AN ABSTRACT OF THE DISSERTATION OF

Kyle R. Pomraning for the degree of Doctor of Philosophy in Molecular and Cellular Biology presented on November 20, 2012

Title: Characterization of *Neurospora crassa* and *Fusarium graminearum* Mutants Defective in Repeat-induced Point Mutation

Abstract approved:

Michael Freitag

Mutation of repetitive DNA by repeat-induced point mutation (RIP) is a process that occurs in many filamentous fungi of the Ascomycota during the sexual cycle. Concurrently, direct DNA repeats are often deleted by homologous recombination at high frequency during the sexual cycle. Thus, the processes of RIP and deletion compete to either mutate or remove repetitive DNA from the genome of filamentous fungi during sexual cycles. Both processes contribute to genome streamlining by controlling proliferation of transposable elements and by limiting expansion of gene families. While the genetic requirements for deletion by homologous recombination are well known, the mechanism behind the specific detection and mutation of repetitive DNA by RIP has yet to be elucidated as only a single gene essential for RIP, *rid*, has been identified.

We have developed *Fusarium graminearum* as a model organism for the study of RIP by showing that it mutates repetitive DNA frequently during the sexual cycle and that the mutations due to RIP are dependent on *rid*. Further, we have sequenced a genetic mapping strain of *F. graminearum* (00-676-2) and identified 62,310 single nucleotide polymorphisms (SNPs) compared to the reference strain (PH-1). The SNP map will be useful for quickly mapping new mutants by bulk segregant analysis and high-throughput sequencing for which bioinformatic tools were specifically developed. The groundwork has thus been laid for

identification of novel RIP mutants in *F. graminearum*, which being homothallic has a major advantage for identification of recessive mutations.

We used a forward genetics approach to shed light on the mechanism of RIP in *Neurospora crassa*. Two *rrr* mutants that dominantly reduce RIP and recombination were characterized and identified as different mutated alleles of the same gene, *rrr-1*^{L496P} and *rrr-1*^{G325N} by bulk segregant analysis and high-throughput sequencing. Bioinformatic characterization suggests RRR-1 belongs to a previously uncharacterized group of dynamin-like proteins, which are generally involved in membrane fission and fusion. RRR-1-GFP localizes to the nuclear membrane, but not DNA, suggesting it affects RIP and recombination frequency indirectly by altering nuclear membrane dynamics during sexual development and thereby altering temporal aspects of RIP and recombination. We used a reverse genetics approach to determine whether high frequency RIP and homologous recombination of repetitive DNA during the sexual cycle are linked mechanistically or spatio-temporally. We tested strains where genes important for deletion by homologous recombination were knocked out and found all to be completely RIP competent except *mre11*, which, while sterile in homozygous deletion crosses, displayed lower RIP frequency in heterozygous crosses. This suggests that *mre11* has roles in homologous recombination as well as non-homologous end joining may be important for RIP. Collectively, this work developed methods for efficiently mapping mutations and identified a novel protein that reduces RIP and recombination frequency but did not identify any mechanistic link between the two processes.

Characterization of *Neurospora crassa* and *Fusarium graminearum* Mutants Defective in Repeat-induced Point Mutation

by

Kyle R. Pomraning

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented November 20, 2012
Commencement June 2013

Doctor of Philosophy dissertation of Kyle R. Pomraning presented on November 20, 2012.

APPROVED:

Major Professor, representing Molecular and Cellular Biology

Director of the Molecular and Cellular Biology Program

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Kyle R. Pomraning, Author

ACKNOWLEDGEMENTS

I would like to express my tremendous gratitude to everyone who has supported me throughout graduate school. My advisor, Michael Freitag, brought me into the light of fungal research and provided the scaffolding for scientific success, without which I could not have completed this thesis. To my friends and colleagues, Erin Bredeweg and Brett Mellbye, I am grateful for constant moral support and enthusiastic scientific discussion. I would also like to express my thanks to all who have toiled beside me at the bench including Lanelle Connolly, Pallavi Phatale, Kristina Smith, Joe Whalen, Jon Galazka, Steve Friedman, Greg Ekberg, Kayly Lembke, Phuong Pham, Mu Feng and the many students from Oregon and abroad who have come through Michael's lab. To Jason Cumbie and Henry Priest I offer my appreciation for assistance with all things computational. My thanks also go out to my graduate committee, Kerry McPhail, Todd Mockler, Steve Strauss and Tom Wolpert for providing a positive environment during the most critical junctures of the graduate experience. To Kelly Vining, Snezhka Oliferendo, Scott Baker, Kevin Mc Cluskey, Monika Schmoll, Fred Bowring, David Catcheside, Eric Selker, Georgi Mitev, Bruce Geller and Buck Wilcox I offer my thanks for fruitful collaborations. To my family, Alex, Don and Mona Pomraning I am grateful for the December beach extravaganza each of the last five years.

CONTRIBUTION OF AUTHORS

Chapter 1: Kyle R. Pomraning wrote the chapter.

Chapter 2: Joseph P. Whalen, Lanelle R. Connolly, Kristina M. Smith, Kyle R. Pomraning and Michael Freitag performed the experiments and data analysis. Kyle R. Pomraning and Michael Freitag wrote the manuscript.

Chapter 3: Sequencing of the Mauriceville strain was done by Kristina M. Smith and Michael Freitag. Fred Bowring isolated the octad strains in David Catcheside's lab at Flinders University, Australia. The strains were sequenced by Michael Freitag and analyzed by Kyle R. Pomraning. All other experiments and analyses were performed by Kyle R. Pomraning. Kyle R. Pomraning and Michael Freitag wrote the manuscript.

Chapter 4: Ann T. Kobsa (Hageman) performed the mutagenesis and identified mutants in Eric Selker's lab at the University of Oregon, Eugene. All other experiments and analyses were performed by Kyle R. Pomraning. Kyle R. Pomraning wrote the chapter.

Chapter 5: Kyle R. Pomraning performed the experiments and wrote the chapter.

Chapter 6: Kyle R. Pomraning performed the experiments and wrote the chapter.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF TABLES

Introduction: repeat-induced gene inactivation in fungi

Kyle R. Pomraning

Manuscript in preparation for *Fungal Biology Reviews*

Genome defense in filamentous fungi

Approximately 13.8 billion years ago (Jarosik, Bennett et al. 2011) the universe began expanding outward from an infinitesimally small singularity. For 99.9999997% of the time since then nothing of significance happened. Primarily atoms flew around and would occasionally collide. This went on until a highly organized group of atoms, which we refer to as a *Homo sapiens*, figured out how to transform another highly organized group of atoms, *Neurospora crassa*, with exogenous DNA (Mishra and Tatum 1973). This is when things finally started to become interesting.

Since then, a number of processes have been described which inactivate genes present in greater than a single copy in the genome. Presumably their evolutionary benefit is protection of the host genome from invading segments of DNA, and this idea has been borne out by the accumulated evidence, as we will see below. The best characterized processes that target repetitive DNA in filamentous fungi include "repeat-induced point mutation" (RIP), which mutates repetitive DNA during the sexual cycle (Selker, Cambareri et al. 1987), "methylation induced premeiotically" (MIP), which methylates the cytosines of repetitive DNA during the sexual cycle (Goyon and Faugeron 1989), cytosine DNA methylation, which has been extensively studied in the vegetative phase of model fungi (Smith, Phatale et al. 2012), and "quelling" which post-transcriptionally silences repetitive DNA through an RNA interference-like mechanism during vegetative growth (Romano and Macino 1992). Quelling will not be discussed further as RNA interference mechanisms in *N. crassa* have been recently reviewed (Li, Chang et al. 2010).

RIP and MIP are fungal specific processes

RIP and MIP occur during the sexual cycle, between fertilization and karyogamy, which precludes identical processes from occurring in archea or bacteria. RIP was the first "genome defense" mechanism discovered in eukaryotes and has been indirectly tested for in many sexual species but never been reported to occur outside of ascomycetous fungi. Transformation of DNA often results in integration of multiple copies of transgenes at a single locus or several loci and in many cases these transgenes become silenced in *cis* or

trans. Most often, gene silencing has been found to be reversible and never associated with mutations in the model organisms favored by geneticists.

There are numerous mechanisms for gene silencing and a detailed discussion is beyond the scope of this introduction to RIP and MIP. Chromatin based silencing mechanisms that are based on Sir proteins in *Saccharomyces cerevisiae* (Hickman, Froyd et al. 2011) or small RNA-based mechanisms in *Schizosaccharomyces pombe* (Cam 2010) will not be further considered here. Likewise, in *Drosophila melanogaster*, a species with little (if any) cytosine methylation, repetitive DNA is susceptible to epigenetic silencing that is generally reversible and therefore not associated with hypermutation during sexual propagation (Dorer and Henikoff 1994; Sabl and Henikoff 1996). Early transformation experiments in the nematode worm *Caenorhabditis elegans*, another species devoid of cytosine methylation, found that high copy number transgenes are not mutated through multiple generations (Stinchcomb, Shaw et al. 1985) suggesting that hypermutation of repeats during the sexual cycle does not occur.

DNA methylation in plants and animals has been extensively studied. In plants repeated segments of DNA can indeed become *de novo* methylated. As discussed below, we still do not know if cytosine methylation is a pre-requisite for RIP or if it sets the stage for cytosine methylation in vegetative tissues of the progeny that issue from crosses in which RIP occurs. In *Arabidopsis thaliana*, for example, integration of multiple copies of the *hph* gene often results in loss of expression (Mittelsten Scheid, Afsar et al. 1994). However, gene inactivation in *A. thaliana* (or any plant studied thus far) is never associated with RIP-like mutations after going through the sexual cycles. Repeats typically result in expression of RNA that can form double-stranded intermediates, which triggers RNA-dependent DNA methylation (Wassenegger, Heimes et al. 1994; Law and Jacobsen 2010). Two types of cytosine DNA methyltransferases, CMT3 and DRM2, will *de novo* methylate repeated DNA segments, while a third, MET1, carries out maintenance methylation (Cao and Jacobsen 2002; Law and Jacobsen 2010). As in *A. thaliana*, it seemed likely that multi-copy genes in mammals were silenced epigenetically, most often by cytosine methylation (Assaad, Tucker et al. 1993), but the exact mechanism for *de novo* methylation of repeats remains unclear. As in plants, there is no evidence for hypermutation by RIP, even though cytidine deaminases,

for example activation-induced deaminase (AID) and apolipoprotein B-editing (APOBEC) enzymes, have been extensively studied in recent years (Zhu 2009; Sanz, Kota et al. 2010; Gazumyan, Bothmer et al. 2012). As far as we know, none of the mammalian cytidine deaminases act during the sexual cycle.

Without conducting transformation experiments for every sexual organism it is impossible to conclude that RIP does not occur outside the fungi, particularly if it occurs at much lower levels than in *N. crassa*, as would be expected for organisms with generally higher genomic repeat content. As we will see below, both RIP and MIP depend on ascomycete-specific cytosine methyltransferase-like proteins (Malagnac, Wendel et al. 1997; Freitag, Williams et al. 2002). These findings support the idea that both phenomena are fungal- and possibly ascomycete-specific.

Discovery of repeat-induced gene inactivation in filamentous fungi

The processes that target repetitive DNA in fungi were unlikely to be noticed until efficient protocols for the introduction of exogenous DNA were developed. Transformation of *N. crassa* was established soon after it had been developed for the budding yeast, *S. cerevisiae* (Mishra and Tatum 1973; Case, Schweizer et al. 1979) and was soon followed by reports of non-Mendelian inheritance. Transformation of *N. crassa* is complicated because the small hyphal fragments stemming from germlings and conidia (i.e., the asexual spores that are used for fungal transformation) contain many nuclei, only one of which is likely to be transformed. This leads to heterokaryotic transformants that when crossed do not appear to behave in a Mendelian manner as many of the nuclei involved in the cross do not contain transgenes. In the literature a number of labs debated whether the non-Mendelian result was due to the DNA being located extrachromosomally on an episomal plasmid, which could be lost during meiosis (Mishra and Tatum 1973; Mishra 1976) or whether the result was due to heterokaryosis (Szabo, Schablik et al. 1978; Szabo and Schablik 1980). However, the original paper describing transformation by Mishra and Tatum (Mishra and Tatum 1973) investigated the stability of transformed DNA in progeny after a backcross, the products of which should be homokaryotic following meiosis, and found transmission of the transgene to progeny to still be unstable, ruling out heterokaryosis but not episomes. Eventually the idea

of the episome was tested by Southern analyses, when researchers found that transformation of the *am*$^{+}$ gene into *am*$^{-}$ strains was so unstable that they could obtain only 1-2% *am*$^{+}$ progeny after meiosis (Grant, Lambowitz et al. 1984). Grant and colleagues were unable to identify episomally located *am* genes in their strains, ruling out the episome hypothesis but they did find copies of *am* that appeared rearranged based on restriction fragment length polymorphism (RFLP). Their conclusion was that either "*recombinant plasmids containing the am gene can replicate autonomously in N. crassa [*… and …*] integrate at multiple sites in the deletion host but in a nonfunctional manner [*… or …*] tandem repeats of the plasmid are integrated into chromosomal DNA but eliminated during meiosis*" (Grant, Lambowitz et al. 1984). The second, simpler interpretation would later turn out to be correct.

Mary Case at the University of Georgia closely followed the non-Mendelian inheritance from transformants containing duplications of the *qa-2* gene and found that crosses with these transformants often resulted in progeny with loss of function of *qa-2* even when both parents had functional *qa-2* genes (Case 1986). The stage was set for a set of incisive experiments carried out by Selker and colleagues at the University of Oregon. They closely examined progeny from crosses containing duplications and found that in many progeny the duplicated region was the expected size but contained different RFLP banding patterns when examined by Southern analyses (Selker, Cambareri et al. 1987). Additionally, the rearranged segments often showed evidence of cytosine methylation (Selker, Cambareri et al. 1987). This process was called "rearrangements induced premeiotically" or RIP. By analyzing individual octads, the products of meiosis from a cross, Selker and colleagues determined that this process occurs prior to the final premeiotic DNA replication since half the products from the meiosis are identical to each other (2:2) segregation. If RIP occurred after the final premeiotic DNA replication the rearrangements would be expected to segregate 1:1:1:1. Additionally, they observed that RIP occurs after fertilization, as they showed that common rearrangements do not occur in progeny from different perithecia, each of which is thought to represent a single fertilization event (Perkins and Barry 1977). Thus, RIP occurs after fertilization but prior to the final premeiotic DNA replication and karyogamy. During this period the parental nuclei of opposite mating type undergo seven to ten rounds of nuclear division (Emerson 1966). Within the nuclear division cycle RIP must

occur during G1 or near the replication fork during S phase since the entire genome would be subject to RIP after DNA replication. It was determined that multiple rounds of RIP can occur in backcrossed strains and that genetically linked and unlinked duplications are susceptible to RIP (Selker, Cambareri et al. 1987). They did not fail to notice that duplicated sequences can recombine to leave a single copy (Selker, Cambareri et al. 1987).

Another phenomenon, dubbed "methylation induced premeiotically" (MIP), was subsequently described in 1989 by Faugeron and colleagues following transformation experiments using an endogenous *met2* gene in *Ascobolus immersus*, which resulted in homologous integration at either the endogenous locus, resulting in single copy transformants, or ectopically, resulting in multi-copy transformants (Faugeron, Goyon et al. 1989). In these studies the multi-copy transformants were found to be silenced by cytosine methylation after going through the sexual cycle but were revertible in subsequent mitoses indicating the DNA had not been mutated (Goyon and Faugeron 1989). Additionally, duplications of different sequence were all independently able to induce methylation, suggesting there is no sequence specific signal (Barry, Faugeron et al. 1993). More detailed analyses using crosses containing one, two or three copies of the *amdS* gene found that MIP occurs prior to the final premeiotic DNA replication and always acts in nuclei with two or three copies without affecting single copy nuclei (Faugeron, Rhounim et al. 1990). This is significant because the single and multi-copy genes reside in nuclei within a common cytoplasm where RNA is thought to be relatively freely diffusible (Faugeron, Rhounim et al. 1990). The only other report of d*e novo* cytosine methylation of repeated DNA prior to karyogamy showed methylation of cytosines at CpG dinucleotides in the basidiomycete *Coprinopsis cinereus* between both tandem and unlinked duplications; this process is also not accompanied by mutation (Freedman and Pukkila 1993).

Duplicated DNA and nearby single copy sequences are targeted by MIP and RIP

Initial characterization found that repetitive DNA is susceptible to inactivation by RIP and MIP regardless of whether some or all of the copies were foreign DNA that has been introduced into the genome (Selker, Cambareri et al. 1987; Selker and Garrett 1988; Faugeron, Goyon et al. 1989; Goyon and Faugeron 1989; Foss, Garrett et al. 1991). Later

efforts found that MIP and RIP appear to act on naturally repetitive DNA without the caveat of having been introduced by transformation, including one mutated 5s rRNA gene ("theta", θ) that is present as a tandem duplication ("zeta-eta", ζ–η) in some *N. crassa* strains (Grayburn and Selker 1989), and in "methylated *Ascobolus* repeated sequences" (*Mars*) in *A. immersus* (Goyon, Rossignol et al. 1996).

Work by Fincham and colleagues investigated crosses of *N. crassa* carrying three copies of the *am* gene and found that much of the time either none, two or all three of the copies were inactivated (Fincham, Connerton et al. 1989). Similar results were found in numerous experiments conducted by Selker and colleagues (Selker, 1990). Similar studies in *A. immersus* crosses involving two or three copies of the *amdS* gene found that MIP results in methylation of both copies when present as a duplication but in methylation of either two or three copies when present as a triplication (Faugeron, Rhounim et al. 1990). Further studies of *met2* duplications in *A. immersus* found gene inactivation was reversible by cloning the methylated gene and propagating it in *e. coli* to get rid of any cytosine methylation, and then reintroducing it into *A. immersus* as a single copy (Rhounim, Rossignol et al. 1992). Thus results from both organisms suggest that both RIP and MIP act specifically on pairs of duplicated DNA sequences and that the pairing can occur multiple times prior to karyogamy during a cross between different pairs of a sequence when present in greater than two copies (Fincham, Connerton et al. 1989; Faugeron, Rhounim et al. 1990).

Investigations into the nature of random mutagenesis found 42 mutants of the *mtr* gene, one allele of which appeared to have undergone extensive RIP; however, no duplication of the gene was detected at any point in the experiment (Watters and Stadler 1995). This result raises the interesting possibility that either (1) RIP infrequently mutates non-repetitive DNA, (2) that DNA is occasionally duplicated, mutated and then recombined into a single copy during pre-meiosis, or (3) that other processes may infrequently generate mutations that leave hallmarks identical to RIP.

Studies on duplications of plasmid fragments found that RIP can occur in single copy sequences flanking duplicated regions up to 930 base pairs (bp) away, as determined by Sanger sequencing, and at least four kilobases (kb) away, as determined by Southern blot analysis (Foss and Selker 1991; Irelan, Hagemann et al. 1994). This spreading of mutations

was studied in greater detail using duplications of 1 – 2 kb inserted next to or up to 2 kb away from a single copy *erg-3* gene and found that only in some duplications directly next to the single copy gene was there enough spreading of RIP to efficiently inactivate *erg-3* (Vyas and Kasbekar 2005). Duplications 500 bp or farther away did not efficiently inactivate the single copy *erg-3* gene but this does not preclude that light RIP may have spread often without causing mutations severe enough to inactivate *erg-3* (Vyas and Kasbekar 2005). Sanger sequencing of alleles from random progeny, regardless of phenotype, would uncover just how much RIP actually spreads outside of a given duplication, but such studies have yet to be undertaken.

Use of RIP for molecular biology

RIP was first used as a tool to create novel mutations or null alleles of genes shortly after its discovery and is particularly useful when working with genes that are lethal as null alleles (Barbato, Calissano et al. 1996). By transforming a second copy of a gene into the genome and allowing the fungus to go through a cross one can obtain many different mutations specifically in the duplicated gene. Random mutations targeted to specific parts of a gene can also be achieved if one duplicates only the part of the gene in question. Since RIP occurs independently within the many fertilization events usually present during a cross and to varying extremes from no RIP to very heavy RIP within the population of ascospores from each fertilization event (Singer, Kuzminova et al. 1995) it is possible to get a very wide range of mutant alleles. This was first done to study the *am*, *acu-8*, *scon-2*, *nmr*, and *chs-1* genes of *N. crassa* (Connerton 1990; Marathe, Connerton et al. 1990; Paietta 1990; Jarai and Marzluf 1991; Yarden and Yanofsky 1991). The numerous methods devised including “sheltered RIP” to recover mutations in essential genes (Harkness, Metzenberg et al. 1994) will not be reviewed further here, though the bulk of references found in a PubMed search for “RIP” constitute papers on individual genes that were mutated by RIP, suggesting how successful a tool this has become for *Neurospora* geneticists.

Length constraints for targeting of duplicated DNA by RIP and MIP

Duplications of various sizes were constructed to characterize the minimum length of sequence homology a repeat must have to be efficiently detected by RIP in *N. crassa* or by MIP in *A. immersus*. The minimum length of a duplication for MIP to occur was determined by testing duplications of various sized *met2* fragments for their ability to induce cytosine methylation. Duplications of 1.1- 5.7 kb all result in targeting of dense cytosine methylation and phenotypic inactivation of a duplicated region (Barry, Faugeron et al. 1993; Goyon, Barry et al. 1996). A smaller duplication (630 bp) showed phenotypic inactivation in ~43% of the asci analyzed but some methylation was present in all progeny suggesting that MIP had occurred but not to the extent necessary to completely silence the reporter gene (Goyon, Barry et al. 1996). Even smaller duplications (317 and 457 bp) were never observed to be phenotypically inactivated by MIP, but methylation sensitive Southern analysis showed that all progeny analyzed with the 457 bp duplication were methylated by MIP, while ~20% of the progeny analyzed with the 317 bp duplication showed signs of methylation putting the size limit for efficient MIP at around 300 bp (Goyon, Barry et al. 1996). Intriguingly, bisulfite sequencing of the methylated regions indicated that longer duplications tend to induce methylation at all C's regardless of context while short duplications retain methylation primarily in the CpG context, differences that could be explained by multiple methyltransferases (*de novo* and maintenance) with different specificities acting at different times (Goyon, Barry et al. 1996).

The minimum length of duplications for RIP to occur was determined by measuring RIP sensitivity using different duplicated alleles of *mtr*. These studies found that a 380 bp duplication and all shorter duplications show no phenotypic evidence of RIP while a 446 bp duplication and all larger duplications show efficient RIP of greater than 70% (Watters, Randall et al. 1999). These studies put the detection limit for efficient RIP at around 400 bp but RIP may inefficiently target smaller duplications and mutate them lightly making sequencing necessary to detect RIP in short duplications and possibly accounting for the slight differences in detection limits found between RIP in *N. crassa* and MIP in *A. immersus*.

Studies of moderate size duplications suggest they are also targeted efficiently, as an 18 kb duplication of the *b2* gene is susceptible to MIP in *A. immersus* (Colot and Rossignol 1995) and a 10 kb duplication containing the *pan-2* gene is susceptible to RIP in *N. crassa*

(Rosa, Folco et al. 2004). If duplications of a few kb are efficiently detected and silenced by mutation or cytosine methylation in RIP and MIP respectively, then it follows that duplication of large segments of any chromosome should be deleterious as many genes ought to be inactivated. This is one explanation for why strains with large duplications infrequently produce viable progeny in backcrosses in *N. crassa* (Perkins, Margolin et al. 1997). Additional studies after the discovery of "meiotic silencing by unpaired DNA" (Aramayo and Metzenberg 1996; Shiu, Raju et al. 2001) revealed that this RNA-based meiotic silencing process can also affect recovery of progeny from crosses with large segmental duplications. The effect of segmental duplications present between different strains on the efficiency of MIP has not yet been tested in *A. immersus*.

In *N. crassa*, duplicated genes from a few progeny that made it through a cross involving a large segmental duplication were tested for mutations by sequencing and found to have mild levels of RIP, confirming that RIP acts upon large duplications in addition to small ones (Perkins, Margolin et al. 1997). RIP has also been measured in small duplications at loci outside of a large segmental duplication. Interestingly, the presence of a large segmental duplication in the genome seems to decrease RIP frequency in the small duplication (*erg-3*), regardless of whether the small and large duplications are in the same nucleus or nuclei of opposite mating type during a cross (Bhat and Kasbekar 2001). However, productivity of crosses heterozygous for large segmental duplications generally have reduced fertility due not only to RIP but also meiotic silencing of unpaired DNA (Shiu, Raju et al. 2001). It is quite possible that low RIP frequency is due to low fertility of the crosses, however, Kasbekar and colleagues argued against this given that in the very small number of progeny from interspecies crosses mutation of *erg-3* appeared to be normal. The necessary experiment to make this argument valid, interspecies crosses from the large segmental duplication containing strain was, however, not performed (Bhat and Kasbekar 2001). Additional strains containing large duplications from 117 kb to greater than 300 kb have been tested for suppression of RIP in small duplications and, interestingly the smallest segmental duplications of 117, 140 and 156 kb do not suppress RIP, while all larger duplications do (Vyas, Ravindran et al. 2006; Perkins, Freitag et al. 2007; Singh and Kasbekar 2008). One hypothesis to explain these effects includes titration of a limited pool of RIP

machinery by large duplications with a size equilibrium around 150 kb (Vyas, Ravindran et al. 2006; Singh and Kasbekar 2008; Singh, Iyer et al. 2009). No molecular data to support this hypothesis have been provided.

Based on these ideas one may imagine that large segmental duplications are also responsible for suppression of RIP phenotypes observed in some wild-collected *Neurospora* strains such as Adiopodoumé, Sugartown, Adiopodoumé-7, Fred, Coon, and Bayan Lepas, which all have been shown to have factors that dominantly suppress RIP (Noubissi, McCluskey et al. 2000; Noubissi, Aparna et al. 2001; Bhat, Noubissi et al. 2003). As described in "Ongoing and future studies" we have been unable to detect any large segmental duplications or expansion of repetitive elements in Adiopodoumé (FGSC430) by high-throughput sequencing of the entire genome arguing against this hypothesis. Instead, it seems more likely that one or more suppressors or modifiers of RIP are present in all wild-collected strains that have been shown to be at least partially deficient in RIP.

Divergence constraints for targeting of duplicated DNA by RIP and MIP

RIP was found to mutate repetitive DNA of decreasing homology in progeny of subsequent generations (Cambareri, Jensen et al. 1989; Cambareri, Singer et al. 1991), showing that the homology of duplicated sequences does not have to be complete for RIP to occur. Nevertheless, the number of progeny that appeared to have undergone RIP decreased in crosses of decreasing homology (Cambareri, Singer et al. 1991). This raises the question of how diverged two sequences can be and still be susceptible to RIP. The *ζ-η* region of *N. crassa*, a naturally occurring tandem duplication of the original θ region that includes a short 5S rRNA gene (Grayburn and Selker 1989), is diverged 14% from the native θ sequence and undergoes RIP only rarely (Selker 1990; Foss, Garrett et al. 1991) while RIP through successive generations found sequences could continue to be detected by RIP below 80% nucleotide identity (Cambareri, Singer et al. 1991). Thus, RIP will continue in successive sexual cycles until sequence identity between two copies decreases below ~85% but it will begin again if such regions become re-duplicated (Cambareri, Singer et al. 1991). Presumably these cycles of duplication and mutagenesis can continue until no cytosines remain. RIP

between diverged sequences with a “normal distribution” of C:G content has not been investigated, nor has the sequence divergence constraints for repeat detection by MIP.

Characterization of the effects of RIP

Early studies uncovered changes in the DNA sequence and cytosine DNA methylation level of duplications that had undergone RIP and were termed “rearrangements induced premeiotically” due to the changes in banding patterns observed on Southern blots (Selker, Cambareri et al. 1987). These changes are in effect a mixture of telltale hallmarks that result from (1) deletion of one of the repeats and all intervening sequence, a phenomenon widely observed in tandem repeats and thought to occur primarily by the single-strand annealing mechanism of double strand break repair (Sugawara and Haber 1992; Lyndaker and Alani 2009), and (2) polarized point mutations, which were used to rename the phenomenon “repeat-induced point mutation” (Cambareri, Jensen et al. 1989). The point mutations are exclusively C:G to T:A transition mutations that, in *Neurospora* and some other taxa, occur primarily when cytosine is in the CpA di-nucleotide context (Cambareri, Jensen et al. 1989). Such mutated sequences are prone to cytosine methylation (Selker, Cambareri et al. 1987) and serve as signals to generate *de novo* cytosine methylation when introduced into the genome (Selker, Jensen et al. 1987; Miao, Freitag et al. 2000; Tamaru and Selker 2003). Bisulfite sequencing found this methylation to be dense and nonsymmetrical (Selker, Fritz et al. 1993) and in contexts beyond that of CpG which is most common in mammals, though recently nonsymmetrical cytosine methylation has also been unambiguously documented in these animals (Lister, Pelizzola et al. 2009). Cytosine methylation can spread from the duplicated sequences into neighboring single copy sequence leading to silencing (Irelan and Selker 1997; Miao, Freitag et al. 2000). DNA methylation associated with RIP was found to be dependent on the degree of mutation by RIP. By looking at *am* alleles with various densities of RIP mutations it was determined that sequences that had undergone light RIP are usually not *de novo* methylated while more heavily RIP mutated alleles are (Singer, Marcotte et al. 1995), suggesting that the degree of RIP is the signal targeted for *de novo* methylation in vegetative cells of *N. crassa*. Further studies on *de novo* methylation signals showed that overall AT-richness and TpA density are factors in how well short DNA segments are *de novo*

methylated (Miao, Freitag et al. 2000; Tamaru and Selker 2003). The two alleles (*am*^{RIP3} *and* *am*^{RIP4}) that were faithfully methylated in the original progeny but which were not able to induce *de novo* methylation when reintroduced into the genome after methylation was stripped either by passage through bacteria or PCR, provided the first evidence for a system that carries out a form of maintenance methylation in *N. crassa* (Singer, Marcotte et al. 1995; Selker, Freitag et al. 2002).

Little systematic research has been carried out on the relationship of RIP and DNA methylation in other fungi. Insufficient density of mutations after RIP, may explain why mutated segments in *Podospora anserina* (Hamann, Feller et al. 2000; Hamann, Feller et al. 2000) and *Fusarium graminearum* (see Chapter 2) do not display cytosine methylation, as these sequences are typically lightly mutated and not as AT-rich as those that are *de novo* methylated in *N. crassa*. In *N. crassa* naturally mutated, AT-rich sequences make up a sizeable fraction of the genome (Galagan, Calvo et al. 2003; Selker, Tountas et al. 2003) and almost all such sequences have 5-methylcytosines (Lewis, Honda et al. 2009; Smith, Phatale et al. 2011). Selker and colleagues showed that this cytosine methylation requires the presence of histone 3 lysine 9 trimethylation and HP1, both hallmarks of constitutive heterochromatin (Tamaru and Selker 2001; Freitag, Hickey et al. 2004; Lewis, Honda et al. 2009).

RIP acts on paired sequences but mutates single strands of DNA independently

Early results from DNA sequencing of RIP mutated segments usually revealed mixtures of C to T and G to A mutations on a given DNA strand; however, occasionally heavily mutated sequences with only C to T or G to A mutations were identified, strongly suggesting that the RIP mutation machinery targets a single strand of DNA in one passage, and that the occurrence of both C to T and G to A mutations on a given strand is caused by recurrence of RIP, or multiple passages of the mutational machinery (Selker 1990; Jarai and Marzluf 1991).

Further evidence for this came from random mutagenesis experiments of the *mtr* gene. Three *mtr* mutants, presumably from the same ascus, were isolated and the *mtr* alleles sequenced. Two of the mutants had identical patterns of C to T mutation as expected if they were mitotic pairs from the mitosis that occurs post-meiotically but prior to ascospore

delineation. The third allele had a different pattern of mutation including additional C to T mutation that had occurred on the opposite strand of DNA. An exceptional ascus containing a pair of ascospores from a cross harboring a duplication of the *ζ-η* flank region which shared a number of common C:G to T:A mutations but where one ascospore had additional C to T mutations on the coding strand while the other had additional G to A mutations on the coding strand (Watters and Stadler 1995; Watters, Randall et al. 1999) lends support to two facts that shed light on the mechanism of RIP, namely that RIP occurs multiple times prior to the last premeiotic DNA replication, and that each round of RIP targets only a single strand of DNA.

RIP and MIP most likely act on cytosine

The fact that RIP and MIP act on a single type of base on individual strands of DNA begs the question of whether the C:G to T:A mutations caused by RIP are C to T or G to A mutations. A number of lines of evidence point to RIP mutating C to T, rather than G to A. First, a number of processes, some enzymatically mediated, result in C to T mutations, whereas no pathways that specifically mutate G to A have been described. Second, the only enzyme essential for RIP, RID, bears homology to known cytosine methyltransferases and is orthologous with Masc1, the only enzyme known to affect MIP in *A. immersus* (Malagnac, Wendel et al. 1997; Freitag, Williams et al. 2002). As a corollary, MIP methylates cytosines, so if MIP and RIP are orthologous processes then RIP should also act upon cytosine. However, enzymatic activities of RID and Masc1 have still not been directly observed, though both the Rossignol and Selker laboratories have tried several *in vitro* assays for DNA methylation and mutagenesis.

Pathways for mutation of C to T include (1) deamination of cytosine to uracil by a cytosine deaminase, (2) deamination of cytosine to uracil by a cytosine methyltransferase under S-adenosylmethionine (SAM) limiting conditions which favor deamination of the unstable intermediate over methylation or (3) methylation of cytosine to 5-methylcytosine followed by deamination of 5-methylcytosine to thymine (Selker 1990; Mautino and Rosa 1998; Sharath, Weinhold et al. 2000). Models to distinguish these three modes of mutation based on their response to SAM availability were developed (Mautino and Rosa 1998) and

eventually tested (Rosa, Folco et al. 2004). Everything else being equal, if the level of SAM is decreased or increased, RIP should not be affected if the single mutating enzyme is a cytosine deaminase since its activity should not directly depend on SAM as a cofactor. If the single enzyme is a methyltransferase acting under SAM limiting conditions, however, RIP frequency should be reduced with higher SAM availability since higher levels of the cofactor are thought to promote methylation.

The third hypothesis requires the existence of two enzymes, whose activities should be tightly coupled. If RIP depends on a cytosine deaminase that specifically recognizes 5-methylcytosine that is generated by a specific cytosine methyltransferase, then RIP frequency should increase with SAM availability as higher levels of SAM result in higher levels of 5-methylcytosine in *N. crassa* (Foss, Roberts et al. 1993; Roberts and Selker 1995; Barra, Mautino et al. 1996; Mautino, Barra et al. 1996). Again, the deaminase in the second step would not be affected since its activity should not depend on SAM as a cofactor. Alternatively, the level of SAM would not affect RIP if the limiting step in the reaction is transfer of the methyl group rather than availability of SAM (Mautino and Rosa 1998), as has been shown for the bacterial and human cytosine methyltransferases M.HpaII (Svedruzic and Reich 2004) and DnmtI (Svedruzic and Reich 2005). Formation of uracil by cytosine methyltransferases has been observed in the bacterial cytosine methyltransferases M.HpaII (Shen, Rideout et al. 1992; Bandaru, Wyszynski et al. 1995; Shen, Zingg et al. 1995) and M.SssI (Bandaru, Wyszynski et al. 1995) but not M.MspI (Zingg, Shen et al. 1998) when the cofactor SAM is limiting or when cofactor analogues are supplied (Zingg, Shen et al. 1996). In the mammalian cytosine methyltransferase Dnmt1, strong interactions between guanine and the CpG recognition site of the enzyme are thought to prevent premature release of the base and solvent access to the active site which leads to deamination (Svedruzic and Reich 2005).

In crosses with mutants that had highly elevated levels of SAM, RIP frequency is reduced, consistent with a model involving a cytosine methyltransferase operating under SAM limiting conditions, which results in formation of an unstable intermediate susceptible to mutation (Mautino and Rosa 1998; Rosa, Folco et al. 2004). However, SAM is important for many metabolic processes in the cell and it cannot be ruled out that the effect is pleiotropic in nature. Additionally, while Rosa and colleagues measured SAM levels in hyphae

during their experiments, it is unknown how SAM levels are affected in the sexual tissue where RIP occurs. The mechanism of action of the Masc1/RID fungal-specific cytosine methyltransferase family has been investigated but no unambiguous enzymatic activity has been detected in assays with purified enzymes produced in bacteria or insect cells (J. L. Rossignol, G. Faugeron, E. U. Selker, G. O. Kothe and M. Freitag, unpublished data). Clearly, additional biochemical reconstitution experiments, including extracts from premeiotic tissues, are required to address the actual enzymatic activity of RID and Masc1.

All three potential mechanisms for RIP outlined above require either one (in the case of a T:G mismatch) or two (in the case of a U:G mismatch) cycles of replication without repair to permanently fix the mutations. *Neurospora* has a uracil DNA-glycosylase as well as all required mismatch repair proteins, though their functionality in the context of U:G, U:A and T:G mismatches has not been directly assayed (Borkovich, Alex et al. 2004). Also, for RIP to work well, these repair pathways must either be globally reduced during the sexual cycle or specifically blocked at sites of mutagenesis. Thus, hypothesis one and two, which both result in U:G mismatches would be further supported if inactivation or over-expression of the uracil DNA-glycosylase affects RIP frequency, whereas hypothesis three would be supported if inactivation or over-expression of the mismatch repair machinery affects RIP frequency.

Mutations in a group of orthologous fungal-specific cytosine methyltransferase-like genes result in enigmatic phenotypes

During the effort to understand cytosine methylation in *A. immersus*, a cytosine methyltransferase-like gene, *masc1*, was identified that when disrupted, reduces MIP in heterozygous crosses (Malagnac, Wendel et al. 1997). This result also suggested Masc1 is the enzyme responsible for methylating cytosines within repetitive DNA as the title of the original paper suggested "*A gene essential for de novo methylation and development in Ascobolus reveals a novel type of eukaryotic DNA methyltransferase structure*" (Malagnac, Wendel et al. 1997). Nevertheless, effects of *masc1* mutation in homozygous crosses could not be tested because these crosses are sterile. Curiously, the block in development occurs before MIP was thought to occur, as after fertilization apothecia are formed and a normal abundance of plurinucleate ascogonial cells is present but croziers with dikaryotic cells are

not produced. This suggests a block between fertilization and pairing of the nuclei of opposite mating type (Malagnac, Wendel et al. 1997).

With the genome sequence of *N. crassa* came the identification of a well conserved homolog of *masc1*, the cytosine methyltransferase-like gene, *rid*, which when disrupted in a homozygous cross results in complete loss of RIP (Freitag, Williams et al. 2002). The ortholog of *rid* in the closely related pseudohomothallic species, *N. tetrasperma*, is also essential for RIP (Bhat, Tamuli et al. 2004) as it is in a more distantly related homothallic sordariomycete, *F. graminearum* (teleomorph: *Gibberella zeae*; see Chapter 2). Inactivation of *rid* does not result in sexual defects, either in development or abundance of ascospores, in any of these three species. Interestingly, the *rid* homolog in *Magnaporthe grisea* is upregulated after DNA damage and is not essential for genomic cytosine methylation in this fungus (Junhyun 2012), however, the requirement of the *rid* for sexual development and RIP in *M. grisea* has yet to be tested.

Aspergillus nidulans, a fungus without active RIP or MIP and without cytosine methylation also has an ortholog of *rid* and *masc1*, named *dmtA* , which, like in *A. immersus* is essential for sexual development (Lee, Freitag et al. 2008). The Aspergilli may represent a case where MIP or RIP was present at one point during evolution as prior analysis of transposons identified regions of the genome that may have been mutated by RIP (Nielsen, Hermansen et al. 2001; Clutterbuck 2004). Nevertheless, if extant, RIP, MIP or cytosine methylation must be very infrequent, as many laboratories have searched intensively for RIP and cytosine methylation in *A. nidulans* over the past 30 years.

High-throughput genome sequencing of many representatives from diverse groups of fungi has made comparative genomics within the fungi, which have relatively small genomes for eukaryotes, a straightforward process. The Joint Genome Institute has made available data generated by cluster analysis for the publically available fungal genomes which aids in the identification of homologous genes (Grigoriev, Nordberg et al. 2012). Cluster analysis places the RID/Masc1 homologs into a single orthologous group present in most if not all of the sequenced species within the Pezizomycotina (Fig. 1.1). The only published genome with a RID homolog that is basal to the Pezizomycotina is from the fission yeast *Saitoella complicata* (Nishida, Hamamoto et al. 2011), a predicted protein that most closely

resembles homologs from the class Eurotiomycete within the Pezizomycotina and may represent a case of horizontal gene transfer. However, few genomes from non-Pezizomycotina Ascomycetes aside from *Saccharomyces* and *Schizosaccharomyces* have been sequenced. Interestingly, *C. cinereus* the only Basidiomycete in which MIP has been reported does not have a cytosine methyltransferase that belongs to the RID/Masc1 family. It would be interesting to see if premeiotic methylation in *C. cinereus* is dependent on either of its two Masc2 homologs.

Unfortunately, even though extensive efforts have been undertaken, enzymatic activity of Masc1 or RID has never been detected, either in extracts or with purified proteins expressed in bacteria or insect cells. It has been proposed that Masc1 may be necessary for sexual development because repetitive genes need to be subjected to MIP in order for sexual development to proceed, which would be similar to genomic imprinting in higher eukaryotes (Malagnac, Wendel et al. 1997). However, homozygous crosses of *A. nidulans dmtA* mutants are also sterile, but neither RIP nor MIP, or indeed any DNA methylation are detectable in this fungus (Lee, Freitag et al. 2008), arguing that DmtA and Masc1 are playing a role besides silencing of repeats that is necessary for fertility. Another hypothesis is that MIP is necessary to counteract deleterious homologous recombination between repeats on non-homologous chromosomes as duplications methylated by MIP are several 100-fold less susceptible to crossing over during meiosis (Maloisel and Rossignol 1998). This hypothesis may also explain why *rid* homologs which mutate rather than methylate duplicated DNA are dispensable for sexual development since, while methylation has to be constantly established to block homologous recombination, mutation may permanently block recombination. Thus loss of *rid* would not have an effect on sexual development as all the “repeats” are mutated to the point where they are less susceptible to homologous recombination.

RIP is not dependent on the homologous recombination machinery

Based on mutation patterns and the behavior of strains with two *versus* three unlinked DNA segments, it was hypothesized that duplications are detected during both MIP and RIP by a DNA pairing process (Faugeron, Rhounim et al. 1990; Selker 1990). To that end a number of genes important for meiotic pairing of homologous DNA having been tested for a

potential role in RIP. The first mutant tested for effects on RIP was *mei-2*, a mutant defective in pairing of homologous chromosomes, and was found to have normal levels of RIP (Foss and Selker 1991). Later, both RIP and premeiotic recombination were tested in *rec-2* backgrounds, which modulate the frequency of recombination during meiosis. This study found that while *rec-2* affects premeiotic recombination frequency, RIP frequency was unchanged (Bowring and Catcheside 1993). The correlation between RIP and premeiotic recombination was further investigated for direct and inverted tandem repeats, as well as genetically unlinked repeats (Irelan, Hagemann et al. 1994). In direct tandem repeats the frequency of recombination (accompanied by eviction of one copy) and RIP are similarly high (i.e., both occur very frequently). Recombination frequency between inverted and unlinked duplications, however, is reduced many orders of magnitude while RIP frequency remains high (Irelan, Hagemann et al. 1994). This disparity in recombination frequency is most harmoniously explained if premeiotic recombination between direct tandem repeats occurs via the single-strand annealing pathway described first in budding yeast (Sugawara and Haber 1992; Lyndaker and Alani 2009). Repair of DNA breaks by single-strand annealing only causes deletions in direct tandem repeats since filling in using a homologous template only occurs in the 5'->3' direction (Fig. 1.2). In Chapter five, I tested a number of genes homologous to those involved in single-strand annealing in yeast and found that RIP frequency is not affected. These studies all suggest that the mechanism of RIP does not involve a recombination intermediate that is generated in the same way as in models based on yeast, but they do not unequivocally exclude the possibility that single-strand intermediates are formed by unknown mechanisms.

A potential link between RIP and premeiotic recombination was again investigated when a *P. anserina* mutant affecting nuclear distribution, *ami1*, was also found to affect both RIP and premeiotic recombination frequency (Bouhouche, Zickler et al. 2004). In Chapter four we show that two mutants in *N. crassa* have a similar phenotype. Interestingly, the French group found that delayed sexual development resulted in increased RIP and premeiotic recombination efficiency (Bouhouche, Zickler et al. 2004), which is consistent with RIP occurring multiple times prior to karyogamy and meiosis. These studies confirmed earlier reports from *N. crassa*, where Selker and colleagues found that ascospores shot

relatively early (~14 days post fertilization) had fewer mutations in duplicated test regions than “late ascospores” collected ~30 days post fertilization (Singer, Kuzminova et al. 1995). In combination, these data suggest that RIP and premeiotic recombination both occur multiple times during the sexual cycle, possibly during a particular phase of replication or nuclear division, however, homologous recombination and RIP may not be mechanistically linked. Single-strand annealing is dependent on single- or double-strand break to initiate the process (Fig. 1.2), which suggests that breaks occur frequently during sexual development since in *N. crassa* hallmarks of single-strand annealing are observed in approximately a fifth of the progeny from a cross with a tandem duplication (Irelan, Hagemann et al. 1994) (and Chapter 4) but infrequently during vegetative growth. Whether RIP or MIP are initiated by DNA breaks is unknown.

Effects of mating type on RIP frequency

It has not escaped notice that duplications are more or less susceptible to RIP depending on the mating type of the strain harboring the duplication. In three independent mapping populations generated by crossing three independently isolated semi-dominant RIP suppressor mutants (N1516, N1522 and Adiopodoumé) to wildtype strains, the progeny that display the highest degree of suppression are invariably of a single mating type (our data in Chapter 4, Noubissi, McCluskey et al. 2000; Bhat, Noubissi et al. 2003, (Tamuli and Kasbekar 2008). There are at least two explanation for this observation: (1) the strains of opposite mating type used to test RIP frequency have intrinsic differences in their repeats, and/or (perhaps more interestingly) (2) the mating type idiomorph or relatively closely linked loci act as quantitative trait loci or “modifiers” for RIP frequency. Hypothesis (1) has evidence to support it as deviations in the nucleotide sequence between repeats decreases RIP frequency (Cambareri, Singer et al. 1991) and sequencing of repeat alleles from RIP tester strains of opposite mating type used in previous publications (Rosa, Folco et al. 2004) has uncovered mutations. Hypothesis (2) can easily be imagined since the *mat* idiomorphs control many aspects of sexual development and may subject nuclei of a certain mating type to fewer or more rounds of RIP. In support of hypothesis (2), mutation of the *mat* genes in *P. anserina*, results in decreased RIP efficiency in biparental progeny as well as decreased fertility

(Arnaise, Zickler et al. 2008). This is most easily explained if the *mat* mutations decrease the duration or the number of times during which the RIP step occurs. Whether the *mat* idiomorphs act as a quantitative trait locus has not been tested and will require careful construction of strains of opposite mating type without a bias in their repeat alleles.

Effects of translesion polymerases on RIP frequency

Another hypothesis to explain RIP is that mutations are made by inaccurate translesion polymerases after damage due specifically to the nature of repetitive DNA. However, the translesion polymerases *pol zeta*, *pol eta*, *pol iota*, *pol kappa* and *rev1* were found to be dispensable for RIP (Tamuli, Ravindran et al. 2006). The dominant RIP defect present in the Adiopodoumé strain had been mapped near *mat*, presumably to *pol zeta* (Tamuli and Kasbekar 2008). The genetic data provided in the paper, and the manner in which the complementation was performed (by back-crossing transformants to an Oak Ridge strain instead of the isogenic *mat a* Adiopodoumé strain), did not rule out the possibility that the dominant suppressor present in Adiopodoumé may indeed be located on the distal side of the *mat* idiomorph. The authors essentially retracted the idea of involvement of translesion polymerases when true complementation assays proved unsuccessful.

Severity of RIP in heterozygous crosses with wild *N. crassa* isolates

The RIP frequency obtained by crossing isolates of *N. crassa* from around the world to a well characterized RIP tester strain suggests that RIP competence is a quantitative trait with a wide range (Noubissi, McCluskey et al. 2000). These studies also confirmed again that RIP frequency is higher in ascospores collected from a cross later in time regardless of genetic background, confirming that RIP is progressive and may occur multiple times during a cross (Singer, Kuzminova et al. 1995). Interestingly, standard lab strains of the Oak Ridge background are typically very competent for RIP when compared to most wild collected *N. crassa* strains, while only a few wildtype strains dominantly suppress RIP, resulting in few *erg-3* mutations that have phenotypic consequences (Noubissi, McCluskey et al. 2000; Noubissi, Aparna et al. 2001; Bhat, Noubissi et al. 2003). Unfortunately, mapping of the dominant loci in strains that suppress RIP has thus far proven difficult, potentially because

multiple loci are necessary for dominance or semi-dominance. Based on the hundreds of crosses with wildtype strains, compared to crosses within the Oak Ridge domesticated laboratory lineage, we should consider the notion that *Neurospora* geneticists may have unwittingly selected for improved RIP by selecting for fast sexual development during the decades the Oak Ridge lineage has been used as laboratory reference strains. This idea can now be tested by high-throughput sequencing and QTL mapping approaches.

Distribution and severity of RIP in the fungi: experimental data

Outside of *N. crassa* and *A. immersus*, active RIP and MIP characterized by hyper-mutation or methylation of repeats during the sexual cycle, respectively, have been experimentally demonstrated in only a handful of fungi but may be widespread among sexual fungi. RIP was also confirmed in *P. anserina* where unmethylated but heavily mutated *Pat* and *Yeti* transposons were observed (Hamann, Feller et al. 2000; Hamann, Feller et al. 2000). A single cDNA from multiple copies of the ectopically inserted *Grisea* gene was sequenced and found to contain a mere three C to T mutations within 1,100 bp analyzed; these mutations were ascribed to RIP (Hamann, Feller et al. 2000). Subsequently, RIP was suggested to have introduced 106 C:G to T:A mutations in a large, 40 kb multi-copy cosmid, but only after four rounds of sexual reproduction (Graia, Lespinet et al. 2001), suggesting that RIP is less efficient in *P. anserina* than *N. crassa*. Analysis of the LTR-retrotransposon *Maggy* present in *M. grisea* identified multiple copies diverged by predominantly, but not exclusively, C:G to T:A mutations characteristic of RIP (Nakayashiki, Nishimoto et al. 1999; Thon, Martin et al. 2004). Active RIP in this fungus was later confirmed through the analysis of inserted *Maggy* and *hph* sequences (Ikeda, Nakayashiki et al. 2002). The next fungus in which active RIP was identified was *Leptosphaeria maculans* when researchers observed inactivation of multiple copies of *hph* after going through a cross. Sequencing identified many C:G to T:A mutations, characteristic of RIP, in the multi-copy genes, again no DNA methylation was associated with the mutated alleles (Idnurm and Howlett 2003). This result demonstrated active RIP for the first time outside of the Sordariomycetes, within the Dothideomycetes. RIP has also been confirmed to be active in a pseudohomothallic species of *Neurospora*, *N. tetrasperma*, and like in *N. crassa*, it is dependent on *rid* (Bhat, Tamuli et

al. 2004). In the homothallic fungus, *F. graminearum*, RIP was confirmed by analyzing its effect on single-copy and duplicated *hph* genes (Cuomo, Guldener et al. 2007). This fungus proved to RIP quite efficiently and generates mutations primarily in CpG and CpA di-nucleotide contexts, similar to *N. crassa*. RIP in a related species, *Nectria haematococca*, was also confirmed by analyzing its effect on duplicated *hph* genes, which were efficiently mutated to the point of inactivation (Coleman, Rounsley et al. 2009). Thus far RIP has been demonstrated only in fungi from the Pezizomycotina, filamentous ascomycetes that all have a single ortholog of *rid*/masc1. MIP, or a system similar to MIP, has thus far only been described in *A. immersus and* the basidiomycete *C. cinerea* but in-depth studies with species other than *A. nidulans* have not been carried out.

Distribution and severity of RIP in the fungi: computational prediction of past RIP

Ancient RIP in many fungi has been predicted based on divergence of repetitive DNA in the genome. Early observations of active RIP uncovered bias toward mutating C in the CpA and CpG dinucleotide contexts in *N. crassa* (Cambareri, Jensen et al. 1989), *P. anserina* (Graia, Lespinet et al. 2001) and *F. graminearum* (Cuomo, Guldener et al. 2007) or in CpA and CpT dinucleotides in *M. grisea* (Ikeda, Nakayashiki et al. 2002). These biases are not complete, as RIP mutates C in all dinucleotide contexts. The mutation bias toward CpA in *N. crassa* was used to develop RIP indices, which are the ratios of the number of TpA over ApT (TpA/ApT), measuring introduction of TpA dinucletides with a correction for A:T-richness (the "product" index), or (CpA+TpG)/(ApC+GpT), measuring depletion of CpA dinucleotides (the "substrate" index) (Margolin, Garrett-Engele et al. 1998). RIP indices can be calculated for sequences in isolation and do not require prior knowledge of a repeat family. The product and substrate index, and a combined index, have been used for genome-wide analyses (Lewis, Honda et al. 2009). A more powerful strategy for determining whether RIP has occurred is to compare multiple sequence alignments of mutated paralogs to an unmutated ancestor sequence which allows determination of the di-nucleotide context of the mutations (Hane and Oliver 2008). This method has been implemented in the RIPCAL software (Hane and Oliver 2008).

RIP indices, RIPCAL and similar strategies have been used to identify putative relics of RIP in many fungi, including some for which a sexual state has never been observed. The presumably asexual fungus *Fusarium oxysporum* has many C:G to T:A mutations between paralogs in its *fot1* transposon family suggesting ancient RIP (Hua-Van, Hericourt et al. 1998; Daboussi, Daviere et al. 2002). Similarly, many C:G to T:A mutations between copies of the *Afut* transposons in *Aspergillus fumigatus* led researchers to speculate that this "asexual" fungus had once the capacity for RIP (Neuveglise, Sarfati et al. 1996). More recently, the sexual state was described lending support to a RIP-like process in *A. fumigatus* (O'Gorman, Fuller et al. 2009). RIP-like mutations are also present between transposons of the presumably asexual fungi *Aspergillus niger* and *Penicillium chrysogenum* (Braumann, van den Berg et al. 2008) suggesting that these fungi either had or still have an undescribed sexual cycle. It needs to be pointed out that in most, if not all, of these cases, the telltale C:G to T:A transition mutations are not the only mutations observed, unlike in RIP that has been observed experimentally after one or a few crosses. The explanation for the appearance of transversions is commonly that duplicated segments can "drift" once inactivated.

In sexual fungi accumulations of RIP-type mutations above that expected by random mutation have been observed in *A. nidulans* within *Dane* and *Mate* transposons (Nielsen, Hermansen et al. 2001; Clutterbuck 2004), in *N. haematococca* within *copia* elements (Shiflett, Enkerli et al. 2002), in *L. maculans* between *Pholy* LTR retrotransposons (Attard, Gout et al. 2005), as well as throughout about a third of the highly repeat-rich genome (Rouxel, Grandaubert et al. 2011). RIP-type mutations have been observed in the basidiomycete *Microbotryum violaceium* where a few *copia* and *helitron* transposable elements display C:G to T:A mutations (Hood, Katawczik et al. 2005; Johnson, Giraud et al. 2010), in *Aspergillus oryzae* within *Tan1* transposons (Montiel, Lee et al. 2006), in *Ophiostoma ulmi* within *Fot1*/*pogo* transposons (Bouvet, Jacobi et al. 2007), in *Colletotrichum cereale* within a number of types of transposons between two clades of the fungus (Crouch, Glasheen et al. 2008), in *Stagonospora nodorum* between *Pholy* retrotransposons and within the rDNA repeats (Attard, Gout et al. 2005; Hane and Oliver 2008), in *Coccidioides immitis* and *Coccidioides posadasii* throughout the repetitive regions of the genomes (Neafsey, Barker et al. 2010), in *Mycosphaerella graminicola* between recently

multiplied paralogs of *dim2* (Dhillon, Cavaletto et al. 2010), in *Grosmannia clavigera* throughout the repetitive regions of the genome (DiGuistini, Wang et al. 2011), in the *Gypsy* elements in *Trichoderma atroviride* as well as the *Copia* and *Tad1* elements in *Trichoderma virens* and *Trichoderma reesei* (Kubicek, Herrera-Estrella et al. 2011), in the genomes of 48 fungi with whole genome sequences available in the subdivision Pezizomycotina (Clutterbuck 2011) and within the four Pucciniomycotina *Microbotryum lychnidis-dioicae*, *Puccinia graminis*, *Melampsora laricis-populina* and *Rhodotorula graminis* but not within the Agaricomycotina or Ustilaginomycotina (Horns, Petit et al. 2012).

It is worth noting that in the vast majority of fungi where putative relics of RIP have been reported the sole type of evidence is divergence between homologous repeat sequences where the divergence is primarily C:G to T:A mutations. Mutation in particular di- or trinucleotide contexts are often stated as evidence for RIP since genuine RIP has been observed to occur preferentially in the CpG and CpA dinucleotide contexts, though it does mutate cytosines in all dinucleotide contexts (Cambareri, Jensen et al. 1989; Cuomo, Guldener et al. 2007). As nonenzymatic deamination of 5-methycytosine to thymine is the cause of the perhaps most common type of mutation, C to T, it would not be surprising that this signature can appear similar to RIP after millennia of evolution. Many organisms have cytosine methyltransferases that target specific sequence contexts for methylation, such as the CpG dinucleotide context in *H. sapiens* and *A. thaliana* (Law and Jacobsen 2010) and Jacobsen, 2010). Depletion of C:G due to nonenzymatic deamination is a slow process that occurs over many generations, whereas RIP occurs at a specific point in time during the sexual cycle at a rate many orders of magnitude higher. In the case of C:G depleted repeat sequences, sequence analyses cannot unambiguously separate these two processes. It is likely that ancient duplications mutated by RIP have undergone nonenzymatic deamination of 5-methylcytosine that was triggered by the original RIP events in *N. crassa*, as remaining cytosines are often targeted for *de novo* methylation by DIM-2. Without direct evidence of high mutation rates in multi-copy DNA during the sexual cycle attributing C:G depletion to RIP is a guess at best.

Evolutionary effects of RIP

The mutation of repetitive DNA by RIP can be quite destructive to coding sequences. RIP and the subsequent silencing by DNA methylation observed in *N. crassa* likely have evolved as a genome-wide defense mechanism against transposable elements (Selker 1990). This idea has been tested directly by using a strain of *N. crassa* from Adiopodoumé in the Ivory Coast, which harbors an active LINE-like retrotransposon, *Tad* (Kinsey 1989; Kinsey 1990). In crosses the active *Tad* element is mutated and inactivated by RIP (Kinsey, Garrett-Engele et al. 1994; Anderson, Tang et al. 2001), supporting the idea that RIP actively defends the genome from mobile elements (Galagan, Calvo et al. 2003; Selker, Tountas et al. 2003; Lewis, Honda et al. 2009). A defense mechanism that targets transposable elements based on sequence homology between duplications begs two immediate questions: (1) Can the fungus evolve by gene duplication, and (2) how can the fungus maintain its rDNA, which generally constitutes numerous tandem repeats that, at ~7 kb for each unit, are well above the minimum size to trigger efficient RIP?

As one answer to the first question, RIP has been considered an evolutionary dead end (Galagan and Selker 2004). Yet, RIP has now been experimentally shown to occur in a phylogenetically diverse range of fungi. The occurrence of RIP in many different species suggests that it does not present a new way to protect or evolve the genome. This may be explained in a number of ways. First, RIP only occurs during the sexual cycle, thus efficient evolution and selection can occur during vegetative growth (Selker, Cambareri et al. 1987). Second, RIP is light overall in some fungi such as *P. anserina* and *M. grisea* (Graia, Lespinet et al. 2001; Ikeda, Nakayashiki et al. 2002) as well as many wild isolates of *N. crassa* (Noubissi, McCluskey et al. 2000; Noubissi, Aparna et al. 2001) and it results in fewer mutations in the earliest ascospores released (Singer, Kuzminova et al. 1995), which allows a spectrum of RIP to occur where the most fit offspring after sex will presumably be those that were mutated just enough to inactivate any invading transposons without mutating the genome as a whole. Even in strains of *N. crassa* and *F. graminearum* in which RIP is efficient, there are always some non-mutated offspring released. Third, the activity of RIP in some regularly transformed fungi, including the aspergilli is undetectable even though the machinery for RIP, at least to the extent that we know about it, appears intact (Lee, Freitag et al. 2008) and the genomes all appear to carry relics of RIP from the past (Neuveglise, Sarfati et al. 1996;

Clutterbuck 2004; Montiel, Lee et al. 2006; Novikova, Fet et al. 2007; Braumann, van den Berg et al. 2008; Clutterbuck 2011). These results suggest that low-level RIP can be selected for. Fourth, closely linked repetitive DNA is not only susceptible to RIP but is also prone to homologous recombination (Selker, Cambareri et al. 1987; Irelan, Hagemann et al. 1994). The ability of repeats to change in copy number sets up a scenario where duplication and hyper-evolution can occur followed by recombination back to a single copy, a scenario which has some evidence for occurring in *N. crassa* (Watters and Stadler 1995), and is relevant for the large cluster of rDNA genes that have been shown to "swell" and "shrink" over the sexual cycle (Butler and Metzenberg 1989; Butler and Metzenberg 1993). Finally, comparative genomics gives us the chance to ask whether "duplication, divergence and deletion" exists and is of evolutionary consequence. Such a scenario may result in short-term expansion of novel gene families, followed by selection and expunging of extra copies by recombination, a feature that many of the current models for RIP lack (Brookfield 2003; Galagan and Selker 2004). At this point we can speculate that this mode of evolution is not important to fungi that use RIP, as they seem to thrive in nature without the presence of extensive gene families.

If fungi that RIP disallow accumulation of closely related gene families, how can they maintain normal rDNA clusters, which are needed for efficient translation? Early on it was realized that either (1) rDNA is somehow actively protected from the RIP machinery, or (2) the rDNA is mutated by RIP like any other genomic region but efficiently repaired to a pre-RIP state. In *N. crassa*, the rDNA cluster regularly changes size, but we have not found much enrichment of RIP-like mutations in the rDNA on linkage group VL by whole genome sequencing (~2x enrichment of C:G to T:A mutation over other types). Most strains have between 150 and 200 copies of the rDNA which are often shorted during sexual development at roughly the same time that tandem duplications undergo intrachromosomal recombination and RIP (Selker, Cambareri et al. 1987; Butler and Metzenberg 1989), suggesting the rDNA repeats are subject to the same instability as the rest of the genome and are not intrinsically protected.

When the rDNA repeats are shortened below their optimum number they expand and can do so mitotically and meiotically, likely through unequal exchange by homologous

recombination (Butler and Metzenberg 1993). This provides a scenario where the average rDNA can undergo RIP and is mutated. These mutated copies would be less able to undergo intrachromosomal recombination as this process is sensitive to homology between repeats, which is diminished by RIP (Cambareri, Singer et al. 1991). Random intrachromosomal recombination could then selectively remove mutated rDNA by bypassing nonhomologous sequences. The small rDNA cluster would eventually be brought back to its optimal size by unequal exchange during mitosis, which has been observed during vegetative growth of hyphae (Rodland and Russell 1982; Russell and Rodland 1986).

In *A. immersus*, the tandem 9 kb repeats in the rDNA cluster show incomplete methylation (Goyon, Rossignol et al. 1996), compared to tandem duplications of similar size that are heavily methylated (Barry, Faugeron et al. 1993). This suggested that the rDNA is protected from MIP or is selectively demethylated following MIP. There is a large body of literature that shows that similar phenomena occur in mammalian and plant rDNA clusters (Preuss and Pikaard 2007).

Evolutionary effects of RIP beyond transposons and rDNA

RIP has the potential to affect genomes of fungi beyond being a simple defense against mobile elements. It has recently been observed to affect avirulence genes in field studies with isolates of the plant pathogen *L. maculans*, which, unlike the majority of coding genes, are dispersed within closely related repetitive segments of the genome (Rouxel, Grandaubert et al. 2011). In these regions, any gene would be susceptible to RIP and interchromosomal recombination between flanking repeats (Idnurm and Howlett 2003; Fudal, Ross et al. 2009) as had been suggested by the early studies in *N. crassa* (Foss, Garrett et al. 1991; Irelan, Hagemann et al. 1994). The combined effects of RIP and (likely premeiotic) recombination were dramatically demonstrated in Australia in 2003, when resistance of canola crops (*Brassica napus*) to *L. maculans* broke down because of inactivation of avirulence genes (Van de Wouw, Cozijnsen et al. 2010). This suggests that repeat-rich regions of the genome, which are often considered a graveyard for mobile elements, can also act as regions for accelerated mutation of genes.

Mutations caused by RIP can affect the overall biology of an organism. For instance, it appears that a mobile element acquired and amplified part of the *dim2* gene, encoding a cytosine methyltransferase (Fig. 1.1), in the Dothideomycete *M. graminicola*. This event occurred fewer than 10,500 years ago and has caused this lineage to lose cytosine methylation in its genome when compared to its close relatives *M. fijiensis* and *Septoria passerinii*, which each have a single functional copy of *dim2* (Dhillon, Cavaletto et al. 2010). Whether this has or will result in deleterious effects on the fitness of the organism remains unknown.

Whole genome sequencing of fungi is unraveling the evolutionary impacts of RIP. The first high-quality draft sequence of *N. crassa* was published in 2003 and revealed that this fungus has a relatively compact genome, especially when compared to higher eukaryotes. *Neurospora* genomes thus have few identical repeats, most "repetitive sequences" are highly mutated transposon relics, and thus there are essentially no multigene families, presumably caused by RIP (Galagan, Calvo et al. 2003). Interestingly, most of the repeats have been compartmentalized into large blocks of centromeric DNA and numerous smaller, 10-20 kb regions at subtelomeres and interspersed euchromatic regions (Lewis, Honda et al. 2009; Smith, Phatale et al. 2011). At the centromeres, DNA segments mutated by RIP may form an essential signal for proper nuclear division in *N. crassa* (Smith, Phatale et al. 2011). Now and in the future, the effects of RIP are being studied by comparison of hundreds of sequenced fungal genomes. It is likely that RIP plays an important role in the streamlining and maintenance of small genomes of filamentous ascomycetes.

Unanswered questions about RIP

A number of mechanistic questions remain with regard to the processes that target repetitive DNA for inactivation, including: (1) how do RIP and MIP specifically recognize and target duplications, (2) what is the precise mechanism of mutation by RIP, (3) how wide-spread is RIP phylogenetically, and (4) what are the long-term evolutionary impacts of RIP?

In this thesis I have attempted to address all of these questions, first by characterizing RIP in the homothallic fungus *F. graminearum* and then by trying to adapt this species to hunt for recessive RIP mutants (Chapter 2), a challenging task that proved in the

end to be more difficult than working with *N. crassa*. I generated molecular and computational tools for mapping of mutations by high-throughput sequencing when I identified the molecular nature of the defect in the single known *Neurospora* cell cycle mutant, *ndc-1* (Chapter 3). This was followed by characterization and mapping of two previously isolated semi-dominant RIP mutants in *N. crassa* (Chapter 4). The nature of the mutants led to a reverse genetics screen to identify more genes involved in duplication detection during the RIP process (Chapter 5). To yield insights into complicated interactions between RIP genotypes and phenotypes I also began to map potential quantitative trait loci in the Adiopodoumé strain (Future directions). Finally, while the evolutionary history of RIP was inferred from presence of RID homologs in sequenced fungal genomes, insight into the mutational mechanism of RIP will be achieved by mutational analysis of *rid* and *masc1* in *N. crassa* (Future directions).

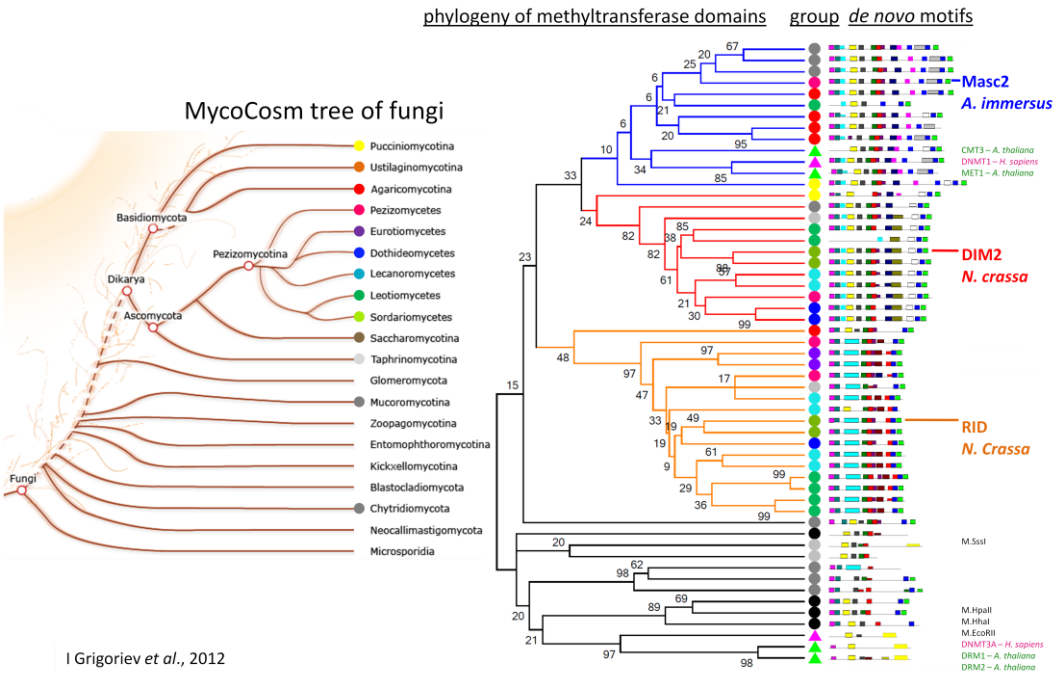


Figure 1.1. Phylogenetic analysis of DNA cytosine methyltransferases in fungi. Three flavors of cytosine methyltransferase, here called Masc2, DIM2 and RID are observed in the Ascomycetes and Basidiomycetes. Of interest is a yeast, *S. complicata*, which is the only organism outside the Pezizomycotina known to have a RID homolog. *A. immersus* has homologs of both the DNMT1 and DIM2 methyltransferases suggesting they are independent families, though DNMT1 may have been lost in most lineages since, even in *A. immersus* DNMT1 seems to be unnecessary for maintenance of genomic cytosine methylation. DNMT1 and DNMT3A from *H. sapiens* and MET1, CMT3, DRM1 and DRM2 from *Arabidopsis thaliana* are included as pink and green triangles respectively. Bacterial cytosine methyltransferases are black circles. DNA cytosine methyltransferases were identified by reciprocal BLAST analysis from two genomes each from each group of fungi with a grayscale or colored circle available at JGI's MycoCosm site. The methyltransferase domains were identified by InterProScan and aligned by Muscle. Evolutionary history was inferred from 283 amino acid positions using 500 bootstraps of the Maximum likelihood method. *De novo* motif analysis of the methyltransferase domains was performed using Meme.

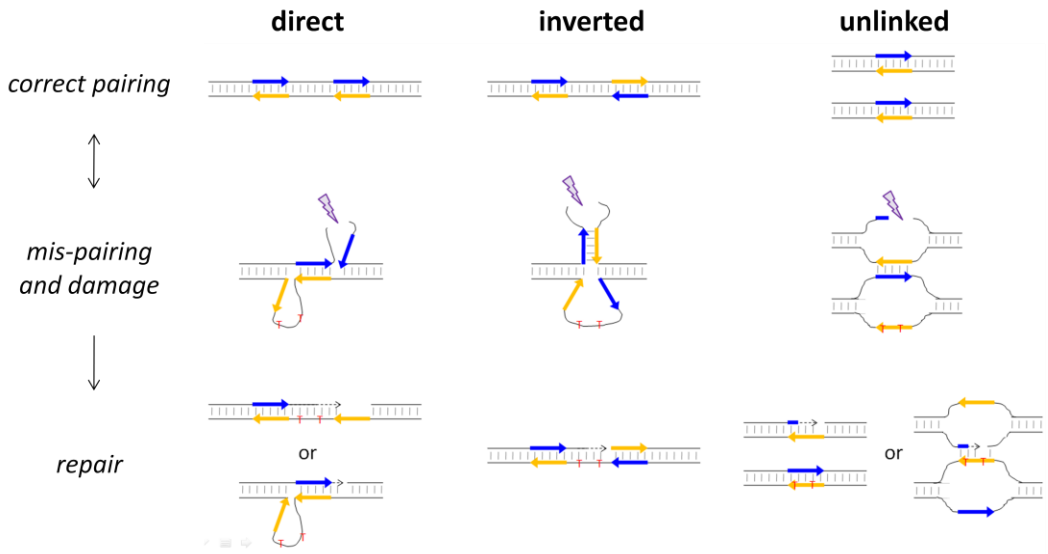


Figure 1.2. Hypothesis for mis-pairing of duplicated DNA as the substrate for RIP, MIP and premeiotic recombination. DNA duplications present as direct and inverted linked repeats or as unlinked repeats have the potential to form similar secondary structures with looped out single-stranded DNA when mis-paired that may be susceptible to methylation and mutation in the case of MIP and RIP or damage as a single- or double-strand break. While the effects of MIP and RIP are predicted to remain in all three types of duplications after repair, a deletion is only predicted to be observed in direct linked repeats when repaired by a process akin to single-strand annealing. Red "T" indicates a RIP mutation while purple lightning indicates a single strand break. Base pairs are not drawn to scale.

Characterization of RIP in *Fusarium graminearum*

Kyle R. Pomraning, Lanelle R. Connolly, Joseph P. Whalen, Kristina M. Smith and Michael Freitag

Parts of this chapter are included in a manuscript accepted for publication as Chapter 6 in *"Fusarium genomics and molecular and cellular biology"* by Robert Proctor and Daren Brown (eds.), Horizon Scientific Press.

Abstract

Multiple mechanisms control genome stability in fungi. To limit the expansion of repeated transposable elements, some filamentous ascomycetes make use of a duplication-dependent mutator system, called "Repeat-Induced Point mutation" (RIP). RIP detects gene-sized duplications and finds homologous copies by an unknown homology search mechanism during sexual development in *Neurospora crassa*. C:G to T:A transition mutations are then introduced, typically into both copies of the DNA segments. In 2007, RIP was first described in the homothallic Ascomycete *Fusarium graminearum* (Cuomo, Guldener et al. 2007). Here we further characterize the activity of RIP in *F. graminearum*, show that the *F. graminearum RIP defective 1 (rid1)* homologue is required for RIP, and begin developing *F. graminearum* as a model organism for the study of RIP and identification of new *rid* mutants.

Introduction

Fungal genomes, like those of other eukaryotes, harbor repetitive sequences that can range from short simple sequence repeats (Karaoglu, Lee et al. 2005; Kim, Booth et al. 2008; Labbe, Murat et al. 2010; Murat, Riccioni et al. 2010) to families of transposable elements (Aleksenko and Clutterbuck 1996; Daboussi 1997; Hua-Van, Daviere et al. 2000; Daboussi and Capy 2003; Clutterbuck 2004; Galagan, Calvo et al. 2005; Wortman, Gilsenan et al. 2009), or functional gene duplications to partial or entire chromosome duplications (Miao, Covert et al. 1991). Fungi also can acquire or evolve "dispensable", "lineage-specific", or "B" chromosomes (Miao, Covert et al. 1991; Ma, van der Does et al. 2010; Rep and Kistler 2010). Repetitive sequences within genomes of a single species are plastic. They can move or be moved around, change in copy number and actively contribute to rearrangements of non-repeat sequence (Daviere, Langin et al. 2001; Rep, van der Does et al. 2005; Bergemann, Lespinet et al. 2008; Ogasawara, Obata et al. 2009; Carr, Tuckwell et al. 2010).

Multiple mechanisms to control the expansion of repeats have been discovered in eukaryotes, the first in the ascomycete *Neurospora crassa* by Selker and colleagues (Selker, Cambareri et al. 1987). This fungus utilizes a genome-wide, DNA-based homology search process that occurs only during the sexual cycle after fertilization but before karyogamy (Fig. 2.1; reviewed in Selker 1990). The phenomenon was first spotted by Southern analyses as a

mixture of gene-sized deletions and RFLPs that revealed putative genome instability, and thus initially referred to as “rearrangements induced premeiotically” (Selker, Cambareri et al. 1987). Subsequently it was shown that premeiotic recombination events can be separated from the mechanism responsible for the observed RFLPs and that duplications of gene-sized regions were required (Selker and Garrett 1988). A DNA pairing-dependent protein complex was hypothesized to introduce the many transition mutations, always C:G to T:A, which were observed, and this observation resulted in a new moniker, “repeat-induced point mutation”, or RIP (Cambareri, Jensen et al. 1989). Previously unmethylated DNA segments that underwent RIP were often found to also acquire methylated cytosines (Selker, Jensen et al. 1987; Selker and Garrett 1988; Singer, Marcotte et al. 1995).

RIP is not restricted to *Neurospora*, as additional filamentous fungi, i.e. *Podospora anserina* (Graia, Lespinet et al. 2001), *Magnaporthe grisea* (Ikeda, Nakayashiki et al. 2002), *Leptosphaeria maculans* (Idnurm and Howlett 2003), *F. graminearum* (Cuomo, Guldener et al. 2007), and *Nectria haematococca* (Coleman, Rounsley et al. 2009) have been experimentally shown to produce the hallmarks of RIP, i.e. reporter gene inactivation, generation of RFLPs and/or cytosine methylation after a cross. RIP or MIP has not been discovered outside of the filamentous fungi, which correlates with the distribution of a family of DNA methyltransferase-like genes that appear to be restricted to the filamentous fungi. At the same time, not all filamentous fungi use RIP or MIP, as studies in *Aspergillus nidulans* (Selker 1990; Lee, Freitag et al. 2008), *Sordaria macrospora* (Nowrousian, Stajich et al. 2010) and *Cochliobolus heterostrophus* (Selker 1990) showed no evidence for either process.

What made the discovery of RIP in *F. graminearum* especially interesting was the fact that this species is homothallic, which has repercussions on the genetic diversity of mutated elements (Fig. 2.1). In heterothallic species (e.g. *N. crassa*), fertilization between strains of two different mating types is required to form a heterokaryotic strain. RIP occurs after fertilization but before karyogamy and will generate transition mutations only in the nucleus that carries duplicated copies of gene-sized segments. If a certain element is only present in one of the mating partners, if RIP is severe, and if the gene in question is recessive, the four meiotic products will segregate as two mutant and two wildtype alleles. Conversely, in a homothallic species that does not require a partner, one would expect RIP to act on all copies

present in the original homokaryotic dikaryon to generate heterokaryotic states by differential mutagenesis of all copies (Fig. 2.1). If again RIP is assumed to be severe and the gene is recessive, all recovered strains should be mutant. Thus, in the most extreme case, one would expect that all progeny issuing from a *F. graminearum* cross with duplications in an essential or selectable gene should be mutated and inviable. This is not observed, however, largely because RIP is hardly ever severe enough to affect all alleles to the point of complete inactivation, and segregation during meiosis will generate duplication strains with one active and one inactivated copy.

While we understand RIP as a phenomenon and as a tool for mutagenesis, we still lack insight into the mechanisms of the homology search and pairing mechanism, as well as the details of the actual mutagenesis. As biochemistry is difficult to carry out on the tissues in question, many attempts have been directed towards the isolation of RIP mutants in laboratory settings or by screening of a large library of wild-collected *Neurospora* strains. At least three wild-collected strains appear to be at least partially defective for RIP, one of those is the Adiopodoumé strain (Noubissi, McCluskey et al. 2000), which also carries the only known active *Neurospora* transposon, *Tad*. One of the difficulties inherent in a mutagenesis scheme in the heterothallic *N. crassa* is the fact that most induced RIP mutants will be recessive and thus complemented by the mating partner after heterokaryon formation subsequent to fertilization. This will result in the isolation of alleles that are for the most part dominant or semi-dominant. Homothallic fungi do not need mating partners, thus a second, more utilitarian, aspect of homothallism is worth noting here. In the homothallic *F. graminearum*, mutagenesis should yield strains that, when selfed, may generate progeny in which RIP defects can be directly observed even when alleles are recessive.

The homothallic *F. graminearum* relies on disseminated ascospores to cause head blight of wheat, thus giving it ample opportunity to accumulate mutations caused by RIP in consecutive sexual cycles that are part of its natural annual life cycle. Sequencing of the *F. graminearum* PH-1 genome revealed a number of repetitive elements (e.g. *Fot1fg*) with many RIP-type transition mutations (Cuomo, Guldener et al. 2007). Strains with two copies of the *hph* gene, encoding hygromycin phosphotransferase that confers resistance to hygromycin (hyg), had been generated for a different study and were selfed to

experimentally confirm that RIP occurs. A high percentage (42%) of progeny had no hyg resistance (hygR) when two copies of *hph* were present, while all progeny with a single copy of *hph* retained resistance (Cuomo, Guldener et al. 2007). This suggested that duplicated *hph* genes were highly unstable. The size of the unlinked *hph* duplication was above the limits described for RIP in *Neurospora*, where ectopically inserted tandem and unlinked repeats of >400 and 1000 base pairs, respectively, can be efficiently detected and mutated (Selker 1990; Watters, Randall et al. 1999). The *hph* alleles from five hyg-sensitive (hygS) *F. graminearum* progeny strains were cloned and sequenced to determine whether loss of resistance could be attributed to point mutations or some other cause. In the five alleles, a total of 348 mutations were identified, 99% of which were C:G to T:A transition mutations (Cuomo, Guldener et al. 2007). In addition, these mutations primarily occurred in the CpA dinucleotide context, similar to what had been observed in *N. crassa* (Selker 1990).

We are taking advantage of the homothallic nature of *F. graminearum* to recover recessive RIP defective (*rid*) mutants. We have further characterized RIP in *F. graminearum* by showing RIP is dependent on *rid1* and that it may not result in *de novo* cytosine methylation of the mutated repeats. We have attempted to make RIP mutants in the PH-1 genetic background and sequenced another strain of *F. graminearum*, 00-676-2 for mapping novel mutants.

Materials and methods

Strains, crosses and growth conditions

All strains were cultured in carboxymethylcellulose (CMC) liquid medium at room temperature on a shaker for 1 week to produce conidia. Conidia were inoculated in yeast extract peptone dextrose (YEPD) liquid medium and grown for 36 hours at room temperature to produce tissue for DNA. Selfings were performed by inoculation of carrot agar (35% carrots, 2% agar) plates with conidia followed by collection of ascospores in sterile water after a few weeks. Tests for hygromycin resistance (hygR) and sensitivity (hygS) were performed on YEPD medium with and without 100 μg/mL hyg. Transformations were carried out by protoplasting day old hyphae as previously described (Watson, Burchat et al. 2008) and selecting for transformants on YEPD agar with 100 μg/mL hyg.

F. graminearum strain PH-1 (FMF1, FGSC9075) is our standard wild-type strain and was used for the *F. graminearum* genome sequencing project (Cuomo, Guldener et al. 2007). Strain 00-676-2 (FGSC9603) (Gale, Bryant et al. 2005) was chosen for genome sequencing to use as a mapping strain. FMF200 (Δ*rid1*::*GFP*; *hph*$^{+}$) and FMF201 (*rid1-flag*; *hph*$^{+}$) were constructed by transforming PH-1 with split marker *rid1*::*GFP*, *hph* and *rid1-flag*, *hph* constructs followed by selfing and selecting hygR progeny. FMF204 (*tdimerRed*$^{+}$; *hph*$^{+}$)$^{\text{multi-copy}}$ and FMF205 (*tdimerRed*$^{+}$; *hph*$^{+}$)$^{\text{single-copy}}$ were constructed by transforming PH-1 with pKP10, a plasmid containing the *tdimerRed* (Campbell, Tour et al. 2002; Freitag and Selker 2005) and *hph* genes followed by selfing and selecting hygR progeny. FMF60 (*hph*^{2x}) was obtained by transforming PH-1 with *hph* and screening for hygR strains with multiple integrations of *hph*. Strains KRP#6 (*hph*$^{+}$), KRP#7 (*hph*$^{+}$, *hph*$^{+}$), KRP#8 (Δ*rid1*::*hph*$^{+}$) and KRP#9 (Δ*rid1*::*hph*$^{+}$, *hph*$^{+}$) were obtained as independent transformants after transforming PH-1 with an *hph* construct targeted to replace the *rid1* homolog (FGSG_08648.3).

Sequencing of RIPped *hph* alleles

FMF60 (*hph*^{2x}) was selfed and presumably RIPped, hygS progeny selected (here called XJW7-). A pool of the duplicated *hph* alleles was amplified from the strains by PCR using primers OMF83 5'-GGCGGAGGCGGCGGAGGCGGAGGCGGAGG and OMF84 5'-CGAGCTCGGATCCATAACTTCGTATAGCA and purified by Qiagen PCR purification (Valencia, CA). These mixed alleles were sequenced by Sanger sequencing using primer OMF84 at the CGRB core labs. The traces were scanned visually for mixed base calls indicative of a mutation present in one allele but not the other.

DNA sequencing

Genomic DNA was extracted from *F. graminearum* tissue as previously described (Pomraning, Smith et al. 2009). Genomic DNA from 00-676-2, was sheared with a microtip-equipped Branson S450A Sonicator (Danbury, CT) by seven successive 10 sec pulses (output: 1.2, duty cycle: 80), followed by 30 sec rest on ice after each cycle. Illumina sequencing libraries were constructed by ligating paired-end adaptors to the sheared DNA and amplifying the library essentially according to our previously published protocol (Pomraning,

Smith et al. 2009). Single-end 40-nt reads were generated on an in-house Illumina GAII machine at the Center for Genome Research and Biocomputing at OSU.

Methylation analysis

Genomic DNA was digested with 5-methylcytosine sensitive (*Sau3AI*) and insensitive (*DpnII*) isoschizomers and analyzed by Southern blot with radiolabelled *hph* probes prepared as described previously (Miao, Freitag et al. 2000). A difference in fragment size obtained by digestion with the two isoschizomers is indicative of DNA methylation.

RIP frequency analysis

RIP frequency is easiest to measure with a convenient phenotypic assay. We make use of the markers *hph* and *tdimerRed* for measuring RIP since mutation of *hph* duplications by RIP leads to loss of hygR while mutation of *tdimerRed* (a direct tandem duplication of the red fluorescent protein gene) by RIP leads to loss of red fluorescence. To test for RIP in strains harboring *hph* duplications the ascospores from a selfing were germinated and tested for growth on YEPD with and without hyg. To test for RIP in strains harboring *tdimerRed* the ascospores are germinated on YEPD and scored for red fluorescence using a fluorescent microscope.

UV mutagenesis

All mutagenesis procedures were carried out in the dark to avoid repair by photoylase. Briefly, 400,000 spores from FMF204 were UV mutagenized in 20 mL of water in a Petri dish in a UVP CL-1000 crosslinker (Upland, CA) at a level to obtain 5% viability (generally 25,000-50,000 $\mu J/cm^2$ depending on the batch of spores). Approximately 4,000 of the mutagenized spores were plated on YEPD in order to obtain ~200 viable spores per plate. These were allowed to recover and grow overnight and then picked into CMC liquid medium to make conidial stocks.

SNP analysis

The 40-nt *F. graminearum* 00-676-2 genomic sequence reads were converted from Solexa FASTQ to Sanger FASTQ quality score format using Mapping and Assembly with Quality (MAQ) v0.7.1 (Li, Ruan et al. 2008). The processed reads were aligned to Assembly 3 of the *F. graminearum* PH-1 reference genome (Cuomo, Guldener et al. 2007) with the MAQ easyrun command (Li, Ruan et al. 2008). Called SNPs were quality-filtered with conservative settings to reduce the likelihood of calling false-positive SNPs. For this, only SNPs called from reads mapped to a single site in the genome with a minimum phred-like quality score of 42, which requires at least 5x coverage, and variant frequency of >= 0.8 were retained.

Results

Characterization of RIP in *F. graminearum*

FMF60, a strain harboring two copies of *hph* was selfed to obtain mutated alleles of *hph*. Hyg^S, presumably RIP mutated progeny were collected and their *hph* genes analyzed for mutations by Sanger sequencing. Between zero and 41 C:G to T:A mutations were identified in each of seven Hyg^S progeny analyzed (Fig. 2.2A) verifying that a RIP-like process is active in *F. graminearum* as found earlier using different strains with multiple copies of *hph* (Cuomo, Guldener et al. 2007). We next examined the dinucleotide context of the C:G to T:A mutations and found that 48% occurred at CpAs and 47% at CpGs (Fig. 2.2B), suggesting less stringent CpA bias than in the previous study (78%; (Cuomo, Guldener et al. 2007)) or in *N. crassa* (74%; (Selker 1990), and M. Freitag, unpublished data). We found one *hph* allele (XJW7-25) with 26 mutations that were all C to T changes on one strand, while the other alleles had mixtures of C:G to T:A transitions.

We next examined the RIP mutated XJW7- progeny for RIP-induced cytosine methylation of the *hph* alleles by Southern blot. In *N. crassa*, DNA methylation is often – but not always – a consequence of RIP (Selker, Cambareri et al. 1987; Singer, Marcotte et al. 1995). There is an almost perfect correlation between AT-rich regions of the genome that bear the signature of RIP (high TpA/ApT content and low CpA content) and the distribution of DNA methylation (Selker, Tountas et al. 2003; Lewis, Honda et al. 2009; Smith, Phatale et al. 2011). We digested genomic DNA with 5-methylcytosine sensitive (*Sau3AI*) and insensitive (*DpnII*) isoschizomers but found no evidence for DNA methylation in any of the *hph* alleles

that had undergone RIP (Fig. 2.3). We also assayed three wildtype *F. graminearum* strains for DNA methylation in the intergenic rDNA spacer regions (Fig. 2.4A) and compared the overall low level of DNA methylation observed to that in *N. crassa* (Fig. 2.4B). Genome-wide methylated-DNA immunoprecipitation suggested an almost complete lack of cytosine methylation in the PH-1 wildtype strain, at least when grown on rich YPD medium (data not shown).

RIP is dependent on *rid1* in *F. graminearum*.

In *N. crassa*, the only gene known to be essential for RIP is *rid* (Freitag, Williams et al. 2002). We replaced the *rid* homolog (FGSG_08648.3; *rid1*) with *hph* in *F. graminearum* (Fig. 2.5A) to determine its importance for RIP and whether the mutant will display a sexual defect. In a parallel experiment, we also replaced *rid1* with the *neo* marker gene (Fig. 2.5B). PCR and Southern analyses of single spore isolates obtained from primary transformants showed that in some *rid1* was either replaced with a single copy of *hph*, KRP#8 (Δ*rid1*::*hph*$^{+}$), while in others, a second copy of *hph* was inserted ectopically in addition to replacing *rid1* with *hph*, KRP#9 (Δ*rid1*::*hph*$^{+}$, *hph*$^{+}$). Additional transformants showed integration of one, KRP#6 (*hph*$^{+}$), or multiple ectopic copies of *hph*, KRP#7 (*hph*$^{+}$, *hph*$^{+}$), without disrupting *rid1* (Fig. 2.5A). The parallel *rid1* deletions that were resistant to G418 also carried *hph* duplications.

All four KRP strains and strains 96, 97 and 98 were selfed and random progeny collected as ascospores. The progeny were tested for RIP phenotypically by measuring their resistance to hygromycin, which is decreased or abolished when *hph* is mutated. Only the strain with multiple copies of *hph* (KRP#7) underwent RIP while KRP#9 which bears multiple copies of *hph* as well as deletion of *rid1* did not RIP (Table 2.1). The same result was obtained with the G418-resistant *hph* duplication strains (Table 2.1). None of the *rid1* mutants (KRP#8 and KRP#9) exhibited sexual defects.

RIP mutant hunt in *F. graminearum*

Deletion of the *rid1* gene showed that RIP mutants can be constructed in *F. graminearum.* Mutations in *N. crassa rid* were recessive, but selfing of *rid1* deletions in *F.*

graminearum resulted in loss of RIP in absence of a mating partner. Thus we were hopeful to be able to isolate recessive RIP mutants in *F. graminearum* after UV mutagenesis.

PH-1 was transformed with a plasmid containing *hph* and *tdimerRed* (a marker containing a tandem duplication of part of *rfp* that undergoes RIP at high frequency in *N. crassa*). This led to the recovery of two useful strains FMF204 (*tdimerRed*$^{+}$; *hph*$^{+}$)$^{\text{multi-copy}}$ and FMF205 (*tdimerRed*$^{+}$; *hph*$^{+}$)$^{\text{single-copy}}$ for a RIP mutant screen. Selfing of FMF205 with a single copy of the plasmid led to rare inactivation of the *tdimerRed* gene while selfing of FMF204 led to frequent inactivation of *tdimerRed* within a week after the first ascospores began to shoot. Conidia from FMF204 were UV mutagenized and isolated. Each mutant was then tested for RIP by allowing it to self, followed by counting the number of ascospores that were RFP^{-} by microscopy. In total, 230 mutants were analyzed, none of which were RIP mutants. Six mutants had defects in sexual development and one had a pigmentation defect. However, frequent loss of RFP signal was later seen in strains derived from FMF204 that also had *rid1* replaced with *neo* (*rid1*::*neo;* (*tdimerRed*$^{+}$; *hph*$^{+}$)$^{\text{multi-copy}}$) suggesting that loss of RFP by recombination between tandem copies of the plasmid may be the main pathway for loss of RFP signal in FMF204.

These findings, combined with the realization that even in homothallic fungi fusion of different segments of one thallus are required for fertility (based on work by Joseph Whalen on two genes required for hyphal fusion, *ham2* and *soft*), led us to abandon the RIP mutant hunt in *F. graminearum*, at least for now. To make a better strain for scoring of RIP mutants it needs to be ensured that PH-1 has multiple unlinked copies of a marker which RIP at high frequency but are not susceptible to loss by recombination. More importantly, the requirement for hyphal fusion for fertility (absent in both *ham2* and *soft* mutants) negates the benefit of working with a homothallic species, as mutated and unmutated nuclei may interact in bulk-irradiated spores and thus result in functionally diploid tissues, as would be the case if a mating partner was required (as in *N. crassa*).

The genome sequence of *F. graminearum* 00-676-2: SNP analysis

The *F. graminearum* 00-676-2 genome was sequenced as single-end 40-nt reads on an Illumina GAII machine. In total, 18,842,228 reads were obtained, 17,005,192 of which

were mapped by MAQ, resulting in a median coverage of ~17-fold across the ~37 Mb *F. graminearum* PH-1 reference genome. We were interested in identifying SNPs between PH-1 and 00-676-2 as a tool for mapping new mutants by bulk segregant analysis and high-throughput sequencing. The *Fusarium* genome project analyzed another *F. graminearum* strain GZ3639 but only to 0.4x coverage, which is not high enough to call many SNPs with confidence (Cuomo, Guldener et al. 2007).

We identified 62,310 high quality SNPs between the PH-1 reference genome and the 00-676-2 genomic reads after quality filtering SNP calls from MAQ alignments, resulting in an average SNP density of 1 SNP per 589 nucleotides. However, the SNPs are heterogeneously spaced with enrichment occurring around ancient telomere fusions (Fig. 2.6) as indicated by chromosome synteny (Cuomo, Guldener et al. 2007) and histone H3 lysine 27 trimethylation (Connolly LR, Smith KM and Freitag M, in preparation) which decreased the number of chromosomes from at least nine in ancient Fusaria to four in *F. graminearum* and closely related species.

Expression analysis of *rid1*

Activity of the only gene known to be essential for RIP, *rid1*, may pinpoint when exactly RIP occurs during the sexual cycle. The coding sequence of *rid1* was replaced with the gene for *gfp* (Fig. 2.7) to analyze its expression and tissue specificity visually, as separation of the sexual tissue from surrounding hyphae for extraction of RNA for northern or qPCR analysis is not really feasible. Green fluorescence was not observable in either vegetative or sexual tissue in FMF200 (Δ*rid1*::*GFP*), suggesting that expression of *rid1* is very low in *F. graminearum* , similar to the case in *N. crassa* (Freitag, Williams et al. 2002).

Tagging of *rid1* for biochemical analysis and to find interacting proteins

Association of RID1 with other proteins is of interest for defining the mechanism of RIP. To that end *rid1* was tagged with a 3xFLAG tag (3xDYKDDDDK) for immunoprecipitation experiments (Fig. 2.7). At present FMF201 (*rid1-flag*) has not been used for flag pull down experiments for two reasons. First, expression from the endogenous *rid* locus is very low which will make it hard to pull down large amounts of RID1-FLAG. Second, at the time the

strain was constructed we could not get *F. graminearum* to go through the sexual cycle in a reproducible manner making it difficult to obtain the tissue necessary for experiments. At present selfings are becoming more standardized as *F. graminearum* is slowly domesticated while the amount of protein necessary for mass-spectroscopy is decreasing making RID1-FLAG immunoprecipitation experiments more feasible.

Discussion

Characterization of RIP in *F. graminearum*

We repeated the experiments of Kistler and colleagues (Cuomo, Guldener et al. 2007) using independent transformants and verified that RIP is active in *F. graminearum*. We obtained slightly lower enrichment of mutations in the CpA dinucleotide context (48% versus 78%), however, we sequenced fewer RIP mutated alleles of *hph* in our study. Interestingly, one allele of *hph* sequenced had 26 C to T changes on a single strand, while the other alleles had mixtures of C:G to T:A transitions. In *N. crassa* the frequency of C to T versus G to A mutations is not random, suggesting RIP acts processively on a single strand of each repeat copy at one time (Selker 1990; Watters, Randall et al. 1999). This hypothesis is also supported by our findings in *F. graminearum*.

Since genomic cytosine methylation is associated with RIP mutated DNA in *N. crassa* (Selker, Tountas et al. 2003; Lewis, Honda et al. 2009) we examined the RIP mutated alleles of *hph* for *de novo* cytosine methylation. We found none. In *Neurospora* heavy RIP induces DNA methylation, while less severely mutated alleles remain unmethylated (Singer, Marcotte et al. 1995). The degree of RIP in the alleles we tested here was relatively light, even in the most heavily mutated allele investigated (33 mutations/kb). This level of RIP would typically be associated with *de novo* DNA methylation in *N. crassa* (Singer, Marcotte et al. 1995; Selker, Tountas et al. 2003). However, frequency of mutation is not the only determinant for *de novo* methylation in *Neurospora*. Studies on *Neurospora* DNA methylation signals suggest that AT content and enrichment for TpA dinucleotides, even in short, ~230 bp fragments increases the likelihood of *de novo* methylation (Miao, Freitag et al. 2000). The best signals for DNA methylation are regions enriched for TAAA or TTAA repeat (Tamaru and Selker 2003). In our study, the wild type *hph* gene was 60% GC, while the wild type *Neurospora am*

gene was 54% GC (Singer, Marcotte et al. 1995). Even in our most heavily mutated *hph* allele, the GC content was only slightly reduced, from 60% to 56.5%. Based solely on their AT content and TpA enrichment, we would expect that our mutated *hph* alleles would also not be methylated in *N. crassa*. Further studies on the fate of DNA subjected to RIP in *F. graminearum* are warranted.

In *N. crassa*, the only gene known to be essential for RIP is *rid* (Freitag, Williams et al. 2002). However, homologs of RID in other fungi display enigmatic phenotypes when inactivated. Strains in which the *rid* homolog of *Ascobolus immersus*, *masc1*, has been mutated showed that Masc1 is important for a similar process, MIP (methylation induced premeiotically), but in addition has a recessive sexual development defect. DmtA, the RID homolog in *Aspergillus nidulans*, a fungus in which neither RIP nor MIP has been observed, also displays a sexual defect (Lee, Freitag et al. 2008). We replaced the *F. graminearum rid* homolog (FGSG_08648.3; *rid1*) with *hph* in to determine its importance for RIP and whether the mutant would display a sexual defect. We obtained strains with single or multiple integrations of *hph* in which some had replaced *rid1*, inactivating it. In strains where *rid1* had been replaced by *hph*, RIP did not inactivate multicopy *hph* alleles suggesting that, like in *N. crassa*, RIP is dependent on *rid1* and that homozygous *rid1* selfings are fertile. This suggests that the function of the *rid* homolog of *F. graminearum* is akin to that of *N. crassa* rather than to that of *A. immersus* or *A. nidulans*. Importantly, deletion of a gene that would generate a recessive mutation in *Neurospora* (Freitag, Williams et al. 2002) resulted in easily observable phenotypes in a selfing of mutant *F. graminearum* strains. This experiment thus constitutes a proof-of-principle study for the likely success of future screens or selections for recessive RIP mutants in *F. graminearum*.

RIP mutant hunt in *F. graminearum*

We attempted to find novel RIP mutants in *F. graminearum* since RIP seems to be dependent on *rid1* (a control for the mutant hunt) and occurs at high frequency like in *N. crassa*. Importantly, *F. graminearum* has the advantage of being homothallic, allowing the recovery of recessive mutants, as no mating partner is required. We decided to use *tdimerRed* as a marker for screening RIP frequency in mutants. Mutation of the *tdimerRed*

marker by RIP leads to loss of red fluorescence in germinated ascospores which can be counted a few hundred at a time on plates. We decided to irradiate to a high death rate and to test survivors one by one for their RIP frequency, a highly laborious process.

Ultimately mutagenesis of the strains harboring the *tdimerRed* allele identified no RIP mutants. The first and foremost reason for this is that *tdimerRed* did not act as a good indicator of RIP in *F. graminearum*. The marker consists of an ~0.7 kb tandem duplication of two monomeric *rfp* genes that undergoes RIP efficiently in *N. crassa* (see Chapter 4). However, it was found to only infrequently undergo RIP when it was integrated in single copy (strain FMF205). Therefore we decided to use FMF204, which had multiple integrations of *tdimerRed*, as the parent for mutagenesis since the RFP signal is lost frequently in progeny from selfings of FMF204. It is still unknown whether the integrations are linked or not and the mutated alleles of *tdimerRed* from selfings of FMF204 were never well characterized. It was later found that inactivation of *tdimerRed* occurred with high frequency even when *rid-1* was replaced by the *neo* marker, suggesting that inactivation of *tdimerRed* was often not due to RIP. If the multiple integrations of *tdimerRed* occurred as linked tandem copies it is likely that inactivation is occurring by the single-strand annealing recombination pathway which affects tandem repeats with high frequency in yeast (Sugawara and Haber 1992) and *N. crassa* (Irelan, Hagemann et al. 1994). In order to avoid this problem in the future strains should be constructed with unlinked duplications of *rfp* which will not be susceptible to single-strand annealing (Irelan, Hagemann et al. 1994).

At this stage the RIP mutant hunt has been unsuccessful but we have (1) learned how to perform efficient UV mutagenesis in *F. graminearum* as many mutants with growth phenotypes were observed (data not shown), (2) improved the reproducibility of selfings (data not shown) by changing experimental methods (e.g. shortening the autoclave time of carrot agar, (3) learned that *F. graminearum* undergoes homologous recombination between direct tandem repeats efficiently which must be avoided to not confuse it with inactivation by RIP and (4) sequenced another wild strain, 00-676-2 which can be used for mapping by high-throughput sequencing. This strain carries a *nit* mutation (Gale, Bryant et al. 2005), which makes it relatively easy to select for recombinant progeny when doing crosses with a homothallic fungus. The progeny can be tested for their phenotype and then pooled for bulk

segregant analysis using the 62,310 high quality SNPs between 00-676-2 and PH-1 identified, similar to a method we developed for *N. crassa* (Pomraning, Smith et al. 2011). This progress sets the stage for successful mutant hunts of any kind in *F. graminearum*.

Acknowledgements

We thank Mark Dasenko, Matthew Peterson, Steve Drake, Chris Sullivan and Scott Givan at the OSU CGRB core facility for assistance with Illumina sequencing and data analysis and Frances Trail (Michigan State University) for *Fusarium* strains. This work was made possible by start-up funds from the OSU Computational and Genome Biology Initiative and grants from the American Cancer Society (RSG-08-030-01-CCG), DOE (DE-FG02-08ER64665) and NIH (P01 GM068087 and R01GM097637) to MF.

Table 2.1. Strains used in this study to assay RIP frequencies. In all cases the RIP frequency (RIP%) is based on tests for Hyg^R on YEPD medium with 100 μg/ml Hyg, typically in 48-well plates. The RIP% is calculated as the number of Hyg^S colonies divided by the total number of strains examined. Failure to isolate Hyg^S strains is denoted as "not detected".

Strain	Genotype	Assay	% RIP (hygS/total)	Notes
IN10	*hph*$^+$	Hyg^R	not detected (0/256)	Cuomo et al. 2007
IS1	*hph*$^+$, *hph*$^+$	Hyg^R	42% (108/256)	Cuomo et al. 2007
KRP#6	*hph*$^+$	Hyg^R	not detected (0/25)	This study
KRP#7	*hph*$^+$, *hph*$^+$	Hyg^R	32% (7/22)	This study
KRP#8	Δ*rid1*::*hph*$^+$	Hyg^R	not detected (0/81)	This study
KRP#9	Δ*rid1*::*hph*$^+$, *hph*$^+$	Hyg^R	not detected (0/47)	This study
LRCX9-1	*hph*^{2ec}	Hyg^R	11% (8/72)	This study
FMF96	Δ*rid1*::*neo*$^+$, *hph*^{2ec}	Hyg^R	not detected (0/128)	This study
FMF97	Δ*rid1*::*neo*$^+$, *hph*^{2ec}	Hyg^R	not detected (0/116)	This study
FMF98	Δ*rid1*::*neo*$^+$, *hph*^{2ec}	Hyg^R	not detected (0/173)	This study

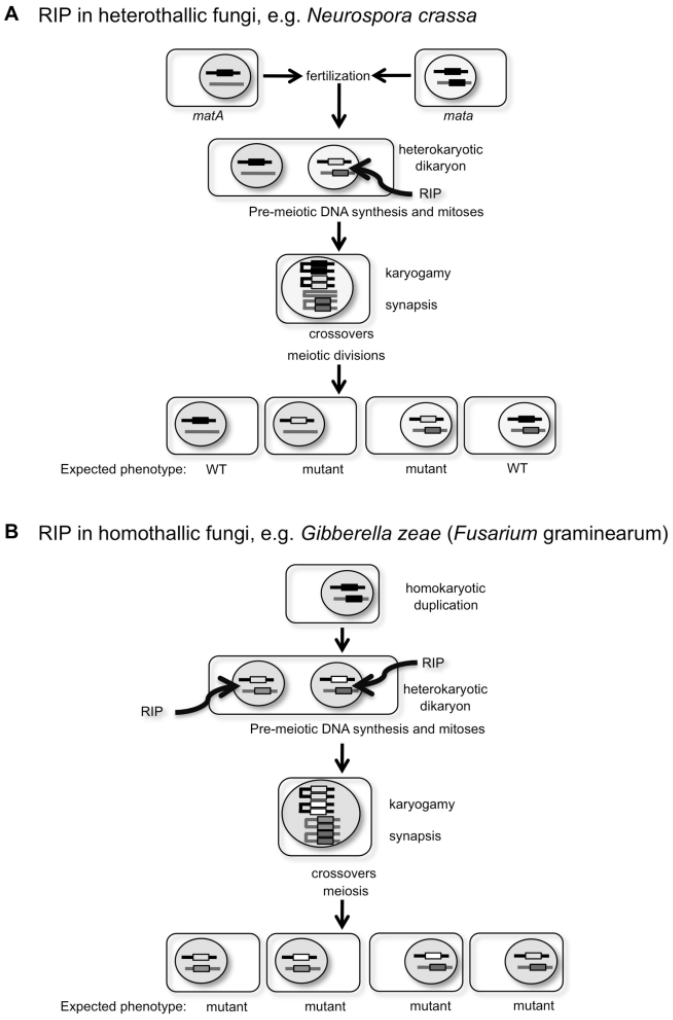
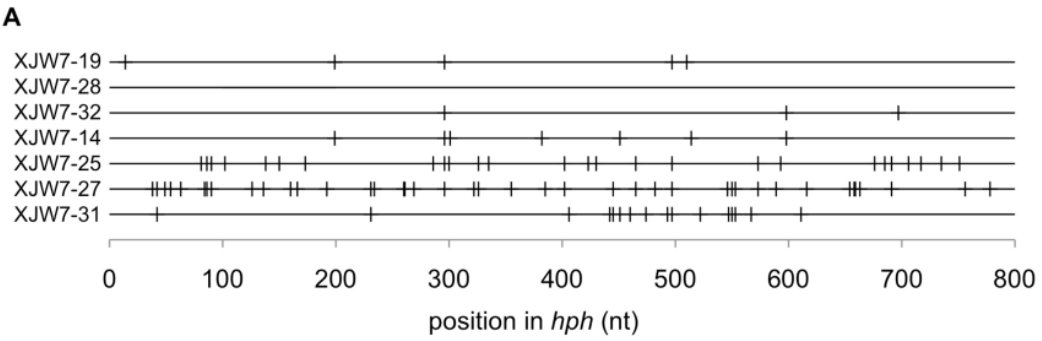


Figure 2.1. Expected differences in RIP in (A) heterothallic and (B) homothallic fungi. In heterothallic fungi (A), mating of two strains is required to generate a heterokaryon. RIP occurs between fertilization and karyogamy and affects only nuclei in which duplications occur. After meiosis, four meiotic products contain segregated copies of the duplications. The black boxes indicate wildtype alleles and shaded grey boxes indicate differentially mutated alleles. If the mutated gene is recessive and if mutation is severe enough to cause phenotypes, two phenotypically wildtype and two phenotypically mutant alleles are recovered after a cross (adapted from (Selker et al. 1987; Cambareri et al. 1989; Selker 1990; Singer et al. 1995)). In homothallic fungi (B), no mating partner is required. After initiation of the sexual phase all nuclei are subject to RIP, but it remains unclear at precisely what stage of development the homology search and mutagenesis occurs. If mutation is severe enough, all four meiotic products are expected to express mutant phenotypes (various shades of gray and white boxes). In both schemes RIP events will escape detection if mutation is light and phenotypes are not easily discernable. DNA sequencing is the only method by which the extent of RIP can be truly quantified.



B Number and context of mutations

Strain	CpA	CpT	CpG	CpC	total
XJW7-19	3	0	2	0	5
XJW7-28	0	0	0	0	0
XJW7-32	1	1	1	0	3
XJW7-14	1	0	6	0	7
XJW7-25	13	1	12	0	26
XJW7-27	24	0	16	1	41
XJW7-31	10	0	6	0	16

Figure 2.2. Sequence analysis of *hph* in progeny from a selfing of an *hph* duplication strain. (A) Position of mutations along the sequenced region of the *hph* duplication. DNA sequencing of seven HygS strains showed variable numbers of mutations. The genomic location of the *hph* duplication was not determined and PCR amplification likely occurred on both alleles. In all but XJW7-25, the results are from mixed alleles, indicated by mixed nucleotide read-outs. In XJW7-25, all mutations were homozygous (only one nucleotide was detected at all positions by sequencing), suggesting that the sequence is derived from a single allele. (B) Tabulation of the number and dinucleotide context of the mutations found in the sequenced portion of *hph* alleles.

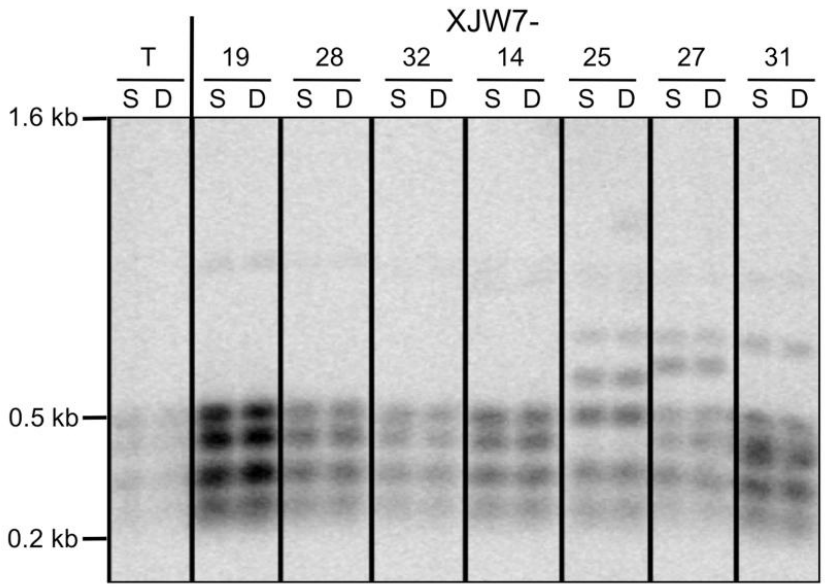


Figure 2.3. Mutations caused by RIP did not result in DNA methylation. Genomic DNA was digested with 5-methylcytosine sensitive *Sau3AI* (S) and insensitive *DpnII* (D) isoschizomers and analyzed by Southern blot with *hph* probe. While mutations are present in all the strains except XJW7-28, in no case did we observe differences in the *Sau3AI*- and *DpnII*-digested DNA, which suggests that RIP mutagenesis was either not severe enough to generate a *de novo* DNA methylation signal, or that DNA methylation is not a consequence of RIP in *F. graminearum*.

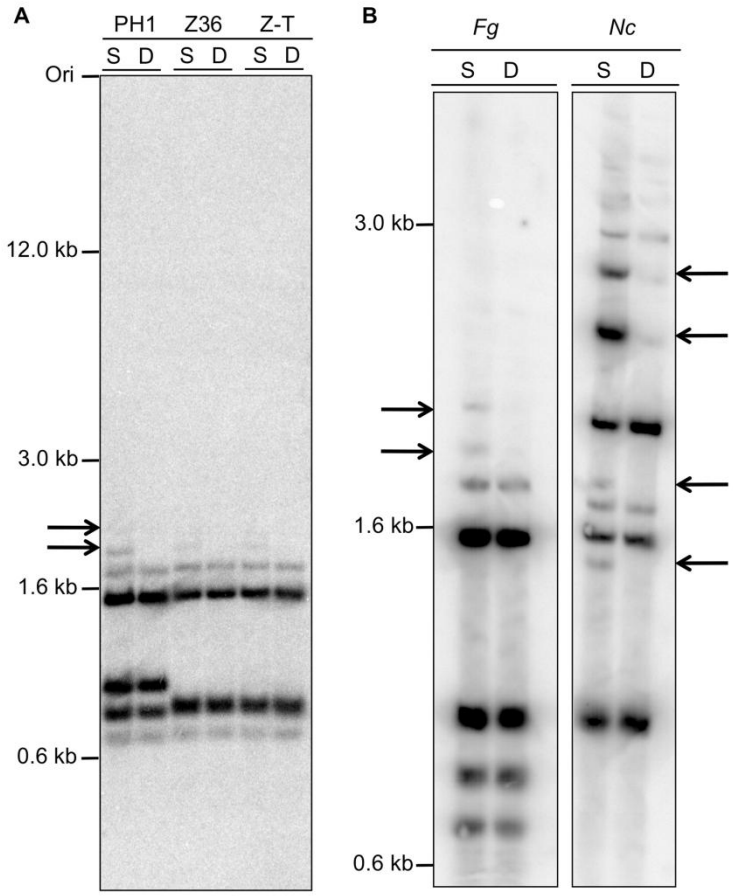


Figure 2.4. Little rDNA methylation in *F. graminearum*. (A) RFLPs in untranscribed rDNA spacer regions of PH1 and Z3639. Genomic DNA from three different *F. graminearum* strains (PH1; Z3639, "Z36"; Z3639-deribed transformant, "Z-T") was digested with the 5-methylcytosine sensitive *Sau3A*I (S) and insensitive *Dpn*II (D) isoschizomers, separated in a 1% agarose gel and blotted to nylon membrane. Blots were probed with a fragment of the rDNA untranscribed spacer derived by PCR of DNA between the large and small ribosomal subunits. Very little DNA methylation was detected (arrows). B. Comparison of rDNA methylation levels in *F. graminearum* and *N. crassa*. Genomic DNA of *F. graminearum* PH1 and *N. crassa* OR74A was treated as in (A). *Fusarium* (*Fg*) and *Neurospora* (*Nc*) samples were run on the same gel and blotted together, membranes were cut after transfer and hybridized with rDNA untranscribed spacer probes derived by PCR from *Fusarium* or *Neurospora* genomic DNA. Blots in (B) were stripped and re-probed with *Fusarium* and *Neurospora hH3* fragments to assess complete digestion (data not shown).

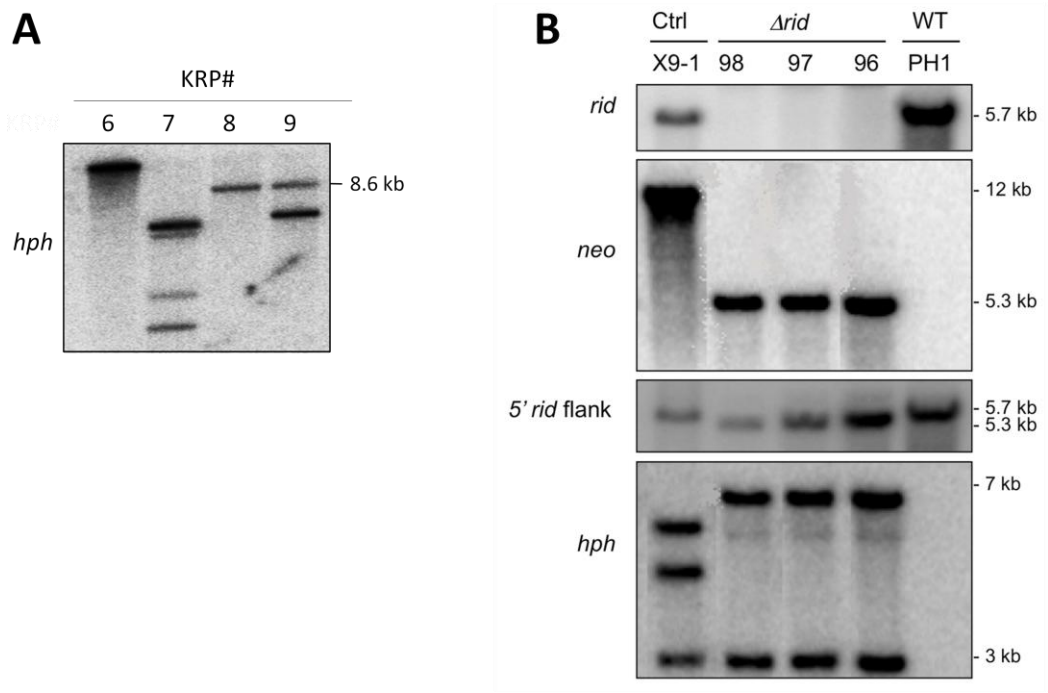


Figure 2.5. Deletion of the *F. graminearum rid* homologue. (A) Genomic DNA from *F. graminearum* strains KRP #6-9 probed with *hph*. A band at 8.6 kb shows *rid1* has been replaced by *hph* and (B) genomic DNA from *F. graminearum* wild type PH1, three independently derived single spore isolates in which *rid1* had been replaced with *neo* (FMF96, 97, 98) and an independent transformant with an ectopic *neo* integration (X9-1) was digested with *Bgl*II, separated in a 1% agarose gel and blotted to nylon membrane. Blots were successively probed with fragments containing the whole coding regions of *rid1*, *neo* or *hph* and the 5' flank of *rid1*. Expected sizes for fragments after correct replacement of *rid1* with *neo* are shown at right. Note the small difference between the fragment size obtained with the 5' *rid1* flank probe (5.7 kb in wild type compared to 5.3 kb in replacement strains). Both the control strain X9-1 and the Δ*rid1* replacement strains have two copies of *hph*.

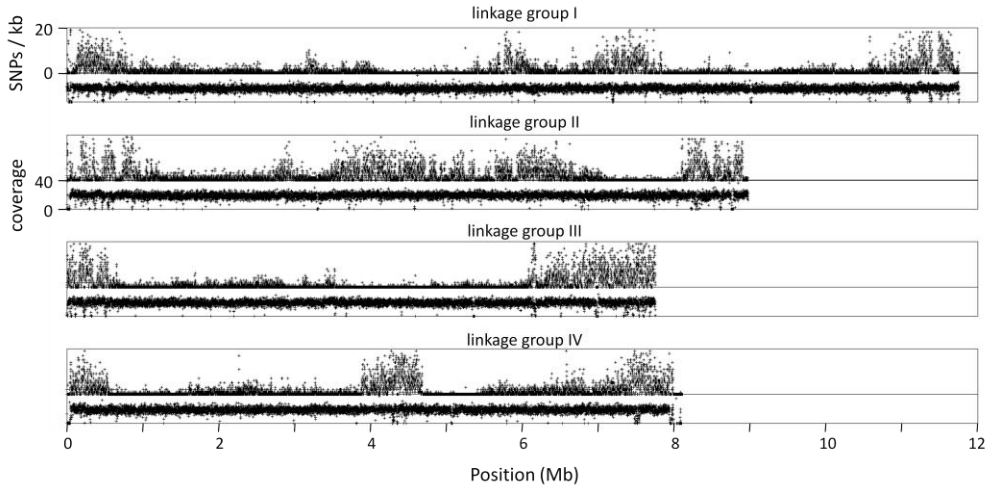


Figure 2.6. Density of 62,310 SNPs between *F. graminearum* PH-1 and 00-676-2. The SNP density (0-20 SNPs/kb) and short read coverage (0-40x) of 00-676-2 Illumina sequencing compared to the PH-1 reference genome. All data points are the mean of 1 kb windows.

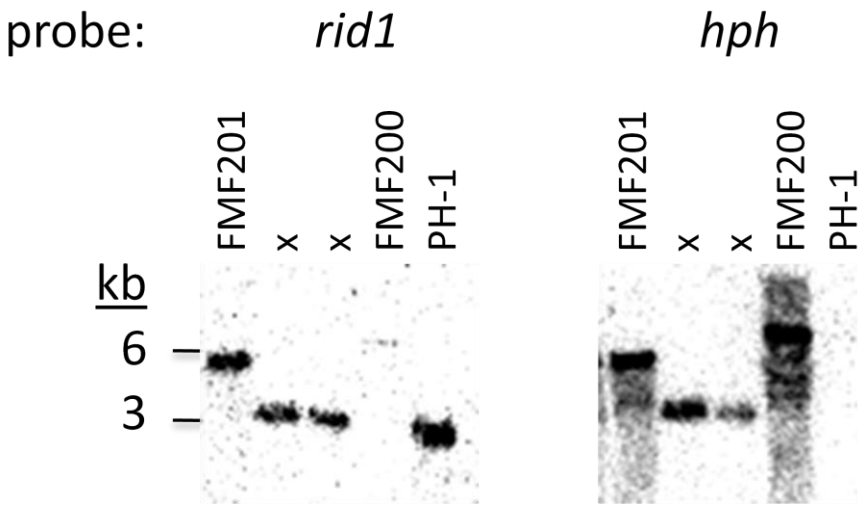


Figure 2.7. Southern analyses of Δ*rid1*::*gfp* and *rid1-flag F. graminearum*. Genomic DNA from homokaryotic strains FMF200 (Δ*rid1*::*GFP*; *hph*$^{+}$), FMF201 (*rid1-flag*; *hph*$^{+}$) and their parent PH-1 was digested with *XhoI*. Southern blots used probes for *hph* and *rid1*. Bands at 5.5 kb and 5.7 kb are expected for integration at the endogenous locus of Δ*rid1*::*GFP*; *hph*$^{+}$ and *rid1-flag*; *hph*$^{+}$ respectively.

Proof of principle of fast mutation mapping by bulk segregant analysis and high-throughput sequencing

Kyle R. Pomraning, Kristina M. Smith Frederick Bowring, David Catcheside and Michael Freitag

Eukaryotic Cell
2011 Jun;10(6):724-733
PMID: 21515825

with additional unpublished work

Abstract

With the advent of high-throughput DNA sequencing it is now straightforward and inexpensive to generate high-density small nucleotide polymorphism (SNP) maps. Here we present 168,579 SNPs between *Neurospora crassa* Mauriceville-1-c (FGSC2225) and OR74A (FGSC987), the strains most typically used by *Neurospora* researchers to carry out mapping crosses. The general map location of a mutation can be identified by a single backcross to a strain enriched in SNPs compared to a standard wild-type strain. Bulk segregant analysis simultaneously increases the likelihood of determining the precise nature of the mutation. We combined high-throughput sequencing with bulk segregant analysis to generate a SNP map between two *N. crassa* strains and expedite mutation mapping. We further demonstrate the utility of the Mauriceville sequence and our approach by mapping the mutation responsible for the only existing temperature-sensitive (ts) cell cycle mutation in *Neurospora*, nuclear division cycle-1 (*ndc-1*). The single T to C point mutation maps to the gene encoding ornithine decarboxylase (ODC), *spe-1* (NCU01271), and changes a Phe to a Ser residue within a highly conserved motif next to the catalytic residue of the enzyme. Further we have used the SNP map to identify crossover points in the four meiotic products of a cross between *N. crassa* Mauriceville-1-c (FGSC2225) and OR74A (FGSC987) and found 1.0±0.7 crossovers per chromosome as well as rare gene conversion events.

Introduction

Single nucleotide polymorphism (SNP) maps between organisms with different genetic backgrounds are useful for identifying point mutations that result in an observable phenotype. By mating a mutant with a strain of different genetic background and then selecting progeny with and without phenotype one is able to map the underlying mutation(s) by bulk segregant analysis (Michelmore, Paran et al. 1991). After undergoing recombination, each progeny exhibiting a specific phenotype will have the same genetic background as the mutant parent in the vicinity of the mutation but a mixture of DNA inherited from either parent in regions of the genome that are unlinked to the mutation. By analyzing DNA from a mixture of phenotypically mutant progeny from the cross it is possible to define the genetic

region containing the mutation, even to identify the mutation itself. This is what is meant by "bulk segregant analysis".

Most recently, molecular mapping strategies based on restriction fragment length polymorphisms (RFLP) (Metzenberg 1984), cleaved amplified polymorphic sequences (CAPS) (Jin, Allan et al. 2007; Lambreghts, Shi et al. 2009) or microarray-based restriction-site-associated DNA mapping (RAD-mapping) (Lewis, Shiver et al. 2007; Baird, Etter et al. 2008) have been used to map mutations in *Neurospora crassa*. These methods can be time-consuming, laborious, and therefore expensive. In addition, once the mutation is mapped to a specific region of the genome additional primer pairs for CAPS mapping need to be ordered and additional SNPs identified to more precisely map the mutation. Typically, researchers will then attempt complementation of phenotypes by candidate genes. Once (or if) complementation was successful, the mutated allele of the candidate gene still needs to be sequenced by traditional Sanger sequencing. In contrast, methods based on microarrays (Brauer, Christianson et al. 2006) or high-throughput DNA sequencing (Blumenstiel, Noll et al. 2009; Birkeland, Jin et al. 2010; Ehrenreich, Torabi et al. 2010; Wenger, Schwartz et al. 2010) make it possible to quickly generate high-density SNP maps that allow direct identification of specific point mutations or indels by sequence analysis. Moreover, this approach does not rely on a specific set of reference strains.

We tested mapping by high-throughput Illumina sequencing in studies with the filamentous fungus *N. crassa*, one of the most widely used genetic model organisms. The two most commonly used genetic backgrounds of *Neurospora* for molecular mapping purposes are OR74A (FGSC987) (Case, Brockman et al. 1965) and Mauriceville-1-c (FGSC2225) (Beauchamp, Horn et al. 1977), hereafter called OR and MV. The OR strain was used for the *N. crassa* genome sequencing project (Galagan, Calvo et al. 2003). The genome has been assembled into seven chromosomes, represented by supercontigs one to seven. There are 13 supercontigs that mostly contain rDNA and AT-rich regions mutated by RIP (K.M. Smith and M. Freitag, unpublished data). We have now sequenced and partially assembled the *N. crassa* MV genome to improve upon the available high-density SNP maps between the OR and MV strains (Lewis, Shiver et al. 2007; Lambreghts, Shi et al. 2009). As proof of principle for mutation mapping by high-throughput sequencing, we sequenced DNA in bulk from

mutant progeny resulting from a cross between MV and the classic *N. crassa* mutant *ndc-1* (*nuclear division cycle-1*; FGSC3441). We identified one specific candidate SNP that likely caused the temperature-sensitive (ts) *ndc-1* defect, which presents as arrest just prior to DNA synthesis during the nuclear division cycle (Serna and Stadler 1978). This single T to C point mutation maps to *spe-1* (NCU01271), the gene encoding ornithine decarboxylase (ODC) (McDougall, Deters et al. 1977; Paulus, Kiyono et al. 1982), and changes a Phe to a Ser residue within a highly conserved motif near the ODC catalytic site. We were able to complement the *ndc-1* phenotypes by growth on spermidine and transformation of the mutant with a wild-type *spe-1* gene. We have also sequenced recombinant progeny in isolation after crossing OR to MV and demonstrate that the SNP map used is highly accurate for analyzing recombination in pure strains.

Materials and methods

Strains, crosses and growth conditions

The MV genome sequence was produced from *N. crassa* Mauriceville-1-c (FGSC2225; NMF37) (Beauchamp, Horn et al. 1977). The octad genome sequences were produced from ordered ascospores collected after crossing Mauriceville-1-c (FGSC2225; NMF37, *matA*) (Beauchamp, Horn et al. 1977) and 74-Ors-6a (FGSC4200, NMF61, *mata*) (Kafer and Fraser 1979). Strains were grown according to standard conditions (Davis 2000) in Vogel's minimal medium (VMM) with 1.5% sucrose at 32°C, except for ts *ndc-1* strains that were grown at 22°C, and *inl* strains that were supplemented with 50 mg/L of myo-inositol. The mapping cross was inoculated with *N. crassa* MV as the recipient strain on synthetic crossing medium supplemented with 0.5% sucrose and 50 mg/L myo-inositol. Conidia from the *ndc-1* mutant (FGSC3441; NMF164) (Serna and Stadler 1978) were added three days later and spread across the recipient culture. After two weeks, random ascospores were recovered from the cross and heat shocked at 65°C for 60 min (standard heat shock) or 60°C for 30 min (gentle heat shock) on VMM with 50 mg/L myo-inositol (VMMI) and FGS (0.5 g/L fructose, 0.5 g/L glucose, 20 g/L sorbose) as the carbon source. After overnight growth at room temperature, two hundred viable progeny were collected that were treated according to each of the heat shock protocols, and incubated in slants with VMMI with 1.5% sucrose. The progeny were

tested for *inl* by spotting conidia on VMM or VMMI with FGS, and for the ts *ndc-1* allele by spotting conidia on VMMI with FGS followed by growth at 22°C or 37°C. Complementation of the *ndc-1* mutant was done on VMMI with 1.5% sucrose and strains were grown at 22°C or 37°C with or without 1 mM spermidine trihydrochloride.

DNA sequencing

Genomic DNA was extracted from *N. crassa* tissue as previously described (Pomraning, Smith et al. 2009). Genomic DNA from MV, the octad strains and different progeny from the MV x *ndc-1 inl* cross was quantified using a NanoDrop2000 spectrophotometer and diluted to 100 ng/ul with TE buffer. DNA from progeny was mixed in two pools, the first containing equal amounts of DNA from 54 progeny (NMF343 to NMF396) with both the *inl* and *ndc-1* mutations, the second containing equal amounts of DNA from nine progeny (NMF334 to NMF342) with only the *inl* mutation. The pooled DNA was placed in a 4°C circulating water bath and sheared into 200-1000 bp fragments in a Bioruptor (Diagenode, Denville, NJ) set on “high” for 18 min (30 sec on/off cycles). MV and octad DNA was sheared with a microtip-equipped Branson S450A Sonicator (Danbury, CT) by seven successive 10 sec pulses (output: 1.2, duty cycle: 80), followed by 30 sec rest on ice after each cycle. Illumina sequencing libraries were constructed by ligating paired-end adaptors, some with barcodes (Table 3.1), to the sheared DNA and amplifying the library essentially according to our previously published protocol (Pomraning, Smith et al. 2009). Paired-end 80-nucleotide (nt) or single-end 40-nt reads were generated on an in-house Illumina GAII at the Center for Genome Research and Biocomputing at OSU for the MV and pooled strains, respectively. Single-end reads of various lengths were generated on an in-house Illumina GAII or HiSeq at the Center for Genome Research and Biocomputing at OSU for the octad strains.

SNP analysis and genome assembly

The 80-nt *N. crassa* MV genomic sequence reads were trimmed to 63-nt and converted from Solexa FASTQ to Sanger FASTQ quality score format using Mapping and Assembly with Quality (MAQ) v0.7.1 (Li, Ruan et al. 2008). The processed reads were aligned to Assembly 10 of the *N. crassa* OR reference genome (Galagan, Calvo et al. 2003) with the

MAQ easyrun command (Li, Ruan et al. 2008). Called SNPs were quality-filtered with conservative settings to reduce the likelihood of calling false-positive SNPs. For this, only SNPs called from reads mapped to a single site in the genome with a minimum phred-like quality score of 42, which requires at least 5x coverage, and N as the second best call were retained. For genome assembly the 80-nt MV reads were trimmed to 76-nt to remove the 4-nt GGGT barcode that was used for Illumina sequencing. Reads were then assembled *de novo* into contigs with Velvet v1.0.12 (Zerbino and Birney 2008). Contig alignments were performed using NUCmer, a part of MUMmer v3.0 (Kurtz, Phillippy et al. 2004).

SNP validation

SNPs were mapped as perfect matches to a dataset of 248,997 *N. crassa* MV expressed sequence tags (ESTs; M. Basturkmen, J. Xu, M. Shi, J. Loros, M. Nelson, M. Henn, C. Kodira, N. Lennon, L. Green, J. Galagan, B. Birren, J. Dunlap, M.S. Sachs, in preparation) of 71-nt reads (35 nt flanking both versions of each SNP that was farther than 35 nt from the nearest called SNP) using Hashmatch (Filichkin, Priest et al. 2009). For validation, 47 SNPs that lie in a *TaqI* restriction endonuclease cleavage site in the OR background but not in the MV background were selected and amplified by PCR from the OR and MV strains with primer pairs obtained from the *Neurospora* Functional Genomics Project (Table 3.1). PCR products were digested with *TaqI* and RFLP patterns analyzed by agarose gel electrophoresis.

Mutant mapping by high-throughput sequencing

The “SNPome”, i.e. all the SNPs between two *N. crassa* strains was generated by running the Perl script “SNPome_builder”. The output SNPome consists of a FASTA-formatted file that contains both versions of each SNP and the reference genome coordinates. SNPome_builder extracts the genomic sequence around each SNP site from the reference genome and makes a FASTA-formatted sequence for each version of the SNPs. When called SNPs are closer than the read length, sequences are concatenated (Fig. 3.1). Processed Illumina reads were mapped to the SNPome as perfect matches with Hashmatch (Filichkin, Priest et al. 2009). A second Perl script, “SNP_profiler” was run on the Hashmatch output to count the number of reads hitting each version of each SNP. This script then

calculates a ratio of reads matching each version of a SNP as a fraction between 0 and 1 to indicate which parental genome was the likely origin of the underlying DNA sequence. SNP profiles were viewed in either the Argo genome browser (R. Engels; www.broadinstitute.org/annotation/argo2/) to identify contiguous genomic regions in which SNPs originated from a specific parental strain or using Gnuplot (http://www.gnuplot.info/).

MAQ (Li, Ruan et al. 2008) was used to identify SNPs between mutant and non-mutant strains in the regions identified by bulk segregant analysis. Exclusion of spurious SNPs as candidates for point mutations was performed by aligning reads from genomic sequences from FGSC322 (Horowitz, Fling et al. 1960) and FGSC3562 (Vigfusson and Weijer 1972) to the OR assembly with MAQ (Li, Ruan et al. 2008) and subtracting SNPs found in these *ndc-1*$^{+}$ *inl*$^{+}$ strains from SNPs found in mutant progeny. The SNPs remaining that had phred-like quality scores of ≥42 were aligned to the OR genome in GFF3 format and visually scanned in Argo for specific regions that generated non-synonymous mutations in predicted or known coding regions.

Sequencing of the *ndc-1* allele

To verify that *ndc-1* is an allele of *spe-1*, partial DNA sequences of the *spe-1*^{ndc} allele were obtained by traditional Sanger sequencing from three *ndc-1* strains (NMF344, NMF345 and NMF346) and a MV sibling (NMF398). We designed primers, within the *spe-1* coding region (OMF2243, OMF2244; Table 3.1), amplified this region from the strains in three independent reactions and carried out DNA sequencing at the CGRB core facility using primer OMF2243.

Complementation of *ndc-1*

The *ndc-1* mutant (FGSC3441; NMF164) was transformed by electroporation with pKP19, a derivative of pGS1, which contains most of the *spe-1* region (Williams, Barnett et al. 1992). pKP19 carries the *hph* gene that can confer resistance to hygromycin integrated between the *BamHI* and *EcoRV* sites of pGS1. Transformants were obtained on VMMI + 100 □g/ml hygromycin B at 22°C. Heterokaryotic transformants were selected and grown at 37°C to score for complementation of the *ndc-1* phenotype. To verify transformation and

complementation, DNA from around the mutated site of *spe-1* was amplified by PCR in three independent reactions using a mixture of three primers (OMF2244, NCU1271.5F and NCU1271.3R; Table 3.1) from the *ndc*-1 mutant (FGSC3441; NMF164) and a heterokaryotic transformant (NMF401). The reactions from each strain were pooled and purified by gel electrophoresis using a Qiaquick gel extraction kit (Qiagen, Valencia, CA) and sequenced with primer OMF2244 (Table 3.1).

Results

Mauriceville genome sequencing

The *N. crassa* MV genome was sequenced as paired-end 80-nt reads on an Illumina GAII. In total, 9,384,230 reads were obtained, resulting in a nominal coverage of ~17-fold across the ~41 Mb *Neurospora* genome. Adapter trimmed 76-mer reads were assembled *de novo* into 10,435 *N. crassa* MV contigs with Velvet (Zerbino and Birney 2008), which resulted in ~80% (34,698,139 nt) coverage of the currently assembled *Neurospora* genome, with a minimum contig length of 200 nt and an n50 of 50,295 nt. MV contigs were aligned to the OR reference genome with NUCmer (Kurtz, Phillippy et al. 2004) and the alignment was visualized as a percent similarity plot (Fig. 3.2, bottom panel). The majority of the contigs are >95% identical to the OR reference genome, with the highest similarity occurring in genic and pericentric regions. In contrast, contigs aligning to the centromere and telomere regions are shortest, as we expected for *de novo* assembly of repetitive DNA. These contigs show the least similarity to OR and are coincident with regions where no MV reads matched the OR reference genome (Fig. 3.2). Whole genome alignments showed that the majority of the OR sequence absent from MV is confined to AT-rich regions. This lack of reads is not simply caused by poor assembly or low coverage of AT-rich MV contigs. Rather, we predicted that many short AT-rich reads from MV are sufficiently different to not match the AT-rich OR regions. To lend support to this idea, we examined repetitive regions previously found by Southern analyses to be present in OR but absent in MV (methylated segments 8:A6, 8:F10 and 8:G3) (Selker, Tountas et al. 2003). Here we show that we found no reads in these regions (Fig. 3.3).

SNP analysis of MV versus OR

It is of particular interest to the *Neurospora* community to identify SNPs between OR and MV because these are the most commonly used strains for molecular mapping. Reasonably detailed RFLP maps (Nelson and Perkins 2000) and primer pairs for CAPS mapping (Jin, Allan et al. 2007; Lambreghts, Shi et al. 2009) have been generated. Nevertheless, even with these techniques several iterations of PCR to obtain fragments covering SNPs are required to map mutations into a small region that lends itself to complementation analyses. Moreover, currently known MV SNPs were found by EST sequencing (Lambreghts, Shi et al. 2009), thus covering only the most conserved regions of the MV and OR genomes. One of our goals was to assemble a more complete SNP map of MV to facilitate mapping of both novel and classical mutations.

We identified 168,579 SNPs between the OR reference genome and MV genomic reads after quality filtering SNP calls from MAQ alignments (Li, Ruan et al. 2008), resulting in an average SNP density of 1 SNP per 243 nt. Of these, 64,066 (38.0%) occur in the coding region of genes leaving the remaining 104,513 in intergenic space. 19,898 (31.1%) of the coding mutations cause non-synonymous changes. The SNPs are spaced relatively homogenously along the coding regions with somewhat decreased SNP density around the centromeres, the pericentric regions (Fig. 3.4A). As mentioned above, few SNPs within centromeric core regions are detected because MV centromeric DNA is different from OR centromeric DNA (Figs. 3.2 and 3.4A). The complete set of SNPs is available on the *Neurospora* homepage (http://www.fgsc.net/*Neurospora*/SNPs/SNP_map.htm).

For validation, SNPs were aligned as either the OR or MV version to a dataset containing 248,997 MV ESTs obtained by traditional Sanger sequencing (M. Basturkmen, J. Xu, M. Shi, J. Loros, M. Nelson, M. Henn, C. Kodira, N. Lennon, L. Green, J. Galagan, B. Birren, J. Dunlap, M.S. Sachs, in preparation). With a requirement of perfect alignment, 29,612 of the SNP calls aligned to the ESTs. As expected, almost all (27,865 or 94.1%) were MV, while 1,747 were OR, suggesting a false-positive SNP call rate of 5.9%. We experimentally confirmed 47 randomly selected SNPs by RFLP after *TaqI* digest and found that in this case all (i.e. >97%) were correctly called SNPs (Fig. 3.4B). The small discrepancy between mapping to

ESTs and experimentally validated SNPs suggests that Sanger sequencing of ESTs resulted in slightly higher error rates in SNP calls than direct Illumina sequencing of genomic DNA.

Molecular mapping of *ndc-1* by bulk segregant analysis

The *ndc-1* strain (NMF164; FGSC3441) carries a ts mutation that results in arrest of the nuclear division cycle just prior to DNA synthesis when grown at 37°C (Serna and Stadler 1978). This strain also carries a linked lesion in myo-inositol-1-phosphatase (*inl*), which requires supplementation with myo-inositol for growth (Serna and Stadler 1978). The *ndc-1* gene had previously been mapped by classical three-point crosses to the left of *arg-4* (NCU10468.4) and *inl* (NCU06666.4) on the right arm of linkage group (LG) V (Fig. 3.5B) (Serna and Stadler 1978). To map *ndc-1* and to uncover the molecular cause of the ts phenotype, we crossed NMF164 with MV (NMF37; FGSC2225) (Beauchamp, Horn et al. 1977). Random ascospores were collected from plate lids and germinated by heat shock at 65°C for 60 min. This treatment yielded virtually no (<0.05%) *ndc-1* progeny, confirming the previous suggestion that normal heat-shock was lethal to *ndc-1* ascospores (Serna and Stadler 1978). Consistent with prior linkage analysis (Serna and Stadler 1978), five (~5%) *inl* progeny were recovered. We changed our conditions to heat shock at 60°C for 30 min, which resulted in recovery of 64 (~32%) *ndc-1* progeny, still somewhat lower than the expected 50%.

We extracted DNA from 61 strains to generate two separate pools of DNA for Illumina sequencing. The first pool consisted of equal amounts of DNA from 54 *ndc-1 inl* progeny, while the second pool consisted of equal amounts of DNA from nine *inl*. This second pool was used to map the recombination track that must contain *ndc-1*, as this region of the genome should have MV genomic sequence at *ndc-1*$^{+}$. In contrast, the region with *inl* should contain DNA from the mutant background, while the remainder of the genome should be a mixture of DNA from both parents. The pool composed of DNA from the 54 *ndc-1 inl* progeny was used to find the specific mutation responsible for the *ndc-1* phenotype.

We generated a “SNPome”, a FASTA file with both versions of each SNP between OR and MV by running “SNPome_builder” (Fig. 3.1). We mapped processed 36-nt reads from the *ndc-1 inl* and *inl* pools to the OR and MV SNPome as perfect matches with Hashmatch

(Filichkin, Priest et al. 2009) and ran SNP_profiler on the Hashmatch output to count the number of reads matching to each version of each SNP (Fig. 3.1). This Perl script generates a ratio of reads matching each version of a SNP as a fraction between 0 and 1 to indicate the amount each parental genome contributes to the progeny in bulk. The resulting SNP profiles were viewed in the Argo genome browser to identify contiguous genomic regions in which SNPs originated from a specific parental strain (Fig. 3.5A).

As described above, the region containing the *ndc-1* locus should contain most – if not all – MV-type SNPs in the pool of nine *inl* strains. By this logic, we narrowed the location of *ndc-1* to a 0.85 Mb region on LG V, the only region in the genome with an extended stretch of MV-only SNPs (Fig. 3.5). We next mapped the *ndc-1 inl* reads to the OR reference genome with MAQ (Li, Ruan et al. 2008) and found 2,724 high-quality SNPs in the region containing *ndc-1* (between 2.70 and 3.55 Mb on LG V). NMF164 (FGSC3441) is of mixed genetic background. We found that the *ndc-1* region is highly polymorphic compared to the OR reference genome (Galagan, Calvo et al. 2003), which accounts for the high number of mutation candidates found. To exclude as many SNPs as possible, we subtracted SNPs that were also present in two additional *Neurospora* strains, FGSC322 and FGSC3562, two non-*ndc-1* strains of mixed and related background (McCluskey, Wiest et al. 2011). This procedure reduced the number of candidate SNPs in the *ndc-1* region to 102. Of these, six SNPs caused non-synonymous mutations within exons.

Three predicted SNPs resulted in changes in protein residues that were not conserved (in NCU01215, NCU01114 and NCU01090). Two SNPs resulted in changes from Thr to Ser or Gly to Glu in regions that were in more conserved protein motifs in NCU01172 and NCU01027, respectively. The most obvious candidate gene, however, was *spe-1* (NCU01271), the gene encoding ornithine decarboxylase (McDougall, Deters et al. 1977; Paulus, Kiyono et al. 1982), as it has well-documented roles during the cell cycle, reviewed in (Igarashi and Kashiwagi 2010). The predicted mutation causes a phenylalanine to serine change at residue 132 (F132S). This phenylalanine is conserved in all ODC proteins we investigated, including those from fungi, vertebrates, flies, trypanosomes, nematodes, cnidarians and plants (Fig. 3.6A). To confirm that the predicted SNP was present in individual strains with the *ndc-1* mutation, we determined the sequence of *spe-1* from the original mutant (NMF164, Fig. 3.7),

three *ndc-1* siblings and one *ndc-1*$^{+}$ sibling from the cross with MV (data not shown). In all cases the DNA sequence from the *spe-1* gene matched that expected from high-throughput sequencing.

Complementation of *ndc-1*

To determine if the T to C mutation resulting in the F132S substitution was the cause of the *ndc-1* phenotype, we carried out two experiments. First, we grew *ndc-1* strains at restrictive temperatures in the presence of 1 mM spermidine trihydrochloride, a metabolic product downstream of the reaction catalyzed by ornithine decarboxylase (Fig. 3.6B). Normal growth of *ndc-1* at 37°C after supplementation (Fig. 3.7) strongly suggested that *ndc-1* is indeed an allele of *spe-1*. Next, we integrated the selectable *hph* marker into pGS1, a plasmid that contains the *spe-1* region and that had been used previously to complement *spe-1* mutants (Williams, Barnett et al. 1992). We transformed NMF164 with this construct, pKP19, and selected for hygromycin resistance. Transformants were tested for growth at 37°C to determine if the ectopically integrated plasmid complemented the ts *ndc-1* phenotype. All (7/7) transformants grew at 37°C, strongly suggesting that *ndc-1* is an allele of *spe-1*. Integration of the plasmid and heterozygosity of the *spe-1* gene in the transformant heterokaryon was confirmed by Sanger sequencing (Fig. 3.7). Based on our results we propose to change *ndc-1* to *spe-1*^{ndc}.

Utility of *ndc-1* to block and release cells in the nuclear division cycle

We observed that *spe-1*^{ndc} strains recovered and resumed growth after shifting flasks or slants from 37°C to 22°C. This paves the way to utilizing *spe-1*^{ndc} to synchronize the nuclear division cycle and DNA synthesis, which has been difficult in *Neurospora*. We grew *ndc-1* strains at 37°C in race tubes and shifted them to 22°C to measure the time lag before linear growth resumed. We found that cells stalled for 28.4 hours took 15.8 hours to recover (Fig. 3.8A) while cells stalled for 13.0 hours took 10.5 hours to recover (Fig. 3.8B). In both cases recovery of three replicate strains occurred simultaneously.

Mapping crossovers in single OR x MV progeny

In *N. crassa*, meiosis results in four ordered nuclei which all undergo a single mitosis to make eight nuclei, each of which develops into a separate ascospore in the ascus making the meiotic products easy to isolate (Fig. 3.9). To demonstrate that analysis of segregation by high-throughput sequencing, which uses many data points with low coverage and quality, works in principle we have independently sequenced all sixteen progeny collected from two asci from a cross between OR and MV.

We sequenced genomic DNA from the sixteen strains on the Illumina GAII and HiSeq platforms and obtained 1,098,036 to 17,103,169 reads per strain. The reads were mapped to the OR and MV SNPome and analyzed in the same manner as the pooled *ndc-1* strains. The strains were analyzed visually for crossovers, which are immediately evident and generally mappable to within a few hundred base pairs depending on where the nearest SNPs are (Fig. 3.10). In recombinant strain E of octad 1, 119,484/168,579 of the SNPs (70.9%) were covered at a depth of 1x or greater. Of these 117,771/119,484 (98.6%) segregated as expected in all the reads covering the SNP. Relatively few (1,662 or 1.4%) of the remaining SNPs showed inconclusive segregation due to reads covering the SNP with a mixture of OR and MV alleles. This frequency is likely due to incorrect base calls by the Illumina sequencing pipeline or repeat regions of the genome. A much smaller fraction, 51 SNPs segregated unexpectedly in all reads covering a SNP. Upon closer inspection at least eight of these resulted from a *bona fide* gene conversion (Fig. 3.11) suggesting the error rate for calling segregation at any particular SNP is less than 0.037%.

Discussion

Mauriceville genome assembly and SNP identification

We sequenced the *N. crassa* MV genome to ~26-fold median coverage and mapped reads to the current assembly of the OR genome to identify SNPs between these two commonly used mapping strains. The highest quality SNP calls are available as a resource for the *Neurospora* community on the *Neurospora* model organism homepage at http://www.fgsc.net/*Neurospora*/SNPs/SNP_map.htm. In addition, we carried out *de novo* assembly of MV reads to generate a MV genome that did not rely on a reference genome. *De novo* assembly statistics suggest that we captured 34.7 Mb or 85% of the ~41 Mb genome in

10,434 contigs. Considering the relatively high amount of near-repetitive DNA (Galagan, Calvo et al. 2003), this compares favorably with the *de novo* assembly of the *Sordaria macrospora* genome (Nowrousian, Stajich et al. 2010), which was reference-guided and included longer reads derived from Roche 454 sequencing.

We found that SNPs are distributed evenly over all seven chromosomes with the exception of large blocks of AT-rich DNA present primarily at the centromeres or dispersed as transposon relics stemming from the action of Repeat-Induced Point mutation (RIP) (Selker, Tountas et al. 2003) (Fig. 3.2). We predicted that these regions are highly disparate in MV when compared to OR, and thus examined both read coverage and *de novo* assembly alignment in these regions. Several selected regions that had been previously shown to be absent from MV by Southern hybridization with OR probes (Selker, Tountas et al. 2003) had few or no reads in MV (Fig. 3.3). The AT-rich reads remaining after alignment of MV reads to OR thus constitute dispersed, subtelomeric and centromeric regions and in further studies we will address the evolution of centromeric DNA. Here we show that these regions are highly variable between MV and OR.

Mapping *ndc-1* by bulk segregant analysis

Mapping a mutation by molecular means can be complicated under ideal conditions when the mutant was isolated in a genetic background for which a reference genome is available. When this is the case, mutants are backcrossed many times to isogenic or near-isogenic strains and pre-mutation and backcrossed mutant strains are compared with the idea of identifying the underlying lesion directly as in (Irvine, Goto et al. 2009). This can be a difficult and time-consuming undertaking, even with comparatively small, usually haploid genomes of fungi (C. Yam, K.R. Pomraning, M. Freitag and S. Oliferenko, in preparation; K.R. Pomraning, M. Schmoll, S. Baker and M. Freitag, in preparation). Direct comparisons of two or even several isogenic strains do not usually establish linkage and, to make matters worse, it is generally not known whether the mutation in question is a point mutation, indel, rearrangement or the result of epigenetic changes. In contrast, doing even a single backcross to a strain of known and overall dissimilar haplotype followed by bulk segregant analysis with a dense SNP map is useful as SNP maps immediately establish linkage of mutations to specific

genomic regions, even if the type of mutation is unknown. It is easiest to find point mutations, as we describe here, but with sufficient coverage or paired-end sequencing of bulk segregant pools we predict it will be straightforward to identify indels or genetic rearrangements by high-throughput sequencing.

We have encountered two problems when using bulk segregant analysis. The first problem occurs when the genetic background of the mutant strain is unknown, as is often the case with classical mutants. One example of this case is discussed here, as *ndc-1* was obtained in a strain of mixed genetic background (Serna and Stadler 1978). We treated NMF164 as OR in the cross with MV but we showed that OR was a poor surrogate parent that left us to consider thousands of SNPs not present in the OR background as potential candidates for the *ndc-1* mutation. To alleviate this problem, we analyzed the ~1 Mb segment on LG V that we predicted to contain *ndc-1* in a diverse set of *Neurospora* wild-type strains with similarly mixed backgrounds (McCluskey, Wiest et al. 2011) with the single criterion being lack of the *ndc-1* phenotype. We reasoned that SNPs present in both wild-type and *ndc-1* strains cannot be responsible for the *ndc-1* phenotype, and thus should be subtracted from the pool of candidate mutations. While not ideal, this strategy will become ever more effective as a wider palette of strains are sequenced and made publically available. If, however, *Neurospora* researchers embark on novel screens and selections for mutations, we recommend to use the OR or MV strains as genetic background to accelerate and simplify future molecular mapping.

If the parental strain prior to mutagenesis is available but is not of OR or MV background, the analysis is simplified by setting mapping crosses with OR (because its assembly is superior to that of MV) and sequencing both the pre-mutagenesis parent and the bulk mutant progeny. SNPs identified in the parent strain are subtracted from those identified by bulk segregant analysis from the pool of mutant progeny. This leaves the subset of SNPs introduced by the mutagenesis as candidates to screen.

The second problem encountered is specific to the mating type region on *N. crassa* LG I. This region is idiomorphic in the opposite mating types (Glass, Grotelueschen et al. 1990; Staben and Yanofsky 1990). By sequencing additional *Neurospora* strains of various backgrounds we found that the mating type region and adjacent large regions of LG I

extending across the centromeric DNA, are typically extremely polymorphic (K.R. Pomraning, K.M. Smith, F.J. Bowring, D. Bell-Pedersen, D.E.A. Catcheside, E.U. Selker and M. Freitag, unpublished data). These mating type specific SNPs make mapping of mutations closely linked to the mating type region difficult. If a mutation is linked to this region, progeny should be pooled by mating type and analyzed with a SNPome generated for either *mata* or *matA* of LG I.

It is tempting to sequence only the pre-mutant parent and mutant strains, as has been done with success to identify single point mutations in fission yeast (Irvine, Goto et al. 2009). In organisms with larger genomes, however, bulk segregant analysis is used (Cuperus, Montgomery et al. 2009; Liu, Chen et al. 2010; Wenger, Schwartz et al. 2010). If the locations for recombination events are assumed to be random during meiosis then the power to identify the region containing a mutation by bulk segregant analysis increases as 1/n, where n is the number of progeny sequenced in bulk, and can be roughly modeled as:

$S = [C / (R+1)] n^{-1}$ (1)

where S is the size of the genomic region sought, C is the size of the mutated chromosome and R is the number of recombination events that occur on the mutated chromosome during meiosis. For *N. crassa* LG V we know the size is approximately 6.4 Mb and we can conservatively estimate that a crossover occurs at least once on each chromosome. Thus, for *N. crassa* LG V:

$S = [6.4 \text{ Mb} / (1+1)] n^{-1}$ (2).

In this study we sequenced two pools containing nine (S = 0.85 Mb, Fig. 3.5) and 54 (S = 0.15 Mb, data not shown) progeny (Fig. 3.12). We were able to narrow the mutation to a region larger than expected from equation (2) (S = 0.36 Mb for nine progeny and S = 0.06 Mb for 54 progeny) suggesting that recombination is slightly suppressed in this region of LG V. By generating and carefully pooling DNA from additional progeny, one can narrow recombination blocks containing the mutation(s) in question to smaller sizes, but of course with diminishing rewards for every new progeny added.

<u>Importance of ornithine decarboxylase (ODC) for the cell cycle</u>

We identified the *ndc-1* mutation as F132S in ODC, the enzyme that catalyzes the initial rate-limiting step of polyamine biosynthesis by converting ornithine into putrescine, which is then converted to spermidine and spermine (Davis 2000) (Fig. 3.6B). Putrescine production is important for the function of eukaryotic cells and with few exceptions, e.g. within the Brassicaceae, which use arginine decarboxylase to synthesize putrescine (Hanfrey, Sommer et al. 2001), eukaryotes utilize a conserved ODC enzyme for putrescine production. To unravel which of the many roles of polyamines contribute to the cell cycle stalling observed in the *ndc-1* allele will be the subject of future work. It is clear that polyamines affect a wide variety of cellular processes by interacting with DNA, RNA, and phospholipids (Watanabe, Kusama-Eguchi et al. 1991; Miyamoto, Kashiwagi et al. 1993). Studies in fission yeast point to a spermidine-requiring modification of lysine to the unique amino acid hypusine in eukaryotic translation initiation factor 5A (Park 2006). This modification is essential for cell proliferation and utilizes much of the free spermidine in the cell under spermidine-limiting conditions (Chattopadhyay, Park et al. 2008). Polyamines interact with DNA as cations by binding to phosphates but also form nuclear aggregates of polyamines which effectively protect DNA from endonucleases *in vitro* (D'Agostino, di Pietro et al. 2006). These complexes stabilize DNA structure through condensation, depletion of which could be catastrophic for genomic DNA packaging within the nucleus. Finally, polyamines are required for normal microtubule assembly (Savarin, Barbet et al. 2010), which are presumably necessary for spindle formation after spindle pole body duplication, roughly the point at which the *ndc-1* allele appears arrested (38). Outside of the cell cycle proper, growing hyphae of filamentous fungi require microtubules for the transport of cargo to the Spitzenkörper at the growing hyphal tip (Uchida, Mourino-Perez et al. 2008). Disruption of this process may result in stalled linear growth observed here.

Application of the *spe-1*^{ndc} allele

At present, *spe-1*^{ndc} is the only ts nuclear division cycle mutant known in *N. crassa*. We observed that growth of the mutant is stalled by incubation for several hours at 37°C but can be restored by transfer back to 22°C (Fig. 3.8). Supplementation with 1 mM spermidine allows the strain to grow at 32°C and 37°C (Fig. 3.7A). While picolinic acid, which rapidly

slows DNA accumulation (Martegani, Levi et al. 1980), and deoxyadenosine, a DNA synthesis inhibitor (Fletcher and Trinci 1980), have been used in the past to synchronize the nuclear division cycle in *N. crassa*, we suggest that shifting *Neurospora* from high to low temperature or supplementation with spermidine is less invasive than the former treatments. To that end, we are currently defining the precise growth conditions to synchronize the cell cycle of *spe-1^{ndc}* strains with the goal to generate a system that will allow synchronized DNA synthesis upon addition of spermidine.

Analysis of single recombinant strains

Sequencing all eight strains from a single ordered octad gives us not only the power to see all the crossovers, which occur at a frequency of 1.0±0.7 crossover per chromosome in the two meiotic events looked at (Fig. 3.10B) but also allows detailed observation of rare recombination events without the need for selection at a specific locus. For example, gene conversion, which was originally described in *N. crassa* as aberrant 6:2 segregation (Mitchell 1955; Mitchell 1955), is evident at roughly half the crossover junctions in the two sequenced octads (Fig. 3.13) as well as in regions not closely associated with crossovers (Fig. 3.11). The octad is also useful for looking at post-mitotic segregation which presents as 5:3 segregation and can generally only be found using mismatch repair mutants.

Acknowledgements

We thank Kevin McCluskey (FGSC, University of Missouri, Kansas City, MO) and Scott Baker (Pacific Northwest National Lab, Richland, WA) for access to *N. crassa* raw sequence data prior to publication and helpful comments on the manuscript. Mark Dasenko, Matthew Peterson and Chris Sullivan at the OSU CGRB core facility assisted with Illumina sequencing. We are indebted to the *Neurospora* Functional Genomics Project for primers and the Fungal Genetics Stock Center for *Neurospora* strains. This work was made possible by start-up funds from the OSU Computational and Genome Biology Initiative and grants from the American Cancer Society (RSG-08-030-01-CCG), DOE (DE-FG02-08ER64665) and NIH (P01 GM068087).

Table 3.1: Primers and adapters used.

Name	sequence (5' to 3' direction)
PE_Adapt1-Phos	GATCGGAAGAGCGGTTCAGCAGGAATGCCGAG
PE_Adapt2	ACACTCTTTCCCTACACGACGCTCTTCCGATCT
PE_PCR1.0	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCT
PE_PCR2.0	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTG AACCGCTCTTCCGATCT
PE_ADAPT1_GGGT	^{Phos}CCCAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAG
PE_ADAPT2_GGGT	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGG*T
NCU01271-5F	CTCGACCAAACCTAACAAGC
NCU01271-3R	TTCTACTTGTCCCTCAGTCC
NCU10501.3F	GACCCAACAGAAGATGAAGC
NCU07445.5R	CCTTTGCTTACCTACGTACC
NCU02175.3F	GCGACGGATATATAGTGACC
NCU02054.3R	CAATGAGTAGGGATCGAACG
NCU10932.5F	AGTGTCAAGTCTCTCAGTGG
NCU02753.3R	CTGATGGACTATGCCATACC
NCU03203.3F	TAGGTCTCTGACTGGTGTGC
NCU03292.3F	CCTAGACTAGATGGAGGTGG
NCU08361.3R	GCCTGTGATCTATGCCTACC
NCU00833.3F	CTCGATGGTTACAAGCAAGG
NCU00754.5R	CTGAACAGCAGAGTATCACC
NCU02886.3R	ATGGGCCTACTACCGATACC
NCU02821.3F	TTGTATGTAATTGAGGCACC
NCU08638.3F	TAACGGAGGTTGAGAGTTGG
NCU03565.5R	CATCACTGAAGAGCAGTAGG
NCU03379.5F	AGCTCTTCCACTTGTGATCG
NCU01778.5F	CACTAGAGCTGAGAGGTTCC
NCU01580.5R	GGTGGTTACAGTGCTAGAGG
NCU08483.3R	ACACACTACTGTAAGCGACC
NCU06183.5R	GGATGTACCAAAAGGGTAGG
NCU06272.5F	GGTGGCTGAAGAGAAATAGG
NCU00160.5R	GGAATCACTACTTACCCTGC
NCU00284.3R	GAATGTCTCTTGTCCAGTGC
NCU07522.5R	GTACAACTATTCGAGGGAGG
NCU08554.5R	AGTAACAGGCCAGTCAGACC
NCU04853.3F	GATGGCTGACAGAGATATGC
NCU10472.5F	GACAGAGAGAGGATTTGAGC

NCU07661.5R	GATCACGGTCTACTTCCTCG
NCU07024.5R	CTCATCTTCCTCATACGTCC
NCU03820.3R	GTCATGGGAGAACATGAAGG
NCU01329.3F	GGTGGTTAACTCAACAGTCG
NCU06653.3F	TCCTCCACTTTGACTCTACC
NCU01471.3R	ATGTTCAGGTAGTGGTGAGC
NCU04126.5F	CCTTGGGTAGCTCTTGAAGC
NCU09957.3R	CGTGCAGTAAGTGATAGTGC
NCU07058.3R	GGTTGGTTTGCTGGTATAGG
NCU05615.3R	GATCCCGACTACCATATAGC
NCU05570.5R	GACAGGGAACTCAAAGTACG
NCU04075.3F	GGAGTGGCTATTGTTCATCG
NCU10986.5R	GTCAGTCGATGACATCAAGG
NCU05989.3R	GAACGTCTACATAAGGCTGG
NCU05122.5F	ATGTAGTGTACGACCGTTGC
NCU05024.3R	TTCTACCTACGTTAGGCTCC
NCU05814.5F	GGCTGAGAAAGATTGAGAGG
NCU06034.5R	GAGAAAGGAGTCAACAGAGG
NCU09037.3R	TACCTCTACTCGGCACATCC
NCU09287.3R	GGTACGTCACCAACAAGTGG
NCU10501.3R	GACTTGGAGAGGTAGTGAGC
NCU07445.5F	TGAGATCTGTCACCTACTGG
NCU02175.3R	GTCAGGATACCTGGTCTTGG
NCU02054.3F	CTATCCTTACCTCCATCTCC
NCU10932.5R	GAGGTAAGAGTGCCTAGTGG
NCU02753.3F	CAAAGTGAGGTAGTCTGTGG
NCU03203.3R	ACACACCCCTACGTACTACC
NCU03292.3R	TTGGTGATGCTCCTTAGTGG
NCU08361.3F	GCACAGAGTGCAACATAACG
NCU00833.3R	GTACATACAGCCTCACATGG
NCU00754.5F	ACTCTTGTGCAGGTCATTCC
NCU02886.3F	GTCTTCCTCATATGGTCTCC
NCU02821.3R	CTCGCTAAACTTTTTCTTCC
NCU08638.3R	AGAGACATTAGGTAGGGAGC
NCU03565.5F	GTCAAGTAGTTTGTGGGTGC
NCU03379.5R	CGTCCAGTACTATCTGCTGC
NCU01778.5R	GATAGACAGGAGACAGAAGG
NCU01580.5F	GTTGGTGTACATGACGTAGG
NCU08483.3F	AGTGTACTTGGGCTTGTACC

NCU06183.5F	CCTACCTACCAACTATGAGC
NCU06272.5R	TTCTGCCATCTCTAGGTACG
NCU00160.5F	AGGGCTACATACTGGACAGG
NCU00284.3F	GTTCTAGATAAGGCCTGACG
NCU07522.5F	GTCACAAGAGAGGACAATCG
NCU08554.5F	TTGTTAGTGGGATGGTGTGG
NCU04853.3R	GGACAGATATTCGATGGACG
NCU10472.5R	ACTACCTAGCTACCGAGAGG
NCU07661.5F	ACTCAGTACCCGACAGTAGC
NCU07024.5F	TTGCTGGAAGACGAAGTAGC
NCU03820.3F	GTACTAGTAAGCCGTTGAGC
NCU01329.3R	AAGAGTTCATAGGTGGCTGG
NCU06653.3R	GGGACATGGTAGTCAAGAGG
NCU01471.3F	CTCTCTGGTTATCGTGTTGG
NCU04126.5R	CGTGAAGTTGTTGGTGAAGC
NCU09957.3F	CTGGGAGATATACGAAGTGG
NCU07058.3F	AGGTCAGGAACTATGCTACC
NCU05615.3F	CCAACTCCCTTAAGTCCTCG
NCU05570.5F	CCGCTAGGATAGATACCAGG
NCU04075.3R	AAGAGGAGTAGGTGAGATGG
NCU10986.5F	ATGTGACACAGTCTGCTACG
NCU05989.3F	TACGTCTTCCATGAACCAGG
NCU05122.5R	TCGTCTTCCTACTCTTCTGC
NCU05024.3F	GGTAGGTTGCCTTACTGACG
NCU05814.5R	GTACTGTAGCGTGTCTCTCC
NCU06034.5F	TGGAAAGTCTCAAGTCGTGG
NCU09037.3F	GTTGTAGACTTCCTCTGTGG
NCU09287.3F	AGTTCATCCTCTTGACGAGC
OMF2243	GGCGATGAGGACACATTCTT
OMF2244	ATCAGTCGAGTCGAGGGAAG
NCU01271.5F	CTCGACCAAACCTAACAAGC
NCU01271.3R	TTCTACTTGTCCCTCAGTCC

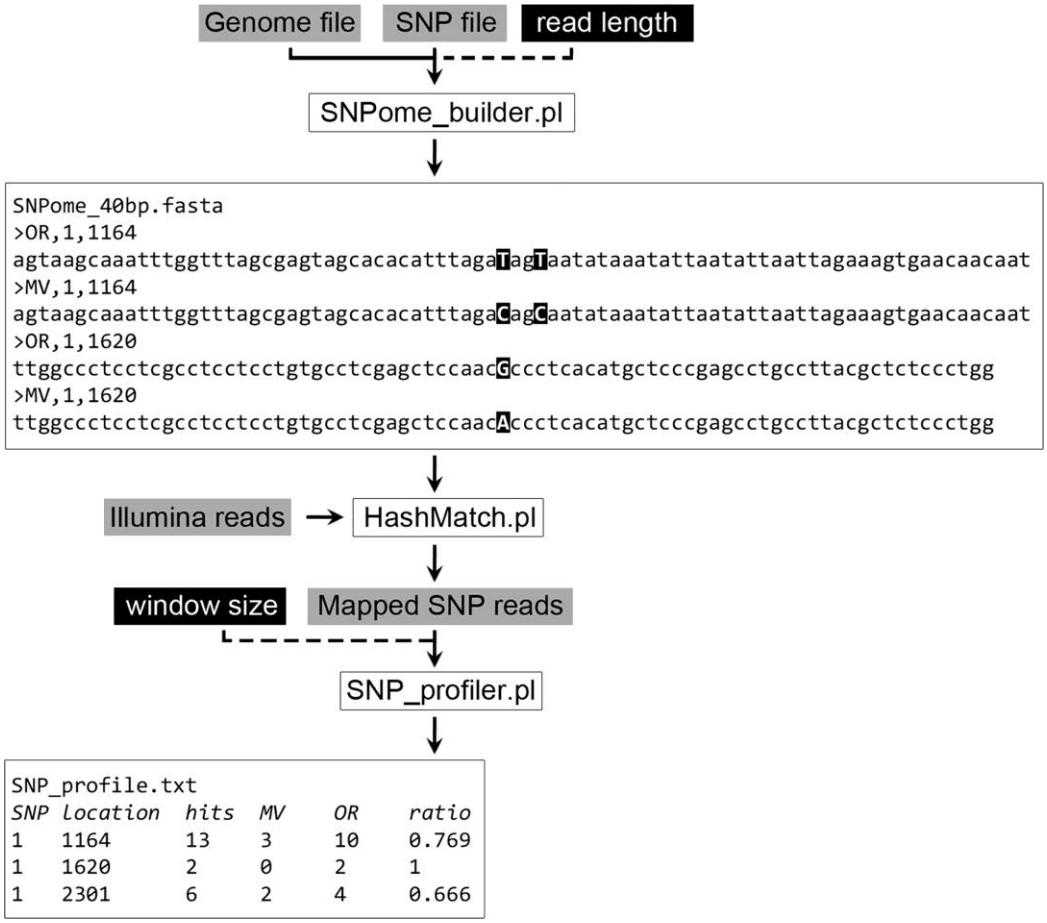


Figure 3.1. Pipeline to generate SNPomes and profile SNPs in bulk segregants. The Perl script SNPome_builder creates a SNPome file (shown here for two example SNPs) given a FASTA formatted genome file, SNP position file and user specified read length. SNPs within the SNPome file between OR and MV are highlighted in capital letters. Hashmatch (Filichkin, Priest et al. 2009) was used to map Illumina high-throughput sequencing reads, the output of which was analyzed by the Perl script SNP_profiler using a 5 kb window size. The output of SNP_profiler indicates the chromosome and position of each SNP, the number of times each version of the SNP was sequenced and the ratio of the versions in a tab delimited file.

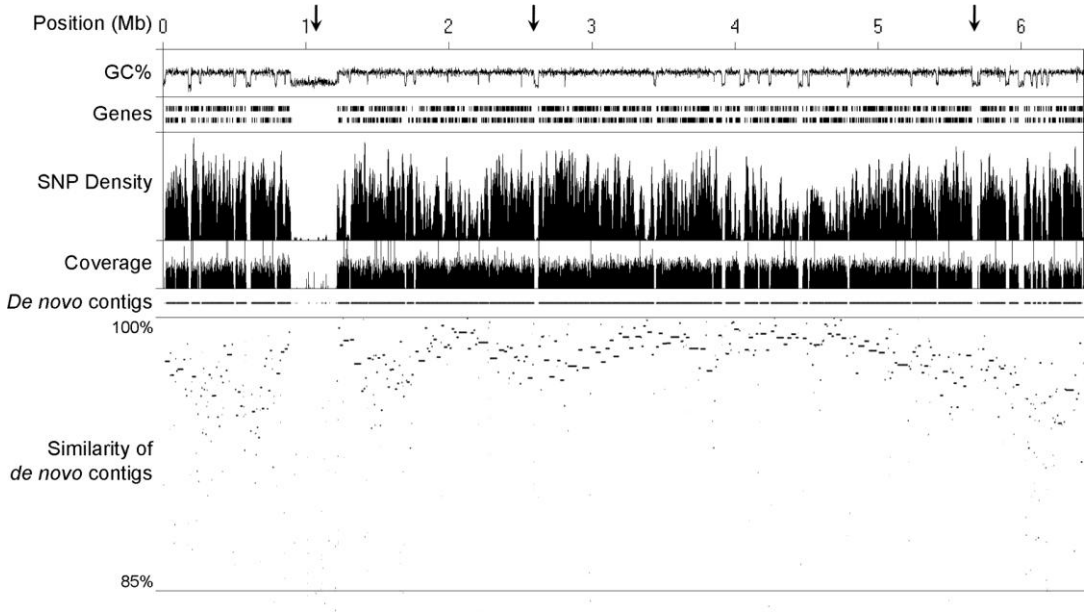


Figure 3.2. Genome sequence of *N. crassa* MV. The SNP density (0-15 SNPs/kb), short read coverage (0-50x) and *de novo* contig coverage of MV Illumina sequencing of LG V was compared to LG V of the OR reference genome. Known and predicted genes and the GC percentage (0-100%) are shown above the mapped reads. Note the absence of mapped reads in most of the centromeric region (large gene sparse block at 1 Mb) and some dispersed AT-rich regions that are composed of RIP-mutated relics of transposons (arrows). The bottom panel shows the similarity of aligned MV contigs that were derived by *de novo* assembly (85-100% nucleotide similarity).

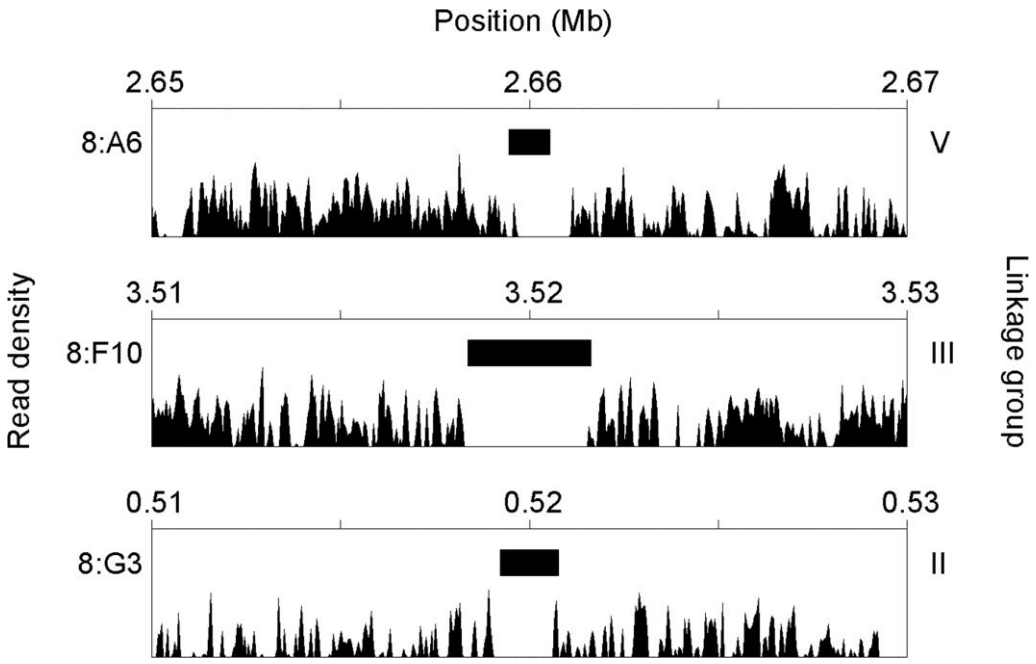


Figure 3.3. Previously identified segments of transposon relics are absent from MV. The top panel shows the short read coverage (0-50x) of the MV sequencing when compared to the OR reference genome for three regions (8:A6, 8:F10 and 8:G3) previously identified to be absent in MV by Southern analyses (Selker, Tountas et al. 2003). The Southern blot probes previously used are indicated by black rectangles.

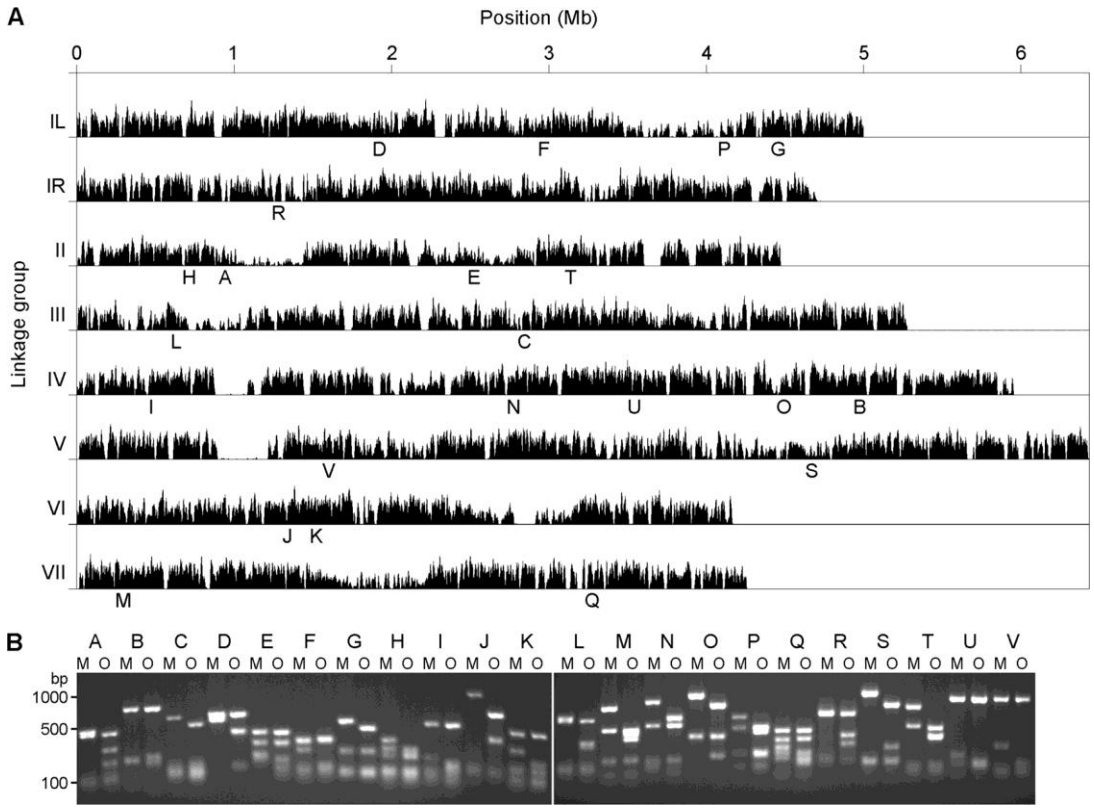


Figure 3.4. SNP map of *N. crassa* MV compared to the OR reference genome. 168,579 SNPs were identified between MV and OR. (A) SNP density plot along all seven of the *Neurospora* chromosomes (0-26 SNPs/kb). (B) A subset of SNPs within a *TaqI* site present in OR (O) but not MV (M) was confirmed by RFLP. Their location in the genome is indicated by letters in (A).

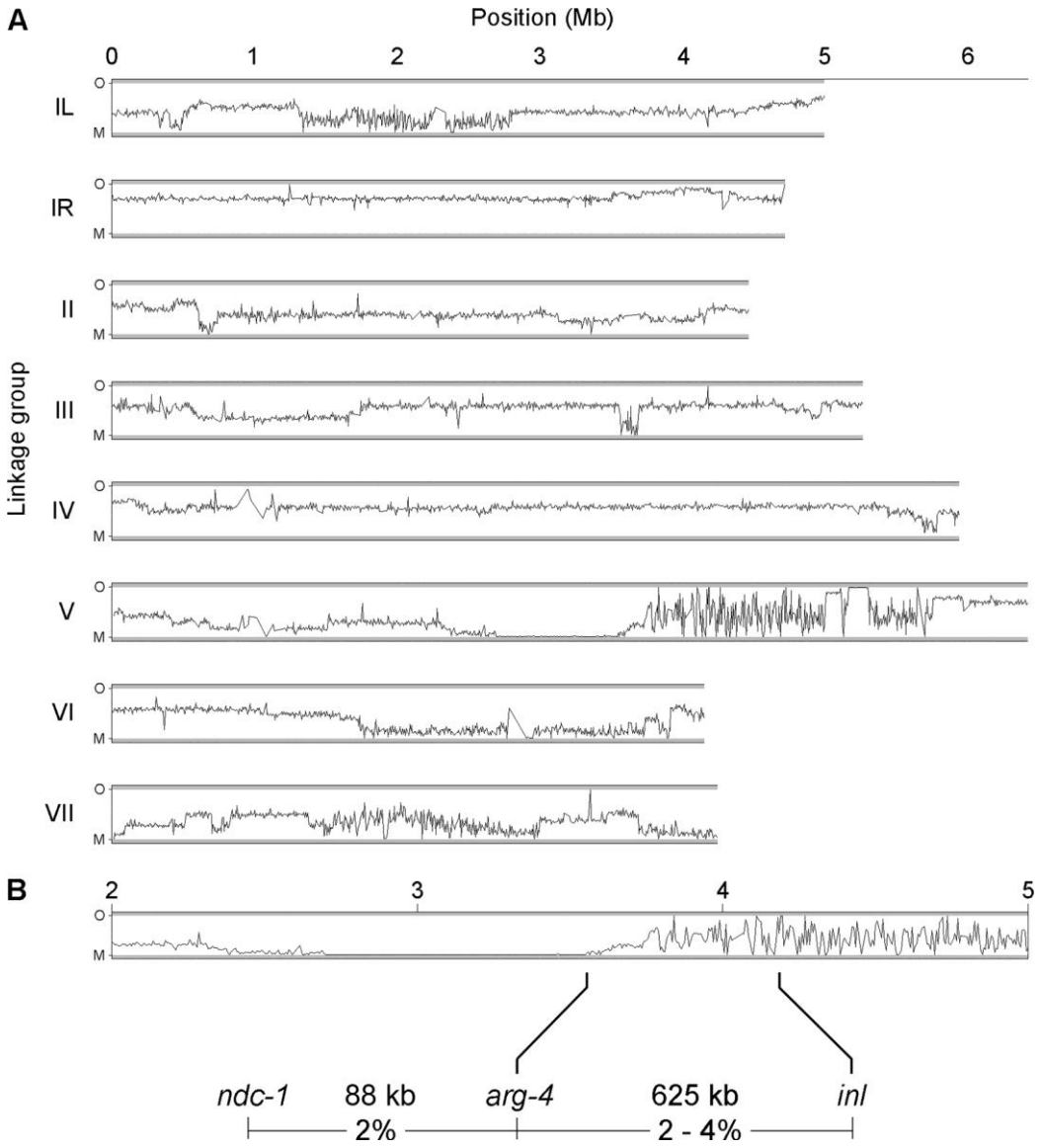


Figure 3.5. Mapping the *ndc-1* mutant. (A) Genomic DNA from 54 *ndc-1*$^{+}$ *inl* bulk progeny was mapped to the OR/MV SNPome. The number of reads obtained from either the OR or MV background was tallied at each SNP and a ratio of OR to MV reads was calculated to indicate the contribution from each genetic background in the bulk progeny. The ratio is shown as 100% OR (O) to 100% MV (M) along all seven *N. crassa* chromosomes as a 5 kb sliding window of the average ratio. Note the 0.85 Mb block on LG V where reads originate entirely from MV; thus this region must contain *ndc-1*. (B) The region of LG V originating from the MV parent is shown in detail. Linkage analysis had previously mapped *ndc-1* to the left of *arg-4* (Serna and Stadler 1978).

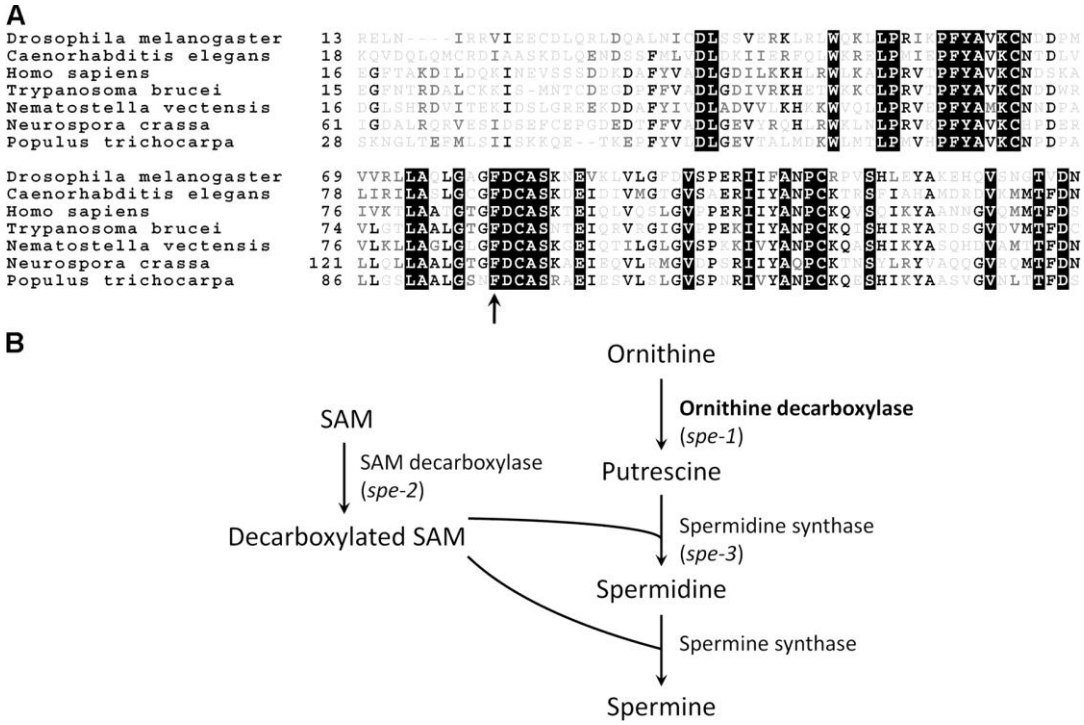


Figure 3.6. The *ndc-1* mutation maps to the gene for ornithine decarboxylase, *spe-1*. (A) We identified a Phe to Ser mutation at residue 132 (F132S, indicated by an arrow) in the *N. crassa* ODC. Alignment of ornithine decarboxylase proteins from a variety of organisms including wildtype *Neurospora* indicates the mutated residue is highly conserved. Only a small region of ODC is shown here. (B) Production of polyamines in *N. crassa*. Mutant alleles are indicated in italics. SAM, S-adenosylmethionine.

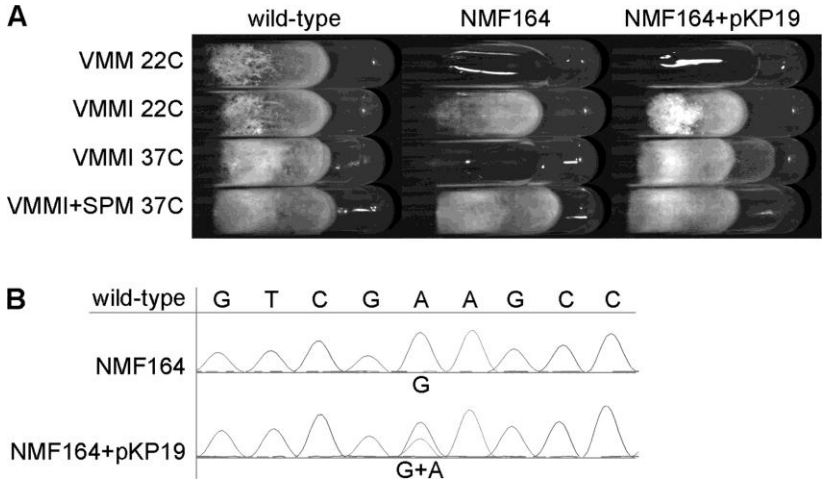


Figure 3.7. The *ndc-1* ts phenotype can be reverted by supplementation with spermidine or integration of a wildtype copy of *spe-1*. (A) VMM slants with or without myo-inositol (VMMI) and spermidine (SPM) were inoculated with conidia from NMF398 (*ndc-1*$^{+}$ *inl*$^{+}$), NMF164 (*ndc-1 inl*) or NMF164 transformed with pKP19, a plasmid containing the wild-type *spe-1* gene (NMF401; *ndc-1*$^{+}$ *inl*). Images were captured after three days of incubation at either 22°C or 37°C. (B) Complementation of *ndc-1* was confirmed by Sanger sequencing. Sequencing peaks for both the wildtype and *ndc-1* allele indicate that both are present in the heterokaryotic transformant. The base(s) present at the SNP site are indicated.

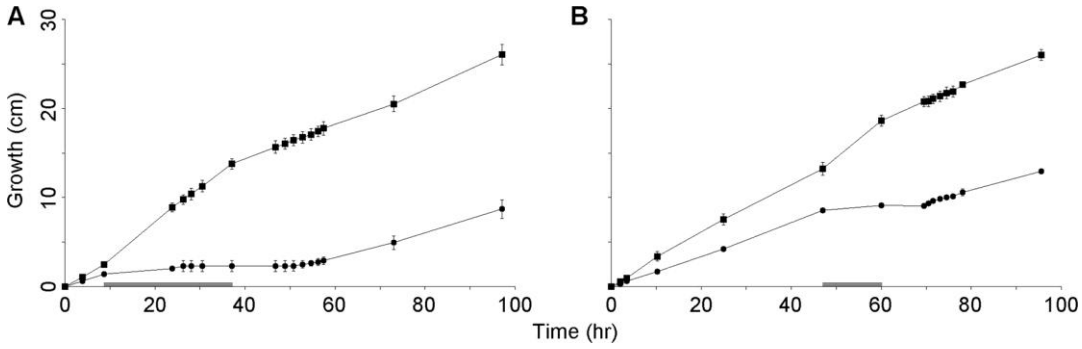


Figure 3.8. Growth arrest and release of ts *spe-1*^{ndc}. Conidia from NMF398 (*ndc-1*$^{+}$ *inl*$^{+}$; squares) or NMF344 (*ndc-1 inl*; circles) were inoculated on VMMI in race tubes at 22°C to monitor linear growth. The conidia were allowed to germinate and establish linear growth prior to transfer to 37°C (indicated by a thick bar on the x-axis). The strains were maintained at 37°C for at least 13 hours and then transferred back to 22°C to determine the lag period before linear growth resumed. Error bars indicate the standard deviation between three replicates.

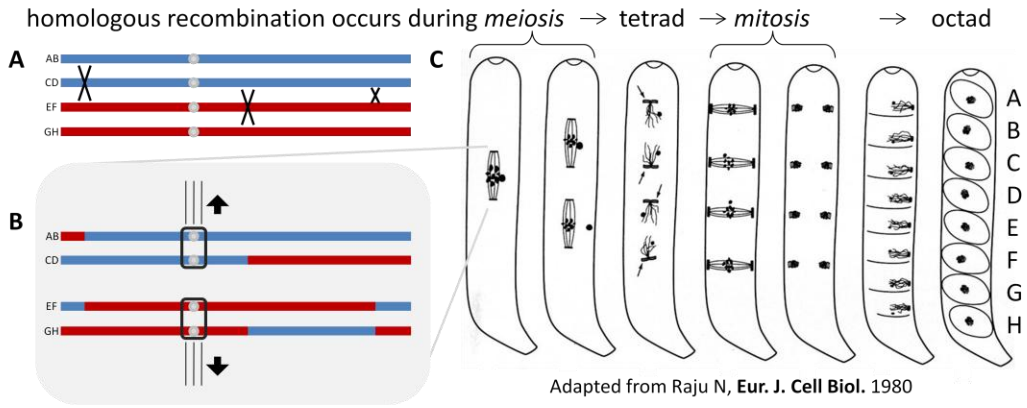


Figure 3.9. Meiotic segregation in *N. crassa*. Two copies of each chromosome, represented by blue and red bars with grey centromeres, from two parents of a cross, each made up of two strands of DNA (e.g. AB) enter meiosis. (A) The chromosomes undergo homologous recombination, sometimes resulting in crossovers and (B) segregate into four recombinants at the tetrad stage. (C) These undergo a single post-meiotic mitosis before ordered ascospore delineation. The eight single strands of DNA that enter meiosis are represented by ascospores A-H.

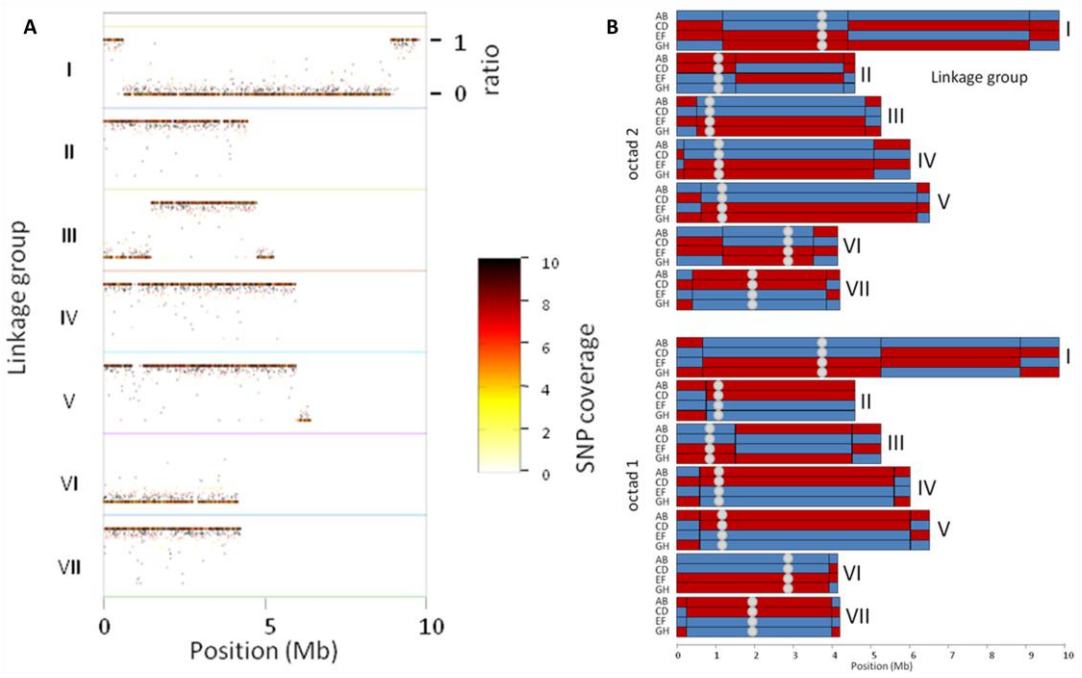


Figure 3.10. SNPome mapping a single recombinant strain. (A) Illumina reads from a *N. crassa* strain purified by ascospore isolation after crossing OR to MV were mapped to the OR and MV SNPome. Each SNPs ratio was plotted at its position along the seven *N. crassa* chromosomes. A ratio of 1 indicates the SNP is OR type, while 0 is MV type. All 119,484 SNPs with > 0x coverage from strain E of octad 1 are shown. (B) Recombination maps of all seven chromosomes for all 16 strains from the two sequenced octads where blue is OR and red is MV. Grey circles indicate the centromeres. The mitotic pairs are collapsed since they are identical at this resolution.

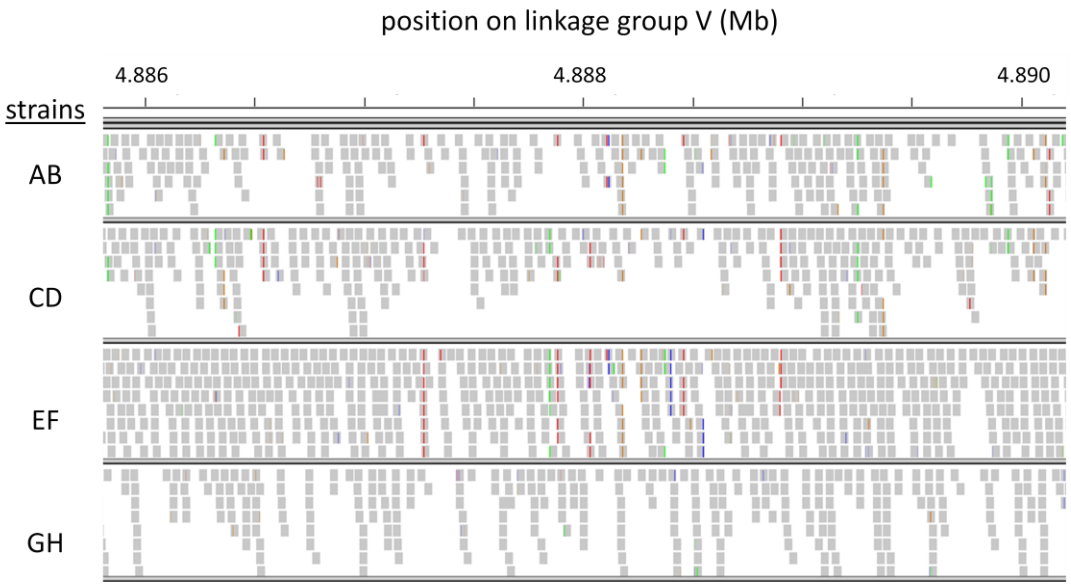


Figure 3.11. Gene conversion event in octad one: Screen shot of Illumina read alignments in the Integrative Genomics Viewer (Robinson, Thorvaldsdottir et al. 2011) on linkage group V around which the chromosomes originate from MV in strains A-D and OR in strains E-H. For ~2 kb the reads are MV type in strains E and F indicating a gene conversion event. SNPs in comparison to the OR reference genome are indicated by colored bands in the reads and segregate 6:2 at the gene conversion and 4:4 elsewhere.

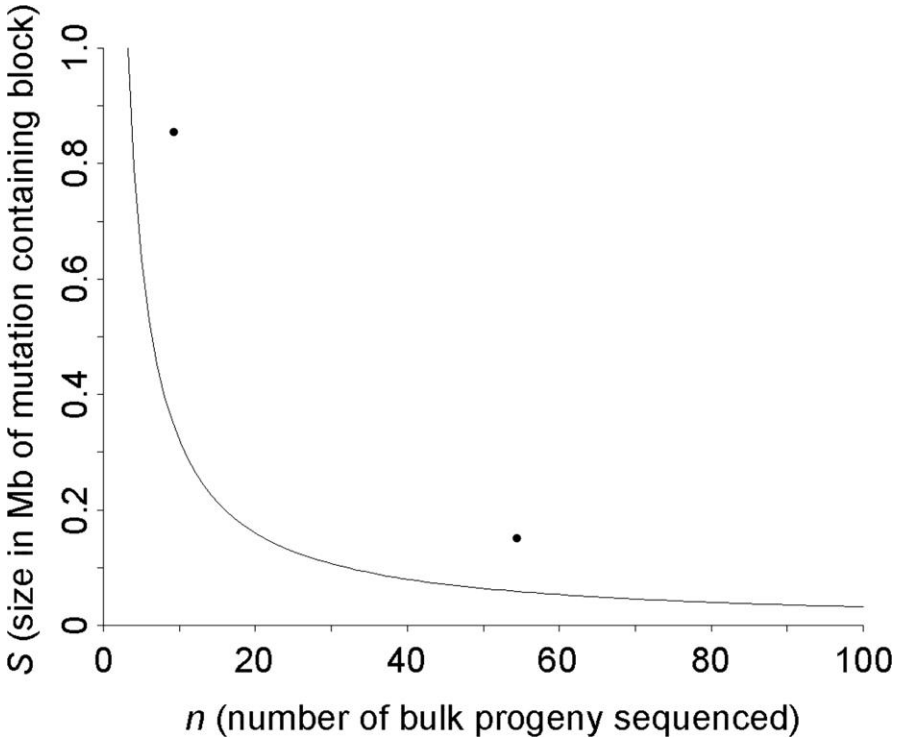


Figure 3.12. Predicted relationship of number of progeny in pools to expected size of mutation-containing region. Equation (1) predicts the relationship between number of progeny sequenced in bulk (n) and the size of the mutation-containing region found (S). It is plotted for *N. crassa* linkage group V assuming size of the chromosome (C) equals 6.4 Mb and the number of recombination events on chromosome V during meiosis (R) equals one. We experimentally narrowed (S) to a 0.85 Mb region using 54 pooled progeny and a 0.15 Mb region using 9 pooled progeny (data not shown) as indicated by two circles.

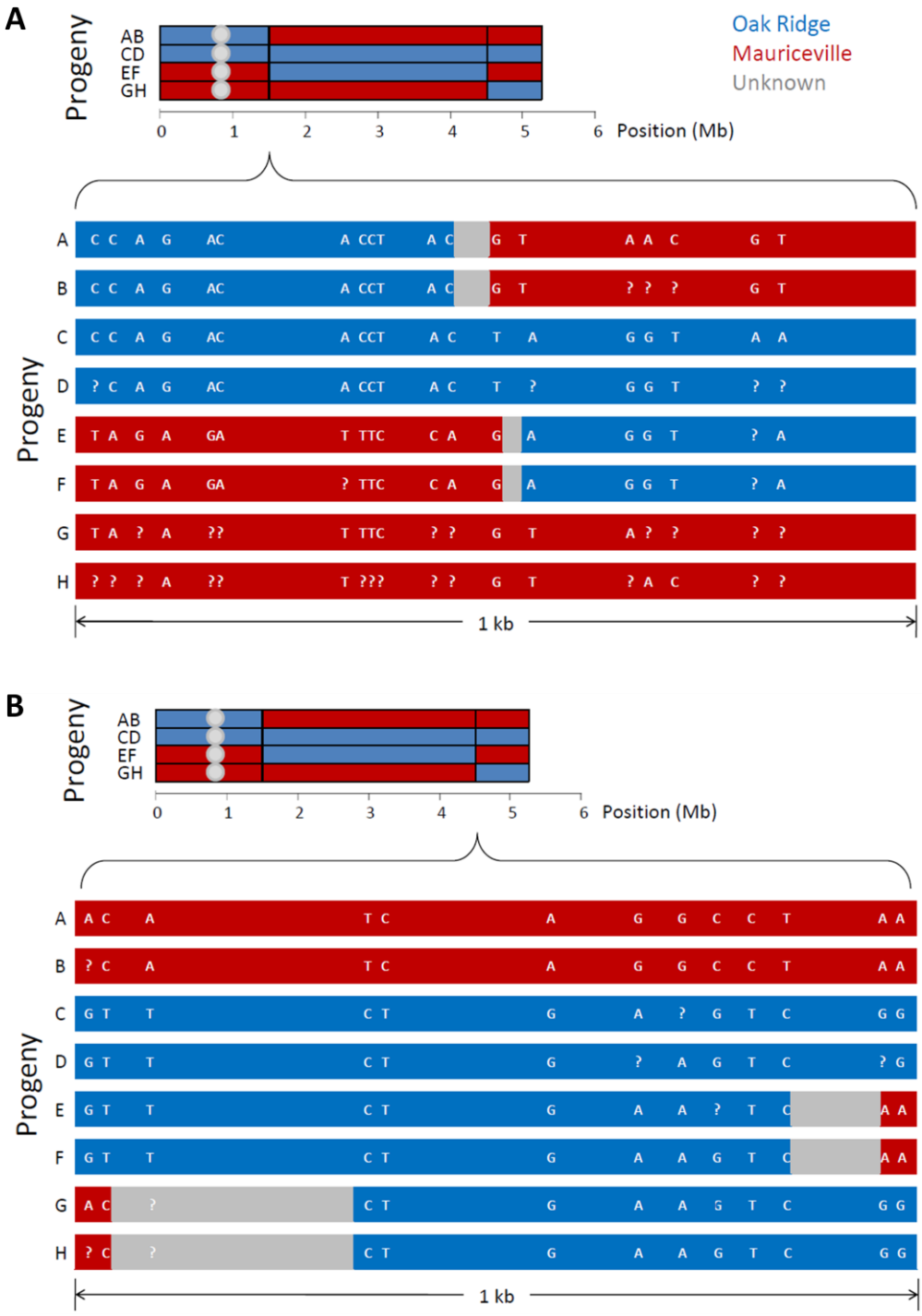


Figure 3.13. Crossovers have short associated gene conversions. The base calls at the OR/MV SNP sites identified by high-throughput sequencing in progeny A-H of octad one are shown for the crossover junctions on linkage group III which occur at (A) 1.4 and (B) 4.5 Mb. Bases without sufficient coverage are indicated by a question mark. Gene conversion is indicated by 6:2 segregation of the progeny.

A large dynamin-like protein affects repeat-induced point mutation and recombination frequency in *Neurospora crassa*

Kyle R. Pomraning, Ann T. Kobsa, Eric U. Selker and Michael Freitag

Manuscript in preparation for *PLoS Genetics*

Abstract

Repeat-induced point mutation (RIP) and premeiotic recombination affect gene-sized duplications in many filamentous fungi. RIP causes C:G to T:A transition mutations while premeiotic recombination can result in loss of repeated DNA segments (Selker 1990). Both processes occur after fertilization but prior to meiosis and can be very efficient, in some cases mutating and/or deleting the duplication in essentially every nucleus that carried it. We employ genetic approaches to elucidate the mechanism of premeiotic recombination and RIP. Here we report the successful identification of semi-dominant mutations that affect both of these processes. We used UV mutagenesis, followed by a screen for reduced inactivation (indicative of RIP) of linked duplications of *hph* and *pan-2*. We identified two mutations by bulk segregant analysis and high-throughput Illumina sequencing. Single point mutations were found in a gene encoding a novel dynamin-like long GTPase, albeit in different conserved domains. The mutated gene, which we call reduced RIP and recombination 1 (*rrr-1*), is expressed in vegetative hyphae and localizes to the nuclear membrane. Both recombination and RIP frequencies are affected, supporting the idea that these processes are mechanistically or temporally linked.

Introduction

Repeat-induced point mutation (RIP) and premeiotic recombination affect repetitive DNA in many filamentous fungi. RIP causes C:G to T:A transition mutations in duplicated DNA regardless of context (Selker, Cambareri et al. 1987; Cambareri, Jensen et al. 1989; Fincham, Connerton et al. 1989; Irelan, Hagemann et al. 1994), while premeiotic recombination can result in loss of repeated DNA segments in direct repeats (Irelan, Hagemann et al. 1994). Both processes occur after fertilization but prior to karyogamy and meiosis. They can be very efficient, in some cases mutating both copies of the duplication and/or deleting the duplicated segment in every nucleus that carried it (Selker 1990). During this period the parental nuclei of opposite mating type are estimated to undergo seven to ten rounds of nuclear division (Emerson 1966). Within the nuclear division cycle, RIP is expected to occur during G1 or S phase since the entire genome would be subject to RIP if it detected duplications after DNA replication. At least in *Neurospora crassa*, RIP has countered the

expansion of gene and transposon families (Galagan and Selker 2004), suggesting that genome streamlining and protection from transposition events may yield long-term benefits to *Neurospora* populations.

The mechanism for duplication detection during RIP remains unknown. Multicopy DNA is not thought to be detected by an RNA mediated mechanism, as the nuclei in which RIP occurs exist in a common cytoplasm where RNA is thought to be relatively freely diffusible. Such RNA would affect unpaired copies of repetitive segments in nuclei with duplications and likely even single-copy segments in nuclei of the opposite mating type. Based on numerous crosses, in which nuclei with single-copy genes have never shown traces of RIP, and evidence for mutation of duplicated segments in pair wise fashion in successive mitoses, this idea has been excluded (Faugeron, Rhounim et al. 1990; Selker 1990). Thus, the current hypothesis is that RIP is dependent on an unknown DNA-pairing mechanism. The mechanism for deletion of tandem repeats, which is dependent on aberrant pairing of repetitive DNA, has been studied thoroughly in yeast in the context of double strand breaks (Sugawara and Haber 1992; Lyndaker and Alani 2009) and was found to be dependent on genes which have orthologues in *N. crassa*.

So far, neither reverse nor forward genetic screens have shed much light on the DNA-pairing mechanism that may be required for RIP. Mutation of a gene, *rid*, that encodes a cytosine methyltransferase-like protein has been shown to abolish RIP in homozygous crosses (Freitag, Williams et al. 2002). In addition, point mutations in histone H3 result in semi-dominant RIP defects and also affect cytosine DNA methylation in the vegetative life stages (M Freitag, AT Kobsa, S Honda, S Friedman and EU Selker, in preparation). One such mutation, $hH3^{dim4}$, was recovered in the same forward genetic screen that resulted in two mutants we further characterize here. We screened for reduced inactivation of linked duplications of *hph* and *pan-2*, resulting in normal dark ascospores and hygromycin-resistant (Hyg^R) progeny, when in crosses with wildtype strains predominantly light brown ascospores were recovered that resulted in hygromycin-sensitive (Hyg^S) cultures. Here we report on the identification and characterization of semi-dominant mutations that affect the frequency of RIP and the frequency of deletion of one or several copies of direct tandem repeats.

Materials and methods

Strains and growth conditions

Neurospora crassa strains were grown at 32 C in Vogel's minimal medium (VMM) with 1.5% sucrose and supplemented with 10 mg/L pantothenate or 500 mg/L histidine when appropriate (Davis 2000). Crosses were performed on modified synthetic crossing medium (SC) (Russo, Sommer et al. 1985) with 0.4% sucrose and supplemented with 10 mg/L pantothenate or 250 mg/L histidine when appropriate(Davis 2000). Ascospores were recovered in water for the scoring of RIP after four weeks and after two weeks for all other crosses. Germination of ascospores was induced by heat shock at 65 C for 1 hour on VMM with FGS (0.5 g/L fructose, 0.5 g/L glucose, 20 g/L sorbose) as the carbon source and 10 mg/L pantothenate or 500 mg/L histidine when appropriate. Strains were tested for mating type by crossing with *fl* strains FGSC4317 and FGSC4347 (Davis 2000). Strains N1477 (*matA*; *pan-2*^{B3}; *pan-2*$^{+}$*-hph*$^{+\ \text{multi-copy}}$) and N1478 (*mata*; *pan-2*^{B3}; *pan-2*$^{+}$*-hph*$^{+\ \text{multi-copy}}$) were constructed by transforming FGSC2248 (*pan-2*^{B3}) with pAH35, a plasmid containing *pan-2* and *hph* sequences. Linear growth assays were performed on VMM with 1.5% sucrose in race tubes at 25C (Davis 2000). The NCU05936 deletion strain from the *N. crassa* single gene deletion collection (see Table 4.1), was crossed to N2868 (*his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$) to obtain deletion strains of both mating type with the *tdimerRed* marker. Progeny were collected and assayed for mating type, hygromycin resistance (Hyg^{R}) and *tdimerRed* phenotypically and for correct integration of *hph* by Southern analysis.

Tagging with Green Fluorescent Protein (GFP)

We made constructs for tagging NCU05936 at the endogenous locus by fusion PCR based on pZero-GFP-loxP-hph-loxP as described previously (Honda and Selker 2009; Smith, Phatale et al. 2011). The 5′ and 3′ flanking regions of NCU05936 were amplified from genomic DNA of N1477 with OMF2721 and OMF2722 and with OMF2723 and OMF2724, respectively (Table 4.2). A fragment containing GFP and *hph* was amplified from pZero-GFP-loxP-hph-loxP with OMF83 and OMF84. The 5′ and 3′ flanks were fused with the GFP-hph cassette by PCR with primers OMF1053 and OMF1054. GFP tagging constructs were transformed into strain N3011 by electroporation (Colot, Park et al. 2006). Hyg^{R} colonies were selected and

backcrossed to N2928 to obtain homokaryotic strains. The homokaryons were confirmed to express NCU05936-GFP from the endogenous locus by Southern blot and western blot using anti-GFP antibody (Santa Cruz, sc-9996). All strains used are summarized in Table 4.1.

Forward genetics screen

Conidia from strains N1477 and N1478 were UV mutagenized and crossed. Because of RIP, crosses of unmutagenized N1477 x N1478 result predominantly in tan (*pan-2*) ascospores. Putative mutants with dominant or semi-dominant defects in RIP were detected by screening for perithecia that ejected predominantly black (*pan-2*$^{+}$) ascospores.

Mutant mapping by bulk segregant analysis and high-throughput sequencing

NMF37 (Mauriceville-1-c), which has at least 168,579 SNPs compared to any other strains used in this study (Pomraning, Smith et al. 2011) was crossed to N1516 and N1522. One hundred recombinant progeny from each cross were collected and tested for their RIP phenotype by crossing to either N2944 (*mat-A*; *rid*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$) or N2945 (*mat-a*; *rid*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$). The progeny from these crosses were germinated and after 18 hours, 100 germinated progeny from each cross were examined for red fluorescence and the percentage of RIP (RIP %) calculated as:

$$\text{RIP \%} = 1 - \frac{2(\text{RFP}^{+}\text{progeny})}{(\text{total progeny})}$$

because only half the progeny are expected to have the *tdimerRed* gene. The progeny with RIP frequency in the lowest quartile from each original mutant were used to generate genomic DNA as described previously (Pomraning, Smith et al. 2009). Approximately 100 ng of DNA from each progeny were pooled in equimolar amounts in a total of 200 uL TE buffer. Each pool, as well as 5 ug of DNA from N1477 and N1478 were separately sheared at 4 C to between 200 and 1000 bp fragment length using a Bioruptor (Diagenode, Liège, Belgium) set on high for 20 minutes with 30 second on / off cycling. The DNA was end repaired and an A overhang added as described previously (Pomraning, Smith et al. 2011). Single-end Illumina sequencing adapters with 5 bp barcodes were ligated to the DNA and libraries containing

~250 bp inserts were prepared and sequenced (single-end, 50 nucleotides) on an Illumina GAII genome analyzer (Illumina, San Diego, USA).

Reads originating from N1477, N1478 and the two RIP mutant pools were parsed by their 5 bp barcodes, which were then removed. Reads were converted from FASTQ to FASTA format and PCR duplicates were removed. The FASTA reads were mapped to a 45 base pair SNPome to determine the contribution of each parent strain in the bulk progeny at every SNP site as described previously (Pomraning, Smith et al. 2011). The output was visualized using Gnuplot v4.4.0 (http://www.gnuplot.info/).

The FASTQ reads were mapped using the maq.pl v0.4.7 easyrun command (Li, Ruan et al. 2008). Mapview files were made by running maq mapview all.map and imported into MapView v3.4 (Bao, Guo et al. 2009). Mapview was used to call SNPs with variant frequency >= 0.8 and coverage >= 3 between the parent strains and pooled mutants and Assembly 10 of the *N. crassa* reference genome (http://www.broadinstitute.org/annotation/genome/*neurospora*/MultiHome.html). SNPs found in either of the parent strains were removed from further analysis. Candidate SNPs were analyzed for coding sequence context using a custom Perl script "SNP_annotator".

Verification of SNPs at *NCU05936*

DNA flanking candidate SNPs was amplified by PCR in three independent reactions from N1516, N1522, N1477 and N1478 using OMF2565 and OMF2566 (Table 4.2). The PCR products were pooled and purified using a Qiagen PCR purification kit (Valencia, USA) and sequenced from both ends using OMF2565 and OMF2566 at the CGRB core lab at Oregon State University.

Bioinformatic characterization of NCU05936

Homologs of NCU05936 within other sequenced fungi of the Ascomycota were found using the Cluster tool from JGI (Run 1697, Cluster 4366, Ascomycota, Feb 2012) (Grigoriev, Nordberg et al. 2012). All 45 proteins within the cluster were aligned by MUSCLE (Edgar 2004) using Mega v5.05 (Tamura, Peterson et al. 2011). To identify all dynamin-like proteins within a given species the dynamin-GTPase domain protein sequence of NCU05936 was used

as bait while blastp and tblastn searching. Hits matching protein sequences were screened for those that had domains identifiable as the dynamin-GTPase domain, middle domain, and GED domain using InterProScan (Quevillon, Silventoinen et al. 2005).

Isolation of *Neurospora* nuclei

Conidia from GFP tagged strains were inoculated in VMM with 2% sucrose and 500 mg/L histidine and grown for 24 hours at room temperature on a shaker. A few grams of germinated conidia were harvested and frozen in liquid nitrogen prior to being ground into a fine powder using a cooled mortar and pestle. All subsequent steps were performed at 4 C. The ground tissue was resuspended in 1mL/g tissue buffer A (1M sorbitol, 7% Ficoll (MW 400,000), 20% glycerol, 5 mM Mg acetate, 3 mM $CaCl_2$, 50 mM Tris-HCl pH 7.5, 3 mM DTT, 10 mg/L pepstatin, 10 mg/L leupeptin and 1 mM PMSF) and filtered through cheese cloth. Two volumes Buffer B (10% glycerol, 5 mM Mg acetate, 25 mM Tris-HCl pH 7.5, 10 mg/L pepstatin, 10 mg/L leupeptin and 1 mM PMSF) was added for each volume of filtrate. This was layered onto buffer A:B (2.5:4) in round bottom Oak Ridge tubes (1.25x of the original volume of buffer A), followed by 7 min of centrifugation in a swinging bucket rotor at 3000 x*g*. Supernatant was layered on buffer D (1 M sucrose, 10% glycerol, 5 mM Mg acetate, 25 mM Tris-HCl pH 7.5, 1 mM DTT, 10 mg/L pepstatin, 10 mg/L leupeptin and 1 mM PMSF; 1/7 the original volume of previous centrifugation) in round bottom Oak Ridge tubes and centrifuged for 15 min in a swinging bucket rotor at 9400 x*g*. The supernatant was removed and nuclei resuspended in nuclei storage buffer (25% glycerol, 5 mM Mg acetate, 0.1 mM EDTA, 25 mM Tris-HCl pH 7.5 and 3 mM DTT) and frozen in liquid nitrogen. To analyze nucleoplasm versus the nuclear membrane, the resuspended nuclei were sonicated in nuclei extraction buffer (5% glycerol, 0.4 M NaCl, 0.1 mM EDTA, 5 mM $MgCl_2$, 20 µM $ZnCl_2$, 15 mM Hepes pH 7.5, 1 mM DTT, 1 mg/L pepstatin, 1 mg/L leupeptin and 0.2 mM PMSF) for 10 min (30 sec on/off cycles) in a Bioruptor (Diagenode, Denville, NJ) set on high and centrifuged at 8000 rpm for 1 min in a table top centrifuge to separate the pellet (membrane) from the supernatant (nucleoplasm). Western blots were performed using anti-GFP antibodies (Santa Cruz, sc-9996).

Results

Two mutants display similar reduced RIP and recombination phenotypes

Two *N. crassa* mutants, N1516 and N1522, were identified by their semi-dominant reduction of inactivation of duplicated *pan-2* and *hph* marker genes in progeny after a cross of their mutagenized parents, N1477 and N1478. We have characterized the fate of the duplicated *pan-2* and *hph* markers in random progeny and found that they are frequently subject to RIP and/or deletion of at least one copy, presumably by a single-strand annealing like mechanism. RIP was observed in 27/48 (56%) random progeny examined while deletion was observed in 13/48 (27%) random progeny examined from a cross between N1477 and N1478 (Fig. 4.1). These numbers are somewhat lower but in general agreement with previous RIP assays and assays for deletion of tandem repeats in *N. crassa* (Selker 1990; Irelan, Hagemann et al. 1994). We found that in heterozygous crosses between either of the mutants, N1516 or N1522, and the wildtype parent N1477 both the frequency of RIP and deletion was decreased (Fig. 4.1) suggesting that the two processes may be linked either mechanistically or spatio-temporally. We thus call this phenotype "reduced RIP and recombination" (RRR) and the mutants *rrr*.

Bulk segregant analysis identifies independent mutations in NCU05936 in *rrr* mutants

Both RIP mutants, N1516 and N1522, share a similar genetic background with the *N. crassa* reference genome, Oak Ridge (OR) and were thus crossed to Mauriceville (MV, NMF37), a strain of dissimilar genetic background with a minimum of 168,579 annotated SNPs compared to the reference genome (Pomraning, Smith et al. 2011). Recombinant progeny in which linkage of the RRR phenotype could be studied were collected and tested for presence of likely *rrr* genes by crossing either mutant to RIP tester strains (N2944 or N2945) that carry *tdimerRed*, which has a short duplication of the *rfp* gene (Campbell, Tour et al. 2002; Freitag and Selker 2005). The progeny segregated into three pools based on their RIP phenotype. The least mutated quartile was considered *rrr* and these strains were mostly *mata*, while the most mutated quartile was considered wildtype (rrr^{+}) and these strains were mostly *matA*. The remaining half of the progeny were intermediate in phenotype and were not considered for subsequent linkage analysis. Based on these data alone, one would have

surmised linkage between *rrr* and *mat.* Linkage between mutations altering RIP frequency and the mating type locus has been observed previously (Tamuli and Kasbekar 2008). These observations are likely caused either by intrinsic differences in susceptibility to RIP in the tester strains of opposite mating type or control of RIP frequency originating within *mat* or a closely linked region. We are in the process of investigating these hypotheses (KR Pomraning and M Freitag, unpublished results).

DNA from the recombinant *rrr* progeny was pooled and sequenced on an Illumina HiSeq2000 sequencer as described previously (Pomraning, Smith et al. 2011), while each of the parent strains N1477 and N1478 was sequenced individually. The two pool and the two parent libraries were sequenced in a single lane using single-end prefix barcoded adapters (Pomraning, Smith et al. 2012). Between 7,867,114 and 18,763,923 45-nt reads were obtained after parsing barcodes for each of the four samples. The reads from the *rrr* pools were mapped to a SNPome generated between the OR and MV strains to map the region responsible for the RIP phenotype by bulk segregant analysis (Pomraning, Smith et al. 2011). In both RIP pools, the region most tightly linked to *rrr* is linkage group (LG) VI from 1.9-3.5 Mb in the N1516 pool and 1.3-3.8 Mb in the N1522 pool (Fig. 4.2A). The parent strains are both highly similar to OR as expected (data not shown).

Reads from the two mutant pools and the two parent strains were mapped to Assembly 10 of the *N. crassa* reference genome using MAQ (Li, Ruan et al. 2008). SNPs were identified using Mapview (Bao, Guo et al. 2009). SNPs unique to the mutants were annotated using the custom Perl script “SNP_annotator”. Within the N1516 mutant pool two nonsynonymous SNPs on LG VI from 1.9-3.5 Mb were identified in genes NCU05547 (A199T) and NCU05936 (L496P). Within the N1522 mutant pool two nonsynonymous SNPs on LG VI from 1.3-3.8 Mb were identified in genes NCU05146 (Q488L) and NCU05936 (G325N) (Fig. 4.2). Thus, NCU05936 seemed the most likely candidate for mutations causing the RRR phenotype, since independent non-synonymous mutations occurred in this gene in both mutants. Sanger sequencing across the segments with the NCU05936 mutations was carried out for the parental strains N1477 and N1478 as well as the mutants N1516 and N1522 (Fig. 4.2B). This verified that both mutants harbor independent mutations in NCU05936, both of

which occur in highly conserved residues in essential GTPase motifs (Fig. 4.2). We propose to call NCU05936 *rrr-1*.

Strains with deletions of *rrr-1* act semi-dominantly to reduce inactivation of duplications

To independently test whether NCU05936 is indeed allelic with *rrr-1* we made use of the *Neurospora* single-gene deletion collection. A strain where NCU05936 had been replaced with *hph* was available (FGSC19906) and was crossed to N2868 to obtain *rrr-1*::*hph* progeny that also carried *tdimerRed*, as a marker to measure the frequency of duplication inactivation (Fig. 4.3). As expected, in crosses wildtype for NCU05936 (NMF488 x NMF489) inactivation of *tdimerRed* occurs with high frequency (81±3%), while inactivation is lower in homozygous *rrr-1*::*hph* crosses (33±5%; NMF501 x NMF502; $p<0.05$).

The lower inactivation frequency in *rrr-1* deletion strains does not correlate with any overt defects in sexual development. The homozygous deletion crosses produce fertilized perithecia at the same time and abundance as wildtype crosses, they begin shooting ascospores at the same time, produce similar numbers of ascospores, and show similar rates of ascospore germination as wildtype crosses (general unquantified observations). In addition, the linear growth rate of wildtype (3.1±0.4 mm/hr) and *rrr-1*::*hph* (3.2±0.3 mm/hr) progeny does not differ. Southern analysis of genomic DNA from progeny with inactivated alleles of *tdimerRed* from a wildtype cross (NMF488 x NMF489) and an *rrr-1*::*hph* cross (NMF501 x NMF502) digested with *Sca*I or *Dpn*II and probed with a *tdimerRed* fragment indicated that in the majority of cases inactivation of *tdimerRed* is due to mutation by RIP rather than deletion, though products characteristic of both processes were observed (data not shown).

Characterization of RRR-1

Homologs of RRR-1 in other fungi are readily identifiable using cluster analysis and were found within the Pezizomycotina but none of the true yeasts (i.e. the Saccharomycotina and the Taphrinomycotina). They belong to a diverse group of large dynamin-like GTPases. Alignment of putative homologs with the RRR-1 sequence showed that both mutations occur in regions of the protein which are conserved evolutionarily (Fig. 4.2). Proteins with similarity

to RRR-1 have been characterized in other organisms including yeast and mammals. To clarify the relationship between RRR-1 and related large dynamin-like proteins that have been better characterized in yeast, such as VPS1 (Rothman, Raymond et al. 1990), MGM1 (Jones and Fangman 1992) and DNM1 (Otsuga, Keegan et al. 1998), RRR-1 protein sequence was used as bait in blastp searches to identify all other large dynamin-like proteins in the genome of *N. crassa* as well as dynamin-like proteins from species in other classes of fungi. This approach revealed that the yeasts *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Schizosaccharomyces japonicus* all encode only three proteins with this domain structure, i.e. homologs of VPS1, MGM1 and DNM1. The Pezizomycotina also have genes that encode homologs of these three proteins but in addition have one copy each of a gene belonging to a class of RRR-1 homologs that forms a discrete cluster, as well as an expanded group of between one (*Tuber melanosporum*) and seven (*Chaetomium globosum*) related dynamin-like proteins (Fig. 4.4). *N. crassa* has four genes in this cluster. Basidiomycetes (*Coprinopsis cinereus*, *Ustilago maydis*, *Malassezia globosa* and *Puccinia graminis* f. sp. *tritici*) and a Zygomycete (*Mucor circinelloides*) that was analyzed here generally have predicted protein products identifiable as the three highly conserved dynamin-like proteins, but RRR-1 homologs or the second novel and expanded dynamin-like proteins are present only in the *C. cinereus* and *P. graminis* f. sp. *tritici* (data not shown).

Localization of RRR-1

RRR-1 was tagged with GFP at the C-terminus of the predicted protein at the endogenous locus. Correct integration was verified by Southern analysis and Sanger sequencing and expression was verified by western blot analysis (Fig. 4.5). Three independently tagged RRR-1-GFP strains were analyzed for their RIP phenotype in heterozygous crosses to NMF488 or NMF489 (83±9%) and none showed significantly different RIP frequency of *tdimerRed* compared to wild type (90±4%; $p>0.05$), while both wild type and the GFP crosses were significantly different from heterozygous crosses with the *rrr-1* mutants (36±19%; $p<0.05$), suggesting that the GFP tag does not affect the function of the RRR-1 protein. We were able to detect the fusion protein by western blot (Fig. 4.5A) but have

so far not been able to detect GFP fluorescence by epifluorescence microscopy in any of the RRR-1-GFP strains we tested (Fig. 4.6B).

We thus focused on fractionation of the cellular components as a method to determine localization. RRR-1-GFP is enriched in the nuclear fraction compared to the two marker proteins H1-GFP and SON-1-GFP, which localize to chromatin or the nuclear pores and cytoplasm, respectively. When the nuclear membrane was separated from the nucleoplasm we found that RRR-1-GFP remains with the membrane, suggesting that RRR-1 is fairly stably associated with the nuclear membrane, at least in *N. crassa* (Fig. 4.6A).

Discussion

Characterization and mapping of *rrr* mutants

Five mutants that displayed semi-dominant or dominant reduction of tandem duplication inactivation were isolated after UV-mutagenesis and screening of ascospores for failure of inactivation of repeated *pan-2* genes (M. Freitag, AT Kobsa, S Honda, S Friedman and EU Selker, in preparation). We report here on two of the originally isolated mutants, N1516 and N1522. Southern analysis of random progeny revealed that during crosses the *pan-2* duplications were subject to point mutations indicative of RIP, as well as deletion indicative of a process akin to premeiotic recombination via single-strand annealing. Bulk-segregant analysis followed by whole genome sequencing of mutant pools mapped the UV induced mutations to the same gene in both mutants, NCU05936, which we have named *rrr-1* (reduced RIP and recombination 1). We verified that mutation of NCU05936 causes the *rrr-1* phenotype by crossing duplications into strains in which NCU05936 had been replaced by *hph*, and by observing decreased inactivation of the duplication by RIP during subsequent homozygous NCU05936::*hph* crosses. Perhaps unexpectedly for deletion alleles, heterozygous crosses with strains harboring NCU05936::*hph* alleles behave semi-dominantly, just like the original *rrr-1* point mutants. This suggests loss of function in even a single parent involved in a cross can cause the dominant phenotype, which may help to explain why this gene was found twice in a screen that uncovered only five semi-dominant or dominant *rrr* mutants.

Function of RRR-1

Domain analysis of the predicted RRR-1 protein revealed that it has a dynamin GTPase domain, a dynamin central domain and a C-terminal GTPase effector domain in the order typically associated with dynamin-like proteins (Praefcke and McMahon 2004). We searched the genomes of ascomycetes to identify the full repertoire of RRR-1 homologs and found a minimum of three highly conserved dynamin-like proteins in every fungal genome analyzed, including those of the budding and fission yeasts. One species, *C. globosum*, has eleven such genes. We constructed a phylogenetic tree using all dynamin-like proteins found to gain insight into the relationships of RRR-1 and dynamin-like proteins described in other organisms. We found that *N. crassa* has orthologues of the three highly conserved dynamin-like proteins represented by DNM1, which is important for mitochondrial and peroxisome fission (Bleazard, McCaffery et al. 1999; Kuravi, Nagotu et al. 2006), MGM1, which acts in mitochondria fusion (Jones and Fangman 1992; Wong, Wagner et al. 2000), and VPS1, which acts during vesicle fission in the trans-Golgi network, vacuole fission and during endocytosis (Rothman, Raymond et al. 1990; Nothwehr, Conibear et al. 1995; Vizeacoumar, Vreden et al. 2006; Smaczynska-de, Allwood et al. 2010). Homokaryotic deletion of any of these three proteins is lethal in *N. crassa*, as no clean replacement strains were obtained after crossing strains with putative deletion alleles that were constructed by the *Neurospora* Functional Genomics Project (data not shown). However, *N. crassa* also has five additional dynamin-like proteins, which apparently are not essential, as homokaryotic strains with deletion alleles can be obtained after crosses. These five proteins fall into two novel classes of large GTPase-like proteins, the RRR-1 cluster and a cluster of four dynamin-like proteins of unknown function.

We have presented evidence to show that *rrr-1* affects the frequency of RIP and premeiotic recombination, which both occur in nuclei with duplicated regions during sexual development, and that *rrr-1* is associated with the nuclear membrane during vegetative growth. Since all other described dynamin-like proteins are involved in membrane fission or fusion and all share important structural features, we hypothesize that *rrr-1* is involved in nuclear membrane dynamics and affects RIP and recombination frequency by altering either the number of times they can occur during the successive mitoses that precede meiosis, or

by altering the period during which these processes can occur during each individual nuclear cycle that precedes meiosis. Simultaneous alteration of RIP and recombination frequency has also been observed in *Podospora anserina ami1* mutants, which display altered nuclear distribution (Bouhouche, Zickler et al. 2004) and, interestingly, delayed sexual development, which resulted in increased RIP and recombination frequency (Bouhouche, Zickler et al. 2004). This is consistent with RIP and recombination occurring multiple times prior to karyogamy and meiosis as observed in *N. crassa* (Singer, Kuzminova et al. 1995; Watters, Randall et al. 1999). If *rrr-1* and *ami1* do not directly affect the mechanisms of RIP and recombination then it stands to reason that at least three processes are spatio-temporally linked, namely membrane fission or fusion, RIP and premeiotic recombination.

Many ascomycete fungi undergo some form of closed mitosis where the nuclear membrane does not completely break down and then reforms, as is typical for some animals. These processes are well-studied in *S. cerevisiae* (Copeland and Snyder 1993), *N. crassa* (Roca, Kuo et al. 2009) and *Aspergillus nidulans* (De Souza, Osmani et al. 2004). A direct role for dynamin-like GTPases that aid in karyokinesis, i.e. division of the nucleus and apportioning of the nuclear membrane to daughter nuclei, has not been considered previously. Dynamins are important for the division of many organelles, as described above for the highly conserved ascomycete dynamin-like proteins, but they are also important for cytokinesis in plants and animals (Wienke, Knetsch et al. 1999; Schafer, Weed et al. 2002; Thompson, Skop et al. 2002; Collings, Gebbie et al. 2008; Mooren, Kotova et al. 2009). In the future, it will be of great interest to explore the precise localization and orientation of RRR-1 in the nuclear membrane to understand the function of non-essential dynamin-like proteins in fungi. We propose here that RRR-1 plays an important role in fission of the nucleus during mitosis, either alone or in concert with one or more of the essential dynamins. The latter possibility seems perhaps more likely, as *rrr-1* is not lethal as one would expect for a protein that is key for an essential cellular function such as mitosis.

In conclusion, we have identified a large dynamin-like protein, RRR-1 that is required for normal frequencies of both RIP and premeiotic recombination in *N. crassa*. We hypothesize that RRR-1 affects RIP and recombination indirectly, likely by altering nuclear

membrane dynamics or karyokinesis during sexual development and thereby altering temporal aspects of RIP and recombination.

Acknowledgements

We thank Mark Dasenko, Matthew Peterson and Chris Sullivan at the OSU CGRB core facility for assistance with Illumina sequencing. We also thank the *Neurospora* Functional Genomics Project (NIH P01 GM068087) for primers and deletion strains. This work was made possible by start-up funds from the OSU Computational and Genome Biology Initiative and grants from the American Cancer Society (RSG-08-030-01-CCG), DOE (DE-FG02-08ER64665) and NIH (P01 GM068087 and R01GM097637) to MF and the NIH (GM025690-22) to EUS.

Table 4.1. *N. crassa* strains used.

Strain #	Genotype	Reference
N1064	*matA*; *pan-2^{B3}*	FGSC2248
N1477	*matA*; *pan-2^{B3}*; *pan-2^{+}-hph$^{+multi-copy}$*	(Rosa, Folco et al. 2004)
N1478	*mata*; *pan-2^{B3}*; *pan-2^{+}-hph$^{+multi-copy}$*	(Rosa, Folco et al. 2004)
N1514	*matA*; *pan-2^{B3}*; *pan-2^{+}-hph$^{+multi-copy}$*; *rrr-1^{L496P}*	this study
N1516	*mata*; *pan-2^{B3}*; *pan-2^{+}-hph$^{+multi-copy}$*; *rrr-1^{L496P}*	this study
N1522	*mata*; *pan-2^{B3}*; *pan-2^{+}-hph$^{+multi-copy}$*; *rrr-1^{G325N}*	this study
N2868	*matA*; *his-3^{+}*::*Pccg-1-tdimerRed^{+}*	this study
N2928	*matA*; *his-3*; *mus-51*::*bar^{+}*	this study
N2944	*matA*; *his-3^{+}*::*Pccg-1-tdimerRed^{+}*; *rid*	this study
N2945	*mata*; *his-3^{+}*::*Pccg-1-tdimerRed^{+}*; *rid*	this study
N3011	*mata*; *his-3*; *mus-51*::*bar^{+}*	FGSC9538
NMF37	*matA*; Mauriceville-1-c	FGSC2225
NMF62	*matA*; *fl*	FGSC4317
NMF63	*mata*; *fl*	FGSC4347
NMF192	*matA*; *his-3^{+}*::*Pccg-1-hH1^{+}-sgfp^{+}*	(Freitag, Hickey et al. 2004)
NMF488	*matA*; *his-3^{+}*::*Pccg-1-tdimerRed^{+}*	this study
NMF489	*mata*; *his-3^{+}*::*Pccg-1-tdimerRed^{+}*	this study
NMF501	*matA*; *his-3^{+}*::*Pccg-1-tdimerRed^{+}*; *rrr-1*(NCU05936)::*hph^{+}*	this study
NMF502	*mata*; *his-3^{+}*::*Pccg-1-tdimerRed^{+}*; *rrr-1*(NCU05936)::*hph^{+}*	this study
NMF528	*mata*; *rrr-1*::*hph^{+}*	FGSC19906
NMF530	*matA*; *his-3*; *mus-51*::*bar^{+}*; *rrr-1-gfp*::*hph^{+}*	this study
NMF531	*mata*; *his-3*; *mus-51*::*bar^{+}*; *rrr-1-gfp*::*hph^{+}*	this study
NMF532	*matA*; *his-3*; *mus-51*::*bar^{+}*; *rrr-1-gfp*::*hph^{+}*	this study
NMF557	*his-3^{+}*::*Pccg-1-son-1^{+}-sgfp^{+}*	this study

Table 4.2. Primers and adapters used.

OMF	Sequence (5' to 3' direction)
83	GGCGGAGGCGGCGGAGGCGGAGGCGGAGG
84	CGAGCTCGGATCCATAACTTCGTATAGCA
1053	AAAAAGCCTGAACTCACCGCGACG
1054	TCGCCTCGCTCCAGTCAATGACC
2565	AAGAGAAGGAGGCACACGG
2566	ACGGGTCTTGACGTTGGTC
2721	TGCTATACGAAGTTATGGATCCGAGCTCGTGAGAGGGAAGTCTAGAAAGT
2722	GCATACACACTCCGCTCTCA
2723	TCGATTTCTTCCCCACTCTG
2724	CCTCCGCCTCCGCCTCCGCCGCCTCCGCCAGCCTCATAGGTCCCAATAG

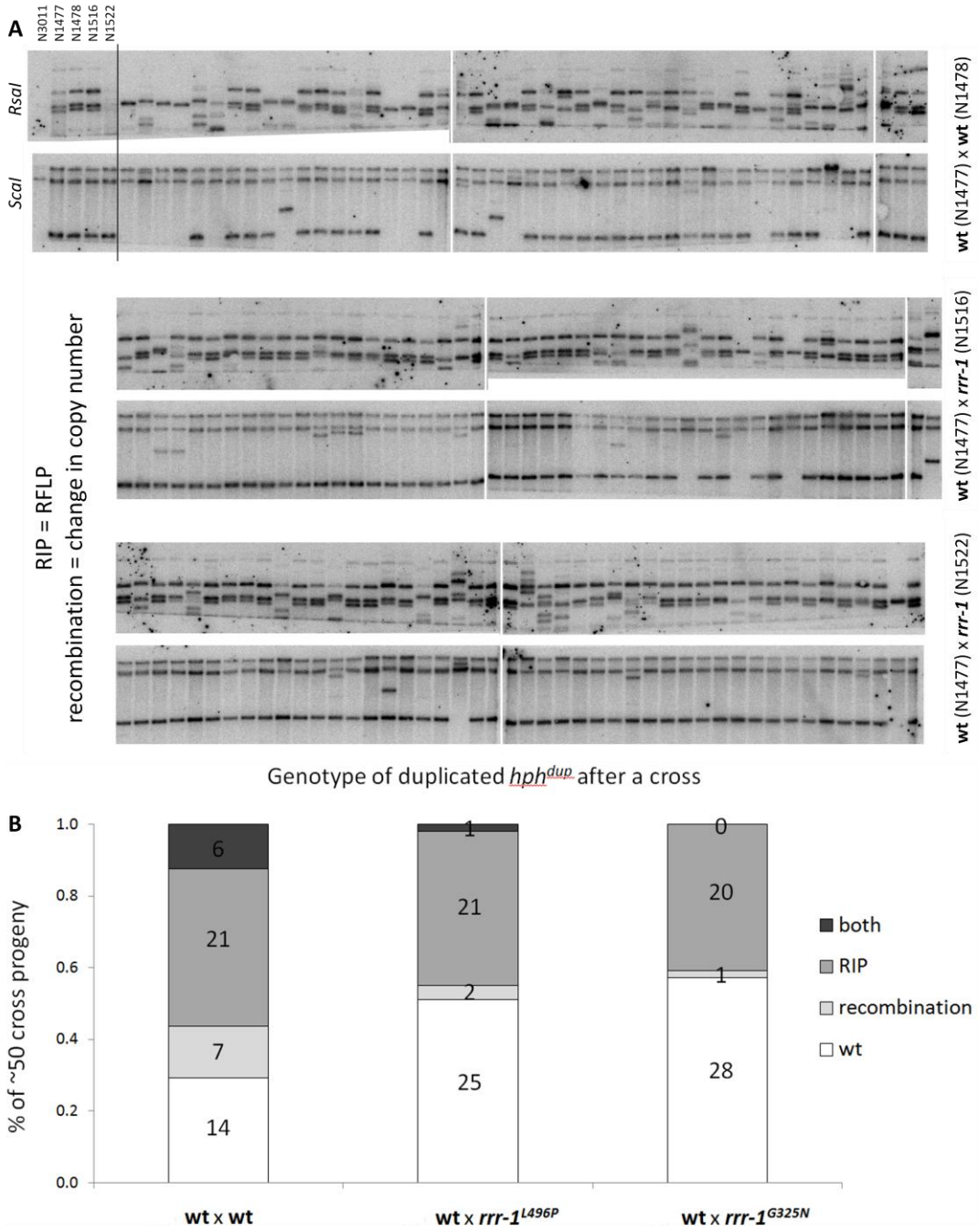


Figure 4.1. Genotype of duplicated *hph* genes in random progeny from wild-type and *rrr*-1 crosses. (A) Southern blot probed with radiolabelled *hph* of genomic DNA from random cross progeny that was digested with *Sca*I or *Rsa*I. Recombination by deletion is indicated by lower copy number of *hph*, indicated primarily by loss of the lower band in the *Sca*I digest, while RIP is indicated by RFLP primarily in the *Rsa*I digest. (B) RIP and recombination frequency is reduced in crosses involving either *rrr-1* allele.

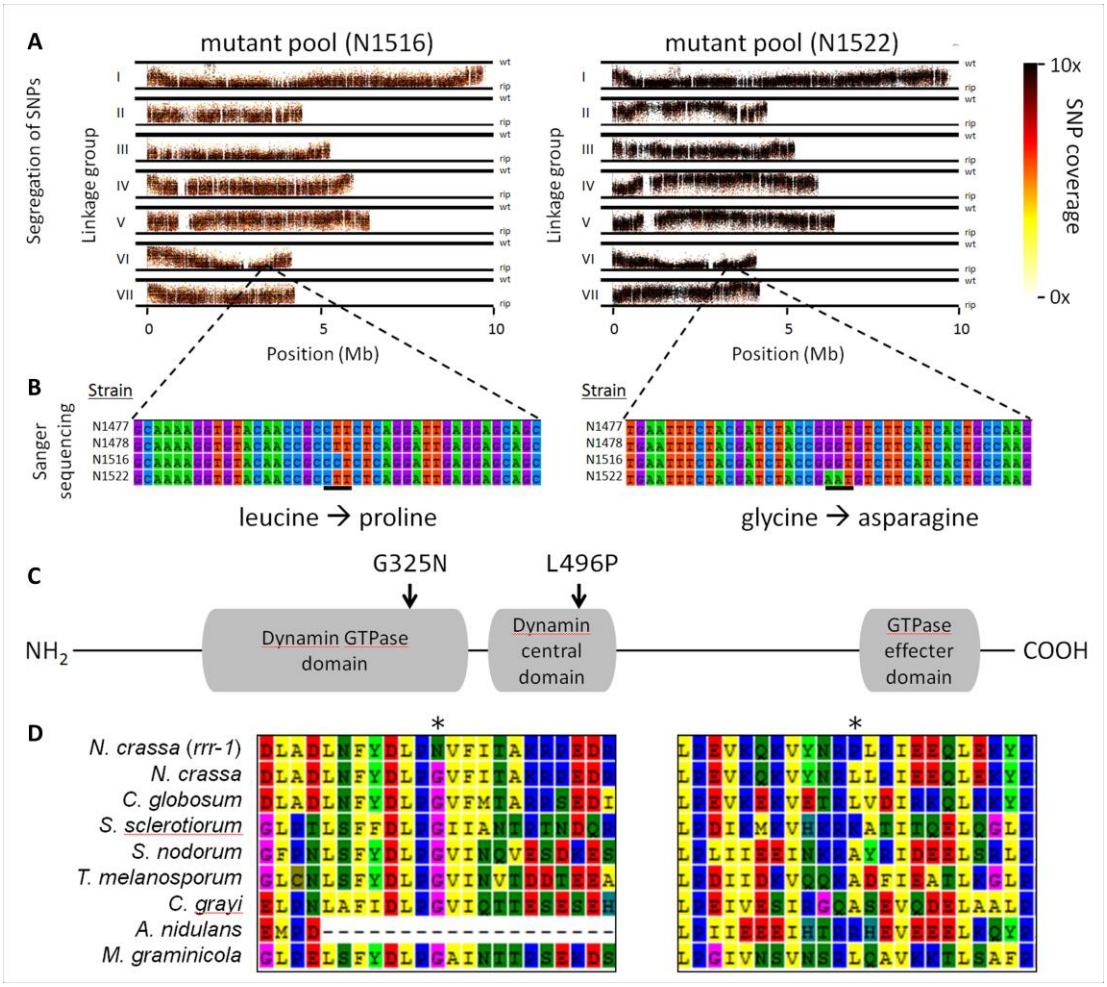


Figure 4.2. Bulk segregant analysis maps "Reduced RIP and Recombination" (*rrr*) to NCU05936. (A) Segregation of 168,579 SNPs along the *N. crassa* chromosomes in bulk *rrr* progeny. (B) Sanger sequencing of independent *rrr* alleles confirms that N1516 and N1522 have different alleles of a single gene, NCU05936 (now called *rrr-1*). (C) Relative location of the *rrr-1* mutations in domains of NCU05936 identified by InterProScan and (D) their conservation among other ascomycetes, where asterisks indicate the mutations.

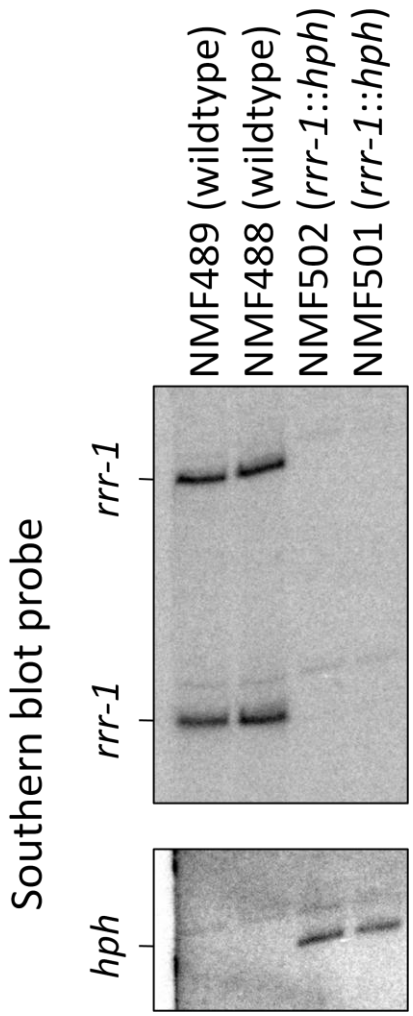


Figure 4.3. Southern analysis confirms *rrr-1*::*hph*; *tdimerRed* strains. Southern blot probed with radiolabelled *hph* or *rrr-1* of genomic DNA from *tdimerRed* strains that was digested with *Cla*I. Correct integration is confirmed by loss of *rrr-1* in the *rrr-1*::*hph* strains.

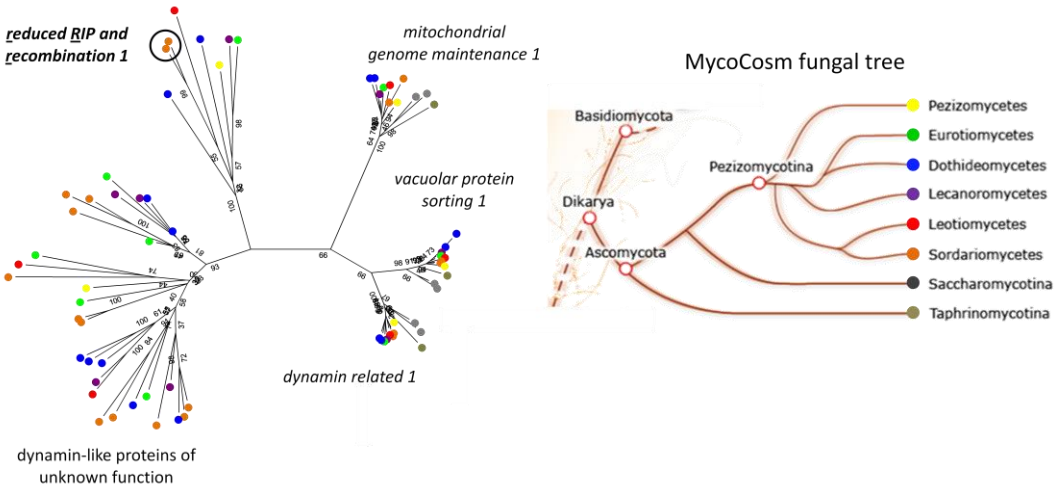


Figure 4.4. Dynamin-like long GTPase families in ascomycetes. RRR-1 homologues belong to a novel, fourth family of dynamin-like long GTPases in fungi (circled for *N. crassa*), which does not have orthologs in budding or fission yeast. Bootstrap consensus tree based on protein models downloaded from MycoCosm (Grigoriev IV et al., 2012) was inferred from 200 replicates generated by the maximum likelihood method. Analyses were conducted in MEGA5 using 866 amino acid positions. Colored circles indicate the class or subphylum.

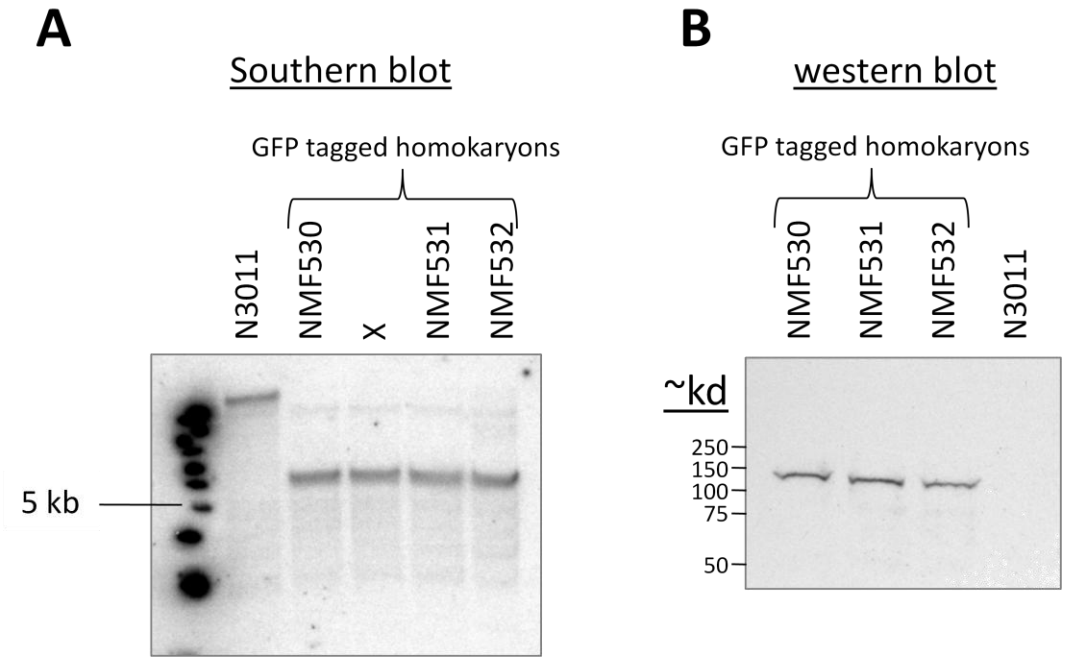


Figure 4.5. Tagging RRR-1 at the endogenous locus with GFP. (A) Southern analysis of *rrr-1-gfp* in backcrossed homokaryotic strains. Integration of *rrr-1-gfp* at the endogenous locus is expected to show up as a band 5.2 kb in length while wildtype *rrr-1* is expected to be > 8 kb when probed with *rrr-1*. (B) Western blot analysis of RRR-1-GFP in backcrossed homokaryotic strains. RRR-1-GFP is expected to be ~138 kd.

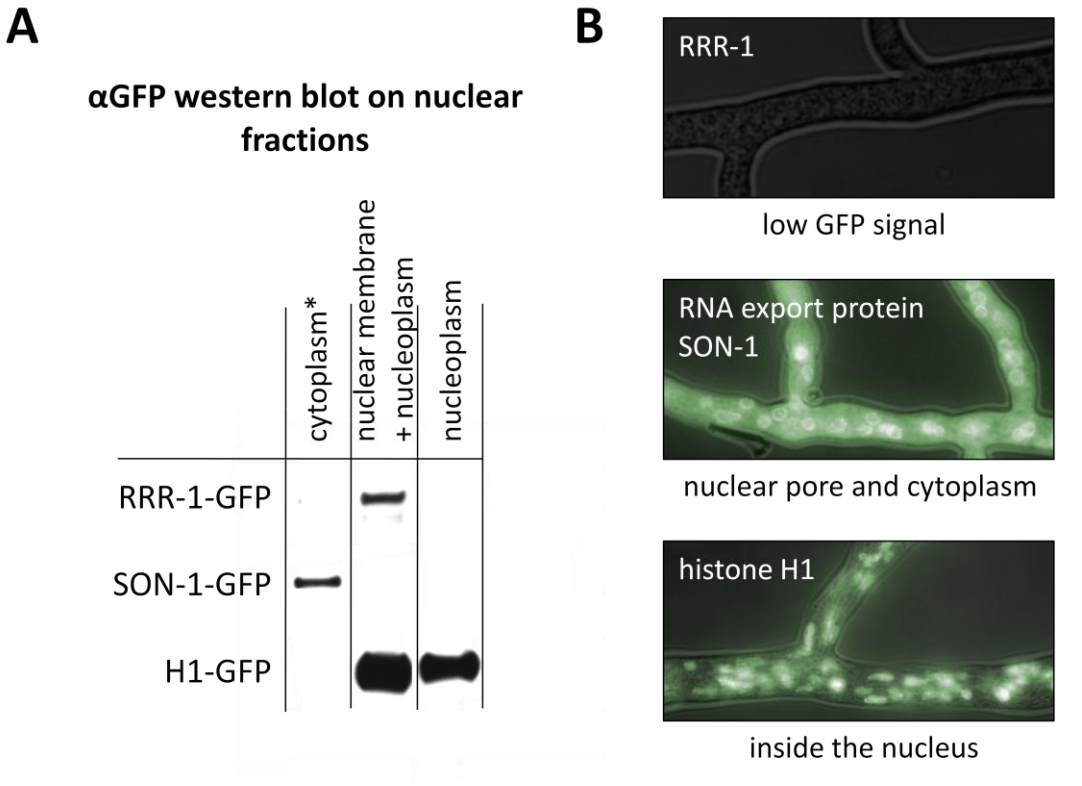


Figure 4.6. RRR-1 localizes to the nuclear membrane. (A) Western blot with anti- GFP antibody on nuclear extracts of three different GFP-tagged strains shows that RRR-1-GFP is enriched in the nuclear membrane fraction. SON-1-GFP is a nuclear pore protein (Roca, Kuo et al. 2009) and H1-GFP is a nuclear protein (Folco, Freitag et al. 2003; Freitag, Hickey et al. 2004). *The "cytoplasm" fraction has been cleared of organelles. (B) Merged fluorescent and phase contrast images of live hyphae from the three strains in (A). Images are shown at identical intensities to show that no RRR-1 signal has been obtained by epifluorescence microscopy.

Effect of deletion of genes important for homologous recombination on RIP frequency

Kyle R. Pomraning and Michael Freitag

Manuscript in preparation for *PLoS ONE*

Abstract

Repeat-induced point mutation (RIP) and premeiotic homologous recombination mutate duplications in many filamentous fungi. RIP is characterized by C:G to T:A point mutations in genetically linked and unlinked repetitive DNA, while loss of DNA segments between linked direct duplications is dependent on homologous recombination (Selker 1990; Irelan, Hagemann et al. 1994). Both processes efficiently mutate DNA during the sexual cycle and contribute to streamlining by inactivating and removing mobile elements (Galagan and Selker 2004). We employ genetic approaches to elucidate the mechanism of premeiotic recombination and RIP. Here we tested the idea that these processes are mechanistically linked by screening the *Neurospora* single gene deletion collection for recombination mutants that show RIP defects.

Introduction

Repetitive DNA is susceptible to mutation during the sexual cycle of some filamentous fungi. In *Neurospora crassa* duplications of at least 400bp are frequently mutated by repeat-induced point mutation (RIP) which causes C:G to T:A point mutations regardless of whether the mutations are genetically linked (Selker, Cambareri et al. 1987; Cambareri, Jensen et al. 1989; Fincham, Connerton et al. 1989; Irelan, Hagemann et al. 1994; Watters, Randall et al. 1999). Premeiotic homologous recombination mutates DNA during the same portion of the lifecycle by causing deletions in direct, closely linked repeats (Irelan, Hagemann et al. 1994). Both processes occur after fertilization but prior to karyogamy and meiosis (Selker 1990) when the nuclei of the ascogenous hyphae are estimated to undergo seven to ten rounds of nuclear division (Emerson 1966).

The mechanism for duplication detection during RIP is unknown but it is hypothesized to be dependent on DNA pairing (Faugeron, Rhounim et al. 1990; Selker 1990). However, *mei-2*, a mutant defective in pairing of homologous chromosomes during meiosis has normal levels of RIP (Foss and Selker 1991). A mutant which alters homologous recombination frequency, *rec-2*, also does not affect RIP frequency (Bowring and Catcheside 1993) suggesting homologous recombination during meiosis and RIP do not share a common mechanism. Careful study of RIP and deletion by homologous recombination frequency

between tandem duplications in either direct or inverted orientation and unlinked duplications determined the duplications are always susceptible to mutation by RIP but only to deletion when they are direct and in tandem (Irelan, Hagemann et al. 1994). This type of frequent deletion by homologous recombination during mitosis can be explained if the pathway responsible is similar to single-strand annealing described first in budding yeast (Sugawara and Haber 1992; Lyndaker and Alani 2009). Repair of DNA breaks by single-strand annealing only causes deletions in direct, closely linked repeats because they are repaired by 5'->3' filling in using a homologous template. If repair proceeds while the repeats are still mis-paired, filling in will proceed across the gap left by the 3' copy of the repeat, effectively deleting it.

The mechanism for deletion of tandem repeats, which is dependent on aberrant pairing of repetitive DNA, has been studied thoroughly in yeast in the context of double strand breaks (Sugawara and Haber 1992; Lyndaker and Alani 2009) and was found to be dependent on genes which have orthologues in *N. crassa*. Briefly, double strand breaks between duplicated segments of DNA are bound by RPA (Alani, Thresher et al. 1992) and homology between the duplicated segments detected by a process dependant on Rad52 (Shinohara, Shinohara et al. 1998; Sugiyama, New et al. 1998). Aberrant pairing between duplicated sequences results in ssDNA overhangs which have to be removed for repair to proceed. Msh2 and Msh3 bind the junction between the paired DNA and the free ssDNA tail (Sugawara, Paques et al. 1997; Surtees and Alani 2006), which is then cleaved off by the Rad1/Rad10 endonuclease (Fishman-Lobell and Haber 1992; Tomkinson, Bardwell et al. 1993).

So far, neither reverse nor forward genetic screens have shed light on the DNA-pairing mechanism that may be required for RIP. Mutation of a gene, *rid*, that encodes a cytosine methyltransferase-like protein has been shown to abolish RIP in homozygous crosses (Freitag, Williams et al. 2002), presumably because this protein is somehow involved in the actual mutagenesis. Here, we have tested the hypothesis that RIP and deletion by homologous recombination are mechanistically linked by screening the *Neurospora* single gene deletion collection for mutants that show RIP defects, starting with deletion mutants that are known or expected to affect recombination pathways.

Materials and methods

Strains and growth conditions

Neurospora crassa strains were grown at 32 C in Vogel's minimal medium (VMM) with 1.5% sucrose (Davis 2000). Crosses were performed on modified synthetic crossing medium (SC) (Russo, Sommer et al. 1985) with 0.4% sucrose. Ascospores were recovered in water for the scoring of RIP after four weeks and after two weeks for all other crosses. Germination of ascospores was induced by heat shock at 65 C for 1 hour on VMM with FGS (0.5 g/L fructose, 0.5 g/L glucose, 20 g/L sorbose) as the carbon source and supplemented with 10 mg/L pantothenate when needed. Strains were tested for mating type by crossing with *fl* strains FGSC4317 and FGSC4347 (Davis 2000). Deletion strains from the *N. crassa* single-gene deletion collection (see Table 5.1) were crossed to N2868, NMF489 or N2945 to obtain deletion strains of both mating type with the *tdimerRed* marker. Progeny were collected and assayed for mating type, hygromycin resistance (Hyg^R) and *tdimerRed* (RFP^+) phenotypically. Assay for *rid* was performed by RFLP on genomic DNA from progeny. After crosses, *tdimerRed* was assumed to be at most lightly mutated by RIP if red fluorescence appeared to be as bright as in progenitor strains.

Southern analysis

Replacement of recombination genes by *hph* was confirmed by Southern analyses. Genomic DNA from *tdimerRed* strains was extracted as previously described (Pomraning, Smith et al. 2009), digested with *Hind*III and transferred to a Biodyne nylon membrane (Pall Life Sciences, Port Washington, NY). Approximately 1 kb DNA fragments expected to be just 5' of the *hph* replacement for each recombination gene to be analyzed were constructed by PCR using primers in Table 5.2. Each fragment was radiolabelled and used to probe recombination mutants derived from the *N. crassa* single-gene deletion collection (Fig. 5.1). Correct integration was confirmed by a shift in band size compared with a strain wildtype for all the recombination genes (19907-1).

Assay for RIP frequency

Strains where genes likely involved in recombination pathways in *N. crassa* had been replaced by *hph* as part of the *Neurospora* Functional Genomics Project were crossed to RIP tester strains that carry *tdimerRed* (a short duplication of the *rfp* gene; (Campbell, Tour et al. 2002; Freitag and Selker 2005) that is frequently (~80%) inactivated by RIP mutations during crosses of *N. crassa* (see Chapter 4) or RIP tester strains that carry *pan-2* duplications that are inactivated by RIP mutations at similar frequencies (Rosa, Folco et al. 2004). Progeny were collected four weeks after fertilization of each cross in H_2O. To quantify inactivation of *tdimerRed*, 100 germinated progeny were analyzed by microscopy for expression of RFP ~18 hours after heat shock. In crosses where both parents were *tdimerRed*$^+$, RIP frequency was calculated as:

$$\text{RIP } \% = 1 - \frac{(\text{RFP}^+\text{progeny})}{(\text{total progeny})}$$

since all the progeny are expected to have the *tdimerRed* gene. In crosses where only one parent was *tdimerRed*$^+$, RIP frequency was calculated as:

$$\text{RIP } \% = 1 - \frac{2(\text{RFP}^+\text{progeny})}{(\text{total progeny})}$$

since half the progeny are expected to have the *tdimerRed* gene. To quantify inactivation of *pan-2*, 100 ascospores were analyzed for black (*pan-2*$^+$) or tan (*pan-2*$^-$) color microscopically and RIP frequency calculated as for heterozygous crosses with *tdimerRed* strains. Mean and standard deviation values of replicates were computed in Excel. Unpaired t-tests were also performed using Excel.

Results

Strain construction

A duplicated marker gene is needed to test for RIP frequency. To that end we crossed strains of *N. crassa* where recombination genes had been replaced with *hph* to strains harboring *tdimerRed*, which is a short duplication of the *rfp* gene (Campbell, Tour et al. 2002; Freitag and Selker 2005), to obtain red fluorescent strains lacking particular genes important for recombination. These strains were then directly assayed for RIP frequency in

subsequent crosses by quantifying inactivation of *tdimerRed*, which is inactivated by RIP at high frequency (~80%; see Chapter 4).

Southern analyses were performed to ensure correct replacement of all recombination genes tested by *hph* in the strains obtained from FGSC. Correct integration was confirmed in all the mutants analyzed (data not shown), except for strain NMF494, which is heterokaryotic, harboring *sae2*$^{+}$ as well as *sae2*::*hph*$^{+}$ nuclei (Fig. 5.1).

Two significant problems with constructing strains in this manner were encountered. First, replacement of many genes important for homologous recombination results in either lethality during vegetative growth or sterility of heterozygous or homozygous deletion crosses in *N. crassa*. For example, we were unable to test deletions for the homolog of replication protein A (RPA; NCU03606.5), which is important for binding ssDNA during homologous recombination (Alani, Thresher et al. 1992; Shinohara, Shinohara et al. 1998; Sugiyama, New et al. 1998), for RIP defects as *rpa*::*hph* appears to be lethal as a homokaryon in *N. crassa* (Nolan, Cecere et al. 2008). Similarly, construction of *rad51*::*hph* and *rad52*::*hph* strains with *tdimerRed*$^{+}$ was difficult because these mutants (NCU02741.5 and NCU04275.5) are nearly sterile in even a heterozygous deletion cross and are defective in homologous recombination and meiosis (Game, Zamb et al. 1980; Prakash, Prakash et al. 1980; Shinohara, Ogawa et al. 1992). Our *tdimerRed* marker and *rad52*::*hph* are located on different chromosomes in *N. crassa* so even with defects in homologous recombination we were able to isolate homokaryotic *rad52*::*hph*; *tdimerRed*$^{+}$ strains. However, *rad51*::*hph* and *tdimerRed* are both located on Linkage Group I in *N. crassa* and we have been so far unable to isolate any strains where a crossover had occurred between the two markers from the few viable progeny obtained after crossing FGSC12433 (*rad51*::*hph*$^{+}$) and N2868 (*his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$). We did not obtain enough viable progeny to determine whether the lack of recombinants was due solely to linkage or whether *rad51*::*hph*$^{+}$ caused an additive effect.

The second problem encountered during strain construction was caused by crossing the duplicated marker gene, *tdimerRed*, with mutant strains to obtain recombinants. Any duplication is susceptible to mutation by RIP as well as partial deletion by premeiotic homologous recombination at high frequency in *N. crassa* (Irelan, Hagemann et al. 1994). To ensure that only few RIP mutations were accumulated we selected only *tdimerRed*$^{+}$ strains

that fluoresced as brightly as strains with unmutated alleles of *tdimerRed*, thus making accumulation of many mutations unlikely. However, one strain used in this study which fluoresced brightly, NMF496, was found to have lost one copy of *rfp* from *tdimerRed* and, as expected, did not undergo RIP in subsequent crosses.

RIP frequency in homologous recombination mutants

Deletion of duplications by recombination and RIP frequency of tandem repeats is affected by *rrr-1*, (Chapter 4) suggesting the two processes may be mechanistically linked. To test this hypothesis we performed a reverse genetics screen of genes known to be involved in deletion of direct tandem repeats by homologous recombination along the homologous sequence (i.e., *rpa*, *rad1*, *rad10*, *msh2* and *rad52*; NCU03606.5, NCU07440.5, NCU07066.5, NCU02230.5, NCU04275.5), described in yeast as single-strand annealing (Sugawara and Haber 1992; Lyndaker and Alani 2009), as well as additional genes important for homologous recombination pathways, including *rad51* (Cheng, Baker et al. 1993), *mre11* (Watanabe, Sakuraba et al. 1997) and *sae2* (Lengsfeld, Rattray et al. 2007).

Three types of crosses were attempted to analyze RIP frequency including (1) heterozygous mutant crosses between a wildtype *tdimerRed*$^{+}$ strain and a recombination mutant *tdimerRed*$^{+}$ strain (Table 5.3), (2) nucleus specific heterozygous mutant crosses between a wildtype *pan-2*$^{+}$*-hph*$^{+multi\text{-}copy}$ strain and a mutant *tdimerRed*$^{+}$ strain (Table 5.4) and (3) homozygous mutant crosses between a recombination mutant *tdimerRed*$^{+}$ strain and a recombination mutant *tdimerRed*$^{+}$ strain (Table 5.5).

Analyzable homozygous deletion crosses for genes involved in single-strand annealing were all competent for RIP (Table 5.5), as inactivation of *tdimerRed* occurred at 78±16% in *rad1* (calculated as twice the observed RIP frequency since only one strain harbored a RIP mutatable *tdimerRed*$^{+}$ gene), 85±5% in *rad10* and 64±4% in *msh2*, compared to 82±2% in wildtype crosses. We were able to obtain a recombinant *rad52*; *tdimerRed*$^{+}$ strain but heterozygous crosses with this deletion strain were nearly sterile (i.e. crosses produced few perithecia which contained ~10-20 viable ascospores each) as reported previously for *rad52* (Kafer and Perlmutter 1980). Homozygous *rad52* crosses were completely sterile, making accurate analysis of RIP frequency impossible. From the few

ascospores collected from heterozygous *rad52* crosses we observed decreased RIP frequency, 50±9%, compared to 82±2% in wildtype crosses (Table 5.4) but as the observed RIP frequency progressively increases during healthy crosses (Singer, Kuzminova et al. 1995) and is increased by mutations that delay sexual development (Bouhouche, Zickler et al. 2004) it cannot be ruled out that the change in frequency is pleiotropic.

Similar to *rad52*, heterozygous deletion crosses with *rad51* (the *recA* homolog in *N. crassa*) were nearly sterile and we were not able to obtain any *rad51*::*hph*; *tdimerRed*, strains. We attempted heterozygous *rad51* crosses (similar to those in Table 5.4 but without *tdimerRed*$^{+}$ in the mutant strain) between N1478 and FGSC12434 and observed RIP frequency of *pan-2*$^{multi\text{-}copy}$ of 14% and 165% in two replicates. These crosses were not healthy, as observed previously when crossing FGSC12433 and N2868. Two replicates of similar wildtype crosses performed in parallel between N1478 and NMF39 gave more reproducible values for RIP of 93% and 100%.

Analysis of RIP frequency in homozygous *mre11* deletion crosses also proved impossible, as they were sterile. Interestingly, RIP frequency in heterozygous *mre11* deletion crosses, which did not display overt sexual defects like *rad51* and *rad52*, was significantly reduced (37±16%) and (63±1%) compared to wildtype (69±6%; $p<0.05$) and (82±2%; $p<0.05$) respectively for both types of heterozygous crosses analyzed. Results from all crosses are detailed in Tables 5.3, 5.4 and 5.5. One of the *sae2* strains used, NMF494, was heterokaryotic, however, all *sae2* heterozygous deletion crosses were RIP competent though somewhat decreased in fertility. To obtain RIP frequencies for *sae2* homozygous deletion crosses we will have to generate new tester strains.

Discussion

Link between RIP and recombination

Concurrent high levels of point mutation by RIP and deletion by homologous recombination of direct tandem DNA repeats has been reported previously (Selker 1990; Irelan, Hagemann et al. 1994) (and Chapter 4) but convincing mechanistic links between the two processes have not been found. Effects of mutation of a few selected components of the homologous recombination and meiotic DNA pairing machineries on RIP in *N. crassa* have

been studied before, but no clear effects have been observed in the case of *rec-2*, which modulates the frequency of recombination during meiosis (Bowring and Catcheside 1993), or *mei-2*, a mutant defective in pairing of homologous chromosomes (Foss and Selker 1991). Careful construction of both direct or inverted tandem duplications, and unlinked duplications revealed that while RIP affects all three types of duplications, deletions occur only when the duplications are direct and in tandem (Irelan, Hagemann et al. 1994). This is most parsimoniously explained if deletions occur by a process akin to single-strand annealing in yeast (Sugawara and Haber 1992; Lyndaker and Alani 2009) as repair of this type can result in deletion only in direct tandem repeats but not inverted or unlinked repeats.

To directly test whether deletion of genes necessary for single-strand annealing affects RIP frequency, we utilized the *Neurospora* single gene replacement collection (Dunlap, Borkovich et al. 2007). We were only able to successfully test genes important for later steps in single-strand annealing, which recognize and remove excess single-stranded DNA, including *msh2*, *rad1* and *rad10* (Lyndaker and Alani 2009). We attempted tests with deletion alleles of genes involved in recognition of repetitive and mis-paired DNA during the early stages of single-strand annealing, such as *rpa* and *rad52.* These genes, however, are important for many cellular processes, including meiosis, and were untestable for RIP as most crosses were sterile or nearly sterile. If there is a mechanistic link between RIP and recombination it likely occurs as a common mis-paired DNA substrate. However we do not know whether recognition of aberrantly paired or broken DNA by *rpa* and *rad52* is an early step in RIP as in single-strand annealing (Lyndaker and Alani 2009). To test this, we will have to construct semi-functional or conditional alleles of some of these genes, an approach that is beyond the scope of this work.

We also successfully tested two genes involved in other recombination processes for RIP defects. Mutation of *sae2* does not result in RIP defects in heterozygous deletion crosses. Sae2 is predicted to be involved in processing of DNA hairpin structures during homologous recombination (Lengsfeld, Rattray et al. 2007). Strains with deletion alleles of *mre11* were sterile in homozygous deletion crosses in *N. crassa* but we found that RIP frequency in heterozygous *mre11* crosses is reduced compared to wildtype, suggesting that Mre11, which

has roles in homologous recombination as well as non-homologous end joining (Watanabe, Sakuraba et al. 1997) may be important for RIP.

In conclusion, we have tested a large number of recombination genes for RIP defects but have not found a mechanistic link between RIP and the later steps of ssDNA processing during homologous recombination by single-strand annealing. We excluded some genes from an essential role in RIP, however, including *rad1*, *rad10*, and *msh2*, which did not show any defects in RIP in our tests. This excludes removal of ssDNA after mis-pairing and DNA damage as a pathway for RIP. This suggests the DNA is mutated directly rather than being removed prior to mutation during repair. Recognition of aberrant DNA structure is likely to be a common theme for RIP and single-strand annealing but genes important for this step were untestable as deletion alleles exhibit poor fertility. However, we did see slightly lower RIP frequency in heterozygous crosses with *mre11*::*hph*$^{+}$ Mre11 is part of the MRN complex (MRX in *S. cerevisiae*) and associates with Rad50 and Nbs1 (encoded by genes NCU00901.5 and NCU06049.5 in *N. crassa*) in the context of DNA repair across a wide phylogenetic range (Farah, Cromie et al. 2005; Waterworth, Altun et al. 2007; Amiard, Charbonnel et al. 2010; Rupnik, Lowndes et al. 2010). To confirm or deny whether the MRN complex is involved in RIP frequency we will test deletion alleles of *rad50* and *nbs1* for defects in RIP.

Acknowledgements

We thank the *Neurospora* Functional Genomics Project (NIH P01 GM068087) for primers and deletion strains. This work was made possible by start-up funds from the OSU Computational and Genome Biology Initiative and grants from the American Cancer Society (RSG-08-030-01-CCG), DOE (DE-FG02-08ER64665) and NIH (P01 GM068087 and R01GM097637).

Table 5.1. *N. crassa* strains used.

Strain #	Genotype	Reference
FGSC12165	*matA*; *rad10*::*hph*$^{+}$	(Dunlap, Borkovich et al. 2007)
FGSC12433	*mata*; *rad51*::*hph*$^{+}$	
FGSC12434	*matA*; *rad51*::*hph*$^{+}$	
FGSC12528	*matA*; *msh2*::*hph*$^{+}$	
FGSC12934	*matA*; *mre11*::*hph*$^{+}$	
FGSC14432	*matA*; *rad1*::*hph*$^{+}$	
FGSC15577	*mata*; *sae2*::*hph*$^{+}$	
FGSC16079	*mata*; *rad52*::*hph*$^{+}$	
N1478	*mata*; *pan-2*^{B3}; *pan-2*$^{+}$-*hph*$^{+multi\text{-}copy}$	(Rosa, Folco et al. 2004)
N2868	*matA*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$	this study
N2945	*mata*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$; *rid*	this study
N2944	*matA*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$; *rid*	this study
NMF39	*matA*	FGSC2489
NMF62	*matA*; *fl*	FGSC4317
NMF63	*mata*; *fl*	FGSC4347
NMF488	*matA*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$	this study
NMF489	*mata*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$	this study
NMF490	*matA*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$; *msh2*(NCU02230)::*hph*$^{+}$	this study
NMF491	*mata*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$; *msh2*(NCU02230)::*hph*$^{+}$	this study
NMF492	*matA*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$; *mre11*(NCU08730)::*hph*$^{+}$	this study
NMF493	*mata*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$; *mre11*(NCU08730)::*hph*$^{+}$	this study
NMF494	*matA*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$; *sae2*(NCU10077)::*hph*$^{+}$ (heterokaryon)	this study
NMF495	*mata*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$; *sae2*(NCU10077)::*hph*$^{+}$	this study
NMF496	*matA*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{mutated\ to\ single\ RFP+}$; *rad1*(NCU07440)::*hph*$^{+}$	this study
NMF497	*mata*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$; *rad1*(NCU07440)::*hph*$^{+}$	this study
NMF498	*matA*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$; *rad10*(NCU07066)::*hph*$^{+}$	this study
NMF499	*mata*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$; *rad10*(NCU07066)::*hph*$^{+}$	this study
NMF500	*matA*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$; *rad52*(NCU04275)::*hph*$^{+}$	this study
19907-1	*mata*; *rid*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$; *rrr-1*(NCU05936)::*hph*$^{+}$	this study

Table 5.2. Primers used for probe construction.

Name	Sequence (5' to 3' direction)
msh2 5'F	TGTTCGCTTCTAGCTTGTCC
msh2 5'R	CTCTCTCCATCTTTGTCACC
mre11 5'F	AAGGAGAGAACTAAGGGAGG
mre11 5'R	GCTGTCCTAGACCATTATGC
sae2 5'F	CAGCTCGCTAGAACTAAAGG
sae2 5'R	GCTAGATCCACCCTGAATCC
rad1 5'F	ACCTACCTACCTTCACATGC
rad1 5'R	GCACAAGAGGAAGTGAGAGC
rad10 5'F	CCACCTTCACATTTACGTCC
rad10 5'R	AGACTGTACAGAACGGTAGC
rad52 5'F	TGGAGGAAAGCTAGAGCAGG
rad52 5'R	GAACTCCACTTCAGTTGTCC

Table 5.3. RIP frequency in heterozygous recombination mutant crosses. For each cross 100 progeny were tested, all crosses were performed at least twice, except the wildtype cross, which was repeated five times. SD is standard deviation.

Cross	Relevant genotype	RIP %	SD
NMF488 x NMF489	*tdimerRed*$^{+}$ x *tdimerRed*$^{+}$	0.82	0.02
NMF488 x NMF495	*tdimerRed*$^{+}$ x *tdimerRed*$^{+}$; *sae2*	0.60	0.01
NMF489 x NMF494	*tdimerRed*$^{+}$ x *tdimerRed*$^{+}$; *sae2* (het)	0.89	0.04
NMF488 x NMF497	*tdimerRed*$^{+}$ x *tdimerRed*$^{+}$; *rad1*	0.74	0.04
NMF489 x NMF496	*tdimerRed*$^{+}$ x *rfp*$^{1x+}$; *rad1*	0.54	0.01
NMF489 x NMF490	*tdimerRed*$^{+}$ x *tdimerRed*$^{+}$; *msh2*	0.87	0.06
NMF489 x NMF492	*tdimerRed*$^{+}$ x *tdimerRed*$^{+}$; *mre11*	0.63	0.01
NMF489 x NMF498	*tdimerRed*$^{+}$ x *tdimerRed*$^{+}$; *rad10*	0.84	0.05
NMF489 x NMF500	*tdimerRed*$^{+}$ x *tdimerRed*$^{+}$; *rad52*	0.50	0.09

Table 5.4. Nucleus specific RIP frequency in heterozygous recombination mutant crosses. For each cross 100 progeny were tested, all crosses were performed in triplicate. SD is standard deviation.

Cross	Relevant genotype	RIP % (*tdimerRed*)	SD (*tdimerRed*)	RIP % (*pan-2*)	SD (*pan-2*)
N1478 x NMF488	*pan-2*$^{multi-copy}$ x *tdimerRed*$^{+}$	0.69	0.06	0.87	0.15
N1478 x NMF490	*pan-2*$^{multi-copy}$ x *tdimerRed*$^{+}$; *msh2*	0.61	0.09	0.83	0.06
N1478 x NMF492	*pan-2*$^{multi-copy}$ x *tdimerRed*$^{+}$; *mre11*	0.37	0.16	0.83	0.15
N1478 x NMF494	*pan-2*$^{multi-copy}$ x *tdimerRed*$^{+}$; *sae2* (het)	0.83	0.05	0.87	0.05
N1478 x NMF496	*pan-2*$^{multi-copy}$ x *rfp*$^{1x+}$; *rad1*	0.05	0.08	0.91	0.12
N1478 x NMF498	*pan-2*$^{multi-copy}$ x *tdimerRed*$^{+}$; *rad10*	0.68	0.12	1.05	0.06

Table 5.5. RIP frequency in homozygous recombination mutant crosses. For each cross 100 progeny were tested, all crosses were performed at least three and upto five times. SD is standard deviation.

Cross	Relevant genotype	RIP %	SD
NMF488 x NMF489	*tdimerRed*$^{+}$ x *tdimerRed*$^{+}$	0.82	0.02
NMF490 x NMF491	*tdimerRed*$^{+}$; *msh2* x *tdimerRed*$^{+}$; *msh2*	0.64	0.04
NMF494 x NMF495	*tdimerRed*$^{+}$; *sae2* x *tdimerRed*$^{+}$; *sae2* (het)	0.75	0.01
NMF496 x NMF497	*tdimerRed*$^{+}$; *rad1* x *rfp*$^{1x+}$; *rad1*	0.39	0.08
NMF498 x NMF499	*tdimerRed*$^{+}$; *rad10* x *tdimerRed*$^{+}$; *rad10*	0.85	0.05

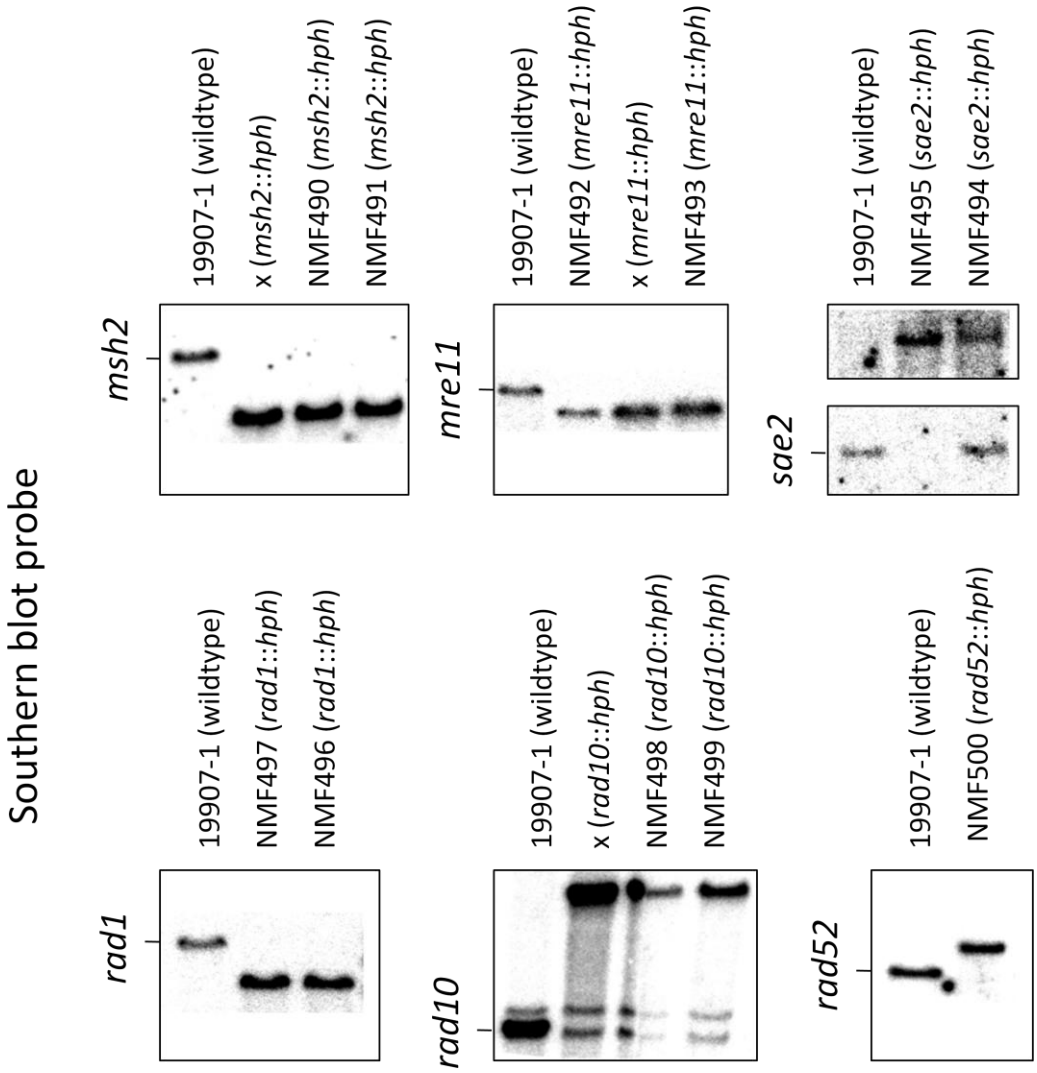


Figure 5.1. Southern analysis confirms replacement of genes with *hph*. Genomic DNA from *tdimerRed* strains derived from recombination mutants from the *N. crassa* single-gene deletion collection was digested with *Hind*III. Southern blot analyses were carried out by probing with radiolabelled fragments 5' of the region replaced by *hph* for each of the indicated genes. Correct integration is confirmed by a shift in band size between strain 19907-1 and the *hph* replacement strains. The expected wildtype size for each probe is indicated. All strains are homokaryotic except NMF494 which appears to be heterokaryotic and has bands indicative of wildtype *sae2* and *sae2*::*hph*. The faint doublet in *rad10*::*hph* strains present at the same fragment size as wildtype *rad10* is nonspecific binding of the *rad10* probe.

General summary and conclusions

Kyle R. Pomraning

Summary and conclusions

The substrate for and result of processes that inactivate repetitive DNA in filamentous fungi during the sexual cycle, which include repeat-induced point mutation (RIP) and methylation induced premeiotically (MIP), have been well characterized in *N. crassa* and *A. immersus* for over two decades. However, how repetitive DNA is detected and specifically targeted remains a mystery. The work described here characterizes RIP and its dependence on *rid1* (*RIP-defective-1*) in *F. graminearum*, a new model fungus for the study of RIP and potentially the isolation of recessive *rid* mutants (Chapter 2). With some help from my co-workers, I applied high-throughput whole genome sequencing to the task of genetically mapping novel mutants. I developed methods for bioinformatic analysis and tested them by mapping recombination junctions by whole genome SNP genotyping of single recombinant strains derived from single octads, the first complete ordered meiosis analyzed in this way from any organism (Chapter 3). My mapping strategy was first applied to identify the only known temperature sensitive cell cycle mutation in *N. crassa* as *spe-1*$^{ndc-1}$ (Chapter 3). I then mapped two previously isolated but novel RIP mutants, which I call *rrr-1* (*reduced RIP and recombination 1*), by means of my new methods (Chapter 4), and I tested a large collection of deletion alleles involved in recombination for effects on RIP and premeiotic recombination. In ongoing work, that will be described in future manuscripts, I am mapping the RIP defect in a wild-collected *N. crassa* strain, and I am constructing mutant alleles of *rid* and *masc1* to test for altered activities of these methyltransferase-like proteins (see "ongoing and future studies" below).

Characterization of RIP in *Fusarium graminearum*

RIP was discovered and first characterized in *N. crassa* (Selker 1990), a heterothallic fungus that requires two strains of opposite mating type to undergo sexual development and meiosis. Since RIP occurs during the sexual cycle, prior to meiosis, construction of recessive *rid* mutants in *N. crassa* is difficult. Mutagenesis strategies can be employed to mutate both parents of a cross followed by screening for differences in RIP frequency (Chapter 4) but the only types of mutants likely to be isolated will be dominant or semi-dominant since the same gene would have to be mutated in both parents of a cross if it acts recessively. To make

isolation of recessive *rid* mutants feasible, we employed *F. graminearum*, a homothallic Sordariomycete relatively closely related to *N. crassa.* Homothallic fungi are able to undergo sexual development without a partner of dissimilar mating type, i.e. they "self". We found that RIP in *F. graminearum* occurs with relatively high frequency and is dependent on *rid-1*, as in *N. crassa*. In contrast to *N. crassa*, however, we did not observe *de novo* methylation of cytosines after RIP (Chapter 2). We sequenced a strain of *F. graminearum* enriched in SNPs compared to the reference strain, which will prove useful for future mapping of mutants. A mutant screen for additional recessive *rid* mutants proved unsuccessful. Our newly gained knowledge about sexual development in *F. graminearum* and another homothallic Sordariomycete, *Sordaria macrospora*, suggests that screens to uncover recessive mutations will be as difficult in these organisms as in *Neurospora*. One reason for this may be the apparent requirement for hyphal fusion before sexual development can continue, suggested by sterile phenotypes observed with mutated alleles of the *F. graminearum soft* and *ham2* and *S. macrospora pro* genes. In the overwhelming majority of all cases, fusion events will result in heterokaryotic mycelia in which mutated and non-mutated nuclei coexist, thus negating the advantage conferred by homothallism. However, if homokaryotic spores are isolated post-mutagenesis and selfed one by one as done in Chapter 2, isolation of novel *rid* mutants is possible, though time consuming, as we have not yet worked out a method of selection.

Mapping RIP mutants in *Neurospora crassa*

Chapter 3 is dedicated to development of the methodology for fast mapping of mutants by bulk segregant analysis and high-throughput sequencing, a strategy applicable to any organism able to undergo meiosis. I showed that this method works well and I revealed some pitfalls by means of a proof-of-principle study that identified the gene with the sole cell cycle mutation ever collected in *N. crassa.* I then utilized this method to map the defects in two novel mutants, which mapped to the same gene, *rrr-1*, so called when I found that not only RIP but also premeiotic recombination frequency is semi-dominantly reduced during sex (Chapter 4). I found that RRR-1 belongs to a previously uncharacterized group of dynamin-like proteins and that it associates with the nuclear membrane. We are in the process of

mapping a third locus that affects RIP frequency in *N. crassa* Adiopodoumé (see ongoing and future studies below), which has remained elusive since its original characterization.

Lastly, I performed a reverse genetics screen to test whether genes involved in recombination are also involved in RIP. This was motivated by the finding that both RIP and recombination were affected in the mutants we identified in the forward genetic screen. We found that deletion of many genes essential for recombination causes low fertility, thus inhibiting the ability to assay for RIP, and that genes involved in the later stages of homologous recombination by single-strand annealing are not essential for RIP (Chapter 4).

Significance

The overall goal of the work presented in this thesis was to elucidate the mechanism by which repetitive DNA is specifically targeted for mutation by RIP. During the course of my work, I helped to characterize RIP in a homothallic fungus to make identification of recessive RIP mutants feasible, I carried out a hunt for recessive RIP mutants, and I developed methods for quickly mapping mutants in both *N. crassa* and *F. graminearum*. I mapped semi-dominant RIP mutations from two independent strains to different positions in a single gene, thus doubling the alleles of genes known to affect RIP in *N. crassa*. I systematically investigated whether RIP and recombination are linked and ruled out a long list of candidate genes involved in recombination. The DNA pairing mechanism is difficult to address by direct biochemical approaches, as the tissue in which RIP occurs is small and difficult to obtain from synchronized structures. One goal has been to make this system amenable to studies during the vegetative growth phase, where large quantities of mycelium for in-depth studies can be obtained. To this end I have begun to dissect the *rid* and *masc1* genes.

Ongoing and future studies

Characterization of *rid* and *masc1* activity

In *N. crassa* silencing is accomplished by RIP of repetitive DNA (Selker, Cambareri et al. 1987; Cambareri, Jensen et al. 1989) and is dependent on *rid* (Freitag, Williams et al. 2002) while in *A. immersus* silencing is induced by MIP of the cytosines of repetitive DNA (Goyon and Faugeron 1989) and is dependent on *masc1* (Malagnac, Wendel et al. 1997). These two

processes bear striking similarity in that (1) both occur during sexual development but prior to meiosis (Selker, Cambareri et al. 1987; Faugeron, Rhounim et al. 1990), (2) both target duplicated sequences of at least 300-400 base pairs in length (Goyon, Barry et al. 1996; Watters, Randall et al. 1999) and (3) both processes seem to target cytosine and are dependent on orthologous cytosine methyltransferase-like genes (Malagnac, Wendel et al. 1997; Freitag, Williams et al. 2002).

To date the enzymatic activity of RID, Masc1, or any other orthologous protein has not been directly measured, leaving the question of just what these proteins do open to debate. As outlined in the introduction, a number of pathways for mutation of cytosine to thymine are possible and include (1) deamination of cytosine to uracil by a cytosine deaminase, (2) deamination of cytosine to uracil by a cytosine methyltransferase under conditions when *S*-adenosylmethionine (SAM) is limiting, which favor deamination of an unstable intermediate over methylation (Sharath, Weinhold et al. 2000), or (3) methylation of cytosine to 5-methylcytosine followed by deamination of 5-methylcytosine to thymine (Selker 1990; Mautino and Rosa 1998). Hypothesis (2) could explain the different outcomes of RIP and MIP if the ability to interact with SAM is different between RID and Masc1 while Hypothesis (3) could explain the difference if *N. crassa* has a 5-methylcytosine deaminase that is not active or present in *A. immersus*.

There is precendence for Hypothesis (2), as formation of uracil by cytosine methyltransferase enzymes has been observed in the bacterial cytosine methyltransferases M.HpaII (Shen, Rideout et al. 1992; Bandaru, Wyszynski et al. 1995; Shen, Zingg et al. 1995) and M.SssI (Bandaru, Wyszynski et al. 1995) but not M.MspI (Zingg, Shen et al. 1998) when the cofactor SAM is limiting or when cofactor analogues are supplied (Zingg, Shen et al. 1996). For the mammalian cytosine methyltransferase Dnmt1, strong interactions between guanine and the CpG recognition site of the enzyme are thought to prevent premature release of the base and solvent access to the active site, which leads to deamination (Svedruzic and Reich 2005). The only major difference between RIP and MIP known is the outcome observed in progeny after a cross, cytosine methylation of repeats by MIP versus C:G to T:A transition mutations in repeats by RIP. How this difference occurs is not known and is the subject of these further studies (see Appendix 1). To determine whether RID and

Masc1 have the same activity or whether methylation versus mutation is caused by differences intrinsic to the methyltransferase-like genes I have begun to perform swapping, mutation and deletion experiments with *rid* and *masc1* genes in *N. crassa* (Fig. A1.1). Outcomes will be assayed as described in Chapter 4 after crossing strains with these constructs to strains with the RIP marker, *tdimerRed*.

Mapping RIP defects in the wild-collected Adiopodoumé strain

A second ongoing project deals with the mapping of QTLs for RIP frequency that have been identified in wild-collected *Neurospora* strains. RIP competence has been tested in a world-wide collection of isolates by crossing each with a RIP tester strain harboring a phenotypically assayable duplication (*erg-3*) and observing the number of progeny that display inactivation of the duplication. Since collected strains have been assayed by crossing to the same strains, any difference in RIP can be attributed to differences in the collected strain. In collected strains RIP frequency acts as a quantitative trait displaying large variation (Noubissi, McCluskey et al. 2000; Noubissi, Aparna et al. 2001; Bhat, Noubissi et al. 2003).

Six strains reproducibly display low or very low frequency inactivation of duplications by RIP (Adiopodoumé, Sugartown, Adiopodoumé-7, Fred, Coon, and Bayan Lepas; (Noubissi, McCluskey et al. 2000; Noubissi, Aparna et al. 2001; Bhat, Noubissi et al. 2003). The Adiopodoumé strain from the Ivory Coast is especially interesting in that it is known to harbor active LINE-like *Tad* retrotransposons (Kinsey 1989; Kinsey 1990) which in the Oak Ridge (OR) reference strain of *N. crassa* have all been inactivated by mutation (Galagan, Calvo et al. 2003; Galagan and Selker 2004). Adiopodoumé appears to be RIP-competent and is able to mutate the multi-copy *Tad* elements during crosses (Kinsey, Garrett-Engele et al. 1994; Anderson, Tang et al. 2001) but at much lower frequency than most *N. crassa* strains.

Mapping the dominant RIP phenotype in Adiopodoumé has been difficult for a number of reasons. First, doing reproducible RIP frequency assays is incredibly time-consuming and second, RIP frequency often appears linked to mating type which is most likely because strains of opposite mating type have to be tested for RIP frequency using different RIP tester strains. The dominant RIP defect in Adiopodoumé was previously mapped to a translesion polymerase very near *mat*, as it had a number of non-synonymous mutations

compared to the reference strain (Tamuli and Kasbekar 2008). However, the phenotype was never properly complemented and we have found by light coverage whole genome sequencing that almost every gene in Adiopodoumé has a number of non-synonymous mutations compared to the reference strain (data not shown). Additionally, the translesion polymerase in question has been shown not to be essential for RIP (Tamuli, Ravindran et al. 2006).

Given this we have attempted to map the dominant RIP trait in Adiopodoumé (FGSC430) by bulk segregant analysis and whole genome sequencing using a small population of 71 recombinant progeny from a cross between *N. crassa* strains from Adiopodoumé and Mauriceville. This method was successfully used previously (see Chapter 4) to map dominant RIP loci in mutants where RIP frequency assays were complicated by an effect originating at or near *mat*.

Appendix 1: Materials related to ongoing and future studies on RIP

Kyle R. Pomraning

Construction of novel *rid* and *masc1* alleles

Strains, crosses and growth conditions

All strains were grown according to standard conditions (Davis 2000) in Vogel's minimal medium (VMM) with 1.5% sucrose at 32°C and supplemented with 500 mg/L of histidine when appropriate. Crosses were performed on synthetic crossing medium (SC) with 0.4% sucrose and supplemented with 250 mg/L histidine when appropriate. For standard crosses, random ascospores were recovered in water, while RIP assay crosses were allowed to mature for four weeks before collection and heat shock at 65°C for 60 min on VMM with 500 mg/L of histidine and FGS (0.5 g/L fructose, 0.5 g/L glucose, 20 g/L sorbose) as the carbon source. After overnight growth at room temperature, viable progeny were collected and incubated in slants with VMM with 2% sucrose and 500 mg/L of histidine. NMF533 (*his-3*; *rid*::*hph*$^{+}$; *mus-52*::*hph*$^{+}$) was constructed by crossing NMF510 (*rid*::*hph*$^{+}$) with N2834 (*his-3*; *mus-52*::*hph*$^{+}$) and screening for hygromycin resistant progeny unable to grow without histidine supplementation.

Plasmid construction

Plasmid EMF755 was constructed by first cloning a 4832 bp PCR fragment of *rid* using primers OMF1344 and OMF1345 into pCR-4Blunt. The length of the 3' *rid* flank was then extended by 557 bp by opening EMF755 with *NotI* and *XhoI* and cloning a *NotI*/*XhoI* genomic PCR fragment containing some of the flanking gene *NCU02033*. A 988 bp PCR containing the *TrpC* promoter followed by *bar*, a gene conferring resistance to phosphinothricin, with flanking *BamHI* sites was PCR amplified using OMF2850 and OMF2872 and fortuitously cloned into the *BglII* site adjacent to the 3' end of *rid* without cutting the *BglII* site further 3' of *rid*. EMF755 (*rid*$^{+}$:*bar*$^{+}$) was Sanger sequenced and used as a base for subsequent plasmid construction. EMF760 (*rid*-R290N^{+}:*bar*$^{+}$), EMF762 (*rid*-C359G^{+}:*bar*$^{+}$), EMF763 (*rid*-H422L^{+}:*bar*$^{+}$), and EMF765 (*rid*-Δ579-845^{+}:*bar*$^{+}$) were constructed by site-directed mutagenesis using primer pairs OMF1508 and OMF1509, OMF1502 and OMF1503, OMF1510 and OMF1511, and OMF2551 and OMF2552 respectively. EMF761 (*rid*-Δ283-574^{+}:*bar*$^{+}$) was constructed by overlap PCR using primers OMF2953 and OMF2954. EMF768 (*rid*-Δ283-574::*masc-1*-226-526^{+}:*bar*$^{+}$) was constructed by first amplifying *masc-1* from *A. immersus*

genomic DNA using primers OMF2951 and OMF2952. This 1.0 kb PCR fragment was purified and used as the primers for megaprimer PCR mutagenesis (Barik 2002) of EMF755. EMF767 (*rid*-Δ1-845::*masc-1*-1-538^{+}:*bar*$^{+}$) was constructed by first amplifying *masc-1* from *A. immersus* genomic DNA using primers OMF2949 and OMF2950. This 1.7 kb PCR fragment was purified and Gibson cloned (Gibson, Young et al. 2009) into an EMF755 PCR fragment with adjoining overhangs amplified using primers OMF3004 and OMF3005. *rid*-Δ16-282^{+}:*bar*$^{+}$) is to be constructed by Gibson cloning (Gibson, Young et al. 2009) using PCR primers OMF3006 and OMF3007. All mutated fragments were Sanger sequenced and subcloned back into EMF755.

Transformation of mutated *rid*

NMF533 (*his-3*; *rid*::*hph*$^{+}$; *mus-52*::*hph*$^{+}$) was confirmed by Southern analyses with radiolabelled *hph* probe and is to be transformed by electroporation with plasmid EMF755 and its derivatives. Transformants will be obtained on VMM with FGS as the carbon source, 500 mg/L of histidine and 10 µL/mL phosphinothricin extract at 22°C. Heterokaryotic transformants are then to be backcrossed to N2834 and assayed by Southern blot to obtain homokaryotic strains.

RIP assay

Pure strains where *rid*::*hph*$^{+}$ had been replaced by homologous recombination with *rid*(mutant):*bar*$^{+}$; are to be tested for their RIP phenotype by crossing to N2944 (*rid*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$). Random ascospores are to be collected four weeks after fertilization in water and germinated by heat shock for 1 hour at 65 C on VMM with FGS as the carbon source. Histidine should not be provided to ensure that only progeny with *Pccg-1-tdimerRed*$^{+}$ at *his-3* are analyzed. After 16 hours 100 germinated progeny from each cross will be examined for red fluorescence and the inactivation % calculated as (RFP^{-} progeny / total germinated progeny). From each cross that displays inactivation, RFP^{-} strains will be collected and analyzed for mutation and cytosine methylation of *tdimerRed* by digesting genomic DNA with 5-methylcytosine sensitive (*Sau3AI*) and insensitive (*DpnII*) isoschizomers and analyzing by Southern blot with p32-*tdimerRed*.

Mapping the mutation(s) causing dominant reduction of RIP in *N. crassa* Adiopodoumé

Strains and growth conditions

N. crassa strains were grown at 32 C for conidia and DNA in Vogel's minimal medium (VMM) with 1.5% sucrose (Davis 2000). Crosses were performed on synthetic crossing medium (SC) with 0.4% sucrose (Davis 2000). Ascospores were recovered in water for the scoring of RIP after four weeks and after two weeks for all other crosses. Germination of ascospores was induced by heat shock at 65 C for 1 hour on VMM with FGS (0.5 g/L fructose, 0.5 g/L glucose, 20 g/L sorbose) as the carbon source. Strains were tested for mating type by crossing with *fl* strains FGSC4317 and FGSC4347 (Davis 2000).

Mutant mapping by bulk segregant analysis and high-throughput sequencing

NMF38 (Mauriceville-1-d) was crossed to FGSC430 (Adiopodoumé). 71 recombinant progeny were collected and twice tested for their RIP phenotype. First by crossing to either N2944 (*mat-A*; *rid*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$) or N2945 (*mat-a*; *rid*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$) and then by crossing to either NMF488 (*mat-A*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$) or NMF489 (*mat-a*; *his-3*$^{+}$::*Pccg-1-tdimerRed*$^{+}$). The progeny from these crosses were germinated and after 18 hours, 100 germinated progeny from each cross were examined for red fluorescence and the RIP % calculated as:

$$\text{RIP }\% = 1 - \frac{2(\text{RFP}^{+}\text{progeny})}{(\text{total progeny})}$$

since only half the progeny are expected to have the *tdimerRed* gene.

Genomic DNA from the progeny of interest as well as the parent strains was prepped as described previously (Pomraning, Smith et al. 2009). The DNA was prepped for Illumina sequencing as described previously (Pomraning, Smith et al. 2011) with the exception that DNA from each strain was ligated to adapters with a different barcode. Single-end 51 bp sequencing was performed on an Illumina HiSeq (Illumina, San Diego, USA) and the reads parsed by adapter. Reads originating from NMF38 and N157 were aligned to the *N. crassa* reference genome using bowtie to identify SNPs unique to each strain which were used to construct a SNPome. Each recombinant progeny strain was then mapped to the SNPome to

determine the contribution of each parent strain at every SNP site as described previously (Pomraning, Smith et al. 2011).

RIP frequency in recombinant progeny

We tested RIP frequency of 71 recombinant progeny after crossing NMF38 (Mauriceville-1-d) to FGSC430 (Adiopodoumé) using RIP tester strains with or without *rid* in the genetic background. The results of the RIP test crosses are summarized in Table A1.2 by strain and Figure A1.2 by mating type. RIP frequency is clearly linked to *mat* in the recombinant progeny, however we believe this is not the only locus dominantly affecting RIP frequency present in Adiopodoumé. Thus we sorted the progeny by mating type and looked for segregation of the RIP frequency trait within each mating type pool.

RIP frequency ranges widely in the *mat-a* strains so we have focused on these for further analysis and whole genome sequencing. We chose four *mat-a* strains that appeared to inherit the low frequency RIP trait from Adiopodoumé (xKP20-22, -24, -29 and -50) and four *mat-a* strains that appeared to inherit the high frequency RIP trait from Mauriceville (xKP20-14, -54, -68 and -70). These eight strains as well as the Adiopodoumé and Mauriceville parents were sequenced independently on an Illumina HiSeq.

Table A1.1. Primers used for plasmid construction.

OMF	Sequence (5' to 3' direction)
1344	CACAACGGAGCAGTTGGTAAT
1345	GAGGTGAACGACAGAAGCAAG
1502	CACTTGTCGCCTCCTGGTCAAGTATGGTCTCC
1503	CGGAGACCATACTTGACCAGGAGGCGACAAG
1508	GCCGGGTGATACAAACGCCGGAGCCGGTGGTG
1509	CCACCGGCTCCGGCGTTTGTATCACCGGC
1510	CATGGATTCACTGACCTTGGCTACTCTGTCCGATGG
1511	CGGACAGAGTAGCCAAGGTCAGTGAATCCATGGAC
2551	TCTTGTCGAATGCTAAGATAAACGGGACA
2552	TGTCCCGTTTATCTTAGCATTCGACAAGA
2850	TACTGGATCCGAAGTTATTCTAGTCGACAG
2872	TACTGGATCCATCTGCAGGTCGACGGATCAG
2949	TTCCCGCTCAGCAACATCTACCCCCCCTCCGCAGAATGTCGGAAAGAAGATACGAGGC
2950	TCGCACCTTTGGGGATCAAGACTTTAGTCTCTACTCCATCCTCTCGAAATCC
2951	CGTGGCCCGAAACAAGTGCCTTCAGATCAGAAGTATACCTTTGGGGATACG
2952	ACGGATTATATTGTCCCGTTTATCTAAGCACTGCCTCAAGGTACTCATAATG
2953	CGTGGCCCGAAACAAGTGCCTTCAGATCAGTGCTTAGATAAACGGGACA
2954	ACGGATTATATTGTCCCGTTTATCTAAGCACTGATCTGAAGGCACTTGTTTC
3004	AGACTAAAGTCTTGATCCCCAA
3005	CATTCTGCGGAGGGGGGGTA
3006	GTTATAGACGACGAAGACGACGTCAAGTATACGGCCGGTGATACATTC
3007	GAATGTATCACCGGCCGTATACTTGACGTCGTCTTCGTCGTCTATAAC

Table A1.2. RIP frequency in recombinant Mauriceville x Adiopodoumé progeny. The first value in the RFP columns indicates the number of RFP^+ and RFP^- ascospores observed after crossing to N2944 or N2945 and the second value indicates the results after crossing to NMF488 or NMF489. N/A indicates the cross was not fertile enough to determine RIP frequency. SD is standard deviation.

Strain	*mat*	RFP+	RFP-	RIP %	SD RIP %
N157-1	A	50 / 15	50 / 15	0%	0%
N157-2	A	49 / 42	51 / 48	4%	3%
NMF38-1	a	25 / 30	75 / 70	45%	7%
NMF38-2	a	22 / 19	78 / 81	59%	4%
xKP20-1	A	42 / n/a	58 / n/a	16%	n/a
xKP20-2	A	54 / 59	46 / 141	17%	35%
xKP20-3	A	46 / 51	54 / 49	3%	7%
xKP20-4	A	53 / 51	47 / 49	-4%	3%
xKP20-5	a	55 / 38	45 / 62	7%	24%
xKP20-6	a	50 / n/a	50 / n/a	0%	n/a
xKP20-7	A	77 / 47	60 / 53	-3%	13%
xKP20-8	A	n/a / 46	n/a / 54	8%	n/a
xKP20-9	A	50 / 13	50 / 16	5%	7%
xKP20-10	a	34 / 31	34 / 69	19%	27%
xKP20-12	A	19 / 42	32 / 58	21%	7%
xKP20-14	a	43 / 27	57 / 73	30%	23%
xKP20-15	A	45 / n/a	55 / n/a	10%	n/a
xKP20-16	A	n/a / 50	n/a / 50	0%	n/a
xKP20-17	a	52 / n/a	48 / n/a	-4%	n/a
xKP20-19	A	n/a / 49	n/a / 51	2%	n/a
xKP20-20	a	n/a / 52	n/a / 48	-4%	n/a
xKP20-21	A	48 / 47	52 / 53	5%	1%
xKP20-22	a	51 / 47	49 / 53	2%	6%
xKP20-23	a	46 / 43	54 / 57	11%	4%
xKP20-24	a	55 / 45	45 / 55	0%	14%
xKP20-26	a	44 / 41	56 / 59	15%	4%
xKP20-27	A	53 / 50	47 / 50	-3%	4%
xKP20-28	A	47 / n/a	53 / n/a	6%	n/a
xKP20-29	a	48 / 48	52 / 52	4%	0%
xKP20-30	A	50 / n/a	50 / n/a	0%	n/a
xKP20-31	A	8 / n/a	8 / n/a	0%	n/a
xKP20-32	A	n/a / 59	n/a / 41	-18%	n/a
xKP20-33	A	48 / 51	52 / 49	1%	4%
xKP20-34	A	43 / n/a	57 / n/a	14%	n/a

xKP20-35	A	38 / 46	62 / 54	16%	11%
xKP20-36	A	55 / 46	45 / 54	-1%	13%
xKP20-37	A	52 / 50	48 / 50	-2%	3%
xKP20-38	A	n/a / 43	n/a / 57	14%	n/a
xKP20-39	A	52 / 52	48 / 48	-4%	0%
xKP20-40	A	14 / n/a	27 / n/a	32%	n/a
xKP20-41	a	41 / 38	59 / 62	21%	4%
xKP20-42	a	39 / 49	61 / 51	12%	14%
xKP20-43	a	43 / 39	57 / 61	18%	6%
xKP20-44	a	43 / 41	57 / 59	16%	3%
xKP20-45	A	46 / 48	54 / 52	6%	3%
xKP20-46	A	49 / 49	51 / 51	2%	0%
xKP20-47	A	54 / 47	46 / 53	-1%	10%
xKP20-48	A	50 / 53	50 / 47	-3%	4%
xKP20-49	a	52 / 44	48 / 56	4%	11%
xKP20-50	a	56 / 44	44 / 56	0%	17%
xKP20-51	a	41 / 35	59 / 65	24%	8%
xKP20-52	a	44 / 47	56 / 53	9%	4%
xKP20-53	a	48 / n/a	52 / n/a	4%	n/a
xKP20-54	a	34 / 39	66 / 61	27%	7%
xKP20-55	A	40 / 45	60 / 55	15%	7%
xKP20-56	a	43 / 42	57 / 58	15%	1%
xKP20-57	a	32 / 51	68 / 49	17%	27%
xKP20-58	a	n/a / 39	n/a / 61	22%	n/a
xKP20-59	a	54 / n/a	46 / n/a	-8%	n/a
xKP20-60	A	44 / 47	56 / 53	9%	4%
xKP20-61	A	52 / 44	48 / 56	4%	11%
xKP20-62	A	50 / 77	50 / 123	12%	16%
xKP20-63	a	44 / n/a	56 / n/a	12%	n/a
xKP20-64	a	44 / 46	56 / 54	10%	3%
xKP20-65	a	43 / 47	57 / 53	10%	6%
xKP20-66	A	68 / n/a	62 / n/a	-5%	n/a
xKP20-67	a	35 / 49	65 / 51	16%	20%
xKP20-68	a	27 / 27	73 / 73	46%	0%
xKP20-69	A	45 / 37	34 / 33	-10%	6%
xKP20-70	a	41 / 30	59 / 70	29%	16%
xKP20-71	a	28 / n/a	45 / n/a	23%	n/a

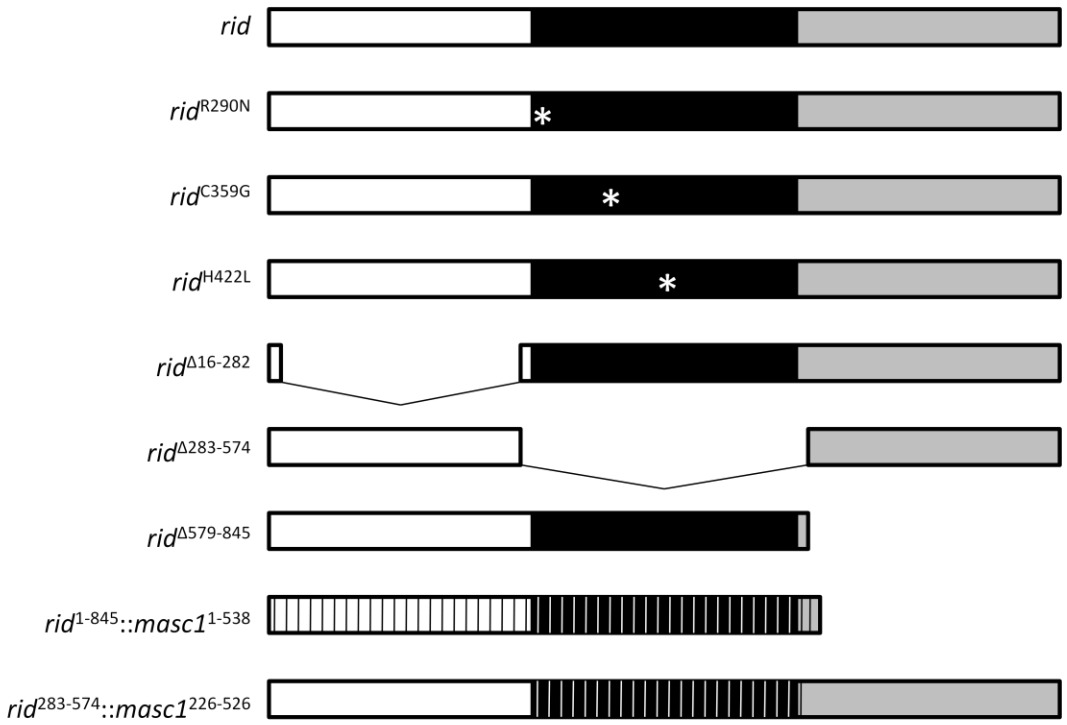


Figure A1.1. Mutated *rid* constructs. The *N. crassa rid* gene has been mutated at specific residues and a series of deletion and replacement mutants made. The black part of the gene represents the methyltransferase domain as identified by InterProScan. Homologous regions of *masc1* from *A. immersus* are indicated by bars. Asterisks indicate point mutations.

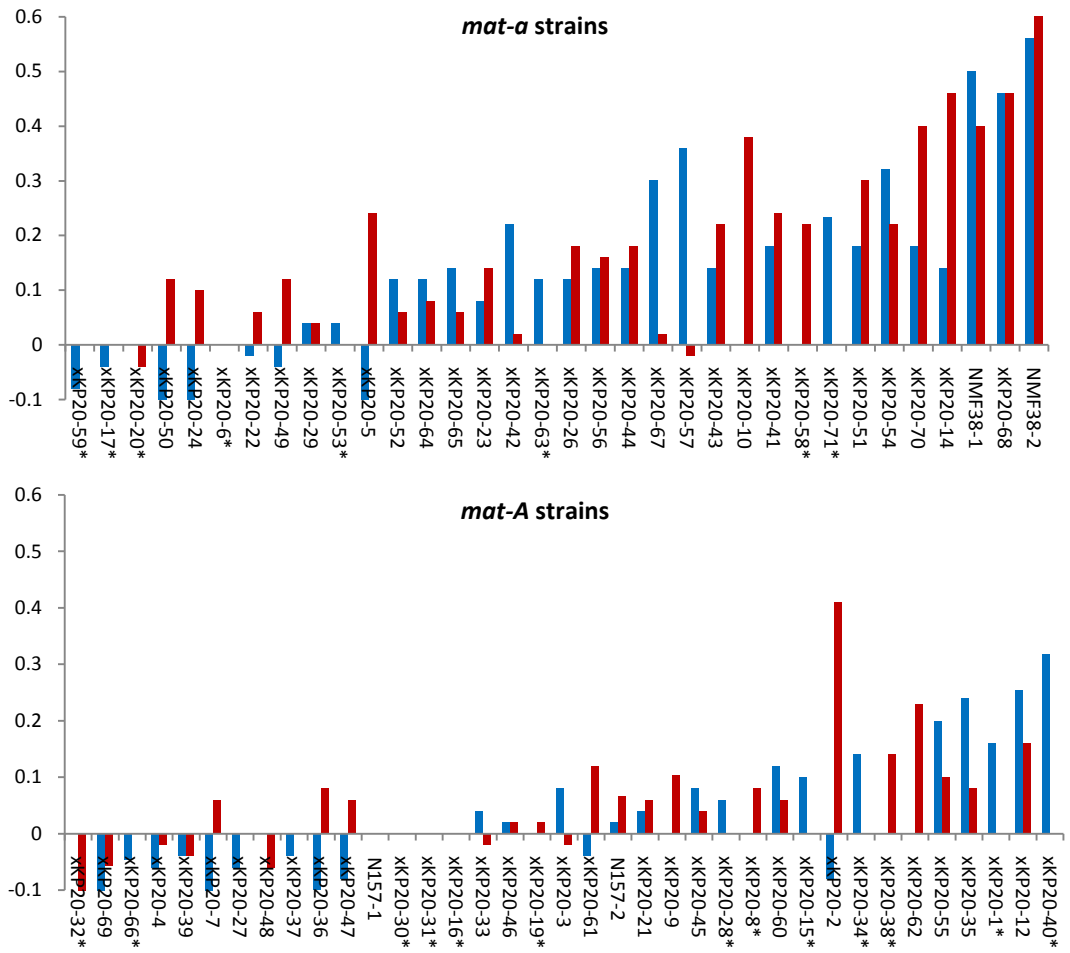


Figure A1.2. RIP frequency in recombinant Mauriceville x Adiopodoumé progeny (xKP20). The y-axis indicates RIP frequency observed after crossing to N2944 or N2945 (blue bars) and NMF488 or NMF489 (red bars). Asterisks indicate strains for which only one RIP test cross yielded enough ascospores to assay RIP frequency. The strains are sorted by mating type and average RIP frequency. Strains xKP20-14, -22, -24, -29, -50 -54, -68, -70, NMF38 and N157 were chosen for high-throughput sequencing.

Appendix 2: List of additional manuscripts published, in press or in preparation

Wilcox BW, **Pomraning KR**, Bredeweg EL, Freitag M and Hays JB. Mutation accumulation in *msh2* and wildtype *Arabidopsis thaliana*. In preparation.

Bowring F, **Pomraning KR**, Catcheside D, Freitag M. High resolution mapping determines the whole genome extent of recombination in the four products of meiosis from an octad of *Neurospora crassa*. In preparation.

Vining KJ, **Pomraning KR**, Wilhelm LJ, Pellegrini M, Yanming DI, Mockler TC, Freitag M, Strauss S. DNA methylation dynamics during *in vitro* propogation of *Populus trichocarpa*. In preparation.

Schmoll M, **Pomraning KR**, Baker S, Freitag M. The sterility locus of industrial *Trichoderma reesei*. In preparation.

Yue H, Yam C, **Pomraning KR,** Freitag M, Oliferenko, S. Characterization of the Bubbles CDP-diacylglycerol synthase in *Schizosaccharomyces pombe*. In preparation.

Pomraning KR, Smith KM, Bredeweg EL, Connolly LR, Phatale PA, Freitag M. Library preparation for rapid Illumina genome sequencing. In Walker JM (Ed.), **Methods Mol Biol.** 2012;944:1-22.

Puckett S, Reese KA, Mitev GM, Mullen V, Johnson RC, **Pomraning KR**, Mellbye BL, Tilley LD, Iversen PL, Freitag M, Geller BL. Bacterial resistance to antisense peptide-phosphorodiamidate morpholino oligomers. **Antimicrob Agents Chemother.** 2012,Sep17 [Epub ahead of print].

Slavov GT, DiFazio SP, Martin J, Schackwitz W, Muchero W, Rodgers-Melnick E, Lipphardt MF, Pennacchio CP, Hellsten U, Pennacchio L, Gunter LE, Ranjan P, Vining K, **Pomraning KR**, Wilhelm LJ, Pellegrini M, Mockler T, Freitag M, Geraldes A, El-Kassaby YA, Mansfield SD, Cronk QCB, Douglas CJ, Strauss SH, Rokhsar D, Tuskan GA. Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. **New Phytol.** 2012;196:713-25.

Vining KJ, **Pomraning KR**, Wilhelm LJ, Priest HD, Pellegrini M, Mockler TC, Freitag M, Strauss S. Dynamic DNA cytosine methylation in the Populus trichocarpa genome: tissue-level variation and relationship to gene expression. **BMC Genomics** 2012;13(1):27.

Smith KM, Phatale PA, Bredeweg EL, Connolly LR, **Pomraning KR**, Freitag M. Epigenetics of filamentous fungi in Epigenetic Regulation and Epigenomics. **Encyclopedia of Molecular Biology and Molecular Medicine**, Wiley-Blackwell. 2012.

Smith KM, Phatale PA, Sullivan CM, **Pomraning KR**, Freitag M. Heterochromatin is required for normal distribution of *Neurospora* CenH3. **Mol Cell Biol.** 2011;31(12):2528-42.

Pomraning KR, Smith KM, Freitag M. Genome-wide high throughput analysis of DNA methylation in eukaryotes. **Methods** 2009;47:142-50.

Cowen JP, Fornari DJ, Shank TM, Love B, Glazer B, Treusch AH, Holmes RC, Soule SA, Baker ET, Tolstoy M, **Pomraning KR**. Volcanic Eruptions at East Pacific Rise Near 9° 50'N. **Eos Trans. AGU** 2007;88(7):81.

Pike DB, Cai S, **Pomraning KR**, Firpo MA, Fisher RJ, Shu X, Prestwich GD, Peattie RA. Heparin-regulated release of growth factors in vitro and angiogenic response *in vivo* to implanted hyaluronan hydrogels containing VEGF and bFGF. **Biomaterials** 2006;27:5242–51.

References

Alani, E., R. Thresher, et al. (1992). "Characterization of DNA-binding and strand-exchange stimulation properties of y-RPA, a yeast single-strand-DNA-binding protein." J Mol Biol **227**(1): 54-71.

Aleksenko, A. and A. J. Clutterbuck (1996). "The plasmid replicator AMA1 in Aspergillus nidulans is an inverted duplication of a low-copy-number dispersed genomic repeat." Mol Microbiol **19**(3): 565-74.

Amiard, S., C. Charbonnel, et al. (2010). "Distinct roles of the ATR kinase and the Mre11-Rad50-Nbs1 complex in the maintenance of chromosomal stability in Arabidopsis." Plant Cell **22**(9): 3020-33.

Anderson, C., Q. Tang, et al. (2001). "Elimination of active tad elements during the sexual phase of the *Neurospora* crassa life cycle." Fungal Genet Biol **33**(1): 49-57.

Aramayo, R. and R. L. Metzenberg (1996). "Meiotic transvection in fungi." Cell **86**(1): 103-13.

Arnaise, S., D. Zickler, et al. (2008). "Mutations in mating-type genes greatly decrease repeat-induced point mutation process in the fungus Podospora anserina." Fungal Genet Biol **45**(3): 207-20.

Assaad, F. F., K. L. Tucker, et al. (1993). "Epigenetic repeat-induced gene silencing (RIGS) in Arabidopsis." Plant Mol Biol **22**(6): 1067-85.

Attard, A., L. Gout, et al. (2005). "Truncated and RIP-degenerated copies of the LTR retrotransposon Pholy are clustered in a pericentromeric region of the Leptosphaeria maculans genome." Fungal Genet Biol **42**(1): 30-41.

Baird, N. A., P. D. Etter, et al. (2008). "Rapid SNP discovery and genetic mapping using sequenced RAD markers." PLoS One **3**(10): e3376.

Bandaru, B., M. Wyszynski, et al. (1995). "HpaII methyltransferase is mutagenic in Escherichia coli." J Bacteriol **177**(10): 2950-2.

Bao, H., H. Guo, et al. (2009). "MapView: visualization of short reads alignment on a desktop computer." Bioinformatics **25**(12): 1554-5.

Barbato, C., M. Calissano, et al. (1996). "Mild RIP-an alternative method for in vivo mutagenesis of the albino-3 gene in *Neurospora* crassa." Mol Gen Genet **252**(4): 353-61.

Barik, S. (2002). "Megaprimer PCR." Methods Mol Biol **192**: 189-96.

Barra, J. L., M. R. Mautino, et al. (1996). "A dominant negative effect of eth-1r, a mutant allele of the *Neurospora* crassa S-adenosylmethionine synthetase-encoding gene conferring resistance to the methionine toxic analogue ethionine." Genetics **144**(4): 1455-62.

Barry, C., G. Faugeron, et al. (1993). "Methylation induced premeiotically in Ascobolus: coextension with DNA repeat lengths and effect on transcript elongation." Proc Natl Acad Sci U S A **90**(10): 4557-61.

Beauchamp, P. M., E. W. Horn, et al. (1977). "Proposed involvement of an internal promoter in regulation and synthesis of mitochondrial and cytoplasmic leucyl-tRNA synthetases of *Neurospora*." Proc Natl Acad Sci U S A **74**(3): 1172-6.

Bergemann, M., O. Lespinet, et al. (2008). "Genome-wide analysis of the *Fusarium* oxysporum mimp family of MITEs and mobilization of both native and de novo created mimps." J Mol Evol **67**(6): 631-42.

Bhat, A. and D. P. Kasbekar (2001). "Escape from repeat-induced point mutation of a gene-sized duplication in *Neurospora* crassa crosses that are heterozygous for a larger chromosome segment duplication." Genetics **157**(4): 1581-90.

Bhat, A., F. K. Noubissi, et al. (2003). "Genetic analysis of wild-isolated *Neurospora* crassa strains identified as dominant suppressors of repeat-induced point mutation." Genetics **164**(3): 947-61.

Bhat, A., R. Tamuli, et al. (2004). "Genetic transformation of *Neurospora* tetrasperma, demonstration of repeat-induced point mutation (RIP) in self-crosses and a screen for recessive RIP-defective mutants." Genetics **167**(3): 1155-64.

Birkeland, S. R., N. Jin, et al. (2010). "Discovery of mutations in Saccharomyces cerevisiae by pooled linkage analysis and whole-genome sequencing." Genetics **186**(4): 1127-37.

Bleazard, W., J. M. McCaffery, et al. (1999). "The dynamin-related GTPase Dnm1 regulates mitochondrial fission in yeast." Nat Cell Biol **1**(5): 298-304.

Blumenstiel, J. P., A. C. Noll, et al. (2009). "Identification of EMS-induced mutations in Drosophila melanogaster by whole-genome sequencing." Genetics **182**(1): 25-32.

Borkovich, K. A., L. A. Alex, 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): 1-108.

Bouhouche, K., D. Zickler, et al. (2004). "Altering a gene involved in nuclear distribution increases the repeat-induced point mutation process in the fungus Podospora anserina." Genetics **167**(1): 151-9.

Bouvet, G. F., V. Jacobi, et al. (2007). "Characterization of three DNA transposons in the Dutch elm disease fungi and evidence of repeat-induced point (RIP) mutations." Fungal Genet Biol **44**(5): 430-43.

Bowring, F. J. and D. E. Catcheside (1993). "The effect of rec-2 on repeat-induced point-mutation (RIP) and recombination events that excise DNA sequence duplications at the his-3 locus in *Neurospora* crassa." Curr Genet **23**(5-6): 496-500.

Brauer, M. J., C. M. Christianson, et al. (2006). "Mapping novel traits by array-assisted bulk segregant analysis in Saccharomyces cerevisiae." Genetics **173**(3): 1813-6.

Braumann, I., M. van den Berg, et al. (2008). "Repeat induced point mutation in two asexual fungi, Aspergillus niger and Penicillium chrysogenum." Curr Genet **53**(5): 287-97.

Brookfield, J. F. (2003). "Genome sequencing: the ripping yarn of the frozen genome." Curr Biol **13**(14): R552-3.

Butler, D. K. and R. L. Metzenberg (1989). "Premeiotic change of nucleolus organizer size in *Neurospora*." Genetics **122**(4): 783-91.

Butler, D. K. and R. L. Metzenberg (1993). "Amplification of the nucleolus organizer region during the sexual phase of *Neurospora* crassa." Chromosoma **102**(8): 519-25.

Cam, H. P. (2010). "Roles of RNAi in chromatin regulation and epigenetic inheritance." Epigenomics **2**(5): 613-26.

Cambareri, E. B., B. C. Jensen, et al. (1989). "Repeat-induced G-C to A-T mutations in *Neurospora*." Science **244**(4912): 1571-5.

Cambareri, E. B., M. J. Singer, et al. (1991). "Recurrence of repeat-induced point mutation (RIP) in *Neurospora* crassa." Genetics **127**(4): 699-710.

Campbell, R. E., O. Tour, et al. (2002). "A monomeric red fluorescent protein." Proc Natl Acad Sci U S A **99**(12): 7877-82.

Cao, X. and S. E. Jacobsen (2002). "Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes." Proc Natl Acad Sci U S A **99 Suppl 4**: 16491-8.

Carr, P. D., D. Tuckwell, et al. (2010). "The transposon impala is activated by low temperatures: use of a controlled transposition system to identify genes critical for viability of Aspergillus fumigatus." Eukaryot Cell **9**(3): 438-48.

Case, M. E. (1986). "Genetical and molecular analyses of qa-2 transformants in *Neurospora* crassa." Genetics **113**(3): 569-87.

Case, M. E., H. E. Brockman, et al. (1965). "Further information on the origin of the Yale and Oak Ridge wild-type strains of *Neurospora* crassa " N. Newsl. **8**: 25-26.

Chattopadhyay, M. K., M. H. Park, et al. (2008). "Hypusine modification for growth is the major function of spermidine in Saccharomyces cerevisiae polyamine auxotrophs grown in limiting spermidine." Proc Natl Acad Sci U S A **105**(18): 6554-9.

Cheng, R., T. I. Baker, et al. (1993). "mei-3, a recombination and repair gene of *Neurospora* crassa, encodes a RecA-like protein." Mutat Res **294**(3): 223-34.

Clutterbuck, A. J. (2004). "MATE transposable elements in Aspergillus nidulans: evidence of repeat-induced point mutation." Fungal Genet Biol **41**(3): 308-16.

Clutterbuck, A. J. (2011). "Genomic evidence of repeat-induced point mutation (RIP) in filamentous ascomycetes." Fungal Genet Biol **48**(3): 306-26.

Coleman, J. J., S. D. Rounsley, et al. (2009). "The genome of Nectria haematococca: contribution of supernumerary chromosomes to gene expansion." PLoS Genet **5**(8): e1000618.

Collings, D. A., L. K. Gebbie, et al. (2008). "Arabidopsis dynamin-like protein DRP1A: a null mutant with widespread defects in endocytosis, cellulose synthesis, cytokinesis, and cell expansion." J Exp Bot **59**(2): 361-76.

Colot, H. V., G. Park, 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**(27): 10352-7.

Colot, V. and J. L. Rossignol (1995). "Isolation of the Ascobolus immersus spore color gene b2 and study in single cells of gene silencing by methylation induced premeiotically." Genetics **141**(4): 1299-314.

Connerton, I. F. (1990). "Premeiotic disruption of the *Neurospora* crassa malate synthase gene by native and divergent DNAs." Mol Gen Genet **223**(2): 319-23.

Copeland, C. S. and M. Snyder (1993). "Nuclear pore complex antigens delineate nuclear envelope dynamics in vegetative and conjugating Saccharomyces cerevisiae." Yeast **9**(3): 235-49.

Crouch, J. A., B. M. Glasheen, et al. (2008). "The evolution of transposon repeat-induced point mutation in the genome of Colletotrichum cereale: reconciling sex, recombination and homoplasy in an ''asexual" pathogen." Fungal Genet Biol **45**(3): 190-206.

Cuomo, C. A., U. Guldener, et al. (2007). "The *Fusarium* graminearum genome reveals a link between localized polymorphism and pathogen specialization." Science **317**(5843): 1400-2.

Cuperus, J. T., T. A. Montgomery, et al. (2009). "Identification of MIR390a precursor processing-defective mutants in Arabidopsis by direct genome sequencing." Proc Natl Acad Sci U S A **107**(1): 466-71.

D'Agostino, L., M. di Pietro, et al. (2006). "Nuclear aggregates of polyamines." IUBMB Life **58**(2): 75-82.

Daboussi, M. J. (1997). "Fungal transposable elements and genome evolution." Genetica **100**(1-3): 253-60.

Daboussi, M. J. and P. Capy (2003). "Transposable elements in filamentous fungi." Annu Rev Microbiol **57**: 275-99.

Daboussi, M. J., J. M. Daviere, et al. (2002). "Evolution of the Fot1 transposons in the genus *Fusarium*: discontinuous distribution and epigenetic inactivation." Mol Biol Evol **19**(4): 510-20.

Daviere, J. M., T. Langin, et al. (2001). "Potential role of transposable elements in the rapid reorganization of the *Fusarium* oxysporum genome." Fungal Genet Biol **34**(3): 177-92.

Davis, R. H. (2000). *Neurospora* : contributions of a model organism. Oxford ; New York, Oxford University Press.

Davis, R. H. (2000). *Neurospora*, contributions of a model organism. New York, Oxford University Press.

De Souza, C. P., A. H. Osmani, et al. (2004). "Partial nuclear pore complex disassembly during closed mitosis in Aspergillus nidulans." Curr Biol **14**(22): 1973-84.

Dhillon, B., J. R. Cavaletto, et al. (2010). "Accidental amplification and inactivation of a methyltransferase gene eliminates cytosine methylation in Mycosphaerella graminicola." Genetics **186**(1): 67-77.

DiGuistini, S., Y. Wang, et al. (2011). "Genome and transcriptome analyses of the mountain pine beetle-fungal symbiont Grosmannia clavigera, a lodgepole pine pathogen." Proc Natl Acad Sci U S A **108**(6): 2504-9.

Dorer, D. R. and S. Henikoff (1994). "Expansions of transgene repeats cause heterochromatin formation and gene silencing in Drosophila." Cell **77**(7): 993-1002.

Dunlap, J. C., K. A. Borkovich, et al. (2007). "Enabling a community to dissect an organism: overview of the *Neurospora* functional genomics project." Adv Genet **57**: 49-96.

Edgar, R. C. (2004). "MUSCLE: multiple sequence alignment with high accuracy and high throughput." Nucleic Acids Res **32**(5): 1792-7.

Ehrenreich, I. M., N. Torabi, et al. (2010). "Dissection of genetically complex traits with extremely large pools of yeast segregants." Nature **464**(7291): 1039-42.

Emerson, S. (1966). Mechanisms of inheritance. The Fungi: An Advanced Treatise. G. Ainsworth and A. Sussman. New York, Academic Press. **2:** 513-566.

Farah, J. A., G. Cromie, et al. (2005). "A novel recombination pathway initiated by the Mre11/Rad50/Nbs1 complex eliminates palindromes during meiosis in Schizosaccharomyces pombe." Genetics **169**(3): 1261-74.

Faugeron, G., C. Goyon, et al. (1989). "Stable allele replacement and unstable non-homologous integration events during transformation of Ascobolus immersus." Gene **76**(1): 109-19.

Faugeron, G., L. Rhounim, et al. (1990). "How does the cell count the number of ectopic copies of a gene in the premeiotic inactivation process acting in Ascobolus immersus?" Genetics **124**(3): 585-91.

Filichkin, S. A., H. D. Priest, et al. (2009). "Genome-wide mapping of alternative splicing in Arabidopsis thaliana." Genome Res **20**(1): 45-58.

Fincham, J. R., I. F. Connerton, et al. (1989). "Premeiotic disruption of duplicated and triplicated copies of the *Neurospora* crassa am (glutamate dehydrogenase) gene." Curr Genet **15**(5): 327-34.

Fishman-Lobell, J. and J. E. Haber (1992). "Removal of nonhomologous DNA ends in double-strand break recombination: the role of the yeast ultraviolet repair gene RAD1." Science **258**(5081): 480-4.

Fletcher, M. H. and A. P. Trinci (1980). "Synchronization of DNA synthesis in *Neurospora* crassa by 2'-deoxyadenosine and spore selection." Arch Microbiol **128**(1): 113-9.

Folco, H. D., M. Freitag, et al. (2003). "Histone H1 Is required for proper regulation of pyruvate decarboxylase gene expression in *Neurospora* crassa." Eukaryot Cell **2**(2): 341-50.

Foss, E. J., P. W. Garrett, et al. (1991). "Specificity of repeat-induced point mutation (RIP) in *Neurospora*: sensitivity of non-*Neurospora* sequences, a natural diverged tandem duplication, and unique DNA adjacent to a duplicated region." Genetics **127**(4): 711-7.

Foss, H. M., C. J. Roberts, et al. (1993). "Abnormal chromosome behavior in *Neurospora* mutants defective in DNA methylation." Science **262**(5140): 1737-41.

Foss, H. M. and E. U. Selker (1991). "Efficient DNA pairing in a *Neurospora* mutant defective in chromosome pairing." Mol Gen Genet **231**(1): 49-52.

Freedman, T. and P. J. Pukkila (1993). "De novo methylation of repeated sequences in Coprinus cinereus." Genetics **135**(2): 357-66.

Freitag, M., P. C. Hickey, et al. (2004). "HP1 is essential for DNA methylation in *neurospora*." Mol Cell **13**(3): 427-34.

Freitag, M., P. C. Hickey, et al. (2004). "GFP as a tool to analyze the organization, dynamics and function of nuclei and microtubules in *Neurospora* crassa." Fungal Genet Biol **41**(10): 897-910.

Freitag, M. and E. U. Selker (2005). "Expression and Visualization of Red Fluorescent Protein (RFP) in *Neurospora* crassa." Fungal Genetics Newsletter **52**: 14-17.

Freitag, M., R. L. Williams, et al. (2002). "A cytosine methyltransferase homologue is essential for repeat-induced point mutation in *Neurospora* crassa." Proc Natl Acad Sci U S A **99**(13): 8802-7.

Fudal, I., S. Ross, et al. (2009). "Repeat-induced point mutation (RIP) as an alternative mechanism of evolution toward virulence in Leptosphaeria maculans." Mol Plant Microbe Interact **22**(8): 932-41.

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

Galagan, J. E., S. E. Calvo, et al. (2005). "Sequencing of Aspergillus nidulans and comparative analysis with A. fumigatus and A. oryzae." Nature **438**(7071): 1105-15.

Galagan, J. E. and E. U. Selker (2004). "RIP: the evolutionary cost of genome defense." Trends Genet **20**(9): 417-23.

Gale, L. R., J. D. Bryant, et al. (2005). "Chromosome complement of the fungal plant pathogen *Fusarium* graminearum based on genetic and physical mapping and cytological observations." Genetics **171**(3): 985-1001.

Game, J. C., T. J. Zamb, et al. (1980). "The Role of Radiation (rad) Genes in Meiotic Recombination in Yeast." Genetics **94**(1): 51-68.

Gazumyan, A., A. Bothmer, et al. (2012). "Activation-induced cytidine deaminase in antibody diversification and chromosome translocation." Adv Cancer Res **113**: 167-90.

Gibson, D. G., L. Young, et al. (2009). "Enzymatic assembly of DNA molecules up to several hundred kilobases." Nat Methods **6**(5): 343-5.

Glass, N. L., J. Grotelueschen, et al. (1990). "*Neurospora* crassa A mating-type region." Proc Natl Acad Sci U S A **87**(13): 4912-6.

Goyon, C., C. Barry, et al. (1996). "Methylation of DNA repeats of decreasing sizes in Ascobolus immersus." Mol Cell Biol **16**(6): 3054-65.

Goyon, C. and G. Faugeron (1989). "Targeted transformation of Ascobolus immersus and de novo methylation of the resulting duplicated DNA sequences." Mol Cell Biol **9**(7): 2818-27.

Goyon, C., J. L. Rossignol, et al. (1996). "Native DNA repeats and methylation in Ascobolus." Nucleic Acids Res **24**(17): 3348-56.

Graia, F., O. Lespinet, et al. (2001). "Genome quality control: RIP (repeat-induced point mutation) comes to Podospora." Mol Microbiol **40**(3): 586-95.

Grant, D. M., A. M. Lambowitz, et al. (1984). "Transformation of *Neurospora* crassa with recombinant plasmids containing the cloned glutamate dehydrogenase (am) gene: evidence for autonomous replication of the transforming plasmid." Mol Cell Biol **4**(10): 2041-51.

Grayburn, W. S. and E. U. Selker (1989). "A natural case of RIP: degeneration of the DNA sequence in an ancestral tandem duplication." Mol Cell Biol **9**(10): 4416-21.

Grigoriev, I. V., H. Nordberg, et al. (2012). "The genome portal of the Department of Energy Joint Genome Institute." Nucleic Acids Res **40**(Database issue): D26-32.

Hamann, A., F. Feller, et al. (2000). "The degenerate DNA transposon Pat and repeat-induced point mutation (RIP) in Podospora anserina." Mol Gen Genet **263**(6): 1061-9.

Hamann, A., F. Feller, et al. (2000). "Yeti--a degenerate gypsy-like LTR retrotransposon in the filamentous ascomycete Podospora anserina." Curr Genet **38**(3): 132-40.

Hane, J. K. and R. P. Oliver (2008). "RIPCAL: a tool for alignment-based analysis of repeat-induced point mutations in fungal genomic sequences." BMC Bioinformatics **9**: 478.

Hanfrey, C., S. Sommer, et al. (2001). "Arabidopsis polyamine biosynthesis: absence of ornithine decarboxylase and the mechanism of arginine decarboxylase activity." Plant J **27**(6): 551-60.

Harkness, T. A., R. L. Metzenberg, et al. (1994). "Inactivation of the *Neurospora* crassa gene encoding the mitochondrial protein import receptor MOM19 by the technique of "sheltered RIP"." Genetics **136**(1): 107-18.

Hickman, M. A., C. A. Froyd, et al. (2011). "Reinventing heterochromatin in budding yeasts: Sir2 and the origin recognition complex take center stage." Eukaryot Cell **10**(9): 1183-92.

Honda, S. and E. U. Selker (2009). "Tools for fungal proteomics: multifunctional *neurospora* vectors for gene replacement, protein expression and protein purification." Genetics **182**(1): 11-23.

Hood, M. E., M. Katawczik, et al. (2005). "Repeat-induced point mutation and the population structure of transposable elements in Microbotryum violaceum." Genetics **170**(3): 1081-9.

Horns, F., E. Petit, et al. (2012). "Patterns of repeat-induced point mutation in transposable elements of basidiomycete fungi." Genome Biol Evol **4**(3): 240-7.

Horowitz, N. H., M. Fling, et al. (1960). "Genetic determination and enzymatic induction of tyrosinase in *Neurospora*." J. Mol. Biol. **2**: 96-104.

Hua-Van, A., J. M. Daviere, et al. (2000). "Genome organization in *Fusarium* oxysporum: clusters of class II transposons." Curr Genet **37**(5): 339-47.

Hua-Van, A., F. Hericourt, et al. (1998). "Three highly divergent subfamilies of the impala transposable element coexist in the genome of the fungus *Fusarium* oxysporum." Mol Gen Genet **259**(4): 354-62.

Idnurm, A. and B. J. Howlett (2003). "Analysis of loss of pathogenicity mutants reveals that repeat-induced point mutations can occur in the Dothideomycete Leptosphaeria maculans." Fungal Genet Biol **39**(1): 31-7.

Igarashi, K. and K. Kashiwagi (2010). "Modulation of cellular function by polyamines." Int J Biochem Cell Biol **42**(1): 39-51.

Ikeda, K., H. Nakayashiki, et al. (2002). "Repeat-induced point mutation (RIP) in Magnaporthe grisea: implications for its sexual cycle in the natural field context." Mol Microbiol **45**(5): 1355-64.

Irelan, J. T., A. T. Hagemann, et al. (1994). "High frequency repeat-induced point mutation (RIP) is not associated with efficient recombination in *Neurospora*." Genetics **138**(4): 1093-103.

Irelan, J. T. and E. U. Selker (1997). "Cytosine methylation associated with repeat-induced point mutation causes epigenetic gene silencing in *Neurospora* crassa." Genetics **146**(2): 509-23.

Irvine, D. V., D. B. Goto, et al. (2009). "Mapping epigenetic mutations in fission yeast using whole-genome next-generation sequencing." Genome Res **19**(6): 1077-83.

Jarai, G. and G. A. Marzluf (1991). "Generation of new mutants of nmr, the negative-acting nitrogen regulatory gene of *Neurospora* crassa, by repeat induced mutation." Curr Genet **20**(4): 283-8.

Jarosik, N., C. L. Bennett, et al. (2011). "Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results." The Astrophysical Journal Supplement Series **192**(14).

Jin, Y., S. Allan, et al. (2007). "Rapid genetic mapping in *Neurospora* crassa." Fungal Genet Biol **44**(6): 455-65.

Johnson, L. J., T. Giraud, et al. (2010). "The impact of genome defense on mobile elements in Microbotryum." Genetica **138**(3): 313-9.

Jones, B. A. and W. L. Fangman (1992). "Mitochondrial DNA maintenance in yeast requires a protein containing a region related to the GTP-binding domain of dynamin." Genes Dev **6**(3): 380-9.

Junhyun, J. (2012). Characterization of experimental evolution and DNA methylation in the rice blast fungus. Raleigh, North Carolina State University. **Functional Genomics:** 164.

Kafer, E. and M. Fraser (1979). "Isolation and genetic analysis of nuclease halo (nuh) mutants of *Neurospora* crassa." Mol Gen Genet **169**(2): 117-27.

Kafer, E. and E. Perlmutter (1980). "Isolation and genetic analysis of MMS-sensitive mus mutants of *neurospora*." Can J Genet Cytol **22**(4): 535-52.

Karaoglu, H., C. M. Lee, et al. (2005). "Survey of simple sequence repeats in completed fungal genomes." Mol Biol Evol **22**(3): 639-49.

Kim, T. S., J. G. Booth, et al. (2008). "Simple sequence repeats in *Neurospora* crassa: distribution, polymorphism and evolutionary inference." BMC Genomics **9**: 31.

Kinsey, J. A. (1989). "Restricted distribution of the Tad transposon in strains of *Neurospora*." Curr Genet **15**(4): 271-5.

Kinsey, J. A. (1990). "Tad, a LINE-like transposable element of *Neurospora*, can transpose between nuclei in heterokaryons." Genetics **126**(2): 317-23.

Kinsey, J. A., P. W. Garrett-Engele, et al. (1994). "The *Neurospora* transposon Tad is sensitive to repeat-induced point mutation (RIP)." Genetics **138**(3): 657-64.

Kubicek, C. P., A. Herrera-Estrella, et al. (2011). "Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of Trichoderma." Genome Biol **12**(4): R40.

Kuravi, K., S. Nagotu, et al. (2006). "Dynamin-related proteins Vps1p and Dnm1p control peroxisome abundance in Saccharomyces cerevisiae." J Cell Sci **119**(Pt 19): 3994-4001.

Kurtz, S., A. Phillippy, et al. (2004). "Versatile and open software for comparing large genomes." Genome Biol **5**(2): R12.

Labbe, J., C. Murat, et al. (2010). "Survey and analysis of simple sequence repeats in the Laccaria bicolor genome, with development of microsatellite markers." Curr Genet **57**(2): 75-88.

Lambreghts, R., M. Shi, et al. (2009). "A high-density single nucleotide polymorphism map for *Neurospora* crassa." Genetics **181**(2): 767-81.

Law, J. A. and S. E. Jacobsen (2010). "Establishing, maintaining and modifying DNA methylation patterns in plants and animals." Nat Rev Genet **11**(3): 204-20.

Lee, D. W., M. Freitag, et al. (2008). "A cytosine methyltransferase homologue is essential for sexual development in Aspergillus nidulans." PLoS One **3**(6): e2531.

Lengsfeld, B. M., A. J. Rattray, et al. (2007). "Sae2 is an endonuclease that processes hairpin DNA cooperatively with the Mre11/Rad50/Xrs2 complex." Mol Cell **28**(4): 638-51.

Lewis, Z. A., S. Honda, et al. (2009). "Relics of repeat-induced point mutation direct heterochromatin formation in *Neurospora* crassa." Genome Res **19**(3): 427-37.

Lewis, Z. A., A. L. Shiver, et al. (2007). "High-density detection of restriction-site-associated DNA markers for rapid mapping of mutated loci in *Neurospora*." Genetics **177**(2): 1163-71.

Li, H., J. Ruan, et al. (2008). "Mapping short DNA sequencing reads and calling variants using mapping quality scores." Genome Res **18**(11): 1851-8.

Li, L., S. S. Chang, et al. (2010). "RNA interference pathways in filamentous fungi." Cell Mol Life Sci **67**(22): 3849-63.

Lister, R., M. Pelizzola, et al. (2009). "Human DNA methylomes at base resolution show widespread epigenomic differences." Nature **462**(7271): 315-22.

Liu, S., H. D. Chen, et al. (2010). "High-throughput genetic mapping of mutants via quantitative single nucleotide polymorphism typing." Genetics **184**(1): 19-26.

Lyndaker, A. M. and E. Alani (2009). "A tale of tails: insights into the coordination of 3' end processing during homologous recombination." Bioessays **31**(3): 315-21.

Ma, L. J., H. C. van der Does, et al. (2010). "Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*." Nature **464**(7287): 367-73.

Malagnac, F., B. Wendel, et al. (1997). "A gene essential for de novo methylation and development in Ascobolus reveals a novel type of eukaryotic DNA methyltransferase structure." Cell **91**(2): 281-90.

Maloisel, L. and J. L. Rossignol (1998). "Suppression of crossing-over by DNA methylation in Ascobolus." Genes Dev **12**(9): 1381-9.

Marathe, S., I. F. Connerton, et al. (1990). "Duplication-induced mutation of a new *Neurospora* gene required for acetate utilization: properties of the mutant and predicted amino acid sequence of the protein product." Mol Cell Biol **10**(6): 2638-44.

Margolin, B. S., P. W. Garrett-Engele, et al. (1998). "A methylated *Neurospora* 5S rRNA pseudogene contains a transposable element inactivated by repeat-induced point mutation." Genetics **149**(4): 1787-97.

Martegani, E., M. Levi, et al. (1980). "Nuclear division cycle in *Neurospora* crassa hyphae under different growth conditions." J Bacteriol **142**(1): 268-75.

Mautino, M. R., J. L. Barra, et al. (1996). "eth-1, the *Neurospora* crassa locus encoding S-adenosylmethionine synthetase: molecular cloning, sequence analysis and in vivo overexpression." Genetics **142**(3): 789-800.

Mautino, M. R. and A. L. Rosa (1998). "Analysis of models involving enzymatic activities for the occurrence of C-->T transition mutations during repeat-induced point mutation (RIP) in *Neurospora* crassa." J Theor Biol **192**(1): 61-71.

McCluskey, K., A. E. Wiest, et al. (2011). "Rediscovery by Whole Genome Sequencing: Classical Mutations and Genome Polymorphisms in *Neurospora* crassa." G3 (Bethesda) **1**(4): 303-16.

McDougall, K. J., J. Deters, et al. (1977). "Isolation of putrescine-requiring mutants of *Neurospora* crassa." Antonie Van Leeuwenhoek **43**(2): 143-51.

Metzenberg, R. L. (1984). "A method for finding the genetic map position of cloned DNA fragments." N. Newsl.(31): 35-39.

Miao, V. P., S. F. Covert, et al. (1991). "A fungal gene for antibiotic resistance on a dispensable ("B") chromosome." Science **254**(5039): 1773-6.

Miao, V. P., M. Freitag, et al. (2000). "Short TpA-rich segments of the zeta-eta region induce DNA methylation in *Neurospora* crassa." J Mol Biol **300**(2): 249-73.

Michelmore, R. W., I. Paran, et al. (1991). "Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations." Proc Natl Acad Sci U S A **88**(21): 9828-32.

Mishra, N. C. and E. L. Tatum (1973). "Non-Mendelian inheritance of DNA-induced inositol independence in *Neurospora*." Proc Natl Acad Sci U S A **70**(12): 3875-9.

Mitchell, M. B. (1955). "ABERRANT RECOMBINATION OF PYRIDOXINE MUTANTS OF *Neurospora*." Proc Natl Acad Sci U S A **41**(4): 215-20.

Mitchell, M. B. (1955). "FURTHER EVIDENCE OF ABERRANT RECOMBINATION IN *Neurospora*." Proc Natl Acad Sci U S A **41**(11): 935-7.

Mittelsten Scheid, O., K. Afsar, et al. (1994). "Gene inactivation in Arabidopsis thaliana is not accompanied by an accumulation of repeat-induced point mutations." Mol Gen Genet **244**(3): 325-30.

Miyamoto, S., K. Kashiwagi, et al. (1993). "Estimation of polyamine distribution and polyamine stimulation of protein synthesis in Escherichia coli." Arch Biochem Biophys **300**(1): 63-8.

Montiel, M. D., H. A. Lee, et al. (2006). "Evidence of RIP (repeat-induced point mutation) in transposase sequences of Aspergillus oryzae." Fungal Genet Biol **43**(6): 439-45.

Mooren, O. L., T. I. Kotova, et al. (2009). "Dynamin2 GTPase and cortactin remodel actin filaments." J Biol Chem **284**(36): 23995-4005.

Murat, C., C. Riccioni, et al. (2010). "Distribution and localization of microsatellites in the Perigord black truffle genome and identification of new molecular markers." Fungal Genet Biol **48**(6): 592-601.

Nakayashiki, H., N. Nishimoto, et al. (1999). "Degenerate MAGGY elements in a subgroup of Pyricularia grisea: a possible example of successful capture of a genetic invader by a fungal genome." Mol Gen Genet **261**(6): 958-66.

Neafsey, D. E., B. M. Barker, et al. (2010). "Population genomic sequencing of Coccidioides fungi reveals recent hybridization and transposon control." Genome Res **20**(7): 938-46.

Nelson, M. A. and D. D. Perkins (2000). "Restriction polymorphism maps of *Neurospora* crassa: 2000 update." Fungal Genet. Newslett. **47**: 25-39.

Neuveglise, C., J. Sarfati, et al. (1996). "Afut1, a retrotransposon-like element from Aspergillus fumigatus." Nucleic Acids Res **24**(8): 1428-34.

Nielsen, M. L., T. D. Hermansen, et al. (2001). "A family of DNA repeats in Aspergillus nidulans has assimilated degenerated retrotransposons." Mol Genet Genomics **265**(5): 883-7.

Nishida, H., M. Hamamoto, et al. (2011). "Draft genome sequencing of the enigmatic yeast Saitoella complicata." J Gen Appl Microbiol **57**(4): 243-6.

Nolan, T., G. Cecere, 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**(2): 532-8.

Nothwehr, S. F., E. Conibear, et al. (1995). "Golgi and vacuolar membrane proteins reach the vacuole in vps1 mutant yeast cells via the plasma membrane." J Cell Biol **129**(1): 35-46.

Noubissi, F. K., K. Aparna, et al. (2001). "Evidence for dominant suppression of repeat-induced point mutation (RIP) in crosses with the wild-isolated *Neurospora* crassa strains Sugartown and Adiopodoume-7." J Genet **80**(2): 55-61.

Noubissi, F. K., K. McCluskey, et al. (2000). "Repeat-induced point mutation (RIP) in crosses with wild-isolated strains of *Neurospora* crassa: evidence for dominant reduction of RIP." Fungal Genet Biol **31**(2): 91-7.

Novikova, O. S., V. Fet, et al. (2007). "[Homology-dependent inactivation of LTR retrotransposons in genomes of Aspergillus fumigatus and A. nidulans]." Mol Biol (Mosk) **41**(6): 973-81.

Nowrousian, M., J. E. Stajich, et al. (2010). "De novo assembly of a 40 Mb eukaryotic genome from short sequence reads: Sordaria macrospora, a model organism for fungal morphogenesis." PLoS Genet **6**(4): e1000891.

O'Gorman, C. M., H. T. Fuller, et al. (2009). "Discovery of a sexual cycle in the opportunistic fungal pathogen Aspergillus fumigatus." Nature **457**(7228): 471-4.

Ogasawara, H., H. Obata, et al. (2009). "Crawler, a novel Tc1/mariner-type transposable element in Aspergillus oryzae transposes under stress conditions." Fungal Genet Biol **46**(6-7): 441-9.

Otsuga, D., B. R. Keegan, et al. (1998). "The dynamin-related GTPase, Dnm1p, controls mitochondrial morphology in yeast." J Cell Biol **143**(2): 333-49.

Paietta, J. V. (1990). "Molecular cloning and analysis of the scon-2 negative regulatory gene of *Neurospora* crassa." Mol Cell Biol **10**(10): 5207-14.

Park, M. H. (2006). "The post-translational synthesis of a polyamine-derived amino acid, hypusine, in the eukaryotic translation initiation factor 5A (eIF5A)." J Biochem **139**(2): 161-9.

Paulus, T. J., P. Kiyono, et al. (1982). "Polyamine-deficient *Neurospora* crassa mutants and synthesis of cadaverine." J Bacteriol **152**(1): 291-7.

Perkins, D. D. and E. G. Barry (1977). "The cytogenetics of *Neurospora*." Adv Genet **19**: 133-285.

Perkins, D. D., M. Freitag, et al. (2007). "Recurrent locus-specific mutation resulting from a cryptic ectopic insertion in *Neurospora*." Genetics **175**(2): 527-44.

Perkins, D. D., B. S. Margolin, et al. (1997). "Occurrence of repeat induced point mutation in long segmental duplications of *Neurospora*." Genetics **147**(1): 125-36.

Pomraning, K. R., K. M. Smith, et al. (2012). Library preparation for rapid Illumina genome sequencing. Methods in Molecular Biology: Fungal Secondary Metabolism. J. M. Walker, Humana Press.

Pomraning, K. R., K. M. Smith, et al. (2009). "Genome-wide high throughput analysis of DNA methylation in eukaryotes." Methods **47**(3): 142-50.

Pomraning, K. R., K. M. Smith, et al. (2011). "Bulk Segregant Analysis Followed by High-Throughput Sequencing Reveals the *Neurospora* Cell Cycle Gene, ndc-1, To Be Allelic with the Gene for Ornithine Decarboxylase, spe-1." Eukaryot Cell **10**(6): 724-33.

Praefcke, G. J. and H. T. McMahon (2004). "The dynamin superfamily: universal membrane tubulation and fission molecules?" Nat Rev Mol Cell Biol **5**(2): 133-47.

Prakash, S., L. Prakash, et al. (1980). "Effects of the RAD52 Gene on Recombination in SACCHAROMYCES CEREVISIAE." Genetics **94**(1): 31-50.

Preuss, S. and C. S. Pikaard (2007). "rRNA gene silencing and nucleolar dominance: insights into a chromosome-scale epigenetic on/off switch." Biochim Biophys Acta **1769**(5-6): 383-92.

Quevillon, E., V. Silventoinen, et al. (2005). "InterProScan: protein domains identifier." Nucleic Acids Res **33**(Web Server issue): W116-20.

Rep, M. and H. C. Kistler (2010). "The genomic organization of plant pathogenicity in *Fusarium* species." Curr Opin Plant Biol **13**(4): 420-6.

Rep, M., H. C. van der Does, et al. (2005). "Drifter, a novel, low copy hAT-like transposon in *Fusarium* oxysporum is activated during starvation." Fungal Genet Biol **42**(6): 546-53.

Rhounim, L., J. L. Rossignol, et al. (1992). "Epimutation of repeated genes in Ascobolus immersus." EMBO J **11**(12): 4451-7.

Roberts, C. J. and E. U. Selker (1995). "Mutations affecting the biosynthesis of S-adenosylmethionine cause reduction of DNA methylation in *Neurospora* crassa." Nucleic Acids Res **23**(23): 4818-26.

Robinson, J. T., H. Thorvaldsdottir, et al. (2011). "Integrative genomics viewer." Nat Biotechnol **29**(1): 24-6.

Roca, M. G., H. C. Kuo, et al. (2009). "Nuclear dynamics, mitosis, and the cytoskeleton during the early stages of colony initiation in *Neurospora* crassa." Eukaryot Cell **9**(8): 1171-83.

Rodland, K. D. and P. J. Russell (1982). "Regulation of ribosomal RNA cistron number in a strain of *Neurospora* crassa with a duplication of the nucleolus organizer region." Biochim Biophys Acta **697**(2): 162-9.

Romano, N. and G. Macino (1992). "Quelling: transient inactivation of gene expression in *Neurospora* crassa by transformation with homologous sequences." Mol Microbiol **6**(22): 3343-53.

Rosa, A. L., H. D. Folco, et al. (2004). "In vivo levels of S-adenosylmethionine modulate C:G to T:A mutations associated with repeat-induced point mutation in *Neurospora* crassa." Mutat Res **548**(1-2): 85-95.

Rothman, J. H., C. K. Raymond, et al. (1990). "A putative GTP binding protein homologous to interferon-inducible Mx proteins performs an essential function in yeast protein sorting." Cell **61**(6): 1063-74.

Rouxel, T., J. Grandaubert, et al. (2011). "Effector diversification within compartments of the Leptosphaeria maculans genome affected by Repeat-Induced Point mutations." Nat Commun **2**: 202.

Rupnik, A., N. F. Lowndes, et al. (2010). "MRN and the race to the break." Chromosoma **119**(2): 115-35.

Russell, P. J. and K. D. Rodland (1986). "Magnification of rRNA gene number in a *Neurospora* crassa strain with a partial deletion of the nucleolus organizer." Chromosoma **93**(4): 337-40.

Russo, V. E. A., T. Sommer, et al. (1985). "A modified Vogel's medium for crossing, mating-type tests and the isolation of female-sterile mutants of *Neurospora* crassa." *Neurospora* Newsletter **32**: 10-11.

Sabl, J. F. and S. Henikoff (1996). "Copy number and orientation determine the susceptibility of a gene to silencing by nearby heterochromatin in Drosophila." Genetics **142**(2): 447-58.

Sanz, L. A., S. K. Kota, et al. (2010). "Genome-wide DNA demethylation in mammals." Genome Biol **11**(3): 110.

Savarin, P., A. Barbet, et al. (2010). "A central role for polyamines in microtubule assembly in cells." Biochem J **430**(1): 151-9.

Schafer, D. A., S. A. Weed, et al. (2002). "Dynamin2 and cortactin regulate actin assembly and filament organization." Curr Biol **12**(21): 1852-7.

Selker, E. U. (1990). "Premeiotic instability of repeated sequences in *Neurospora* crassa." Annu Rev Genet **24**: 579-613.

Selker, E. U., E. B. Cambareri, et al. (1987). "Rearrangement of duplicated DNA in specialized cells of *Neurospora*." Cell **51**(5): 741-52.

Selker, E. U., M. Freitag, et al. (2002). "Induction and maintenance of nonsymmetrical DNA methylation in *Neurospora*." Proc Natl Acad Sci U S A **99 Suppl 4**: 16485-90.

Selker, E. U., D. Y. Fritz, et al. (1993). "Dense nonsymmetrical DNA methylation resulting from repeat-induced point mutation in *Neurospora*." Science **262**(5140): 1724-8.

Selker, E. U. and P. W. Garrett (1988). "DNA sequence duplications trigger gene inactivation in *Neurospora* crassa." Proc Natl Acad Sci U S A **85**(18): 6870-4.

Selker, E. U., B. C. Jensen, et al. (1987). "A portable signal causing faithful DNA methylation de novo in *Neurospora* crassa." Science **238**(4823): 48-53.

Selker, E. U., N. A. Tountas, et al. (2003). "The methylated component of the *Neurospora* crassa genome." Nature **422**(6934): 893-7.

Serna, L. and D. Stadler (1978). "Nuclear division cycle in germinating conidia of *Neurospora* crassa." J Bacteriol **136**(1): 341-51.

Sharath, A. N., E. Weinhold, et al. (2000). "Reviving a dead enzyme: cytosine deaminations promoted by an inactive DNA methyltransferase and an S-adenosylmethionine analogue." Biochemistry **39**(47): 14611-6.

Shen, J. C., W. M. Rideout, 3rd, et al. (1992). "High frequency mutagenesis by a DNA methyltransferase." Cell **71**(7): 1073-80.

Shen, J. C., J. M. Zingg, et al. (1995). "A mutant HpaII methyltransferase functions as a mutator enzyme." Nucleic Acids Res **23**(21): 4275-82.

Shiflett, A. M., J. Enkerli, et al. (2002). "Nht2, a copia LTR retrotransposon from a conditionally dispensable chromosome in Nectria haematococca." Curr Genet **41**(2): 99-106.

Shinohara, A., H. Ogawa, et al. (1992). "Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein." Cell **69**(3): 457-70.

Shinohara, A., M. Shinohara, et al. (1998). "Rad52 forms ring structures and co-operates with RPA in single-strand DNA annealing." Genes Cells **3**(3): 145-56.

Shiu, P. K., N. B. Raju, et al. (2001). "Meiotic silencing by unpaired DNA." Cell **107**(7): 905-16.

Singer, M. J., E. A. Kuzminova, et al. (1995). "Different frequencing of RIP among early vs. late ascospores of *Neurospora* crassa." Fungal Genet Newslett **42**: 74-75.

Singer, M. J., B. A. Marcotte, et al. (1995). "DNA methylation associated with repeat-induced point mutation in *Neurospora* crassa." Mol Cell Biol **15**(10): 5586-97.

Singh, P. K., S. V. Iyer, et al. (2009). "Chromosome segment duplications in *Neurospora* crassa: barren crosses beget fertile science." Bioessays **31**(2): 209-19.

Singh, P. K. and D. P. Kasbekar (2008). "Titration of repeat-induced point mutation (RIP) by chromosome segment duplications in *Neurospora* crassa." Genetica **134**(3): 267-75.

Smaczynska-de, R., II, E. G. Allwood, et al. (2010). "A role for the dynamin-like protein Vps1 during endocytosis in yeast." J Cell Sci **123**(Pt 20): 3496-506.

Smith, K., P. Phatale, et al. (2012). Epigenetics of filamentous fungi. Encyclopedia of Molecular Cell Biology and Molecular Medicine, Wiley-Blackwell.

Smith, K. M., P. A. Phatale, et al. (2011). "Heterochromatin is required for normal distribution of *Neurospora* crassa CenH3." Mol Cell Biol **31**(12): 2528-42.

Staben, C. and C. Yanofsky (1990). "*Neurospora* crassa a mating-type region." Proc Natl Acad Sci U S A **87**(13): 4917-21.

Stinchcomb, D. T., J. E. Shaw, et al. (1985). "Extrachromosomal DNA transformation of Caenorhabditis elegans." Mol Cell Biol **5**(12): 3484-96.

Sugawara, N. and J. E. Haber (1992). "Characterization of double-strand break-induced recombination: homology requirements and single-stranded DNA formation." Mol Cell Biol **12**(2): 563-75.

Sugawara, N., F. Paques, et al. (1997). "Role of Saccharomyces cerevisiae Msh2 and Msh3 repair proteins in double-strand break-induced recombination." Proc Natl Acad Sci U S A **94**(17): 9214-9.

Sugiyama, T., J. H. New, et al. (1998). "DNA annealing by RAD52 protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA." Proc Natl Acad Sci U S A **95**(11): 6049-54.

Surtees, J. A. and E. Alani (2006). "Mismatch repair factor MSH2-MSH3 binds and alters the conformation of branched DNA structures predicted to form during genetic recombination." J Mol Biol **360**(3): 523-36.

Svedruzic, Z. M. and N. O. Reich (2004). "The mechanism of target base attack in DNA cytosine carbon 5 methylation." Biochemistry **43**(36): 11460-73.

Svedruzic, Z. M. and N. O. Reich (2005). "DNA cytosine C5 methyltransferase Dnmt1: catalysis-dependent release of allosteric inhibition." Biochemistry **44**(27): 9472-85.

Tamaru, H. and E. U. Selker (2001). "A histone H3 methyltransferase controls DNA methylation in *Neurospora* crassa." Nature **414**(6861): 277-83.

Tamaru, H. and E. U. Selker (2003). "Synthesis of signals for de novo DNA methylation in *Neurospora* crassa." Mol Cell Biol **23**(7): 2379-94.

Tamuli, R. and D. P. Kasbekar (2008). "Dominant suppression of repeat-induced point mutation in *Neurospora* crassa by a variant catalytic subunit of DNA polymerase zeta." Genetics **178**(3): 1169-76.

Tamuli, R., C. Ravindran, et al. (2006). "Translesion DNA polymerases Pol zeta, Pol eta, Pol iota, Pol kappa and Rev1 are not essential for repeat-induced point mutation in *Neurospora* crassa." J Biosci **31**(5): 557-64.

Tamura, K., D. Peterson, et al. (2011). "MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods." Mol Biol Evol **28**(10): 2731-9.

Thompson, H. M., A. R. Skop, et al. (2002). "The large GTPase dynamin associates with the spindle midzone and is required for cytokinesis." Curr Biol **12**(24): 2111-7.

Thon, M. R., S. L. Martin, et al. (2004). "BAC end sequences and a physical map reveal transposable element content and clustering patterns in the genome of Magnaporthe grisea." Fungal Genet Biol **41**(7): 657-66.

Tomkinson, A. E., A. J. Bardwell, et al. (1993). "Yeast DNA repair and recombination proteins Rad1 and Rad10 constitute a single-stranded-DNA endonuclease." Nature **362**(6423): 860-2.

Uchida, M., R. R. Mourino-Perez, et al. (2008). "Microtubule dynamics and the role of molecular motors in *Neurospora* crassa." Fungal Genet Biol **45**(5): 683-92.

Van de Wouw, A. P., A. J. Cozijnsen, et al. (2010). "Evolution of linked avirulence effectors in Leptosphaeria maculans is affected by genomic environment and exposure to resistance genes in host plants." PLoS Pathog **6**(11): e1001180.

Vigfusson, N. V. and J. Weijer (1972). "Sexuality in *Neurospora* crassa. II. Genes affecting the sexual development cycle." Genet Res **19**(3): 205-11.

Vizeacoumar, F. J., W. N. Vreden, et al. (2006). "The dynamin-like protein Vps1p of the yeast Saccharomyces cerevisiae associates with peroxisomes in a Pex19p-dependent manner." J Biol Chem **281**(18): 12817-23.

Vyas, M. and D. P. Kasbekar (2005). "Collateral damage: spread of repeat-induced point mutation from a duplicated DNA sequence into an adjoining single-copy gene in *Neurospora* crassa." J Biosci **30**(1): 15-20.

Vyas, M., C. Ravindran, et al. (2006). "Chromosome segment duplications in *Neurospora* crassa and their effects on repeat-induced point mutation and meiotic silencing by unpaired DNA." Genetics **172**(3): 1511-9.

Wassenegger, M., S. Heimes, et al. (1994). "RNA-directed de novo methylation of genomic sequences in plants." Cell **76**(3): 567-76.

Watanabe, K., Y. Sakuraba, et al. (1997). "Genetic and molecular characterization of *Neurospora* crassa mus-23: a gene involved in recombinational repair." Mol Gen Genet **256**(4): 436-45.

Watanabe, S., K. Kusama-Eguchi, et al. (1991). "Estimation of polyamine binding to macromolecules and ATP in bovine lymphocytes and rat liver." J Biol Chem **266**(31): 20803-9.

Waterworth, W. M., C. Altun, et al. (2007). "NBS1 is involved in DNA repair and plays a synergistic role with ATM in mediating meiotic homologous recombination in plants." Plant J **52**(1): 41-52.

Watson, R. J., S. Burchat, et al. (2008). "A model for integration of DNA into the genome during transformation of *Fusarium* graminearum." Fungal Genet Biol **45**(10): 1348-63.

Watters, M. K., T. A. Randall, et al. (1999). "Action of repeat-induced point mutation on both strands of a duplex and on tandem duplications of various sizes in *Neurospora*." Genetics **153**(2): 705-14.

Watters, M. K. and D. R. Stadler (1995). "Spontaneous mutation during the sexual cycle of *Neurospora* crassa." Genetics **139**(1): 137-45.

Wenger, J. W., K. Schwartz, et al. (2010). "Bulk segregant analysis by high-throughput sequencing reveals a novel xylose utilization gene from Saccharomyces cerevisiae." PLoS Genet **6**(5): e1000942.

Wienke, D. C., M. L. Knetsch, et al. (1999). "Disruption of a dynamin homologue affects endocytosis, organelle morphology, and cytokinesis in Dictyostelium discoideum." Mol Biol Cell **10**(1): 225-43.

Williams, L. J., G. R. Barnett, et al. (1992). "Ornithine decarboxylase gene of *Neurospora* crassa: isolation, sequence, and polyamine-mediated regulation of its mRNA." Mol Cell Biol **12**(1): 347-59.

Wong, E. D., J. A. Wagner, et al. (2000). "The dynamin-related GTPase, Mgm1p, is an intermembrane space protein required for maintenance of fusion competent mitochondria." J Cell Biol **151**(2): 341-52.

Wortman, J. R., J. M. Gilsenan, et al. (2009). "The 2008 update of the Aspergillus nidulans genome annotation: a community effort." Fungal Genet Biol **46 Suppl 1**: S2-13.

Yarden, O. and C. Yanofsky (1991). "Chitin synthase 1 plays a major role in cell wall biogenesis in *Neurospora* crassa." Genes Dev **5**(12B): 2420-30.

Zerbino, D. R. and E. Birney (2008). "Velvet: algorithms for de novo short read assembly using de Bruijn graphs." Genome Res **18**(5): 821-9.

Zhu, J. K. (2009). "Active DNA demethylation mediated by DNA glycosylases." Annu Rev Genet **43**: 143-66.

Zingg, J. M., J. C. Shen, et al. (1998). "Enzyme-mediated cytosine deamination by the bacterial methyltransferase M.MspI." Biochem J **332 (Pt 1)**: 223-30.

Zingg, J. M., J. C. Shen, et al. (1996). "Methylation inhibitors can increase the rate of cytosine deamination by (cytosine-5)-DNA methyltransferase." Nucleic Acids Res **24**(16): 3267-75.