asexual spore (conidium) germinates, forming branch filaments called hyphae. The hyphal system spreads out to form a mycelium, which then produces aerial hyphae or conidiophores where new conidia are formed. Each conidium can start new vegetative growth or fertilize strains of the opposite mating type to start sexual development (ARAMAYO and SELKER 2013).

N. crassa possesses two mating types referred to as "*A*" and "*a*", both of which can act either as female or male cells during sexual development. Detection of nutrient limitation activates the sexual phase by inducing the formation of female structures (protoperithecia). When a specialized hyphal structure (trichogyne) comes out from a protoperithecium of one mating type and finds a conidium from the opposite mating type, fertilization occurs and sexual development proceeds. Plasmogamy initiates the development of perithecia or fruiting bodies, which is the sexual apparatus. After plasmogamy, the nuclei from both mating types coexist in heterokaryotic tissue. Multiple mitotic divisions occur until the nuclei are sorted into dikaryotic tissue in which each cell has one nucleus from each mating type. The crozier—a hook-shaped cell precursor of the ascus cell—is formed and the two nuclei get into proximity and undergo several mitotic divisions. A coordinated mitosis and septum formation yield two

uninucleated cells and one ascus mother cell. Inside the ascus, two haploid nuclei from opposite mating types fuse together, forming a transient diploid zygote, where meiotic *trans*-sensing takes place (BORKOVICH *et al.* 2004). After karyogamy two meiotic and one post-meiotic mitotic divisions occur. As a result, 8 individual haploid ascospores are formed inside each ascus arrayed in an order that reflects their lineage (RAJU 1980; RAJU and LESLIE 1992). Once activated, MSUD is maintained through sexual development until ascospores are compartmentalized. Up to 200 asci are developed inside each perithecium. Once mature, the ascospores are ejected from the ascus finalizing the sexual phase (Figure 1.4) (ARAMAYO *et al.* 1996). Meiotic segregation and recombination can be studied in *Neurospora* by analyzing individual asci or random spores ejected from numerous asci (ARAMAYO and SELKER 2013). Gene products responsible for ascospores phenotypes such as pigmentation (dark color versus white) or shape (spindle versus round) can be used as reporters of chromosome segregation and gene regulation during sexual development (Raju 2009).
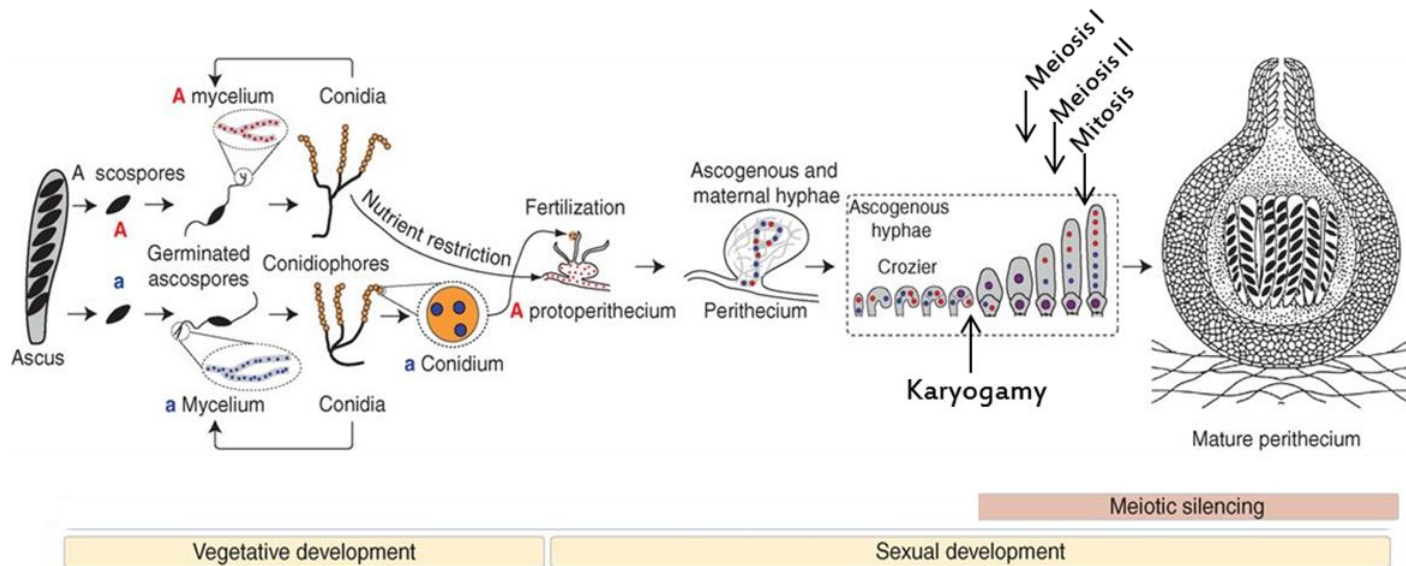
Figure 1.4 Life cycle of *N. crassa*. Half of developed ascospores are mating type A (red) and half are mating type a (blue). Sexual spores (ascospores) and vegetative spores (conidia) germinate and build mycelia, from which conidiophores rise up, forming conidia. Under low nitrogen conditions mycellium from either mating type form the female structure (protoperithecium). Once the protoperithecium is fertilized by the male structure (conidia) from the opposite mating type, the perithecium or fruiting body is formed. After fertilization, ascogenous hyphae with nuclei from both mating types undergo karyogamy, meiosis I, meiosis II and a post-meiotic mitosis division. As a result, eight individual ascospores are formed inside each ascus. More than 200 asci are held in each perithecium. The time period where meiotic silencing mechanism is functional is shown. Figure adapted from (ARAMAYO and SELKER 2013).

Also, *N. crassa* is a perfect eukaryotic model system for RNAi studies. In addition to MSUD, another RNAi mechanism is induced by repetitive transgenic sequences during the vegetative cycle called quelling (ROMANO and MACINO 1992). This demonstrates the diversity of RNAi phenomena in this organism. The identification of components in both, MSUD and quelling mechanisms have contributed to the general understanding of RNAi pathways not only in filamentous fungi but also in metazoans.

In MSUD, nuclear and perinuclear stages are necessary for meiotic silencing completion. At the beginning of meiosis, unpaired DNA is detected by meiotic *trans*-sensing (ARAMAYO and METZENBERG 1996). This step is indispensable for activation of the perinuclear phase where the actual destruction of the transcripts expressed from the unpaired region (SHIU *et al.* 2001).

Meiotic *trans*-sensing

In *N. crassa*, genes sense each other during meiotic chromosomal pairing through the meiotic *trans*-sensing mechanism. Here, unpaired DNA regions are detected and a silencing signal (*e.g.,* aberrant RNA (aRNA)) is synthesized. aRNA is a diffusible signal that activates meiotic silencing machinery and triggers gene silencing during meiosis. In 1996 Aramayo and Metzenberg demonstrated that deletion mutants of *Asm-1* (*Ascospore maturation-1*) gene—whose product is required for the formation of female structures and ascospore

maturation—are ascus-dominant. *Asm-1$^\Delta$* deletion strains carrying ectopic DNA copies of this gene rescued the normal phenotype in vegetative but not in sexual phase. Opposite to that, strains carrying frameshift alleles (*asm-1^{fs}*) had the same vegetative defects as *Asm-1$^\Delta$* strains but had different effects on sexual development. Hence, several gene replacement experiments demonstrated that for proper expression of the gene, *asm-1$^+$* must be paired to its allelic counterpart before meiotic division occurs (Figure 1.5) (ARAMAYO and METZENBERG 1996; ARAMAYO and SELKER 2013).

The proposed hypothesis is that meiotic *trans*-sensing is the mechanism that detects unpaired DNA in the zygotic cell immediately after karyogamy, but before the first meiotic division begins. If an unpaired chromosomal region is detected via *trans*-sensing, meiotic silencing is activated and a sequence-specific signal, such as aRNA, is produced and maintained throughout the post-meiotic mitotic division, ensuring the silencing of unpaired genes and all genes sharing homology in the genome (ARAMAYO *et al.* 1996; SHIU *et al.* 2001). If the unpaired DNA encodes a protein that is required for normal meiosis progression and/or ascospore development, MSUD will arrest the sexual cycle at the stage where the gene product of the silenced gene has a function (SHIU and METZENBERG 2002).

Figure 1.5 Discovery of meiotic *trans*-sensing. Genetic experiments are illustrated. *Ascospore maturation-1* (*Asm-1*) gene was used as reporter gene. Genotype of the haploid parents is shown for each cross. Red, chromosomes from *mating type A*; blue, chromosomes from *mating type a*. Violet boxes represent the diploid cells. Asci phenotypes are shown. Black represents mature (viable) ascospores; white represents immature (inviable) ascospores. Key genetic experiments illustrate the characterization of meiotic *trans*-sensing and meiotic silencing. (A) Wild-type cross. (B) Crosses between *asm-1*$^+$ strain and *asm-1*^{fs} strain produced 4:4 segregation of viable and inviable ascospores, suggesting that *asm-1*^{fs} is recessive. (C) Interestingly, crosses between deletion mutant (*Asm-1*$^\Delta$) and wild-type (*asm-1*$^+$) result in nearly all the ascopores being inviable, including those carrying the wild-type allele. Therefore, *Asm-1*$^\Delta$ allele was ascus-dominant, contrasting with the recessive behavior in vegetative cells. Haplo-insufficiency could be the cause of the observed dominance. (D) For testing this possibility, the authors evaluated crosses between wild-type strain with deletion strains carrying ectopical functional copies (*asm-1*$^+$ X *Asm-1*$^\Delta$, *asm-1*$^+$ (ectopic)). The functional ectopic copy of the gene failed to rescue the spore maturation. It was proposed that an interaction between alleles at homologous chromosomal positions was required for normal gene expression. (E) The hypothesis was tested by crossing strains which both carried an ectopic copy of the gene in an *Asm-1*$^\Delta$ background (*Asm-1*$^\Delta$, *asm-1*$^+$ (ectopic) X *Asm-1*$^\Delta$, *asm-1*$^+$ (ectopic)). Indeed, the two ectopic alleles localized at the same homologous chromosomal position produced normal progeny. (F) Finally, it was demonstrated that meiotic silencing results from the presence of unpaired alleles rather than from absence of paired ones. For that, crosses between wild-type strains with strains carrying an ectopic copy of the gene plus the wild-type allele were tested. All ascospores were inviable, indicating that silencing occurred Figure adapted from (ARAMAYO and SELKER 2013).
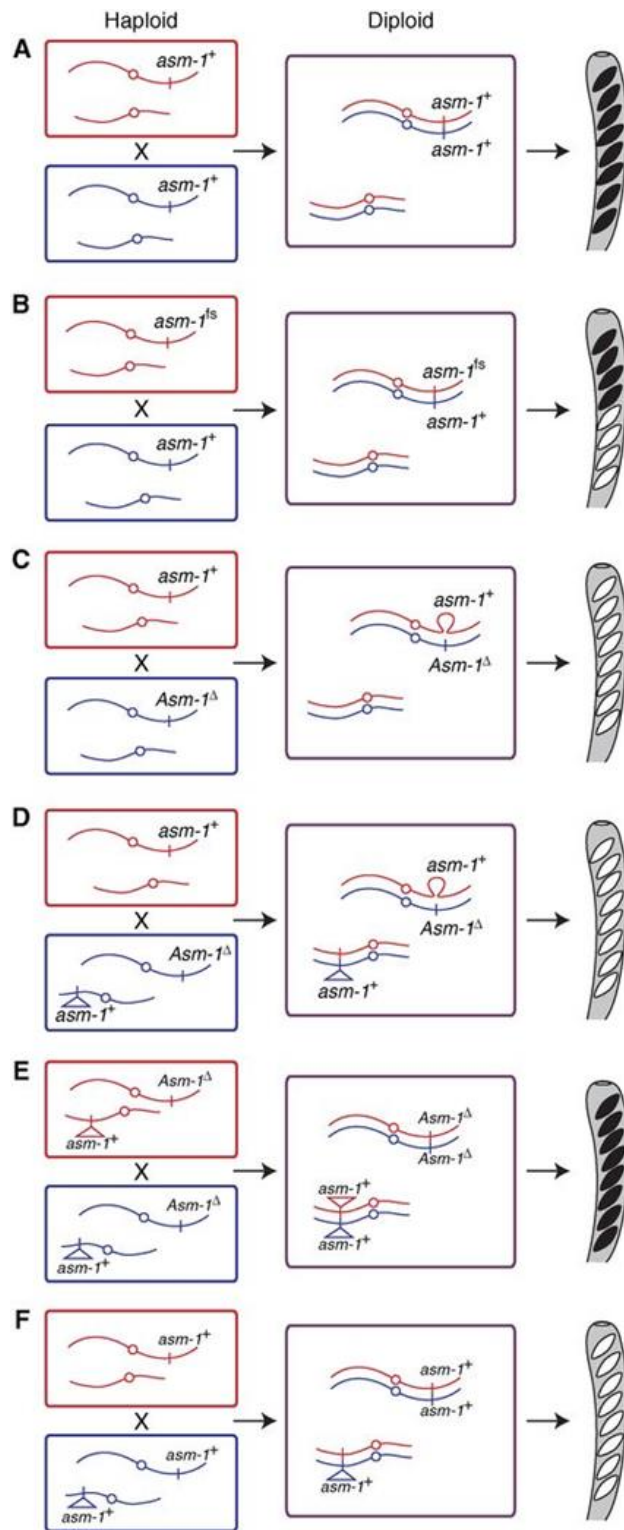
Unpaired DNA must fulfill some molecular requirements in order to be recognized and silenced: 1) A minimum region of DNA (~700 nucleotides) must be unpaired; 2) the length of the unpaired region is proportional to the efficiency of silencing; 3) unpaired DNA must have homology to the actual gene transcript; 4) intergenic regions do not trigger silencing of the reporter gene; and 5) gene silencing does not required the presence of promoters in the unpaired region (LEE *et al.* 2004). In addition, the silencing signal produced by unpaired DNA is very specific and does not propagate onto adjacent paired regions (KUTIL *et al.* 2003). Together, these observations support the idea that MSUD is a post-transcriptional gene regulation that involves RNA molecules. These observations also raise the possibility that unpaired DNA is transcribed by an unconventional mechanism, in which promoter elements are not needed (LEE *et al.* 2004).

Although meiotic *trans*-sensing and meiotic silencing are coupled processes, they work independently. By studying homeology or partial homology, it was demonstrated that DNA methylation affects chromosome sensing without having an effect on silencing (PRATT *et al.* 2004).

Meiotic Silencing

Components of RNA- interference (RNAi) Mechanism

Numerous gene products essential for MSUD have been identified using forward and reverse genetics; the genes are referred to as suppressors of meiotic silencing because mutations in these genes suppress the silencing of

unpaired DNA. Some of these suppressors are homologous to proteins that participate in regulation of gene expression via RNA interference (RNAi) pathways, including post-transcriptional gene silencing processes (PTGS) (MELLO and CONTE 2004). These proteins include: RNA dependent RNA polymerase (RdRP/SAD-1) (SHIU *et al.* 2001), Argonaute-like protein (SMS-2) (LEE *et al.* 2003), and Dicer-like protein (DCL-1) (ALEXANDER *et al.* 2008). These RNAi-related proteins are highly conserved and found in a variety of organisms. The participation of these three proteins in MSUD suggests that the mechanism of silencing unpaired DNA may involve the production of small interfering RNAs. Therefore, a link between MSUD and RNAi was established and it was proposed that MSUD is a PTGS mechanism (LEE *et al.* 2004).

In *N. crassa,* RdRP is encoded by the S*ad-1* (*Suppressor of ascus dominance-1*) gene. Mutations in S*ad-1* (*Sad-1*$^{\Delta}$) suppress "ascus-dominant" phenotypes, like the one observed in crosses with heterozygous *Asm-1*$^{\Delta}$ deletion mutant (Figure 1.5 (C)). It was proposed that mutations in this gene suppress meiotic silencing and subsequently ascus-dominance (SHIU *et al.* 2001). SAD-1 homologous proteins in other organisms are required for double-strand RNA (dsRNA) synthesis and have been implicated in PTGS, in which mRNA of targeted genes are degraded (COGONI and MACINO 1999a; SMARDON *et al.* 2000; CARTHEW 2001). In addition, a second RdRP in *N. crassa*, encoded by the *qde-1* (*quelling defective-1*) gene is also involved in MSUD (Aramayo Lab unpublished data). QDE-1 protein was originally reported as a main player in

quelling, another PTGS mechanism that silences genes homologous to transgenes in *N. crassa* vegetative cells (COGONI and MACINO 1999a). Transcripts of s*ad-1*[+] are only detected in cells undergoing sexual development and their expression is independent of the presence of unpaired DNA. Furthermore, homozygous *Sad-1* mutant crosses are sterile. This evidence indicates that SAD-1 not only plays a role in MSUD, but also may have another function required for normal meiotic progression (SHIU and METZENBERG 2002). SAD-1 protein is localized in the perinuclear region, where it has been hypothesized that actual degradation of the transcripts occurs (SHIU *et al.* 2006).

Along that line, an Argonaute-like protein is also required for MSUD, and it is encoded by the *Sms-2* (*Suppressor of meiotic silencing-2)* gene (LEE *et al.* 2003). Argonaute proteins are highly conserved and are the catalytic components of the protein complexes responsible for the silencing of gene expression through RNA-silencing pathways (MEISTER 2013). Argonaute binds directly to sequence-specific small RNAs (siRNAs), which guide the cleavage of targeted mRNAs via the RNA-induced silencing complex (RISC) by complementary base pairing (KAWAMATA and TOMARI 2010). In *N. crassa,* the SMS-2 protein seems to have a role exclusively during the sexual phase. *Sms-2* mutant strains proceed normally through vegetative growth. However, they show sexual development impairment, in which homozygous mutant crosses are completely barren and arrested in meiotic prophase. Heterozygous crosses reduce meiotic silencing to a low level, indicating that *Sms-2* behaves as a

dominant suppressor (Lee *et al.* 2003). Interestingly, by using fluorescent fusion proteins, it was established that the Argonaute protein (SMS-2) and the RNA polymerase (RdRP; SAD-1) protein co-localize in the perinuclear cellular region in *N. crassa* (Alexander *et al.* 2008).

By testing the behavior of single and double *Sms-2* and *Sad-1* mutants in the presence of unpaired DNA—and calculating the percentage of mature ascospores as an indirect form of estimating suppression of meiotic silencing—Lee *et al.* demonstrated that *Sms-2* and *Sad-1* are both necessary, but not sufficient, for meiotic silencing (Lee *et al.* 2003).

The last key component of RNAi or PTGS mechanism is the RNaseIII-type endonuclease Dicer, which cleaves long double-stranded RNA (dsRNA) molecules to produce small interference RNAs (siRNAs) (Bernstein *et al.* 2001). In *N. crassa*, the *Dcl-1*/*Sms-3* (*dicer like protein-1*/*Suppressor of meiotic silencing-3*) gene encodes a Dicer enzyme required for meiotic silencing (Lee *et al.* 2004; Alexander *et al.* 2008). The *dcl-1*+ gene is expressed during vegetative growth and sexual development. In fact, DCL-1 is one of two Dicer enzymes needed for quelling (Catalanotto *et al.* 2004). Sexual development arrests at an early stage in crosses homologous for *Dcl-1*$^\Delta$ mutation.

Other Suppressors of Meiotic Silencing

Other gene products with unknown function are required for meiotic silencing. *Sad-2* (*Suppressor of ascus dominance-2*) dominant mutant

32

suppresses the meiotic silencing of unpaired loci with similar efficiency as *Sad-1$^\Delta$*. SAD-2 protein is also required for normal meiosis progression and ascospore formation. Homozygous crosses for *Sad-2* mutant are blocked at prophase I. Moreover, expression of the *sad-2$^+$* gene is limited to the sexual phase, from pre-karyogamy until diplotene stage in meiosis (SHIU *et al.* 2006). The RNA polymerase (*i.e.,*SAD-1) and SAD-2 proteins co-localize in the perinuclear region throughout meiotic prophase and physical interaction between them was observed (SHIU *et al.* 2006; BARDIYA *et al.* 2008). It has been proposed that SAD-2 may recruit the RNA polymerase to the perinucleus, based on the observation that SAD-1 distribution was scattered in the cytoplasm in the *Sad-2* mutant strain. This observation also implies that the proper localization of SAD-1 may be important for its activity; however, SAD-2 function remains undetermined (SHIU *et al.* 2006).

The gene product of *qip* (*quelling defective-2 interacting protein*) is involved in meiotic silencing and necessary for normal sexual development in *Neurospora*. QIP localizes in the perinuclear region, as does SMS-2 (LEE *et al.* 2010a; XIAO *et al.* 2010). QIP, an exonuclease responsible for the degradation of one strand of the siRNAs duplex (passenger strand) which is already bound to the Argonaute protein, was originally described as an essential protein for quelling (MAITI *et al.* 2007). Like the majority of suppressors of meiotic silencing, a homozygous cross for *Qip$^\Delta$* deletion mutant is completely barren and meiosis is arrested at an early stage.

Parallel analyses testing the involvement of qip^+ in meiotic silencing produced different results in terms of how efficiently qip mutant alleles suppress meiotic silencing. When silencing is induced by homeologous alleles of the reporter gene, Qip^Δ acts as a dominant suppressor of meiotic silencing (LEE *et al.* 2010a). However, Qip^Δ does not act as a dominant suppressor when silencing is induced by a deletion of one copy of the reporter gene (XIAO *et al.* 2010). These results suggest that homeology—partial correspondence between homologous chromosomal regions—and heterology conditions are treated differently and each one activates a specific meiotic silencing response. Interestingly, this discrepancy revealed an important aspect of meiotic silencing: the nature of the unpaired DNA has an effect on the meiotic silencing response. This evidence suggests that the cell can differentiate between diverse classes of unpaired conditions (homeology versus heterology) and that they are silenced in a different fashion via MSUD (R. Millimaki and R. Aramayo, unpublished data).

Other suppressors of meiotic silencing with perinuclear localization have been identified. However, their molecular function and involvement in meiotic silencing are still unknown (Table 1.1).

In 2008 Pratt reported the first suppressor that has a nuclear localization called *Sms-4* (*Suppressor of meiotic silencing*-4) (PRATT 2008). *Sms-4* gene product is predicted to bind RNA, and is required for meiotic silencing, but not for sexual development. Crosses homozygous for *Sms-4* mutant do not block meiosis. *Sms-4* transcripts are detected in all stages of *N. crassa* vegetative and

sexual development (PRATT 2008). Recently, another nuclear protein was reported and named SAD-5 (Suppressor of Ascus Dominance-5). However, its function is undetermined (HAMMOND *et al.* 2013b). Because of their cellular localization, it has been proposed that SMS-4 and SAD-5 may be involved in meiotic *trans*-sensing and unpaired DNA detection. It was also proposed that SMS-4 may play a role connecting meiotic *trans*-sensing and meiotic silencing by transporting RNA molecules from the nucleus to the perinuclear region (PRATT 2008).

Many other suppressors of meiotic silencing have been identified in the Aramayo Lab using a reverse genetic screen coupled to the *N. crassa* knockout library (COLOT *et al.* 2006). Suppression of meiotic silencing is determined using reporter genes (*e.g., Rsp* (*Round spore*) gene). The percentage of spores with wild-type phenotype (*e.g.,* spindle shape) compared to mutant phenotype (*e.g.,* round shape) is calculated from crosses between a heterozygous knockout mutant strain and an unpaired reporter gene strain. When MSUD is active, crosses involving an unpaired reporter gene produce nearly 100% mutant phenotype spores. When MSUD is suppressed, due to mutations in a gene involved in meiotic silencing, the percentage of mutant phenotype decreases.

From those characterized suppressors of meiotic silencing, a significant number are localized in the nucleus. It is reasonable to think that they are involved in the initial nuclear stage of meiotic silencing, participating in the recognition of the unpaired DNA regions. The second group of suppressors is

localized in the perinuclear region (Table 1.1) (Aramayo Lab unpublished data). A model for meiotic silencing is presented in Figure 1.6.

The perinuclear localization of some suppressors, including SAD-1, SAD-2, SMS-2, DCL-1, and QIP, suggests the possibility of a protein complex being formed in this cellular space. It also indicates that the perinuclear region is the center of RNAi activity for MSUD. This is reminiscent of another perinuclear structure observed in the germline of diverse eukaryotes, *i.e.* germline granules (EDDY 1975). These bodies are referred to as P granules in *C. elegans* (UPDIKE and STROME 2010), nuage in *D. melanogaster* (LIM and KAI 2007), or chromatoid bodies in mammals (KOTAJA and SASSONE-CORSI 2007). Ribonucleoprotein complexes localize to germline granules and participate in mRNA transport and translation control. More specifically, they play a role in transposon repression (SOPER *et al.* 2008). In *Drosophila* and mouse, RNAi-related proteins co-localize in germline granules, suggesting that this structure functions as an intracellular focal domain where RNA processing occurs (KOTAJA and SASSONE-CORSI 2007; PANE *et al.* 2007).

Table 1.1 Suppressors of meiotic silencing in *N. crassa*

Gene Name	Accession #	Molecular Function	Cellular Localization	Reference
Sad-1	NCU02178	RdRP	Perinuclear	(SHIU ET AL. 2001)
Sad-2	NCU04294	Unknown	Perinuclear	(SHIU ET AL. 2006)
Sad-3/Sms-7	NCU09211	Helicase	Perinuclear	(HAMMOND ET AL. 2011)
Sad-4/Sms-10	NCU01591	Unknown	Perinuclear	(HAMMOND ET AL. 2013B)
Sad-5	NCU06147	Unknown	Nuclear	(HAMMOND ET AL. 2013B)
Dcl-1/Sms-3	NCU08270	Dicer	Perinuclear	(LEE ET AL. 2004; ALEXANDER ET AL. 2008)
Qip	NCU00076	Exonuclease	Perinuclear	(LEE *et al.* 2010a; XIAO *et al.* 2010)
Sms-1	NCU02495	mRNA splicing	Nuclear	Aramayo *et. al.* Unpublished
Sms-2	NCU09434	Argonaute	Perinuclear	(LEE ET AL. 2003)
Sms-4	NCU01310	RNA Binding	Nuclear	(PRATT 2008)
Sms-5	NCU02088	SET domain	Perinuclear/Nuclear	Aramayo *et. al.* Unpublished
Sms-6	NCU04083	Unknown	Nuclear	Aramayo *et. al.* Unpublished
Sms-8	NCU04236	Unknow	Nuclear	Aramayo *et. al.* Unpublished
Sms-9	NCU06190	Helicase	Nuclear	Aramayo *et. al.* Unpublished
Sms-11	NCU01917	RNA recognition	Nuclear	Aramayo *et. al.* Unpublished
Sms-12	NCU08504	Unknown	Unknown	Aramayo *et. al.* Unpublished
Sms-13	NCU09064	Kinase	Perinuclear	Aramayo *et. al.* Unpublished
Sms-14	NCU05246	Helicase	Nuclear	Aramayo *et. al.* Unpublished
Sms-15	NCU06316	Argonaute siRNA chaperone	Perinuclear	Aramayo *et. al.* Unpublished
Sms-16	NCU09093	Helicase	Perinuclear	Aramayo et. al. Unpublished
Sms-17	NCU07579	Unknown	Nuclear	Aramayo et. al. Unpublished
Qde-1	NCU07534	D/RdRP	Nuclear	(COGONI and MACINO 1997)
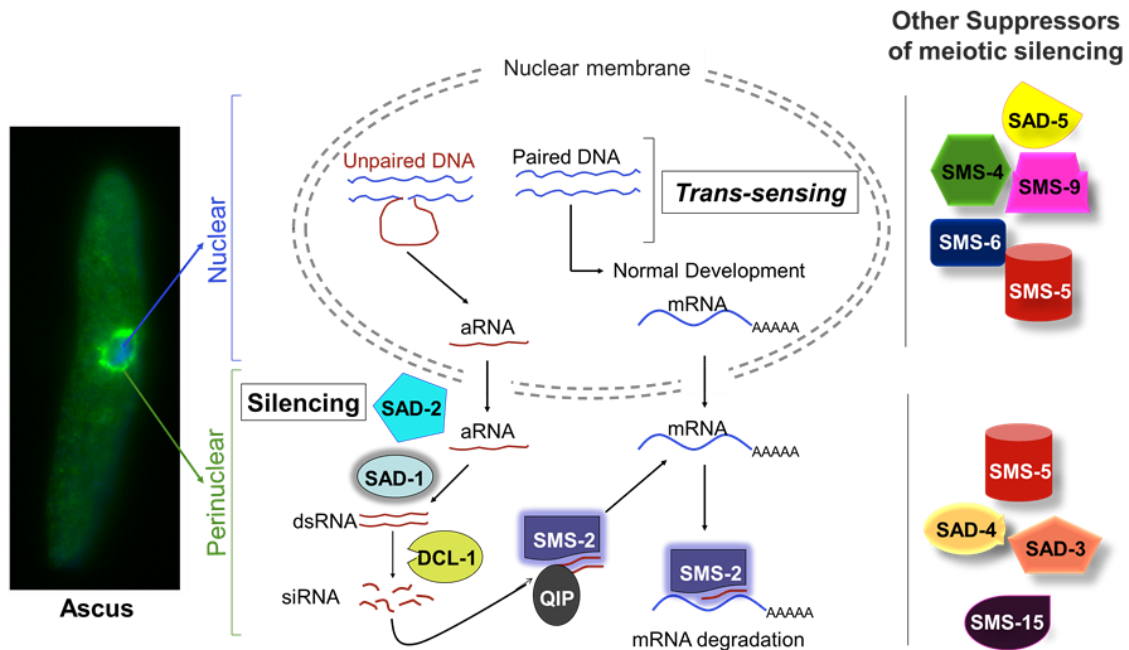
Figure 1.6 Model for meiotic *trans*-sensing and meiotic silencing. Left, an image of a sexual-spore bearing cell (ascus) at early meiosis. Inside this cell, the nucleus (blue), delineated by its nuclear membrane, is surrounded by a perinuclear structure (green). Right, close-up of nuclear and perinuclear cellular compartments. After karyogamy, homologous chromosomes are sensed (*trans*-sensing) and paired. In the nucleus, aberrant RNA (aRNA) is produced as a result of unpaired DNA detection. aRNA migrates to the perinuclear region where components of the meiotic silencing apparatus are localized. RdRP SAD-1 synthesizes dsRNA using aRNA as a template. Dicer SMS-3 dices dsRNA, forming siRNA duplexes, which are loaded into Arognaute SMS-2. Exonuclease QIP degrades the sense strand of the siRNA duplex. SMS-2, in partnership with the siRNA guide strand, identifies targeted transcript (mRNA) by sequence complementary and degrades it. SAD-2 is required for recruiting SAD-1, and probably other suppressors of meiotic silencing, to the perinuclear region. Other nuclear and perinuclear suppressors have been identified; however, their functions are still unknown (SHIU *et al.* 2006; ALEXANDER *et al.* 2008; ARAMAYO and PRATT 2010; HAMMOND *et al.* 2013b).

A great number of suppressors of meiotic silencing are essential for sexual development and ascospore formation. Therefore, although heterozygous crosses for suppressor mutants may develop normally, homozygous crosses are barren, stopping sexual development at different stages. For instance, S*ad-1* or *Sad-2* homozygous crosses are arrested in meiotic prophase I after asci are formed; however *dcl-1* or *qip* homozygous crosses arrest at a much earlier stages. This evidence suggests that some suppressors are important for ascus development (SHIU *et al.* 2006; ALEXANDER *et al.* 2008; LEE *et al.* 2010a). However, suppressors of meiotic silencing that are homozygous fertile (*i.e., Sad-4*, *sad-5*, *Sms-4,* and *Sms-5*) demonstrate that meiotic silencing and sexual development are uncoupled processes.

Small-interfering RNAs (siRNAs) in MSUD

Because MSUD is an RNAi-like mechanism, siRNA molecules are expected to be involved in this process. It was hypothesized that after unpaired DNA regions are detected, aberrant RNAs (aRNAs) form and translocate from the nucleus to the perinuclear region. In the perinucleus, they serve as templates for the synthesis of dsRNAs, which then are processed into siRNA. Argonaute protein, in partnership with siRNA, detects and silences targeted transcripts. Hammond *et al.* presented the first evidence of siRNAs associated with MSUD by sequencing and comparing RNA extracted from sexual tissue of wild-type (rsp^+ X rsp^+) and reporter-unpaired crosses (rsp^Δ X rsp^+). A significant

39

increase in small RNAs targeting the unpaired region were observed in the cross carrying the deletion. These small RNAs were called MSUD-associated small interfering RNAs (masiRNAs) (HAMMOND *et al.* 2013a). Most masiRNAs are around 25 nt in length and have a nucleotide bias towards Uridine (U) at the 5' end, which are similar features to small RNA molecules in other systems. The majority of masiRNAs are anti-sense with respect to the unpaired transcript, and therefore are complementary (CZECH and HANNON 2011).

RESEARCH AIMS

Important efforts have been made in order to determine the mechanism by which unpaired DNA is silenced during meiosis in *N. crassa*. It has been proposed that two individual and sequential processes work together. First, cells recognize the lack of homology or the presence of unpaired DNA regions via *trans*-sensing at a very early stage of meiosis. Second, the unpaired DNA recognition triggers the activation of the silencing machinery which silences all transcripts homologous to the unpaired region (*i.e.,* meiotic silencing). Despite the numerous gene products that have been determined to be involved in MSUD, very little is known about the molecular mechanisms underlying this intricate process.

Due to the similarities found between meiotic silencing and RNAi mechanisms, much is known and also speculated about the functions of the suppressors and the regulation of silencing in the perinuclear region. However,

less is known about the initial nuclear stage and how detection of unpaired DNA in the nucleus activates the meiotic silencing machinery in the perinucleus. It is reasonable to speculate that a molecular connection between the two stages must exist, especially because they occur in different cellular compartments. Several questions need to be answered to fulfill the proposed meiotic silencing model (Figure 1.6). Some of these questions are: How is the unpaired DNA recognized at the beginning of meiosis? How are aRNAs synthesized? How are aRNAs translocated to the perinuclear region? What is the nuclear molecular signal responsible for the activation of the perinuclear silencing machinery?

In order to answered those questions and continue with the study of the molecular mechanisms and regulations behind meiotic silencing, it is critical to combine different approaches, such as genetic, biochemical, bioinformatics, and proteomics analyses. Thanks to previous genetic studies in conjunction with bioinformatics analyses, several components required in the silencing process have been identified. However, there is little data relating to biochemical and molecular interactions during meiotic silencing. The difficulties associated with the scarcity of the sexual tissue, the manipulation and disruption of sexual structures, and the low protein yield obtained from sexual material are the main reasons for the lack of biochemical data. Despite these difficulties, there is a necessity for incorporating biochemical analyses in the study of meiotic silencing in an attempt to get a better understanding of its molecular basis.

It is of great relevance to understand how the meiotic silencing process is regulated in fungi. This knowledge can be extrapolated to higher eukaryotes, allowing for a greater comprehension of other meiotic silencing processes, such as MSCI and MSUC. This would also help to elucidate the process by which meiotic silencing errors cause sterility in humans.

In addition to the biochemical analyses, proteomics studies are also required in order to get a better understanding of the diverse biological and molecular mechanisms observed in cells. Data produced from the study of genes and proteins are complementary and will yield a more complete panorama of how biological processes are regulated. Therefore, a greater understanding of the proteome of model organisms such as *Neurospora* is crucial not only to uncover molecular processes exclusive to the organism, but also because discoveries and knowledge generated from these models can be used to understand more specialized biological processes in higher organisms.

My dissertation research focused on determining the proteome of sexual growth for *N. crassa*. Establishing the proteome of sexual development will not only deepen our understanding of phenomena related to meiosis and sexual development, including MSUD, but it would also provide important insights into the process of cell differentiation in eukaryotes.

The long term goal of this work is to understand the molecular regulation of meiotic *trans*-sensing and silencing mechanisms in *N. crassa*. The specific objectives of my research were to: 1) Introduce biochemical approaches to the

study of meiotic silencing, 2) Determine the role that suppressor SMS-5 plays in meiotic silencing, 3) Establish a molecular connection between meiotic *trans*-sensing and meiotic silencing, and 4) Generate the first proteomic data set and functional catalog for *N. crassa* sexual development that will serve as a reference for the study of phenomena related to meiosis and sexual differentiation.

CHAPTER II

CHROMATIN REMODELING AND SIGNALING TRANSDUCTION

PATHWAYS ASSOCIATED WITH MEIOTIC SILENCING

INTRODUCTION

Meiotic Silencing by Unpaired DNA (MSUD) is a process in *Neurospora* that silences transcripts synthesized from an unpaired DNA region and any other transcripts with homology to the unpaired region (KELLY and ARAMAYO 2007). This process is composed of two stages: an initial nuclear stage, where unpaired DNA is detected by *trans*-sensing of homologous chromosomes (ARAMAYO and METZENBERG 1996), and a perinuclear stage where the actual silencing of the transcripts occurs. Silencing of the transcripts takes place in a sequence-specific manner. In the perinucleus, an antisense RNA molecule is required for the recognition and degradation of the target transcripts. Shiu and colleagues proposed that the antisense RNA is synthesized by the action of an RNA-dependent RNA polymerase (SAD-1), which synthesizes dsRNA (SHIU *et al.* 2001; SHIU *et al.* 2006). Recently, Hammond et al. reported evidence for the presence of antisense RNA fragments originated from an unpaired reporter gene (HAMMOND *et al.* 2013a).

The RNA molecule used by the RNA polymerase as a template for producing dsRNA is unknown. One possible scenario is that the synthesis of the RNA template occurs during the initial stage in the nucleus. The meiotic

silencing model proposes that an aberrant RNA (aRNA) molecule is synthesized from the detected unpaired DNA. This aRNA acts as a diffusible signal that is exported from the nucleus (PRATT 2008). Then, aRNA serves as the template for the synthesis of dsRNA, hence activating meiotic silencing (see Chapter I). Although most of the components involved in the silencing stage at the perinuclear region have been described, less is known about the nuclear stage. Therefore, the molecular basis behind the production of the aRNA and the system responsible for its transportation to the perinucleus needs to be determined.

If aRNA is synthesized directly from unpaired DNA, a special mark or signature characteristic of the unpaired DNA would be expected. Therefore, the machinery responsible for the production of aRNAs may recognize the DNA region that needs to be specially transcribed. Diverse organisms use epigenetic marks for similar purposes. Chemical modifications (*e.g.,* DNA, RNA and protein methylation) and structural modifications (*i.e.,* chromatin remodeling) are examples of epigenetic markers and are involved in gene regulation and silencing processes (BERNSTEIN *et al.* 2007; BONASIO *et al.* 2010).

In plants, DNA methylation and chromatin structure are common regulators of post-transcriptional gene silencing (PTGS) (MOREL *et al.* 2000). Methylated DNA transgenes trigger PTGS of homologous endogenous genes and inhibition of DNA methylation represses PTGS (KOVARIK *et al.* 2000). Therefore, it has been proposed that DNA methylation acts as a special

signature on the transgenes that may induce biosynthesis of aRNAs (BAULCOMBE 1996). In mammals, DNA methylation is associated with different phenomena, including X-chromosome inactivation and suppression of repetitive elements (BOURC'HIS and BESTOR 2004; BASU and ZHANG 2011). In *Neurospora*, DNA methylation contributes to the regulation of expression of repeated sequences. In the filamentous fungus, DNA methylation is closely tied to a genome defense mechanism named RIP, for Repeat Induced Point Mutation. RIP detects and introduces mutations to repeated sequences in the haploid phase of early sexual development. Thereafter, mutated sequences are recognized and methylated during vegetative growth (SELKER *et al.* 2003; GALAGAN and SELKER 2004).

DNA is not the only molecule that could be chemically labeled. RNA could also be chemically modified. More than a hundred chemical modifications are reported in different types of cellular RNAs (CANTARA *et al.* 2011), yet the role of these modifications are still unknown. It has been proposed that RNA modifications function as regulators of gene expression at the post-transcriptional level (HE 2010). Recently, a connection between RNA methylation and the correct processing of non-coding RNAs (ncRNAs) into small regulatory RNAs was found. These small RNAs participate in post-transcriptional regulation of gene expression in humans (HUSSAIN *et al.* 2013a). In addition, one of the components of mouse's chromatoid body—RNA-processing organelle (Chapter I)— is an RNA methyltransferase (HUSSAIN *et al.* 2013b).

Protein methylation is also an important chemical modification. SET domain methyltransferases catalyze the methylation of lysine residues in specific proteins, including histones. The SET domain is an evolutionary conserved sequence and is found in proteins of diverse functions ranging from yeast to mammals (DILLON *et al.* 2005; THORSTENSEN *et al.* 2011). Methylation of proteins is also associated with regulation of gene expression and chromatin conformation (DEL RIZZO and TRIEVEL 2011). For instance, histone methylations act epigenetically to activate or repress gene expression (KOUZARIDES 2007). In addition, a connection between different chemical modifications has been observed. For example, in *Neurospora*, histone methylation controls DNA methylation (TAMARU and SELKER 2001).

In the Aramayo lab, it was found that a SET-domain protein is a suppressor of meiotic silencing (*i.e.,* SMS-5). It is therefore reasonable to hypothesize that chemical modifications may be a key factor in meiotic silencing. Epigenetic signals like DNA or protein methylation might be responsible for labeling unpaired DNA/chromatin regions and activating aRNA synthesis. It is also possible that chemical modifications may confer stability to the aRNAs during their translocation to the perinuclear region. To date, DNA methylation by DIM-2 is the only chemical modification tested in meiotic silencing. Pratt *et al.* found that DIM-2-dependent DNA methylation affects meiotic *trans*-sensing—by enhancing RIP alleles recognition—but not meiotic silencing (PRATT *et al.* 2004). This observation suggests that DIM-2-dependent DNA methylation does not

work as an epigenetic signal for the unpaired DNA. Even though all known cytosine methylation in *N. crassa* are the product of DIM-2, we cannot discard the possibility that other undescribed DNA methyltransferase would be responsible for DNA methylation during meiosis.

It is of great interest to study whether DNA, RNA and/or protein chemical modifications are part of the molecular mechanism driving meiotic silencing. The SMS-5 protein is an excellent candidate to start with for this investigation. SMS-5 is a SET-domain protein that is not essential for sexual development, but is required for meiotic silencing.

In a previous study, histone methyltransferase (HMTase) activity was not detected when purified SMS-5 protein was tested (Lee and Aramayo, unpublished). This result suggests that SMS-5 has a different protein target for methylation or a different function. The aim of this research was to study the role that SMS-5 plays in meiotic silencing. I hypothesize that SMS-5 is involved in the initial nuclear stage of unpaired DNA silencing. In order to determine SMS-5 function, I identified possible SMS-5 target protein(s). I developed a set of biochemical experiments that allowed me to study protein interactions in *N. crassa* sexual tissue. I constructed and affinity purified a GST-SMS-5 recombinant protein. This purified protein was used for pulling-down SMS-5-interacting proteins from whole cell extract from *N. crassa* sexual tissue. SMS-5-binding proteins were identified by mass spectrometry (MS) analysis. Interactions were verified by *in vitro* binding assay and protein binding partners

for SMS-5 were established. Determining SMS-5 interacting partners revealed other aspects of the molecular mechanism of the nuclear stage of meiotic silencing.

RESULTS

Preliminary Data

In *N. crassa*, the product of the NCU02088 gene (Broad Institute) is required for silencing unpaired reporter genes, but is not needed for meiosis. The gene was named *Sms-5* (S*uppressor of meiotic silencing-5*). Crosses homozygous for *Sms-5* mutants undergo normal meiosis (Lee and Aramayo; unpublished data). This is different from what has been observed with the majority of meiotic silencing suppressors, suggesting not only that the SMS-5 protein function is directly involved in meiotic silencing, but also that meiotic silencing is not required for meiosis or ascus development.

SMS-5 is Localized to Both the Nucleus and the Perinucleus

Meiotic silencing occurs in two sequential stages at different cellular compartments. In the nucleus, unpaired DNA is detected via *trans*-sensing; then in the perinucleus, the transcripts generated from the unpaired DNA are silencing. Therefore, to determine the cellular localization of the protein is the first step in order to establish in which stage of meiotic silencing the protein participates. Cellular localization of the SMS-5 protein was tested using

fluorescent reporter genes fused to *sms-5+* (*i.e., sms-5+::mKO+* and *sms-5+::sGFP+*). Fusion proteins expressed at the canonical position were partially functional in meiotic silencing, but the fluorescent signal was not observed (Lee and Aramayo, unpublished data). These results suggested that cellular SMS-5 accumulation might be low under normal conditions, due either to low protein biosynthesis or high protein turnover rate. Another possibility is that the fusion could affect normal SMS-5 protein biosynthesis and/or regulation. To overcome the possibility of low protein biosynthesis, the fusion construct was introduced into an ectopic position under the regulation of the *ccg-1* (*clock controlled gene-1*) promoter (HONDA and SELKER 2009). Although a fluorescent signal was detected in the perinuclear region, the signal was very weak (Figure 2.1 (A)) (Lee and Aramayo, unpublished data).

Based on previous results, we cannot discard the possibility that SMS-5 localizes in the nucleus; but cannot be detected due to technical limitations. For that reason, I used bioinformatics analysis to predict SMS-5 subcellular localization using PredictProtein automatic service (http://www.predictprotein.org) (GOLDBERG *et al.* 2012). In addition, a nuclear localization signal (NLS) analysis using the cNLS Mapper software predicted a bipartite NLS at the C-terminus of SMS-5. The bipartite signal received a score of 3.9 (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi). Using this software, proteins with a score between 3 and 5 are predicted to likely localize to both the nucleus and the cytoplasm (KOSUGI *et al.* 2009) (Figure 2.1 (B)).

Together these results suggest that SMS-5 may be present in both cellular compartments: nucleus and perinucleus.
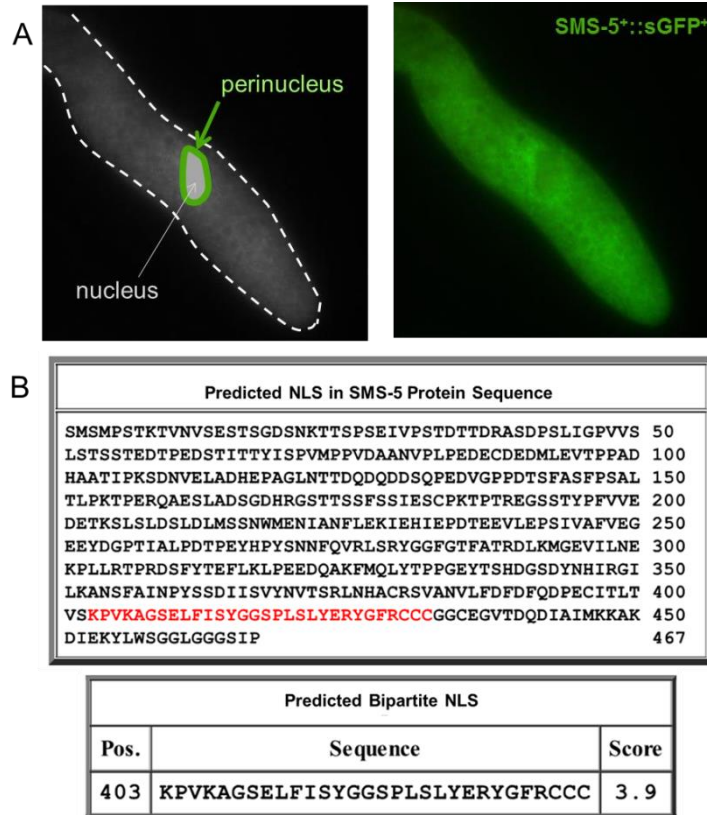
Figure 2.1 SMS-5 cellular localization. (A) Displayed image is from Prophase I. An individual ascus expressing the fusion protein SMS-5$^+$::sGFP$^+$ is shown. Left image is a cartoon of the right picture. (B) Predicted Nuclear Localization Signal (NLS) in SMS-5 using cNLS Mapper.

SMS-5 is a Novel SET Domain Hypothetical Protein Found in Some Ascomycetes and a few Basidiomycetes Fungi

The translated SMS-5 sequence consists of 465 amino acid residues with a molecular weight of 50.6 kDa. Although SMS-5 has a putative SET domain, it is a novel hypothetical fungal protein only found in a few members of the subkingdom Dikarya. A BLAST analysis of SMS-5 against non-redundant protein sequences displayed the highest percentage of sequence identity within Ascomycetes, especially within the order of Sordariales, *i.e., N. tetrasperma* (84%) and *Sordaria macrospora* (58%). Nevertheless, a conserved region exists among other *Sordaryomycetes* at the SET domain position (250-450 amino acid residues). Additionally, very low sequence similarity was observed with few Basidiomycetes (Figure 2.2; Table 2.1).

BLAST analysis using the UniProt/SwissProt protein database revealed very poor homology of SMS-5 with proteins from other eukaryotes. This data suggests that the SET domain of SMS-5 is conserved among Ascomycetes; however, the entire SMS-5 protein is highly conserved among some Sordariales fungi.
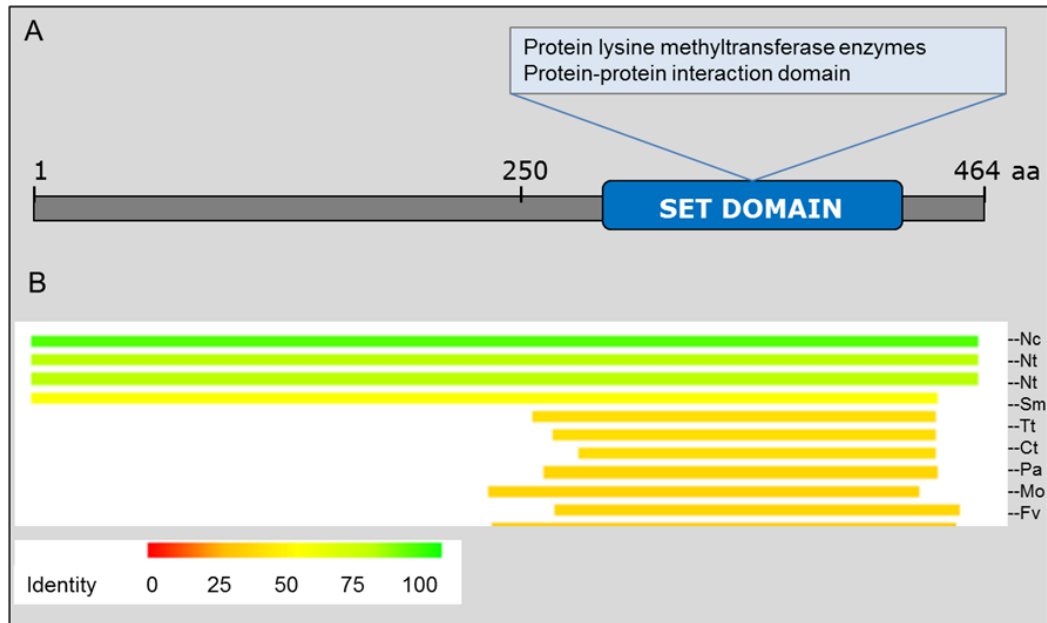
Figure 2.2 SMS-5 protein sequence analysis. (A) Conserved domain analysis results. A putative SET domain was detected at the C-terminus (~ 130 amino acid residues). (B) BLAST results. Database sequence hits are shown aligned to the query sequence (SMS-5) displayed in (A). Of the aligned sequences, the most similar are shown closest to SMS-5. Scoring matches are represented in colors. Initials of organism's names are shown next to the aligned sequences. Nc, *Neurospora crassa*; Nt, *Neurospora tetrasperma*; Sm, *Sordaria macrospora*; Tt, *Thielavia terrestris*; Ct, *Chaetomium thermophilum*; Pa, *Podospora anserina*; Mo, *Magnaporthe oryzae*; Fv *Fusarium verticillioides*.

Table 2.1 SMS-5 and homologous proteins

Organism Scientific Name	Source[a]	ID Number[b]	Score	% Identity[c]	Division
Neurospora crassa	Broad	NCU02088.7	947	100	Ascomycetes
Neurospora tetrasperma	Broad	NEUTE1DRAFT_54310	774	84	Ascomycetes
Sordaria macrospora	Broad	SMAC_04985	506	58	Ascomycetes
Thielavia terrestris	Broad	THITE_2091683	138	40	Ascomycetes
Chaetomium thermophilum	NCBI	XP_006696199.1|	135	39	Ascomycetes
Myceliophthora thermophila	Broad	MYCTH_2122106	128	40	Ascomycetes
Grosmannia clavigera	NCBI	GL629765.1	118	37	Ascomycetes
Gaeumannomyces graminis	Broad	GGTG_00363	114	33	Ascomycetes
Ophiostoma piceae	NCBI	EPE06191.1	112	36	Ascomycetes
Magnaporthe oryzae	Broad	MGG_00673	110	32	Ascomycetes
Podospora anserina	NCBI	XP_001912535.1	105	35	Ascomycetes
Fusarium verticillioides	Broad	FVEG_16600	102	34	Ascomycetes
Trichoderma virens	Broad	TRIVIDRAFT_157176	74	46	Ascomycetes
Trichoderma reesei	NCBI	ETS07049.1	73	43	Ascomycetes
Heterobasidion irregulare	NCBI	ETW78294.1	73	29	Basidiomycetes
Ceriporiopsis subvermispora	NCBI	EMD31248.1	62	24	Basidiomycetes
Arthroderma benhamiae	NCBI	XP_003016650.1	60	28	Basidiomycetes

[a]Broad = Broad Institute (http://www.broadinstitute.org/); NCBI = National Center of Biotechnology Information (http://www.ncbi.nlm.nih.gov/)
[b]ID Number according to the source
[c]Percentage of identical amino acid residues

Although SMS-5 is a unidomain protein, the SET domain is frequently present in multi-domain nuclear proteins with diverse functions (JENUWEIN *et al.* 1998). This domain is associated with protein lysine methyltransferases. The most studied ones are histone methyltransferases (HMTs); however, SET-domain proteins also methylate non-histone proteins (DILLON *et al.* 2005). For instance, human SET7/9 methylates the transcription factor TAF10, resulting in an increase affinity for RNA polymerase II and transcriptional activation of TAF10-dependent genes (KOUSKOUTI *et al.* 2004). SET7/9 also methylates the tumor-suppressor protein p53 and increases its stability (CHUIKOV *et al.* 2004). In addition, the SET domain is a protein-protein interaction domain; it mediates conserved interactions with dual-specific lipid phosphatases (CUI *et al.* 1998). A subset of SET domains, called PR domains, is involved in protein-protein interaction (HUANG *et al.* 1998). Because of this diversity in SET domain function, knowing the binding partners for SMS-5 may help in uncovering the function of this suppressor.

Identification of Potential Protein Binding Partners for SMS-5

Due to the cellular localization of SMS-5 and the predicted functions of the SET domain, I hypothesize that SMS-5 is an important component for the initial nuclear stage of meiotic silencing. This protein might be involved in the synthesis and/or processing of aRNAs. It may also be the connection between the nuclear and the perinuclear stages of this silencing mechanism. Therefore, it

was my interest to investigate the role that this SET-domain protein has in a mechanism like meiotic silencing. For that purpose, I developed a biochemical strategy to identify potential protein binding partners for SMS-5 (Figure 2.3). I used a pull-down biochemical assay as a discovery method for identifying unknown interactions.

Glutathione S-transferase (GST) tag was fused to *Sms-5* and cloned into pGEX-6P-1 vector. GST-SMS-5 recombinant protein was expressed in *Escherichia coli* cells. Protein was purified by affinity chromatography and immobilized in a solid matrix of glutathione agarose beads (GST-SMS-5 column). As a control column, referred to as GST column, only GST was purified and immobilized. Whole cell extract from *N. crassa* sexual tissue was loaded into the experimental and control columns. Affinity purification of the SMS-5 interacting proteins was achieved by large-scale capture and elution of binding proteins with reduced glutathione and PreScission protease. Proteins were resolved by SDS-PAGE and stained with Coomassie blue. The elute fraction from the GST-SMS-5 column contained several protein bands that were not present in the GST column (Figure 2.4). The high number of observed bands suggested that—in addition to the proteins that directly interact with SMS-5—proteins that indirectly interact were also detected.
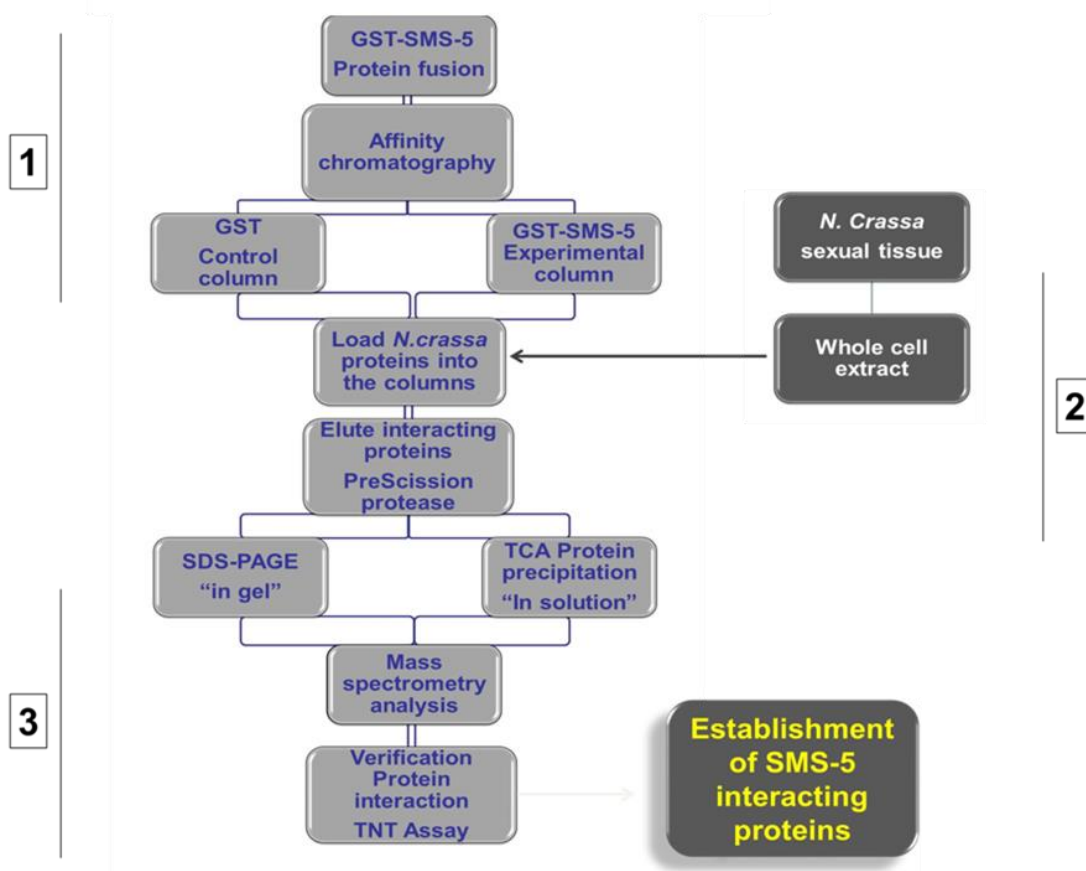
Figure 2.3 Strategy for identifying possible protein binding partners for SMS-5. The strategy was split in three modules: 1) Preparation of affinity columns. Recombinant proteins were expressed in *E.* coli cells. GST-SMS-5 (experimental) and GST (control) proteins were purified by affinity chromatography and immobilized in glutathione agarose beads; 2) Pull-down assay. Whole cell extract from *N. crassa* sexual tissue was prepared and loaded into the two columns. After several washes, binding proteins were eluted; 3) Protein identification. Half of the elution fraction was resolved in a protein gel, while the proteins in the other half were precipitated and concentrated. The two samples ("in gel" and "in solution") were analyzed by mass spectrometry (MS) and followed by protein identification. Protein interactions were verified using an *in vitro* transcription and translation (TNT) assay.
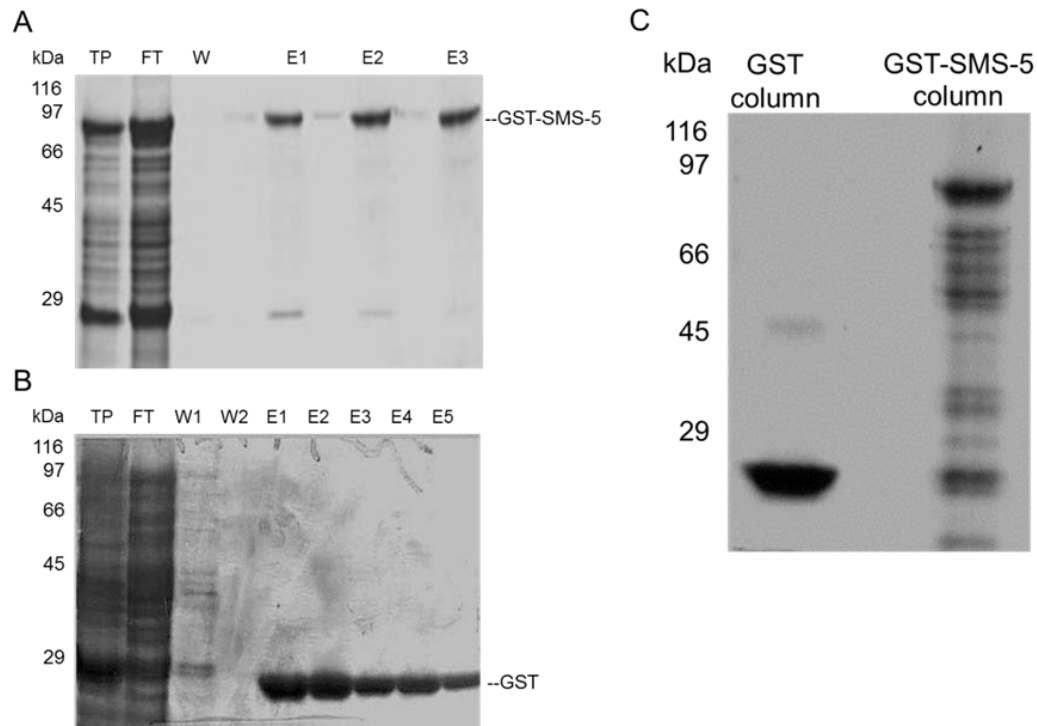
Figure 2.4 Affinity protein purification and pull-down assay results. SDS-PAGE analysis of the affinity chromatography and purification of (A) recombinant protein (GST-SMS-5) and (B) control (GST). TP, total protein; FT, flow-through; W, wash; E, elution fraction. (C) SDS-PAGE analysis of one of the pull-down assays. Whole cell extract from *N. crassa* sexual tissue was passed through the GST and GST-SMS-5-immobilized columns. Non-bound proteins were washed off the column and the binding proteins were eluted. Elution fractions were resolved by SDS-PAGE and visualized by Coomassie blue staining.

A fraction of eluted proteins was resolved by SDS-PAGE (*i.e.,* in gel samples); the remaining fraction was precipitated and proteins concentrated (*i.e.,* in solution samples). Proteins isolated from both the in gel and in solution samples were identified by mass spectrometry analysis. Results from three independent experiments were analyzed (Figure 2.5; Table 2.2).
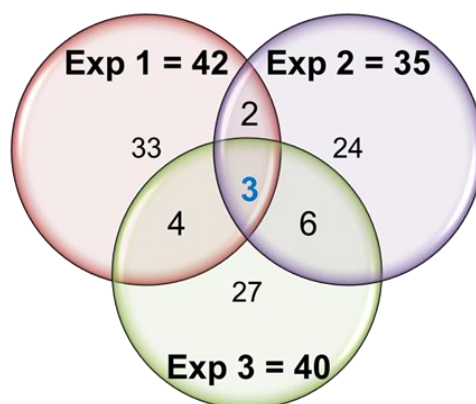
Figure 2.5 Venn diagram showing the results from three independent pull-down experiments. Proteins that bound specifically to GST-SMS-5 column and identified by mass spectrometry with a Confident Interval (C.I. %) > 95% are represented (Appendix A; Table A.1).

I determined potential SMS-5-interacting proteins by comparing results from the two columns, GST-SMS-5 and GST. Only proteins bound specifically to the GST-SMS-5 column were used for further analysis. These proteins were classified depending on the number of experiments where they were detected. A total of 99 proteins were detected with a Total Ion Score and/or Protein Score Confidence Interval (C.I.%) higher than 95% (Appendix A; Table A.1). From these proteins, only three were consistently pulled-down in all three experiments, including the bait protein SMS-5. The other two proteins are: 1) the histone acetylase complex subunit PAF400 (NCU01379); a *Neurospora* homolog of the yeast Tra1 protein, and 2) the cytosolic regulator Pianissimo (NCU07854). In addition, 12 proteins were found in two pull-downs (Table 2.2).

Table 2.2 Proteins recovered from more than one pull-down assay

Sequence[a] ID #	Protein Description	Frequency[b]	Experiment 1 C.I. %[c]		Experiment 2 C.I. %[c]		Experiment 3 C.I. %[c]	
			In solution[d]	In gel[e]	In solution[d]	In gel[e]	In solution[d]	In gel[e]
NCU02088	Suppressor of Meiotic Silecing-5 SMS-5	3	100	99.8	100	99.2	100	100
NCU01379	histone acetylase complex subunit PAF400	3	100		99.9	22	99.6	
NCU07854	cytosolic regulator Pianissimo	3		97.5	99.9		99.5	
NCU01680	plasma membrane ATPase-1 PMA-1	2	100				100	
NCU08936	clock-controlled gene-15 CCG-15	2	100				100	
NCU01323	cohesin complex subunit	2	100				99.8	
NCU10021	high affinity glucose transporter-1 HGT-1	2			100		100	
NCU05488	RNA-binding protein Vip1	2			99.9	18.3	100	
NCU07554	chromosome segregation protein SudA	2	99.9		99.9			
NCU04865	polyketide synthase-3	2	99.9	0		70.4		
NCU01634	histone H4-1	2		99.9			99.9	0
NCU08600	hypothetical protein	2			99.9		99.3	
NCU10346	hypothetical protein	2			99.9		98.5	
NCU06701	cephalosporin C regulator 1 RFX	2				57.1	99.8	
NCU03072	hypothetical protein	2				0	99.1	

[a]Broad = Broad Institute (http://www.broadinstitute.org/).
[b]Frequency, total number of times the protein was detected.
[c]C.I. %, Total Ion Score and/or Protein Score Confidence Interval. "0" under this column means that the protein was detected but the C.I. % was zero or lower.
[d]In solution, eluted proteins were precipitated and concentrated. The whole protein mix was enzymatically digested followed by MS identification.
[e]In gel, eluted proteins were separated by SDS-PAGE. Protein bands were cut from the gel, enzymatically digest followed by MS identification.

Interestingly, although the HMTase assay showed no methylation of histones in the presence of SMS-5, the histone H4 was recovered as a possible protein binding partner for SMS-5. Assuming that indeed an interaction between SMS-5 and hH4 exists, the negative result previously observed in the HMTase assay could be explained by two possibilities. First, SMS-5 has another function different from transferring methyl groups to histones. Second, SMS-5 might need an unknown cofactor to transfer the methyl group. A cofactor requirement has been observed in other systems (TAN *et al.* 2006).

From this exploratory phase, I expected to recover new suppressors of meiotic silencing and reveal possible interactions between the already known components of this silencing phenomenon. Suppressors SMS-2 (Argonaute protein) and SMS-9 (a helicase ortholog of the human ATRX protein) were present in one of the pull-downs. In addition, the DNA helicase QDE-3 was also identified. QDE-3 is required for quelling, the other PTGS mechanism present in *N. crassa*. (Appendix A; Table A.1). Although these proteins were only recovered once, detecting other suppressors supported the biochemical strategy applied in this study.

For further analysis, I selected three of the possible SMS-5-interacting proteins. The criteria for selecting these proteins were: 1) frequency of protein recovery from the pull-down assays and 2) protein predicted function. Therefore, I sought to determine the interactions between SMS-5 and the *Neurospora* homolog of the yeast Tra1 protein PAF400, and the cytosolic regulator

Pianissimo. Both proteins were detected with C.I. % > 95% in all three experiments. Because meiotic silencing is a mechanism that involves RNA molecules, the third protein selected was an RNA-binding protein identified from two pull-down assays.

Testing the Involvement of an RNA-Binding Protein in Meiotic Silencing

A protein similar to the RNA-binding protein Vip1 in *Schizosaccharomyces pombe* was identified and selected as a possible interacting partner for SMS-5. In *Neurospora*, this RNA-binding protein is NCU05488 (from now on, referred to as RBP). RBP was detected twice from the pull-down assays and characterized by MS with C.I. % > 99.9%. The translated sequence consists of 284 amino acids with a molecular weight of 30.3 kDa. RBP has an RNA recognition motif domain (*i.e.,* RRM domain) at its N-terminus (Figure 2.6). RRM-containing proteins bind single-stranded RNAs and are involved in post-transcriptional gene expression processes, including mRNA and rRNA processing, RNA export and RNA stability in different eukaryotes (DAUBNER *et al.* 2013). The RRM domain is not only involved in RNA recognition but also in protein-protein interaction and is one of the most abundant domains in eukaryotes (CLERY *et al.* 2008).
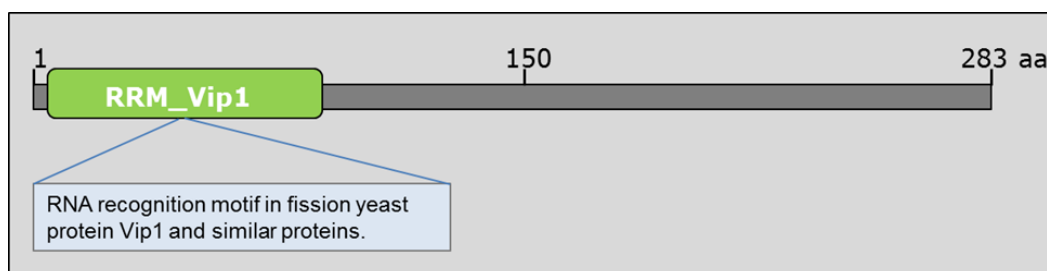
Figure 2.6 Protein structure of the RNA-binding protein NCU05488. The protein has an RNA recognition motif (RRM) domain localized at the N-terminal region. RRM domain is highly abundant in eukaryotes and is found in proteins involved in mRNA and rRNA processing, RNA export and RNA stability.

To study protein conservation and also infer a possible protein function for this RNA-binding protein, I did a protein BLAST analysis. First, the RBP protein sequence was compared against the non-redundant protein sequence database, but excluding all fungal proteins. Higher match scores were not evident. Only low protein similarities were observed with proteins from few amoebas (kingdom protozoa) and some eudicotyledons (kingdom plantae). Protein BLAST analysis against the entire non-redundant protein database showed high sequence similarity with other fungal proteins, especially from Ascomycota. This result suggested protein conservation between Ascomycota RRM-domain proteins. The highest percentage of protein identity was found among Sordariales (*i.e., N. tetrasperma* (99%) and *S. macrospora* (92%)). Hence, NCU05488 had the same conservation pattern observed with SMS-5,

suggesting that both proteins may evolve together and might be involved in similar mechanisms in these fungi.

Because my hypothesis states that SMS-5 participates in the initial nuclear stage of meiotic silencing, and the nuclear stage is associated with synthesis and translocation of aRNAs, it is reasonable to expect RNA binding molecules associated directly or indirectly with SMS-5. For that reason, I sought to evaluate whether RBP (NCU05488) was involved in meiotic silencing. To do that, I followed a genetic strategy.

A strain deletion mutant for $NCU05488^\Delta$ is available at the *N. crassa* knockout library (*i.e.,* FGSC21956), indicating that this is not an essential protein for vegetative development. I evaluated whether the *NCU05488* gene was essential for sexual development. For that purpose, I constructed strains of both mating types carrying the deletion at *NCU05488* locus and studied the phenotype and ascospore production of homozygous crosses (Figure 2.7; cross 4). For comparison, wild-type crosses were also analyzed. Sexual development and ascospore production proceeded normally on those crosses, suggesting that *NCU05488* gene product was not essential for sexual development.

Then, I investigated whether NCU05488 is required for meiotic silencing. If this were the case, then I expected to observe a suppression of silencing of an unpaired reporter gene in crosses carrying $NCU05488^\Delta$ deletion mutant. To test this possibility, I evaluated meiotic silencing of an unpaired reporter gene in heterozygous and homozygous crosses for $NCU05488^\Delta$ deletion mutant. I used

Round spore (*Rsp*) gene as a reporter. To create the unpaired condition and induce meiotic silencing, I worked with a dominant RIP allele of the *Rsp* gene (*i.e.*, *Rsp*^{RIP93}) (PRATT *et al.* 2004). RIP alleles carry G:C to A:T transition mutations and cytosine methylations. Therefore, RIP alleles are homeologous to their counterpart wild-type alleles. All crosses performed in this study are listed in table 2.3.

I observed ascospore morphology and quantified the number of spindle-shaped asocospores (wild-type phenotype) versus round-shaped ascospores (mutant phenotype) produced in each cross (Figure 2.7). Wild-type crosses served as control to show there was not silencing in paired DNA conditions. It produced 100% of spindle spores (*rsp*$^+$ X *rsp*$^+$; cross 1). Homozygous crosses for *Rsp*^{RIP93} served as control to show that *Rsp* gene product is responsible for spindle-shaped spore morphology. It produced almost 0% of spindle spores (*Rsp*^{RIP93} X *Rsp*^{RIP93}; cross 2). Heterozygous crosses for *Rsp*^{RIP93} served as control to show meiotic silencing in the presence of unpaired DNA. It produced 13.5% of spindle spores, demonstrating that silencing of a fraction of the transcripts from the wild-type allele occurred (*Rsp*^{RIP93} X *rsp*$^+$; cross 3). Homozygous crosses for *NCU05488*$^{\Delta}$ allowed evaluating the requirement of *NCU05488* gene product for sexual development. Cross 4 (*NCU05488*$^{\Delta}$ X *NCU05488*$^{\Delta}$) and cross 5 (Rspr^{RIP93}; NCU05488$^{\Delta}$ X Rspr^{RIP93}; NCU05488$^{\Delta}$) both ended with abundant ascospores production, indicating that NCU05488 is not essential for sexual growth.
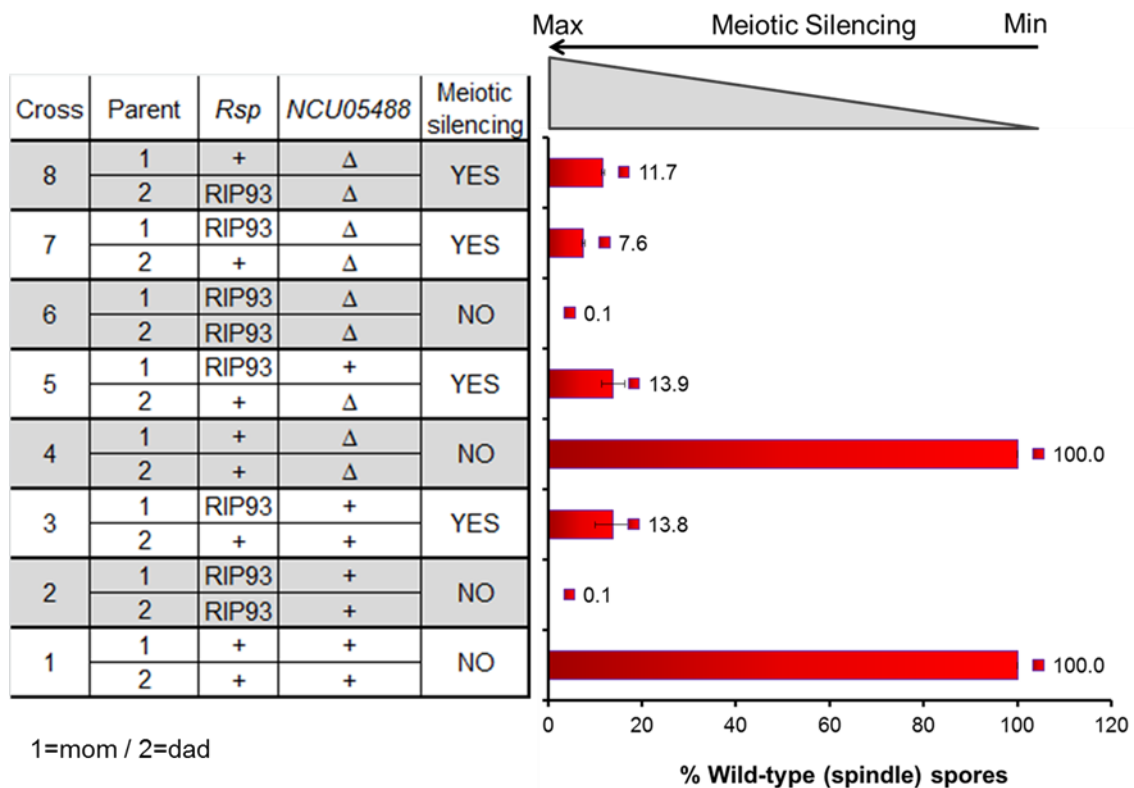
Figure 2.7 The RNA binding protein NCU05488 is not required for silencing homeologies. Rsp^{RIP93} allele was used as a reporter of meiotic silencing. Wild-typed crosses and Rsp^{RIP93} homozygous crosses were used as controls. Heterozygous and homozygous crosses for $NCU05488^{\Delta}$ were evaluated in the presence and absence of the unpaired reporter gene. Each percentage number indicated is the average of three crosses with a mean of 1430 ascospores counted per cross.

Table 2.3 Crosses analyzed during the evaluation of the involvement of NCU05488 in meiotic silencing

Cross Number	Strains		Relevant Diploid Genotype	Brief Description
	Female (mating type A)	**Male (mating type a)**		
1	FGSC2489	FGSC2490	wt / wt	Control. Paired rsp^+ allele; No meiotic silencing
2	RPNCR074	RPNCR073	Rsp^{RIP93} / Rsp^{RIP93}	Control. Paired Rsp^{RIP93} allele; No meiotic silencing
3	RPNCR074	FGSC2490	Rsp^{RIP93} / wt	Control. Unpaired Rsp^{RIP93} allele; Meiotic silencing
4	VSNCR063	VSNCR060	$NCU05488^{\Delta}$ / $NCU05488^{\Delta}$	Test $NCU05488$ requirement for sexual development
5	VSNCR068	VSNCR067	Rsp^{RIP93}; $NCU05488^{\Delta}$ / Rsp^{RIP93}; $NCU05488^{\Delta}$	Test $NCU05488$ requirement for sexual development in Rsp^{RIP93} homozygous condition
6	RPNCR074	FGSC21956	Rsp^{RIP93} / $NCU05488^{\Delta}$	Test meiotic silencing in $NCU05488^{\Delta}$ heterozygous condition; unpaired Rsp^{RIP93} allele
7	VSNCR068	VSNCR060	Rsp^{RIP93}; $NCU05488^{\Delta}$ / $NCU05488^{\Delta}$	Test meiotic silencing in $NCU05488^{\Delta}$ homozygous condition; unpaired Rsp^{RIP93} maternal allele
8	VSNCR063	VSNCR067	$NCU05488^{\Delta}$ / Rsp^{RIP93}; $NCU05488^{\Delta}$	Test meiotic silencing in $NCU05488^{\Delta}$ homozygous condition; unpaired Rsp^{RIP93} paternal allele

To test for the effect of heterozygous $NCU05488^\Delta$ deletion mutant in meiotic silencing, cross 6 (Rsp^{RIP93} X $NCU05488^\Delta$) was analyzed. If NCU05488 was involved in meiotic silencing, I expected an increase in the number of spindle spores compared to cross 3. However, I observed a similar percentage of spindle spores (13.9%). This data indicated that heterozygous $NCU05488^\Delta$ did not have an effect on meiotic silencing. There were two explanations for the results observed in heterozygous crosses. First, NCU05488 was not involved in meiotic silencing. Second, the presence of one copy of the wild-type $NCU0588^+$ gene was enough to proceed with meiotic silencing, indicating that the gene itself escaped from silencing. If this were the case, a basal expression level of this gene produced enough protein to participate in meiotic silencing. To test these possibilities, homozygous crosses for $NCU05488^\Delta$ were analyzed in the presence of Rsp^{RIP} alleles. Those were cross 7 (Rsp^{RIP93}; $NCU05488^\Delta$ / $NCU05488^\Delta$) and cross 8 ($NCU05488^\Delta$ / Rsp^{RIP93}; $NCU05488^\Delta$). There was no increase in the number of spindle spores in these crosses. In fact, I observed a decrease in spindle spore number. The difference was not statistically significant. Therefore, I conclude that the RNA-binding protein NCU05488 was not required for silencing Rsp^{RIP93} unpaired DNA.

Although NCU05488 gene product is not necessary for silencing homeologies (RIP alleles), it is still possible that this protein play a role in silencing other types of unpairedness, such as heterology (product of a genetic

deletion or insertion). Therefore, I continued investigating the interaction between SMS-5 and NCU05488.

To determine physical interactions between proteins, I performed *in vitro* protein binding and co-immunoprecipitation (co-IP) assays. Proteins were expressed in rabbit reticulocyte lysate (RRL) as T7-tagged proteins and/or ^{35}S-methionine radiolabeled proteins. Both proteins were incubated together to allow interaction and then immunoprecipitated using T7-tag antibody beads. Precipitates were collected and analyzed by SDS-PAGE. T7-tagged protein binding partners were detected by autoradiography. As the negative control, radiolabeled proteins were incubated with the beads in the absence of the T7-tagged protein. As the positive control, Ku70-Ku80 protein interaction was evaluated (RIHA *et al.* 2002).

SMS-5 and NCU05488 were expressed in RRL as T7-tagged protein and ^{35}S-methionine radiolabeled protein, respectively. After mixing both proteins and incubated together, T7-tagged SMS-5 was immunoprecipitated using T7-tag antibody beads. Co-immunoprecipitation of radiolabeled NCU05488 was evaluated by SDS-PAGE and autoradiography. As shown in Figure 2.8, radiolabeled NCU05488 was not precipitated in the presence of T7-tagged SMS-5 (lanes 4 and 6). In control reactions, no interaction was seen between T7 antibody and untagged NCU05488 (Figure 2.8, lane 2). As positive control, interaction between radiolabeled Ku70 and T7-tagged Ku80 was observed

(Figure 2.8, lanes 11-14). Therefore, I conclude that NCU05488 and SMS-5 do

not interact directly *in vitro*.
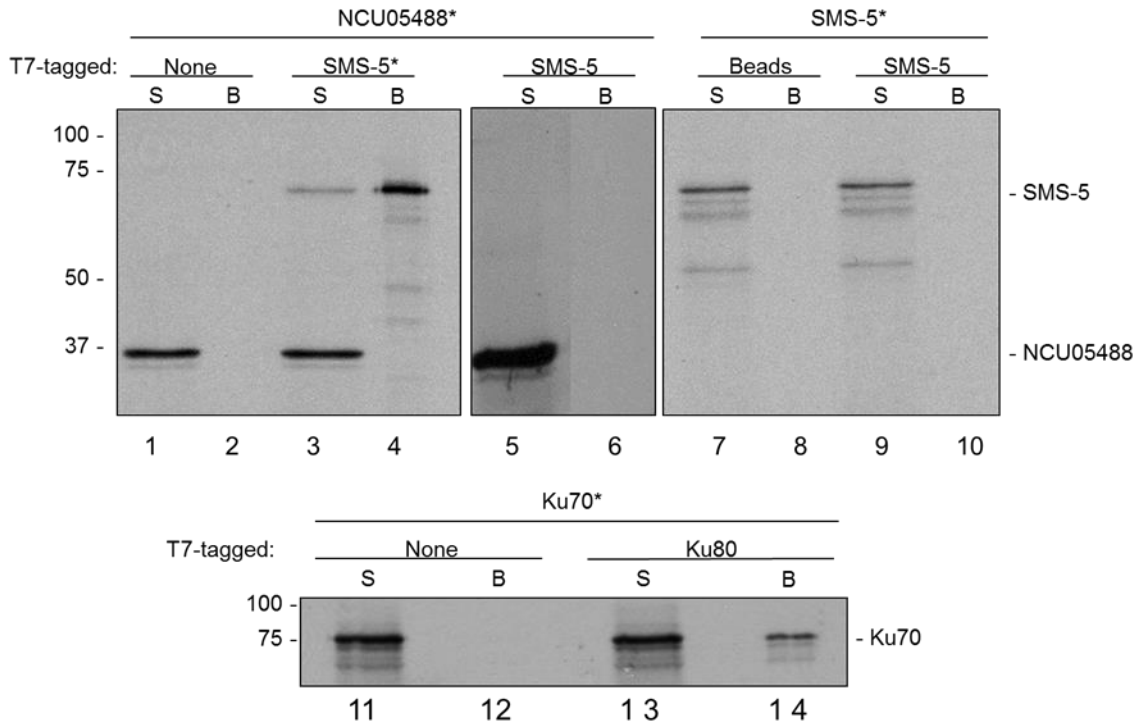
Figure 2.8 *In vitro* SMS-5 binding assay. Co-IP was conducted with ^{35}S-methionine labeled (asterisk) and T7-tagged protein expressed in RRL. Radiolabeled protein is co-immunoprecipitated on T7-beads when bound to a tagged partner. Precipitates were collected and analyzed on 10% SDS-PAGE, followed by autoradiography. S, supernatant; B, beads. Radiolabeled NCU05488 is not co-immunoprecipitated with T7-tagged SMS-5. Ku70-Ku80 interaction served as Co-IP positive control.

Many multi-protein complexes contain homodimers, such as nucleosome (BENTLEY *et al.* 1984) and proteasome (BOCHTLER *et al.* 1999). It has been demonstrated that homodimers have twice as many interacting partners than non-self-interacting proteins (ISPOLATOV *et al.* 2005). To test whether SMS-5 forms homodimers that contribute to the interaction with other proteins, SMS-5 was expressed in RRL as T7-tagged proteins or ^{35}S-methionine radiolabeled proteins. No interaction between radiolabeled SMS-5 and T7 antibody was observed in the control reaction (Figure 2.8, lane 8). Because radiolabeled SMS-5 was not precipitated with T7-tagged SMS-5 (Figure 2.8, lane 10), I conclude that SMS-5 does not form homodimers.

An indirect protein interaction between SMS-5 and NCU05488 could be the reason for the negative result observed in the *in vitro* co-IP. Indirect interactions can be detected by pull-down assays when multi-protein complexes are precipitated. However, the *in vitro* protein binding assay tested for direct physical interaction between the two proteins. Another possibility could be the necessity of post-translational protein modifications that induce protein interactions. Because I used an *in vitro* system for protein synthesis, post-translational modifications were not taking in consideration. Together these results suggested that the homolog of RNA binding protein Vip1 was not a direct protein binding partner for SMS-5 *in vitro* and it was not required for silencing of unpaired *Rsp*^{RIP93} allele.

71

The Histone Acetylase Complex Subunit PAF400 is a Protein Binding Partner for SMS-5

Two proteins were considered strong candidates as interacting partners for SMS-5, because they were detected in all three affinity purification assays. One of them was the homolog of the histone acetylase complex subunit PAF400 (NCU01379). Recently, in a global analysis of protein kinases in *N. crassa*, this protein was classified as an atypical serine/threonine-protein kinase (*stk-18* gene) (PARK *et al.* 2011). For simplicity, I will refer to this protein as PAF400.

The translated PAF400 sequence consists of 3896 amino acid residues with a molecular weight of 443.3 kDa. The predicted amino acid sequence of PAF400 was compared to the UniProt/SwisProt protein database. BLAST analysis revealed that PAF400 is homologous to Tra1/TRRAP proteins, which are highly conserved and present in several eukaryotes, including yeast, flies, humans, and mice. PAF400 is between 26% and 36% identical to its homologs. However, when conservative substitutions are allowed, the total similarities between PAF400 and its homologs are higher than 50%. This data suggested that PAF400 is related to Tra1/TRRAP proteins (Table 2.4). PAF400 homologous proteins are major components of the nucleosomal acetyltransferase protein complexes SAGA and NuA4 (GRANT *et al.* 1998; MCMAHON *et al.* 1998; BROWN *et al.* 2001). Both complexes have important roles in chromatin remodeling, transcriptional regulation and DNA double-strand breaks (DSBs) repair (BIRD *et al.* 2002; ROBERT *et al.* 2006).

Table 2.4 *N. crassa* histone acetylase complex subunit PAF400 and homologs

Organism Scientific Name	Protein name	Sequence length (aa)	Uniprot Accession #	Identity (%)	Similarity (%)
Neurospora crassa	Histone acetylase complex subunit PAF400	3896	Q7S7K6	100	100
Saccharomyces cerevisiae	Transcription-associated protein 1 (Tra1)	3744	P38811	36	56
Schizosaccharomyces pombe	Transcription-associated protein 1 (Tra1)	3699	Q9HFE8	34	52
Schizosaccharomyces pombe	Uncharacterized PI3/PI4-kinase family protein (Tra2)	3655	Q10064	34	53
Drosophyla melanogaster	Transcription-associated protein 1 (dTRA1)	3803	Q8I8U7	27	45
Homo sapiens	Transformation/transcription domain-associated protein (TRRAP)	3859	Q9Y4A5	28	47
Mus musculus	Transformation/transcription domain-associated protein (Trrap)	2565	Q80YV3	22	45
Dictyostelium discoideum	Probable transcription-associated protein 1 (Tra1)	4582	Q54T85	28	46

Tra1/TRRAP proteins are very large proteins and belong to the phosphatidylinositol 3 kinase-related kinase (PIKK) family of proteins, which regulate a diverse set of signaling pathways (HELMLINGER *et al.* 2011). The PIKK family includes the key cellular regulators serine/threonine-protein kinases, such as ATM, ATR, DNA-PKcs, and TOR (LOVEJOY and CORTEZ 2009).

PIKK proteins consist of three distinct domains found at the extreme C-terminus of these very large proteins: The FAT domain (named after representative proteins sharing the domain (FRAP, ATM and TRRAP)), PI3-kinase catalytic domain, and FATC (FAT-C-terminus) domain (BOSOTTI *et al.* 2000). Tra1/TRRAP proteins also have these three domains; however, the PI3-kinase domain is catalytically inactive since it lacks the DXXXXN and DFG motifs, critical residues necessary for protein phosphorylation (MCMAHON *et al.* 1998). However, the PI3-kinase domain in Tra1/TRRAP proteins conserves the three-dimensional folding structure that has been proposed to be necessary for mediating protein-protein interaction (MUTIU *et al.* 2007). A phylogenetic analysis using several representative PIKK proteins demonstrated that PAF400 cluster with Tra1/TRRAP sub-group (Figure 2.9).
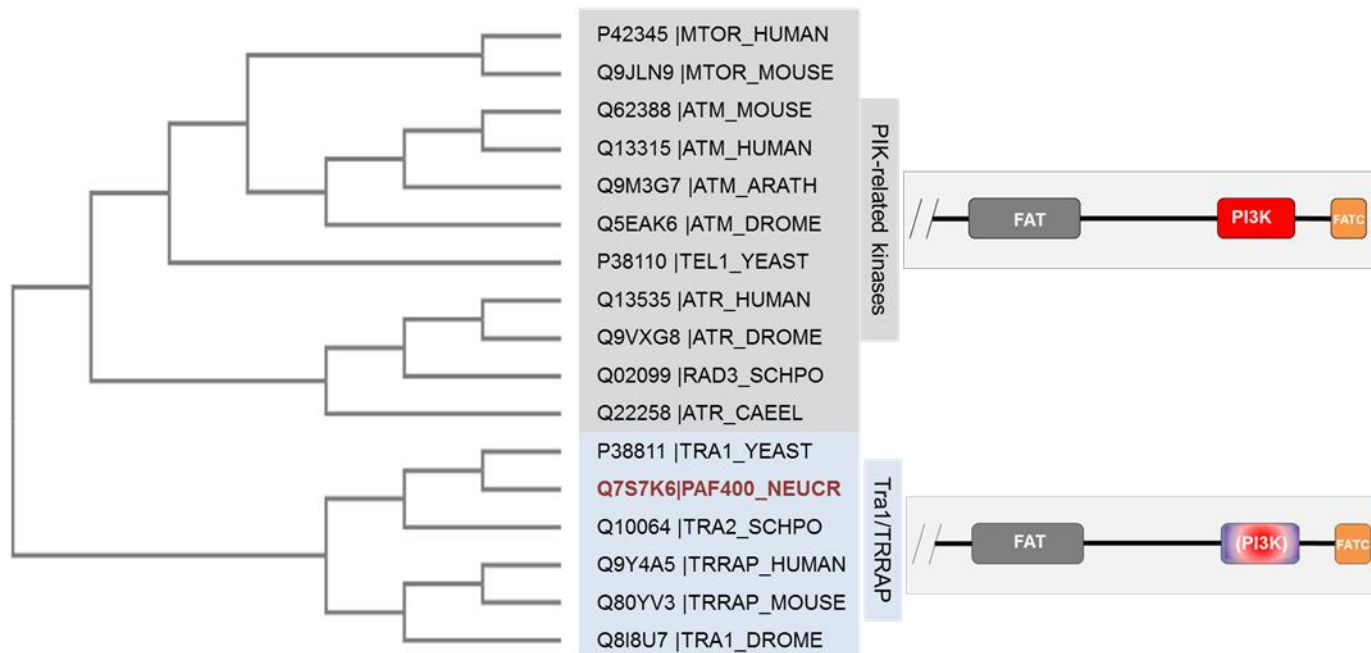
Figure 2.9 Phylogenetic tree and domain architecture of the PIKK protein family. Several representative proteins of the phosphatidylinositol 3 kinase-related kinase (PIKK) family, including TRRAP proteins are aligned. All members share the same domain architecture (right); however, the kinase domain is catalytically inactive in the Tra1/TRRAP protein subgroup. The lack of activity is denoted by a hollow in the domain (PI3K). *N. crassa* PAF400 protein clustered with Tra1/TRRAP proteins. The phylogenetic tree was generated by ClustalW2 Multiple Sequence Alignment and Phylogeny (LARKIN *et al.* 2007) using UPGMA clustering method with pairwise gap removal and PAM distances. UniProt identifiers are described in front of the tree with the name of the protein and the organism. The name of some organisms is abbreviated: *Arabidopsis thaliana,* ARATH; *Drosophyla melanogaster,* DROME; *Schizosaccharomyces pombe*, SCHPO; *Caenorhabditis elegans*, CAEEL; *Neurospora crassa*, NEUCR.

75

In addition, the conserved domain analysis of PAF400 showed that all three domains were present and localized at the C-terminus. This analysis also predicted that the PI3-kinase domain lacked the catalytic residues (Figure 2.10). Therefore, it was conclude that PAF400 belongs to the PIKK family, more specific to the Tra1/TRRAP protein sub-group.
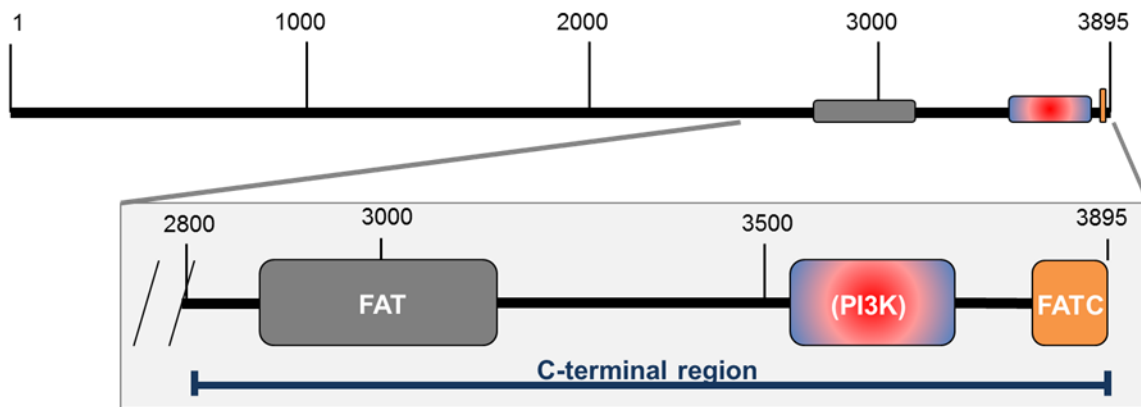
Figure 2.10 Protein structure of the *N. crassa* histone acetylase complex subunit PAF400. PAF400 is related to Tra1/TRRAP proteins, members of the PIK-related protein kinases (PIKK) family. PIKK family members consist of three distinct domains: FAT (FRAP, ATM and TRRAP) domain (gray); PI3K catalytic domain, which is inactive in Tra1/TRRAP (red); and FATC (FAT-C-terminus) domain (orange). Top, representation of the entire 3895 amino acid residues of PAF400. Bottom, zoom to the C-terminal region, where the domains are localized.

All PIKK proteins are nuclear localized. Therefore, I wanted to determine whether PAF400 was a nuclear protein. As expected, PAF400 contains a predicted monopartite nuclear localization signal consensus at amino acids 2051-2061 (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi) (KOSUGI *et al.* 2009). In order to confirm nuclear localization, a fluorescent reporter gene was fused to *Paf400* gene (*sGFP*[+]::*Paf400*[+]) at the canonical position. Nuclear fluorescent signal was observed under the microscope, confirming that PAF400 is a nuclear protein (Figure 2.11).
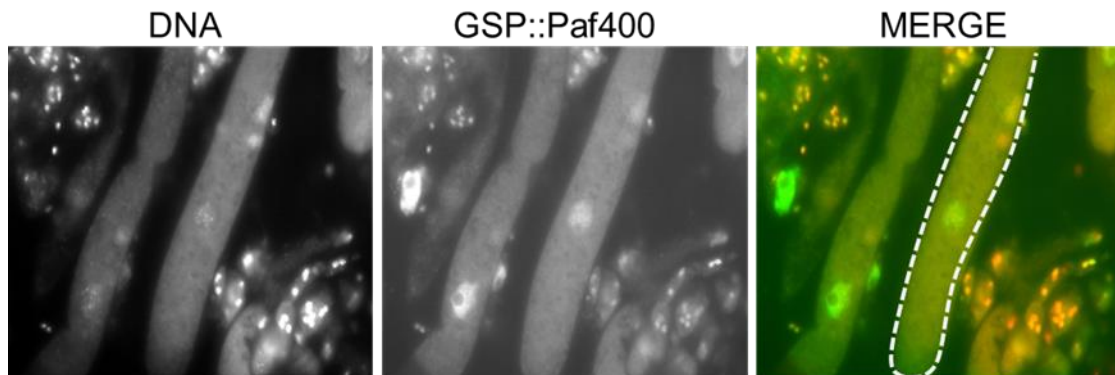
Figure 2.11 PAF400 is a nuclear protein. Localization of PAF400, relative to the nuclear DNA is shown. Displayed images are from Prophase I asci expressing GFP::Paf400. DNA was stained with dye Hoechst 33258. One of the ascus is outlined in white in the merged image (Pictures by Dong Lee).

In order to confirm interaction between SMS-5 and PAF400, I performed the *in vitro* co-IP assay, previously described. The large size of PAF400 made it difficult to express the entire protein *in vitro*. It has been demonstrated that the domains localized in the C-terminal region of Tra1/TRRAP proteins are necessary for interaction with other proteins (BOSOTTI *et al.* 2000; MUTIU *et al.* 2007). For that reason, I cloned and tested the last 1100 amino acids at the C-terminal region (Figure 2.10). This fragment included all three predicted domains.

PAF400 (C-ter) and SMS-5 were expressed in rabbit reticulocyte lysate as T7-tagged protein and ^{35}S-methionine radiolabeled protein, respectively (Figure 2.12A), or vice versa (Figure 2.12B). Radiolabeled proteins were incubated with the T7 antibody as the negative control. Ku70-Ku80 protein interaction was used as the positive control (Figure 2.12C) (RIHA *et al.* 2002). As shown in Figure 2.12A, radiolabeled SMS-5 was detected in the PAF400 (C-ter) IP (lane 4) with a signal just above background with the control reaction (lane2). Along that line, PAF400 (C-ter) bound and was precipitated with T7-tagged SMS-5 (Figure 2.12B, lane 2).
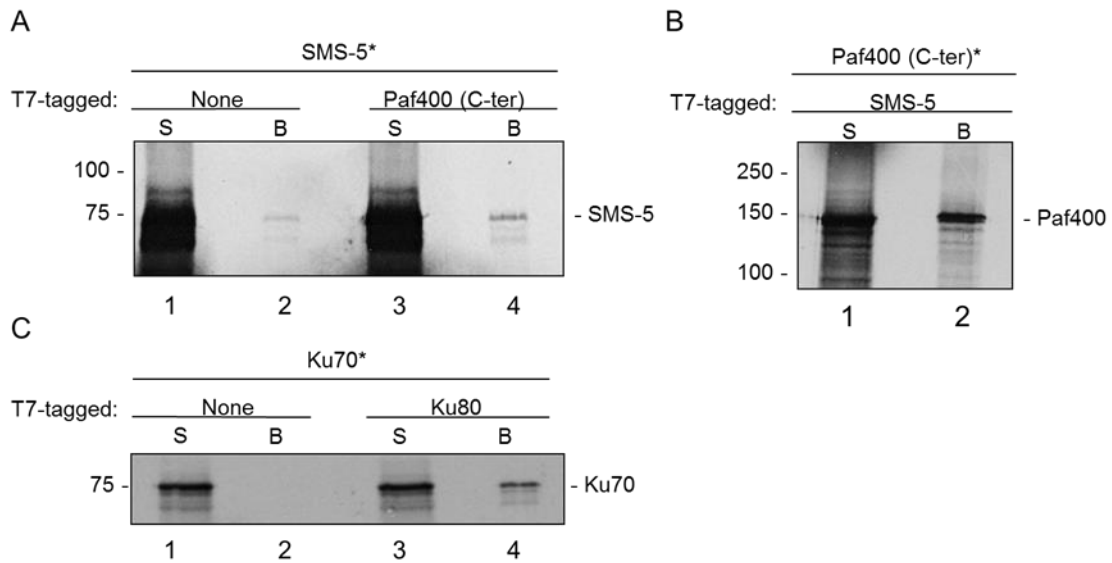
Figure 2.12 SMS-5 interacts with the PAF400 C-terminal region. Co-immunoprecipitation was conducted with ^{35}S-methionine labeled (asterisk) and T7-tagged protein expressed in RRL. Labeled protein is co-immunoprecipitated on T7-beads when bound to a tagged partner. S, supernatant; B, beads. (A) PAF400 C-terminus (C-ter) was T7-tagged and served as the bait protein; SMS-5 was radiolabbeled and acted as the prey protein. (B) SMS-5 was T7-tagged and served as the bait protein; PAF400 (C-ter) was radiolabbeled and acted as the prey protein. (C) Ku70-Ku80 interaction served as positive control.

The reciprocal co-immunoprecipitation of SMS-5 and PAF400 confirmed the interaction originally detected by the pull-down assays. Therefore, I conclude that the homolog of the histone acetylase complex subunit PAF400 is a protein binding partner for SMS-5. The approaches used to study further the role of PAF400 in meiotic silencing depend on whether *Paf400* is an essential gene. *Paf400* deletion mutant strains were never recovered in the lab and a mutant strain is not available in the *N. crassa* mutant collection. Together, these

observations suggest that the deletion of PAF400 is lethal and that PAF400 is an essential protein in *Neurospora*. As it happens, Tra1/TRRAP are essential proteins for cellular viability in yeast and mammalian cells (KNUTSON and HAHN 2011). For that reason, it was not possible to evaluate silencing of unpaired DNA during sexual growth in the absence of PAF400.

Cytosolic Regulator Pianissimo is a Protein Binding Partner for SMS-5

In the exploratory phase conducted to establish possible interacting partners for SMS-5, the cytosolic regulator Pianissimo (NCU07854) was also a strong candidate. Pianissimo was consistently pulled-down with SMS-5 in all affinity purifications. Until now, the function of Pianissimo has not been determined in *Neurospora*. Therefore, I used a set of bioinformatics analyses to determine the level of conservation of this protein sequence. By establishing homologous proteins, I expected to gain new insights into the possible function(s) Pianissimo could have, especially during meiotic silencing.

The translated Pianissimo sequence consists of 1547 amino acid residues with a molecular weight of 172.4 kDa. I conducted a protein BLAST analysis using UniProt/SwisProt protein database. The result revealed that Pianissimo is highly conserved in evolution, showing similarity of 47-59% with homologous proteins (Table 2.5). In addition, BLAST analysis against the non-redundant protein sequences (nr) database showed high conservation of Pianissimo protein among numerous Ascomycetes and Basidiomycetes;

however the majority of these proteins are reported as hypothetical with unknown function.

Pianissimo homologs belong to a highly conserved protein family, named the Pianissimo family. Members of this family are found in diverse organisms, including slime mold, yeast, mouse, and human. Pianissimo shares homology with the previously described PiaA protein, which is implicated in cAMP-induced cell migration by activation of adenylyl cyclase via Ras signaling in *Dictiostelium discoidieum* (CHEN et al. 1997); Ste20 protein requires for fertility in *S. pombe* (HILTI et al. 1999); Avo3/Tsc11 protein in *Saccharomyces cerivisiae* (LOEWITH *et al.* 2002); and RICTOR protein in mammals (SARBASSOV *et al.* 2004).

Avo3 and RICTOR proteins have a conserved interaction with the protein kinase TOR—member of the PIKK protein family—in the multi-protein complex "Target of Rapamycin Complex 2" (TORC2). It has been showed that TORC2 complex participates in signaling transduction by phosphorylating AGC (PKA, PKG, and PKC) protein kinases and AKT or protein kinase B (OH and JACINTO 2011). In mammals and budding yeast, TORC2 complex regulates the organization of the actin cytoskeleton (LOEWITH *et al.* 2002; JACINTO *et al.* 2004). In fission yeast, TORC2 complex is required to maintain genome integrity by protecting the cells from DNA damage during S phase (SCHONBRUN *et al.* 2013). Furthermore, TORC2 interacts with ribosomes and has been implicated in co-translational processing or maturation of nascent polypeptides in mammals cells (OH and JACINTO 2011).

Table 2.5 *N. crassa* Pianissimo protein and homologs

Organism Scientific Name	Protein name	Sequence length (aa)	Uniprot Accession #	Identity (%)	Similarity (%)
Neurospora crassa	Cytosolic regulator Pianissimo	1541	Q7SBV4	100	100
Schizosaccharomyces pombe	Target of rapamycin complex 2 subunit Ste20	1309	Q09743	36	59
Saccharomyces cerivisiae	Target of rapamycin complex 2 subunit Tsc11 (Avo3)	1430	P40061	27	47
Dictyostelium discoideum	Protein pianissimo A PiaA	1148	O77203	27	49
Homo sapiens	Rapamycin-insensitive companion of mTOR (hAVO3 / RICTOR)	1708	Q6R327	29	49
Mus musculus	Rapamycin-insensitive companion of mTOR (mAVO3 / Rictor)	1708	Q6QI06	29	50

To establish Pianissimo conserved domains, an analysis of the protein sequence was done and five domains were predicted. These domains are: HR1 at the N-terminus; RICTOR domains that represent conserved sections along the mammalian RICTOR proteins (RICTOR N, RICTOR M, and RICTOR V); and the Ras exchanger motif or REM domain (Figure 2.13).
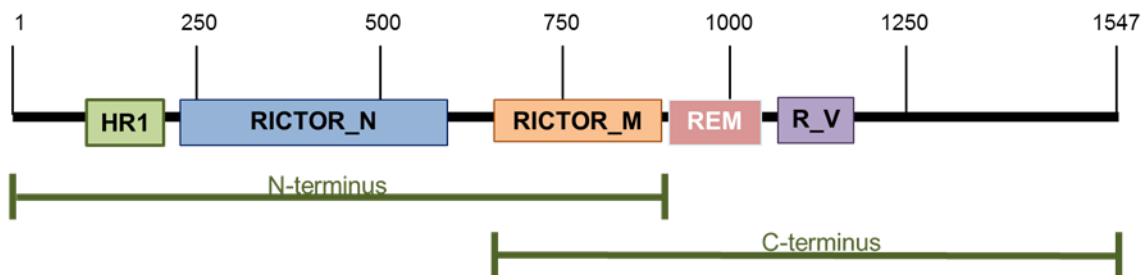
Figure 2.13 Protein structure of *N. crassa* Pianissimo. Pianissimo belongs to the Pianissimo protein family. Members of this family are subunits of the TORC2 complex. Five conserved domains were predicted. RICTOR N (light blue), RICTOR M (orange), REM (pink), and RICTOR V (purple) are present in all Pianissimo members. HR1 domain is only present in Pianissimo proteins of Dikarya organisms. Top, representation of the entire 1547 amino acid residues of Pianissimo. Bottom, the two fragments used for *in vitro* co-IP are shown and referred as N-terminus and C-terminus.

Domains architecture was shared among members of the Pianissimo family. The observed high protein conservation implied that Pianissimo may be required for essential cellular functions in *N. crassa*, as it has been demonstrated in other organisms (Ho *et al.* 2005; Jones *et al.* 2009).

Interestingly, the HR1 domain was only present in Pianissimo proteins of the fungal subkingdom Dikarya. In general, HR1 is a Rho-binding domain frequently found in Rho effector proteins such as protein kinases, lipid kinases, and scaffold proteins (BISHOP and HALL 2000). The GTP-bound form of Rho protein activates HR1-domain-containing proteins (FLYNN *et al.* 1998). The fungal specificity showed by HR1 domain suggested that this domain may be required for an additional Pianissimo function, exclusive to Ascomycetes and Basidiomycetes. Together these data suggested that Pianissimo is a highly conserved protein that belongs to the Pianissimo family and may be a subunit of TORC2 complex.

Pianissimo homologs are reported to be mainly cytolosolic proteins. However, the mammalian homolog RICTOR protein has been detected not only in the cytoplasm, but also in the nucleus (ROSNER and HENGSTSCHLAGER 2008). To predict cellular localization of *N. crassa* Pianissimo, I performed a bioinformatics analysis. A cytoplasmic localization was predicted by PredictProtein server (ROST *et al.* 2004). However, a nuclear localization signal (NLS) analysis using the cNLS Mapper software predicted a monopartite NLS in the middle of the protein, at residue 583, suggesting that Pianissimo is partially localized to the nucleus (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi).

To verify the cellular localization of Pianissimo, several attempts to fuse GFP to the C-terminal or N-terminal regions of the protein were made. However,

strains carrying the construct were not recovered. This result suggested that the fusions were not functional and that disruption of Pianissimo may have lethal effects in *N. crassa*. Along that line, Avo3 is an essential protein that is required for the regulation of the actin cytoskeleton through the TORC2 complex in *S. cerevisiae* (HO *et al.* 2005). In addition, disruption of Rictor in mice causes embryonic lethality (SARBASSOV *et al.* 2004). Although I did not obtain experimental data that demonstrates cellular localization of Pianissimo, the bioinformatics analysis and cellular localization of homologous proteins suggested that Pianissimo might be localized mainly in the cytoplasm, but partially in the nucleus.

In order to verify the interaction between SMS-5 and Pianissimo, I applied the same *in vitro* co-IP strategy described above. Due to the large size of the protein, I tested two overlapping fragments of Pianissimo. They were called the N-terminal half and the C-terminal half (Figure 2.13). Proteins were expressed in RRL; SMS-5 as ^{35}S-methionine radiolabeled protein. Both, N-terminal and C-terminal Pianissimo fragments were expressed as T7-tagged proteins. Proteins were precipitated using T7-tag antibody beads. Precipitates were collected and analyzed by SDS-PAGE and binding proteins were detected by autoradiography. The results showed that Pianissimo is an interacting partner for SMS-5 (Figure 2.14). Both, N-terminal and C-terminal Pianissimo fragments interacted with SMS-5. However, N-terminal region showed a strong interaction, as it was exhibited by the intensity of SMS-5 band in the autoradiography. This

result confirmed the interaction originally detected by the pull-down assays. Therefore, I conclude that the homolog of the cytosolic regulator Pianissimo/RICTOR protein is a protein binding partner for SMS-5.
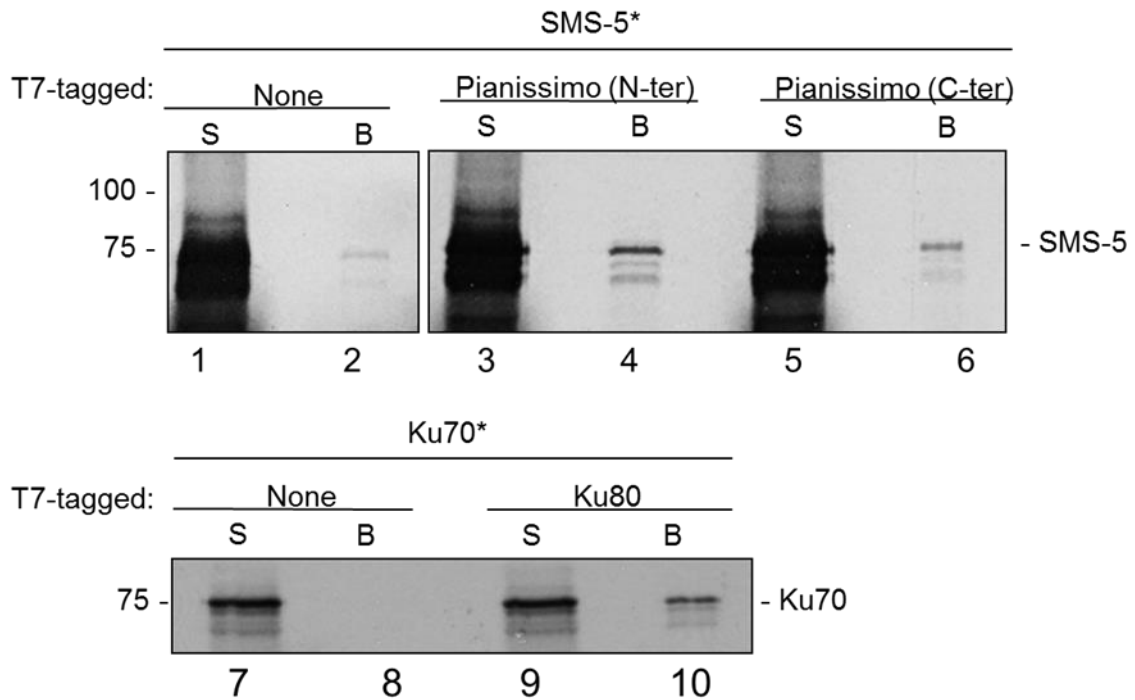
Figure 2.14 Pianissimo is a protein binding partner for SMS-5. Co-immunoprecipitation was conducted with ^{35}S-methionine labeled (asterisk) and T7-tagged protein expressed in RRL. Labeled protein is co-immunoprecipitated on T7-beads when bound to a tagged partner. S, supernatant; B, beads. Ku70-Ku80 interaction served as positive control.

Homologs of Pianissimo are essential proteins in several organisms. Avo3 is an essential protein for *Saccharomyces cervisiae* (WULLSCHLEGER *et al.* 2005). Ste20 is essential for sexual differentiation and meiosis in *S. pombe* (HILTI *et al.* 1999). Rictor is essential for the development of both embryonic and extra-embryonic tissues in mouse (SHIOTA *et al.* 2006). In *N. crassa*, Pianissimo may be also an essential protein. A deletion mutant strain was never recovered in the laboratory and a mutant strain for Pianissimo is not available in the *N. crassa* knockout collection. For that reason, it was not possible to test its direct involvement in meiotic silencing.

DISCUSSION

Biochemical work on *N. crassa* sexual tissue is technically challenging. Disruption of the tissue for extraction of meiotic cells is not trivial. A large amount of tissue is required to obtain a sufficient concentration of total protein for biochemical assays, requiring an enormous amount of media, reagents and other materials, as well as a lot of laboratory space. As a consequence, there is a lack of studies using biochemical approaches in sexual development in *Neurospora*. I developed a strategy that allowed me to grow a substantial amount of sexual tissue using limited resources and laboratory space. As a result, protein yield was improved. To my knowledge, this is the first biochemical study of *N. crassa* protein interactions during sexual development. This

standardized procedure could be used for further proteomic and biochemical studies in filamentous fungi during sexual differentiation.

In this study, I sought to reveal the function of SMS-5 by uncovering its protein binding partners to provide insight into the role of SMS-5 in meiotic silencing. For that purpose, I found that the nuclear Tra1/TRRAP homologous protein PAF400 and Pianissimo/RICTOR homologous protein are binding partners for SMS-5 during sexual development in *N. crassa*. Although Pianissimo is predicted to be mainly a cytoplasmic protein, it might be localized partially in the nucleus. If this were the case, it is likely that a nuclear protein interaction occurs between SMS-5 and Pianissimo. Alternatively, SMS-5 may interact in the nucleus with PAF400 and in the cytosol with Pianissimo.

I was not been able to test directly the involvement of these two proteins in meiotic silencing because PAF400 and Pianissimo are both essential proteins for *N. crassa*. This lethality may also be the reason that these two proteins were not detected as suppressor of meiotic silencing in previous genetic screens. The fact that two essential proteins were uncovered in my study shows the relevance of combining genetic and biochemical approaches when investigating a molecular mechanism. However, further studies are required in order to establish their functions in meiotic silencing.

Although it was not possible to analyze directly meiotic silencing in a PAF400 or Pianissimo mutant genetic background, indirect evidence indicate that both proteins play a role in meiotic silencing. These two proteins interact

directly with SMS-5, which in turn is a suppressor of meiotic silencing. Furthermore, *Sms-5* mutant strains do not have any other evident phenotype; both vegetative and sexual growth proceed normally. This observation suggests that SMS-5 does not participate in other cellular processes and may be only involved in meiotic silencing. Therefore, the direct interaction of PAF400 and Pianissimo with SMS-5 implies a participation of these two essential proteins in MSUD.

What is the Role of PAF400 in Meiotic Silencing?

PAF400 homologous proteins Tra1/TRRAP are essential proteins, which act as an important scaffold for several histone acetylase (HAT) complexes, including SAGA and NuA4 (GRANT *et al.* 1998). These complexes are targeted to gene regulatory regions by transcription activators and are associated with chromatin remodeling and transcriptional regulation. Tra1/TRRAP is the component responsible for the recruitment of these complexes to chromatin during transcription (MURR *et al.* 2007).

PAF400 and its homologs are essential proteins. It is has been proposed that the lethality observed in Tra1 and TRRAP mutations is due to disruption of NuA4 complex function(s). Several features support this hypothesis. Tra1/TRRAP is the only component shared by SAGA and NuA4 complexes. Tra1/TRRAP is the only subunit essential for viability in the SAGA complex (KNUTSON and HAHN 2011). In addition, Tra1 is not essential in *S. pombe.*

However, fission yeast has two paralogous proteins, Tra1, which associates specifically with SAGA, and Tra2, which associates specifically with NuA4. In contrast to Tra1, Tra2 is required for viability (HELMLINGER *et al.* 2011).

The function of Tra1/TRRAP proteins is not limited to transcriptional control. Acetylation of histone H4 by NuA4 complex is required for DNA DSBs repair (BIRD *et al.* 2002). It was demonstrated that histone acetylation by Trrap-containing complex modulates loading of repair proteins to DNA DSBs in Mouse Embryonic Fibroblast cells (MEFs) (MURR *et al.* 2006). Furthermore, TRRAP is also associated with non-HAT complexes such as MRN (MRE11, RAD50, and NBS1) complex, which play an important role in the initial processing of DSBs prior to repair (Chapter I). TRRAP-MRN complex is involved in detection and repair of DNA DSBs (Robert 2006). Thus, TRRAP contributes to DNA DSBs repair through its participation in both HAT and non-HAT complexes. Together these observations suggested that Tra1/TRRAP proteins may function as a link between DSBs signaling, repair, and chromatin remodeling.

The participation of a Tra1/TRRAP homologous protein in meiotic silencing is consistent with the hypothesis that PAF400 together with SMS-5 are involved in chromatin remodeling around the unpaired DNA during meiosis. This chromatin modification would help to relax the chromatin conformation around the unpaired region and allow the accessibility of other proteins or complexes, such as the transcriptional machinery. This could be an essential step in the synthesis of aRNAs from the unpaired DNA. It is possible that an accumulation

of DSBs occurs in the unpaired DNA. In this view, DSBs would work as a mark for the unpaired region to activate the nuclear stage of meiotic silencing.

SPO11 protein is responsible for the formation of DSBs during meiosis, an essential step for initiation of meiotic recombination (KEENEY 2001). However, Pratt demonstrated that meiotic silencing proceeds normally in the presence of a mutant *spo11*[RIP] allele, suggesting that SPO11-dependent DSBs are dispensable for meiotic silencing. Therefore, it was concluded that recombination and silencing are not related in *Neurospora* (PRATT 2008). It was recently showed that there is a spo11-independent mechanism for initiation of recombination in *Neurospora* (BOWRING *et al.* 2013), suggesting that SPO11-independent DSBs may be formed during meiosis. Therefore, it is possible that SPO11-independent DSBs participate in unpaired DNA recognition and activation of meiotic silencing.

I hypothesize that PAF400 may function during meiotic silencing as a scaffolding protein that brings different proteins together to the unpaired DNA region. One of these proteins could be SMS-5. Together, this multi-protein complex could be involved in the biosynthesis of aRNAs by remodeling the chromatin structure surrounding the unpaired DNA to facilitate access to transcriptional machinery. If this were the case, PAF400 may function as a molecular link between chromatin modification and the activation of the nuclear stage of meiotic silencing.

What is the Role of Pianissimo in Meiotic Silencing?

Pianissimo is a RICTOR-like protein. In mammals and other organisms, RICTOR is a core component that interacts with the catalytic subunit TOR in the TORC2 complex. In budding yeast and mammals the TORC2 complex is an important regulator of cytoskeleton that function to mediate phosphorylation of Protein Kinase C (PKC) (CYBULSKI and HALL 2009). TORC2 complex regulates many kinases (MATSUO *et al.* 2003; LEE *et al.* 2005b; JONES *et al.* 2009) and is proposed to participate in signaling transduction pathways. However, the molecular mechanism by which TORC2 works and the mechanism in which it participates remains unclear.

RICTOR proteins have also been found to associate with other proteins independently of TOR kinase. For that reason, it has been hypothesized that RICTOR proteins participates in other biological processes in which TORC2 complex is not involved (Oh 2011). However, its function still remains elusive.

Finding the interaction between SMS-5 and Pianissimo suggested that signaling transduction pathways could be involved in the regulation of meiotic silencing. It would be of interest to explore whether Pianissimo participates in meiotic silencing as a component of the TORC2 complex or if Pianissimo and PAF400 are components of a different and new multi-protein complex together with SMS-5. The possibility of direct interaction between PAF400 and Pianissimo needs to be explored.

What is the Role of SMS-5 in Meiotic Silencing?

A histone methyltransferase analysis did not show histone methylation in the presence of purified SMS-5 protein, suggesting that SMS-5 does not have a HMTase activity. Interestingly, histone H4 protein was found in association with GST-SMS-5 affinity column. This result could reflect an indirect interaction between these two proteins. In fact, histone H4 could be a direct target for PAF400-containing complex, such as NuA4 complex, which acetylases histone H4 during chromatin remodeling. Although this is an attractive explanation, a direct interaction between SMS-5 and histone H4 must be tested. It has been demonstrated that histone H4 methylation plays important roles to ensure genome integrity, especially during DNA damage repair, in several organisms (JORGENSEN *et al.* 2013). Therefore, a new evaluation of SMS-5 HMTase activity should be made.

On the other hand, assuming that SMS-5 is indeed a HMTase, the negative result observed by Lee (Lee and Aramayo, unpublished data) could be explained by a couple of possibilities. For examples, SMS-5 may need the presence of an activator or cofactor to transfer the methyl group to its target protein. An activator requirement has been observed in other systems, such as zebrafish, in which heat shock protein 90 (HSP90) acts as a cofactor of the HMTase SmyD1 (TAN *et al.* 2006). Because the HMTase assay was performed with purified SMS-5 in the absence of any cellular extract, the lack of a cofactor in the reaction could explain previous results.

93

HSP90 was not recovered from the initial pull-down assays. However, another heat shock protein HSP70 was identified in all affinity purifications. HSP70 is also a chaperone protein that assists in protein folding. This protein was not taken into consideration for further analysis because HSP70 was recovered from both GST-SMS-5 and GST columns. Even though HSP70 did not show specificity for SMS-5, it would be interesting to establish whether HSP70 does interact with SMS-5 and test whether it is a cofactor for HMTase activity.

Direct interaction between SMS-5 and two other proteins, PAF400 and Pianissimo, were verified in this study. Although a biochemical function has not been demonstrated for SMS-5 and its SET domain, the possibility that SMS-5 works as a chemical modifier of non-histone proteins still needs to be tested. Based on the results obtained in this study, these non-histone proteins could be PAF400 and/or Pianissimo. It would be of great interest to investigate protein post-translational modifications in PAF400 and Pianissimo in the presence of SMS-5.

In order to know whether PAF400 and Pianissimo have methylation sites, which could be the targets for SMS-5, I performed a protein methylation site prediction analysis using PMeS program (SHI *et al.* 2012). The analysis showed that both PAF400 and Pianissimo have predicted methylation sites. PAF400 has 8 arginine residues that may be methylated; one of them has a high statistical probability (>0.9 support vector machine (SVM) probability). The program also

predicted 3 methylation sites for Pianissimo, one lysine residue and 2 arginine residues with >0.6 SVM probability (Table 2.6).

Table 2.6 Protein Methylation Sites Prediction using PMeS[a]

Protein	Position Site	Flanking residues	SVM[b] Probability
PAF400	364	HIINFNF-**R**-KIFLPKI	0.6013
PAF400	720	KKSAILL-**R**-RLFKLAFM	0.6440
PAF400	837	PHLSYLM-**R**-RPLVVALR	0.6415
PAF400	844	RPLYYAL-**R**-RAGTELVG	0.6486
PAF400	1330	PIFAKPL-**R**-ALAFGIQ	0.9416
PAF400	1815	EPAKGGQ-**R**-FLDRAVI	0.6168
PAF400	2679	AKAFPEC-**R**-RLPPHVLK	0.6308
PAF400	3743	PHKFNIA-**R**-GSGNIWG	0.6031
PIANISSIMO	566	LLAIFV-**K**-SGVVQGL	0.6902
PIANISSIMO	845	IQMLDRW-**R**-IFNMMYR	0.6475
PIANISSIMO	907	ALRKYAT-**R**-PRISTQG	0.6787

[a]PMeS: Prediction Methylation Site (http://bioinfo.ncu.edu.cn/inquiries_PMeS.aspx)
[b]SVM refers to support vector machine probability

The identification of these predicted methylation sites in both SMS-5-interacting proteins, prompts the hypothesis that SMS-5 is a protein methyltransferase. Protein methylation is an important post-translational modification involved in several biological process, including signaling, RNA processing and transport, transcription, and DNA repair (PAHLICH *et al.* 2006; PAIK *et al.* 2007). Protein methylation, especially at arginine residues could promote or inhibit protein interactions (LEE *et al.* 2005a). Now that putative targets for SMS-5 have been determined, it will be possible to test SMS-5

biochemical function and establish whether SMS-5 has a methyltransferase activity when interacting with these proteins.

The RNA-binding Protein NCU05488 was not Essential for Silencing of Rsp^{RIP93} Allele

The RNA-binding protein NCU05488 was recovered in the affinity purification assays. However, the direct interaction between SMS-5 and NCU05488 was not observed using *in vitro* co-IP assays. In addition, this RNA-binding protein was not required for silencing the unpaired reporter gene. The most obvious conclusion would be that NCU05488 was recovered in the affinity purification assays as a false positive. Therefore, NCU05488 is not a real binding partner for SMS-5 and is not required for meiotic silencing. However, there are other possibilities that need to be considered.

NCU05488 was recovered from the pull-down assays, suggesting an interaction with SMS-5. This could be an indirect interaction, in which SMS-5 and NCU05488 share a protein partner. NCU05488 and SMS-5 may be components of the same multi-protein complex and for that reason both proteins were pulled-down together in the affinity purification assays. Because the co-IP was designed to test direct protein interactions, indirect interactions would yield negative results. Alternatively, a tag on the protein could prevent interaction *in vitro*, giving a negative result.

Participation of an RNA-binding protein in meiotic silencing seems to be necessary; RNA molecules are involved not only in the initial nuclear phase, but also in the perinuclear stage, where actual silencing of transcripts occurs. Now that direct interacting partners for SMS-5 were established, I proposed to test whether NCU05488 interacts with PAF400 and/or Pianissimo. If this were the case, that could explain the affinity purification results.

The involvement of NCU05488 in meiotic silencing was tested using a homeology RIP allele of the reporter gene *Rsp*. It has been observed that the nature of the unpaired DNA (*i.e.,* homeology, heterology or deletion) and the reporter gene (*e.g., Rsp, Asm-1*) used to evaluate meiotic silencing have an effect on the meiotic silencing response (Chapter I) (LEE *et al.* 2010a; XIAO *et al.* 2010). Therefore, it would be worthwhile to test whether NCU05488 is involved in silencing different unpaired reporter genes under different unpairedness conditions.

Proposed Model for the Nuclear Phase of Meiotic Silencing

The results presented above lead me to propose a simple threshold hypothesis for the activation of the nuclear phase of meiotic silencing. This model includes possible roles for SMS-5 in conjunction with PAF400 and Pianissimo after the detection of unpaired DNA during meiosis, and how the nuclear and the perinuclear phase of meiotic silencing could be connected.

The model is displayed and described in Figure 2.15. (a) After karyogamy homologous chromosomes come into proximity, pair and proceed to evaluate the molecular identity of their pairing partners via *trans*-sensing mechanism. (b) Programmed DNA DSBs are formed by SPO11. SPO11-independent DSBs or other DNA damage signal may also be present in early meiosis. This SPO11-independent breaks might be the product of persistent unrepaired DSBs, forming during DNA replication (BOWRING *et al.* 2013). Synapsis initiates and DSBs are repaired by meiotic homologous recombination. However, breaks localized in the unpaired DNA region cannot be repaired due to lack of homology. Hence, DSBs accumulate and this accumulation becomes the signal that activates the initial nuclear phase of meiotic silencing. (c) Reaching a threshold of unrepaired DNA damage stimulates the recruitment of several molecules to the unpaired region. The threshold hypothesis fits well into the requirements for meiotic silencing activation. A minimum region of DNA must be unpaired (~700 nucleotides) and the length of the unpaired region is proportional to the efficiency of silencing (Chapter I).

PAF400-containing complexes, such as NuA4-like complex with histone acetyltransferase activity binds to the chromatin surrounding sites of DSBs and initiates chromatin remodeling, possibly by histone H4 acetylation (MURR *et al.* 2006). (d) Chromatin relaxation facilitates the accessibility for other proteins to the unpaired DNA region, such as the MRN complex.

Figure 2.15 Proposed model for the activation of *trans*-sensing. Possible functions for SMS-5, PAF400, and Pianissimo are described. Interactions between one pair of homologous chromosomes (red and blue) are schematically represented. Sister chromatids, which are products of pre-meiotic DNA replication, are shown as double red or blue lines. Unpaired DNA region (red bubble) could be formed as consequence of gene insertion in the red chromosome or gene deletion in the blue chromosome. (a) After karyogamy homologous chromosomes pair and sense. (b) SPO11-dependent and independent DNA DSBs are formed (stars). DSBs remain unrepaired and accumulate in the unpaired DNA region. Accumulation of DSBs may be the signal that activates meiotic silencing nuclear phase. (c) Reaching a threshold of unrepaired DSBs stimulates the recruitment of NuA4-like complex to the unpaired DNA region. NuA4-like complex, formed by PAF400 and a histone acetyltransferase (HAT), binds to the chromatin surrounding sites of DSBs and initiates chromatin remodeling. (d) Chromatin relaxation facilitates the accessibility for other proteins to the unpaired DNA region, such as the components of the MRN complex. MRN complex is responsible for DNA damage repair through non-homologous end-joining (NHEJ) process. (e) Proteins involve in biosynthesis of aRNA are also recruited to the chromatin, including an RNA polymerase (POL). (f) Next, methylation of PAF400 and Pianissimo (PIA) by SMS-5 induces structural changes in these proteins allowing them to interact. (g) PAF400-Pianissimo interaction activates a signal transduction cascade mediated by protein kinases (PK) and proteins phosphorylation (P). The signaling transduction triggers activation of the meiotic silencing machinery (MS) in the perinuclear region. aRNA also migrates to the perinuclear region and incorporates into the meiotic silencing apparatus. Dashed lines represent the nuclear membrane.
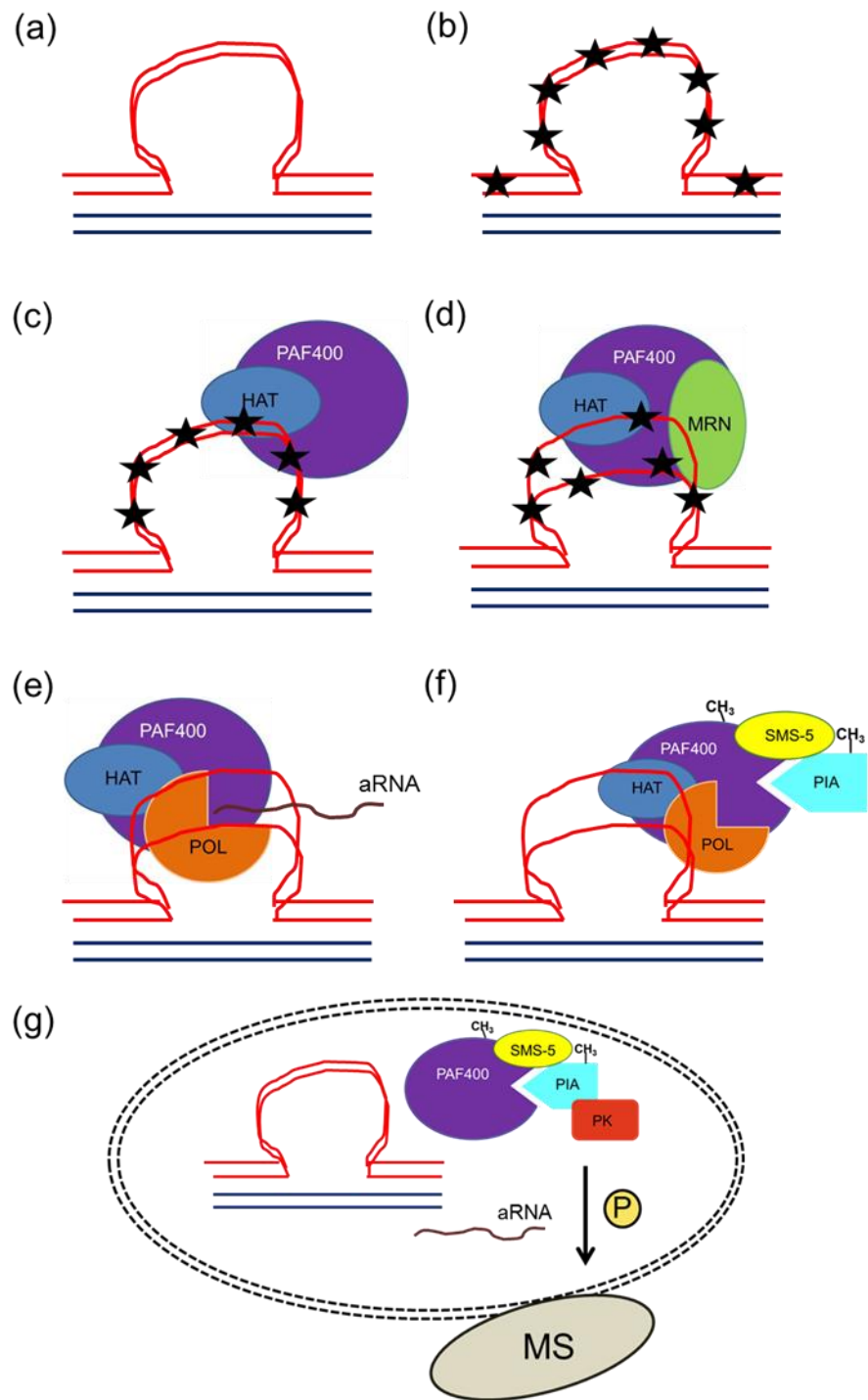
MRN complex is responsible for DNA damage repair through non-homologous end-joining (NHEJ) process (ROBERT *et al.* 2006). In addition, proteins involved in the generation of aRNA may also be recruited to the remodeled chromatin. PAF400 homologs are involved in transcriptional activation by recruiting the basal transcriptional machinery to DNA (SAGA complex) (MUTIU *et al.* 2007). (e) Using the same mechanism, PAF400-containing complex could bind the transcription machinery and mediate loading on the unpaired DNA region. This machinery would then be responsible for the transcription of an aRNA molecule from the unpaired DNA. A nuclear DNA/RNA-dependent RNA polymerase QDE-1 could be the catalytic component of this transcription (LEE *et al.* 2010b). QDE-1 is responsible for the generation of aRNAs and dsRNAs during the silencing of transgenes via quelling. Furthermore, QDE-1 has an effect on meiotic silencing when the mechanism is induced by homeologies. In quelling, QDE-1 interacts with QDE-3 a DNA helicase, also involved in DNA damage repaired. QDE-3 was precipitated once from the GST-SMS-5 column.

I also propose that (f) methylation of PAF400 and Pianissimo (PIA) by SMS-5 induces structural changes in these proteins allowing them to interact. (g) PAF400-Pianissimo interaction activates a signal transduction cascade mediated by protein kinases that triggers activation of the meiotic silencing machinery (MS) in the perinuclear region. aRNA also migrates to the perinuclear

region and incorporates into the meiotic silencing apparatus where the real silencing of the transcripts occurs (Figure 2.15).

Conclusions

This work has uncovered new and unexpected aspects regarding meiotic silencing. I determined two protein binding partners for the suppressor of meiotic silencing SMS-5. One of these partners belongs to the PIKK protein family and is associated with chromatin remodeling and DNA damage repair pathways. The other protein belongs to the Pianissimo protein family and is involved in signaling transduction mechanisms. Both proteins, PAF400 and Pianissimo are essential proteins, which explained why they were not detected as suppressor of meiotic silencing previously during the genetic screen. Finding essential proteins in this study points out the strength of combining genetic and biochemical approaches in order to study molecular mechanisms.

Now that protein binding partners for SMS-5 have been found, I hypothesize the biochemical function that SMS-5 may have as an essential component of meiotic silencing. SMS-5 may be a protein methyltransferase that regulates protein interaction and function of its targets during meiotic silencing. Although additional studies are required to test this hypothesis, the interaction between SMS-5-PAF400 and SMS-5-Pianissimo not only provide new insights into our understanding of the role SMS-5 could be executed in MSUD, but also

connects chromatin remodeling and signaling transduction pathways with meiotic silencing.

MATERIALS AND METHODS

Strains, Plasmids and Oligonucleotides

All *N. crassa* strains used in this study are described in Table A.3 (Appendix A). FGSC2490 (*mating type a*) strain and RANCR49A strain (*mating type A*) were used for setting up sexual crosses, induction of sexual growth, and extraction of endogenous proteins from sexual tissue. FGSC21956 strain was crossed with RANCR49A, KBNCR05A and RPNCR74A strains to construct VSNCR strains carrying the deletion mutant at the *NCU05488* allele.

Standard *Neurospora* culturing techniques (DAVIS 1970) were used throughout the study, except for the preparation of sexual tissue for protein extraction (described later on the document under "Preparation of *N. crassa* Sexual Tissue). Vegetative mycelium was cultivated in Vogel's Medium with 2% sucrose and sexual development was induced in Westergaard's Medium with 1.5% sucrose. The formulas for Vogel's Medium N and the Westergaard's Medium have been described by Davis and de Serres (1970) (DAVIS 1970)

Escherichia coli BL21 (DE3) cells were the host for all bacterial manipulation. All the plasmids used in this study are described in Table A.4 (Appendix A). All the oligonucleotides used in this study are described in Table A.5 (Appendix A)

The expression vector pGEX-6P-1 was used to construct and express GST-SMS-5 fusion protein and GST tag. The expression vector pET-28a(+) was used to construct and express T7-tagged proteins. The expression vector pCITE-4a(+) was used to clone and express ^{35}S-methionine radiolabeled proteins.

Purification of GST Fusion Proteins

GST and GST-SMS-5 recombinant proteins were expressed in *Escherichia coli* BL21 (DE3) cells. GST fusion proteins were purified based on the method described by Kellogg and Moazed (KELLOGG and MOAZED 2002). 8 g of *E. coli* cells expressing the recombinant protein were ground in liquid nitrogen. 5 volumes of room temperature phosphate buffer saline (PBS) containing 0.5% (v/v) Tween 20, 1 m*M* phenylmethylsulfonyl fluoride (PMSF), and 1 *M* NaCl was added to the cells. The cell powder was resuspended by stirring in a cold room for 10 min. To reduce viscosity the lysate was sonicated for 20 seconds and Dithiothreitol (DTT) was added to a final concentration of 10 m*M*. Lysate was centrifuged for 2 hours at 15,000 rpm in a Beckman JA-20 rotor. The supernatant was loaded onto a 5 ml pre-equilibrated glutathione agarose column (SIGMA) over a period of 2 hours. 50 ml of washing buffer 1, containing 1X PBS pH 7.4, 0.05% Tween 20, 0.5 m*M* DTT, and 0.25 *M* KCl, were passed through the column. Then, the column was washed with 25 ml of washing buffer 2, containing 1X PBS pH 7.4, 0.5 m*M* DTT, and 0.25 *M* KCl. GST fusion protein

was eluted in seven fractions of 0.8 ml of elution buffer, containing 50 mM Tris, pH 8.1, 0.25 M KCl and 5 mM reduced glutathione. The fractions were analyzed in 10% SDS-PAGE, and the peak fractions were combined and dialyzed into 50 mM HEPES, pH 7.6, 150 mM KCl, and 30% (v/v) glycerol. Protein concentration was calculated using Bradford assay (OLSON and MARKWELL 2007). Affinity columns were prepared by loading 12-15 mg of purified recombinant protein into 2.5 ml pre-equilibrated glutathione agarose column (SIGMA).

Preparation of *N. crassa* Sexual Tissue

Protein extraction from the sexual stage of *Neurospora* is challenging, due to the physical characteristics of the tissue. The small size of the sexual structure (perithecium) and the thickness of its wall make it very difficult to disrupt the tissue and extract all meiotic cells. As a result, a large amount of material is required to obtain a sufficient concentration of proteins. For that reason, I developed a strategy that allowed me to growth a substantial amount of sexual tissue using limited resources and laboratory space.

Instead of regular petri dishes, I used Pyrex glass pie plates (9-1/2-inch) to grow *N.* crassa under mating conditions. 200 ml of Westergaard's Medium was poured per plate and 10 plates were used in each extraction. Two pie plates were coupled together in a way that one plate served as the lid for the other plate. Therefore, a reduced spaced for incubation was achieved. Sexual growth is activated in the presence of light. Uniform distribution of light was assured to

both, the bottom and the top plates, due to the large surface area of the plate and the transparency of the glass. By using these conditions, I obtained approximately 7 g of sexual tissue per pie plate, for a total of ~ 70 g of tissue per extraction. This represents 6 times more tissue, comparing to the growth in traditional petri dishes using the same volume of media (Figure 2.16).

Westergaard's solid media was covered with stripes of pre-washed cellophane (7 X 1.5"). The cellophane enhanced the harvesting of the tissue by facilitating the separation of the tissue from the media. Three plugs, containing approximately the same amount of vegetative cells of RANCR49A strain, were used to inoculate each pie plate. Plates were incubated at 25°C with constant light for 6 days to induce female structure development. Following that, plates were fertilized with FGCS2490 strain by spreading male conidia suspension on top of the female tissue. Plates were incubated at 25°C for another 4 days to allow sexual tissue to form before protein extraction

The cellophane stripes were prepared as follow: cellophane was cut into stripes (7 X 1.5"), placed into 1 liter of deionized water and boiled in the microwave for 10 minutes. This step was repeated three times after replacing the water. The washed cellophane was transferred to a beaker with deionized water and autoclaved for 20 min.

Figure 2.16 Diagram for the preparation of *N. crassa* sexual tissue. 1) Glass pie plates (9" X 1.2") are washed with abundant hot water and sterilized. 2) 200 ml of sterile Westergaard's Media are poured per plate and 10 plates are prepared for each protein extraction. 3) A stripe of pre-washed/sterile cellophane (grey square) is laid on the surface of the media and the media is inoculated with three plugs (orange cylinders) of vegetative cells of the desired strain. Plugs are equally distributed in the plate. 4) Two inoculated pie plates are coupled together in a way that one plate serves as the lid for the other plate. 5) Each paired plate is placed into a closed plastic container to avoid dehydration and incubated for 6 days at 25°C to induce growing of maternal sexual tissue (orange). 6) After the six days, maternal tissue is fertilized with conidial suspension from the strain of the opposite mating type. Using a cotton swab, the conidial suspension is carefully spread over the area covered with the cellophane. 7) Paired plates are coupled and incubated for another 4 days at 25°C. 8) At the fourth day after fertilization, perithecia (sexual tissue; black area) are peeled off from the cellophane, frozen in liquid nitrogen, ground using a mortar and pestle, and transferred into a clean tube. 9) The tissue harvested from each pie plate is weighed. 10) Tissue from all ten plates is combined and placed into a pre-cooled Mill machine plastic container that already has a pre-cooled stainless-steel impactor. 11) Tissue is pulverized in the pre-cooled Mill machine. 12) Pulverized tissue is transferred to a clean tube and is frozen until ready for protein extraction.

| 1 | 2 | 3 | 4 |

Pour media   Add cellophane Inoculate   Couple two pie plates One at the top, one at the bottom

| 8 | 7 | 6 | 5 |

Weigh the tissue   Harvest sexual tissue   Incubate at 25°C 4 days   Fertilization using cotton swab   Incubate at 25°C 6 days

| 9 | 10 | 11 | 12 |

~7 g of sexual tissue/plate

Transfer tissue to a pre-cooled plastic container   Ground to powder Pre-cooled Mill machine   Collect pulverized tissue   Protein extraction

108

Protein Extraction from Sexual Tissue for Affinity Chromatography

Sexual tissue was peeled off from the cellophane, frozen in liquid nitrogen and then ground using a mortar and pestle. Ground tissue was placed into pre-cooled Mill machine-plastic containers with a pre-cooled stainless-steel impactor and pulverized in the Mill machine (SPEX 6850 Freezer/Mill). The cycles were as follow: 10 minutes pre-cooling, 2 minutes milling, 1 minute pre-cooling and 2 minutes milling. All the following steps were performed at 4°C. The frozen powder was weighed and transferred to a clean beaker. The powder was resuspended into 1.5 volumes of extraction buffer (50 mM HEPES pH 7.4, 137 mM NaCl, 10% Glycerol, 0.1% NP-40, 1X Protease Inhibitor Cocktail (PIC), and 1 mM PMSF) by stirring in a cold room for 10 minutes. The lysate was filtered using pre-treated and pre-cooled cheesecloth/Miracloth. The crude extract was centrifuged at 12,000 rpm for 15 minutes in a Beckman JA14 rotor. The supernatant was transferred to a clean pre-cooled conical tube and total protein concentration was calculated using Bradford assay (OLSON and MARKWELL 2007). Approximately 70 mg of total protein from whole cell extract was load into the GST-SMS-5 or GST affinity column.

Affinity Chromatography

The whole procedure was done in a cold room at 4°C. *N. crassa* cell extract was loaded in parallel on both, GST and GST-SMS-5 columns. Once the first drops of flow-through appeared, the flow from the column was closed to let

the proteins interact overnight. To optimize binding interactions, sample loading was run at a flow-rate of about 10 ml per hour. Flow-through aliquots were taken at the beginning and at the end of the experiment. The column was washed with 25 ml of washing buffer 1 (1X PBS pH 7.4, 0.05% Tween 20, 0.5 mM DTT, and 0.25 M KCl). Then, 10 ml of washing buffer 2 (1X PBS pH 7.4, 0.5 mM DTT, and 0.25 M KCl) were passed through the column. A last wash was done by passing 10 ml of 1X PBS pH 7.4. Samples from each wash step were taken and total protein concentration calculated using Bradford assay. If total protein was detected in the last wash faction, another 10 ml of 1X PBS pH 7.4 were passed through the column, until all the unbound proteins were washed out. For the elution of binding proteins, PreScission Protease enzyme (GE Healthcare) was used to cleave the proteolytic site localized in between the GST-tag and the fusion protein. Therefore, binding proteins were eluted without the GST-tag. The column was equilibrated with 20 ml of PreScission Protease Buffer (PPB), containing 50mM *Tris*-HCl pH 8, 100 mM NaCl, 1 mM Ethylenediaminetetraacetic acid (EDTA) and 1 mM DTT. 125 U of protease were resuspended in 2 ml of PPB and added to the column with closed flow. The column was kept close overnight to allow the catalytic reaction to occur. Proteins were eluted from the column by adding another 2 ml of PPB. The elution fraction was passed into a second 1 ml glutathione agarose column (GSTtrap column) and eluted by adding an additional 1 ml of PPB. A total of 4 ml represented the elution protein fraction. The purpose of the GSTtrap was to get rid of any extra

GST-tag. Hence, it was reduced the possibility for GST to interfere with the mass spectrometry analysis. Elution fraction was divided into two samples; one sample was precipitated and resolved on SDS-PAGE. Proteins were excised from the gel and identified by mass spectrometry analysis (In gel samples). The second elution sample was precipitated and directly submitted for mass spectrometry analysis (In solution sample).

Gel Electrophoresis and Identification of Bound Proteins

Proteins from flow-through, wash, and elution fractions were precipitated with Trichloroacetic acid (TCA). For that, sodium deoxycholate (DOC) was added to the sample to a final concentration of 0.2%, mixed and incubated on ice for 30 minutes. 15% of Trichloroacetic acid (TCA) was added to the mix and incubated for 1 hour on ice. The sample was centrifuged 10 min at 13,000 g. After that, the supernatant was aspirated and the pellet was washed with 1 ml of pre-cold acetone. After 5 min of incubation at room temperature, the sample was spun down and the pellet air-dried. Pellet was resuspended into 50 µl of 1X SDS-loading buffer (Laemmli buffer). Protein suspension was left at room temperature for 1 hour and then boiled for 5 min. 10 µl of sample were resolved in 10% SDS-PAGE. Proteins were visualized by Coomassie blue staining. Samples from GST and GST-SMS-5 affinity columns were compared.

In vitro Binding Assay

For *in vitro* co-immunoprecipitation, *Sms-5*, *NCU05488*, *Paf400* C-terminal region, and *Pianissimo* N-terminal and C-terminal regions cDNAs were cloned into pET-28a and pCITE-4a vectors (Novagen). pET-28a (T7-tag fusion) and pCITE-4a (untagged) constructs were expressed using rabbit reticulocyte lysate (RRL) according to manufacturer's instructions (Promega). Expression occurred in the absence or presence of ^{35}S-methionine (PerkinElmer). Cyclohexamide (2 mg/ml) was added to stop translation reaction. Translation of T7-tagged proteins was verified in the presence of [^{35}S] methionine on a small aliquot from the same master mix. T7-tagged and untagged radiolabeled proteins were immunoprecipitated using agarose beads (Novagen) containing the T7 monoclonal antibody (Bryan *et al.* 2000).

20 µl of T7-agarose beads were washed 4 times in 1.5 ml of Wash Buffer-100 (20 m*M* Tris acetate, pH 7.5, 10% glycerol, 1 m*M* EDTA, 5 m*M* MgCl2, 0.1% Nonidet P-40, 1 m*M* DTT, 100 m*M* potassium glutamate), centrifuging at 1,500 ×*g* for 1 min between washes. The beads were incubated 3 times with 1 ml of Blocking Buffer (20 M*m* Tris acetate, pH 7.5, 10% glycerol, 1 m*M* EDTA, 5 m*M* MgCl2, 0.1% Nonidet P-40, 1 m*M* DTT, 100 m*M* potassium glutamate, 0.5 mg/ml lysozyme, 0.5 mg/ml bovine serum albumin, 0.05 mg/ml glycogen) for 1 hour at 4 °C with agitation.

20 µl of T7-tagged and 60 µl of untagged radiolabeled proteins were combined and incubated at 30°C for 20 min, followed by 20 min at 4°C. 240 µl of

Blocking Buffer were added to the protein mix. The solution was incubated for 1 hour at 4 °C and centrifuged at 16,000 × g for 10 min at 4 °C to remove any particulates. The supernatant was transferred to a tube containing 20 µl of blocked beads. Proteins and beads were incubated overnight at 4 °C under agitation. The beads were washed 7 times in 750 µl of Wash Buffer-400 (20 mM Tris acetate, pH 7.5, 10% glycerol, 1 mM EDTA, 5 mM MgCl$_2$, 0.1% Nonidet P-40, 1 mM DTT, 400 mM potassium glutamate), 2 times in 750 µl of TMG (10 mM Tris acetate, pH 8.0, 1 mM MgCl$_2$, 10% glycerol), and resuspended in 20 µl of TMG. 7.5 µl of 5 X SDS-loading buffer were added to the precipitated and supernatant samples, boiled for 3 min, and resolved in a 10% SDS-PAGE. The gel was fixed in 25% isopropyl alcohol, 10% acetic acid for 30 min, dried at 80 °C, and exposed to an autoradiography film overnight.

Meiotic Silencing Assay

For testing the involvement of NCU05488 in meiotic silencing, strains containing the deletion mutant at *NCU05488* allele were constructed under different genetic backgrounds. Deletion of the *NCU05488* gene was verified by PCR amplification using oligonucleotides that annealed upstream and downstream of the gene. *Rsp*^{RIP93} genotype was verified by DNA digestion with restriction enzymes and Southern blot analysis (PRATT 2008).

Control and experimental crosses were set up under conditions that induced sexual development. Strains from opposite mating type were point

inoculated in Westergaard's Medium and incubated in the presence of light at 25 °C for 16 days. Once ascospores were shooting into the plate's lid, spores were harvested in 0.5 ml of water and immediately quantified under the microscope using a hemocytometer. Pictures from different field of view were taking and used for spore quantification. An average of 1500 spores were counted per cross. Three independent crosses were analyzed.

CHAPTER III

COMPARATIVE PROTEOMIC ANALYSIS OF TWO STAGES OF SEXUAL DEVELOPMENT IN *Neurospora crassa*

INTRODUCTION

N. crassa is a powerful eukaryotic model system to study cell differentiation and phenomena related to meiosis and sexual development, largely because of its ability to sense environmental changes and adapt to them by switching between the vegetative and the sexual growth state (ARAMAYO and SELKER 2013). Under specific environmental conditions, sexual differentiation is activated and several sexual structures develop (Figure 3.1). While numerous studies have used these model system to study development (NELSON *et al.* 1997b; KIM and NELSON 2005; NOWROUSIAN *et al.* 2007; LICHIUS *et al.* 2012), our understanding is still limited regarding the regulation and molecular control of the morphological and physiological changes observed during the sexual cycle of *N. crassa*.
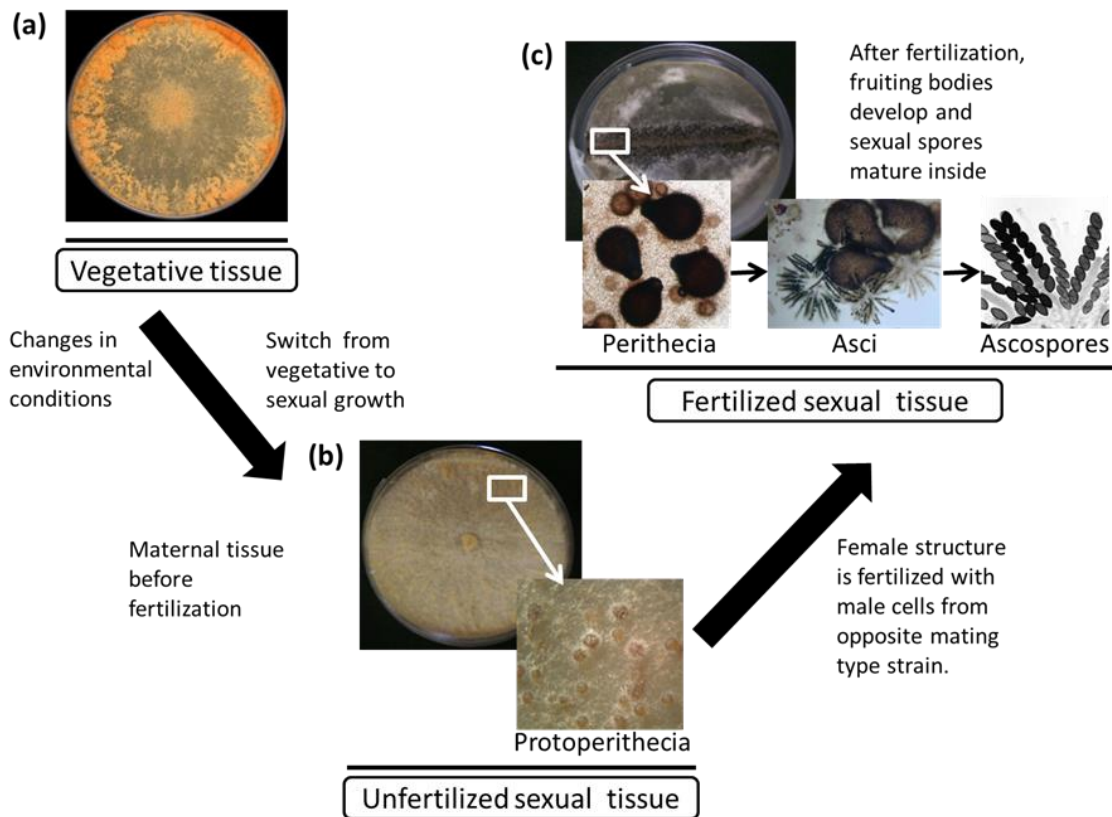
Figure 3.1 *N. crassa* development. *N. crassa* undergoes (a) vegetative or (b,c) sexual development depending on growing conditions. Under nitrogen starvation, exposure to light and low temperature, sexual differentiation is activated. Sexual growth is divided into two main stages: (b) pre-fertilization (unfertilized sexual tissue), in which female reproductive structures called protoperithecia are formed, and (c) post-fertilization (fertilized sexual tissue), in which perithecia, asci and spores developed.

While important efforts have been made to establish genes that are expressed during sexual growth, still little is known in terms of protein content and metabolic processes associated with sexual differentiation. Over 200 genes associated with various sexual development stages have been identified by

116

genetic mutations that alter normal sexual cycle progression in *N. crassa* (JOHNSON 1978; DELANGE and GRIFFITHS 1980; PERKINS *et al.* 1982; NELSON and METZENBERG 1992; RAJU 1992). Despite this research, molecular function directly related to sexual growth has been established for only a few of these gene products (JOHNSON 1979; GLASS *et al.* 1990; STABEN and YANOFSKY 1990; NELSON *et al.* 1997b; KIM *et al.* 2002; NOWROUSIAN *et al.* 2007). More importantly, the majority of the mutant genes with a defect in sexual development are likely not directly involved in sexual growth. Analysis of expressed sequences tags (ESTs) gave the first approximation of genes specifically expressed at different stages of the sexual cycle, such as in unfertilized tissue and/or during fruiting body maturation (NELSON *et al.* 1997a; DOLAN *et al.* 2000). Subsequently, transcriptome analyses were used to determine specific gene expression patterns under different growth conditions, including those that activate sexual growth (LI *et al.* 2005; PARK *et al.* 2008; LICHIUS *et al.* 2012; WANG *et al.* 2012). Those studies revealed the numerous genes and identified molecules that are potentially relevant for the development of sexual structures. Despite this, little is known about protein accumulation at the different stages and a global proteomic approach has not been implemented. While transcriptome analyses have the potential to identify differential protein expression, various studies have shown poor correlations between mRNA and protein abundance (GREENBAUM *et al.* 2003). Therefore, it is clear that analyses of mRNA and protein accumulation are

117

complementary: both are required for a complete understanding of how the cell works.

With plenty of genomic data for *Neurospora*, further functional research can provide more information at the protein and metabolic levels, which would enable better understanding of fundamental processes such as cell differentiation and sexual development.

Although the complete genome of *N. crassa* is sequenced (http://www.broadinstitute.org/annotation/genome/neurospora/MultiHome.html) (GALAGAN *et al.* 2003), and around 10,000 protein-coding genes have been predicted (BORKOVICH *et al.* 2004), functional classification of *Neurospora* proteins has remained challenging. The Munich Information Center for Proteins Sequences (MIPS) *Neurospora crassa* Genome Database (MNCDB) (http://mips.helmholtz-muenchen.de/genre/proj/ncrassa/) presents information on the molecular structure and functional network on the completely sequenced *N. crassa* genome; yet as to date, only 43% of predicted proteins have functional annotation according to MIPS Functional Catalogue (FunCat) (RUEPP *et al.* 2004). Furthermore, a complete Gene Ontology database for *Neurospora* is still missing. Until now, a general proteomic approach has not been applied for studying sexual development in *N. crassa*; mainly because of the technical difficulties associated with the scarcity of the sexual tissue and the challenges associated with tissue disruption for protein extraction. Therefore, new

approaches must be implemented in order to extend the functional annotation for the *N. crassa* proteome.

The goal of this study was to develop a comprehensive proteomics data set and a functional catalogue for *N. crassa* sexual development. To overcome the challenges associated with amount and disruption of the sexual tissue, I developed a culturing strategy for growing a considerable amount of tissue and applied a mechanical disruption that yielded a higher breakage of sexual structures. I then used a mass spectrometry-based proteomic approach followed by functional annotation and metabolic pathways analyses to visualize the potential molecular differences between unfertilized and fertilized sexual tissues.

In this study, a total of 841 distinct proteins were identified, hereafter referred as the proteome of *N. crassa* sexual tissue. Enrichment analysis revealed metabolic processes related to the morphological changes observed in sexual development, such as significant over-representation of biosynthesis of secondary metabolites and carbohydrate degradation during fruiting body formation. To the best of my knowledge, this is the first study that uses proteomics to investigate the potential biochemical differences between two stages of sexual development in filamentous fungi. This functional catalogue will serve as a reference tool for further studies related to sexual development not only in *N. crassa*, but also in other filamentous ascomycetes. The proteome, functional annotation and comparative proteomic analyses presented in this study, combined with the available gene expression data, can provide insights

119

into the biological processes activated at different developmental stages of *N. crassa*.

RESULTS

Proteome of *N. crassa* Sexual Development

The proteome of *N. crassa* sexual development was determined by identifying proteins extracted from two stages of *N. crassa* sexual growth: an early stage, before fertilization (unfertilized tissue) and a middle stage, four days after fertilization (fertilized tissues). Proteins were identified by two independent mass spectrometry analyses. The workflow for the identification and analyses of the proteome is summarized in Figure 3.2. From unfertilized sexual tissue, 620 proteins were characterized; 676 proteins were characterized from tissue harvested four days post-fertilization. The complete list of proteins identified in each stage and the protein confidence score are shown in Table B.1 (Appendix B). Combining the two data sets resulted in a final list of 841 distinct proteins that I referred to as the proteome of *N. crassa* sexual development. Unique proteins were divided into three sub-groups depending on tissue-specificity: (i) common proteins group (CP), which were detected in both sexual developmental stages and contained 455 proteins (54%), (ii) unfertilized-specific proteins (USP) with 165 proteins (19.6%), and (iii) fertilized-specific proteins (FSP) with 221 proteins (26.3%) (Figure 3.3).
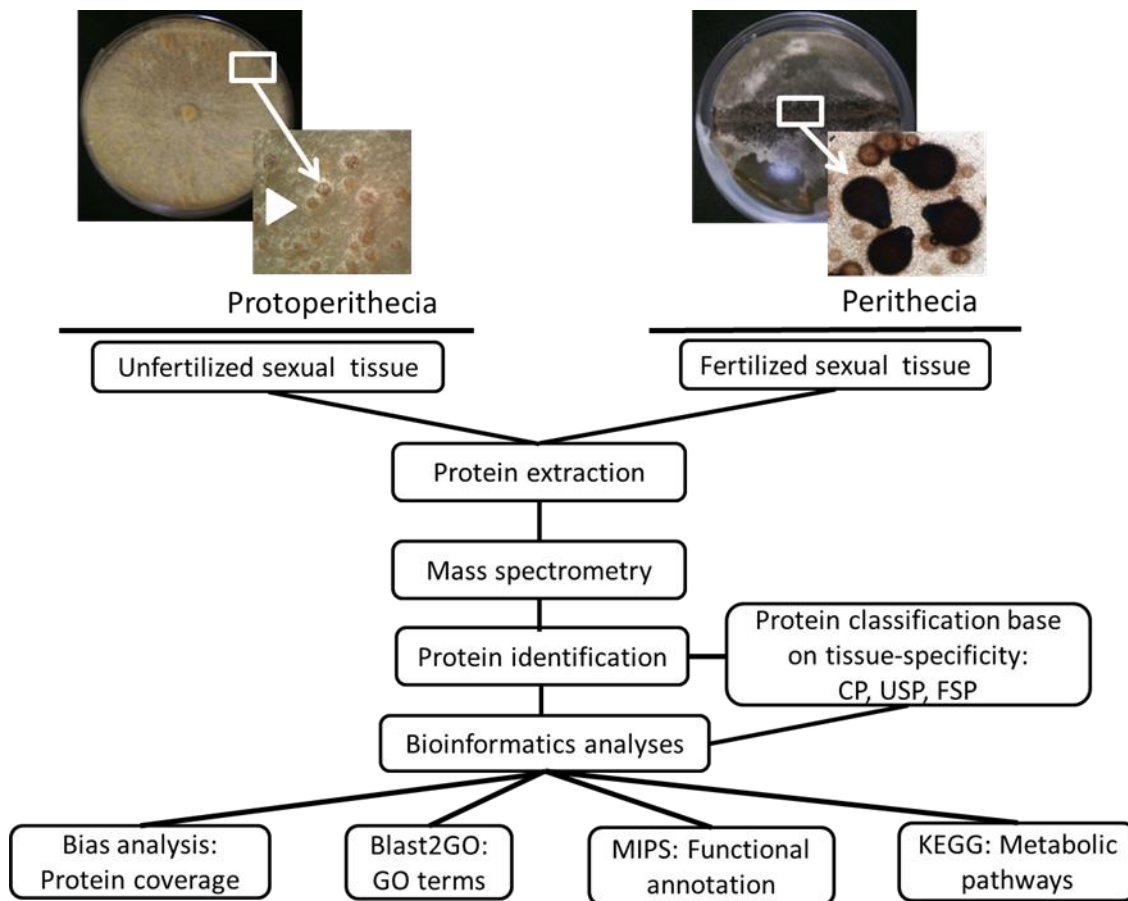
Figure 3.2 Workflow of the comparative proteomic analysis. The proteome of sexual development was determined by identifying proteins extracted from two different tissues: 6 days old unfertilized sexual tissue in which protoperithecia structures were formed (head arrow), and fertilized tissue yielded 4 days after fertilization, in which mature perithecia were developed. Mass spectrometry (MS) analyses were done to the extracted proteins and the MS data were searched against the *N. crassa* database for protein identification. Proteins were classified depend on tissue-specificity in: Common proteins (CP) identified in both sexual tissues; unfertilized-specific proteins (USP) identified only in unfertilized tissue, and fertilized-specific proteisn (FSP) detected exclusively in the fertilized tissue. Bioinformatics analyses were done to determined protein coverage, Gene Ontology (GO) terms by Blast2GO, Functional annotation via MIPS FunCat, and metabolic pathways and genetic processes using Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis.
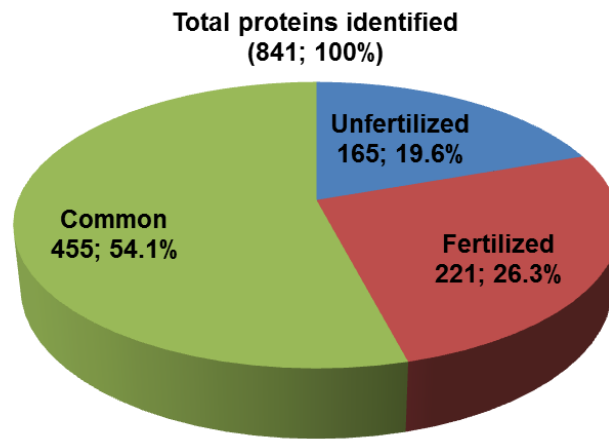
Figure 3.3 Distribution of the proteome of *N. crassa* sexual development. 841 distinct proteins were classified depending on tissue-specificity (unfertilized or fertilized) then divided into three protein sub-groups: common proteins, which were present in both tissues (CP) (green), unfertilized-specific proteins (USP) (blue), and fertilized-specific proteins (FSP) (red).

Bias Analysis

To assess the extent of proteome coverage achieved by the proteins identified in this study, I performed bias analysis of the protein length and Gene Ontology (GO) terms. These analyses revealed undetected proteins that were likely expressed at low levels, only in specialized cells, or only at particular developmental stages, such as during vegetative growth. The identified proteins from sexual development cover 8.4% of the proteins predicted from the total *N. crassa* genome. Because *Neurospora* spends most of its life cycle in the haploid stages of the vegetative cycle, it was not surprising to find low genome protein

coverage from proteins expressed during sexual growth.

By comparing the ratio of identified and expected proteins, I detected biases against short proteins (<100 amino acids) and long proteins (> 1000 amino acids) (Figure 3.4). This result is expected because short proteins yield less peptides and hence are less likely to be detected by LC-MS/MS. The bias against long proteins demonstrated that small proteins (100-500 amino acids) were more abundant than long proteins in the studied tissues. Similar results were observed with the fission yeast proteome (GUNARATNE *et al.* 2013) and my results are in accordance with the biosynthetic cost minimization hypothesis. This hypothesis states that smaller proteins tends to be more highly expressed than large proteins (WARRINGER and BLOMBERG 2006).

To analyze bias in terms of functional annotation, GO terms were assigned to the identified proteins using Blast2GO software coupled with the InterProScan program (CONESA *et al.* 2005; GOTZ *et al.* 2008). A total of 92% (774) of the identified proteins received at least one GO term. GO terms are divided in three categories, biological process, molecular function and cellular components. By analyzing the distribution of GO terms, I found that under the category of biological process, *metabolic processes* (87.2%) and *cellular processes* (61.6%) had the highest percentages of proteins (Figure 3.5A). Under *metabolic process*, a notorious over-representation of GO terms associated with protein synthesis, such as *protein metabolic process*, *gene expression* and
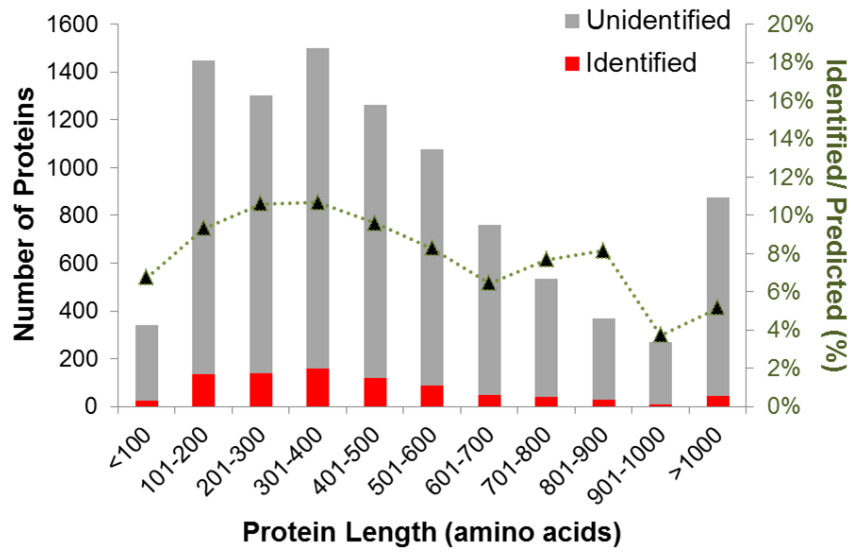
Figure 3.4 Bias analysis of protein length. The distribution of protein length for all *N. crassa* predicted proteins (10,066) was plotted as a histogram (grey bars, corresponding to the left vertical axis). The analysis was repeated for the proteins identified in this study (841) (red bars overlapping grey bars). The percentages of identified proteins in comparison to the total predicted for each histogram bar was calculated and is depicted as green triangles (right vertical axis), connected with lines for clarity.

translation, were observed (Figure 3.6). In addition, metabolic processes related to *growth*, *reproduction*, and *embryo development* showed high frequency in the proteome when compared with the genome. Interestingly, *carbohydrate metabolism* and *secondary metabolic process* were also observed as relevant biological processes during *N. crassa* sexual development (Figure 3.6). The molecular function GO term *binding* was also highly represented (77.3%). From this group, *protein binding*, *carbohydrate binding*, and *RNA binding* GO terms

showed high frequency compared with the total predicted proteins (Figure 3.5B; Figure 3.7). In addition, the analysis of cellular component categories showed that our set of data included proteins localized not only in the cytoplasm, but also in different intracellular compartments, including the nucleus (Figure 3.5C; Figure 3.8).

No large biases against specific biological process GO terms were found (Figure 3.6). However, biases were observed against molecular function GO categories related to transcription factor activity (GO:0001071; 0003700) (Figure 3.7). In addition, biases against proteins under the cellular component GO term *nuclear chromosome* (GO:0000228) were also detected (Figure 3.8) (Table B.2). Biases against transcription factors and chromosome-binding proteins could be explained by the low abundance that characterizes these proteins (JIANG *et al.* 2009). Taken together, the available data suggest that the majority of annotated GO categories were represented by the proteome of *N. crassa* sexual development. However, we cannot rule out that some proteins, such as poorly soluble, low-abundance, very short or very long proteins are missing in the dataset. This could be the result of potential technical limitations, such as protein/peptides that are difficult to extract and/or detect using our procedure.

Figure 3.5 Gene Ontololgy (GO) terms assigned to the proteome of *N. crassa* sexual development. Major GO terms associated with (A) biological process, (B) molecular function and (C) cellular component identified in the proteome of *N. crassa* sexual development. Percentages refer to the ratios of mapped proteins for each term to the total number of sexual developmental proteins identified with GO annotation.

Biological Process GO Terms

127

Figure 3.5 Continued.

Cellular Component GO Term
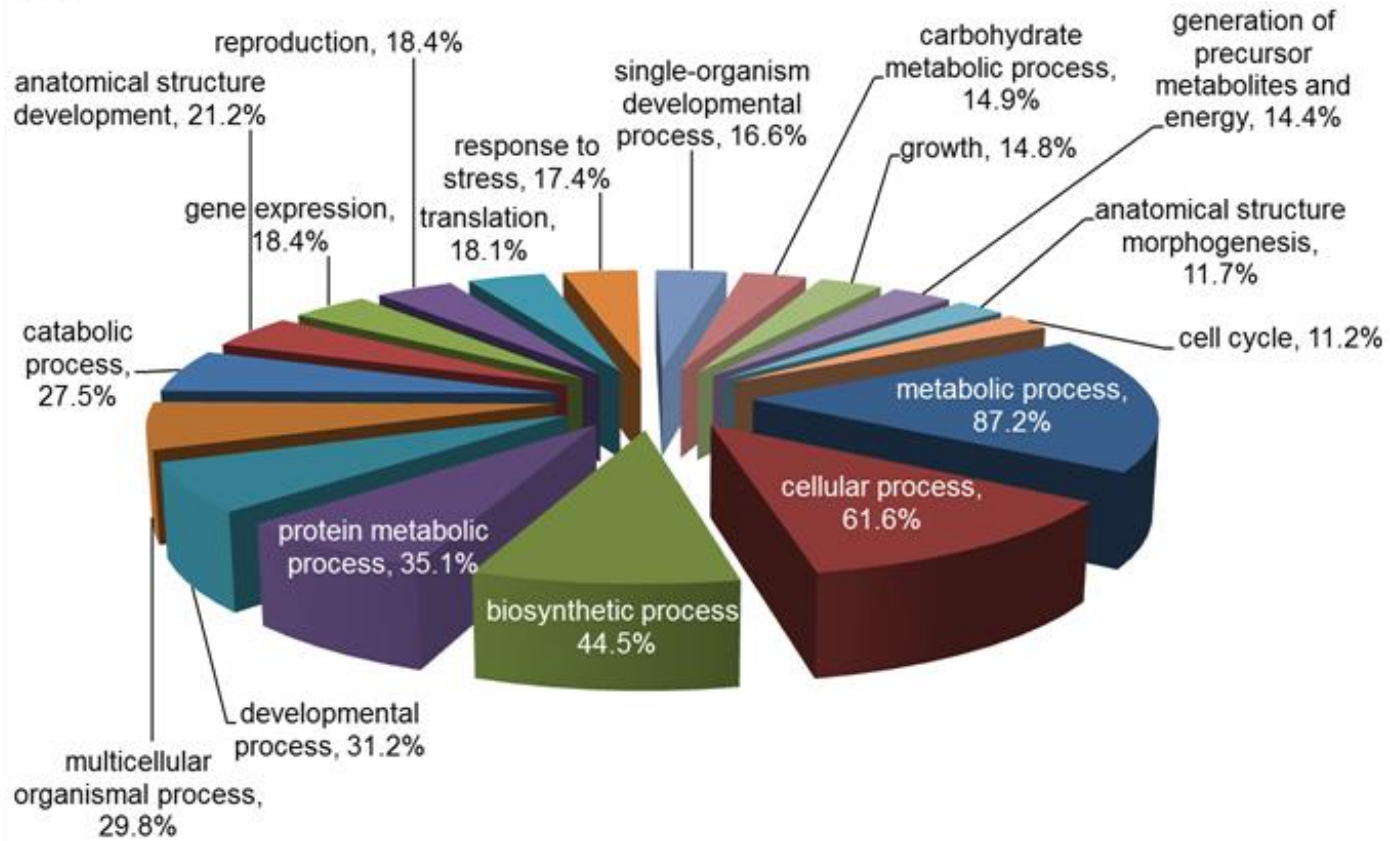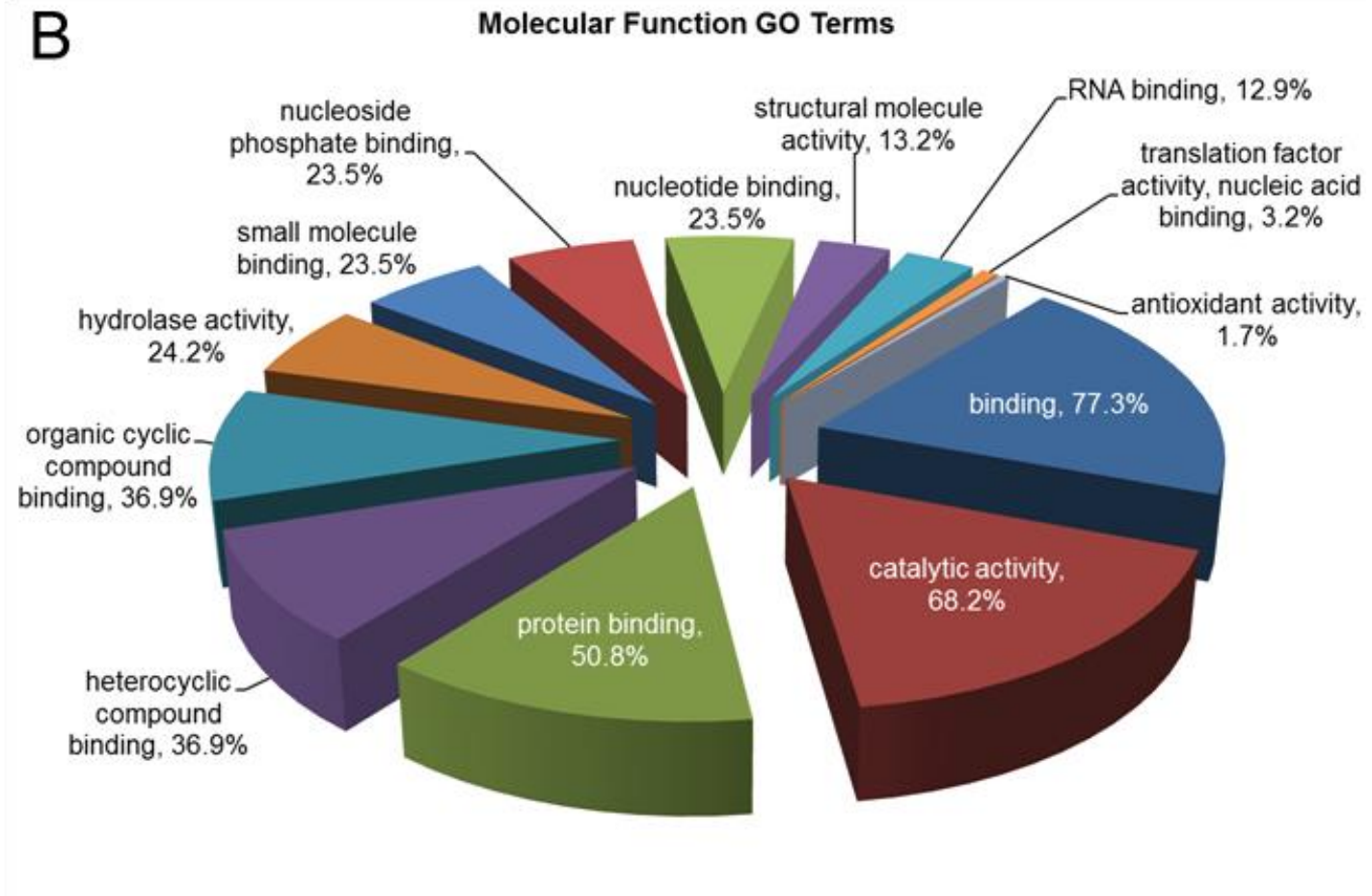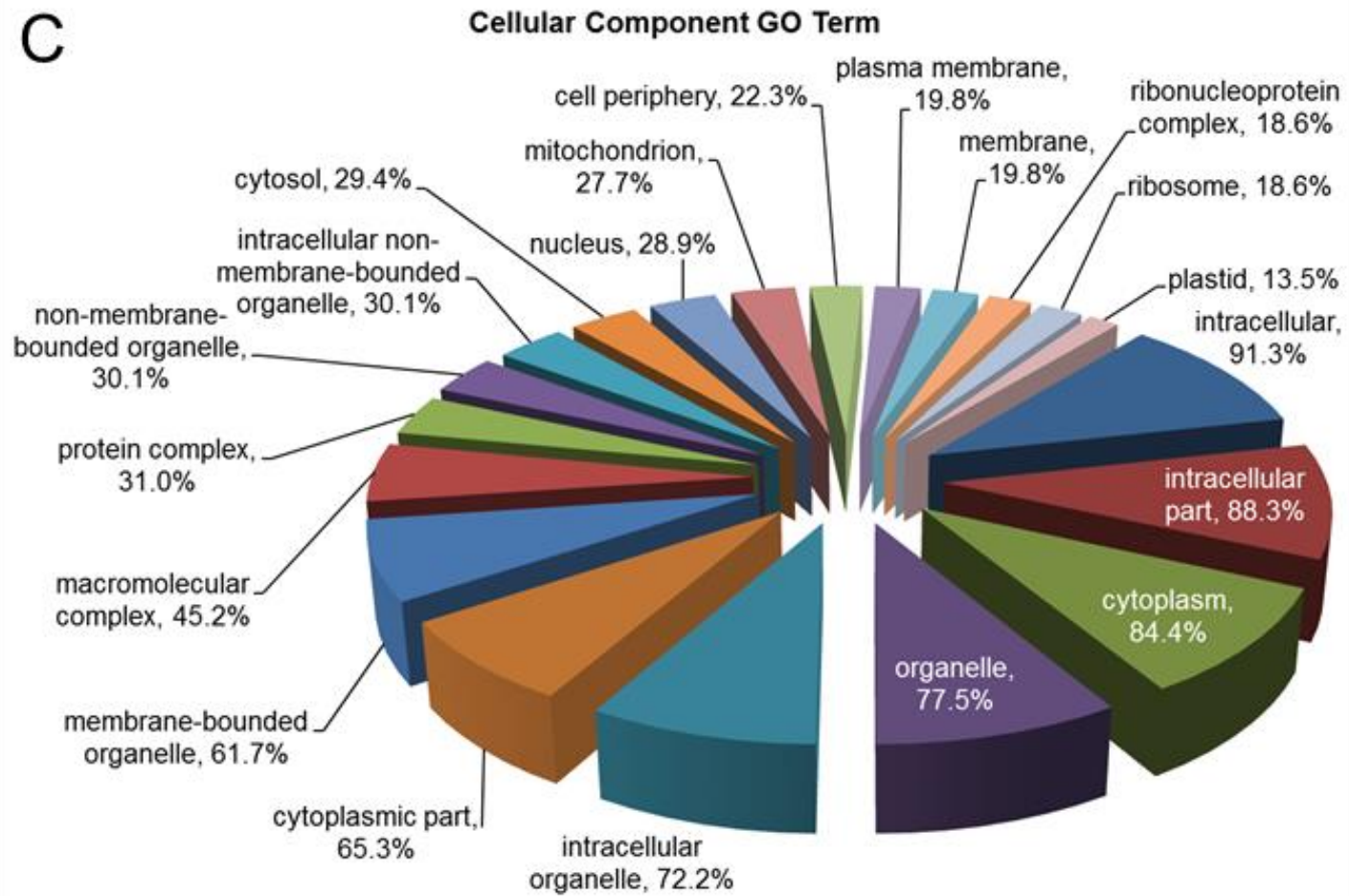
Figure 3.5 Continued.

Figure 3.6 Biological Process (BP) GO terms bias analysis. Distribution of annotated BP GO terms of all proteins identified in this study (proteome; red bars) in comparison to total predicted proteins (genome; grey bars) according to Table B.2 (Appendix B). GO terms are separated based on the $\log_{10}$ of the frequency. Frequency= ((Number of detected proteins associated with corresponding GO term/ Number of total detected proteins annotated with BP GO-terms)/(Number of genome predicted proteins associated with corresponding GO term/ Number of total genome predicted proteins annotated with BP GO-terms)).

signal transduction
regulation of cellular process
cellular response to stimulus
cellular protein modification..
protein modification process
macromolecule modification
DNA metabolic process
nucleic acid metabolic..
symbiosis, encompassing..
interspecies interaction..
multi-organism process
regulation of gene expression
regulation of..
regulation of metabolic..

Genom
Proteor

0%   5%   10%   1
Percentage of Protein

metabolic process
organic substance..
single-organism process
regulation of biological..
response to stimulus
establishment of localization
cellular component..
organic cyclic compound..
heterocycle metabolic..
cellular aromatic..
organelle organization
single-organism metabolic..
signaling
cell differentiation
protein transport
establishment of protein..
macromolecule localization
cell proliferation
multi-organism cellular..

Genome
Proteome

0% 20% 40% 60% 80%100%
Percentage of Proteins

biosynthetic process
protein metabolic process
single-multicellular..
multicellular organismal..
cellular macromolecule..
anatomical structure..
gene expression
macromolecule..
organic substance..
response to stress
single-organism..
growth
anatomical structure..
response to abiotic..
single-organism transport
cell death
secondary metabolic..
response to external..
cell-cell signaling
cellular homeostasis
cell growth
cytoplasm organization

Genome
Proteome

0% 10% 20% 30% 40% 50%
Percentage of Proteins

Biological Process

-0.2   0   0.1   0.6

log Frequency (%Proteome / %Genome)

Figure 3.7 Molecular Function (MF) GO terms bias analysis. Distribution of annotated MF GO terms of all identified proteins (proteome; red bars) in comparison to total predicted proteins (genome; grey bars) according to Table B.2 (Appendix B). GO terms are separated based on the $\log_{10}$ of the frequency. Frequency= ((Number of detected proteins associated with corresponding GO term/ Number of total detected proteins annotated with MF GO-terms)/(Number of genome predicted proteins associated with corresponding GO term/ Number of total genome predicted proteins annotated with MF GO-terms)).
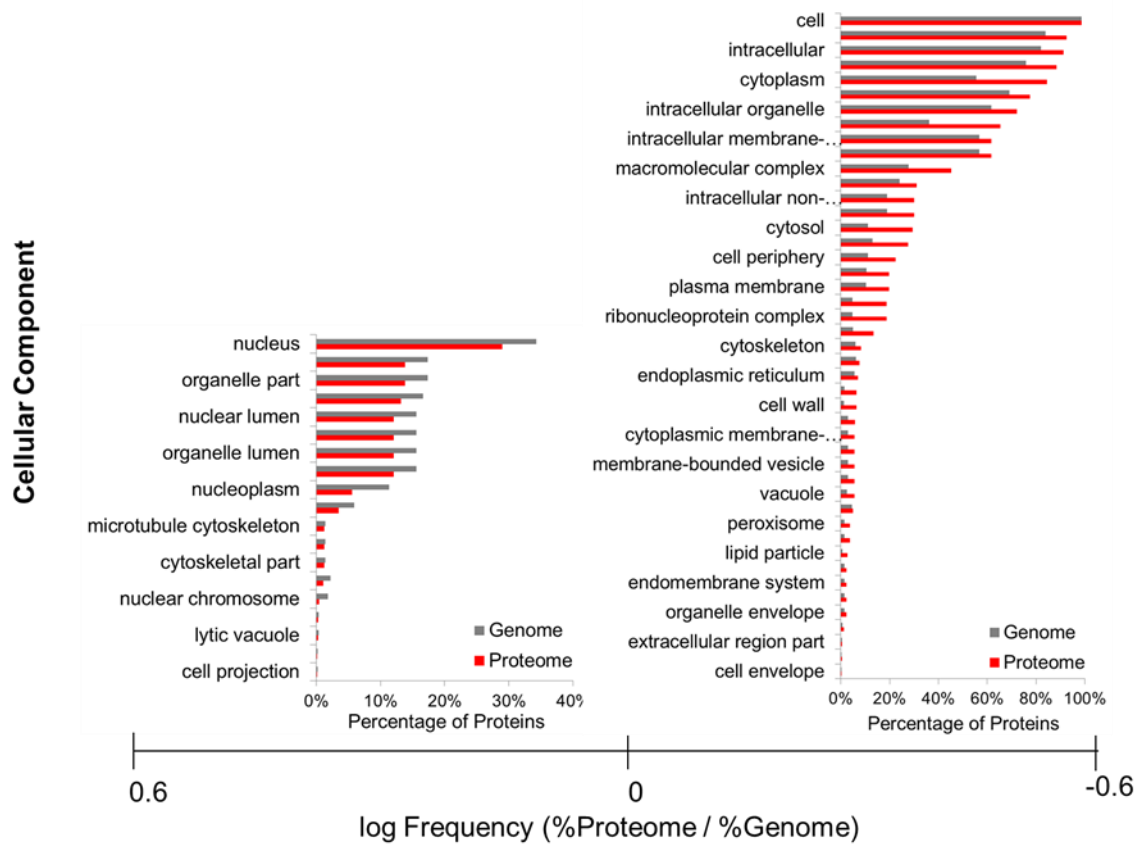
Figure 3.8 Cellular Component (CC) GO terms bias analysis. Distribution of annotated CC GO terms of all identified proteins (proteome; red bars) in comparison to total predicted proteins (genome; grey bars) according to Table B.2 (Appendix B). GO terms are separated base on the $\log_{10}$ of the frequency. Frequency= ((Number of detected proteins associated with corresponding GO term/ Number of total detected proteins annotated with CC GO-terms)/(Number of genome predicted proteins associated with corresponding GO term/ Number of total genome predicted proteins annotated with CC GO-terms)).

Functional Annotation Analysis

In order to determine the molecular processes that regulate sexual development in *N. crassa,* I performed a functional analysis using the *N. crassa* functional annotation established by MIPS Functional Catalogue (FunCat: http://mips.helmholtz-muenchen.de/funcatDB/). FunCat is a classification system enabling the description of proteins based on their biological function. Contrary to Gene Ontology, FunCat is organized into a smaller number of categories; biological and molecular processes are described in 28 main functional categories (RUEPP *et al.* 2004). This compact classification gives a more general picture of the cellular processes and therefore facilitates the interpretation of enrichment analysis. From approximately 10,000 predicted coding genes in *N. crassa*, 4213 proteins are classified by FunCat in at least one functional category. Out of the 841 unique proteins identified in the two stages of sexual development, 689 proteins (83%) were functionally annotated using FunCat (Table 3.1). Interestingly, when I analyzed each tissue-specific group of proteins separately, I observed a slight increase in the number of annotated proteins in the unfertilized and common groups (80.6% and 87.5%, respectively) compared with the fertilized group (75.6%) (Table 3.1). This result is consistent with previous observations in which mycelial and perithecial EST libraries had lower percentage of annotated genes and greater percentages of cDNAs encoding apparently novel genes (unannotated proteins) compared with conidial or unfertilized tissue libraries (DOLAN *et al.* 2000). The observation that fertilized-

specific proteins contain lower numbers of annotated proteins shows the potential that studies in fungal sexual structures have for the discovery of novel proteins.

Table 3.1 Summary of *N. crassa* proteins functionally classified via MIPS FunCat

FunCat Classification	Genome Proteins[a]	Sexual Development Proteome			
		Unfertilized Proteins[b]	Fertilized Proteins[c]	Common Proteins[d]	Total[e]
Classified	4213 (41.8%)	133 (80.6%)	167 (75.6%)	398 (87.5%)	698 (83.0%)
Unclassified	5853 (58.2%)	32 (19.4%)	54 (24.4%)	57 (12.5%)	143 (17.0%)
Total	10066	165	221	455	841

[a]Total genome proteins predicted in *N. crassa*
[b]Proteins identified in this study present only in unfertilized tissue
[c]Proteins identified in this study present only in fertilized tissue
[d]Proteins identified in this study present in both tissues
[e]Total proteins identified in this study

Interestingly, the percentage of unclassified proteins in the total experimental data (17%) was not as elevated as the percentage from the total genomic pool (58.2%). The same is true for other previously reported studies. For example, out of the 358 mitochondrial proteins reported in the *N. crassa* mitochondrial proteome, 58 (16.1%) were FunCat unclassified (KEEPING *et al.* 2011). This difference between the number of proteins annotated from experimental and predicted data could reflect some limitations with protein-coding gene predictions.

Distribution of the total number of identified proteins into main FunCat functional categories and the percentage of observed proteins compared to the percentage of expected proteins from the whole genome are summarized in Table 3.2. Due to the large differences detected between the percentages of observed and predicted unclassified proteins discussed above, I calculated and compared only percentages of functionally classified proteins.

The majority of proteins reported in this study were involved in *metabolism* and/or had a *binding function or cofactor requirement*. To establish if categories were over- or under-represented, I performed a *Z*-test analysis based on the difference between the proportions of observed and expected proteins. The results revealed statistically significant (*P*-value <0.01) over-representation of proteins involved in *metabolism, energy*, *protein synthesis*, *cell fate*, and *cell type differentiation*, among others. This analysis also revealed an under-representation or depletion of proteins involved in *transcription process*, especially in *RNA processing* (Table 3.2).

Table 3.2 Distribution of Main Functional Categories among Identified Proteins Compared with the Total Predicted Genome Proteins

Main Functional Categories According to FunCat	Total Predicted Genome Proteins		Total Identified Proteins		Z-test	P-value
	Expected Proteins	(%)	Observed Proteins	(%)		
01 Metabolism	1740	41.3	344	49.3	3.952	7.74E-05
02 Energy	375	8.9	162	23.2	11.219	3.28E-29
10 Cell cycle and DNA processing	622	14.8	101	14.5	-0.203	8.39E-01
11 Transcription	725	17.2	71	10.2	-4.672	2.98E-06
12 Protein synthesis	362	8.6	134	19.2	8.612	7.16E-18
14 Protein fate (folding, modification, destination)	862	20.5	182	26.1	3.358	7.86E-04
16 Protein with binding function or cofactor requirement (structural or catalytic)	1375	32.6	406	58.2	12.994	1.32E-38
18 Regulation of metabolism and protein function	216	5.1	52	7.4	2.502	1.23E-02
20 Cellular transport, transport facilities and transport routes	970	23.0	186	26.6	2.090	3.66E-02
30 Cellular communication/ signal transduction mechanism	286	6.8	50	7.2	0.363	7.16E-01
32 Cell rescue, defense and virulence	594	14.1	150	21.5	5.044	4.55E-07
34 Interaction with environment	473	11.2	97	13.9	2.040	4.14E-02
36 Systemic interaction with environment	7	0.2	3	0.4	1.431	1.52E-01
40 Cell fate	241	5.7	75	10.7	5.011	5.42E-07
41 Development (systemic)	45	1.1	14	2.0	2.106	3.52E-02
42 Biogenesis of cellular components	523	12.4	122	17.5	3.669	2.43E-04
43 Cell type differentiation	280	6.6	72	10.3	3.481	5.00E-04
Total proteins[a]	**4213**		**698**			

[a]Total functionally classified proteins
Over-represented category (p<0.01)
Under-represented category (p<0.01)

The enrichment analysis of functional categories determined by MIPS FunCat was consistent with the results obtained previously with Blast2GO analysis. Both analyses showed that most of the proteins from the proteome of sexual development in *N. crassa* participated in metabolic processes, protein synthesis, and/or had a function binding with other molecules. The consistency of both analyses suggested that Blast2GO could be used as an effective tool to improve the functional annotation of the complete genome of *N. crassa*.

To determine whether the enrichment/depletion pattern observed was specific for the sexual development proteome, I compared our protein dataset with gene expression data reported for a different developmental and cellular stage: conidial germination during vegetative growth. The comparison with genetic data was necessary because proteomic data is lacking. I used the conidial germination transcription profile data set reported for 1-4 hours post-germination (KASUGA *et al.* 2005) and performed a functional analysis via FunCat. I observed enrichment in *protein synthesis* category and depletion in *metabolism* category (data not shown), as was reported originally by Kasuga *et al.* Therefore, I conclude that two different sets of gene products from *N. crassa* generate different functional enrichment/depletion patterns.

This result suggested that the functional enrichment observed depended on growth conditions and development stages. If this hypothesis were true, I expected to observe enrichment in similar functional categories when comparing similar data sets. To test the hypothesis, I compared the proteomic data

generated in this study with previously reported genomic data, in which gene expression during fruiting body development was analyzed (WANG *et al.* 2012). Both proteomic and genomic studies were performed under similar growth conditions that induced sexual development. In the genomic study, the authors observed a statistically significant enrichment in genes whose products are related to *metabolism*, *proteins with binding function*, *protein fate*, *cellular transport*, *cell fate* and *cell type differentiation* (Table S4 from (WANG *et al.* 2012)). These results are consistent with my observations. Therefore, I conclude that the functional enrichment pattern observed in the present study reflects the protein population of sexual development.

Wang *et al*. also reported enrichment in genes involved in transcriptional processes (WANG *et al.* 2012). However, I observed depletions in this category, mainly in proteins involved in RNA processing. This finding could be explained because gene expression does not always correlate with protein synthesis, as both transcription and translation are highly regulated processes (HAIDER and PAL 2013). Genes can be transcribed but not necessarily translated in parallel.

Other biological aspects could explain the discrepancy observed in the transcriptional category between genomic and proteomic data. These aspects are: protein concentration and protein turnover. Protein concentration can affect protein population results because a threshold is required for protein detection via mass spectrometry analysis. Therefore, low-abundance proteins would not be detected in proteomic analyses, unless the method used for protein

extraction targets those low concentrated proteins. An example of low-abundance proteins is transcription factors (TFs). TFs are expressed in very low concentrations in most of the cells, thus making their identification difficult in proteomic studies (JIANG *et al.* 2009).

Tissue-Specific Proteins Pathway Analysis

Because the proteome of sexual development showed a specific enrichment pattern in molecular functions, I wanted to establish whether each stage of sexual development also presents a distinctive over-representation of functional categories. To do this, I functionally analyzed all three sub-groups of proteins individually (common, unfertilized-specific and fertilized-specific proteins). All main functional categories and sub-categories showing statistically significant enrichment of proteins (P-value<0.01) in at least one of the protein sub-groups are summarized in Table 3.3. This analysis revealed that each stage of development (unfertilized and fertilized) has enrichment on different functional categories. The difference observed in the enrichment pattern could explain the molecular and biological processes specific for each stage. For example, the formation of female structures in unfertilized tissue and the maturation of the fruiting body that takes place after fertilization.

Table 3.3 Protein Functional Analysis of Different Stages of *N. crassa* Sexual Development

ID Functional Categories	Genome		Unfertilized				Fertilized				Common			
	Exp[a]	%	Obs[b]	%	Z-score	P-value	Obs[c]	%	Z-score	P-value	Obs[d]	%	Z-score	P-value
01 Metabolism	1740	41.3	60	45.1	0.9	0.380	98	58.7	4.5	0.000	189	47.5	2.4	0.017
01.01 Amino acid metabolism	319	7.6	16	12.0	1.9	0.058	23	13.8	2.9	0.003	45	11.3	2.6	0.008
01.01.03.02.02 degradation of glutamate	16	0.4	1	0.8	0.7	0.498	4	2.4	3.8	0.000	5	1.3	2.5	0.013
01.01.05.01 metabolism of polyamines	15	0.4	0	0.0	-0.7	0.491	3	1.8	2.9	0.004	3	0.8	1.2	0.224
01.01.05.03 metabolism of urea (urea cycle)	10	0.2	3	2.3	4.2	0.000	2	1.2	2.3	0.020	2	0.5	1.0	0.321
01.02 nitrogen, sulfur and selenium metabolism	150	3.6	12	9.0	3.3	0.001	9	5.4	1.2	0.215	15	3.8	0.2	0.831
01.05 C-compound and carbohydrate metabolism	759	18.0	27	20.3	0.7	0.500	58	34.7	5.4	**0.000**	105	26.4	4.1	0.000
01.05.03 polysaccharide metabolism	169	4.0	1	0.8	-1.9	0.056	17	10.2	3.9	**0.000**	12	3.0	1.0	0.328
01.05.11 aromate metabolism	40	0.9	3	2.3	1.5	0.134	6	3.6	3.3	**0.001**	3	0.8	0.4	0.698
01.06.05 fatty acid metabolism	71	1.7	7	5.3	3.1	0.002	6	3.6	1.8	0.066	19	4.8	4.3	0.000
01.20 secondary metabolism	245	5.8	12	9.0	1.5	0.123	24	14.4	4.5	**0.000**	21	5.3	0.4	0.659
01.20.01 metabolism of primary metabolic sugar derivatives	17	0.4	0	0.0	-0.7	0.463	4	2.4	3.7	**0.000**	0	0.0	1.3	0.204
01.20.35 metabolism of secondary products derived from L-phenylalanine and L-tyrosine	33	0.8	1	0.8	0.0	0.968	7	4.2	4.5	**0.000**	1	0.3	1.2	0.236

Table 3.3 Continued.

ID Functional Categories	Genome Exp[a]	Genome %	Unfertilized Obs[b]	Unfertilized %	Unfertilized Z-score	Unfertilized P-value	Fertilized Obs[c]	Fertilized %	Fertilized Z-score	Fertilized P-value	Common Obs[d]	Common %	Common Z-score	Common P-value
02 Energy	375	8.9	26	19.5	4.2	0.000	26	15.6	2.9	0.003	111	27.9	11.8	0.000
02.11 electron transport and membrane-associated energy conservation	116	2.8	9	6.8	2.7	0.006	8	4.8	1.6	0.120	31	7.8	5.5	0.000
02.13 respiration	138	3.3	9	6.8	2.2	0.028	7	4.2	0.6	0.516	36	9.0	5.8	0.000
02.19 metabolism of energy reserves (e.g. glycogen, trehalose)	40	0.9	0	0.0	-1.1	0.259	5	3.0	2.6	0.010	14	3.5	4.6	0.000
02.25 oxidation of fatty acids	38	0.9	4	3.0	2.4	0.015	5	3.0	2.7	0.007	5	1.3	0.7	0.482
12 Protein synthesis	362	8.6	23	17.3	3.5	**0.001**	15	9.0	0.2	0.860	98	24.6	10.2	0.000
12.01 ribosome biogenesis	221	5.2	9	6.8	0.8	0.440	7	4.2	-0.6	0.548	70	17.6	9.7	0.000
12.01.01 ribosomal proteins	152	3.6	8	6.0	1.5	0.147	6	3.6	0.0	0.992	65	16.3	11.5	0.000
12.04 translation	188	4.5	21	15.8	6.0	**0.000**	12	7.2	1.7	0.098	84	21.1	13.5	0.000
12.04.01 translation initiation	50	1.2	10	7.5	6.2	**0.000**	2	1.2	0.0	0.990	7	1.8	1.0	0.324
12.10 aminoacyl-tRNA-synthetases	38	0.9	4	3.0	2.4	0.015	2	1.2	0.4	0.694	13	3.3	4.3	0.000
16 Protein with binding function or cofactor requirement (structural or catalytic)	1375	32.6	71	53.4	5.0	0.000	71	42.5	2.7	0.008	269	67.6	13.9	0.000
16.01 protein binding	678	16.1	37	27.8	3.6	0.000	39	23.4	2.5	0.013	168	42.2	12.9	0.000
16.02 peptide binding	9	0.2	2	1.5	2.9	0.004	1	0.6	1.0	0.306	3	0.8	2.0	0.043
16.07 structural protein binding	30	0.7	2	1.5	1.1	0.293	5	3.0	3.2	0.001	14	3.5	5.5	0.000

Table 3.3 Continued.

ID Functional Categories	Genome		Unfertilized				Fertilized				Common			
	Exp[a]	%	Obs[b]	%	Z-score	P-value	Obs[c]	%	Z-score	P-value	Obs[d]	%	Z-score	P-value
16.13 C-compound binding	21	0.5	1	0.8	0.4	0.685	4	2.4	3.2	0.001	6	1.5	2.5	0.012
20.09.07.03 ER to Golgi transport	55	1.3	2	1.5	0.2	0.843	7	4.2	3.1	0.002	7	1.8	0.8	0.453
32 Cell rescue, defense and virulence	594	14.1	25	18.8	1.5	0.127	34	20.4	2.3	0.024	92	23.1	4.8	0.000
32.01 stress response	351	8.3	19	14.3	2.4	0.015	26	15.6	3.3	0.001	74	18.6	6.8	0.000
36 Systemic interaction with environment	7	0.2	2	1.5	3.3	**0.001**	1	0.6	1.3	0.199	1	0.3	0.4	0.697
36.20 fungal specific systemic sensing and response	4	0.1	2	1.5	4.3	**0.000**	1	0.6	1.9	0.059	1	0.3	0.9	0.365
40 Cell fate	241	5.7	15	11.3	2.7	**0.007**	14	8.4	1.4	0.149	47	11.8	4.8	0.000
40.01 cell growth / morphogenesis	186	4.4	13	9.8	2.9	**0.004**	12	7.2	1.7	0.091	30	7.5	2.8	0.005
40.10 cell death	63	1.5	5	3.8	2.1	0.038	3	1.8	0.3	0.754	22	5.5	5.7	0.000
41 Development (systemic)	45	1.1	0	0.0	-1.2	0.231	2	1.2	0.2	0.873	12	3.0	3.4	0.001
41.01 fungal/microorganismic development	45	1.1	0	0.0	-1.2	0.231	2	1.2	0.2	0.873	12	3.0	3.4	0.001
42 Biogenesis of cellular components	523	12.4	18	13.5	0.4	0.700	30	18.0	2.1	0.034	74	18.6	3.5	0.000
42.10 nucleus	110	2.6	4	3.0	0.3	0.778	9	5.4	2.2	0.030	20	5.0	2.8	0.005

Table 3.3 Continued.

ID Functional Categories	Genome		Unfertilized				Fertilized				Common			
	Exp[a]	%	Obs[b]	%	Z-score	P-value	Obs[c]	%	Z-score	P-value	Obs[d]	%	Z-score	P-value
43 Cell type differentiation	280	6.6	11	8.3	0.7	0.461	17	10.2	1.8	0.075	45	11.3	3.5	0.001
43.01.03.09 development of asco- basidio- or zygospore	93	2.2	7	5.3	2.3	0.021	10	6.0	3.2	**0.002**	18	4.5	2.9	0.004
Total[e]	**4213**		**133**				**167**				**398**			

[a]Exp referes to the total number of genome predicted proteins that were expected to be annotated in the functional category
[b]Obs referes to the total number of unfertilized-specific proteins identified that were annotated in the functional category
[c]Obs referes to the total number of fertilized-specific proteins identified that were annotated in the functional category
[d]Obs referes to the total number of identified common proteins that were annotated in the functional category
[e]Total number of proteins functionally annotated in each protein group
Over-represented category (p<0.01)

I further investigated the metabolic pathways that undergo changes in sexual development. The KEGG pathway database has been used as a reference for the systematic analysis of gene function and for mapping cellular processes and organism behaviors from molecular data sets (KANEHISA *et al.* 2008). In this study metabolic pathways were determined for each protein group using the KEGG Automatic Annotation Server (KAAS). From the proteome, a total of 452 proteins were mapped to 97 metabolic pathways, 17 genetic processes, and 13 cellular processes. In addition, enzyme codes (EC) were assigned to 266 proteins. Complete annotation of the proteome of *N. crassa* sexual development, showing the metabolic pathways and enzyme codes determined by KAAS are displayed in Table B.1. Then, I sought to determine whether specific pathways were associated with different stages of sexual growth. For that, I compared the most predominant metabolic pathways and/or genetic processes for each protein sub-group. In all three protein sub-groups, the *biosynthesis of secondary metabolites* was the pathway with the highest representation of identified proteins, suggesting fundamental functions for these compounds during sexual development in *N. crassa* (Figure 3.9).

Figure 3.9 Top 15 KEGG pathways. Metabolic pathways and genetic processes, for each protein sub-group from the proteome of *N. crassa* sexual development. (A) Common proteins. (B) Unfertilized-specific proteins. (C) Fertilized-specific proteins. (D) Predicted total genome proteins.

A Common Proteins Sub-group

B Unfertilized-Specific Protein Sub-group

C Fertilized-Specific Protein Sub-group

D Genome Predicted Proteins

147
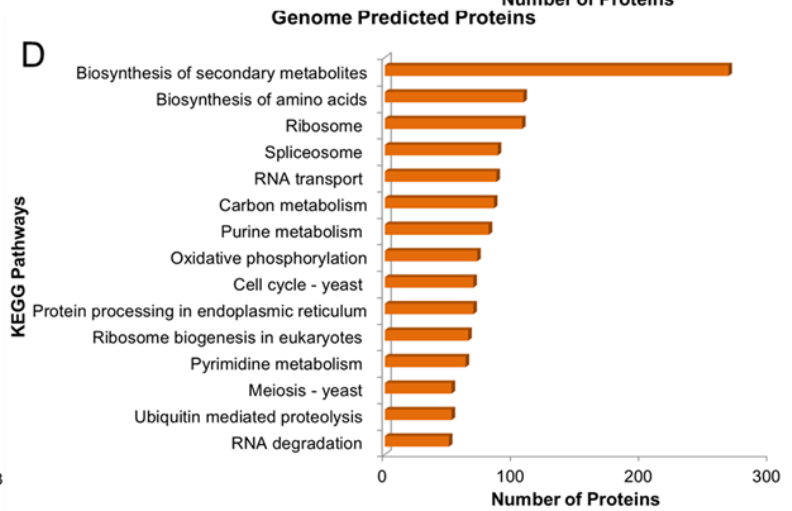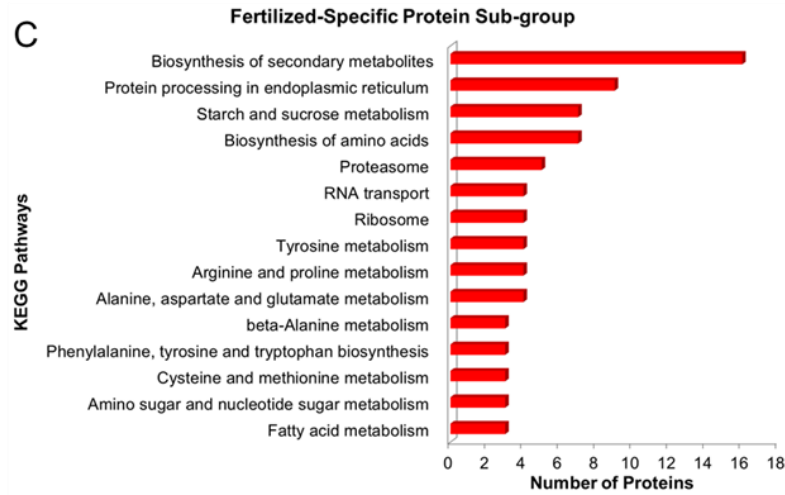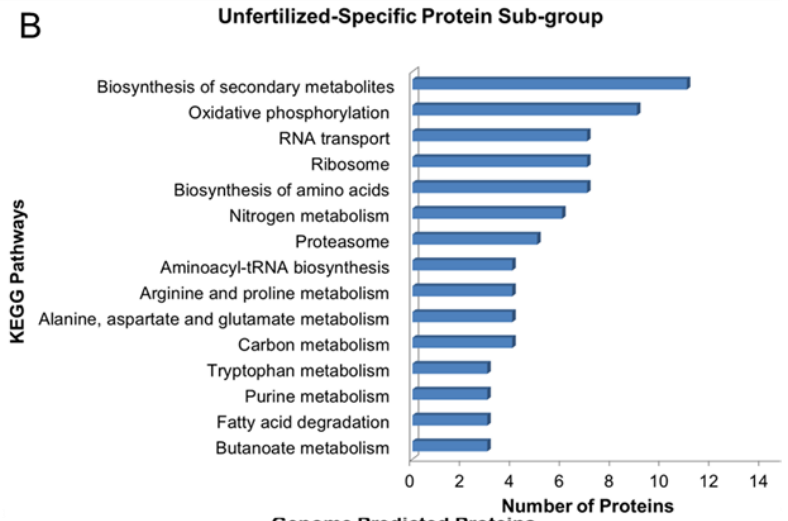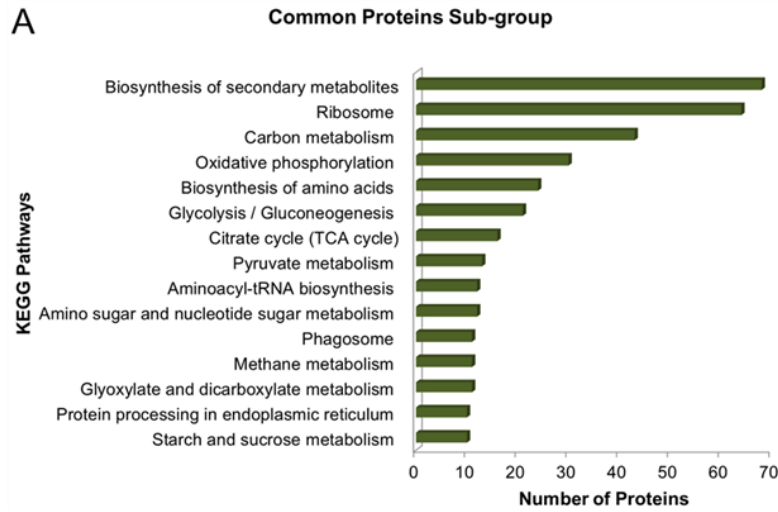
Protein Synthesis and Energy Metabolism are Enriched Processes throughout Sexual Development

Proteins classified as common were detected in both tissues, suggesting their role in vital processes required for normal cellular growth such as the synthesis of essential molecules, energy production, mechanical support, transport of small molecules and storage. In addition, over-representation of functional categories in this protein sub-group may reveal details about the biological and cellular processes associated with sexual growth in *N. crassa*. Such as, biological processes that lead cells to switch from vegetative to sexual cycle. Functional analysis results showed statistically significant enrichment in *proteins with binding function*, mainly *proteins that bind to other proteins*, proteins required for *translation*, *ribosomal proteins*, and proteins involved in *ribosomal biogenesis*. Also over-representation of *amino acid metabolism* was revealed (Figure 3.10A). Together, this result indicates an active translation process throughout sexual development. The functional analysis data was consistent with the KEGG pathway analysis. After *biosynthesis of secondary metabolites*, the genetic processes *ribosome* and *protein processing in endoplasmic reticulum*, jointly with metabolic pathways such as *biosynthesis of amino acids* and *aminoacyl-tRNA biosynthesis*, were the most represented categories by the common proteins (Figure 3.9A). Together, these results demonstrated that active protein synthesis occurred before and after fertilization during sexual development.

Figure 3.10 Functional classification and enrichment analysis of common proteins. Functional categories with statistically significant enrichment by common protein sub-group (green bars) (*P*-value <0.01) compared with the total predicted genome proteins (gray bars). (A) Categories with percentages of observed proteins equal or higher than 10%. (B) Categories with percentages of observed proteins below 10%.

The functional analysis of common proteins also showed that proteins involved in *cell growth and morphogenesis*, *cell death*, *fungal and ascospore development*, and *stress response* were enriched (Figure 3.10B). It is known that changes of some abiotic factors in the environment, like low nitrogen concentration, lower temperature, and increased light exposure will result in activation of the sexual cycle. Therefore, genetic and morphological changes are expected as a response to stressed situations. The only category under-represented was *transcription*, more specifically *RNA processing*. As I mentioned above, this result could be explained by the low-abundance characteristic of some proteins involved in transcription regulation, such as transcription factors.

Cell differentiation demands energy; therefore pathways associated with energy production represented by common proteins were expected. Via KEGG pathway analysis we found that *oxidative phosphorylation*, *glycolysis/gluconeogenesis*, and *pyruvate metabolism* were highly represented by common proteins. Carbon utilization was also enriched in this group of proteins, as *carbon metabolism*, *citrate cycle*, and *methane metabolism* were part of the top 15 pathways. The *citrate cycle* (*TCA cycle*) is a central metabolic pathway that completes the oxidative degradation of three essential molecules, carbohydrates, fatty acids, and amino acids, which are all required for essential biological and cellular processes (Figure 3.9A). Therefore, I conclude that essential biological processes were represented by the common group of

proteins. To reach any further conclusion in terms of sexual cycle specificity, it would be necessary to compare these results with vegetative mycelium proteomic data and determine which categories are over-represented only during sexual progress.

Systemic Interaction with the Environment and Morphogenesis are Processes Enriched in Unfertilized Tissue

MIPS functional analysis of unfertilized-specific proteins showed statistically significant over-representation in proteins involved in processes such as *protein synthesis*, specifically *translation*. Other functional categories that were enriched were *proteins with binding function*, especially those that bind to other proteins and peptides (Figure 3.11A). This active protein metabolism in the unfertilized tissue was also revealed by the KEGG pathway analysis, in which *RNA transport*, *ribosome*, *biosynthesis of amino acids*, and *aminoacyl-tRNA biosynthesis* were the metabolic pathways/genetic processes represented (Figure 3.9B). Together, these data are consistent with previously reported gene expression data from unfertilized tissue (WANG *et al.* 2012). In addition, KEGG analysis revealed that after the *biosynthesis of secondary metabolites*, *oxidative phosphorylation* was the second metabolic pathway represented by unfertilized-specific proteins (Figure 3.9B). Oxidative phosphorylation is necessary for ATP production. Thus, large amounts of energy may be required for the morphological changes, including protoperithecia formation, which occurs once

sexual growth is activated.

Because switching from vegetative to sexual growth depends on environmental changes, I expected to observe enrichment of proteins that participate in sensing and response to those changes. As expected, I observed enrichment in the categories of *specific systemic sensing and response*, *urea cycle* and *nitrogen /sulfur metabolism.* In addition, it has been shown that the fatty acid composition of sexual tissue differs greatly from the composition in asexual tissue, and those changes correlate with several events through sexual development (GOODRICH-TANRIKULU *et al.* 1998). Acordingly, *fatty acid metabolism* was another enriched category in this group of proteins. Due to morphological changes associated with protoperithecia formation, also expected to observe enrichment in proteins associated with cell fate, and this prediction was confirmed, especially in the subcategory of *cell growth/morphogenesis* (Figure 3.11B).
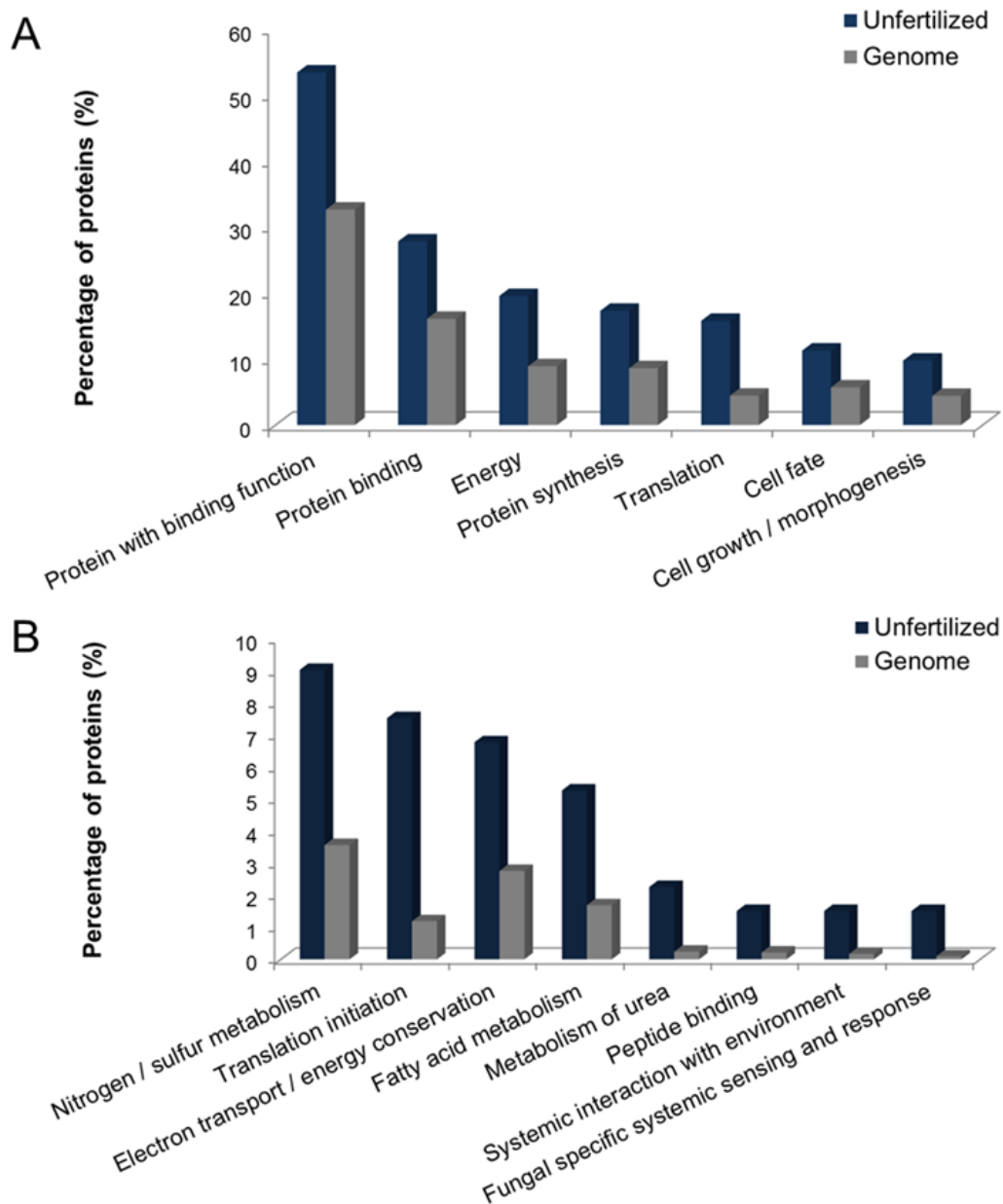
Figure 3.11 Functional classification and enrichment analysis of unfertilized-specific proteins. Functional categories with statistically significant enrichment by unfertilized-specific protein sub-group (blue bars) (P-value <0.01) compared with the total predicted genome proteins (gray bars). (A) Categories with percentage of observed proteins equal or higher than 10%. (B) Categories with percentage of observed proteins below 10%.

Secondary Metabolites Synthesis, Carbohydrates Metabolism, and Cell Type Differentiation are Enriched Processes in Fertilized Tissue

Once fertilization occurs, cellular morphogenesis and tissue differentiation is observed: perithicium is formed, meiosis and post-meiotic mitosis divisions occur, and ascospores develop. Therefore, I expected that proteins related to energy production, response to stress, and cell type differentiation would be present in fertilized tissue. As expected, the MIPS functional analysis of fertilized-specific proteins showed an enrichment in categories such as *proteins with binding function or cofactor requirement*, *protein fate*, *energy*, *stress respond*, *cell type differentiation*, *cellular transport* and various subcategories under *metabolism* (Figure 3.12). Consistent with my observations, enrichment in *stress response*, *proteins with binding function* and *cell fate* categories were reported in a functional analysis of genes up-regulated at 96 h after fertilization (WANG *et al.* 2012). Over-representations of proteins involved in diverse metabolic processes, including *metabolism of amino acids*, *carbohydrates*, *polysaccharides* and *secondary metabolites* were also statistically significant (P-value <0.01) (Figure 3.12A).
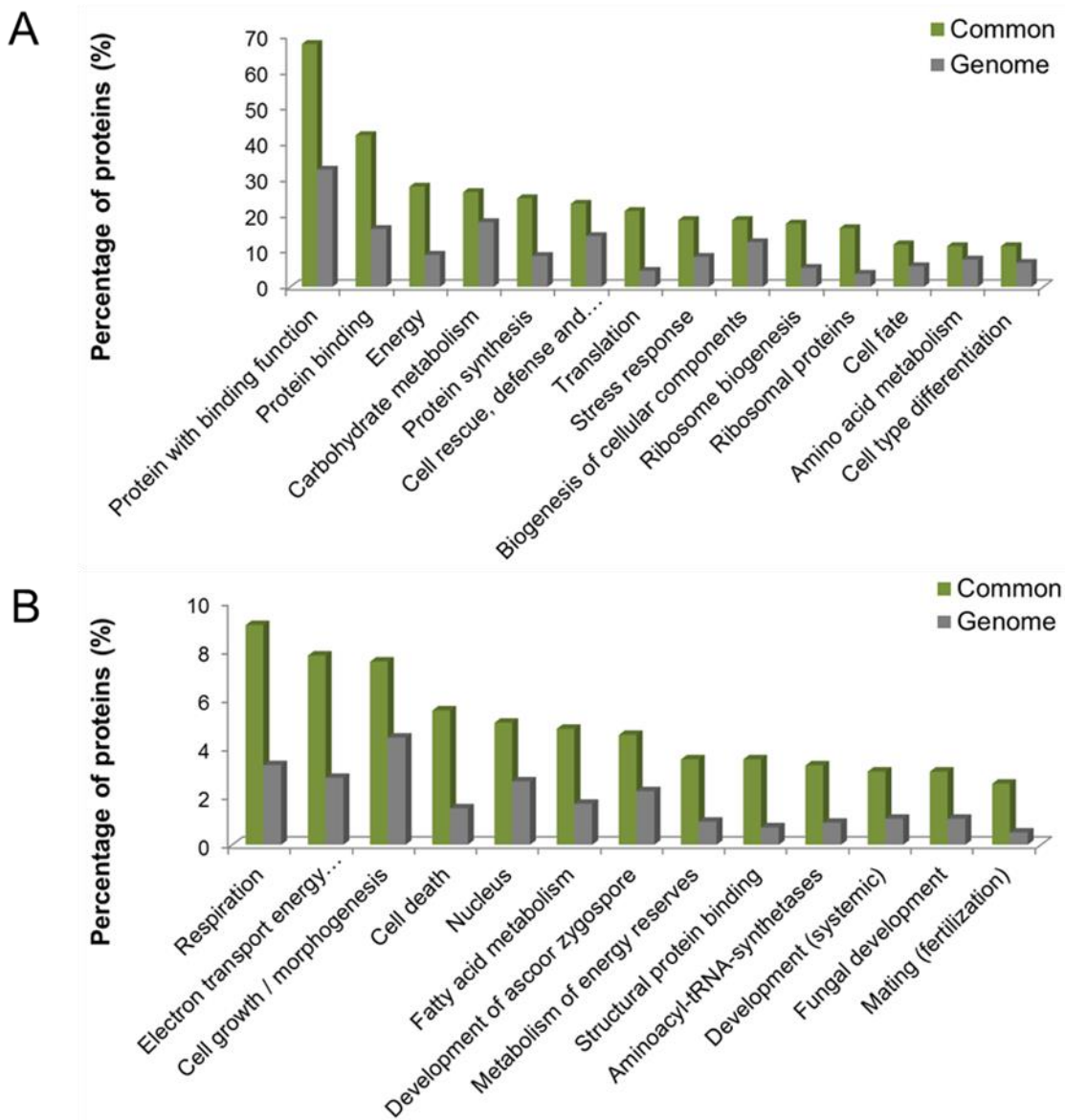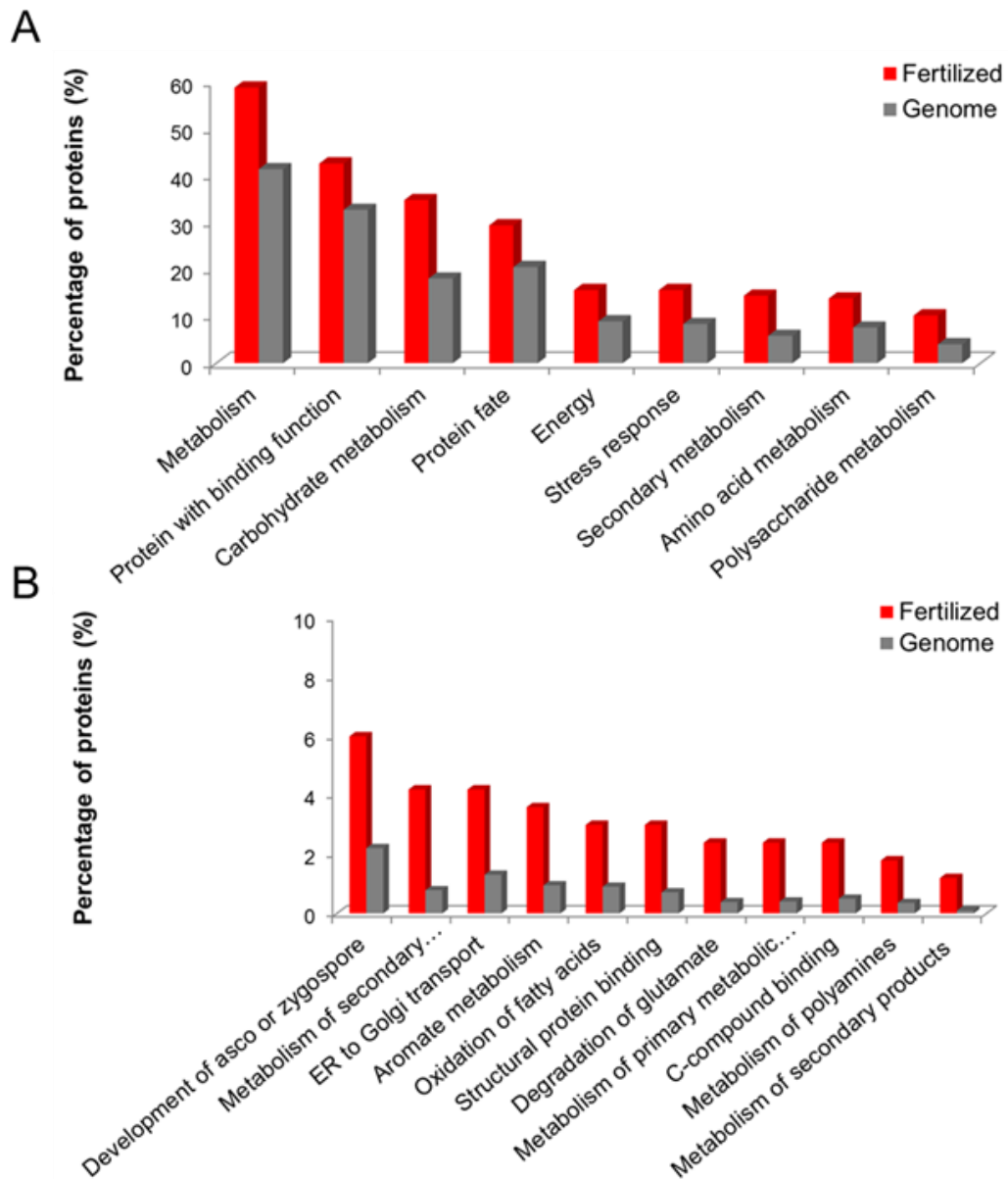
Figure 3.12 Functional classification and enrichment analysis of fertilized-specific proteins. Functional categories with statistically significant enrichment by fertilized-specific protein sub-group (red bars) (*P*-value <0.01) compared with the total predicted genome proteins (gray bars). (A) Categories with percentage of observed proteins equal or higher than 10%. (B) Categories with percentage of observed proteins below 10%.

Likewise, results observed from KEGG pathway analysis were in agreement with functional analysis. After *biosynthesis of secondary metabolites*, the subsequent metabolic pathway represented was *starch and sucrose metabolism*, which is crucial not only for the allocation of carbon resources but also for the initiation of hexose-based sugar signals (Figure 3.9C). Interestingly, *tyrosine metabolism* was also represented. Tyrosine is the precursor to the pigment melanin (MASON 1948; EISENMAN and CASADEVALL 2012), which is responsible for the black color of the sexual structures, perithecia and ascospores, in *N. crassa.* Melanin offers protection from UV rays, solar radiation, desiccation, enzymatic lysis and extreme environmental conditions (PAL *et al.* 2014).

Carbohydrate and Polysaccharide Metabolism are Important Processes for Ascus Maturation and Ascospore Development

Several enzymes with cellulose degradation activity were found exclusively in fertilized tissue. For example, four β-glucosidases out of the seven predicted in *N. crassa*—including, cellobiohydrolase I and II, exoglucanase 3, and endoglucanase II and IV—were identified in tissue harvested four days post-fertilization. In addition, two sugar transporters from the cellodextrin transport system (cellodextrin-1 and -2) were also accumulated after fertilization (Table 3.4). The high content of cellulase enzymes explains the enrichment of categories related to carbohydrate and polysaccharide metabolism observed in

fertilized tissue. Interestingly, previous reports have mainly focused on the cellulose system in mycelia (asexual cycle) (TIAN *et al.* 2009; GALAZKA *et al.* 2010; WU *et al.* 2013a). Recently, the secretome of vegetative cells of *N. crassa* during growth in microcrystalline cellulose was reported, and the core component of the cellulase system responsible for 43% of total cellulose degradation activity was established (PHILLIPS *et al.* 2011). I found that the same core components of the secretome are expressed in tissue harvested four days after fertilization, and the same proteins were undetected in unfertilized tissue. This evidence suggests that cellulases play an important role in fruiting body formation in *N. crassa.* In addition to the cellulase complex and the enrichment in carbohydrate metabolism, over-representation of proteins involved with *secondary metabolites* derived from sugar metabolism were also observed as described below.

Table 3.4 Proteins with cellulose degradation activity found specifically in fertilized tissue in *N. crassa*

Accession #	Protein name	Predicted function	Reference
NCU00130	GH1-1	β-glucosidases	(Wu *et al.* 2013b)
NCU04952[a]	GH3-4	β-glucosidases	(PHILLIPS *et al.* 2011)
NCU00577	GH3-5	β-glucosidases	(Wu *et al.* 2013b)
NCU07487	GH3-6	β-glucosidases	(Wu *et al.* 2013b)
NCU07340[a]	CBHI	Cellobiohydrolase	(PHILLIPS *et al.* 2011)
NCU09680[a]	CBHII	Cellobiohydrolase	(PHILLIPS *et al.* 2011)
NCU07190	GH6-3	Exoglucanase 3	(PHILLIPS *et al.* 2011)
NCU00762[a]	EG-II	Endoglucanase II	(PHILLIPS *et al.* 2011)
NCU01050	EG-IV	Endoglucanase IV	(PHILLIPS *et al.* 2011)
NCU00801	CDT-1	Cellodextrin-1	(GALAZKA *et al.* 2010)
NCU08114	CDT-2	Cellodextrin-2	(GALAZKA *et al.* 2010)

[a]These four proteins are responsible for 43% of total cellulase activity observed in *N. crassa* vegetative cells (PHILLIPS *et al.* 2011)

First Evidence of DHN-melanin Biosynthesis in *N. crassa* Sexual Development

Another important metabolic pathway that appeared enriched in fertilized-specific proteins was *secondary metabolism*, particularly *metabolism of primary metabolic sugar derivatives*, *secondary products derived from L-glutamic acid*, *L-proline and L-ornithine*, and *secondary products derived from L-phenylalanine and L-tyrosine* (Figure 3.12B). This result was consistent with the results obtained from KEGG analysis (Figure 3.9C). Filamentous fungi are well known for the production of a wide range of secondary metabolites (SMs); in many cases the benefits of these compounds on the organism's growth and development are unknown. However, the majority of these compounds are of medical, industrial and/or agricultural importance (KELLER *et al.* 2005). Biosynthesis of these natural products has been associated with fungal

158

development and cell differentiation, such as fruiting body formation and sporulation in *Aspergillus* (Calvo *et al.* 2002).

One of the most described secondary metabolites in filamentous fungi is Melanin, a component of the fungal cell wall required for protection of sexual spores and fruiting bodies (Eisenman and Casadevall 2012). Two metabolic pathways are involved in the biosynthesis of melanin: L-3,4-dihydroxyphenylalanine (DOPA) and 1,8-dihydroxynaphthalene (DHN). Synthesis of DOPA-melanin requires the enzyme tyrosinase that catalyzes the convertion of tyrosin to DOPA (L-3,4-dihydroxyphenylalanine), which then is polymerized to form melanin. (Jacobson 2000). On the other hand, several enzymes catalyze DHN-melanin synthesis, such as polyketide synthase, hydroxynaphthalene reductase, and ascytalone dehydratase (Figure 3.13). DHN-melanin has been associated mainly with pathogenic fungi where this macromolecule serves as a virulence factor. Genes encoding DHN-melanin enzymes have been found organized in a gene cluster in several fungi, such as *Aspergillus fumigatus* and *Penicillium marneffei* (Tsai *et al.* 1999; Woo *et al.* 2010), or dispersed in the genome as in *Colletotrichum lagenarium* and *Sordaria macrospora* (Tsuji *et al.* 2003; Engh *et al.* 2007).
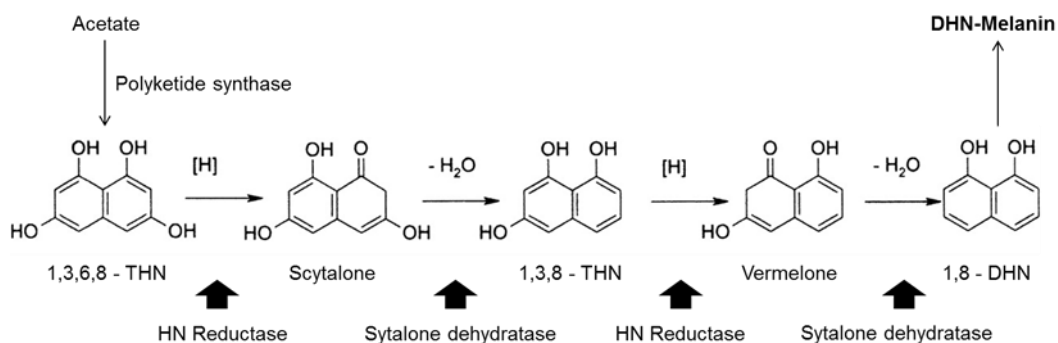
Figure 3.13 DHN-melanin biosynthesis pathway. Tetrahydroxynaphtalene (1,3,6,8 – TNH); Trihydroxynaphtalene (1,3,8 – THN); Dihydrohynaphtalene (1,8 – DHN). Enzymes are displayed in blue and the steps where they are involved are indicated by head arrows (Figure adapted from (BUTLER *et al.* 2009)).

Previous genetic and biochemical evidences, such as identification of tyrosinase activity under conditions favoring sexual growth, suggested that *N. crassa* produces melanin in the cell walls of perithecia and ascospores through DOPA-melanin pathway (LERCH 1983; FREE 2013). However, analysis of the *N. crassa* genome revealed the presence of putative proteins with similarity to the DHN-melanin biosynthesis enzymes described in *A. fumigatus*. This observation suggested that *N. crassa* is also able to produce DHN-melanin (GALAGAN *et al.* 2003; ENGH *et al.* 2007). Despite the computational predictions, no further studies have been done in this matter.

In this study, I identified different components responsible for DHN-melanin metabolism. All of them were detected exclusively from fertilized tissue and were classified under the category of secondary metabolism derived from L-

phenylalanine and L-tyrosine by MIPS FunCat. These proteins are type I polyketide synthase PER-1 (NCU03584), tetrahydroxynaphthalene reductase-1 TNR-1 (NCU09390), tetrahydroxynaphthalene reductase-2 TNR-2 (NCU06905), scytalone dehydratase (NCU07823), pigment biosynthesis protein AYG1 (NCU01903), and conidial pigment biosynthesis protein related to AYG1 (NCU05821).

MAPK Pathways are Involved in Sexual Differentiation

Mitogen-activated protein kinases (MAPK) pathways are critical downstream components of signal transduction pathways in eukaryotic organisms. They consist of three serine/threonine protein kinases (MAP3K, MAP2K, and MAPK) that act sequentially, culminating in phosphorylation of target proteins that regulate different cellular processes (CHEN and THORNER 2007). In filamentous fungi, MAPK pathways play an important role in growth, development, and pathogenesis. Nine MAPK proteins organized in three MAPK pathways have been identified in *N. crassa*. These pathways are i) the pheromone response (PR) pathway, ii) the cell wall integrity pathway, and iii) the osmoregulatory (OS) pathway (BORKOVICH *et al.* 2004).

In this study, I detected components of the PR and OS pathways during sexual development. Osmotic sensitive-2 OS-2 (NCU07024), a component of the OS pathway, was identified in fertilized tissue. The mitogen-activated protein kinase MAK-2 (NCU02393), a component of the PR pathway, was detected in

unfertilized tissue.

Several downstream targeted proteins of the MAK-2 protein kinase called Mak-2 Kinase-Regulated proteins (MRK) were also identified in this study. Pyridoxin 4-dehydrogenase MKR-3 (NCU02930), MKR-5 (NCU07449), and cupin domain-containing protein MKR-6 (NCU02919) were detected in both sexual stages and classified as common proteins. MKR-5 is an aerial hyphae development-related protein that plays an important role in cell-type differentiation in *N. crassa* (LI *et al.* 2005). *mrk-3* and *mrk-6* genes are found in a gene cluster on chromosome I in *N. crassa* along with a gene encoding a conserved hypothetical protein (NCU02921), also identified in this study in both sexual tissues. The other two members of the gene cluster are *mrk-2* (NCU02923) and a putative polyketide synthase *pks-6* (NCU02918). The orthologous cluster in *Sordaria macrospora* represents a putative polyketide biosynthesis pathway that is strongly up-regulated during fruiting body formation (NOWROUSIAN 2009). The same could be true in *N. crassa*.

DISCUSSION

I have generated the first proteomics data set for *N. crassa* under conditions that induced sexual development and performed a comparative proteomics analysis between two different phases of sexual differentiation: before and after fertilization.

Biases against short (<100 amino acids), long (> 1000 amino acids) and low-abundance proteins were detected. One possibility to explain this result is a limitation in mass spectrometry: short proteins yield lower numbers of peptides, reducing the probability for being detected by mass spectrometry analysis. On the other hand, this observed bias against large peptides may be an accurate reflection of the proteome. According to the biosynthetic cost minimization hypothesis (WARRINGER and BLOMBERG 2006), small proteins (between 100 and 600 amino acids) are more likely to be expressed than large proteins, thereby reducing the burden of protein synthesis on the cell. In addition, short mRNAs are also more efficiently transcribed and translated; they are also more stable than long transcripts (LACKNER et al. 2012). Therefore, small proteins tend to be more abundant and are more likely to be detected. Despite the biases detected, this study achieved a comprehensive representation of the sexual growth process in terms of protein population: the majority of annotated GO categories were represented by the proteome of N. crassa sexual development when compared with the total annotated predicted proteins.

Functional annotation analysis revealed that proteins synthesis and energy metabolism, are important processes that are required during the course of sexual development. However, each stage of sexual growth evidently has a distinctive enrichment pattern of functional categories. While categories such as *systemic interaction with environment* and *cell growth and morphogenesis* were over-represented in the unfertilized tissue, categories like *carbohydrate*

metabolism and *biosynthesis of secondary metabolites* were predominant among fertilized-specific proteins. These results were consistently observed with the three approaches I used to analyze the data (Blast2GO, MIPS Functional Catalogue and KEGG pathways).

Cellulase Activity During Fruiting Body Maturation

The proteins that constitute the core component of the cellulase system were found exclusively in the fertilized tissue. This data suggests that cellulases play an important role in fruiting body formation in *N. crassa.* In nature, *N. crassa* produces perithecia embedded within colonized plant tissue (PERKINS 2002). It is therefore not surprising to find that production of cell-wall-degrading enzymes may be associated with perithecia development. The cellulase system has been extensively studied in the secretome of vegetative cells; however, to my knowledge, no previous studies have reported cellulase activity in fungal sexual structures. Finding the proteins that compose the core secretome of *N. crassa* expressed during maturation of fertilized sexual structures, opens new possibilities for exploring fungal degradation of plant biomass and ultimately production of biofuels. It will be of interest to determine biomass degradation efficiency by the secretome produced during sexual development.

Does DHN-melanin have a Function During *N. crassa* Sexual Development?

Melanin is an important component of the cell wall of fungal sexual structures and is produced by two independent pathways: DHN and DOPA. DOPA-melanin has been frequently associated with sexual development in *Neurospora*. However, I found that all of the proteins required for DHN-melanin biosynthesis were present exclusively in fertilized tissue during perithecia maturation. My findings are in agreement with the hypothesis that *N. crassa* is able to produce DHN-melanin and, as it was demonstrated in *S. macrospora*, melanin biosynthesis is linked to fruiting body development (ENGH *et al.* 2007).

Tyrosinase—the key enzyme for DOPA-melanin synthesis— was not detected in this study, neither before nor after fertilization. Different possibilities could explain this result. First, tyrosinase may not be produced during sexual development. However, previous observations demonstrated that in the absence of the gene that encodes this enzyme, the formation of perithecia is compromised. Tyrosinase may also be a low-abundance protein; hence, not detected by mass spectrometry analysis. Although tyrosinase was not detected in our data, I cannot discard the possibility that DOPA-melanin is also produced during sexual development, but at different time points from the ones evaluated in this study. However, this is the first report showing physical evidence that all the proteins/enzymes required for the synthesis of DHN-melanin were detected exclusively in fertilized tissue, suggesting that DHN-melanin may be synthesized

during sexual development in *N. crassa*. Therefore, it would be of interest to study the molecular genetics of melanin biosynthesis in *N. crassa* to determine whether DHN and DOPA-melanin play similar roles during fruiting body maturation, or if they are formed at different stages of sexual development.

MAPK Pathways are Involved in Sexual Differentiation

I detected components of the pheromone response (PR) and the osmoregulatory (OS) MAPK pathways during the course of sexual development. Osmotic sensitive-2 OS-2 (NCU07024) is a component of the OS pathway and was identified in fertilized tissue. OS-2 is homologous to HOG1 in *S. cerevisiae*, which is a protein kinase involved in a signal transduction pathway that is activated by changes in the osmolarity of the environment (O'ROURKE and HERSKOWITZ 2004). In *N. crassa*, it has been demonstrated that the *os-2* gene product is necessary in osmoregulation, conidial integrity, fungicide sensitivity, protoperithecia development and female fertility (ZHANG *et al.* 2002; JONES *et al.* 2007).

The mitogen-activated protein kinase MAK-2 (NCU02393), a component of PR pathway, was detected in unfertilized tissue. MAK-2 is homologous to FUS3/KSS1 in *Saccharomyces cerevisiae.* FUS3, together with KSS1, are the final kinases in the signal transduction cascade regulating activation/repression of the mating and filamentation pathways, induced by pheromones and nitrogen/carbon limitation (TEDFORD *et al.* 1997). In *N. crassa*, *mak-2* mutant

strains showed abnormal filamentous growth and development of aerial hyphae; in addition, protoperithecia formation was compromised due to female sterility, which arrested sexual development in an early stage. Another observation was de-repression of conidial morphogenesis during sexual development (LI *et al.* 2005). Previous genetic evidence and current proteomic data suggest an important role for the *mak-2* gene product in protoperithecia formation.

Interestingly, several Mak-2 Kinase-Regulated proteins (MRK), whose genes are organized in a cluster, were detected. This gene cluster represents a putative polyketide biosynthesis pathway that is strongly up-regulated during fruiting body formation in the closest filamentous fungus *S. macrospora* (NOWROUSIAN 2009). Our data shows the first evidence of proteins from the cluster being accumulated not only in unfertilized tissue, but also in *N. crassa* tissue 4 days after fertilization. Even though not all gene products from the cluster were detected in the two stages analyzed, our results are consistent with previously reported transcription data (LI *et al.* 2005; NOWROUSIAN 2009). This result indicates that secondary metabolites, including polyketides, may play an important role in protoperithecia formation and sexual development in *N. crassa*. Previously, it was shown that in the absence of the *mak-2* gene product, the expression of genes from this cluster is reduced (LI *et al.* 2005). Combining transcription and proteomic results, the data suggest that genes from the cluster are regulated by the MAPK pathway MAK-2. Finding this cluster represented in the proteome of sexual development in *N. crassa*, and the regulatory protein

167

kinase MAK-2, suggests a connection between MAPK pathways, polyketide biosynthesis, secondary metabolites and sexual differentiation in filamentous fungi.

Other Proteins Required for Sexual Development

Although various genes required for sexual development in *N. crassa* have been described in the past two decades (NELSON and METZENBERG 1992; NELSON *et al.* 1997b; KIM and NELSON 2005), most of this information comes from gene expression data obtained under specific conditions and not protein measurement data. Therefore, this work represents the first comparative proteomic data generated from sexual development tissues. A goal of this study was to consolidate the genetic and the proteomic information relevant to the gene products that play an important role during sexual differentiation. To do this, I compared our proteomic data with transcription data published in the past. In this study, I identified several proteins that were expressed in one or both stages of sexual development. The genes for some of these proteins were previously reported as required for normal sexual growth. However, for a number of proteins, I present the first evidence of their detection in tissues where they were not previously reported. This finding has important implications in the understanding of the progression of sexual development. I discuss the most relevant examples below.

The protein ascus development-1 ASD-1 (NCU05598) is a rhamnogalacturonase B (RGase B) enzyme essential for normal sexual development and ascospore delineation (NELSON *et al.* 1997b). Evidence of ASD-1 protein accumulation at 5 and 7 days after fertilization was shown by Nelson *et al*. In addition, neither protein nor transcript from the *asd-1* gene was detected under vegetative growth (NELSON *et al.* 1997b). A previous study concluded that the *asd-1* gene product plays an essential role in the late states of sexual development, especially in ascospore production (NELSON *et al.* 1997b). In contrast, I detected ASD-1 protein not only in fertilized tissue, but also in unfertilized tissue. Therefore, this is the first report that reveals the expression of the ASD-1 protein in sexual growth prior to fertilization, bringing to light the importance of ASD-1 in early sexual development before fertilization,

Abundant perithecial protein APP (NCU04533) was described as highly expressed and specific for perithecia—constituting 35% of total perithecia protein—but was not considered essential for sexual growth in *Sordaria macrospora* and *N. crassa* (NOWROUSIAN *et al.* 2007). Previously, it was reported that *N. crassa app* transcripts were present prior to fertilization, but the protein accumulated only after fertilization (NOWROUSIAN *et al.* 2007). My data conflicts with this conclusion. APP is not a perithecia-specific protein as was reported previously. In this study, APP was identified prior to and after perithecia formation, suggesting that APP is not only present after fertilization but also in early stages of protoperithecia formation. The discrepancy between these

results could be explained by the differences in time points analyzed by the two studies. I analyzed 6 day old unfertilized tissue; however, 10-, 13-, 14-, and 16-day- old unfertilized cells were studied by Nowrousian *et al*. (NOWROUSIAN *et al.* 2007). Combining previous data and the data of this study, it is possible to imagine a scenario in which APP protein is conditionally degraded in unfertilized tissue. If fertilization does not occur, at some point after 6 days of growth the protein is destroyed; however, if fertilization occurs, this protein accumulates during perithecia development. Although not essential for sexual development, APP is an interesting protein to study because it is specific and abundant under conditions for sexual growth, indicating a possible role in a process that only occurs during sexual cycle (*e.g.*, meiosis).

Catalase 1 CAT-1 (NCU08791) was detected in both tissues, and is one of the two major fungal-specific *Neurospora* catalases required for oxidative stress response. CAT-1 is predicted to be involved in the production of secondary metabolites. In filamentous fungi experimental evidence has shown that secondary metabolism is triggered by oxidative stress (ROZE *et al.* 2011). CAT-1 activity has been extensively studied in asexual cycle differentiation and found predominantly in conidial germination and early mycelia growth (PERAZA and HANSBERG 2002). The first connection between CAT-1 and sexual differentiation was proposed in 2007 by Yamashita *et al.,* where it was shown that OS-2 MAP kinase, in combination with an unknown regulatory system, were required for the regulation of CAT-1 protein accumulation in conidial

development. The authors suggested that this combined regulation of CAT-1 could be also involved in sexual differentionl (YAMASHITA *et al.* 2007). To the best of my knowledge, there are no reports of *cat-1* transcripts or CAT-1 protein accumulation during sexual development in *N.* crassa. However, I found CAT-1 in both sexual stages. It would be of interest to study the role of CAT-1 in cell differentiation during sexual growth in *N. crassa* to determine if there is any connection with production of secondary metabolites.

ATP citrate lyase (ACL) is involved in the formation of cytosolic acetyl-CoA, an essential metabolite required in protein acetylation and intermediary carbon and energy metabolism, including biosynthetic pathways such as fatty acid and sterol formation. ACL is the link between the metabolism of carbohydrates and the production of fatty acids. In animals, ACL isozymes are encoded by one gene, but in *Pezizomycotina* fungi two different subunits are encoded by two separate genes, and they seem to be clustered in the genome (NOWROUSIAN *et al.* 2000). As was found in *S. macrospora*, *Gibberella pulicaris*, and *A. nidulans*, *acl1* and *acl2* are clustered genes that are divergently transcribed in *N. crassa*, with the key difference being a very short dubious predicted gene in the middle. In the present study, ACL1 and ACL2 proteins (NCU06785 and NCU06783 – E.C.2.3.3.8) were found in both unfertilized and fertilized tissues of sexual growth.

The first molecular analysis of ACL1 in fungi was done in *S. macrospora* (NOWROUSIAN *et al.* 1999), in which *acl1* mutants exhibited defects in fruiting

171

body maturation. A time course analysis of protein accumulation revealed that ACL1 was specifically induced at the beginning of the sexual cycle. Therefore, it was hypothesized that the crucial role of ACL1 in *S. macrospora* was to produce certain amounts of acetyl-CoA and its derivatives during the early stages as a prerequisite for later perithecia maturation (NOWROUSIAN *et al.* 1999). In *A. nidulans,* a complete loss of sexual development was found in *acl* deletion strains (HYNES and MURRAY 2010).

Conversion of cytoplasmic acetyl-CoA to malonyl-CoA, by the enzyme acetyl-CoA carboxylase (ACC), is an important step in the biosynthesis of fatty acids and various secondary metabolites. In this study, acetyl-CoA carboxylase (NCU08535) was also detected in both stages of sexual development. Changes in the composition of fatty acid during sexual development have long been reported in *N. crassa*, showing higher concentrations of oleate in developing asci and ascospores than in perithecial wall tissues (GOODRICH-TANRIKULU *et al.* 1998). As oleate is a metabolic derivative of acetyl-CoA, this could explain the relevance of enzymes like ACL and ACC during sexual growth and ascospore formation in *N. crassa.* However, adding oleate to the media did not induce fruiting body formation in *acl* mutants in *A. nidulans*, suggesting the possibility that ACL has an additional role during sexual development (HYNES and MURRAY 2010).

Acetyl-CoA is also the substrate for acetylation of proteins including histones. A recent study demonstrated how glucose availability determined

global histone acetylation through an ACL-dependent pathway in mammalian cells differentiation, and how this histone ACL-dependent acetylation selectively affected the expression of genes involved in glucose metabolism and macromolecular synthesis (WELLEN *et al.* 2009).

In this study, I report the first evidence of the presence of ACLs enzymes in two stages of sexual development in *N. crassa*. *acl1* or *acl2* transcripts were not reported in the transcriptional profile analysis during fruiting body formation, suggesting that the genes were not expressed continuously throughout sexual development, as only genes expressed in all time points (between 0 and 140 h post-fertilization) were reported (WANG *et al.* 2012). It will be of interest to determine the function of ACL1 and ACL2 during sexual development in a model organism such as *N. crassa*. Determining whether the function of ACLs during sexual development is related to fatty acid biosynthesis and/or with histone acetylation and gene regulation will have an impact not only in fungal biology, and also in connecting metabolic pathways to cellular differentiation and development of higher organisms. Establishing ACLs cellular localization will help to reveal the molecular function, cytosol localization for fatty acid biosynthesis or nuclear localization for histone acetylation. In addition, determining if a regulatory mechanism is involved in the divergent transcription of these clustered genes will contribute in understanding the molecular genetics behind sexual development.

Are Suppressors of Meiotic Silence Represented in the Proteome?

In *N. crassa*, mRNAs transcribed from an unpaired DNA region, and from regions with high sequence similarity to the unpaired DNA, are destroyed by a meiotic silencing mechanism (Aramayo and Selker 2013). Several suppressors of meiotic silencing have been identified by genetic screens (Chapter I) and two new components were found in a biochemical screen (Chapter II). The general goal in my research was to gain insights into the molecular regulation of this mechanism (Chapter II). For that reason, I wanted to know if relevant information about suppressors of meiotic silencing could be extracted from the proteomic data reported in this study.

To understand the molecular function of suppressor of meiotic silencing proteins, it is useful to define protein expression and accumulation of these proteins at different points of development. Therefore, I wondered whether suppressor proteins were represented in the proteome of sexual development. In this study, Argonaute-like protein SMS-2 (NCU09434) was detected in both unfertilized and fertilized tissues; however, it was the only suppressor observed. One explanation for this result is that suppressors are likely low-abundance proteins: thus making their identification difficult in proteomic studies. Another possibility—although not mutually exclusive—is that components of meiotic silencing, at least some of them, are produced transiently in early meiosis. Therefore, by the time I harvested the fertilized tissue most of the suppressors may be already degraded. To test this possibility, it would be necessary to

analyze proteins from tissues harvested at different time points of development, especially tissue harvested immediately after fertilization when the two nuclei fused and unpaired DNA is detected.

This study represents the first attempt to introduce proteomics in the study of sexual development. The standardized protocol presented here could be applied for further studies, such as the analysis of suppressors throughout different time points of development.

Conclusion

This study presents a comparative proteomics analysis and provides evidence of differential metabolic features between two stages of sexual development: before fertilization and 4 days after fertilization. A comprehensive proteomics data set, composed of 841 proteins, and a functional catalogue, which includes GO terms, metabolic pathways maps and enzyme codes was developed. I found that sexual development is enriched in secondary metabolites biosynthesis, especially after fertilization. All the proteins/enzymes responsible for the synthesis of DHN-melanin, an important secondary metabolite, were detected in fertilized tissue. To date, DOPA-melanin was the only form of melanin reported for *N. crassa*. However, this study provides the first evidence that supports the hypothesis that DHN-melanin is also produced by *N. crassa*. Furthermore, the data suggests that DHN-melanin is synthesized only after fertilization during fruiting body maturation.

In addition, this work provides the first evidence of protein accumulation during sexual growth for several genes, which participation in sexual development was previously suggested. Therefore, my data complement previous reports and extend the knowledge about sexual proteins.

The analysis of the proteome of sexual development in *N. crassa* contributes to the understanding of the molecular processes occurring during sexual growth not only in *Neurospora*, but also in other filamentous fungi.

MATERIALS AND METHODS

Strains and Growth Conditions

Standard *Neurospora* culturing techniques (DAVIS 1970) were used throughout the study, except for the preparation of sexual tissue for protein extraction (described below). Vegetative mycelium was cultivated in Vogel's Medium N with 2% sucrose, and sexual development was induced in Westergaard's Medium with 1.5% sucrose. The formulas for both Vogel's and the Westergaard's Media have been previously described by Davis and de Serres (1970) (DAVIS 1970). The laboratory *N. crassa fluffy* mutant strain RANC49 (*mating type A*) and the wild-type strain FGSC2490 (*mating type a*), obtained from the Fungal Genetics Stock Center, were used as female and male strains, respectively.

Tissue Preparation

Protein extraction from the sexual stages of *Neurospora* can be challenging, due to the physical characteristics of the tissue. The small size of the sexual structure (protoperithecium and perithecium) and the thickness of its wall make it very difficult to disrupt the tissue and extract all meiotic cells. As a result, a large amount of biological material is required to obtain a useful concentration of proteins. To overcome these challenges, I developed an innovative strategy that allowed us to growth a considerable amount of sexual tissue using limited resources and laboratory space (Chapter II; Figure 2.16).

Instead of regular petri dishes, Pyrex glass pie plates (9-1/2-inch) were used to grow *N. crassa* under mating conditions. Two pie plates were coupled together in a way that one plate served as the lid for the other plate. Therefore, a reduced spaced for incubation was achieved. Uniform distribution of light was assured to both the bottom and the top plates, due to the large area of the plate and the transparency of the glass. 200 ml of Westergaard's Media was poured per plate and twenty plates were prepared per extraction. Westergaard's solid media was covered with stripes of pre-washed cellophane (7" X 1.5"). For the female strain inoculum, three plugs from RANC49 strain grown on supplemented solid Vogel's media were distributed to each plate. Plates were incubated at 25°C with constant light for 6 days to induce female structure development. To collect the tissue from unfertilized plates, tissue was scraped from the cellophane of 10 plates using a sterile spatula and proteins extracted. The other

ten plates were fertilized with FGCS2490 strain by spreading male conidia suspended in Vogel's Media. These were incubated at 25°C for an additional 4 days before protein extraction (to obtain fertilized tissue).

Cellophane Preparation

Cellophane was cut into stripes (7 X 1.5"), placed into 1 liter of deionized water and boiled in the microwave for 10 minutes. This step was repeated three times after first replacing the water with fresh water and boiling again. The washed cellophane was transferred to a beaker with deionized water and autoclaved for 20 min. Before inoculation, cellophane stripes were laid on top of Westergaard's solid media plates. The cellophane facilitated the harvesting of the tissue, because sexual tissue could then be peeled from the surface of the cellophane.

Protein Extraction

Unfertilized and fertilized sexual tissues were peeled from the cellophane, frozen in liquid nitrogen and then ground using a mortar and pestle. Each tissue collection was placed into pre-cooled Mill machine-plastic containers and ground to powder in a Mill machine (SPEX 6850 Freezer/Mill) using the following procedure: 10 minutes pre-cooling, 2 minutes milling, 1 minute pre-cooling and 2 minutes milling. All of the following steps were performed at 4°C. The frozen powder was weighed and transferred to 1.5 volumes of extraction buffer that

contained 50 mM HEPES pH 7.4, 137 mM NaCl, 10% Glycerol, 0.1% NP-40, 1X

PIC, 1 mM PMSF. The powder was immediately re-suspended by stirring in a

cold room for 10 minutes. The tissue was filtered using pre-treated and pre-

cooled cheesecloth/Miracloth. The lysate was centrifuged in a JA20 rotor

(Beckman) at 15,000 g for 15 minutes. The supernatant was then transferred to

a clean pre-cooled conical tube and total protein concentration was calculated

by Coomassie (Bradford) protein assay (Thermo Scientific). A total of 1 mg

protein was transferred to a 1.5 ml eppendorf tube and the proteins were

precipitated by adding sodium deoxycholate (DOC) to a final concentration of

0.2%, mixed and incubated on ice for 30 minutes. Then 15% of trichloroacetic

acid (TCA) was added to the mixture and incubated for 1 hour on ice. The

sample was centrifuged 10 minutes at 13,000 g and the pellet was washed with

pre-cooled acetone. After 5 min of incubation at room temperature, the sample

was spun down and the pellet was air-dried prior to downstream mass

spectrometry analysis. For the proteomic analyses, two biological replicates of

protein samples from two different experiments were obtained for each condition

analyzed.

Liquid Chromatography Mass Spectrometry

Two analytic replicates were analyzed. Each protein sample was reduced, alkylated and enzymatically digested using trypsin. Peptides were then fractionated using an in-house isoelectric point based electrophoretic device (LIM *et al.* 2007). Usially, 500ug of the protein digested was loaded into the MSWIFT device and peptides were separated according to their isoelectric point into 6 fractions using buffering membranes with the following pH values: 2.9, 4.3, 5.2, 6.6, 7.5, 9.5 and 11, as previously described (COLOGNA *et al.* 2010). An aliquot from each fraction (estimated to be 1-3ug) was then further separated by reversed phase liquid chromatography separation coupled off-line with MALDI-MS/MS as previously described (ROSAS-ACOSTA *et al.* 2005). The column eluent was mixed with the MALDI matrix (7 mg mL-1 α-cyano-4-hydroxycinnamic acid, 60% (v/v) acetonitrile, 10 mM ammonium dihydrogen phosphate, 10% isopropanol) via a mixing tee and the resulting mixture was spotted onto a MALDI target via a robotic spotting device (Probot, ThermoScientific). MALDI mass spectra were acquired using a 4800 Proteomics Analyzer (Applied Biosystems, Framingham, MA). External calibration was performed using the standard peptides bradykinin fragment 2-9 and adrenocorticotropic hormone fragment 18-39. Collision-induced dissociation (CID) spectra were acquired using air as the collision gas (medium pressure setting) and at 2 kV of collision energy. Tandem mass spectrometry (MS/MS) data were searched against the *Neurospora crassa* database (downloaded 04/30/2013, www.broadinstitute.org)

using the ProteinPilot Software v. 3.0 and the Paragron™ Algorithm (Applied Biosystems, Framingham, MA).

Proteins comprising one or more peptides with high confidence score (>95%) and a low false discovery rate (FDR <5%) were considered positively identified. Distinct proteins were classified in three protein sub-groups depending on tissue-specificity: common proteins (CP), unfertilized-specific proteins (USP), and fertilized-specific proteins (FSP).

Gene Ontology (GO) Annotation via Blast2GO

Gene Ontology (GO) annotation was performed using InterProScan program along with the Blast2GO algorithm (V 2.7.0) (http://www.blast2go.com/b2ghome) (CONESA and GOTZ 2008; GOTZ *et al.* 2008). For annotation, the default configuration settings were used and the proteins were searched against the Swiss-Prot protein database (June 2013). Generic GO terms were then retrieved using the GO-slim option.

Functional Annotation and Statistical Analysis

The MIPS Functional Catalogue (FunCat) annotation scheme (http://mips.helmholtz-muenchen.de/funcatDB/) was applied to the complete list of protein identifications and the three protein sub-groups (CP, USP and FSP) to organize proteins according to their cellular and molecular functions (RUEPP *et al.* 2004). Enrichment or depletion of functional categories across protein sub-

groups compared with the genome was evaluated using a Z-score analysis based on the differences between the observed and expected proportions, given that the sample sizes and the protein frequency for each category is known. The *P*-value was determined by calculating the standard normal distribution of *Z*-score (RIVALS *et al.* 2007).

KEGG Pathways Analysis

The FASTA protein sequences of identified proteins were searched against KEGG GENES database (KANEHISA *et al.* 2008; OKUDA *et al.* 2008) using BLASTP program via KEGG Automatic Annotation Server (KAAS) (http://www.genome.jp/tools/kaas/). The corresponding KEGG pathways and enzyme codes (EC) information for each protein sub-group were extracted.

CHAPTER IV

CONCLUSIONS AND FUTURE DIRECTIONS

SUMMARY

In the Research Aims section of Chapter I, four specific objectives were outlined for my work. These objectives became the focus of Chapter II and III. In Chapter II, I present the results for the first biochemical work developed as a discovery tool for the identification of unknown protein interactions during sexual development in *N. crassa*. I standardized a biochemical strategy that allowed me to establish protein-binding partners for the suppressor of meiotic silencing SMS-5. These interacting partners, PAF400 and Pianissimo, represent new molecular components involved in the meiotic silencing mechanism. These results led me to propose a molecular role for SMS-5 and its interacting proteins in the silencing of unpaired DNA. Interactions between these three components, SMS-5, PAF400, and Pianissimo establish a connection between chromatin remodeling, DNA repair, signaling transduction pathways and meiotic silencing.

SMS-5 may be a protein methyltransferase that regulates protein interaction and function of its targets during meiotic silencing. Based on the molecular function described for PAF400 homologous proteins, I proposed that PAF400 acts as a scaffold protein (GRANT *et al.* 1998; MCMAHON *et al.* 1998; ALLARD *et al.* 1999; BOSOTTI *et al.* 2000; MUTIU *et al.* 2007; KNUTSON and HAHN 2011) that mediates the recruitments of different components to the unpaired

DNA region after been detected via *trans*-sensing. PAF400 may be involved in chromatin remodeling of the unpaired region (GRANT *et al.* 1998; DOYON and COTE 2004; MURR *et al.* 2007). Chromatin relaxation ensures accessibility of the transcriptional machinery responsible for the synthesis of the aRNA molecules. On the other hand, Pianissimo may be involved in activating a signal transduction pathway (CYBULSKI and HALL 2009; OH and JACINTO 2011) that could be the key molecular process that connects nuclear events and activation of the perinuclear stage of silencing. The data presented here also suggest a mechanism that connects the nuclear and the perinuclear stages required for silencing of unpaired DNA during meiosis. Altogether this study has provided new insights into the molecular events associated with the initial nuclear stage of meiotic silencing.

In Chapter III, I describe the experiments and data analyses used to develop a comprehensive proteomics data set and a functional catalogue for *N. crassa* sexual development. I used a global proteomics approach and comparative protein functional analysis to investigate the potential molecular differences between two stages of sexual development in filamentous fungi. This study provides evidence of differential metabolic features between two stages of sexual development, before and after fertilization. It also provides direct evidence of differential sexual expression of proteins associated with secondary metabolites biosynthesis and cellulase activity, indicating that both processes are required in fruiting body maturation. A functional catalogue of sexual

development proteins in *N. crassa* will serve as a reference tool for further studies related to sexual development not only in *N. crassa*, but also in other filamentous ascomycetes.

CONCLUSIONS

Improving Protein Yield from Sexual Tissue

In order to use biochemical and proteomics approaches into the study of sexual differentiation and meiosis-related events in *N. crassa*, it was necessary to ensure that a significant concentration of total proteins can be obtained from the sexual tissue. Obtaining a high protein yield that could be used for further biochemical and proteomics analyses was a critical initial step in my research. Accordingly, I needed to improve the culturing technique for growing sexual tissue and the disruption technique for protein extraction. For that purpose, I developed a cost/space-efficiency culturing strategy for growing considerable amount of sexual tissue that involved the use of glass pie plates (9-1/2- inch) instead of the traditional petri dishes. I also combined the classical mortar and pestle trituration of the tissue with mechanical disruption using a Mill machine for breaking the sexual structures. As a result, protein yield was increased to an optimal level for further biochemical protein work. Using this strategy, I obtained enough biological material from sexual tissues to study protein-protein interactions (Chapter II). The increased protein yield also allowed me to determine the proteome of sexual development by analyzing protein populations

from two stages of sexual growth (Chapter III). These new standardized procedures for growing sexual tissue, tissue disruption, and protein extraction represent a significant technical improvement for further proteomic and biochemical studies in filamentous fungi during sexual differentiation.

Introducing Biochemical Approaches to the Study of Meiotic Silencing

There was a need for introducing biochemical analyses to the study of meiotic silencing in *N. crassa*. Despite the numerous gene products determined to be involved in MSUD, very little is known about the molecular mechanisms underlying this intricate process. In addition, the precise function executed by most of the suppressor of meiotic silencing proteins is unknown. This is especially true for the nuclear suppressors and for the nuclear events that precede and lead to silencing. Therefore, I was interested in using biochemical approaches that allow me to define protein interactions and functions of suppressors. My ultimate goal was to gain insights into the molecular process of meiotic silencing.

A preliminary step in elucidating protein function is to determine interacting partners that could also be the biochemical targets. I sought to determine the function of the suppressor SMS-5, a SET domain-containing protein. To that end, I used affinity purification and pull-down assays. The combination of both assays allowed me to identify protein-binding partners for

SMS-5, to reveal new molecular components involved in the silencing process of unpaired DNA and to propose a biochemical function for SMS-5.

SMS-5 is a SET–domain containing protein, suggesting that SMS-5 is likely a protein methyltransferase (MTase). A preliminary test found no evidence to support that SMS-5 has a histone MTase (HMTase) activity. This result raised the possibility that SMS-5 functions as a non-histone protein MTase. In order to test that possibility it was critical to know binding partners for SMS-5, which could also be SMS-5's biochemical targets.

I found two non-histone proteins that physically interact with SMS-5; these proteins are PAF400 and Pianissimo. This result suggests that SMS-5 could act as a non-histone protein MTase. Protein methylation is an important post-translational modification involved in several biological process, including signaling, RNA processing and transport, transcription, and DNA repair (PAHLICH *et al.* 2006; PAIK *et al.* 2007). Protein methylation, especially at arginine residues, is known to promote or inhibit protein interactions (LEE *et al.* 2005a). Therefore, I hypothesize that SMS-5 is required for methylation of PAF400 and Pianissimo and this post-translational modification is a critical step in the nuclear events of meiotic silencing.

Now that targets for SMS-5 have been determined, it would be possible to test SMS-5 biochemical function and establish whether SMS-5 has MTase activity when interacting with PAF400 and/or Pianissimo. To that end, methyltransferase activity of purified SMS-5 should be tested in the presence of

purified PAF400 and Pianissimo. In this research, I confirmed that both proteins have predicted methylation sites. Therefore, the assay must include proteins carrying mutations on those sites.

At the end of Chapter II, I present a model summarizing the participation of PAF400 and Pianissimo in meiotic silencing. Following is a detail description of how each of these proteins is involved in silencing, specifically in meiotic *trans*-sensing, and the connection between nuclear and perinuclear events. I also propose experimental approaches in order to test the model.

A Chromatin Remodeling Pathway May Participate in Unpaired DNA Recognition

One of the protein-binding partners for SMS-5, PAF400 protein, is member of the Tra1/TRRAP family of proteins. PAF400 is a large polypeptide associated with histone acetyltransferase (HAT) complexes, such as SAGA and NuA4 (KNUTSON and HAHN 2011). These complexes are required for chromatin remodeling by histone acetylation. Once chromatin acquires a relaxed conformation, the SAGA complex is involved in transcriptional activation (GRANT *et al.* 1998), and the NuA4 complex is associated with DNA damage repair (MURR *et al.* 2006). In both cases the scaffolding action of Tra1/TRRAP is essential for the recruitment of the different components to the chromatin (MURR *et al.* 2007).

Although I was not able to test directly the involvement of PAF400 in meiotic silencing because it is an essential protein, the physical interaction of PAF400 with SMS-5 provides indirect evidence that PAF400 participates in this silencing phenomenon. Because of PAF400 is a nuclear protein and its homologous proteins participate in chromatin remodeling and DNA-involved molecular processes, I hypothesize that PAF400 is an important component for meiotic *trans*-sensing by participating in the unpaired DNA detection and/or transcription of aberrant RNA (aRNA) molecules from the unpaired DNA region.

Model for the Involvement of PAF400 in the Detection of Unpaired DNA Regions and the Transcription of aRNAs

In *Neurospora*, unpaired DNA regions are detected early in meiosis, which triggers meiotic silencing activation. It has been proposed that detection of unpaired regions occurs through a meiotic *trans*-sensing mechanism (ARAMAYO and METZENBERG 1996), after homologous chromosomes are paired. However, the molecular mechanism behind meiotic *trans*-sensing is still unknown.

Pratt made the first attempt to determine the regulation of meiotic *trans*-sensing by testing the connection between *trans*-sensing and homologous recombination (PRATT 2008). In that study, the efficiency of meiotic silencing was measured in the absence of SPO11, which is responsible for DNA double-strand breaks (DSBs) formation and initiation of recombination. The results showed that SPO-11-dependent DNA DSBs were dispensable for meiotic silencing.

Therefore, it was concluded that recombination/synapsis and silencing were not related in *Neurospora*, (PRATT 2008). Neverthelees, it was recently reported that there is a SPO11-independent mechanism for initiation of recombination in *Neurospora*, suggesting that SPO11-independent DSBs may be also present during meiosis. (BOWRING *et al.* 2013). Accordingly, "spontaneous" DSBs and SPO11-independent DNA repair foci were observed in mammalian meiocytes (CAROFIGLIO *et al.* 2013). Together these findings raise a key question: Are SPO11-independent DNA DSBs required for unpaired DNA recognition and therefore activation of meiotic silencing?

An interesting connection between DNA DSBs repair and PAF400 homologous proteins has been observed. Homologs of PAF400, Tra1 and TRRAP are involved in the repair of DNA DSBs. Histone H4 acetylation by TRRAP-containing complex NuA4 is critical for non-homologous end-joining (NHEJ) repair of DSBs (BIRD *et al.* 2002). It has been shown that TRRAP modulates loading of MRN repair complex to the chromatin surrounding sites of DNA DSBs (MURR *et al.* 2006; ROBERT *et al.* 2006). All together, these observations lead me to propose that PAF400 is an essential component that participates in the recognition of unpaired DNA regions through detection of unrepaired DSBs

I propose the following model for the recognition of unpaired DNA region in early meiosis in *N. crassa*. After fertilization, the two nuclei of opposite mating type fuse forming a transient diploid cell that starts meiosis division. During the

leptotene stage of early meiosis, homologous chromosomes come into proximity, pair and sense. The topoisomerase SPO11 initiates the formation of DNA DSBs that activates homologous recombination. In addition to the SPO11-dependent DSBs, it is possible that SPO11-independent DNA damage is also present at this stage. It has been proposed that these spontaneous DNA DSBs are persistent breaks formed during DNA-replication in the S-phase (CAROFIGLIO *et al.* 2013). Another possibility is that they are formed by an unknown enzyme also early in meiosis. Both SPO11-dependent and -independent DSBs are processed and repaired via homologous recombination. However, due to lack of homology, DSBs localized in the unpaired DNA region remain unrepaired. Once a threshold of unrepaired DSBs is reached, activation of meiotic silencing initiates and PAF400-containing complex (NuA4-like complex) is recruited to the unpaired DNA region. It is known that the length of the unpaired region is proportional to the efficiency of silencing (Chapter I). Therefore, the larger the unpaired DNA region, the more extended the accumulation of unrepaired DSBs, and the more efficient the silencing. The histone acetylase activity of NuA4 complex mediates chromatin remodeling at the proximities of the DSBs by histone H4 acetylation. This relaxed form of chromatin facilitates the interaction of the unpaired DNA with other proteins, including the components of the MRN repair complex, which are responsible for DSBs repair by the NHEJ mechanism (Figure 4.1).

★ SPO11-independent DSBs ★ SPO11-dependent DSBs
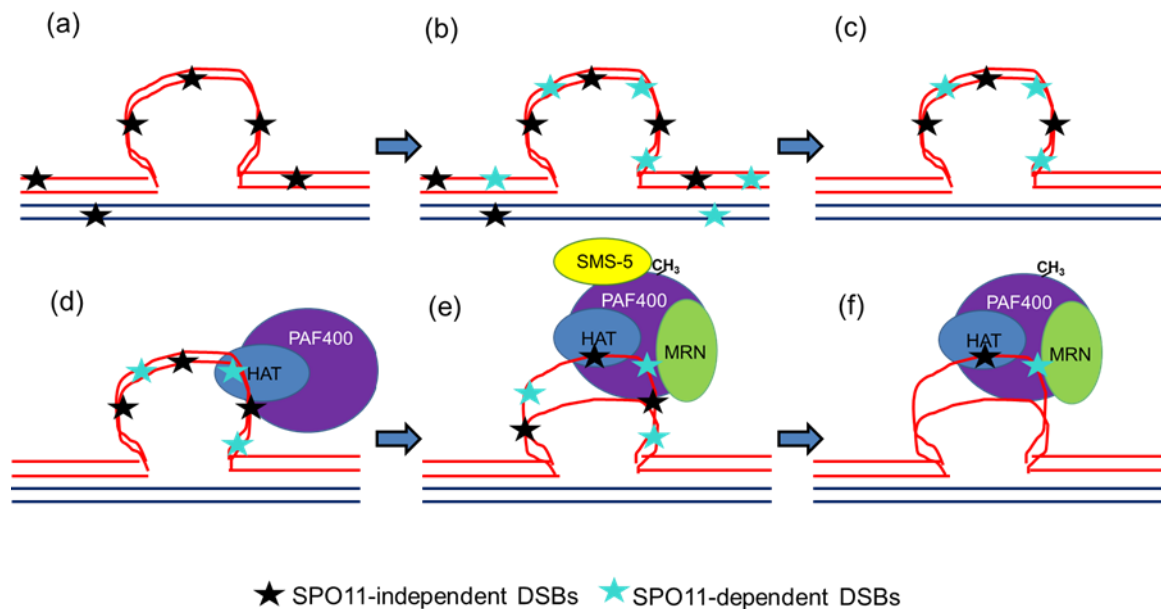
Figure 4.1 Proposed model for PAF400 involvement in the unpaired DNA recognition in *N. crassa* early meiosis. Interactions between one pair of homologous chromosomes (red and blue) are schematically represented. Sister chromatids, which are products of pre-meiotic DNA replication, are shown as double red or blue lines. SPO11-independent DNA DSBs (black star) may be persistent breaks formed during DNA replication. Unpaired DNA region (red bubble) is formed as consequence of gene insertion in the red chromosome or deletion in the blue chromosome. (a) After karyogamy homologous chromosomes pair and sense. (b) SPO11-dependent DNA DSBs are formed (blue stars) and recombination starts. (c) DSBs are repaired via homologous recombination in the paired DNA regions. However, DSBs remain unrepaired and accumulate in the unpaired DNA region. Accumulation of DSBs may be the signal that activates meiotic silencing nuclear phase. (d) Reaching a threshold of unrepaired DSBs stimulates the recruitment of NuA4-like complex to the unpaired DNA region. NuA4-like complex, formed by PAF400 and a histone acetyltransferase (HAT), binds to the chromatin surrounding sites of DSBs and initiates chromatin remodeling. (e) Chromatin relaxation and the presence of PAF400 facilitate the recruitment of other proteins to the unpaired DNA region, such as the components of the MRN repair complex. SMS-5 may be involved in regulating protein interactions by adding methyl group(s) to PAF400. (f) MRN complex is responsible for DNA damage repair through non-homologous end-joining (NHEJ) process.

192

SMS-5 methyltransferase activity may be required for regulating PAF400 protein interactions. PAF400 is a very large polypeptide that acts as a scaffold protein at the proximity of unrepaired DSBs. Switching between interactions with NuA4 components to interactions with MRN components may need special post-translational modifications like methylation (Figure 4.1).

PAF400 homologs are also involved in transcriptional activation when bound to the SAGA complex (MUTIU *et al.* 2007). SAGA, another HAT complex, is recruited to promoter regions where it is required for relaxing chromatin structure by acetylation of histone H3. Therefore, I propose that chromatin relaxation and the presence of PAF400 homologous protein facilitate the recruitment of the transcriptional machinery (KOUTELOU *et al.* 2010).

I hypothesize that PAF400 is also involved in the transcription of the aRNA molecule, which is then translocated to the perinuclear compartment and mediates the actual silencing of mRNA transcripts. I propose that the scaffolding activity of PAF400 may modulate the recruitment of components necessary for aRNA synthesis. Because the unpaired DNA would serve as the template for the synthesis of aRNA, it is possible that the transcription process starts after DNA is repaired (Figure 4.2). Methylation of PAF400 by SMS-5 could be also required for regulating protein interactions with the transcriptional machinery.
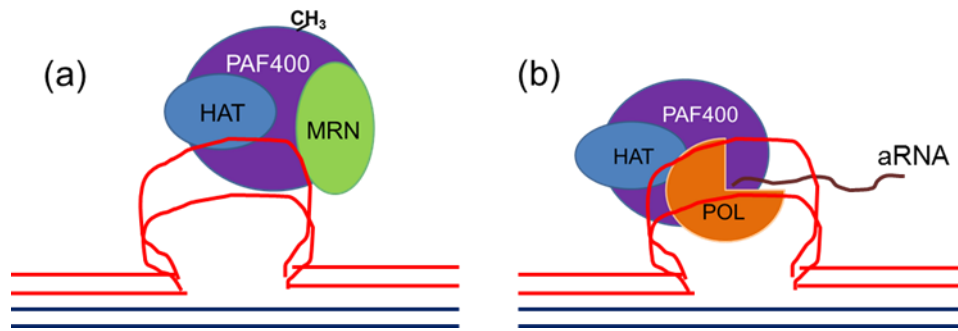
Figure 4.2 Proposed model for PAF400 involvement in the biosynthesis of aberrant RNAs. Interactions between one pair of homologous chromosomes (red and blue) are schematically represented. Sister chromatids, which are products of pre-meiotic DNA replication, are shown as double red or blue lines. Unpaired DNA region (red bubble). (a) Once DNA DSBs are repair by MRN complex, the complex dissociates from PAF400 facilitating the recruitment of new molecular components. At this point, unpaired DNA is ready to serve as template for transcription of aRNAs. (b) Proteins required for transcription of aRNAs are recruited by PAF400 to the unpaired DNA. aRNA is synthesized and then exported to the perinuclear region where it is incorporated into the silencing machinery.

Interestingly, recent observations have revealed that the nuclear DNA/RNA-dependent RNA polymerase QDE-1 has an effect on the meiotic silencing response of homeology regions (Millimaki and Aramayo, unpublished data). In *Neurospora* vegetative cells, QDE-1 is required for the generation of aRNAs and dsRNA, which are essential for silencing transgenes via quelling (LEE *et al.* 2010b). It has been demonstrated that QDE-1 needs to interact with replication protein A (RPA) and DNA helicase QDE-3 to produce dsRNAs (NOLAN *et al.* 2008). RPA and QDE-3 proteins participate in DNA damage and stress responses (COGONI and MACINO 1999b). Excitingly, in my work QDE-1

and QDE-3 were pulled-down once by GST-SMS-5 fusion protein (Appendix A; Table A.1). Together, these data suggest that QDE-1-QDE-3 may be involved in the nuclear stage of meiotic silencing. However, several studies must be conducted to gain new insights into its molecular function.

PAF400 may be involved in detection of unpaired DNA and transcription of aRNAs by mediating the recruitment of different complexes. Finding an interaction between the suppressor SMS-5 and the scaffolding protein PAF400 is an important step towards understanding the molecular mechanism of meiotic silencing. This protein interaction represents the connection between chromatin remodeling, DNA damage repair, and meiotic silencing mechanism. However, to test this proposed model, additional studies are required and several questions need to be answered. Is PAF400 methylated by SMS-5? Is this post-translational modification required for meiotic silencing? Is PAF400 an essential component of the NuA4-like and SAGA-like complexes in *N. crassa*? Is NuA4-like complex required for DNA DSBs repaired at the unpaired DNA? Is SAGA-like complex required for activation of transcription of aRNAs?

In order to answer those questions, the first step is to determine whether PAF400 is methylated by SMS-5. To do this, an *in vitro* protein MTase assay needs to be conducted to both purified proteins. In this study, an analysis that predicts methylation sites determined that PAF400 has 8 arginine residues that are likely to be methylated. Therefore, PAF400 mutant proteins carrying single-

amino-acid substitution mutations in each of those residues need to be purified and included in the MTase assay.

The next step would be to test whether PAF400 methylation by SMS-5 is required for meiotic silencing. PAF400 is an essential protein, indicating that its biological function is required for other cellular process in addition to meiotic silencing. For that reason, it would be necessary to determine mutations that impair methylation by SMS-5 without compromising cellular viability. Finding these mutations would allow us to test directly the involvement of PAF400 in meiotic silencing. To conduct this experiment, *N. crassa* strains carrying mutations in previously found methylation sites need to be constructed. PAF400 mutant strains would be tested for cellular viability, and viable mutant strains would be used for analyzing meiotic silencing efficiency.

It is possible to get negative results from MTase assay, indicating that SMS-5 does not methylate PAF400. A negative result would suggest that SMS-5 has a different function, such as mediating protein interactions between PAF400 and other components. If this were the case, it would be essential to identify PAF400 protein regions that are required for binding SMS-5 and vice versa. That could be accomplished by generating a series of internal deletions spanning the entire length of PAF400 and SMS-5 testing them for impaired protein binding. After finding mutations that disrupt PAF400-SMS-5 interaction, meiotic silencing efficiency can be measured. A caveat is that meiotic silencing can be tested only if mutant strains are viable.

Once the nature of PAF400 and SMS-5 interaction is established, it would be useful to determine other protein-binding partners for PAF400, particularly the catalytic subunits of the HAT complexes. In the models presented in Figures 4.1 and 4.2, I propose that PAF400 binds a HAT protein responsible for histone acetylation and chromatin remodeling. I also propose that PAF400 could be associated with components of MRN complex which will be responsible for DNA damage repaired (ROBERT *et al.* 2006). An alternative possibility is that PAF400 interacts with proteins required for transcription, as it has been reported for SAGA complex in other organisms (RODRIGUEZ-NAVARRO 2009; KOUTELOU *et al.* 2010; HELMLINGER *et al.* 2011).

To test all these possibilities, it is necessary to determine the protein complexes in which PAF400 is a critical component. To do this, *in vivo* immunoprecipitation (IP) and immunoblotting assays in conjunction to mass spectrometry analysis must be performed. However, the PAF400 interaction with different proteins may occur sequentially during meiosis. For that reason, IP needs to be performed at different time points in the course of meiosis. By determining other protein-binding partners for PAF400, it would be possible to establish if NuA4-like and/or SAGA-like complexes are related to meiotic silencing.

Once the catalytic subunit that interacts with PAF400 is established, it would be necessary to determine whether the catalytic activity of this protein is required for meiotic silencing. Therefore, point mutations that disrupt the

197

catalytic site need to be constructed, and then meiotic silencing would be analyzed in the presence of these mutants.

My model proposes that PAF400 is recruited to the unpaired DNA region after a threshold of DNA DSBs is reached. Therefore, a direct interaction between the unpaired DNA and PAF400 is expected. Using chromatin immunoprecipitation (ChIP) assays, it would be possible to test this hypothesis.

A Signaling Transduction Pathway May Connect the Nuclear and the Perinuclear Phases of Meiotic Silencing

I identified a second protein-binding partner for SMS-5, Pianissimo. Pianissimo is a homolog of Aveo3/RICTOR proteins. Pianissimo homologous proteins are essential components of the TORC2 complex, which is involved in several signaling transduction pathways in numerous organisms (CYBULSKI and HALL 2009). Pianissimo homologs have also been found associated to other proteins, supporting the notion that it could be mediated other functions outside of TORC2 (Oh 2011). However, its function still remains elusive.

Although I was not been able to test directly the involvement of Pianissimo in meiotic silencing, its physical interaction with SMS-5 provides indirect evidence for its participation. Finding this interaction between SMS-5 and Pianissimo suggests that signaling transduction pathways could be involved in the regulation of meiotic silencing. Interestingly, Pianissimo homologs have been detected in both the cytosol and in the nucleus. Computational analysis

predicts that Pianissimo could be also localized in both cellular compartments. This observation suggests that Pianissimo protein and its homologs may be shuttling proteins continuously back and forth between the nucleus and the cytoplasm. I hypothesize that Pianissimo protein is involved in the activation of the meiotic silencing machinery in the perinucleus by activating a signal transduction cascade once unpaired DNA is detected and aRNAs synthesized in the nucleus.

It would be of interest to explore whether Pianissimo participates in meiotic silencing as a component of the TORC2 complex or whether Pianissimo and PAF400 are components of a different and new multi-protein complex together with SMS-5.

Model for the Involvement of Pianissimo in Meiotic Silencing

Pianissimo protein possesses an HR1 domain—Rho binding domain—at the N-terminus, and a REM domain—Ras exchange motif—at the C-terminus. The presence of these two domains indicates that Pianissimo is an effector protein that is regulated by small G proteins (molecular switches) (HALL 2012). On the other hand, Pianissimo homologs bind directly with protein kinase TOR, and together, they are the core components of TORC2 complex. This complex is responsible for phosphorylation of several substrates, including AGC protein kinases, and participates in signal transduction cascades (CYBULSKI and HALL 2009). TOR protein kinase is a member of the phosphatidylinositol-3-kinase-

199

related kinases (PIKK), which function as serine/threonine protein kinases. Interestingly, PAF400–the other protein-binding partner for SMS-5––is also a PIKK protein (Chapter II). Because of the high structural similarities observed among PIKK members (Figure 2.9), it is possible that Pianissimo interacts with PAF400. Therefore, I hypothesize that Pianissimo binds to PAF400 during the nuclear events of meiotic silencing. Together all these observations, lead me to hypothesize that Pianissimo is responsible for coupling the nuclear signal generated after recognition of unpaired DNA to a signaling cascade that triggers meiotic silencing in the perinuclear region.

The model I propose for the function of Pianissimo in meiotic silencing is as follows: PAF400 and Pianissimo proteins interact in the proximity of the unpaired DNA while aRNA is still synthesized. Methylation of PAF400 and/or Pianissimo by SMS-5 modulates this protein interaction. Interaction between PAF400 and Pianissimo represents the stimulus that activates the signaling cascade; then the signaling cascade in conjunction with the aRNA trigger meiotic silencing in the perinuclear region. A small G protein may mediate activation of the effector protein Pianissimo by binding to GTP and release of GDP. Pianissimo-GTP-binding form recruits and interacts with a protein kinase (TOR-like protein). This protein kinase is responsible for the phosphorylation of downstream effector proteins, and the generation of the phosphorylation signaling cascade that goes from the nucleus to the perinucleus. Finally, the signal is received in the perinuclear region and the silencing machinery is

activated, ready to couple the aRNA that was exported from the nucleus (Figure 4.3).

The interaction between SMS-5, Pianissimo and PAF400 represents the connection between meiotic silencing and signaling transduction pathways. Finding Pianissimo as a possible component of meiotic silencing creates new possibilities to explain the molecular connection that must exist between meiotic *trans*-sensing in the nucleus and activation of silencing in the perinucleus. It would be interesting to explore whether a signaling cascade mediated by kinases is responsible for activation of the meiotic silencing complex. In order to do this, additional studies are necessary.

The model presented in Figure 4.3 gives an attractive explanation of how a nuclear event triggers a perinuclear action. However, several gaps in the model need to be filled. The direct binding between SMS-5 and Pianissimo suggests that Pianissimo is a possible target protein for the biochemical action of SMS-5, which is predicted to be methylation. Pianissimo has three predicted methylation sites. It would be important to test whether those sites are methylation targets for SMS-5. Therefore, a protein MTase *in vitro* assay is recommended.
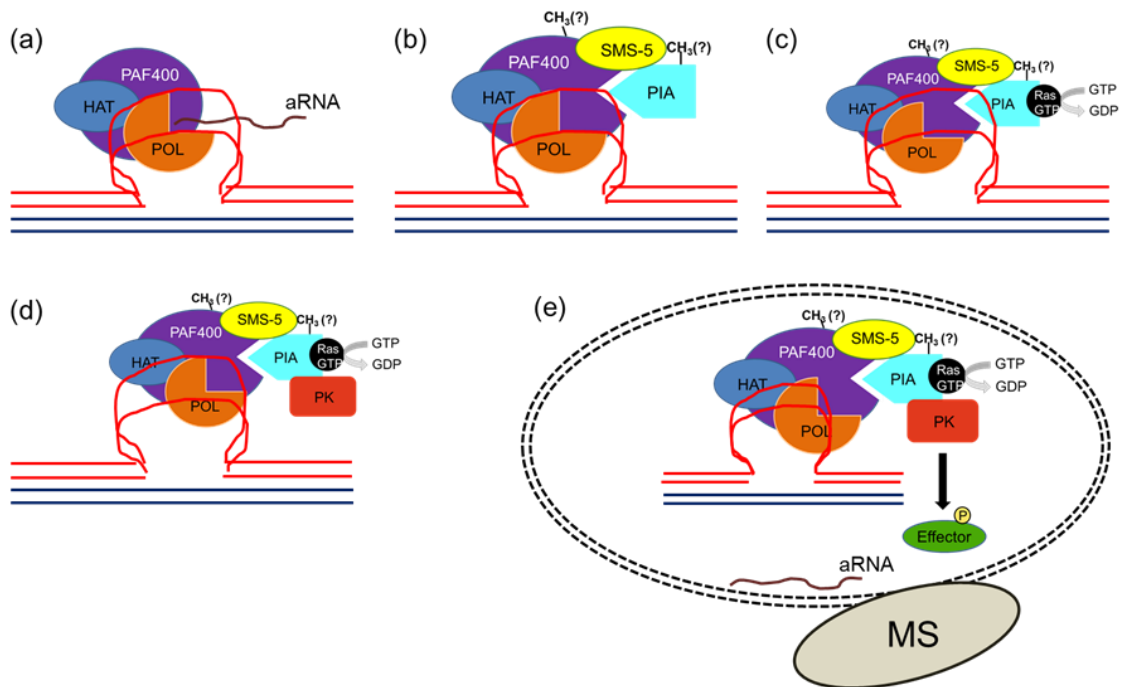
Figure 4.3 Proposed model for Pianissimo function in meiotic silencing as an activator of a signaling cascade. Interactions between one pair of homologous chromosomes (red and blue) are schematically represented. Sister chromatids, which are products of pre-meiotic DNA replication, are shown as double red or blue lines. Unpaired DNA region (red bubble). (a) Unpaired DNA is detected and by the scaffolding action of PAF400, chromatin is relaxed, DSBs are repair, and aRNA is synthesized (Figure 4.1 and 4.2). (b) Interaction between PAF400 and Pianissimo is observed while PAF400 is still associated with the unpaired DNA. Methylation (CH_3) of PAF400 and/or Pianissimo by SMS-5 may modulate this protein interaction. Interaction between PAF400 and Pianissimo represents the stimulus that activates the signal cascade. (c) Ras protein may mediate activation of the effector protein Pianissimo by binding to GTP and release of GDP. (d) Pianissimo-RasGTP-binding form binds to a protein kinase (PK) (TOR-like protein). (e) This protein kinase is responsible for the phosphorylation (P) of downstream effector proteins, and the generation of the phosphorylation signaling cascade (Black vertical arrow). Finally, the signal and the aRNA are received by the meiotic silencing (MS) components in the perinuclear region, and silencing is activated.

Once the Pianissimo region required for SMS-5 interaction is established, additional experiments could be executed. For example, it would be possible to analyze meiotic silencing efficiency when the interaction between SMS-5 and Pianissimo is compromised. Because Pianissimo is an essential protein, it is not possible to perform experiments that involve Pianissimo deletion mutant *in vivo*. It would be ideal to find a Pianissimo separation of function mutant in which all biological functions of the protein remain intact, except for the interaction with SMS-5. If the mutant with this characteristic is identified, it would be possible to find whether Pianissimo is a direct component of meiotic silencing.

Pianissimo homologs interact with a protein kinase that belongs to the PIKK family in several organisms. Therefore, it is reasonable to hypothesize that Pianissimo interacts with PAF400, another PIKK protein. It would be important to find whether Pianissimo interacts directly with PAF400. To do this, *in vitro* protein synthesis and immunoprecipitation assays could be performed.

Assuming that Pianissimo and PAF400 directly interact, it would be also necessary to establish whether that interaction is regulated by SMS-5. For that purpose, an *in vitro* MTassay in combination with co-IP experiments could be designed. Alternatively, an *in vivo* experiment could be done by using cell extracts from wild-type and *Sms-5* mutant for immunoprecipitation analysis. If SMS-5 is required for the interaction of Pianissimo and PAF400, it is expected that both co-immunoprecipitated from the wild-type cellular extract, but not from SMS-5-depleted cell extract.

Determining other protein-binding partners for Pianissimo will also be informative. For example, Pianissimo homologs interact with the catalytic subunit of TORC2 complex, the protein kinase TOR. It is worth mentioning that *N. crassa* TOR kinase was also recovered from my pull-down assays; however, it was also present in the control GST column. This could be an example of a false negative result. Establishing whether Pianissimo binds to a protein kinase would have significant implications for meiotic silencing regulation. Once the interacting partners of Pianissimo are determined, additional experiments to test the biochemical actions of those proteins could be performed.

Testing whether Pianissimo is an effector protein regulated by a small G protein would help to more firmly establish the link to signal transduction pathway. Computational evidence suggests that this is the case. HR1 domain and REM domain have been predicted to be part of Pianissimo; both domains are associated with regulation by small G proteins. To test this possibility, a GTPase activation assay could be implemented.

Proteome of Sexual Development and Meiotic Silencing

The proteome of sexual development was constructed by analyzing the protein representation of two critical points of sexual growth: maternal tissue before fertilization and four days after fertilization. One of the purposes of constructing a proteome database is to determine the protein population present under a specific condition or at a particular time point in development. Functional

analysis of the proteins can provide new insights into the cellular and molecular processes that occur under those specific conditions.

From the proteome of sexual development, I expected to extract more information related to protein accumulation of suppressors of meiotic silencing. However, the Argonaute-like protein SMS-2 was the only suppressor identified. This result reflects the transient nature of the proteins involved in meiotic silencing. That is, suppressors are synthesized in a small period of time, which may accumulate only very early in meiosis. The lack representation of suppressors of meiotic silencing in the proteome could be also due to low-abundance. Therefore, technical limitations are responsible for this result.

In order to improve proteome and peptide coverage, it would be necessary to address different experimental aspects. First, more developmental time points need to be included in the analysis. For example, biological samples just after fertilization and after karyogamy are needed because it is early after karyogamy that unpaired DNA regions are detected and meiotic silencing is activated. Therefore, it is expected that the silencing machinery would be active at those points. In addition, high proteome and peptide coverage would be achieved by combining different strategies. That is, combination of sample fractionation, high resolution mass spectrometry analysis and combined databases search for protein identification.

REFERENCES

Alexander, W. G., N. B. Raju, H. Xiao, T. M. Hammond, T. D. Perdue *et al.*, 2008 DCL-1 colocalizes with other components of the MSUD machinery and is required for silencing. Fungal Genet Biol 45**:** 719-727.

Allard, S., R. T. Utley, J. Savard, A. Clarke, P. Grant *et al.*, 1999 NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. EMBO J 18**:** 5108-5119.

Aramayo, R., and R. L. Metzenberg, 1996 Meiotic transvection in fungi. Cell 86**:** 103-113.

Aramayo, R., Y. Peleg, R. Addison and R. Metzenberg, 1996 Asm-1+, a Neurospora crassa gene related to transcriptional regulators of fungal development. Genetics 144**:** 991-1003.

Aramayo, R., and R. J. Pratt, 2010 Meiotic *trans*-sensing and silencing in *Neurospora*, pp. 132-144 in *Cellular and Molecular Biology of the Filamentous Fungi*, edited by K. A. Borkovich and D. J. Ebole. ASM Press, Washington, DC.

Aramayo, R., and E. U. Selker, 2013 Neurospora crassa, a model system for epigenetics research. Cold Spring Harb Perspect Biol 5**:** a017921.

Ashley, T., A. P. Gaeth, L. B. Creemers, A. M. Hack and D. G. de Rooij, 2004 Correlation of meiotic events in testis sections and microspreads of mouse spermatocytes relative to the mid-pachytene checkpoint. Chromosoma 113**:** 126-136.

Baarends, W. M., E. Wassenaar, R. van der Laan, J. Hoogerbrugge, E. Sleddens-Linkels *et al.*, 2005 Silencing of unpaired chromatin and histone H2A ubiquitination in mammalian meiosis. Mol Cell Biol 25**:** 1041-1053.

Bardiya, N., W. G. Alexander, T. D. Perdue, E. G. Barry, R. L. Metzenberg *et al.*, 2008 Characterization of interactions between and among components of the meiotic silencing by unpaired DNA machinery in Neurospora crassa using bimolecular fluorescence complementation. Genetics 178**:** 593-596.

Basu, R., and L. F. Zhang, 2011 X chromosome inactivation: a silence that needs to be broken. Genesis 49**:** 821-834.

206

Baudat, F., Y. Imai and B. de Massy, 2013 Meiotic recombination in mammals: localization and regulation. Nat Rev Genet 14**:** 794-806.

Baulcombe, D. C., 1996 RNA as a target and an initiator of post-transcriptional gene silencing in transgenic plants. Plant Mol Biol 32**:** 79-88.

Bean, C. J., C. E. Schaner and W. G. Kelly, 2004 Meiotic pairing and imprinted X chromatin assembly in Caenorhabditis elegans. Nat Genet 36**:** 100-105.

Bentley, G. A., A. Lewit-Bentley, J. T. Finch, A. D. Podjarny and M. Roth, 1984 Crystal structure of the nucleosome core particle at 16 A resolution. J Mol Biol 176**:** 55-75.

Bernstein, B. E., A. Meissner and E. S. Lander, 2007 The mammalian epigenome. Cell 128**:** 669-681.

Bernstein, E., A. A. Caudy, S. M. Hammond and G. J. Hannon, 2001 Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409**:** 363-366.

Bhalla, N., and A. F. Dernburg, 2005 A conserved checkpoint monitors meiotic chromosome synapsis in Caenorhabditis elegans. Science 310**:** 1683-1686.

Bird, A. W., D. Y. Yu, M. G. Pray-Grant, Q. Qiu, K. E. Harmon *et al.*, 2002 Acetylation of histone H4 by Esa1 is required for DNA double-strand break repair. Nature 419**:** 411-415.

Bishop, A. L., and A. Hall, 2000 Rho GTPases and their effector proteins. Biochem J 348 Pt 2**:** 241-255.

Bochtler, M., L. Ditzel, M. Groll, C. Hartmann and R. Huber, 1999 The proteasome. Annu Rev Biophys Biomol Struct 28**:** 295-317.

Bonasio, R., S. Tu and D. Reinberg, 2010 Molecular signals of epigenetic states. Science 330**:** 612-616.

Borkovich, K. A., L. A. Alex, O. Yarden, M. Freitag, G. E. Turner *et al.*, 2004 Lessons from the genome sequence of Neurospora crassa: tracing the path from genomic blueprint to multicellular organism. Microbiol Mol Biol Rev 68**:** 1-108.

Borner, G. V., N. Kleckner and N. Hunter, 2004 Crossover/noncrossover differentiation, synaptonemal complex formation, and regulatory

surveillance at the leptotene/zygotene transition of meiosis. Cell 117**:** 29-45.

Bosotti, R., A. Isacchi and E. L. Sonnhammer, 2000 FAT: a novel domain in PIK-related kinases. Trends Biochem Sci 25**:** 225-227.

Bourc'his, D., and T. H. Bestor, 2004 Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature 431**:** 96-99.

Bowring, F. J., P. J. Yeadon and D. E. Catcheside, 2013 Residual recombination in Neurospora crassa spo11 deletion homozygotes occurs during meiosis. Mol Genet Genomics 288**:** 437-444.

Brown, C. E., L. Howe, K. Sousa, S. C. Alley, M. J. Carrozza *et al.*, 2001 Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit. Science 292**:** 2333-2337.

Bryan, T. M., K. J. Goodrich and T. R. Cech, 2000 A mutant of Tetrahymena telomerase reverse transcriptase with increased processivity. J Biol Chem 275**:** 24199-24207.

Burgoyne, P. S., S. K. Mahadevaiah and J. M. Turner, 2009 The consequences of asynapsis for mammalian meiosis. Nat Rev Genet 10**:** 207-216.

Butler, M. J., R. B. Gardiner and A. W. Day, 2009 Melanin synthesis by Sclerotinia sclerotiorum. Mycologia 101**:** 296-304.

Cabrero, J., M. Teruel, F. D. Carmona, R. Jimenez and J. P. Camacho, 2007 Histone H3 lysine 9 acetylation pattern suggests that X and B chromosomes are silenced during entire male meiosis in a grasshopper. Cytogenet Genome Res 119**:** 135-142.

Calvo, A. M., R. A. Wilson, J. W. Bok and N. P. Keller, 2002 Relationship between secondary metabolism and fungal development. Microbiol Mol Biol Rev 66**:** 447-459.

Cantara, W. A., P. F. Crain, J. Rozenski, J. A. McCloskey, K. A. Harris *et al.*, 2011 The RNA Modification Database, RNAMDB: 2011 update. Nucleic Acids Res 39**:** D195-201.

Carofiglio, F., A. Inagaki, S. de Vries, E. Wassenaar, S. Schoenmakers *et al.*, 2013 SPO11-independent DNA repair foci and their role in meiotic silencing. PLoS Genet 9**:** e1003538.

Carthew, R. W., 2001 Gene silencing by double-stranded RNA. Curr Opin Cell Biol 13**:** 244-248.

Catalanotto, C., M. Pallotta, P. ReFalo, M. S. Sachs, L. Vayssie *et al.*, 2004 Redundancy of the two dicer genes in transgene-induced posttranscriptional gene silencing in Neurospora crassa. Mol Cell Biol 24**:** 2536-2545.

Celeste, A., S. Petersen, P. J. Romanienko, O. Fernandez-Capetillo, H. T. Chen *et al.*, 2002 Genomic instability in mice lacking histone H2AX. Science 296**:** 922-927.

Checchi, P. M., and J. Engebrecht, 2011 Heteromorphic sex chromosomes: navigating meiosis without a homologous partner. Mol Reprod Dev 78**:** 623-632.

Chen, M. Y., Y. Long and P. N. Devreotes, 1997 A novel cytosolic regulator, Pianissimo, is required for chemoattractant receptor and G protein-mediated activation of the 12 transmembrane domain adenylyl cyclase in Dictyostelium. Genes Dev 11**:** 3218-3231.

Chen, R. E., and J. Thorner, 2007 Function and regulation in MAPK signaling pathways: lessons learned from the yeast Saccharomyces cerevisiae. Biochim Biophys Acta 1773**:** 1311-1340.

Chuikov, S., J. K. Kurash, J. R. Wilson, B. Xiao, N. Justin *et al.*, 2004 Regulation of p53 activity through lysine methylation. Nature 432**:** 353-360.

Clery, A., M. Blatter and F. H. Allain, 2008 RNA recognition motifs: boring? Not quite. Curr Opin Struct Biol 18**:** 290-298.

Cogoni, C., and G. Macino, 1997 Isolation of quelling-defective (qde) mutants impaired in posttranscriptional transgene-induced gene silencing in Neurospora crassa. Proc Natl Acad Sci U S A 94**:** 10233-10238.

Cogoni, C., and G. Macino, 1999a Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. Nature 399**:** 166-169.

Cogoni, C., and G. Macino, 1999b Posttranscriptional gene silencing in Neurospora by a RecQ DNA helicase. Science 286**:** 2342-2344.

Cohen, P. E., S. E. Pollack and J. W. Pollard, 2006 Genetic analysis of chromosome pairing, recombination, and cell cycle control during first meiotic prophase in mammals. Endocr Rev 27**:** 398-426.

Cologna, S. M., W. K. Russell, P. J. Lim, G. Vigh and D. H. Russell, 2010 Combining isoelectric point-based fractionation, liquid chromatography and mass spectrometry to improve peptide detection and protein identification. J Am Soc Mass Spectrom 21**:** 1612-1619.

Colot, H. V., G. Park, G. E. Turner, C. Ringelberg, C. M. Crew *et al.*, 2006 A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. Proc Natl Acad Sci U S A 103**:** 10352-10357.

Conesa, A., and S. Gotz, 2008 Blast2GO: A comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics 2008**:** 619832.

Conesa, A., S. Gotz, J. M. Garcia-Gomez, J. Terol, M. Talon *et al.*, 2005 Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21**:** 3674-3676.

Costa, Y., R. M. Speed, P. Gautier, C. A. Semple, K. Maratou *et al.*, 2006 Mouse MAELSTROM: the link between meiotic silencing of unsynapsed chromatin and microRNA pathway? Hum Mol Genet 15**:** 2324-2334.

Cui, X., I. De Vivo, R. Slany, A. Miyamoto, R. Firestein *et al.*, 1998 Association of SET domain and myotubularin-related proteins modulates growth control. Nat Genet 18**:** 331-337.

Cybulski, N., and M. N. Hall, 2009 TOR complex 2: a signaling pathway of its own. Trends Biochem Sci 34**:** 620-627.

Czech, B., and G. J. Hannon, 2011 Small RNA sorting: matchmaking for Argonautes. Nat Rev Genet 12**:** 19-31.

Daubner, G. M., A. Clery and F. H. Allain, 2013 RRM-RNA recognition: NMR or crystallography...and new findings. Curr Opin Struct Biol 23**:** 100-108.

Davis, L., and G. R. Smith, 2001 Meiotic recombination and chromosome segregation in Schizosaccharomyces pombe. Proc Natl Acad Sci U S A 98**:** 8395-8402.

Davis, R. H., and D. D. Perkins, 2002 Timeline: Neurospora: a model of model microbes. Nat Rev Genet 3**:** 397-403.

Davis, R. H. A. D. S., F.J., 1970 Genetic and microbiological research techniques for *Neurospora crassa*. Meth. Enzymol 17**:** 79-143.

Del Rizzo, P. A., and R. C. Trievel, 2011 Substrate and product specificities of SET domain methyltransferases. Epigenetics 6**:** 1059-1067.

DeLange, A. M., and A. J. Griffiths, 1980 Meiosis in Neurospora crassa. I. The isolation of recessive mutants defective in the production of viable ascospores. Genetics 96**:** 367-378.

Dernburg, A. F., K. McDonald, G. Moulder, R. Barstead, M. Dresser *et al.*, 1998 Meiotic recombination in C. elegans initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. Cell 94**:** 387-398.

Dillon, S. C., X. Zhang, R. C. Trievel and X. Cheng, 2005 The SET-domain protein superfamily: protein lysine methyltransferases. Genome Biol 6**:** 227.

Dolan, P. L., D. O. Natvig and M. A. Nelson, 2000 Neurospora proteome 2000. Fungal Genet. Newsl. 47**:** 7-24.

Doyon, Y., and J. Cote, 2004 The highly conserved and multifunctional NuA4 HAT complex. Curr Opin Genet Dev 14**:** 147-154.

Eddy, E. M., 1975 Germ plasm and the differentiation of the germ cell line. Int Rev Cytol 43**:** 229-280.

Eisenman, H. C., and A. Casadevall, 2012 Synthesis and assembly of fungal melanin. Appl Microbiol Biotechnol 93**:** 931-940.

Engh, I., M. Nowrousian and U. Kuck, 2007 Regulation of melanin biosynthesis via the dihydroxynaphthalene pathway is dependent on sexual development in the ascomycete Sordaria macrospora. FEMS Microbiol Lett 275**:** 62-70.

Findley, S. D., M. Tamanaha, N. J. Clegg and H. Ruohola-Baker, 2003 Maelstrom, a Drosophila spindle-class gene, encodes a protein that colocalizes with Vasa and RDE1/AGO1 homolog, Aubergine, in nuage. Development 130**:** 859-871.

Flynn, P., H. Mellor, R. Palmer, G. Panayotou and P. J. Parker, 1998 Multiple interactions of PRK1 with RhoA. Functional assignment of the Hr1 repeat motif. J Biol Chem 273**:** 2698-2705.

Free, S. J., 2013 Fungal cell wall organization and biosynthesis. Adv Genet 81**:** 33-82.

Galagan, J. E., S. E. Calvo, K. A. Borkovich, E. U. Selker, N. D. Read *et al.*, 2003 The genome sequence of the filamentous fungus Neurospora crassa. Nature 422**:** 859-868.

Galagan, J. E., and E. U. Selker, 2004 RIP: the evolutionary cost of genome defense. Trends in Genetics 20**:** 417-423.

Galazka, J. M., C. Tian, W. T. Beeson, B. Martinez, N. L. Glass *et al.*, 2010 Cellodextrin transport in yeast for improved biofuel production. Science 330**:** 84-86.

Ghabrial, A., and T. Schupbach, 1999 Activation of a meiotic checkpoint regulates translation of Gurken during Drosophila oogenesis. Nat Cell Biol 1**:** 354-357.

Glass, N. L., J. Grotelueschen and R. L. Metzenberg, 1990 Neurospora crassa A mating-type region. Proc Natl Acad Sci U S A 87**:** 4912-4916.

Goldberg, T., T. Hamp and B. Rost, 2012 LocTree2 predicts localization for all domains of life. Bioinformatics 28**:** i458-i465.

Goodrich-Tanrikulu, M., K. Howe, A. Stafford and M. A. Nelson, 1998 Changes in fatty acid composition of Neurospora crassa accompany sexual development and ascospore germination. Microbiology 144 (Pt 7)**:** 1713-1720.

Gotz, S., J. M. Garcia-Gomez, J. Terol, T. D. Williams, S. H. Nagaraj *et al.*, 2008 High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res 36**:** 3420-3435.

Grant, P. A., D. Schieltz, M. G. Pray-Grant, J. R. Yates, 3rd and J. L. Workman, 1998 The ATM-related cofactor Tra1 is a component of the purified SAGA complex. Mol Cell 2**:** 863-867.

Greenbaum, D., C. Colangelo, K. Williams and M. Gerstein, 2003 Comparing protein abundance and mRNA expression levels on a genomic scale. Genome Biol 4**:** 117.

Guillon, H., and B. de Massy, 2002 An initiation site for meiotic crossing-over and gene conversion in the mouse. Nat Genet 32**:** 296-299.

Gunaratne, J., A. Schmidt, A. Quandt, S. P. Neo, O. S. Sarac *et al.*, 2013 Extensive mass spectrometry-based analysis of the fission yeast proteome: the Schizosaccharomyces pombe PeptideAtlas. Mol Cell Proteomics 12**:** 1741-1751.

Haider, S., and R. Pal, 2013 Integrated analysis of transcriptomic and proteomic data. Curr Genomics 14**:** 91-110.

Hall, A., 2012 Rho family GTPases. Biochem Soc Trans 40**:** 1378-1382.

Hammond, T. M., W. G. Spollen, L. M. Decker, S. M. Blake, G. K. Springer *et al.*, 2013a Identification of small RNAs associated with meiotic silencing by unpaired DNA. Genetics 194**:** 279-284.

Hammond, T. M., H. Xiao, E. C. Boone, L. M. Decker, S. A. Lee *et al.*, 2013b Novel proteins required for meiotic silencing by unpaired DNA and siRNA generation in Neurospora crassa. Genetics 194**:** 91-100.

Hammond, T. M., H. Xiao, E. C. Boone, T. D. Perdue, P. J. Pukkila *et al.*, 2011 SAD-3, a Putative Helicase Required for Meiotic Silencing by Unpaired DNA, Interacts with Other Components of the Silencing Machinery. G3 (Bethesda) 1**:** 369-376.

Hamoir, G., 1992 The discovery of meiosis by E. Van Beneden, a breakthrough in the morphological phase of heredity. Int J Dev Biol 36**:** 9-15.

Handel, M. A., 2004 The XY body: a specialized meiotic chromatin domain. Exp Cell Res 296**:** 57-63.

Harper, L., I. Golubovskaya and W. Z. Cande, 2004 A bouquet of chromosomes. J Cell Sci 117**:** 4025-4032.

He, C., 2010 Grand challenge commentary: RNA epigenetics? Nat Chem Biol 6**:** 863-865.

Helmlinger, D., S. Marguerat, J. Villen, D. L. Swaney, S. P. Gygi *et al.*, 2011 Tra1 has specific regulatory roles, rather than global functions, within the SAGA co-activator complex. EMBO J 30**:** 2843-2852.

Hilti, N., D. Baumann, A. M. Schweingruber, P. Bigler and M. E. Schweingruber, 1999 Gene ste20 controls amiloride sensitivity and fertility in Schizosaccharomyces pombe. Curr Genet 35**:** 585-592.

Ho, H. L., Y. S. Shiau and M. Y. Chen, 2005 Saccharomyces cerevisiae TSC11/AVO3 participates in regulating cell integrity and functionally interacts with components of the Tor2 complex. Curr Genet 47**:** 273-288.

Hochwagen, A., and A. Amon, 2006 Checking your breaks: surveillance mechanisms of meiotic recombination. Curr Biol 16**:** R217-228.

Homolka, D., P. Jansa and J. Forejt, 2012 Genetically enhanced asynapsis of autosomal chromatin promotes transcriptional dysregulation and meiotic failure. Chromosoma 121**:** 91-104.

Honda, S., and E. U. Selker, 2009 Tools for fungal proteomics: multifunctional neurospora vectors for gene replacement, protein expression and protein purification. Genetics 182**:** 11-23.

Horandl, E., 2009 A combinational theory for maintenance of sex. Heredity (Edinb) 103**:** 445-457.

Huang, S., G. Shao and L. Liu, 1998 The PR domain of the Rb-binding zinc finger protein RIZ1 is a protein binding interface and is related to the SET domain functioning in chromatin-mediated gene expression. J Biol Chem 273**:** 15933-15939.

Hussain, S., A. A. Sajini, S. Blanco, S. Dietmann, P. Lombard *et al.*, 2013a NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. Cell Rep 4**:** 255-261.

Hussain, S., F. Tuorto, S. Menon, S. Blanco, C. Cox *et al.*, 2013b The mouse cytosine-5 RNA methyltransferase NSun2 is a component of the chromatoid body and required for testis differentiation. Mol Cell Biol 33**:** 1561-1570.

Hynes, M. J., and S. L. Murray, 2010 ATP-citrate lyase is required for production of cytosolic acetyl coenzyme A and development in Aspergillus nidulans. Eukaryot Cell 9**:** 1039-1048.

Inagaki, A., S. Schoenmakers and W. M. Baarends, 2010 DNA double strand break repair, chromosome synapsis and transcriptional silencing in meiosis. Epigenetics 5**:** 255-266.

Ispolatov, I., A. Yuryev, I. Mazo and S. Maslov, 2005 Binding properties and evolution of homodimers in protein-protein interaction networks. Nucleic Acids Res 33**:** 3629-3635.

Jacinto, E., R. Loewith, A. Schmidt, S. Lin, M. A. Ruegg *et al.*, 2004 Mammalian TOR complex 2 controls the actin cytoskeleton and is rapamycin insensitive. Nat Cell Biol 6**:** 1122-1128.

Jacobson, E. S., 2000 Pathogenic roles for fungal melanins. Clin Microbiol Rev 13**:** 708-717.

214

Jenuwein, T., G. Laible, R. Dorn and G. Reuter, 1998 SET domain proteins modulate chromatin domains in eu- and heterochromatin. Cell Mol Life Sci 54**:** 80-93.

Jiang, D., H. W. Jarrett and W. E. Haskins, 2009 Methods for proteomic analysis of transcription factors. J Chromatogr A 1216**:** 6881-6889.

Johnson, T. E., 1978 Isolation and characterization of perithecial development mutants in neurospora. Genetics 88**:** 27-47.

Johnson, T. E., 1979 A neurospora mutation that arrests perithecial development as either male or female parent. Genetics 92**:** 1107-1120.

Jones, C. A., S. E. Greer-Phillips and K. A. Borkovich, 2007 The response regulator RRG-1 functions upstream of a mitogen-activated protein kinase pathway impacting asexual development, female fertility, osmotic stress, and fungicide resistance in Neurospora crassa. Mol Biol Cell 18**:** 2123-2136.

Jones, K. T., E. R. Greer, D. Pearce and K. Ashrafi, 2009 Rictor/TORC2 regulates Caenorhabditis elegans fat storage, body size, and development through sgk-1. PLoS Biol 7**:** e60.

Jorgensen, S., G. Schotta and C. S. Sorensen, 2013 Histone H4 lysine 20 methylation: key player in epigenetic regulation of genomic integrity. Nucleic Acids Res 41**:** 2797-2806.

Kanehisa, M., M. Araki, S. Goto, M. Hattori, M. Hirakawa *et al.*, 2008 KEGG for linking genomes to life and the environment. Nucleic Acids Res 36**:** D480-484.

Kasuga, T., J. P. Townsend, C. Tian, L. B. Gilbert, G. Mannhaupt *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**:** 6469-6485.

Kawamata, T., and Y. Tomari, 2010 Making RISC. Trends Biochem Sci 35**:** 368-376.

Keeney, S., 2001 Mechanism and control of meiotic recombination initiation. Curr Top Dev Biol 52**:** 1-53.

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

Keeping, A., D. Deabreu, M. Dibernardo and R. A. Collins, 2011 Gel-based mass spectrometric and computational approaches to the mitochondrial proteome of Neurospora. Fungal Genet Biol 48**:** 526-536.

Keller, N. P., G. Turner and J. W. Bennett, 2005 Fungal secondary metabolism - from biochemistry to genomics. Nat Rev Microbiol 3**:** 937-947.

Kellogg, D. R., and D. Moazed, 2002 Protein- and immunoaffinity purification of multiprotein complexes. Methods Enzymol 351**:** 172-183.

Kelly, W. G., 2006 Standing guard: Perinuclear localization of an RNA-dependent RNA polymerase. Proc Natl Acad Sci U S A 103**:** 2007-2008.

Kelly, W. G., and R. Aramayo, 2007 Meiotic silencing and the epigenetics of sex. Chromosome Res 15**:** 633-651.

Kim, H., R. L. Metzenberg and M. A. Nelson, 2002 Multiple functions of mfa-1, a putative pheromone precursor gene of Neurospora crassa. Eukaryot Cell 1**:** 987-999.

Kim, H., and M. A. Nelson, 2005 Molecular and functional analyses of poi-2, a novel gene highly expressed in sexual and perithecial tissues of Neurospora crassa. Eukaryot Cell 4**:** 900-910.

Knutson, B. A., and S. Hahn, 2011 Domains of Tra1 important for activator recruitment and transcription coactivator functions of SAGA and NuA4 complexes. Mol Cell Biol 31**:** 818-831.

Kosugi, S., M. Hasebe, M. Tomita and H. Yanagawa, 2009 Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. Proc Natl Acad Sci U S A 106**:** 10171-10176.

Kotaja, N., and P. Sassone-Corsi, 2007 The chromatoid body: a germ-cell-specific RNA-processing centre. Nat Rev Mol Cell Biol 8**:** 85-90.

Kouskouti, A., E. Scheer, A. Staub, L. Tora and I. Talianidis, 2004 Gene-specific modulation of TAF10 function by SET9-mediated methylation. Mol Cell 14**:** 175-182.

Koutelou, E., C. L. Hirsch and S. Y. Dent, 2010 Multiple faces of the SAGA complex. Curr Opin Cell Biol 22**:** 374-382.

Kouzarides, T., 2007 Chromatin modifications and their function. Cell 128**:** 693-705.

Kovarik, A., H. Van Houdt, A. Holy and A. Depicker, 2000 Drug-induced hypomethylation of a posttranscriptionally silenced transgene locus of tobacco leads to partial release of silencing. FEBS Lett 467**:** 47-51.

Kutil, B. L., K. Y. Seong and R. Aramayo, 2003 Unpaired genes do not silence their paired neighbors. Curr Genet 43**:** 425-432.

Lackner, D. H., M. W. Schmidt, S. Wu, D. A. Wolf and J. Bahler, 2012 Regulation of transcriptome, translation, and proteome in response to environmental stress in fission yeast. Genome Biol 13**:** R25.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan *et al.*, 2007 Clustal W and Clustal X version 2.0. Bioinformatics 23**:** 2947-2948.

Lee, D. W., R. Millimaki and R. Aramayo, 2010a QIP, a component of the vegetative RNA silencing pathway, is essential for meiosis and suppresses meiotic silencing in Neurospora crassa. Genetics 186**:** 127-133.

Lee, D. W., R. J. Pratt, M. McLaughlin and R. Aramayo, 2003 An argonaute-like protein is required for meiotic silencing. Genetics 164**:** 821-828.

Lee, D. W., K. Y. Seong, R. J. Pratt, K. Baker and R. Aramayo, 2004 Properties of unpaired DNA required for efficient silencing in Neurospora crassa. Genetics 167**:** 131-150.

Lee, D. Y., C. Teyssier, B. D. Strahl and M. R. Stallcup, 2005a Role of protein methylation in regulation of transcription. Endocr Rev 26**:** 147-170.

Lee, H. C., A. P. Aalto, Q. Yang, S. S. Chang, G. Huang *et al.*, 2010b The DNA/RNA-dependent RNA polymerase QDE-1 generates aberrant RNA and dsRNA for RNAi in a process requiring replication protein A and a DNA helicase. PLoS Biol 8**:** e1000496.

Lee, S., F. I. Comer, A. Sasaki, I. X. McLeod, Y. Duong *et al.*, 2005b TOR complex 2 integrates cell movement during chemotaxis and signal relay in Dictyostelium. Mol Biol Cell 16**:** 4572-4583.

Lerch, K., 1983 Neurospora tyrosinase: structural, spectroscopic and catalytic properties. Mol Cell Biochem 52**:** 125-138.

Li, D., P. Bobrowicz, H. H. Wilkinson and D. J. Ebbole, 2005 A mitogen-activated protein kinase pathway essential for mating and contributing to vegetative growth in Neurospora crassa. Genetics 170**:** 1091-1104.

217

Lichius, A., K. M. Lord, C. E. Jeffree, R. Oborny, P. Boonyarungsrit *et al.*, 2012 Importance of MAP kinases during protoperithecial morphogenesis in Neurospora crassa. PLoS One 7**:** e42565.

Lim, A. K., and T. Kai, 2007 Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster. Proc Natl Acad Sci U S A 104**:** 6714-6719.

Lim, P., R. North and G. Vigh, 2007 Rapid isoelectric trapping in a micropreparative-scale multicompartment electrolyzer. Electrophoresis 28**:** 1851-1859.

Loewith, R., E. Jacinto, S. Wullschleger, A. Lorberg, J. L. Crespo *et al.*, 2002 Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth control. Mol Cell 10**:** 457-468.

Lovejoy, C. A., and D. Cortez, 2009 Common mechanisms of PIKK regulation. DNA Repair (Amst) 8**:** 1004-1008.

MacQueen, A. J., and A. Hochwagen, 2011 Checkpoint mechanisms: the puppet masters of meiotic prophase. Trends Cell Biol 21**:** 393-400.

Mahadevaiah, S. K., D. Bourc'his, D. G. de Rooij, T. H. Bestor, J. M. Turner *et al.*, 2008 Extensive meiotic asynapsis in mice antagonises meiotic silencing of unsynapsed chromatin and consequently disrupts meiotic sex chromosome inactivation. J Cell Biol 182**:** 263-276.

Mahadevaiah, S. K., J. M. Turner, F. Baudat, E. P. Rogakou, P. de Boer *et al.*, 2001 Recombinational DNA double-strand breaks in mice precede synapsis. Nat Genet 27**:** 271-276.

Maiti, M., H. C. Lee and Y. Liu, 2007 QIP, a putative exonuclease, interacts with the Neurospora Argonaute protein and facilitates conversion of duplex siRNA into single strands. Genes Dev 21**:** 590-600.

Marcon, E., and P. B. Moens, 2005 The evolution of meiosis: recruitment and modification of somatic DNA-repair proteins. Bioessays 27**:** 795-808.

Mason, H. S., 1948 The chemistry of melanin; mechanism of the oxidation of dihydroxyphenylalanine by tyrosinase. J Biol Chem 172**:** 83-99.

Matsuo, T., Y. Kubo, Y. Watanabe and M. Yamamoto, 2003 Schizosaccharomyces pombe AGC family kinase Gad8p forms a conserved signaling module with TOR and PDK1-like kinases. EMBO J 22**:** 3073-3083.

McKee, B. D., and M. A. Handel, 1993 Sex chromosomes, recombination, and chromatin conformation. Chromosoma 102**:** 71-80.

McKim, K. S., B. L. Green-Marroquin, J. J. Sekelsky, G. Chin, C. Steinberg *et al.*, 1998 Meiotic synapsis in the absence of recombination. Science 279**:** 876-878.

McMahon, S. B., H. A. Van Buskirk, K. A. Dugan, T. D. Copeland and M. D. Cole, 1998 The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. Cell 94**:** 363-374.

Meiklejohn, C. D., and Y. Tao, 2010 Genetic conflict and sex chromosome evolution. Trends Ecol Evol 25**:** 215-223.

Meister, G., 2013 Argonaute proteins: functional insights and emerging roles. Nat Rev Genet 14**:** 447-459.

Mello, C. C., and D. Conte, Jr., 2004 Revealing the world of RNA interference. Nature 431**:** 338-342.

Monesi, V., 1965 Differential rate of ribonucleic acid synthesis in the autosomes and sex chromosomes during male meiosis in the mouse. Chromosoma 17**:** 11-21.

Morel, J. B., P. Mourrain, C. Beclin and H. Vaucheret, 2000 DNA methylation and chromatin structure affect transcriptional and post-transcriptional transgene silencing in Arabidopsis. Curr Biol 10**:** 1591-1594.

Murr, R., J. I. Loizou, Y. G. Yang, C. Cuenin, H. Li *et al.*, 2006 Histone acetylation by Trrap-Tip60 modulates loading of repair proteins and repair of DNA double-strand breaks. Nat Cell Biol 8**:** 91-99.

Murr, R., T. Vaissiere, C. Sawan, V. Shukla and Z. Herceg, 2007 Orchestration of chromatin-based processes: mind the TRRAP. Oncogene 26**:** 5358-5372.

Mutiu, A. I., S. M. Hoke, J. Genereaux, C. Hannam, K. MacKenzie *et al.*, 2007 Structure/function analysis of the phosphatidylinositol-3-kinase domain of yeast tra1. Genetics 177**:** 151-166.

Namekawa, S. H., and J. T. Lee, 2009 XY and ZW: is meiotic sex chromosome inactivation the rule in evolution? PLoS Genet 5**:** e1000493.

Namekawa, S. H., P. J. Park, L. F. Zhang, J. E. Shima, J. R. McCarrey *et al.*, 2006 Postmeiotic sex chromatin in the male germline of mice. Curr Biol 16**:** 660-667.

Namekawa, S. H., J. L. VandeBerg, J. R. McCarrey and J. T. Lee, 2007 Sex chromosome silencing in the marsupial male germ line. Proc Natl Acad Sci U S A 104**:** 9730-9735.

Naumova, A. K., S. Fayer, J. Leung, K. A. Boateng, R. D. Camerini-Otero *et al.*, 2013 Dynamics of response to asynapsis and meiotic silencing in spermatocytes from Robertsonian translocation carriers. PLoS One 8**:** e75970.

Nelson, M. A., S. Kang, E. L. Braun, M. E. Crawford, P. L. Dolan *et al.*, 1997a Expressed sequences from conidial, mycelial, and sexual stages of Neurospora crassa. Fungal Genet Biol 21**:** 348-363.

Nelson, M. A., S. T. Merino and R. L. Metzenberg, 1997b A putative rhamnogalacturonase required for sexual development of Neurospora crassa. Genetics 146**:** 531-540.

Nelson, M. A., and R. L. Metzenberg, 1992 Sexual development genes of Neurospora crassa. Genetics 132**:** 149-162.

Nolan, T., G. Cecere, C. Mancone, T. Alonzi, M. Tripodi *et al.*, 2008 The RNA-dependent RNA polymerase essential for post-transcriptional gene silencing in Neurospora crassa interacts with replication protein A. Nucleic Acids Res 36**:** 532-538.

Nowrousian, M., 2009 A novel polyketide biosynthesis gene cluster is involved in fruiting body morphogenesis in the filamentous fungi Sordaria macrospora and Neurospora crassa. Curr Genet 55**:** 185-198.

Nowrousian, M., U. Kuck, K. Loser and K. M. Weltring, 2000 The fungal acl1 and acl2 genes encode two polypeptides with homology to the N- and C-terminal parts of the animal ATP citrate lyase polypeptide. Curr Genet 37**:** 189-193.

Nowrousian, M., S. Masloff, S. Poggeler and U. Kuck, 1999 Cell differentiation during sexual development of the fungus Sordaria macrospora requires ATP citrate lyase activity. Mol Cell Biol 19**:** 450-460.

Nowrousian, M., M. Piotrowski and U. Kuck, 2007 Multiple layers of temporal and spatial control regulate accumulation of the fruiting body-specific

protein APP in Sordaria macrospora and Neurospora crassa. Fungal Genet Biol 44**:** 602-614.

O'Rourke, S. M., and I. Herskowitz, 2004 Unique and redundant roles for HOG MAPK pathway components as revealed by whole-genome expression analysis. Mol Biol Cell 15**:** 532-542.

Oh, W. J., and E. Jacinto, 2011 mTOR complex 2 signaling and functions. Cell Cycle 10**:** 2305-2316.

Okuda, S., T. Yamada, M. Hamajima, M. Itoh, T. Katayama *et al.*, 2008 KEGG Atlas mapping for global analysis of metabolic pathways. Nucleic Acids Res 36**:** W423-426.

Olson, B. J., and J. Markwell, 2007 Assays for determination of protein concentration. Curr Protoc Pharmacol Appendix 3**:** 3A.

Page, S. L., and R. S. Hawley, 2003 Chromosome choreography: the meiotic ballet. Science 301**:** 785-789.

Page, S. L., and R. S. Hawley, 2004 The genetics and molecular biology of the synaptonemal complex. Annu Rev Cell Dev Biol 20**:** 525-558.

Pahlich, S., R. P. Zakaryan and H. Gehring, 2006 Protein arginine methylation: Cellular functions and methods of analysis. Biochim Biophys Acta 1764**:** 1890-1903.

Paik, W. K., D. C. Paik and S. Kim, 2007 Historical review: the field of protein methylation. Trends Biochem Sci 32**:** 146-152.

Pal, A. K., D. U. Gajjar and A. R. Vasavada, 2014 DOPA and DHN pathway orchestrate melanin synthesis in Aspergillus species. Med Mycol 52**:** 10-18.

Pan, X., L. Zhai, R. Sun, X. Li and X. Zeng, 2005 INI1/hSNF5/BAF47 represses c-fos transcription via a histone deacetylase-dependent manner. Biochem Biophys Res Commun 337**:** 1052-1058.

Pane, A., K. Wehr and T. Schupbach, 2007 zucchini and squash encode two putative nucleases required for rasiRNA production in the Drosophila germline. Dev Cell 12**:** 851-862.

Park, G., S. Pan and K. A. Borkovich, 2008 Mitogen-activated protein kinase cascade required for regulation of development and secondary metabolism in Neurospora crassa. Eukaryot Cell 7**:** 2113-2122.

Park, G., J. A. Servin, G. E. Turner, L. Altamirano, H. V. Colot *et al.*, 2011 Global analysis of serine-threonine protein kinase genes in Neurospora crassa. Eukaryot Cell 10**:** 1553-1564.

Peraza, L., and W. Hansberg, 2002 Neurospora crassa catalases, singlet oxygen and cell differentiation. Biol Chem 383**:** 569-575.

Perkins, D. D., 2002 Neurospora perithecia: The first sighting. Fungal Genet. Newsl 49**:** 9-10.

Perkins, D. D., A. Radford, D. Newmeyer and M. Bjorkman, 1982 Chromosomal loci of Neurospora crassa. Microbiol Rev 46**:** 426-570.

Petronczki, M., M. F. Siomos and K. Nasmyth, 2003 Un menage a quatre: the molecular biology of chromosome segregation in meiosis. Cell 112**:** 423-440.

Phadnis, N., R. W. Hyppa and G. R. Smith, 2011 New and old ways to control meiotic recombination. Trends Genet 27**:** 411-421.

Phillips, C. M., A. T. Iavarone and M. A. Marletta, 2011 Quantitative proteomic approach for cellulose degradation by Neurospora crassa. J Proteome Res 10**:** 4177-4185.

Pratt, R. J., 2008 Meiotic *trans*-sensing and Meiotic Silencing in *Neurospora crassa*, pp. 1-183 in *Doctoral Dissertation*. Texas A&M University, College Station, TX.

Pratt, R. J., D. W. Lee and R. Aramayo, 2004 DNA methylation affects meiotic trans-sensing, not meiotic silencing, in Neurospora. Genetics 168**:** 1925-1935.

Raju, N. B., 1980 Meiosis and ascospore genesis in Neurospora. Eur J Cell Biol 23**:** 208-223.

Raju, N. B., 1992 Genetic-Control of the Sexual Cycle in Neurospora. Mycological Research 96**:** 241-262.

Raju, N. B., 2009 Neurospora as a model fungus for studies in cytogenetics and sexual biology at Stanford. J Biosci 34**:** 139-159.

Raju, N. B., and J. F. Leslie, 1992 Cytology of recessive sexual-phase mutants from wild strains of Neurospora crassa. Genome 35**:** 815-826.

Rayman, J. B., Y. Takahashi, V. B. Indjeian, J. H. Dannenberg, S. Catchpole *et al.*, 2002 E2F mediates cell cycle-dependent transcriptional repression in vivo by recruitment of an HDAC1/mSin3B corepressor complex. Genes Dev 16**:** 933-947.

Riha, K., J. M. Watson, J. Parkey and D. E. Shippen, 2002 Telomere length deregulation and enhanced sensitivity to genotoxic stress in Arabidopsis mutants deficient in Ku70. EMBO J 21**:** 2819-2826.

Rivals, I., L. Personnaz, L. Taing and M. C. Potier, 2007 Enrichment or depletion of a GO category within a class of genes: which test? Bioinformatics 23**:** 401-407.

Robert, F., S. Hardy, Z. Nagy, C. Baldeyron, R. Murr *et al.*, 2006 The transcriptional histone acetyltransferase cofactor TRRAP associates with the MRN repair complex and plays a role in DNA double-strand break repair. Mol Cell Biol 26**:** 402-412.

Rodriguez-Navarro, S., 2009 Insights into SAGA function during gene expression. EMBO Rep 10**:** 843-850.

Roeder, G. S., and J. M. Bailis, 2000 The pachytene checkpoint. Trends Genet 16**:** 395-403.

Romano, N., and G. Macino, 1992 Quelling: transient inactivation of gene expression in Neurospora crassa by transformation with homologous sequences. Mol Microbiol 6**:** 3343-3353.

Rosas-Acosta, G., W. K. Russell, A. Deyrieux, D. H. Russell and V. G. Wilson, 2005 A universal strategy for proteomic studies of SUMO and other ubiquitin-like modifiers. Mol Cell Proteomics 4**:** 56-72.

Rosner, M., and M. Hengstschlager, 2008 Cytoplasmic and nuclear distribution of the protein complexes mTORC1 and mTORC2: rapamycin triggers dephosphorylation and delocalization of the mTORC2 components rictor and sin1. Hum Mol Genet 17**:** 2934-2948.

Rost, B., G. Yachdav and J. Liu, 2004 The PredictProtein server. Nucleic Acids Res 32**:** W321-326.

Royo, H., G. Polikiewicz, S. K. Mahadevaiah, H. Prosser, M. Mitchell *et al.*, 2010 Evidence that meiotic sex chromosome inactivation is essential for male fertility. Curr Biol 20**:** 2117-2123.

Roze, L. V., A. Chanda, J. Wee, D. Awad and J. E. Linz, 2011 Stress-related transcription factor AtfB integrates secondary metabolism with oxidative stress response in aspergilli. J Biol Chem 286**:** 35137-35148.

Ruepp, A., A. Zollner, D. Maier, K. Albermann, J. Hani *et al.*, 2004 The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. Nucleic Acids Res 32**:** 5539-5545.

Sarbassov, D. D., S. M. Ali, D. H. Kim, D. A. Guertin, R. R. Latek *et al.*, 2004 Rictor, a novel binding partner of mTOR, defines a rapamycin-insensitive and raptor-independent pathway that regulates the cytoskeleton. Curr Biol 14**:** 1296-1302.

Schimenti, J., 2005 Synapsis or silence. Nat Genet 37**:** 11-13.

Schoenmakers, S., E. Wassenaar, J. W. Hoogerbrugge, J. S. Laven, J. A. Grootegoed *et al.*, 2009 Female meiotic sex chromosome inactivation in chicken. PLoS Genet 5**:** e1000466.

Schoenmakers, S., E. Wassenaar, W. A. van Cappellen, A. A. Derijck, P. de Boer *et al.*, 2008 Increased frequency of asynapsis and associated meiotic silencing of heterologous chromatin in the presence of irradiation-induced extra DNA double strand breaks. Dev Biol 317**:** 270-281.

Schonbrun, M., M. Kolesnikov, M. Kupiec and R. Weisman, 2013 TORC2 is required to maintain genome stability during S phase in fission yeast. J Biol Chem 288**:** 19649-19660.

Selker, E. U., N. A. Tountas, S. H. Cross, B. S. Margolin, J. G. Murphy *et al.*, 2003 The methylated component of the Neurospora crassa genome. Nature 422**:** 893-897.

Shi, S. P., J. D. Qiu, X. Y. Sun, S. B. Suo, S. Y. Huang *et al.*, 2012 PMeS: prediction of methylation sites based on enhanced feature encoding scheme. PLoS One 7**:** e38772.

Shiota, C., J. T. Woo, J. Lindner, K. D. Shelton and M. A. Magnuson, 2006 Multiallelic disruption of the rictor gene in mice reveals that mTOR complex 2 is essential for fetal growth and viability. Dev Cell 11**:** 583-589.

Shiu, P. K., and R. L. Metzenberg, 2002 Meiotic silencing by unpaired DNA: properties, regulation and suppression. Genetics 161**:** 1483-1495.

Shiu, P. K., N. B. Raju, D. Zickler and R. L. Metzenberg, 2001 Meiotic silencing by unpaired DNA. Cell 107**:** 905-916.

Shiu, P. K., D. Zickler, N. B. Raju, G. Ruprich-Robert and R. L. Metzenberg, 2006 SAD-2 is required for meiotic silencing by unpaired DNA and perinuclear localization of SAD-1 RNA-directed RNA polymerase. Proc Natl Acad Sci U S A 103**:** 2243-2248.

Smardon, A., J. M. Spoerke, S. C. Stacey, M. E. Klein, N. Mackin *et al.*, 2000 EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Curr Biol 10**:** 169-178.

Soper, S. F., G. W. van der Heijden, T. C. Hardiman, M. Goodheart, S. L. Martin *et al.*, 2008 Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis. Dev Cell 15**:** 285-297.

Staben, C., and C. Yanofsky, 1990 Neurospora crassa a mating-type region. Proc Natl Acad Sci U S A 87**:** 4917-4921.

Sun, H., D. Treco, N. P. Schultes and J. W. Szostak, 1989 Double-strand breaks at an initiation site for meiotic gene conversion. Nature 338**:** 87-90.

Tamaru, H., and E. U. Selker, 2001 A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. Nature 414**:** 277-283.

Tan, X., J. Rotllant, H. Li, P. De Deyne and S. J. Du, 2006 SmyD1, a histone methyltransferase, is required for myofibril organization and muscle contraction in zebrafish embryos. Proc Natl Acad Sci U S A 103**:** 2713-2718.

Tedford, K., S. Kim, D. Sa, K. Stevens and M. Tyers, 1997 Regulation of the mating pheromone and invasive growth responses in yeast by two MAP kinase substrates. Curr Biol 7**:** 228-238.

Thorstensen, T., P. E. Grini and R. B. Aalen, 2011 SET domain proteins in plant development. Biochim Biophys Acta 1809**:** 407-420.

Tian, C., W. T. Beeson, A. T. Iavarone, J. Sun, M. A. Marletta *et al.*, 2009 Systems analysis of plant cell wall degradation by the model filamentous fungus Neurospora crassa. Proc Natl Acad Sci U S A 106**:** 22157-22162.

Tsai, H. F., M. H. Wheeler, Y. C. Chang and K. J. Kwon-Chung, 1999 A developmentally regulated gene cluster involved in conidial pigment biosynthesis in Aspergillus fumigatus. J Bacteriol 181**:** 6469-6477.

Tsai, J. H., and B. D. McKee, 2011 Homologous pairing and the role of pairing centers in meiosis. J Cell Sci 124**:** 1955-1963.

Tsuji, G., T. Sugahara, I. Fujii, Y. Mori, Y. Ebizuka *et al.*, 2003 Evidence for involvement of two naphthol reductases in the first reduction step of melanin biosynthesis pathway of Colletotrichum lagenarium. Mycol Res 107**:** 854-860.

Turner, J. M., 2007 Meiotic sex chromosome inactivation. Development 134**:** 1823-1831.

Turner, J. M., O. Aprelikova, X. Xu, R. Wang, S. Kim *et al.*, 2004 BRCA1, histone H2AX phosphorylation, and male meiotic sex chromosome inactivation. Curr Biol 14**:** 2135-2142.

Turner, J. M., S. K. Mahadevaiah, P. J. Ellis, M. J. Mitchell and P. S. Burgoyne, 2006 Pachytene asynapsis drives meiotic sex chromosome inactivation and leads to substantial postmeiotic repression in spermatids. Dev Cell 10**:** 521-529.

Turner, J. M., S. K. Mahadevaiah, O. Fernandez-Capetillo, A. Nussenzweig, X. Xu *et al.*, 2005 Silencing of unsynapsed meiotic chromosomes in the mouse. Nat Genet 37**:** 41-47.

Updike, D., and S. Strome, 2010 P granule assembly and function in Caenorhabditis elegans germ cells. J Androl 31**:** 53-60.

van der Heijden, G. W., A. A. Derijck, E. Posfai, M. Giele, P. Pelczar *et al.*, 2007 Chromosome-wide nucleosome replacement and H3.3 incorporation during mammalian meiotic sex chromosome inactivation. Nat Genet 39**:** 251-258.

Wang, Z., N. Lehr, F. Trail and J. P. Townsend, 2012 Differential impact of nutrition on developmental and metabolic gene expression during fruiting body development in Neurospora crassa. Fungal Genet Biol 49**:** 405-413.

Warringer, J., and A. Blomberg, 2006 Evolutionary constraints on yeast protein size. BMC Evol Biol 6**:** 61.

Wellen, K. E., G. Hatzivassiliou, U. M. Sachdeva, T. V. Bui, J. R. Cross *et al.*, 2009 ATP-citrate lyase links cellular metabolism to histone acetylation. Science 324**:** 1076-1080.

Woo, P. C., E. W. Tam, K. T. Chong, J. J. Cai, E. T. Tung *et al.*, 2010 High diversity of polyketide synthase genes and the melanin biosynthesis gene cluster in Penicillium marneffei. FEBS J 277**:** 3750-3758.

Wu, W., A. Hildebrand, T. Kasuga, X. Xiong and Z. Fan, 2013a Direct cellobiose production from cellulose using sextuple beta-glucosidase gene deletion Neurospora crassa mutants. Enzyme Microb Technol 52**:** 184-189.

Wu, W., T. Kasuga, X. Xiong, D. Ma and Z. Fan, 2013b Location and contribution of individual beta-glucosidase from Neurospora crassa to total beta-glucosidase activity. Arch Microbiol 195**:** 823-829.

Wullschleger, S., R. Loewith, W. Oppliger and M. N. Hall, 2005 Molecular organization of target of rapamycin complex 2. J Biol Chem 280**:** 30697-30704.

Xiao, H., W. G. Alexander, T. M. Hammond, E. C. Boone, T. D. Perdue *et al.*, 2010 QIP, a protein that converts duplex siRNA into single strands, is required for meiotic silencing by unpaired DNA. Genetics 186**:** 119-126.

Yamamoto, A., and Y. Hiraoka, 2001 How do meiotic chromosomes meet their homologous partners?: lessons from fission yeast. Bioessays 23**:** 526-533.

Yamashita, K., A. Shiozawa, S. Banno, F. Fukumori, A. Ichiishi *et al.*, 2007 Involvement of OS-2 MAP kinase in regulation of the large-subunit catalases CAT-1 and CAT-3 in Neurospora crassa. Genes Genet Syst 82**:** 301-310.

Zhang, Y., R. Lamm, C. Pillonel, S. Lam and J. R. Xu, 2002 Osmoregulation and fungicide resistance: the Neurospora crassa os-2 gene encodes a HOG1 mitogen-activated protein kinase homologue. Appl Environ Microbiol 68**:** 532-538.

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

APPENDIX A

Table A.1: Mass spectrometry results from three independent pull-down experiments (Chapter II)

Broad[a] ID #	Protein Description	Frequency[b]	Experiment 1 C.I. %[c]		Experiment 2 C.I. %[c]		Experiment 3 C.I. %[c]	
			In solution[d]	In gel[e]	In solution[d]	In gel[e]	In solution[d]	In gel[e]
NCU02088	Suppressor of meiotic silecing-5	3	100	99.8	100	99.2	100	100
NCU01379	histone acetylase complex subunit Paf400	3	100		99.9	22	99.6	
NCU07854	cytosolic regulator Pianissimo	3		97.5	99.9		99.5	
NCU01680	plasma membrane ATPase-1	2	100				100	
NCU08936	clock-controlled gene-15	2	100				100	
NCU01323	cohesin complex subunit	2	100				99.8	
NCU10021	high affinity glucose transporter-1	2			100		100	
NCU05488	RNA-binding protein Vip1	2			99.9	18.3	100	
NCU07554	chromosome segregation protein SudA	2	99.9		99.9			
NCU04865	polyketide synthase-3	2	99.9	0		70.4		
NCU01634	histone H4-1	2		99.9			99.9	0
NCU08600	hypothetical protein	2			99.9		99.3	
NCU10346	hypothetical protein	2			99.9		98.5	
NCU06701	cephalosporin C regulator 1	2				57.1	99.8	
NCU03072	hypothetical protein	2				0	99.1	
NCU05803	translational activator	1	100					
NCU06977	hypothetical protein	1	100					

Table A.1: Continued.

Broad[a] ID #	Protein Description	Frequency[b]	Experiment 1 C.I. %[c]		Experiment 2 C.I. %[c]		Experiment 3 C.I. %[c]	
			In solution[d]	In gel[e]	In solution[d]	In gel[e]	In solution[d]	In gel[e]
NCU09477	ADP, ATP carrier protein	1	100					
NCU03234	ubiquitin-protein ligase E3 component	1	100					
NCU08114	cellodextrin transport-2	1	100					
NCU10066	coatomer alpha subunit	1	100					
NCU09263	anchored cell wall protein-4	1	100					
NCU06283	conserved hypothetical protein	1	100					
NCU01878	vesicle-mediated transporter	1	100					
NCU04102	hypothetical protein	1	100					
NCU05989	hypothetical protein	1	99.9					
NCU08957	hypothetical protein	1	99.9					
NCU04051	protein GCN20 - ABC transporter	1	99.9					
NCU03116	Ras GTPase activating protein	1	99.9					
NCU06518	NADH-cytochrome b5 reductase 2	1	99.9					
NCU04021	conserved hypothetical protein - ABC transporter	1	99.9					
NCU02539	cell division control protein 54	1	99.9					
NCU09481	hypothetical protein	1	99.9					
NCU08889	hypothetical protein	1	99.9					
NCU08598	quelling-defective-3	1	99.9					
NCU08919	chromatin remodelling factor 2-1	1	99.9					
NCU06935	hypothetical protein	1	99.9					
NCU07894	oligopeptide transporter 2	1	99.9					
NCU17239	probable apsB protein	1	99.9					
NCU02314	hypothetical protein	1	99.9					
NCU01823	two-component sensor protein histidine protein kinase	1	99.9					

Table A.1: Continued.

Broad[a] ID #	Protein Description	Frequency[b]	Experiment 1 C.I. %[c]		Experiment 2 C.I. %[c]		Experiment 3 C.I. %[c]	
			In solution[d]	In gel[e]	In solution[d]	In gel[e]	In solution[d]	In gel[e]
NCU08262	hypothetical protein	1	99.9					
NCU02437	histone H2A	1		99.9				
NCU05693	interferon-induced GTP-binding protein Mx2	1		99.9				
NCU10142	hypothetical protein	1		99.6				
NCU02953	hypothetical protein	1		96.1				
NCU07648	hypothetical protein	1		96				
NCU02793	hypothetical protein	1			99.9			
NCU01288	hypothetical protein	1			99.9			
NCU16491	fungal specific transcription factor domain-containing protein	1			99.9			
NCU05191	hypothetical protein	1			99.9			
NCU07129	amino-acid permease inda1	1			99.9			
NCU02075	heat shock protein 70-2	1			99.9			
NCU05028	kinesin	1			99.9			
NCU07378	serine/threonine protein kinase-12	1			99.9			
NCU07690	methylenetetrahydrofolate reductase 1	1			99.9			
NCU03220	hypothetical protein	1			99.9			
NCU07027	glycogen phosphorylase	1			99.9			
NCU03477	hypothetical protein	1			99.9			
NCU04117	ATP-dependent permease MDL2	1			99.9			
NCU08452	hypothetical protein	1			99.9			
NCU03534	hypothetical protein	1			99.9			
NCU01895	response regulator-1	1			99.9			
NCU04143	serine/threonine protein kinase-26	1			99.9			
NCU01053	hypothetical protein	1			99.9			

Table A.1: Continued.

Broad[a] ID #	Protein Description	Frequency[b]	Experiment 1 C.I. %[c]		Experiment 2 C.I. %[c]		Experiment 3 C.I. %[c]	
			In solution[d]	In gel[e]	In solution[d]	In gel[e]	In solution[d]	In gel[e]
NCU07369	hypothetical protein	1			99.9			
NCU02994	hypothetical protein	1			99.9			
NCU11309	integral membrane protein	1				99.8		
NCU00940	hypothetical protein	1				97.5		
NCU07984	chromosome segregation protein	1				96.2		
NCU09049	hypothetical protein	1					100	
NCU01552	ribosomal protein S28	1					99.9	
NCU02514	ATPase-1	1					99.9	
NCU09883	hypothetical protein	1					99.9	
NCU00019	Fork head domain 1	1					99.9	
NCU16593	hypothetical protein	1					99.9	
NCU03530	anchored cell wall protein-6	1					99.8	
NCU04137	vacuolar protein sorting-associated protein Vps5	1					99.8	
NCU04757	hypothetical protein	1					99.7	
NCU06468	midasin	1					99.6	
NCU03277	peroxin 10	1					99.6	
NCU01615	hypothetical protein	1					99.5	
NCU09642	high affinity sulfate transporter 1	1					99.5	
NCU01839	carboxylesterase	1					99.5	
NCU03592	P-type ATPase	1					99.5	
NCU06054	squalene synthetase	1					99.4	
NCU00582	cryptochrome	1					99.4	
NCU03857	tricarboxylic acid-5	1					99.2	
NCU03787	hypothetical protein	1					99	
NCU01869	hypothetical protein	1					98.9	

Table A.1: Continued.

Broad[a] ID #	Protein Description	Frequency[b]	Experiment 1 C.I. %[c]		Experiment 2 C.I. %[c]		Experiment 3 C.I. %[c]	
			In solution[d]	In gel[e]	In solution[d]	In gel[e]	In solution[d]	In gel[e]
NCU04105	hypothetical protein	1					98.9	
NCU09814	hypothetical protein	1					98.9	
NCU09596	phytanoyl-CoA dioxygenase	1					98.8	
NCU16656	hypothetical protein	1					98.7	
NCU06103	mitochondrial translation initiation factor	1					98.5	
NCU03523	serine/threonine protein kinase-22	1					98.4	
NCU08468	actin-interacting protein	1					98.3	
NCU09434	Suppressor of meiotic silecing-2 (Argonaute protein)[f]	1		83.7				
NCU06190	Suppressor of meiotic silencing-9 (Helicase)[f]	1				33.6		

[a]Broad = Broad Institute (http://www.broadinstitute.org/).
[b]Frequency, total number of times the protein was detected.
[c]C.I. %, Total Ion Score and/or Protein Score Confidence Interval. "0" under this column means that the protein was detected but the C.I. % was zero or lower.
[d]In solution, eluted proteins were precipitated and concentrated. The whole protein mix was enzymatically digested followed by MS identification.
[e]In gel, eluted proteins were separated by SDS-PAGE. Protein bands were cut from the gel, enzymatically digest followed by MS identification.

Table A.2: Gene Ontology (GO) Terms Assigned to the Affinity Purified Proteins Using Blast2GO (Chapter II)

Sequence ID #	Protein Description	GO molecular function
NCU02088	suppressor of meiotic silecing-5	protein binding;
NCU01379	histone acetylase complex subunit Paf400	transferase activity; protein binding; kinase activity;
NCU07854	cytosolic regulator Pianissimo	binding;
NCU01680	plasma membrane ATPase-1	hydrolase activity; transporter activity; nucleotide binding; protein binding; binding;
NCU08936	clock-controlled gene-15	
NCU01323	cohesin complex subunit	protein binding; chromatin binding; nucleotide binding;
NCU10021	high affinity glucose transporter-1	transporter activity;
NCU05488	RNA-binding protein Vip1	nucleic acid binding; nucleotide binding;
NCU07554	chromosome segregation protein SudA	protein binding; structural molecule activity; nucleotide binding;
NCU04865	polyketide synthase-3	transferase activity; binding; catalytic activity; protein binding;
NCU01634	histone H4-1	DNA binding; protein binding;
NCU08600	hypothetical protein	
NCU10346	hypothetical protein	chromatin binding; DNA binding;
NCU06701	cephalosporin C regulator 1	transcription regulator activity; DNA binding;
NCU03072	hypothetical protein	

Table A.2: Continued.

Sequence ID #	Protein Description	GO biological process
NCU02088	suppressor of meiotic silecing-5	
NCU01379	histone acetylase complex subunit Paf400	multicellular organismal development; DNA metabolic process; metabolic process; regulation of biological process; response to stress; biosynthetic process; cell differentiation; organelle organization; cellular protein modification process;
NCU07854	cytosolic regulator Pianissimo	multicellular organismal development; regulation of biological process; cytoskeleton organization; metabolic process; signal transduction;
NCU01680	plasma membrane ATPase-1	nucleobase-containing compound metabolic process; biosynthetic process; ion transport; catabolic process; generation of precursor metabolites and energy;
NCU08936	clock-controlled gene-15	
NCU01323	cohesin complex subunit	regulation of biological process; cell cycle; signal transduction; response to stress; nucleobase-containing compound metabolic process; organelle organization; response to abiotic stimulus; DNA metabolic process; biosynthetic process; transport;
NCU10021	high affinity glucose transporter-1	transport;
NCU05488	RNA-binding protein Vip1	
NCU07554	chromosome segregation protein SudA	cytoskeleton organization; cell cycle; signal transduction; biological_process; organelle organization; biosynthetic process; regulation of biological process; DNA metabolic process;
NCU04865	polyketide synthase-3	biosynthetic process; metabolic process;
NCU01634	histone H4-1	multicellular organismal development; cell differentiation; regulation of biological process; organelle organization; signal transduction;
NCU08600	hypothetical protein	

Table A.2: Continued.

Sequence ID #	Protein Description	GO biological process
NCU10346	hypothetical protein	
NCU06701	cephalosporin C regulator 1	regulation of biological process;
NCU03072	hypothetical protein	

Sequence ID #	Protein Description	GO cellular component
NCU02088	suppressor of meiotic silecing-5	
NCU01379	histone acetylase complex subunit Paf400	nucleoplasm; protein complex;
NCU07854	cytosolic regulator Pianissimo	intracellular;
NCU01680	plasma membrane ATPase-1	cell; plasma membrane;
NCU08936	clock-controlled gene-15	cell;
NCU01323	cohesin complex subunit	cytoplasm; chromosome; protein complex; nuclear chromosome; cell;
NCU10021	high affinity glucose transporter-1	cell;
NCU05488	RNA-binding protein Vip1	
NCU07554	chromosome segregation protein SudA	chromosome; proteinaceous extracellular matrix; cytoplasm; nucleus; cytoskeleton; protein complex;
NCU04865	polyketide synthase-3	
NCU01634	histone H4-1	chromosome; nucleoplasm;
NCU08600	hypothetical protein	
NCU10346	hypothetical protein	chromosome;
NCU06701	cephalosporin C regulator 1	
NCU03072	hypothetical protein	

Table A.3: Fungal strains used in the analysis of the involvement of NCU05488 in meiotic silencing (Chapter II)

Name[a]	Genotype[b]	Origin
FGS2489	*A*	FGSC[c]
FGS2490	a	FGSC[c]
FGSC21956	*NCU05488$^{\Delta}$; a*	FGSC[c]
KBCNR05A	*his-3$^{1\text{-}234\text{-}723}$, Rsp^{RIP93}; fl^{P} A*	Aramayo Lab Collection
RANCR49A	*fl^{P} A*	Aramayo Lab Collection
RPNCR73A	*Rsp^{RIP93}; inl^{89601} a*	Aramayo Lab Collection
RPNCR74A	*his-3$^{1\text{-}234\text{-}723}$, Rsp^{RIP93}; inl^{89601} A*	Aramayo Lab Collection
VSNCR60A	*NCU05488$^{\Delta}$; a*	Aramayo Lab Collection
VSNCR63A	*NCU05488D; a*	Aramayo Lab Collection
VSNCR67A	*Rsp^{RIP93} ; NCU05488$^{\Delta}$ a*	Aramayo Lab Collection
VSNCR68A	*his-3$^{1\text{-}234\text{-}723}$, Rsp^{RIP93}; NCU05488$^{\Delta}$ A*	Aramayo Lab Collection

[a]KBNC, RANC, RPNC and VSNC indicate strains constructed for Kevin D. Baker, Rodolfo Aramayo, Robert J. Pratt, and Victoria Suescún, respectively.
[b]Detailed description of all loci can be found at "The Neurospora e-Compemdium" (www.fgsc.net)
[c]FGSC, indicates strains obtained from the Fungal Genetics Stock Center

Table A.4: Plasmids used in the identification of protein-binding partners for SMS-5 project (Chapter II)

Name	Description	Origin
pGEX-6P-1	Expression vector for construction of GST-tagged proteins	GE Healthcare
pET-28a(+)	Expression vector for construction of T7-tagged proteins	Novagen
pCITE-4a(+)	Expression vector for construction of radiolabeled proteins	Novagen
pDL362	pGEX-6P-1_SMS-5. Expression GST-SMS-5	D. Lee, Aramayo Lab
pAVS011	pCITE-4a(+)_SMS-5. Expression of radiolabeled SMS-5	V. Suescún; Aramayo Lab
pAVS012	pET28a(+)_SMS-5. Expression of T7-tagged SMS-5	V. Suescún; Aramayo Lab
pAVS017	pET28a(+)_NCU05488. Expression of T7-tagged NCU05488	V. Suescún; Aramayo Lab
pAVS019	pCITE-4a(+)_NCU05488. Expression of radiolabeled NCU05488	V. Suescún; Aramayo Lab
pAVS020	pET28a(+)_Ku70. Expression of T7-tagged Ku70	V. Suescún; Aramayo Lab
pAVS021	pET28a(+)_Ku80. Expression of T7-tagged Ku80	V. Suescún; Aramayo Lab
pAVS022	pCITE-4a(+)_Ku70. Expression of radiolabeled Ku70	V. Suescún; Aramayo Lab
pAVS023	pCITE-4a(+)_Ku80. Expression of radiolabeled Ku80	V. Suescún; Aramayo Lab
pAVS051	pET28a-Pianissimo (N-ter). Expression of T7-tagged Pianissimo (N-ter)	V. Suescún; Aramayo Lab
pAVS054	pET28a-Pianissimo (C-ter). Expression of T7-tagged Pianissimo (C-ter)	V. Suescún; Aramayo Lab
pAVS055	pCITE-Pianissimo (C-ter). Expression of radiolabeled Pianissimo (C-ter)	V. Suescún; Aramayo Lab
pAVS056	pET28a-PAF400 (C-ter)	V. Suescún; Aramayo Lab
pAVS058	pCITE-PAF400 (C-ter). Expression of radiolabeled PAF400 (C-ter)	V. Suescún; Aramayo Lab

Table A.5: Oligonucleotides used in the identification of protein-binding partners for SMS-5 project (Chapter II)

Name[a]	Description[b]	Sequence
OVS095	T7-PROMOTER PRIMER (pET28a)	TAATACGACTCACTATAGGG
OVS096	T7-TERMINATOR PRIMER (pET28a)	GCTAGTTATTGCTCAGCGG
OVS097	pCITE primer (pCITE)	GGGGACGTGGTTTTCCTTTG
OVS098	T3 PROMOTER PRIMER (pCITE)	ATTAACCCTCACTAAAGGG
OVS0103	NCU05488 - F (EcoRI)	CTGCGGAATTCATGTCTACAGTCTAC
OVS0104	NCU05488 - R (NotI)	ATTGCGGCCGCTTACTGGGGAATCTT
OVS0120	pCITE F	TGCTTTACATGTGTTTAGTCG
OVS0121	pCITE R	CTCACTAAAGGGAACAAAAG
OVS0122	NCU05488-F	AACGAACAACAGCAAACA
OVS0123	NCU05488-R	TCATGAAATCAGCCAGAC
OVS0124	NCU07854-F N-ter (BamHI)	AGATGGATCCATGGCTGGGCCATCCACCAT
OVS0125	NCU07854-R N-ter (NotI)	ATGCGGCCGCGAGAAGTTTGTTATCCGA
OVS126	NCU07854-F C-ter (BamHI)	AGATGGATCCACGCAAAAGTACGTCAGAGT
OVS127	NCU07854-R C-ter (NotI)	ATGCGGCCGCCTACCTAAACGGCCCTCTGA
OVS137	pET28 upstream primer	CCATCGCCGCTTCCACTT
OVS142	NCU07854-F C-ter (BamHI)	AGGTGGATCCTCGGAGGTCAAGAGCACAAG
OVS143	NCU07854-R C-ter (NotI)	ATGCGGCCGCCCGCCACGCTACCTAAA
OVS144	NCU01379-R C-term (NotI)	ATGCGGCCGCCAGATACGGCATCCACAA
OVS145	NCU01379-F N-term (SacI)	GCCAGAGCTCATGGCGAACCTAATTGACGATG
OVS146	NCU01379-R N-term (NotI)	ATCGGCGGCCGCGATCAGCTCAATAGACCT

[a]Oligonucleotides were named following Aramayo Lab nomenclature based on the person that design the primers.
OVS = Oligonucleotide_Victoria_Suescún
[b]Description of the region/gene where the oligonucleotide annealed.

APPENDIX B

Table B.1: Proteins identified from two stages of *N. crassa* sexual development via mass spectrometry analysis and discussed in Chapter III are included in a separate file.

Table B.2: GO terms bias analysis of identified proteins (Chapter III)

Biologicla Process								
GO_ID	GO_TERM	Proteome			Genome			Log$_{10}$ (frequency[c])
		#PGOT[a]	#PBP[b]	(%)	#PGOT	#PBP	(%)	
GO:0008037	cell recognition	0	717	0.0%	7	5321	0.1%	#NUM!
GO:0048519	negative regulation of biological process	0	717	0.0%	5	5321	0.1%	#NUM!
GO:0048583	regulation of response to stimulus	0	717	0.0%	5	5321	0.1%	#NUM!
GO:0023051	regulation of signaling	0	717	0.0%	5	5321	0.1%	#NUM!
GO:0006091	generation of precursor metabolites and energy	103	717	14.4%	219	5321	4.1%	0.54
GO:0044249	cellular biosynthetic process	130	717	18.1%	303	5321	5.7%	0.50
GO:1901576	organic substance biosynthetic process	130	717	18.1%	303	5321	5.7%	0.50
GO:0034645	cellular macromolecule biosynthetic process	130	717	18.1%	303	5321	5.7%	0.50
GO:0009059	macromolecule biosynthetic process	130	717	18.1%	303	5321	5.7%	0.50
GO:0006412	translation	130	717	18.1%	303	5321	5.7%	0.50
GO:0010467	gene expression	132	717	18.4%	344	5321	6.5%	0.45
GO:0007028	cytoplasm organization	4	717	0.6%	13	5321	0.2%	0.36
GO:0040007	growth	106	717	14.8%	372	5321	7.0%	0.33
GO:0044767	single-organism developmental process	119	717	16.6%	461	5321	8.7%	0.28
GO:0009790	embryo development	119	717	16.6%	461	5321	8.7%	0.28
GO:0019748	secondary metabolic process	33	717	4.6%	128	5321	2.4%	0.28
GO:0007267	cell-cell signaling	24	717	3.3%	98	5321	1.8%	0.26
GO:0016049	cell growth	17	717	2.4%	70	5321	1.3%	0.26
GO:0000003	reproduction	132	717	18.4%	560	5321	10.5%	0.24
GO:0007010	cytoskeleton organization	49	717	6.8%	208	5321	3.9%	0.24
GO:0009607	response to biotic stimulus	30	717	4.2%	130	5321	2.4%	0.23
GO:0048856	anatomical structure development	152	717	21.2%	664	5321	12.5%	0.23
GO:0009058	biosynthetic process	319	717	44.5%	1408	5321	26.5%	0.23
GO:0009056	catabolic process	197	717	27.5%	877	5321	16.5%	0.22
GO:0019538	protein metabolic process	252	717	35.1%	1159	5321	21.8%	0.21

Table B.2: Continued.

Biological Process								
		Proteome			Genome			Log$_{10}$
GO_ID	GO_TERM	#PGOT[a]	#PBP[b]	(%)	#PGOT	#PBP	(%)	(frequency[c])
GO:0005975	carbohydrate metabolic process	107	717	14.9%	493	5321	9.3%	0.21
GO:0044267	cellular protein metabolic process	176	717	24.5%	857	5321	16.1%	0.18
GO:0044765	single-organism transport	51	717	7.1%	253	5321	4.8%	0.17
GO:0006811	ion transport	51	717	7.1%	253	5321	4.8%	0.17
GO:0009605	response to external stimulus	28	717	3.9%	141	5321	2.6%	0.17
GO:0007610	behavior	24	717	3.3%	122	5321	2.3%	0.16
GO:0008219	cell death	49	717	6.8%	254	5321	4.8%	0.16
GO:0009653	anatomical structure morphogenesis	84	717	11.7%	436	5321	8.2%	0.16
GO:0016265	death	49	717	6.8%	255	5321	4.8%	0.15
GO:0043170	macromolecule metabolic process	266	717	37.1%	1400	5321	26.3%	0.15
GO:0032502	developmental process	224	717	31.2%	1193	5321	22.4%	0.14
GO:0009628	response to abiotic stimulus	53	717	7.4%	283	5321	5.3%	0.14
GO:0032501	multicellular organismal process	214	717	29.8%	1153	5321	21.7%	0.14
GO:0007275	multicellular organismal development	214	717	29.8%	1153	5321	21.7%	0.14
GO:0044707	single-multicellular organism process	214	717	29.8%	1153	5321	21.7%	0.14
GO:0044237	cellular metabolic process	333	717	46.4%	1833	5321	34.4%	0.13
GO:0007049	cell cycle	80	717	11.2%	443	5321	8.3%	0.13
GO:0065008	regulation of biological quality	18	717	2.5%	100	5321	1.9%	0.13
GO:0019725	cellular homeostasis	18	717	2.5%	100	5321	1.9%	0.13
GO:0042592	homeostatic process	18	717	2.5%	100	5321	1.9%	0.13
GO:0007005	mitochondrion organization	11	717	1.5%	62	5321	1.2%	0.12
GO:0006950	response to stress	125	717	17.4%	710	5321	13.3%	0.12
GO:0044260	cellular macromolecule metabolic process	191	717	26.6%	1096	5321	20.6%	0.11
GO:0071704	organic substance metabolic process	448	717	62.5%	2673	5321	50.2%	0.09
GO:0044238	primary metabolic process	462	717	64.4%	2759	5321	51.9%	0.09
GO:0009987	cellular process	442	717	61.6%	2652	5321	49.8%	0.09
GO:0044764	multi-organism cellular process	4	717	0.6%	24	5321	0.5%	0.09

Table B.2: Continued.

Biological Process								
		Proteome			**Genome**			**Log₁₀**
GO_ID	**GO_TERM**	**#PGOT[a]**	**#PBP[b]**	**(%)**	**#PGOT**	**#PBP**	**(%)**	**(frequency[c])**
GO:0016032	viral process	4	717	0.6%	24	5321	0.5%	0.09
GO:0033036	macromolecule localization	49	717	6.8%	297	5321	5.6%	0.09
GO:0008104	protein localization	49	717	6.8%	297	5321	5.6%	0.09
GO:0045184	establishment of protein localization	49	717	6.8%	297	5321	5.6%	0.09
GO:0071702	organic substance transport	49	717	6.8%	297	5321	5.6%	0.09
GO:0015031	protein transport	49	717	6.8%	297	5321	5.6%	0.09
GO:0008152	metabolic process	625	717	87.2%	3791	5321	71.2%	0.09
GO:0044699	single-organism process	333	717	46.4%	2044	5321	38.4%	0.08
GO:0044710	single-organism metabolic process	91	717	12.7%	562	5321	10.6%	0.08
GO:0009719	response to endogenous stimulus	33	717	4.6%	204	5321	3.8%	0.08
GO:0050896	response to stimulus	196	717	27.3%	1251	5321	23.5%	0.07
GO:0048869	cellular developmental process	62	717	8.6%	408	5321	7.7%	0.05
GO:0030154	cell differentiation	62	717	8.6%	408	5321	7.7%	0.05
GO:0044763	single-organism cellular process	225	717	31.4%	1487	5321	27.9%	0.05
GO:0006996	organelle organization	94	717	13.1%	625	5321	11.7%	0.05
GO:0050789	regulation of biological process	250	717	34.9%	1701	5321	32.0%	0.04
GO:0065007	biological regulation	251	717	35.0%	1733	5321	32.6%	0.03
GO:0071840	cellular component organization or biogenesis	156	717	21.8%	1082	5321	20.3%	0.03
GO:0016043	cellular component organization	156	717	21.8%	1082	5321	20.3%	0.03
GO:0008283	cell proliferation	29	717	4.0%	202	5321	3.8%	0.03
GO:0051179	localization	177	717	24.7%	1239	5321	23.3%	0.03
GO:0051234	establishment of localization	177	717	24.7%	1239	5321	23.3%	0.03
GO:0006810	transport	177	717	24.7%	1239	5321	23.3%	0.03
GO:0006629	lipid metabolic process	66	717	9.2%	466	5321	8.8%	0.02
GO:0006807	nitrogen compound metabolic process	149	717	20.8%	1062	5321	20.0%	0.02
GO:0006725	cellular aromatic compound metabolic process	149	717	20.8%	1062	5321	20.0%	0.02

Table B.2: Continued.

Biological Process								
GO_ID	GO_TERM	Proteome			Genome			Log$_{10}$ (frequency[c])
		#PGOT[a]	#PBP[b]	(%)	#PGOT	#PBP	(%)	
GO:0034641	cellular nitrogen compound metabolic process	149	717	20.8%	1062	5321	20.0%	0.02
GO:0046483	heterocycle metabolic process	149	717	20.8%	1062	5321	20.0%	0.02
GO:0006139	nucleobase-containing compound metabolic process	149	717	20.8%	1062	5321	20.0%	0.02
GO:1901360	organic cyclic compound metabolic process	149	717	20.8%	1062	5321	20.0%	0.02
GO:0023052	signaling	86	717	12.0%	621	5321	11.7%	0.01
GO:0044700	single organism signaling	86	717	12.0%	621	5321	11.7%	0.01
GO:0007154	cell communication	93	717	13.0%	675	5321	12.7%	0.01
GO:0051716	cellular response to stimulus	71	717	9.9%	553	5321	10.4%	-0.02
GO:0050794	regulation of cellular process	71	717	9.9%	553	5321	10.4%	-0.02
GO:0007165	signal transduction	71	717	9.9%	553	5321	10.4%	-0.02
GO:0051704	multi-organism process	6	717	0.8%	60	5321	1.1%	-0.13
GO:0044419	interspecies interaction between organisms	6	717	0.8%	60	5321	1.1%	-0.13
GO:0019222	regulation of metabolic process	5	717	0.7%	50	5321	0.9%	-0.13
GO:0044403	symbiosis, encompassing mutualism through parasitism	6	717	0.8%	60	5321	1.1%	-0.13
GO:0060255	regulation of macromolecule metabolic process	5	717	0.7%	50	5321	0.9%	-0.13
GO:0010468	regulation of gene expression	5	717	0.7%	50	5321	0.9%	-0.13
GO:0043412	macromolecule modification	55	717	7.7%	580	5321	10.9%	-0.15
GO:0036211	protein modification process	55	717	7.7%	580	5321	10.9%	-0.15
GO:0006464	cellular protein modification process	55	717	7.7%	580	5321	10.9%	-0.15
GO:0090304	nucleic acid metabolic process	28	717	3.9%	323	5321	6.1%	-0.19
GO:0006259	DNA metabolic process	28	717	3.9%	323	5321	6.1%	-0.19

Table B.2: Continued.

Molecular Function								
		Proteome			Genome			Log (frequency[c])
GOID	GOTERM	#PGOT[a]	#PMF[b]	(%)	#PGOT	#PMF	(%)	
GO:0031386	protein tag	0	727	0.0%	3	5681	0.1%	#NUM!
GO:0019825	oxygen binding	0	727	0.0%	4	5681	0.1%	#NUM!
GO:0016817	hydrolase activity, acting on acid anhydrides	0	727	0.0%	29	5681	0.5%	#NUM!
GO:0045735	nutrient reservoir activity	1	727	0.1%	1	5681	0.0%	0.89
GO:0005198	structural molecule activity	96	727	13.2%	179	5681	3.2%	0.62
GO:0016209	antioxidant activity	12	727	1.7%	25	5681	0.4%	0.57
GO:0008135	translation factor activity, nucleic acid binding	23	727	3.2%	53	5681	0.9%	0.53
GO:0045182	translation regulator activity	3	727	0.4%	8	5681	0.1%	0.47
GO:0003779	actin binding	11	727	1.5%	40	5681	0.7%	0.33
GO:0030246	carbohydrate binding	24	727	3.3%	89	5681	1.6%	0.32
GO:0003723	RNA binding	94	727	12.9%	356	5681	6.3%	0.31
GO:0008233	peptidase activity	50	727	6.9%	199	5681	3.5%	0.29
GO:0009055	electron carrier activity	22	727	3.0%	100	5681	1.8%	0.24
GO:0005509	calcium ion binding	13	727	1.8%	60	5681	1.1%	0.23
GO:0043167	ion binding	13	727	1.8%	61	5681	1.1%	0.22
GO:0043169	cation binding	13	727	1.8%	61	5681	1.1%	0.22
GO:0046872	metal ion binding	13	727	1.8%	61	5681	1.1%	0.22
GO:0008092	cytoskeletal protein binding	22	727	3.0%	109	5681	1.9%	0.20
GO:0005102	receptor binding	14	727	1.9%	72	5681	1.3%	0.18
GO:0036094	small molecule binding	171	727	23.5%	1082	5681	19.0%	0.09

Table B.2: Continued.

Molecular Function								
GOID	GOTERM	Proteome			Genome			Log (frequency[c])
		#PGOT[a]	#PMF[b]	(%)	#PGOT	#PMF	(%)	
GO:1901265	nucleoside phosphate binding	171	727	23.5%	1082	5681	19.0%	0.09
GO:0000166	nucleotide binding	171	727	23.5%	1082	5681	19.0%	0.09
GO:0008289	lipid binding	15	727	2.1%	98	5681	1.7%	0.08
GO:0005515	protein binding	369	727	50.8%	2456	5681	43.2%	0.07
GO:0003824	catalytic activity	496	727	68.2%	3343	5681	58.8%	0.06
GO:1901363	heterocyclic compound binding	268	727	36.9%	1866	5681	32.8%	0.05
GO:0097159	organic cyclic compound binding	268	727	36.9%	1866	5681	32.8%	0.05
GO:0016787	hydrolase activity	176	727	24.2%	1275	5681	22.4%	0.03
GO:0005488	binding	562	727	77.3%	4237	5681	74.6%	0.02
GO:0060089	molecular transducer activity	9	727	1.2%	73	5681	1.3%	-0.02
GO:0004871	signal transducer activity	9	727	1.2%	73	5681	1.3%	-0.02
GO:0003676	nucleic acid binding	130	727	17.9%	1074	5681	18.9%	-0.02
GO:0005215	transporter activity	51	727	7.0%	433	5681	7.6%	-0.04
GO:0022892	substrate-specific transporter activity	3	727	0.4%	26	5681	0.5%	-0.04
GO:0022857	transmembrane transporter activity	3	727	0.4%	26	5681	0.5%	-0.04
GO:0022803	passive transmembrane transporter activity	3	727	0.4%	26	5681	0.5%	-0.04
GO:0022891	substrate-specific transmembrane transporter activity	3	727	0.4%	26	5681	0.5%	-0.04
GO:0015267	channel activity	3	727	0.4%	26	5681	0.5%	-0.04
GO:0015075	ion transmembrane transporter activity	3	727	0.4%	26	5681	0.5%	-0.04
GO:0022838	substrate-specific channel activity	3	727	0.4%	26	5681	0.5%	-0.04

Table B.2: Continued.

Molecular Function								
GOID	GOTERM	Proteome			Genome			Log (frequency[c])
		#PGOT[a]	#PMF[b]	(%)	#PGOT	#PMF	(%)	
GO:0005216	ion channel activity	3	727	0.4%	26	5681	0.5%	-0.04
GO:0030234	enzyme regulator activity	20	727	2.8%	175	5681	3.1%	-0.05
GO:0016740	transferase activity	102	727	14.0%	921	5681	16.2%	-0.06
GO:0042578	phosphoric ester hydrolase activity	5	727	0.7%	51	5681	0.9%	-0.12
GO:0016791	phosphatase activity	5	727	0.7%	51	5681	0.9%	-0.12
GO:0016772	transferase activity, transferring phosphorus-containing groups	28	727	3.9%	316	5681	5.6%	-0.16
GO:0016301	kinase activity	28	727	3.9%	316	5681	5.6%	-0.16
GO:0004872	receptor activity	4	727	0.6%	46	5681	0.8%	-0.17
GO:0016788	hydrolase activity, acting on ester bonds	12	727	1.7%	141	5681	2.5%	-0.18
GO:0004518	nuclease activity	7	727	1.0%	91	5681	1.6%	-0.22
GO:0016773	phosphotransferase activity, alcohol group as acceptor	16	727	2.2%	219	5681	3.9%	-0.24
GO:0004672	protein kinase activity	16	727	2.2%	219	5681	3.9%	-0.24
GO:0003682	chromatin binding	4	727	0.6%	73	5681	1.3%	-0.37
GO:0003677	DNA binding	28	727	3.9%	561	5681	9.9%	-0.41
GO:0001071	nucleic acid binding transcription factor activity	5	727	0.7%	132	5681	2.3%	-0.53
GO:0003700	sequence-specific DNA binding transcription factor activity	5	727	0.7%	132	5681	2.3%	-0.53

[a]#PGOT: #Proteins_in_specific GO_Term
[b]#PBP: #Proteins_with_Biological Process_GO_Terms or
[b]#PMF: #Proteins_with_Molecular Function_GO_Terms
[c]frequency= Proteome%/Genome%