Epigenetics of the plant pathogen *Zymoseptoria tritici*

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem
Fachbereich Biologie
der Philipps-Universität Marburg
vorgelegt von

Klaas Schotanus
aus Leeuwarden, Niederlande

Marburg/Lahn im Juni 2014

Die Untersuchungen der vorliegenden Arbeit wurden von Anfang Oktober 2010 bis Ende April 2014 unter Betreuung von Frau Dr. Eva H. Stukenbrock in Marburg am Max-Planck-Institut für terrestrische Mikrobiologie in der Abteilung Fungal Biodiversity durchgeführt.

Vom Fachbereich Biologie
der Philipps-Universität Marburg als Dissertation
angenommen am: 18.09.2014
Erstgutachterin: Frau Dr. Eva H. Stukenbrock
Zweitgutachter: Herr Prof. Dr. Torsten Waldminghaus
Tag der mündlichen Prüfung: 10.12.2014

Declaration

I hereby declare that the dissertation entitled "Epigenetics of the plant pathogen *Zymoseptoria tritici*" submitted to the Department of Biology, Philipps-Universität Marburg, Germany, is the original and independent work carried out by me under the guidance of the PhD supervisor, and the dissertation is not formed previously on the basis of any award of Degree, Diploma or other similar titles.

_____ _____

(Date and Place) **(Klaas Schotanus)**

Summary

The genome of the fungal wheat pathogen *Zymoseptoria tritic*i consists of thirteen essential chromosomes and several so-called dispensable chromosomes. These dispensable chromosomes encode only 6% of the protein coding genes of *Z. tritici.* To date no genes involved in pathogenicity are described on the dispensable chromosomes which can be lost after meiosis or mitotic cell division without any apparent effect on fitness.

To investigate the underlying molecular basis of instability of the dispensable chromosomes, the epigenetic components of both the essential and dispensable chromosomes were characterized here. Chromatin immunoprecipitation and sequencing of DNA associated with the centromere specific histone (CenH3) was conducted to identify the centromeres of *Z. tritici.* It was shown that the centromeres of *Z. tritici* are small, sequence independent and lack any conserved motif. The centromeres are AT-rich, but not located in the most abundant AT-rich region of the chromosomes, and the centromeric organization is similar for both essential and dispensable chromosomes. To study centromere dynamic, parental and progeny strains derived from a meiotic cross were included in the study. The centromeres of these strains were shown to be conserved among *Z. tritici* strains. The deletion of the centromere of the dispensable chromosome 14 resulted in several strains were chromosome 14 was completely lost, while only a single strain was identified with a neocentromere on chromosome 14.

The chromatin content of both types of chromosomes was also investigated. Three histone modifications specific for either euchromatin or heterochromatin were characterized. The essential chromosomes are enriched with euchromatin while the dispensable chromosomes are mainly heterochromatic. Several repeat rich regions with low gene density were also enriched with heterochromatin on the essential chromosomes. One particularly large region of 780 kb of the essential chromosome 7 was in addition found to be enriched with facultative heterochromatin. Genes in this region are silenced both during axenic and infectious growth. Based on the obtained results, it can be concluded that the difference between the essential and dispensable chromosomes cannot be associated with the centromeres. However, differences in the chromatin states is a main difference between the two types of chromosomes.

To investigate the hemibiotrophic lifestyle switch in *Z. tritici* the epigenetic component of infectious growth was studied with a focus on RNA interference (RNAi). Five mutant strains of several proteins involved in the RNAi pathway were created. It could be demonstrated that Dicer and Argonaute genes play a role during the formation of asexual fruiting bodies called pycnidia. In contrast to the Dicer gene, the Argonaute genes show an unusual degree of sequence variation among *Z. tritici* strains. Collectively, the work presented here underlines the importance of epigenetics in both genome stability as well as pathogenicity in the fungal pathogen *Z. tritici*.

Zusammenfassung

Das Genom des hemibiotrophen weizenpathogenen Pilzes *Zymoseptoria tritici* besteht aus dreizehn essentiellen und mehreren sogenannten akzessorischen Chromosomen. Diese akzessorischen Chromosomen kodieren lediglich für 6% der Gene von *Z. tritici* und können nach einer Meiose oder einer Mitose ohne einen Einfluss auf Wachstum, Paarung und Pathogenität verloren gehen.

Zur Untersuchung der molekularen Instabilität der akzessorischen Chromosomen, wurden in dieser Arbeit die epigenetischen Komponenten der essentiellen und akzessorischen Chromosomen charakterisiert. Eine Immunopräzipitation des Chromatins sowie eine Sequenzierung der DNA, welche mit den Zentromer spezifischen Histonen (CenH3) assoziiert ist, wurden durchgeführt um die Zentromere in *Z. tritici* zu identifizieren. Es konnte gezeigt werden, dass die Zentromere von *Z. tritici* vergleichsweise klein und sequenzunabhängig sind sowie keine konservierten Motive beinhalten. Sie sind AT-reich aber befinden sich dennoch nicht in den AT-reichsten Regionen der Chromosomen und es konnte kein Unterschied zwischen den Zentromeren der essentiellen und akzessorischen Chromosomen festgestellt werden. Um die Dynamik der Zentromere zu untersuchen, wurden Nachkommen einer meiotischen Kreuzung sowie die parentalen Stämme benutzt, wobei gezeigt werden konnte, dass die Zentromere in allen verwendeten *Z. tritici* Stämmen konserviert sind. Eine Zentromer-Deletion des akzessorischen Chromosoms 14 resultierte in mehreren Stämmen, die dieses Chromosom verloren haben während ein einzelner Stamm identifiziert wurde, der ein Neozentromer gebildet hat.

Des Weiteren wurde der Chromatingehalt von beiden Chromosomentypen untersucht. Dazu wurden drei für Euchromatin oder Heterochromatin spezifische Histon-Modifikationen charakterisiert. Es wurde gezeigt, dass die essentiellen Chromosomen mit Euchromatin angereichert sind, während die akzessorischen Chromosomen überwiegend heterochromatisch sind. Auf den essentiellen Chromosomen gibt es aber auch repetitive Regionen mit geringer Gendichte, die reich an Heterochromatin sind. Eine besonders große 780 kbp Region auf Chromosom 7 ist reich an fakultativem Heterochromatin. Gene in dieser Region sind sowohl in axenischer Kultur als auch während der Infektion der Wirtspflanze stillgelegt. Auf diesen Ergebnissen basierend kann festgestellt werden, dass der Unterschied zwischen essentiellen und akzessorischen Chromosomen nicht mit den Zentromeren in Verbindung steht. Die beiden Chromosomentypen unterscheiden sich aber stark in ihrem Chromatingehalt.

Der Einfluss der Epigenetik auf den Übergang des Pilzes von der biotrophen zur nekrotrophen Phase während der Pflanzeninfektion wurde ebenfalls untersucht. Der Fokus lag dabei auf der RNA Interferenz (RNAi). Fünf verschiedene Proteine des RNAi Systems wurden deletiert und es konnte gezeigt werden, dass sowohl das Dicer als auch die Argonaut-Gene in der Produktion der asexuellen Fruchtkörper (Pyknidien) involviert sind. Im Gegensatz zu dem Dicer kodierenden Gen, zeigen die Argonaut-Gene eine ungewöhnlich hohe Sequenz Variation innerhalb der verwendeten *Z. tritici* Stämme. Die hier präsentierte Arbeit unterstreicht damit die Relevanz der Epigenetik bei der Genom-Stabilität als auch der Pathogenität von *Z. tritici.*

Abbreviations

Δ	Deletion
5'U	5' Uridine
AFLP	Amplified fragment length polymorphism
Ago	Argonaute
AT-content	Adenine-thymine content
Blast	Basic local alignment search tool
bp	Base pair
cDNA	Complementary DNA
cen	Centromere
CenH3	Centromere specific histone H3
ChIP	Chromatin immunoprecipitation
C-terminal	Carboxyl-terminal
Dcl	Dicer-like
DNA	Deoxyribonucleic acid
dpi	Days post infections
Fig	Figure
GFP	Green fluorescence protein
H3K4me2	Di-methylation of Lysine 4 of histone 3
H3K9me3	Tri-methylation of Lysine 9 of histone 3
H3K27me3	Tri-methylation of Lysine 27 of histone 3
IPO323	*Zymoseptoria tritici* strain IPO323
kb	Kilo basepairs
Mb	Mega basepairs
min	Minutes
Mnase	Micrococcal nuclease
mRNA	Messenger RNA
nt	Nucleotide
N-terminal	Start of a protein sequence
ORF	Open reading frame
PAMP	Pathogen-associated molecular pattern
PCR	Polymerase chain reaction
qPCR	Quantitative real-time PCR
RNA	Ribonucleic acid
RNAi	RNA interference

Table of Contents

Chapter 1: Introduction

Chapter 1.1: *Zymoseptoria tritici*

Zymoseptoria tritici (synonym *Mycosphaerella graminicola*) is a plant pathogenic fungus which causes septoria tritici blotch on wheat (*Triticum aestivum*) (Goodwin et al., 2011). *Z. tritici* is a hemibiotropich pathogen (Goodwin et al., 2004). After penetration through the plant stomata, the pathogen grows intercellularly and symptom-less (Duncan and Howard, 2000). After 10 to 12 days the pathogen undergoes a drastic life style switch resulting in necrosis and chlorosis and the formation of asexual fruiting bodies, so-called pycnidia (Duncan and Howard, 2000). Approximately 28 days after the initial infection, the infected region is completely necrotic and the fruiting bodies mature (Duncan and Howard, 2000). Sexual reproduction occurs on plant debris between strains of opposite mating types resulting in the production of pseudothecia and ascospores (Ponomarenko et al., 2011). Ascospores can be spread by the wind and serve as the primary source of inoculum. Secondary infections are initiated by the pycnidiospores, that are spread by rain splash between neighboring plants (Ponomarenko et al., 2011).

Population genetics and comparative genomics have shown that *Z. tritici* originates from the Middle East and that the fungus co-evolved with its host (Stukenbrock et al., 2007). The closely related species *Z. pseudotritici* and *Z. ardabiliae* still occur on wild grass species in the center of origin (Stukenbrock et al., 2007). The divergence of the species occurred recently, and comparative genomics showed that selection mainly occurred on putative secreted proteins, which are the result from host specialization.

The genome structure of *Z. tritici* isolates can be highly variable (McDonald and Martinez, 1991) & (Mehrabi et al., 2007). Electrophoretic separation of chromosomes using pulsed-field gel electrophoresis have shown that there is a high rate of chromosome length polymorphism as well presence-absence variation of chromosomes between isolates. The 13 largest chromosomes are the so-called essential chromosomes present in all isolates. Variation in the length of these chromosomes has also been observed (McDonald and Martinez, 1991) & (Mehrabi et al., 2007). In addition to the 13 essential chromosomes, *Z. tritici* has a set of dispensable chromosomes (Goodwin et al., 2011). The number and length of these dispensable chromosomes varies between isolates (McDonald and Martinez, 1991) & (Mehrabi et al., 2007). In general, the dispensable chromosomes are the smallest chromosomes and can be lost without any apparent effect on fitness (Wittenberg et al., 2009) & (Goodwin et al., 2011) & (Croll et al., 2013). To date no genes involved in pathogenicity have been identified on the dispensable chromosomes (Kellner et al., 2014) & (Goodwin et al., 2011). Nevertheless, the dispensable chromosomes most likely contain genes with

a beneficial effect on fitness under some conditions. Such "conditional" advantageous effects could drive the presence-absence variation of chromosomes through balancing selection (Croll and McDonald, 2012). Comparative genomics between *Z. tritici* and the closely related species *Z. pseudotritici, Z. ardabiliae* and *Z. brevis* have revealed that dispensable chromosomes are also present in other *Zymoseptoria* species indicating that the origin of the dispensable chromosome precedes the speciation of *Z. tritici* (Stukenbrock et al., 2011). The reference strain IPO323 has eight dispensable chromosomes, the largest number of dispensable chromosomes known in the *Zymoseptoria* clade. IPO323 has a genome of ~40 Mb and the 21 chromosomes ranging from 6 to 0.4 Mb (Goodwin et al., 2011). The genome of IPO323 is almost completely sequenced from telomere to telomere, except for one telomere of chromosome 21 and small gaps of 1.4 and 4.5 kb at chromosome 18 (Goodwin et al., 2011). No differences in telomere sequence were described between the essential and dispensable chromosomes. Compared to the essential chromosomes, the dispensable chromosomes are shorter in length, have a lower gene density (mean: 6%) and have shorter genes (Mean 1.3 kb and 0.8 kb respectively).

Croll et al showed that the average exon length of genes located on the essential chromosomes (517 bp) is almost twice as long as the average exon length of genes of the dispensable chromosomes (314 bp) (Croll et al., 2013). Other features of the dispensable chromosomes are, a higher AT content and an enrichment in repetitive elements compared to the essential chromosomes (Goodwin et al., 2011). Van der Burgt et al however showed that pseudogenes are randomly distributed in the genome and not enriched on the dispensable chromosomes (van der Burgt et al., 2014). Kellner et al, showed that the gene expression is lacking at the dispensable chromosomes (Kellner et al., 2014).

Chapter 1.2: Dispensable chromosomes

Like *Z. tritici*, several other fungal species have dispensable chromosomes in their genome, in addition to an "essential" chromosome set (Covert, 1998). The function or origin of these dispensable chromosomes are in many cases not clear. The dispensable chromosomes are defined as B-chromosomes in plants, insects and animals (Houben and Carchilan, 2012). The nomenclature to describe dispensable chromosomes in fungi is diffuse and several terms have been used: mini chromosomes (in yeast) and supernumerary chromosomes (in *Nectria haematococca*), dispensable chromosomes (in *Z. tritici*) or lineage specific chromosomes (in *Fusarium oxysporum*) in plant pathogens (Poláková et al., 2009) & (Coleman et al., 2009) & (Goodwin et al., 2011) & (Ma et al., 2010). All of these terms, describe small, non-essential chromosomes. In this thesis the term "dispensable" will be used.

In spite of the presence of dispensable chromosomes in fungi, it is not known how these dispensable chromosomes originate and why they are maintained or frequently lost (Covert, 1998) & (Croll and McDonald, 2012). Chromosomes obtain structural stability by the presences of functional centromeres and telomeres. Instability of centromeres on the dispensable chromosomes could result in improper chromosome segregation during meiosis. Information about the structure, molecular and epigenetically properties of dispensable chromosomes in fungi are lacking and until now not well studied. Insight into these would however allow us to understand better their functional role and the molecular basis of "dispensability".

The dispensable chromosomes (chromosome 14, 15 and 17) of the plant pathogen *Nectria haematococca* mating population VI (MPVI*)* (asexual name *Fusarium solani*) are inherited in non-Mendelian manner and can be lost without any apparent effect on fitness when grown axenically in media (Coleman et al., 2009). On axenic media a small amount of genes of the dispensable chromosomes are expressed (Coleman et al., 2009). Unlike in *Z. tritici*, the genes of the dispensable chromosomes of *N. haematococca* MPVI are coding for pathogenicity, antibiotic resistance, and have functions in metabolism (Coleman et al., 2009). Coleman et al furthermore hypothesized that dispensable chromosomes promote chromosome plasticity in *N. haematococca* MPVI to promote rapid adaptive evolution of pathogens allowing them to explore a variety of environmental niches and host genotypes (Coleman et al., 2009). The dispensable chromosomes of *N. haematococca* MPVI are up to 1.57 Mb in length, which makes them larger than the dispensable chromosome of *Z. tritici* (Goodwin et al., 2011) &. (Coleman et al., 2009) However, they share other common features with the dispensable chromosomes of *Z. tritici* such as an enrichment of repetitive DNA, a

higher AT content, shorter genes and a lower gene density compared to the essential chromosomes (Coleman et al., 2009).

Dispensable chromosomes which provide pathogenicity to a specific host are also found in the *Fusarium* species complex. One species (*Fusarium oxysporum*) has four dispensable chromosomes (Chromosome 3, 6, 14 and 15), that are enriched with transposon sequences (74% of the total length) and lack any house-keeping genes (Ma et al., 2010) & (Schmidt et al., 2013).

Expression studies of *F. oxysporum* in planta have shown that all known pathogenicity related genes are located on chromosome 14, one of the dispensable chromosomes. The relevance of these genes was confirmed with in-planta experiments using strains without the dispensable chromosomes. In absence of chromosome 14 the strains are non-pathogenic on tomato plants (Ma et al., 2010). With chromosome 14 they can however infect the tomato plants. Interestingly, the expression of a specific effector protein class is still regulated by a transcription factor located on an essential chromosome, suggesting that the chromosomes contain sequences which have co-evolved.

Leptosphaeria maculans has one dispensable chromosome of 0.73 Mb (Balesdent et al., 2013). Like the other fungal dispensable chromosomes, the dispensable chromosome of *L. maculans* is enriched with transposable elements, has a higher AT content in comparison to the essential chromosomes, and only has a few genes (Rouxel et al., 2011). Besides 36 putative genes, *Leptosphaeria maculans* has one confirmed effector gene on the dispensable chromosome (Balesdent et al., 2013). This effector gene, *AvrLm11*, is small, has no homology to proteins in any database and is up-regulated during infection. The gene was transformed into an avirulent strain and this resulted in a virulent strain. The authors conclude that the gene is essential for the infection of particular *Brassiceae* cultivars.

Phylogenetic sequence analysis of the essential and dispensable chromosomes of *F. oxysporum* with other *Fusarium* species suggested that the dispensable chromosome 14 originate from a *Fusarium* species (Ma et al., 2010).

In *N. haematococca* MPVI a small part of dispensable chromosome 14, shows homology to three short regions of 50 kb of the essential chromosome 6 (Coleman et al., 2009). The remaining sequences of the dispensable chromosome lack any similarity to the essential chromosomes. Comparative genomics showed that genes on the essential chromosomes of *N. haematococca* MPVI are conserved when compared to the homologous genes in *F. graminearum* (Coleman et al., 2009). However, the dispensable chromosomes lack homologs in the *F. graminearum* genome (Coleman et al., 2009). Interestingly, the genes of the dispensable chromosome 14 have homology to genes of *Aspergillus* species, indicating that the dispensable chromosomes could be acquired from horizontal

gene or chromosome transfer (Coleman et al., 2009). One report showed that a dispensable chromosome was retained in the genome of *N. haematococca* MPVI after 2,5 years of cell transfer on axenic media suggesting mitotic stability (VanEtten et al., 1998).

Wittenberg and co-workers performed a crossing experiment between two *Z. tritici* strains isolated from bread wheat to test inheritance of dispensable chromosomes. The crossing experiment resulted in 68 isolated progenies, from which 8 isolates were twins. One parent strain lacked chromosome 20 and as result of lack of homologues recombination the chromosome was absent in a large proportion of the progeny strains (Wittenberg et al., 2009). AFLP and PCR analysis showed that besides chromosome 20, also chromosome 15, 17, 18 and 21 were frequently lost in the progeny population. Not only loss of chromosomes was demonstrated, also disomy was observed. An additional cross between a bread wheat isolate and a durum wheat isolate was performed and this resulted in 148 progeny isolates (23 twins). Here, both parents had 21 chromosomes. The authors showed that some of the progeny lack one of more of the dispensable chromosomes.

Similar events in *Z. tritici* have been been shown by Croll et al, which identified several chromosome length polymorphism by analyzing 98 isolates collected from four regional populations (Croll et al., 2013). Several large AT rich regions of the dispensable chromosomes varying from 100-700 kb, flanked by repeat rich DNA, were lacking in one or more isolates. This suggests that the dispensable chromosomes are also variable in terms of structure. The authors speculate that this is the result of chromosome breakage. In addition to the absent of whole regions, the loss of whole dispensable chromosomes were also observed.

Chromosome length polymorphism were not only found in the dispensable chromosome, also chromosome 10 and 13 had regions which were absent in some of the isolates. In contrast to what has been observed for the dispensable chromosomes, the lack of a complete essential chromosome has never been shown.

Croll et al performed several crosses with strains isolated in Switzerland (Croll et al., 2013). As shown in Wittenberg et al, the loss of several dispensable chromosomes was observed (Wittenberg et al., 2009). Chromosome 19 showed a variation in length of 300 kb, while chromosome 15 and 21 always had the same length. One cross, wherein one parent lacked chromosome 17, resulted in progeny strains for which chromosome 17 was almost twice as large than in the parent strain. Interestingly, the comparison of parent and progeny strains showed a region of 77 kb added to the end of the chromosome. The authors showed that the extension of the specific chromosome was the product of a partial duplication event, most likely resulting from of a chromosome breakage and fusion mechanism (Croll et al., 2013).

In the pathogen *Cochliobolus carbonum* a dispensable region of DNA is part of an essential chromosome, thereby a partially dispensable chromosome (Ahn and Walton, 1996). The authors hypothesize that a 1.4-1.5 Mb dispensable chromosome has been fused with a large essential chromosome resulting in a 3.5 Mb chromosome. Similarly in *F. oxysporum,* two regions at the end of chromosome 1 and 2 lack synteny to the genome of the closely related *Fusarium verticillioides* (Ma et al., 2010). As for the dispensable chromosomes these regions are enriched for transposable elements (Ma et al., 2010). This suggests that a fragment of a dispensable chromosome was fused to an essential chromosome in the past.

The chromosomal organization of the human pathogen *Candida glabrata* is also very dynamic, however the fungus is clonal (Poláková et al., 2009). After infection, the chromosome content can rearrange within a few days and this can lead to populations of two or three different karyotypes in one patient. The formation of novel chromosomes was also described. De novo chromosomes were formed by a segment duplication including the centromeres. The new chromosomes were 120-200 kb in length and had telomeres. In addition to the de novo chromosome formation, also chromosome fusions were observed in *C. glabrata*. One fused chromosome was formed after the fusion of two complete chromosomes. The other fused chromosome consisted of two large segments from two chromosomes. From one of the original chromosomes a segment including the centromere was deleted and the remaining segment was fused to another chromosome (Poláková et al., 2009). It is clear that dispensable chromosomes in fungi can be highly dynamic and undergo considerable structural changes (Covert, 1998).

Several mechanisms have been proposed to describe chromosome rearrangements which result in the formation of small chromosome parts that could evolve to chromosomes. Important mechanisms include:

* Formation of new chromosomes. New chromosomes can occur by the paring of homologous (repetitive) regions of either a sister chromosome or from the same chromosome. This can lead to a circular chromosome or leads to the formation of a new chromosome which can become stable after telomere and neocentromere formation (Zhang et al., 2008).

* Deletion of chromosomal segments. Recombination between two chromosomes which differ in length can result in the loss of a chromosome segment and lead to variation in chromosome lengths. Similar events occur when homologues regions of two non-homologous chromosomes recombine leading to loss of chromosomal segments.

* Chromosome breakage near the centromeres of two non-homologous chromosomes can result in a new chromosome through the fusion of telomeres (Han et al., 2006).

Another type of chromosomal fusion is a Robertsonian translocation, formed between two acrocentric chromosomes (Jones, 1998). On acrocentric chromosomes, the centromeres are located close to the telomere and the region between the centromere and the telomere contain few or no genes. In the case of a Robertsonian translocation, the short arms (p-arm) of two acrocentric chromosomes breaks, both long arms (q-arms) fuse, which results in the forming of one large chromosome. Both the p-arms are lost, this is due to the lack of genes on these chromosome segments.

It is hypothesized that the B-chromosomes in plants are derived from A-chromosomes. Evidence that B-chromosomes originate from A-chromosomes was shown by a whole genome comparison of *Z. mays* individuals. Lamb et al hypothesized that large DNA fragments, flanked by transposons were transferred from A-chromosomes to the B-chromosomes (Lamb et al., 2005). Another study has described examples where the B-chromosomes have been formed after missegragation of chromosomes after interspecific mating resulting in chromosomal rearrangements (Riddle and Birchler, 2003). Zhang and co-workers showed that the crossing of the two grasses *Elymus trachycaulus* and *Triticum aestivum* resulted in the formation of a hybrid essential chromosome which contained chromosomal fragments of both species (Zhang et al., 2008).

Dhar and co-workers showed the “birth” of a new B-chromosome in *Plantago lagopus* (Dhar et al., 2002). The B-chromosome arose in accession with a trisomic chromosome. The new chromosome underwent several chromosomal rearrangements. The new chromosome was shown to have a functional centromere, as well as telomeres and to be enriched with heterochromatin. These characteristics are typical for B-chromosome formation (Dhar et al., 2002).

As discussed above dispensable chromosomes exist in many species and can originate through various mechanisms. Due to the presence of virulence factors on the dispensable chromosome of fungi like *F. oxysporum, N. haematococca* MPVI and *L. maculans* dispensable chromosomes can be important for the fungus-host interaction (Ma et al., 2010) & (Coleman et al., 2009) & (Rouxel et al., 2011). The dispensable chromosomes of *Z. tritici* encode 6% of the total genes (Goodwin et al., 2011). To date, genes involved in pathogenicity have not been described on these dispensable chromosomes (Kellner et al., 2014). The lack of gene transcription, higher frequency of transposons and the frequent rearrangement suggest that the chromosomal organization of the dispensable chromosomes differs from the essential chromosomes in *Z. tritici*. One of the factors which influence chromosome organization as well as transcription are histone modifications (Campos and Reinberg, 2009).

Chapter 1.3: Histone modifications

Eukaryotic chromosomes are composed of DNA which is organized and structured by a protein complex to form chromatin (Campos and Reinberg, 2009). Chromatin consists of DNA wrapped around nucleosomes. The nucleosome is composed of an octamer of histones formed by two histone (H2A-H2B) dimers and a histone $(H3\text{-}H4)_2$ tetramer (Figure 1.3.1). Approximately 150 bp of DNA is wrapped around each nucleosome. An additional linker histone (H1) binds to the internucleusomal DNA (linker DNA) which is located between two nucleosomes.

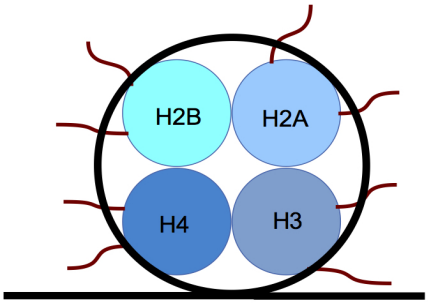

Figure 1.3.1: DNA is wrapped around a histone complex. The DNA (in black) is wrapped around an octamer of histones (H2B, H2A, H3 and H4). The histone tails are shown in red.

The binding between the nucleosome and the DNA can change. The "degree" of binding influences the gene transcription. A more relaxed chromatin state makes the DNA accessible to transcription factors. This type of chromosome structure is known as euchromatin. Opposed to euchromatin, heterochromatin consists of densely wrapped DNA. This prevents the binding of transcription factors and the transcription of genes.

A change in the chromatin state can result from post-transcriptional modifications of histone tails (Campos and Reinberg, 2009). In contrast to the histone protein complexes, the histone tails are unstructured and almost all of the polar residues and some prolines can be subject of post-transcriptional modifications. Examples of such histone modifications are acetylation, methylation, phosphorylation, ubiquitylation, sumoylation, ADP ribosylation, deimination and proline isomerization (Figure 1.3.2) (Bártová et al., 2008). All of these modifications can affect gene transcription.

In euchromatic regions, the histones are acetylated or methylated (or both) (Schübeler et al., 2004). Both the acetyl and methyl groups act like docking sites for respectively bromodomains and chromodomains which affect transcription (Rando and Winston, 2012). In addition to the docking function, the negatively charged acetyl group results in a change of the net charge of the histone and influence the chromatin interactions which again influence transcription.

Mono-methylation of the lysine at the 4th position of the tail of histone 3 (H3K4me) and mono-methylation of lysine 9 of the tail of histone 3 (H3K9me), lysine 27 of the tail of histone 3 (H3K27me), lysine 79 of the tail of histone 3 (H3K79me) and lysine 20 of the tail of histone 4 (H4K20me) are hallmarks for euchromatin, this is also true for di-methylation of lysine 79 of the tail of histone 3 (H3K79me2) and tri-methylation of lysine 4 of the tail of histone 3 (H3K4me3) and H3K36. Finally acetylation of lysine 9 of the tail of histone 3 (H3K9ac), H3K14ac and H3K27ac are also found in euchromatic regions (Bannister and Kouzarides, 2011) & (Freitag, 2014).

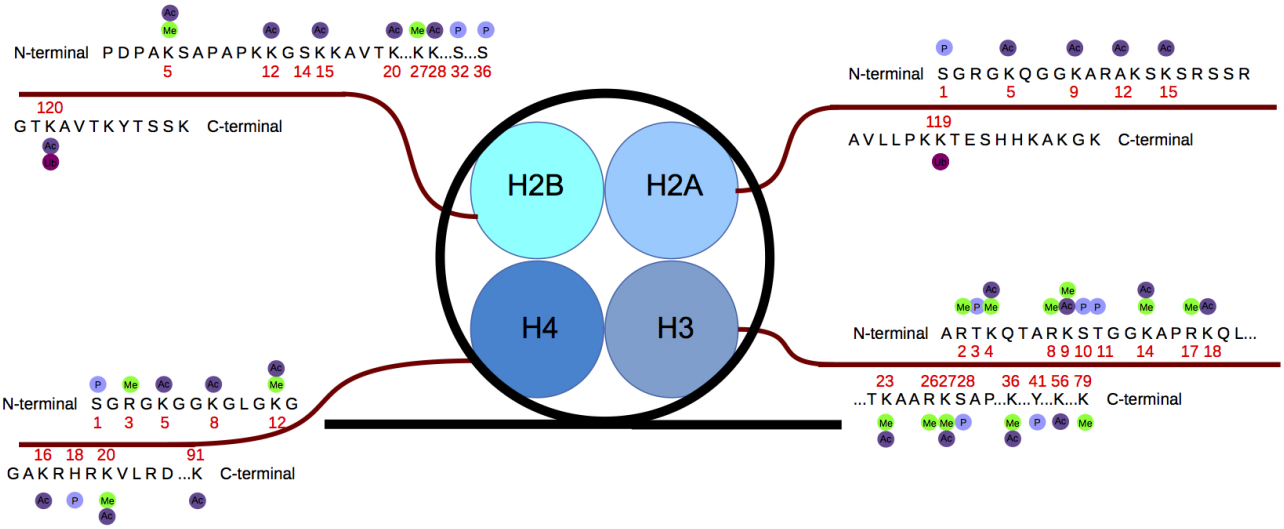

Figure 1.3.2: Histone modifications at the tails of histone proteins. The four histone protein tails are shown in red. The amino-acids are numbered from the N-terminal to the C-terminal side and gaps are shown with dots. Amino-acids which of the subject to modifications are numbered in red and the specific modifications are shown in circles on top or below the amino-acid. (Figure is modified from: (Rodríguez-Paredes and Esteller, 2011)).

Heterochromatin is enriched in gene poor and repetitive regions (Peng and Karpen, 2008)& (Luijsterburg et al., 2009) & (Dinant and Luijsterburg, 2009). In most of the species studied, constitutive heterochromatin is found around centromeres and telomeres (Smith et al., 2011) & (Stimpson and Sullivan, 2010) & (Smith et al., 2012). The typical histone modifications associated to heterochromatin are tri-methylation on lysine 9 of histone H3 (H3K9me3) and H4K20me3(Grewal and Jia, 2007). The lysine methyltransferase KMT1 is required for tri-methylation of H3K9 and H3K9me3 is maintained by the heterochromatic protein 1 (HP1) (Campos and Reinberg, 2009). An important function of heterochromatin is to reduce random recombination by "silencing" of repetitive sequences and transposons (Grewal and Jia, 2007). Not only coding sequences are suppressed, it has also been shown that non-coding RNA, which originate from heterochromatic regions, are suppressed (Gullerova and Proudfoot, 2008) & (Zofall

et al., 2009).

Heterochromatin is also able to spread in the genome and influence the transcription of adjacent regions. In chicken it was shown that spreading of heterochromatin is disrupted by a DNA binding protein that associates with short euchromatic regions surrounding heterochromatin (Gaszner and Felsenfeld, 2006). Also, in *S. cerevisiae* and in *S. pombe* similar phenomena were described.

In *N. crassa* Lewis et al showed that HP1 co-localizes with Dim2 (DNA methyltransferase) and that there is a correlation between DNA methylation and heterochromatin (Lewis et al., 2009). The heterochromatin spreading in *N. crassa* is controlled by a DNA methylation modulator-2 (DMM-2), proposed to bind with heterochromatin protein 1(HP1) to H3K9me3 at sequences which are the subject of a genome defence mechanism, resulting in the so-called repeat induced point mutations (RIP). During meiosis, homologous duplicated sequences undergo C:G to T:A mutations into both copies of the duplicated sequences (Galagan and Selker, 2004). Deletion of DMM-2 (and also DMM-1) resulted in heterochromatin spreading in euchromatic neighboring regions (Honda et al., 2010).

Facultative heterochromatin is found in regions where gene expression is silenced under some conditions while under other conditions gene expression is induced (Palmer and Keller, 2010). Facultative heterochromatin is characterized by di- and tri-mehtylation of H3K27 (Campos and Reinberg, 2009). Tri-methylation of H3K27 occurs by the lysine methyltransferase KMT6 (Campos and Reinberg, 2009). Facultative heterochromatic regions are present at the inactive X-chromosome of humans and Hox genes. In fungi, many secondary metabolite clusters are found within facultative heterochromatin regions (Palmer and Keller, 2010). When grown on axenic media, H3K27me3 affects one third of the genome of *F. graminearum* and most of genes located in these regions lack transcription (Connolly et al., 2013). In a comparative study of *F. graminearum* and *N. haematococca MPVI* it was found that the evolutionary conserved sequences in these two pathogens are enriched with H3K4me2. While species specific sequences in *F. graminearum* are enriched with H3K27me3 (Connolly et al., 2013). In contrast to *F. graminearum,* H3K27me3 covers only 6.8% of the whole genome of *N. crassa.* H3K27me3 is enriched in regions near the telomeres and, as in *F. graminearum,* also in regions with species specific genes (Jamieson et al., 2013). Interestingly, not all fungi have H3K27me3, prominent examples of fungi which lack this histone mark are *S. cerevisiae* and *S. pombe* (Connolly et al., 2013).

The expression of a gene depends on an array of histone modifications. A single histone modification will not influence whether a gene is transcribed or not. In a study where Wang and colleagues compared 12.541 gene promoters they showed that a minimum of 22 histone

modifications at the promoter regions were required for gene transcription. However all the 22 histone modifications are hallmarks for euchromatin, variation within the histone modifications is possible and it is not necessary to have the same specific histone marks for efficient gene transcription (Wang et al., 2008) & (Barski et al., 2007).

One third of the B-chromosome of *Zea mays* is euchromatic although the DNA is non-coding(Yin et al., 2011). Interestingly, in the middle of the euchromatic regions is the B-chromosome centromere repeat ZmB. The other two third of the B-chromosome is heterochromatic and contains repetitive elements, the centromere and telomeres. This is similar to the chromatin organization of the B-chromosomes of another plant species *Brachycome dichromosomatica*. Interestingly, the levels of heterochromatin of these B-chromosomes are comparable to the A-chromosomes, suggesting that the B-chromosomes are not enriched for heterochromatin although less transcribed (Marschner et al., 2007).

Protein-DNA interactions such as histone modifications can be studied by chromatin immunoprecipitation (ChIP) (Collas, 2010). The protein of interest is immunoprecipated and the associated DNA can be identified by several techniques: the most common used methods are PCR, qPCR, micro-arrays or sequencing of the DNA by next generation sequencing techniques. The immunoprecipitation consists of several steps. The first step is the cross-linking of the DNA and proteins by formaldehyde. The second step is fragmentation of the protein bound DNA by sonication or micrococcal nuclease (Mnase) treatment. This is followed by immunoprecipitation with an antibody that can either directly target the protein of interest or target a tag which is fused to the DNA binding protein. This can be followed by a secondary antibody step if necessary. After the immunoprecipitation, the cross-link between DNA and the protein can be decross-linked and subsequently the DNA will be purified. In case of sequencing of the associated, DNA adapters are ligated to the purified DNA, the DNA is side-selected and purified before the sequencing.

There are additional histones like H2A.Z and the centromere specific histone H3 (CenH3 or Cse4 homolog)), in addition to the canonical histones described above (Campos and Reinberg, 2009). Centromere composition and location is also studied by ChIP-seq, for example by targeting the CenH3 histone (Smith et al., 2011).

Chapter 1.4: Centromeres

During mitosis and meiosis the kinetochore mediates the translocation of the chromosomes in the nucleus (Vagnarelli et al., 2008). The kinetochore is connected to the chromosomes by binding to the centromere, and more specifically to the centromere specific histone H3 (CenH3) (CENP-A in humans or Cse4 in *Saccharomyces cerevisiae*) (Vagnarelli et al., 2008). The CenH3 histone is the only epigenetic mark for a functional centromere (Figure 1.4.1) (Allshire and Karpen, 2008).

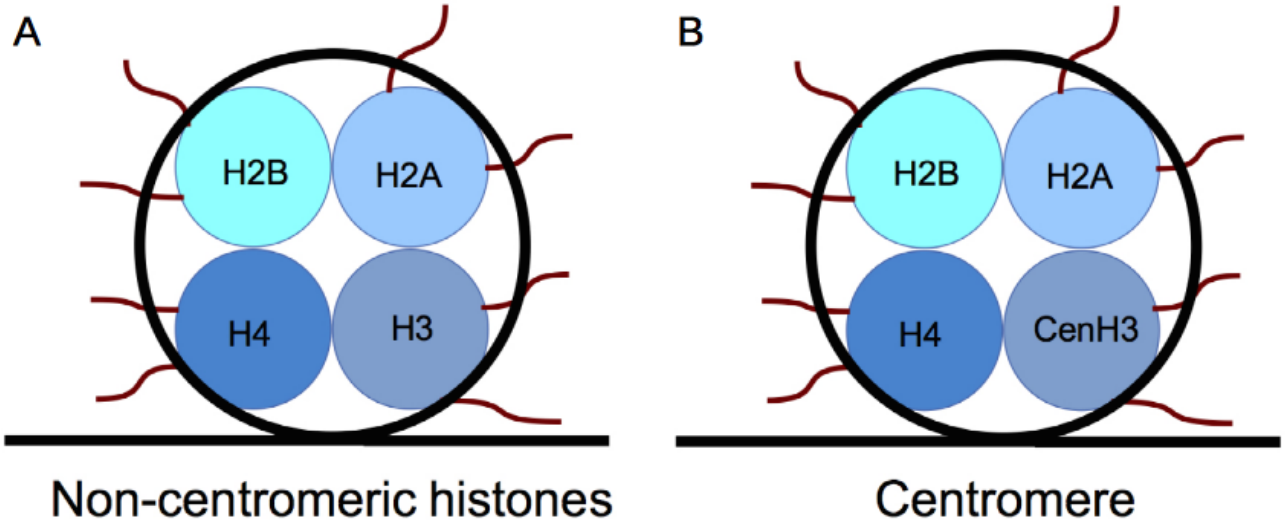

Figure 1.4.1: Centromeres are specified by a centromere specific histone H3. The DNA (in black) is wrapped around an octamer of histones (A). The histone tails are shown in red. At the centromere, histone H3 is replaced by a centromeric specific histone H3 (B).

Almost all centromeres are sequence independent. The only known exceptions are the centromeres of *Saccharomyces cerevisiae*, *S. paradoxus* and closely related budding yeasts (Clarke and Carbon, 1980) & (Rhind et al., 2011). These species have so-called point centromeres consisting of short, highly structured sequences of 125 bp (Hieter et al., 1985). The budding yeast centromeres consist of three elements (Centromeric DNA elements, CDE): A conserved eight base-pair motif forms CDE1 and a non-conserved AT rich sequence with a length of ~80 bp that forms CDE2. CDE3 consists of a 25 bp conserved motif with an imperfect palindrome. CDE3 is essential for the centromere function. Deletion or introduction of point mutations in this domain lead to non-functional centromeres (Ng and Carbon, 1987). Deletion of CDE1 and CDE2 has some impact, but the centromere remains active and stable (Cumberledge and Carbon, 1987).

All other described centromeres in eukaryotes are so called regional centromeres. Regional centromeres are independent of the underlying nucleotide sequences and are only characterized by the presence of CenH3. Lefrançois et al showed that the point centromeres of *S. cerevisiae* could

originate from regional centromeres. Over expression of the CenH3 (Cse4) of *S. cerevisiae* led to the identification of 23 centromeric-like regions (Lefrançois et al., 2013). Most of these regions are located in intergenic regions within 25 kb of the centromere. They are enriched with AT nucleotides, ~90 bp long and not bound to the kinetochore proteins. The sequence are not conserved, as in the CDEs. The authors hypothesized that the centromere-like regions are remnants of regional centromeres. Accordingly, the centromeres of *S. cerevisiae* originally were regional centromeres which have evolved to point centromeres. During the process the heterochromatin was lost and as result the centromeric repeats diverged. This would have allowed the integration of a self propagating element that could have lead to the specification of centromeric proteins.

Next to *S. cerevisiae* there are only a few species in the fungal kingdom in which centromeres are characterized. These include the model organisms *Candida albicans, Schizosaccharomyces pombe, Neurospora crassa* and *Fusarium graminearum* (Figure 1.4.2) (Sanyal et al., 2004) & (Clarke et al., 1986) & (Smith et al., 2011) & (Smith et al., 2012). In other fungal species the centromeres are predicted by sequence homology of centromere surrounding regions, but not molecularly verified (Smith et al., 2012).

C. albicans has small regional centromeres that lack any structural organization. The CenH3 binding region is approximately 3-4.5 kb (2-3.3 kb Ketel), AT-rich and not located in repetitive regions, although seven of the eight centromeres are located close to repetitive regions (Sanyal et al., 2004) & (Ketel et al., 2009). Some of these repetitive regions have a palindromic structure flanking the centromeres. Genes next to centromeres show expression levels similar to the rest of genome. This is due to the lack of pericentric heterochromatin at most of the centromeres. Only centromere 5 shows repetitive elements in this region.

The centromeres of *S. pombe* are larger (40-110 kb) than in *C. albicans,* but also structured as in *S. cerevisiae* (Wood et al., 2002). The three *S. pombe* centromeres exist of an inner core of ~10-15 kb. The inner core of chromosome 1 and 3 consist of almost identical 3.3 kb long regions. However, the inner core of chromosome 2 has no homology to the centromeric regions of chromosome 1 and 3. Deletion of the whole inner core results in lack of a functional centromere suggesting an essential function (Ishii et al., 2008). However, partial deletion of only small parts of the inner core has no effect on the centromeric function. This and the lack of homology between the three centromeres suggest that the essential part of the centromeres is small or degenerated. The inner core is flanked by inner and outer repeat domains of 10-60 kb. These domains are not conserved and consist of palindromic sequences.

N. crassa has long centromeres with a length of 150–300 kb (Smith et al., 2011). The centromeres

are located in AT-rich sequences with some relics of transposable elements. However, the centromeres are repeat rich, any organization as found in centromeres of *S. pombe,* is lacking. The lack of organization is most likely due to RIP, which has introduced point mutations into the repetitive sequences of the centromeres. The surrounding region of the centromeres are enriched with heterochromatin and cytosine DNA methylation and lack H3K4me2.

The centromeres of *F. graminearum* have also been identified by ChIP-seq of DNA bound to the CenH3 histone protein (Smith et al., 2012). The genome assembly of *F. graminearum* however consists of several thousand scaffolds and the centromere specific sequences cannot be mapped to the genome (Ma et al., 2010). The genome of *F. graminearum* exemplifies the limitations of ChIP-seq approaches for repeat rich resequenced genomes.

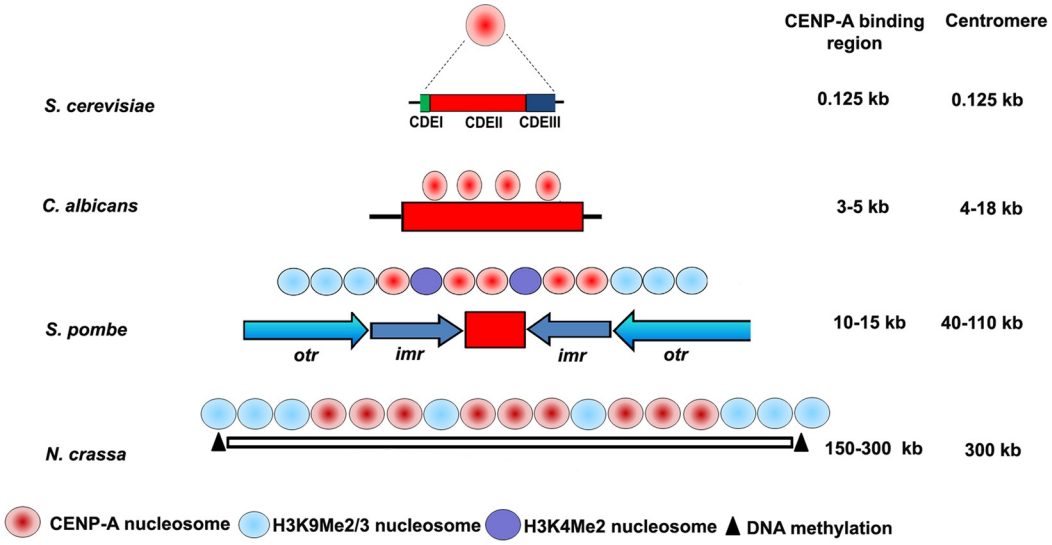

Figure 1.4.2: Centromeres of fungi (Figure modified from: (Roy and Sanyal, 2011)).

S. cerevisiae has short point centromeres and the centromere is structured. The other fungi have regional centromeres and lack any sequence structure and conserved domains, except for *S. pombe* where the inner core is flanked by outer and inner repeat domains.

The centromeres of A chromosomes of *Zea mays* are enriched with a repetitive element. In B-chromosomes, however, this element is not specific for the centromere, and is located along the whole B-chromosome. Page et al showed similarity between a region of chromosome 4 and the B-chromosome centromere repeat ZmB (Page et al., 2001). ZmB is a repetitive element (ZmB) of 1.4 kb which covers over 9 Mb. Two other repetitive elements are found in smaller regions (700 kb) of the ZmB enriched sequence, interestingly this 700 kb region co-localized with the binding of CenH3. The homology suggests that the B-chromosome centromere could originate from chromosome 4. Although the sequential similarity between chromosome 4 and ZmB, the region of chromosome 4 does not associated with CenH3. The centromeres of the B-chromosome have also

been characterized by functional studies. The deletion of sequences of the centromere resulted in the identification of the functional domain. Lamb and co-workers have shown that the B-chromosome centromere consists of significantly less amounts of CenH3 compared to the centromeres of A-chromosomes (Lamb et al., 2005). The authors hypothesize that this is the result of an accelerated mutation rate at the A-chromosome centromeres. Due the frequent impairing of the B-chromosomes, the B-chromosomes get lost in the population and this results in a lower selection pressure in-comparison to the A-chromosomes.

Padmanabhan et al compared the centromeres of the closely related candida species, *C. albicans* and *Candida dubliniensis* (Padmanabhan et al., 2008). The overall sequence conservation between the two species is high. Nevertheless, the CenH3 associated regions to other species are highly diverged. Still, the chromosomal locations of the centromeres are conserved between species. The ORF free regions between the centromeres and the coding sequences show an increase in sequence conservation towards the side where genes are predicted. The genes adjacent to the centromeres show an increased identity of 81–99 % at the amino acid level. This suggest that the centromeric sequence are under less selective constraints and have a higher mutation rate than the non-centromeric regions.

A high evolutionary rate of centromeres has also been shown by comparison of the centromeric regions of several *Schizosaccharomyces* species. Rhind et al compared the genomes of four members of the *Schizosaccharomycetes* (Rhind et al., 2011). The centromeres here were identified based on sequence similarity to the centromeres of *S. pombe*. The centromeres of *S. pombe* and *Schizosaccharomyces octosporous* are homologous and both lack transposable elements, while centromeres in *Schizosaccharomyces japonicus* on the other hand are enriched with transposable elements (Rhind et al., 2011). This underlies the variation in centromere composition even among species of the same genus.

Mutations or (partial) deletions of a centromere can result in the formation of a neocentromere (Burrack and Berman, 2012). A newly formed centromere is fully functional and like the original centromeres associated with all known centromeric proteins. As most centromeres are regional centromeres, the centromeres are independent of a specific DNA sequence and neocentromere could in principle occur at a random location along the chromosome (Burrack and Berman, 2012). However, in most species where neocentromere formation has been studied, newly formed centromeres were located in regions enriched with heterochromatin.

In *C. albicans,* Ketel et al showed that replacement of centromere 5 (cen5) with a selection marker resulted in the formation of a neocentromere on the chromosome (Ketel et al., 2009). As *C.*

albicans is diploid, the authors selected for strains which lacked cen5 of both sister pairs of chromosome 5 and obtained four strains where the original cen5 was absent. Growth curves for several cen5 deletion strains showed that cen5 is not essential as there was no difference in the fitness between the cen5 deletion strains and the wild-type. In general, the neocentromere formation occurred in the close proximity of the original centromeric locus or adjacent to the left telomere. Like the original centromeres, the neocentromere are not enriched for DNA motives, but are still located close to repetitive sequences. Compared to the original centromeres, the neocentromere showed a lower amount of enrichment of CenH3.

Also Thakur et al, induced neocentromere formation in *C. albicans* by several experiments (Thakur and Sanyal, 2013). In the first experiment, the centromeric region was extended by the integration of an URA3 gene inside the centromeric region. This integration did not result in chromosome losses or chromosome rearrangements. The second experiment was the replacement of the core domain of centromere 7 (cen7) with an URA3 gene. Besides neocentromere formation close to the original centromere locus, chromosomal rearrangements and the loss of chromosome 7 were observed. The length of the neocentromere is similar to the original centromere (3–4 kb). The third experiment was the deletion of the whole centromere 7. None of the transformed strains showed chromosome abnormalities. Neocentromere formation occurred at the same position where also the neocentromere formed after the deletion of the core domain. Additionally, a neocentromere was identified 13 kb from the original centromere. The final experiment consisted of the deletion of a 30 kb region, including the centromeres and flanking regions. Neocentromere formed adjacent to the deleted region. Similar to the results of Ketel et al, the neocentromeres show lower levels of CenH3 enrichment in comparison to the wild-type centromere. After reintroducing the sequence of cen7 the chromosome could become dicentric. However, ChIP experiments of gene-converted derivates showed that the neocentromere is not enriched anymore with CenH3 and the functional centromere was again the original centromere. Suggesting a preference of a particular chromosome location.

To confirm and compare the results to the outcome of Ketel et al, Thakur et al also deleted the cen5. They obtained neocentromeres close to the wild-type cen5 with a longer length than the wild-type cen5. As mentioned, the chromosomal location of the centromeres are conserved between *C. albicans* and *C. dubliniensis*. To study if the neocentromere formation is conserved between both species, the cen7 of *C. dubliniensis* was deleted by Thakur et al. Five mutants where obtained and neocentromeres were found close to the deleted wild-type centromeres.

Although the centromeres of *C. albicans* are defined as epigenetic features, both studies shown that neocentromere formation in *C. albicans* occurs close to the original centromere position and that

the neocentromeres still have common characteristic similar to the original centromeres.

In *S. Pombe,* Ishii and co-workers deleted the whole centromere, including the outer heterochromatic regions of chromosome 1 (Ishii et al., 2008). After the introduction of two loxP sites, flanking the centromere and the expression of Cre recombinase, the centromere could be excised from the chromosome. The excised sequence corresponding to the centromere, formed an additional circular chromosome. PFGE and subsequently Southern blots showed that the transformed cells lost centromere 1 (cen1) and subsequently neocentromere formation occurred.

ChIP experiments showed that the neocentromere formed mainly at regions close to both telomeres. These regions contain several ORFs and are not enriched with repetitive elements. As in *C. albicans,* the newly formed neocentromeres silence the genes which are located inside the neocentromere. The newly formed neocentromeres of *S. pombe* have the same length as the original centromeres and are located near heterochromatic regions. Some transformed strains also underwent chromosomal rearrangements in addition to neocentromere formation. The rearrangements occurred between the telomeres of chromosome 1 and 2 and resulted in large chromosomes. The fused chromosomes lacked neocentromere formation and the original centromere of chromosome 2 was still active.

To test if heterochromatin is required for neocentromere formation, the cen1 was deleted in several heterochromatin-deficient strains. In comparison to the wild-type, there was a significant decrease of neocentromere formation and increase of chromosome fusion. Nevertheless, neocentromere formations occurred in the absences of heterochromatin.

It is not known if centromere stability or composition can be involved in genome dynamic of fungal plant pathogen, e.g. through an effect on the dispensable chromosomes.

Centromere and telomeric regions of *S. pombe* and *N. crassa* are enriched with small-RNAs as opposed to the centromeres of *C. albicans* (Volpe et al., 2003) & (Lee et al., 2010) & (Drinnenberg et al., 2009). This suggests that small-RNAs are important for the formation of centromeres or the histone modifications which are associated with the centromeres. Not only are small-RNAs important for the genome stability, they also have functions in many more cell processes as is described in the following chapter (Chang et al., 2012) & (Billmyre et al., 2013).

Chapter 1.5: Proteins involved in the RNAi-pathways

RNA interference (RNAi) is a conserved mechanism that effects gene expression by post-translation or transcriptional silencing (Ketting, 2011). It also has been shown that small-RNAs are involved in defense mechanisms against pathogens and repetitive elements and transposons (Malone and Hannon, 2009). RNAi was described in *Caenorhabditis elegans* in 1998, but RNAi related pathways were already described in plants and fungi in respectively 1990 and 1992 (Fire et al., 1998) & (Napoli et al., 1990) & (Romano and Macino, 1992).

Small RNAs are processed by Dicer and Argonaute proteins in most cases (Wilson and Doudna, 2013). Gene silencing is mediated by small-RNAs (20-30 nt) which are "diced" by a Dicer protein (containing a ribonuclease III (RNase III) endonuclease) and subsequently bound by an Argonaute protein (Wilson and Doudna, 2013). Once bound by the Argonaute, complementary RNA sequences are identified by homology and are degraded. The RNAi mechanism is conserved in the eukaryotic kingdom however there are many species independent, alternative pathways to produce a broad range of small-RNAs which all have distinct functions (Mallory and Vaucheret, 2010).

In the fungal kingdom, insight into RNAi pathways are mainly obtained from studies in *Neurospora crassa*, *Cryptococcus neoformans* and Schizosaccharomyces *pombe* (Dang et al., 2014) & (Liu et al., 2002) & (Martienssen et al., 2005). Although the RNAi mechanism is present in all fungal families, several fungal species have lost the ability to produce small-RNAs (Nicolás et al., 2013). The best studied example is *Saccharomyces cerevisiae* (Drinnenberg et al., 2009). *S. cerevisiae* lacks homologs of both the Dicer and Argonaute encoding genes. Interestingly, the RNAi mechanism is still present in the closely related yeast *Saccharomyces castellii.* RNAi in *S. cerevisiae* can be restored by introducing Dicer and Argonaute genes of *S. castellii* (Drinnenberg et al., 2009). It was also shown that the loss of RNAi in *S. cerevisiae* correlates with the acquisition of a so-called Killer virus in the *S. cerevisiae* genome (Drinnenberg et al., 2011). Strains with Killer have an advantage because they can inhibit the grow of non Killer strains which grow in close proximity. By re-introducing the RNAi mechanism, *S. cerevisiae* obtained a defense mechanism against viruses and this resulted in the loss of the Killer system.

Another yeast, *Candida albicans* lacks a conventional Dicer (Bernstein et al., 2011). This is compromised by the presence of an RNase III nuclease, which, in addition to the dicing function, also is essential for ribosome and spliceosome biogenesis.

In *N. crassa* two major RNAi pathways are identified, one vegetative and one sexually dependent pathway (Dang et al., 2014). Both have a specific biogenesis to produce small-RNAs (Figure 1.5.1).

The vegetative pathway is better known as quelling and is a post-transcriptional mechanism of gene silencing active during vegetative growth (Romano and Macino, 1992). Quelling was discovered by Ramano and Macino when they performed a transformation of Albino-1 and Albino-3 genes resulting in the reduction of expression of the specific genes.

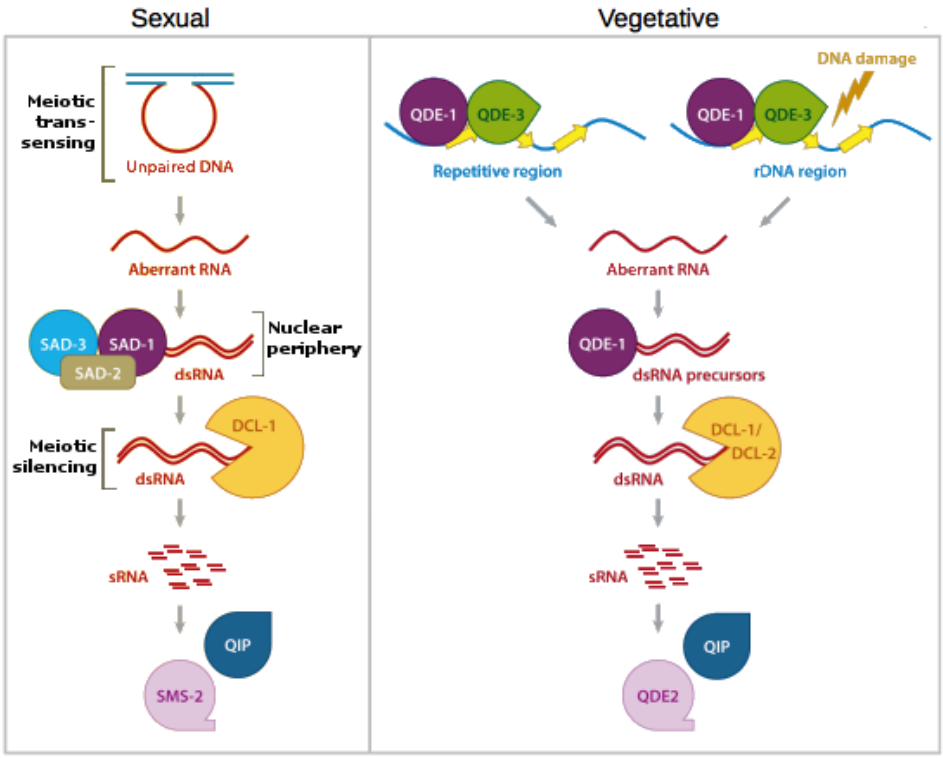

Figure 1.5.1 The RNAi pathways of *N. crassa* are divided in a sexual and vegetative stage (Figure modified from: (Chang et al., 2012)). During the sexual stage, unpaired DNA is recognized and aberrant RNA is produced. The aberrant RNA is converted into dsRNA, diced into siRNAs by Dcl-1 and loaded on the Argonaute (Sms-2). During the vegetative stage, aberrant RNA is produced as a result of the recognition of repetitive sequences. Besides this aberrant RNA could also derive from ribosomal DNA. The aberrant RNA is converted into dsRNA and processed by either Dcl-1 or Dcl-2. The small-RNAs are loaded on the Argonaute Qde-2 and transported to the target sequence.

The quelling pathway is activated by (single) DNA repeats, either from repetitive sequences or from the ribosomal DNA. The aberrant DNA is transcribed into aberrant RNA (aRNA) by QDE-1 (RNA polymerase) and the aRNA is processed into 25nt small interfering RNA (siRNA) by a Dicer-like protein (Chang et al., 2012). The siRNAs are bound by the QDE-2 (an Argonaute protein) and subsequently the siRNA degrades the homologous messenger RNA (mRNA).

The sexual RNAi pathway is triggered by unpaired DNA (Meiotic silencing by unpaired DNA (MSUD)), resulting from the absence of sequence paring of homologous chromosomes (Hammond et al., 2013). The unpaired DNA is transcribed into aRNA which leads to dsRNA produced by SAD-1 (RNA polymerase). Dicer-like 1 cleaves the dsRNA and the resulting small noncoding RNAs (sRNAs) are loaded on SMS-2 (Argonaute). In contrast to quelling, only Dicer-like 1 is essential

and Dicer-like 2 is not involved in the sexual RNAi pathway. The final step is the silencing of homologous DNA by post-transcriptional silencing.

In addition to the siRNAs and sRNAs, several other classes of small-RNAs have been identified in *N. crassa.* For example qiRNA (QDE-2 interacting small-RNAs), these are 21-23 nt long small-RNAs which are produced as a reaction to DNA damage and are the product of a pathway similar to the quelling pathway (Lee et al., 2009) & (Lee et al., 2010). The most obscure class of small-RNAs in *N. crassa* are the Dicer-independent small interference RNAs (disiRNAs) (Lee et al., 2010) & (Dang et al., 2013). These are 22 nt long small-RNAs which originate from non-repetive DNA loci. Strikingly, non of the known Dicer-like and Argonaute genes are involved in the biogenesis of this class of small-RNAs. One class of small-RNAs are the micro-RNAs (miRNA) with a size of 25 nt. The biogenesis of these small-RNAs is well studied in plants and animals (Carthew and Sontheimer, 2009). Only recently miRNAs were discovered in fungi first in *N. crassa* and later in *C. neoformans* (Lee et al., 2010) & (Jiang et al., 2012). In *N. crassa* miRNAs are QDE-2 dependent, have a strong 5' U bias, and as in animals and plants, the miRNAs originate from precursor-miRNAs. The *N. crassa* miRNAs allow mismatches when targeting homologous sequences as miRNA targeting in animals. At least four independent pathways to synthesize miRNA in *N. crassa* have been described (Lee et al., 2010).

In the pathogen *C. neoformans,* two miRNAs, miR1 (22 nt) and miR2 (18 nt) originating from a pre-miRNA of +/- 70 nt have been identified (Jiang et al., 2012). The biogenesis of these miRNAs is unknown. Similar to *N. crassa, C. neoformans* also has an RNAi pathway which shows similarities with quelling and depends on Dicer and Argonaute (Wang et al., 2012). The SIS (sex-induced silencing) pathway functions as a post-transcriptional silencing mechanism where the efficiency depends on the copy number of the gene (Wang et al., 2012). Despite the name, sex-induced silencing, the SIS pathway is also active during vegetative growth, although the frequency is significantly lower.

Exonic-siRNAs (ex-siRNAs) have been identified in the fungus *Mucor circinelloides* (Nicolas et al., 2010). The ex-siRNAs are synthesized from exons and target the mRNA of which they originate. Until now four classes of ex-siRNAs have been characterized and have been classified based on their biogenesis pathway. Piwi-interacting RNAs (piRNAs) have not been identified in fungi, nevertheless Dicer independent small-RNA have been found in *N. crassa* and in *S. pombe* (Lee et al., 2010) & (Halic and Moazed, 2010). In *S. pombe*, priRNAs (primal small-RNAs) are Dicer independent and originate from the 3' UTR of cellular transcripts and centromeric repeats (Halic and Moazed, 2010).

In 2002, Volpe et al showed that in *S. pombe* the formation of heterochromatin depends on small-RNAs (Volpe et al., 2002). Heterochromatic DNA is defined by methylation of lysine 9 of histone H3 (H3K9) (See also chapter 1.3). Due to this specific histone mark, the DNA is densely condensed and consequently the genes are transcriptionally silenced. In *S. pombe*, heterochromatic DNA is distributed at the telomeres, centromeres and mating type loci (Volpe et al., 2003). In *S. pombe,* siRNAs originates from nascent RNA of the pericentromeric regions and are loaded on the RITS complex (RNA-induced transcriptional silencing complex) consisting of an Argonaute and a chromodomain containing protein connecting the RITS complex to heterochromatin. The siRNA which is bound by the RITS complex acts as a primer. Subsequently, an RNA-dependent RNA polymerase starts to synthesize dsRNA of the earlier mentioned nascent RNA. The dsRNA is diced by Dicer into small-RNAs and loaded onto the Argonaute of the RITS complex. In Dicer and Argonaute mutants of *S. pombe,* transcription at the centromeres and telomeres have been described. More over, reduced H3K9 methylation and increased H3K4 levels can result from Dicer and Argonaute deletions. All of these results suggest that the heterochromatin formation and RNAi in *S. pombe* are closely connected.

Most of the studied plant pathogenic fungi still have a functional RNAi system (Billmyre et al., 2013). RNAi is sometimes used experimentally, as a tool to silence transcripts, if gene deletion is difficult. Most insight into the function of RNAi in plant pathogenic fungi is based on the sequencing and analyses of small-RNAs.

In the genome paper of *Z. tritici,* Goodwin et al suggested the presence of 418 micro-RNAs in the *Z. tritici* genome (Goodwin et al., 2011). This prediction was based on a computational comparison of pre-microRNAs from the RFAM database with the genome of *Z. tritici.* According to the authors, there is a particular enrichment of putative micro-RNAs on the dispensable chromosomes.

In the ascomycete pathogen *Sclerotinia sclerotiorum,* computational evidence for presence of miRNA was also reported and experimental evidence for two potential miRNAs was shown by Northern blot (Zhou et al., 2012). Differential expression of the putative miRNAs during plant infection was shown in *S. sclerotiorum*. Suggesting a role of small-RNAs in pathogenicity. Differential expression of small-RNAs under specific conditions was also shown in the rice blast pathogen *Magnaporthe oryzae* (Nunes et al., 2011). In this fungus, apperessoria are formed during initial infection and small-RNAs could be isolated from these cells. The apperessoria specific small-RNAs have a size of 28-35 nt. Interestingly, small-RNAs of 18-23 nt are expressed from mycelia. These small-RNAa are mainly derived from transposable elements and other intergenic regions.

The RNAi system of the oomycete plant pathogen *Phytophthora infestans* has been characterized

by Vetukuri et al (Vetukuri et al., 2011). *P. infestans* has one Dicer and 5 Argonaute genes. The Dicer is up-regulated 24 hours post infection, this is in contrast to the Ago1/2 which shows up-regulation during the germination of cysts. To verify the involvement of the Dicer and Ago1/2 in gene silencing, a transient silencing experiment was performed. A marker gene was silenced in the wild-type and this resulted in an almost complete absence of the mRNA of the marker gene. The mRNA levels of the maker gene were restored when either the Dicer or Ago1/2 in this strain was silenced.

Small-RNAs of *P. infestans* were sequenced from cells growing at different conditions (Vetukuri et al., 2012). The majority of the small-RNAs map to transposable elements and belong to three groups. The first group contains small-RNAs with a size of 21nt are mainly derived from Crinkler effector genes (CRN) and transposable elements. The second (25-26nt) and third group (32nt) mainly map to transposable elements. Specific categories of transposable elements were shown to be associated with one of the small-RNA groups. Two almost identical pathogenicity effector related genes (CRNs) were studied in more detail. It was found that the small-RNAs were enriched at three positions of the effector genes, mainly where they are more variable. The relevance of small-RNAs in pathogenicity was shown by the comparison of small-RNAs derived from two strains. One of the strains shows reduced pathogenicity on potato. This strains has small-RNAs mapped to two RxLR effector coding genes. However, in the pathogenic strain small-RNAs were absent from the effector encoding sequence. Vetukuri and co-workers concluded that the RNAi mechanism plays a role in virulence and that transcription of effector genes can be particularly affected by specific small-RNAs.

The RNAi machinery in plants is well studied, in particular for the model species *Arabidopsis thaliana* (Mallory and Vaucheret, 2010). One of the focuses has been the use of RNAi as a defense mechanism. Many studies have connected miRNA syntheses in *A. thaliana* to a defense mechanism against bacteria, fungi and viruses (Pumplin and Voinnet, 2013) & (Katiyar-Agarwal and Jin, 2010). Host-pathogen interaction can also occur at the level of small-RNAs. An example is the bacteria flagellin derived PAMP (flg22) which induce miRNA393 expression in the host. It was shown that *Pseudomonas syringae* secrets effectors (AvrPto and AvrPtoB) that are able to interfere with the host RNAi pathways to reduce the expression of miRNA393 (Zhang et al., 2011). Also infections with *Agrobacterium tumefaciens* leads to reduced levels of miRNA393 and miRNA167 in *A. thaliana* (Zhang et al., 2011). Ellendorff et al showed that the RNAi machinery of *A. thaliana* plays a role in the defense mechanism against pathogenic fungi (Ellendorff et al., 2009). The authors tested the involvement of multiple components of the RNAi machinery by comparing the infection

rate of different *A. thaliana* mutants and several fungal pathogens. Interestingly, only the infection rate of *Verticillium dahliae* was reduced. The most dramatic effect was observed in Argonaute 1, Dicer1 and Hen1 mutants. Other examples of host-pathogen interactions at the level of small-RNAs involve the fungus *Cronartium quercuum f. sp. fusiforme* that induces miRNA syntheses in developing xylem tissue of the host loblolly pine (*Pinus taeda*) (Lu et al., 2007). Infection of *Erysiphe graminis* in *Triticum aestivum* resulted in the down-regulation of a wheat derived small-RNA (miRNA156) (Gupta et al., 2012). Consequently, there was an up regulation of a particular mRNA, a putative transcription factor (Ta3711) that involves in the wheat defense system.

An intriguing question is how small-RNAs can be transferred from one species to another. Transfer of RNA between species has been shown in plants, where RNA hairpins from the parasitic plant *Triphysaria versicolor* were expressed in the tissue of the host plant lettuce and *A. thaliana* (Westwood et al., 2009).

Still, little is known about the role of small-RNAs in plant pathogenic fungi during infection. Only recently Weiberg et al have shown that small-RNAs play a role in the infection process of the host plants (*A. thaliana* and *Solanum lycopersicum*) of *Botrytis cinerea* (Weiberg et al., 2013). In the experiment of Weiberg and colleagues, both hosts were infected with *B. cinerea* and small-RNAs isolated from the infected plants at several time-points during infection. Deep sequencing of the small-RNAs resulted in 73 *B. cinerea* small-RNAs which were up regulated during the infection of both host plants. These fungal derived small-RNAs map to the genomes of both host plants. To investigate the role of small-RNAs during infection, three small-RNAs of *B. cinerea* were selected for further investigation. These specific small-RNAs were found to suppress genes involved in host defenses of *A. thaliana. S*everal experiments have been performed to verify if the down regulation of the selected host plant genes is the result of the silencing by small-RNAs of *B. cinerea*. Firstly, a plant artificial miRNA vector was used to express the three *B. cinerea* small-RNAs in *A. thaliana.* This resulted in the down-regulation of the host defense genes. Secondly, a Dicer mutant of *B. cinerea* was created and used for infection of *A. thaliana*. This resulted in a strong reduction of virulence. To investigate how the small-RNA of *B. cinerea* can hijack the RNAi system of the host plant, an Ago1 mutant of *A. thaliana* was infected with *B. cinerea*. This also resulted in reduced virulence in comparison to infection of wild type *A. thaliana* plants. These results confirm that small-RNAs of *B. cinerea* hijack the RNAi system of the host plant to suppress target genes involved in host defenses. However it is still unknown how the delivery mechanism of the *B. cinerea* small-RNAs to its host plant works.

Aims of the study

Zymoseptoria tritici provides an excellent model to study the underlying molecular mechanisms of host-pathogen interactions and pathogen adaptation. Furthermore the particular genome structure of this fungus allows studies of genome evolution and genome plasticity in fungal plant pathogens. So far the epigenetic component of *Z. tritici* has been unknown.

One aim of this study was to identify and characterize the centromeres of *Z. tritici.* More specifically, to compare centromeres of essential and dispensable chromosomes.
By ChIP-seq of CenH3 associated DNA the centromeres could be identified and mapped to the genome.
The second aim was to characterize histone modifications of the essential and dispensable chromosomes. ChIP-seq of three histone modifications have led to the characterization of the chromatin organization.
Finally, the third aim was to study and to identify if the RNAi mechanism in *Z. tritici* is involved in host-pathogen interactions. Several mutants of genes involved in the RNAi mechanism were created and tested on planta.

Chapter 2: Results

Chapter 2.1: Histone modifications of essential and dispensable chromosomes in *Z. tritici*

Chromatin

To determine the chromatin status of the chromosomes of Zymoseptoria *tritici*, ChIP-seq experiments targeting histone modifications of the tail of histone 3 were performed. The histone modifications which were the subject of this study were H3K4me2, H3K9me3 and H3K27me3. After the ChIP experiments the fragments were sequenced by Illumina sequencing and the reads were mapped to the *Z. tritici IPO323* genome (Table 2.1.3) (Goodwin et al., 2011).

The chromatin organization of chromosome 3 is illustrated in Figure 2.1.1. Chromosome 3 is an essential chromosome with a total length of 3.5 Mb (Table 2.1.1). The occurrence of euchromatin correlates with the absence of heterochromatin as expected.

There is a striking difference in the chromatin organization between the essential and dispensable chromosomes (Figure 2.1.1 and 2.1.2). Essential chromosomes are mainly enriched with H3K4me2. In contrast, dispensable chromosomes have only a few regions with are enriched with H3K4me2. Instead, dispensable chromosomes are enriched with H3K9me3 and H3K27me3 modifications. This corresponds to the observed differences in gene density, gene transcription and repeat content between the two types of chromosomes.

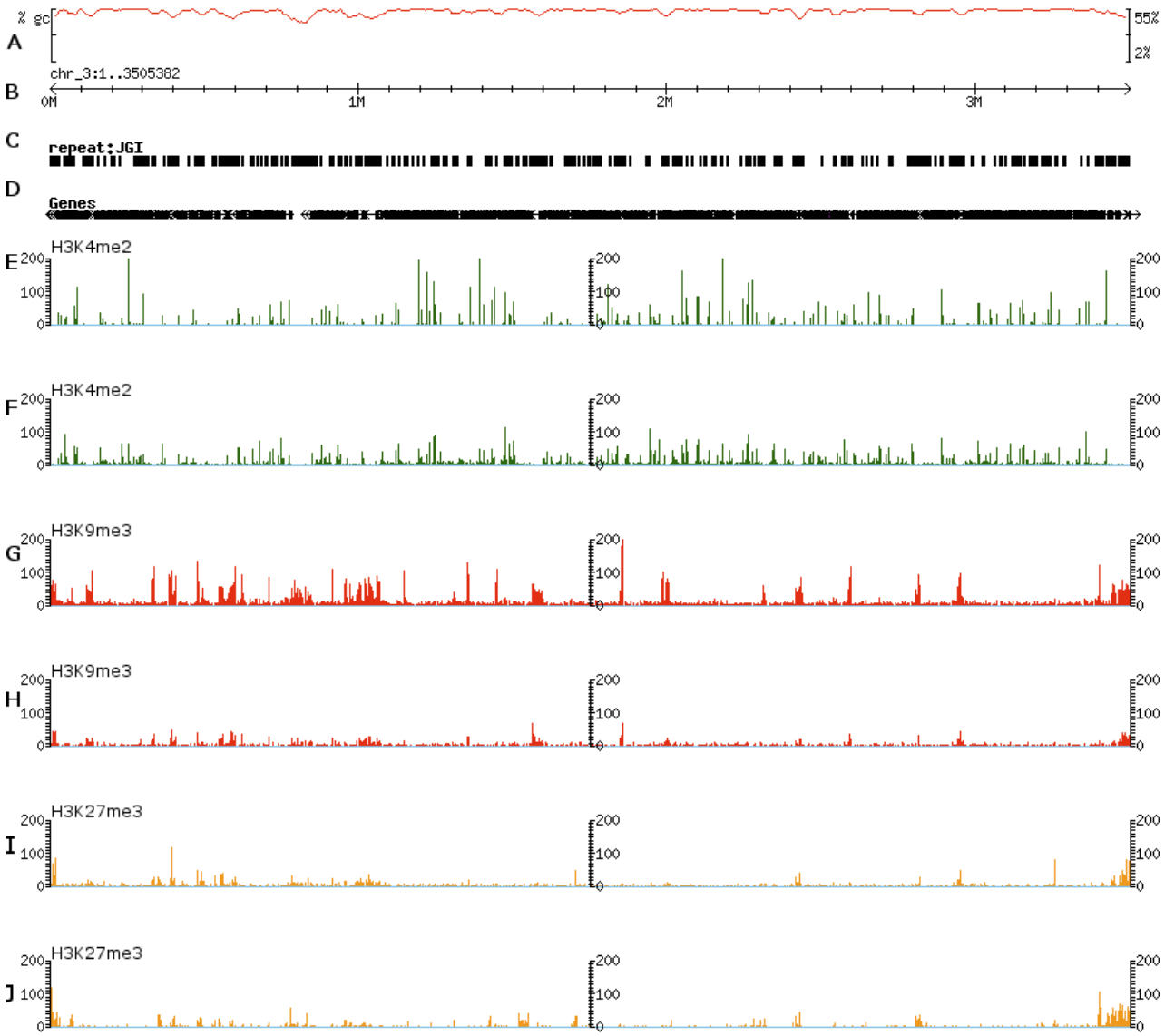

Figure 2.1.1: Chromatin organization of the essential Chromosome 3.

A) GC enrichment, B) Chromosome length, C) Amounts of repeats, D) Gene predications based on the JGI annotation, E & F) Coverage of H3K4me2 reads, G & H) Coverage of H3K9me3 reads, I & J) Coverage of H3K27me3 reads.

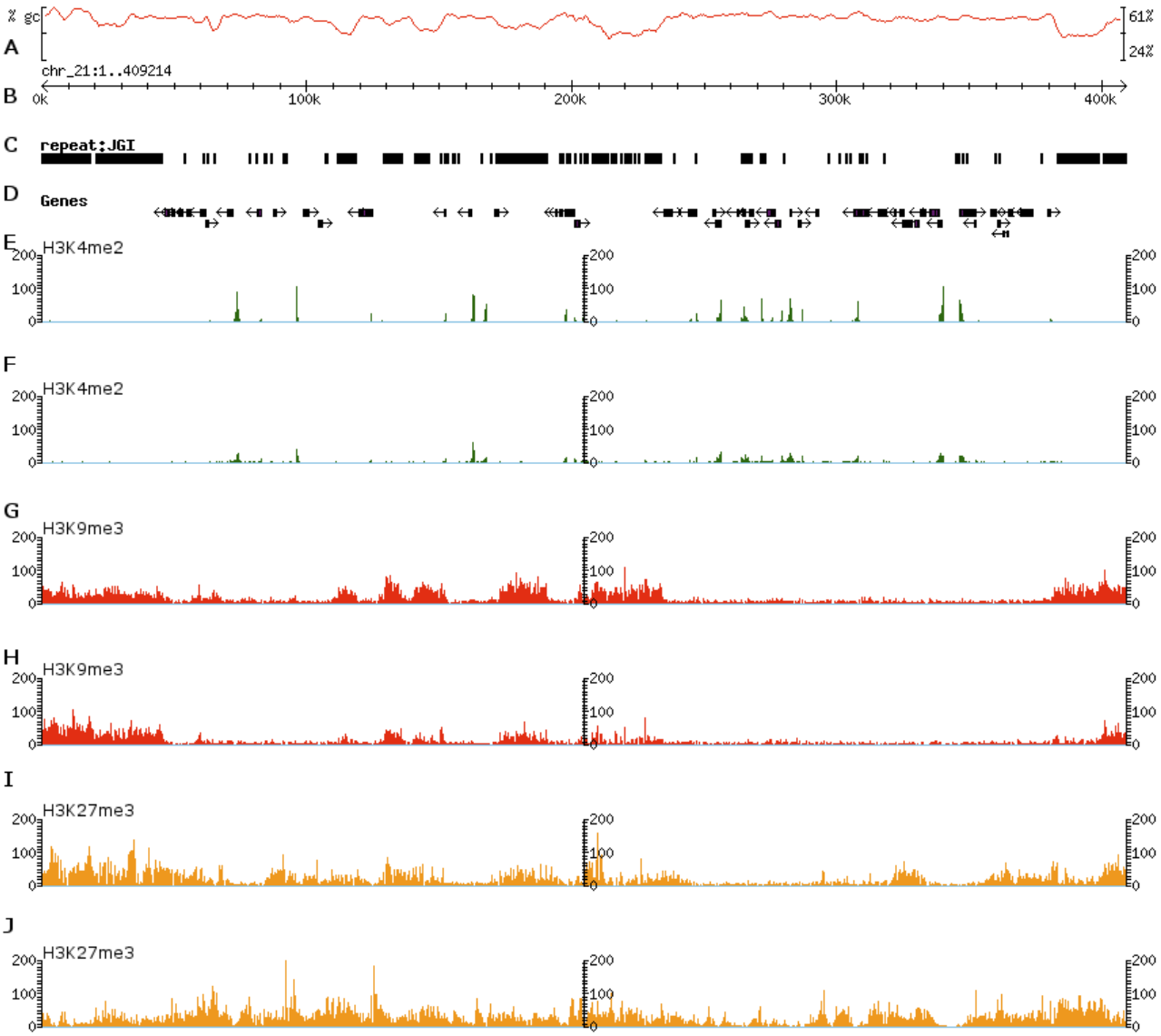

Figure 2.1.2: Chromatin organization of the dispensable Chromosome 21.

A) GC enrichment, B) Chromosome length, C) Amounts of repeats, D) Gene predications based on the JGI annotation, E & F) Coverage of H3K4me2 reads, G & H) Coverage of H3K9me3 reads, I & J) Coverage of H3K27me3 reads.

Heterochromatic islands

The genome of *Z. tritici* has several regions that are enriched with AT nucleotides (Table 2.1.1). The average length of these AT-rich (>51%) regions is 46 kb. In IPO323 24 regions were characterized here. Twenty of these AT-rich regions are located on the essential chromosomes and are highly enriched in repeats, lack any gene predication and are not centromeric. The chromatin organization of these regions are different from the rest of the chromosome. However, the surrounding regions still have H3K4me2 marks and encode genes, the entire AT-rich region lacks H3K4me2 and are highly enriched with the H3K9me3 histone modification (Figure 2.1.3 and figure 2.1.4).

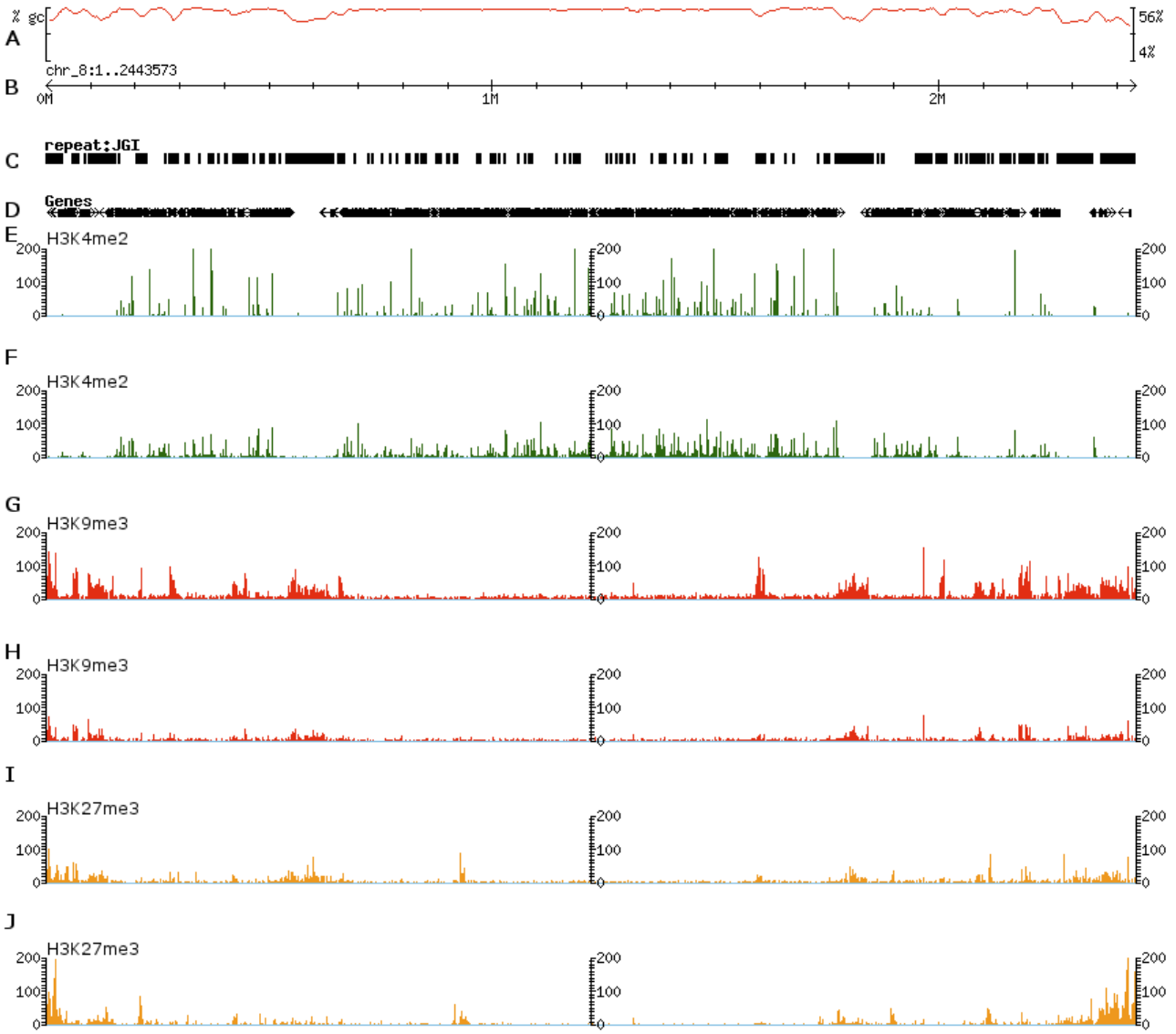

Figure 2.1.3: Chromatin organization of the essential Chromosome 8.

A) GC enrichment, B) Chromosome length, C) Amounts of repeats, D) Gene predications based on the JGI annotation, E & F) Coverage of H3K4me2 reads, G & H) Coverage of H3K9me3 reads, I & J) Coverage of H3K27me3 reads.

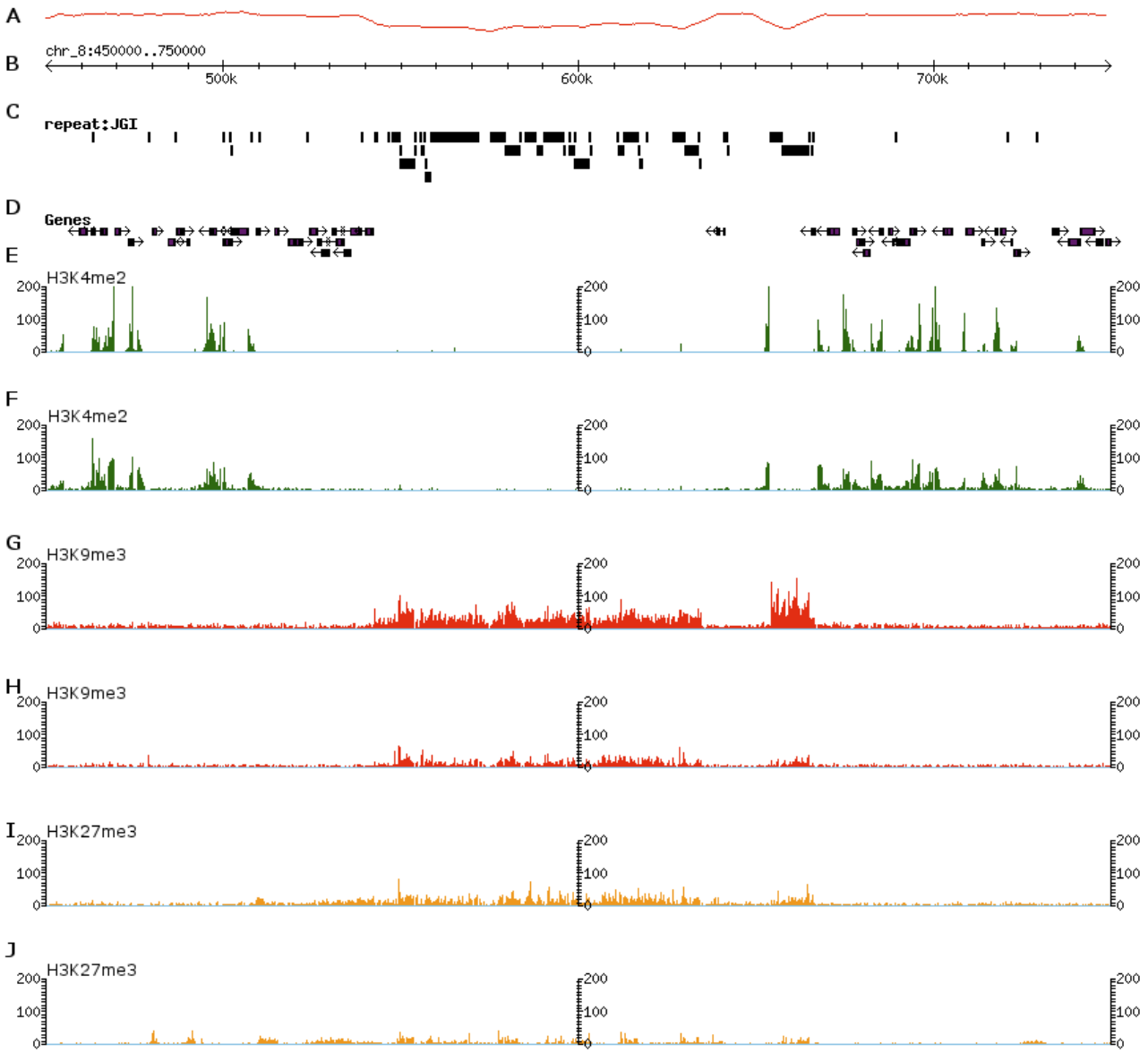

Figure 2.1.4: Details of a heterochromatic island of chromosome 8.

A) GC enrichment, B) Chromosome length, C) Amounts of repeats, D) Gene predications based on the JGI annotation, E & F) Coverage of H3K4me2 reads, G & H) Coverage of H3K9me3 reads, I & J) Coverage of H3K27me3 reads.

Table 2.1.1: Statistics and overview of heterochromatic islands of *Z. tritici IPO323.*

The first column shows the chromosome length and the average AT content of the chromosome. The second column shows the number of heterochromatic islands on the specific chromosome. The third column shows the chromosomal position and the length of the heterochromatic island. The fourth and last column shows the AT content of the heterochromatic island.

	Chromosome length (Mb) / AT-content	Number of heterochromatic islands	Coordinates and length (kb) of heterochromatin islands	AT content of the heterochromatic islands
Chr1	6.09 / 46.9 %	1	Chr_1:2761000..2799999 (34)	56.6 %
Chr2	3.86 / 47.6 %	1	Chr_2:3202000..3306999 (105)	54.3 %
Chr3	3.51 / 47.4 %	1	Chr_3:778000..844999 (67)	57.3 %
Chr4	2.88 / 47.8 %	1	Chr_4:1528000..1604999 (77)	57.7 %
Chr5	2.68 / 48.0 %	1	Chr_5:911000..988999 (78)	56.4 %
Chr5	2.68 / 48.0 %	2	Chr_5:1764000..1818999 (55)	56.3 %
Chr6	2.68 / 48.6 %	1	Chr_6:402000..468999 (67)	56.4 %
Chr7	2.67 / 47.2 %	1	Chr_7:751000..806999 (56)	55.6 %
Chr8	2.44 / 48.3 %	1	Chr_8:543000..634999 (92)	56.7 %
Chr8	2.44 / 48.3 %	2	Chr_8:1776000..1837999 (62)	55.4 %
Chr8	2.44 / 48.3 %	3	Chr_8:2178000..2207999 (30)	51.1 %
Chr8	2.44 / 48.3 %	4	Chr_8:2270000..2344999 (75)	58.3 %
Chr9	2.14 / 48.5 %	1	Chr_9:948000..975999 (28)	55.4 %
Chr9	2.14 / 48.5 %	2	Chr_9:1050000..1081999 (32)	51.9 %
Chr9	2.14 / 48.5 %	3	Chr_9:1380000..1457999 (78)	54.6 %
Chr10	1.68 / 47.5 %	1	Chr_10:861000..880999 (20)	51.3 %
Chr10	1.68 / 47.5 %	2	Chr_10:1399000..1432999 (34)	55.8 %
Chr10	1.68 / 47.5 %	3	Chr_10:1551000..1584999 (34)	57.0%
Chr11	1.62 / 47.2 %	1	Chr_11:1045000..1064999 (20)	54.9 %
Chr12	1.46 / 47.7 %	1	Chr_12:76000..118999 (43)	54.3 %
Chr12	1.46 / 47.7 %	2	Chr_12:410000..421999 (12)	53.5 %
Chr12	1.46 / 47.7 %	3	Chr_12:606000..624999 (19)	55.0%
Chr13	1.19 / 48.0 %	1	Chr_13:315000..340999 (26)	54.6 %
Chr14	0.77 / 51.5 %	1	Chr_14:213000..253999 (41)	58.5 %
Chr16	0.61 / 48.5 %	1	Chr_16:168000..188999 (21)	55.9 %
Chr20	0.47 / 48.5 %	1	Chr_20:287000..326999 (40)	57.9 %

A region enriched with facultative heterochromatin

In addition to the heterochromatic islands, one large region of chromosome 7 shows complete absence of gene transcription (Kellner et al., 2014). This region is approximately 780 kb long and is located on the right chromosome arm of chromosome 7 (Figure 2.1.5). As for the heterochromatic islands, the H3K4me2 mark is absent in this region. However, there is a high enrichment of H3K27me3 (facultative heterochromatin).

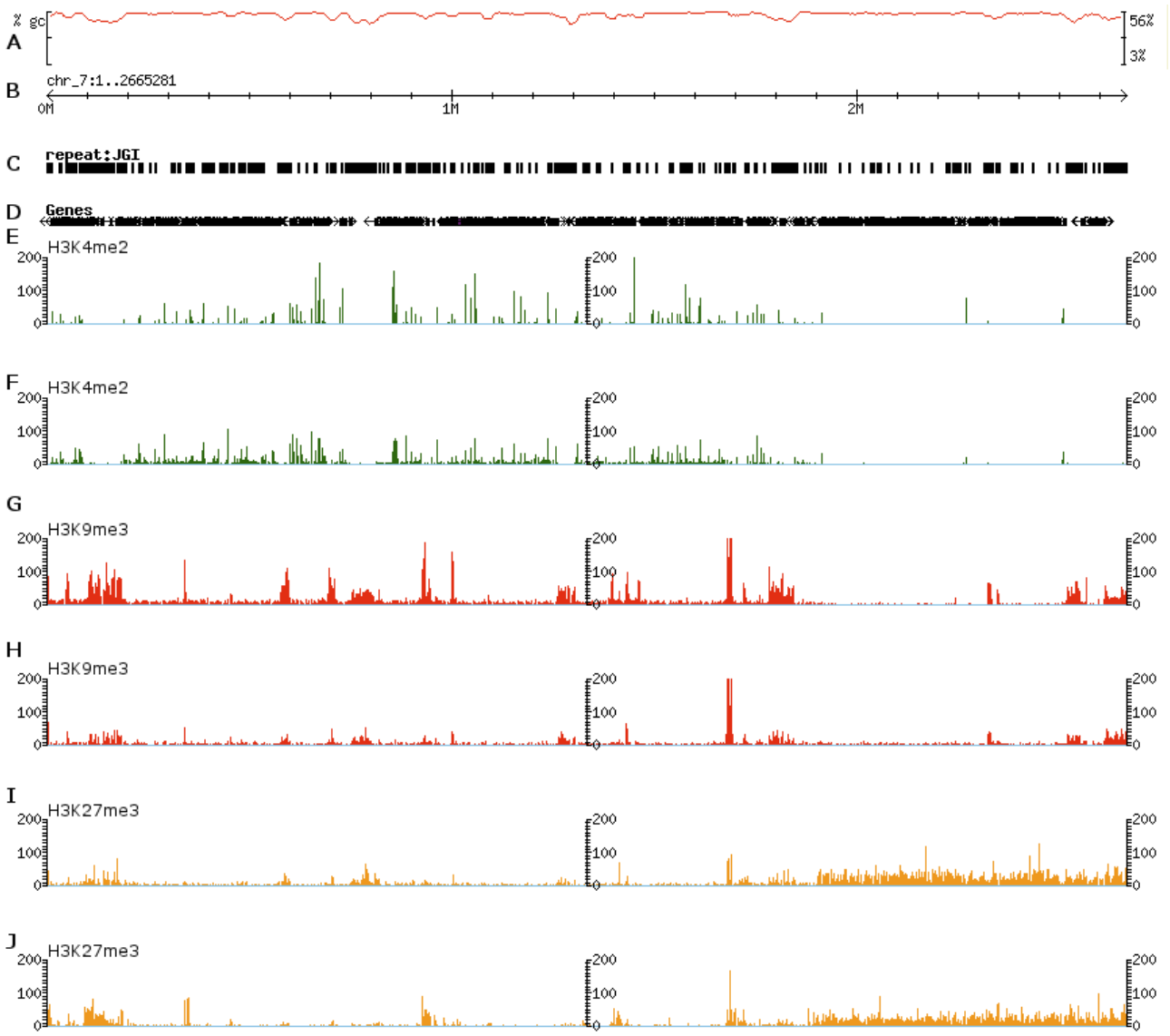

Figure 2.1.5: Details of the region enriched with H3K27me3 of chromosome 7.

A) GC enrichment, B) Chromosome length, C) Amounts of repeats, D) Gene predications based on the JGI annotation, E & F) Coverage of H3K4me2 reads, G & H) Coverage of H3K9me3 reads, I & J) Coverage of H3K27me3 reads.

Chromatin structure after meiosis

To study how epigenetic marks are inherited, four additional *Z. tritici* strains were used including parental and progeny strains from a meiotic cross (Table 2.2.1). Progeny strains 01151 and 02133 are derived from a cross between IPO323 and IPO95052 performed on bread and durum wheat respectively (Wittenberg et al., 2009). Strain IPO94269 is an additional *Z. tritici* isolate collected in the Netherlands.

The genomes of IPO95052, IPO94269 and the progeny strain 02133 are not completely assembled and consist of many contigs (Kema, unpublished).

The genome of the progeny strain 01151 has not been sequenced, and the ChIP-seq reads obtained from 01151 were mapped to the genome of *Z. tritici* IPO323.

Table 2.1.2: Details of the genomes used for to study epigenetic inheritance.

Progeny strains 01151 and 02133 are derived from a crossing experiment between IPO323 and IPO95052 ((Wittenberg et al., 2009) & Kema, unpublished).

	IPO323	**IPO95052**	**IPO94269**	**01151 (323 x 95052)**	**02133 (323 x 95052)**
ID number of GFP-CenH3 strain	Zt121	Zt139	Zt137	Zt136	Zt138
Origin	The Netherlands	Algeria	The Netherlands		
Host plant	Bread wheat	Durum wheat	Bread wheat	Bread wheat	Durum wheat
# Chr	Δ Chr18	none	Δ Chr20	NA	Δ Chr15, 18 & 21
Contigs	21 chromosomes	3951	201.520	NA	3639

The amount of mapped reads after filtering and quality controls is summarized in table 2.1.3. Less reads were obtained from *IPO95052*. This is consistent in all four ChIP-seq experiments, either due to the MNase treatment or insufficient of input DNA.

Table 2.1.3: Summary of ChIP-seq data after quality filtering and mapping Illumina reads to the genomes. Three different antibodies were used in the ChIP-seq experiment. For each strain the reads were mapped to the specific genome, except for 01151 which was mapped to the genome of IPO323.

Strain	Replicate	Genotype	Antibody	Number of reads (x10 ^{6})	Number of mapped reads (x10 ^{6})	% mapped
Z. tritici IPO323	1	Wild type	H3K4me2	8.29	7.47	90.10
Z. tritici IPO323	2	Wild type	H3K4me2	6.41	5.81	90.59
Z. tritici IPO323	1	Wild type	H3K9me3	13.62	12.42	91.19
Z. tritici IPO323	2	Wild type	H3K9me3	5.88	5.29	90.00
Z. tritici IPO323	1	Wild type	H3K27me3	7.47	6.47	86.69
Z. tritici IPO323	2	Wild type	H3K27me3	7.06	6.40	90.72
Z. tritici IPO323 ΔKU70	1	ΔKU70 GFP-CenH3	H3K9ac	3.42	2.28	66.59
Z. tritici IPO323 ΔKU70	1	ΔKU70 GFP-CenH3	H3K9me3	0.99	0.69	69.75
Z. tritici IPO323 ΔKU70	2	ΔKU70 GFP-CenH3	H3K9me3	1.23	0.86	69.73
Z. tritici IPO323 GFP-CenH3	1	GFP-CenH3	H3K4me2	26.59	24.28	91.31
Z. tritici IPO323 GFP-CenH3	2	GFP-CenH3	H3K4me2	18.91	17.49	92.50
Z. tritici 01151 GFP-CenH3	1	GFP-CenH3	H3K4me2	11.26	9.71	86.28
Z. tritici 01151 GFP-CenH3	2	GFP-CenH3	H3K4me2	10.39	9.35	90.03
Z. tritici 02133 GFP-CenH3	1	GFP-CenH3	H3K9me3	8.92	3.02	33.80
Z. tritici 02133 GFP-CenH3	2	GFP-CenH3	H3K9me3	12.88	4.31	33.45
Z. tritici 02133 GFP-CenH3	1	GFP-CenH3	H3K27me3	8.82	5.09	57.73
Z. tritici 02133 GFP-CenH3	2	GFP-CenH3	H3K27me3	14.03	8.95	63.79
Z. tritici IPO94269 GFP-CenH3	1	GFP-CenH3	H3K4me2	21.76	19.75	90.76
Z. tritici IPO94269 GFP-CenH3	2	GFP-CenH3	H3K4me2	22.41	20.23	90.25
Z. tritici IPO95052 GFP-CenH3	1	GFP-CenH3	H3K9me3	5.94	1.35	22.73
Z. tritici IPO95052 GFP-CenH3	2	GFP-CenH3	H3K9me3	6.60	1.05	15.95
Z. tritici IPO95052 GFP-CenH3	1	GFP-CenH3	H3K27me3	12.81	2.13	16.60
Z. tritici IPO95052 GFP-CenH3	2	GFP-CenH3	H3K27me3	10.72	0.73	6.79

Figure 2.1.6 shows histone marks of chromosome 3 in the five *Z. tritici* strains. All histone modifications were mapped to the IPO323 genome. Apart from chromosome 3, also essential chromosome 7 and dispensable chromosome 21 were compared for histone modifications (Figure 2.1.7 and figure 2.1.8). The patterns of histone marks are conserved in the strains.

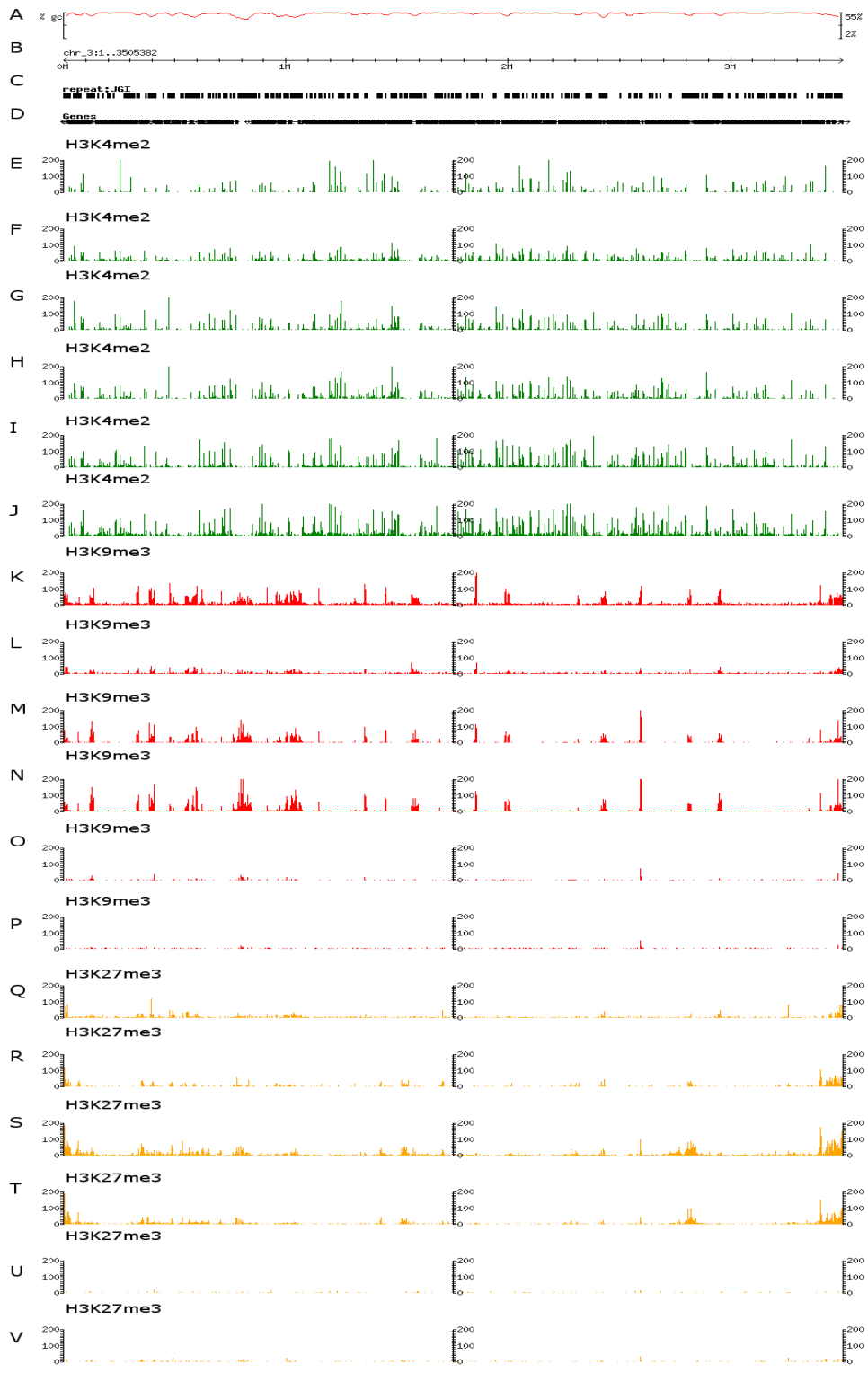

Figure 2.1.6: Chromatin modifications of chromosome 3.

All histone modifications were mapped to the IPO323 genome. A) GC enrichment, B) Chromosome length, C) Amounts of repeats, D) Gene predications based on the JGI annotation, E & F) Coverage of H3K4me2 reads for IPO323, G &H) Coverage of H3K4me2 reads for 01151, I & J) Coverage of H3K4me2 reads for IPO94269, K & L) Coverage of H3K9me3 reads for IPO323, M & N) Coverage of H3K9me3 reads for 02133, O & P) Coverage of H3K9me3 reads for IPO95052, Q & R) Coverage of H3K27me3 reads for IPO323, S & T) Coverage of H3K27me3 reads for 02133, U & V) Coverage of H3K27me3 reads for IPO95052.

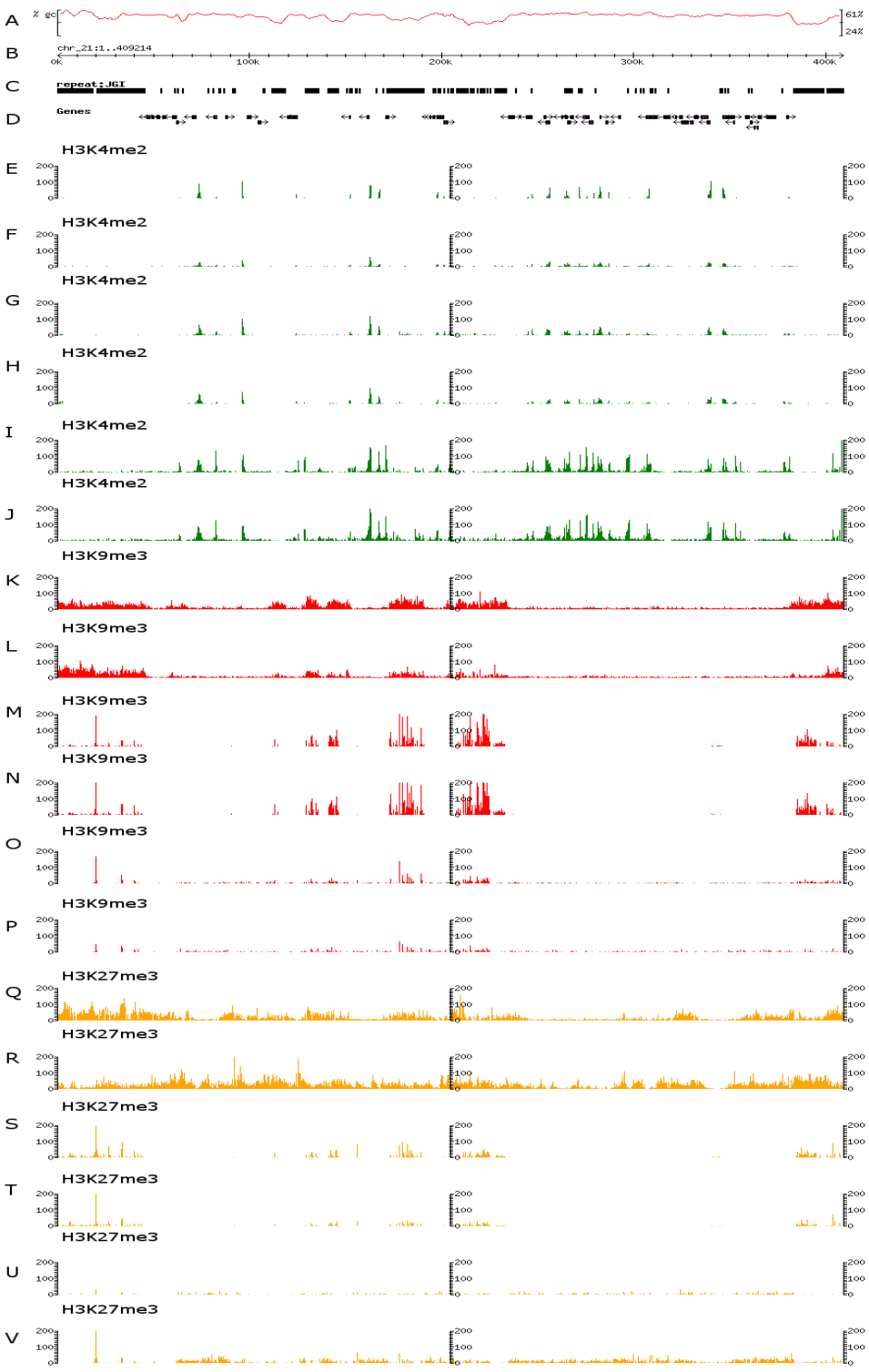

Figure 2.1.7: Chromatin modifications of chromosome 21.

All histone modifications were mapped to the IPO323 genome. A) GC enrichment, B) Chromosome length, C) Amounts of repeats, D) Gene predications based on the JGI annotation, E & F) Coverage of H3K4me2 reads for IPO323, G &H) Coverage of H3K4me2 reads for 01151, I & J) Coverage of H3K4me2 reads for IPO94269, K & L) Coverage of H3K9me3 reads for IPO323, M & N) Coverage of H3K9me3 reads for 02133, O & P) Coverage of H3K9me3 reads for IPO95052, Q & R) Coverage of H3K27me3 reads for IPO323, S & T) Coverage of H3K27me3 reads for 02133, U & V) Coverage of H3K27me3 reads for IPO95052.

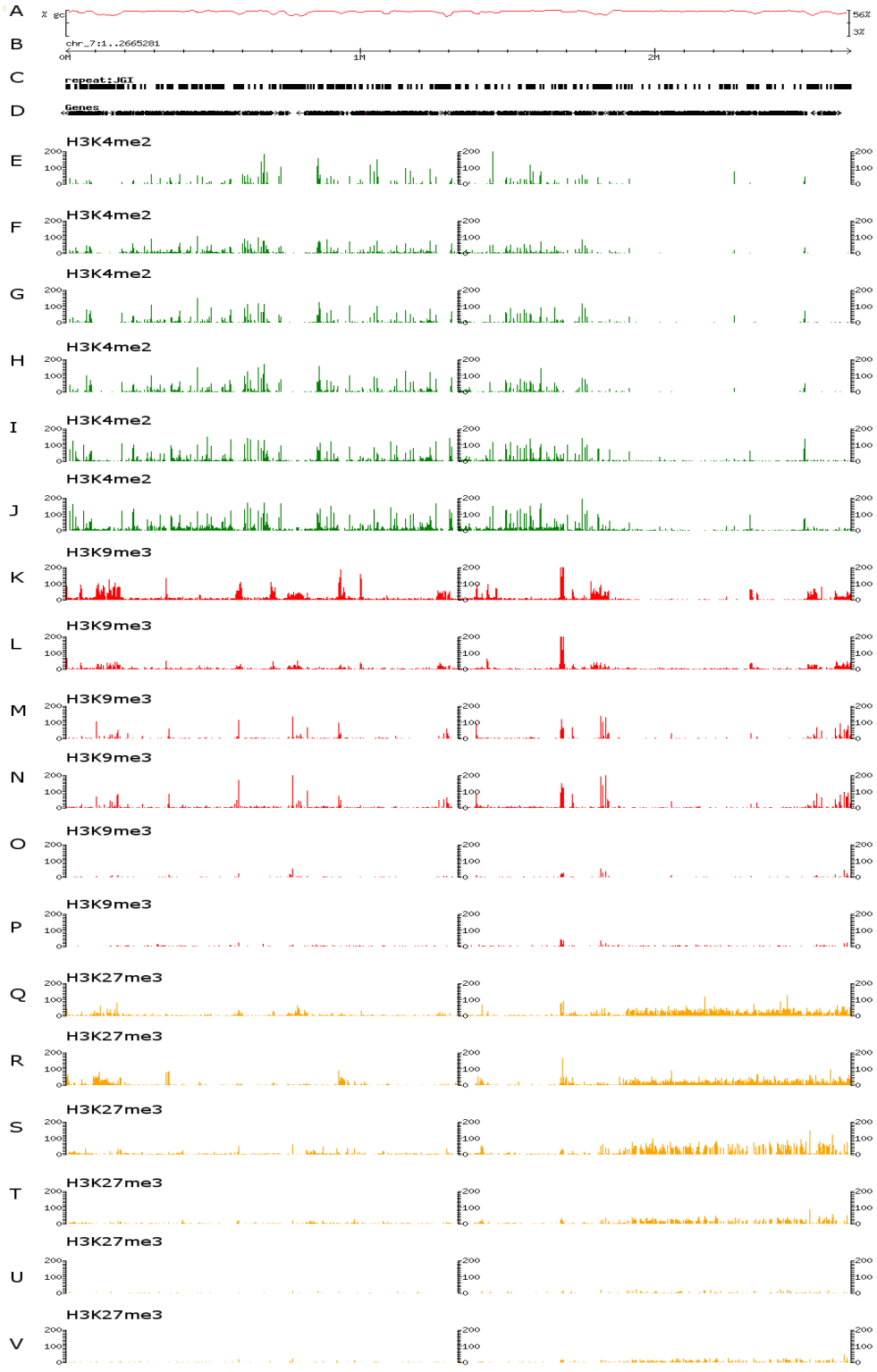

Figure 2.1.8: Chromatin modifications of chromosome 7.

All histone modifications were mapped to the IPO323 genome. A) GC enrichment, B) Chromosome length, C) Amounts of repeats, D) Gene predications based on the JGI annotation, E & F) Coverage of H3K4me2 reads for IPO323, G &H) Coverage of H3K4me2 reads for 01151, I & J) Coverage of H3K4me2 reads for IPO94269, K & L) Coverage of H3K9me3 reads for IPO323, M & N) Coverage of H3K9me3 reads for 02133, O & P) Coverage of H3K9me3 reads for IPO95052, Q & R) Coverage of H3K27me3 reads for IPO323, S & T) Coverage of H3K27me3 reads for 02133, U & V) Coverage of H3K27me3 reads for IPO95052.

Chapter 2.2: Characterization of centromeres in *Zymoseptoria tritici*

GFP-CenH3

To identify the centromeres of essential and dispensable chromosomes in *Zymoseptoria tritici,* several ChIP-seq based experiments were conducted.

First, a nucleotide Blast search using the CenH3 homolog of *Neurospora crassa* (NCU00145) as input resulted in the identification of the *Z. tritici* CenH3 in the genome of *Z. tritici* IPO323 (Smith et al., 2011) & (Goodwin et al., 2011). A binary vector, including the gene encoding the CenH3 protein with an N-terminal GFP tag, approximately one kb of up and down stream flanking sequence, and a hygromycin resistance cassette was created by overlap PCR. The GFP-CenH3 construct was integrated into two *Z. tritici* strains (IPO323 ΔChr18 and IPO323 ΔKu70 (Bowler et al., 2010) by *Agrobacterium tumefaciens* mediated transformation. The correct integration of the construct in both strains was verified by an initial PCR screen and subsequently Southern blot analyses. Further more Western blot analyses showed the correct translation of the GFP-CenH3 gene product (Data not shown). Finally, fluorescence microscopy confirmed that the GFP-CenH3 was localized inside the nucleus of *Z. tritici*.

In *N. crassa* and *Fusarium graminearum* the GFP tagged CenH3 protein is visible as one distinct bright spot (Smith, unpublished) & (Smith et al., 2012). This is in contrast to the CenH3 of *Z. tritici,* where multiple fluorescence foci inside the nucleus can be observed (Figure 2.2.1).

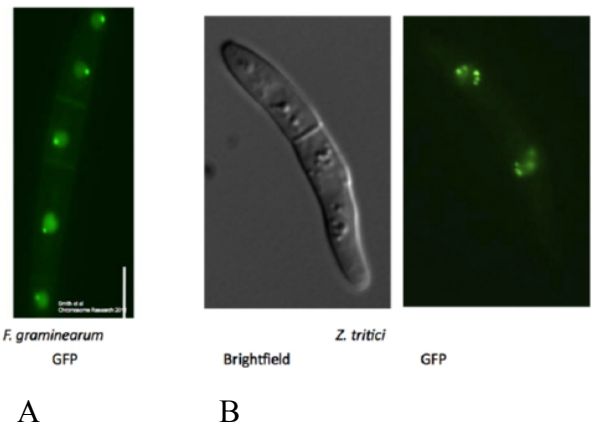

Figure 2.2.1: Localization of the CenH3 in *Z. tritici.*

A: In *F. graminearum* the CenH3 shows one bright spot inside the nucleus (Smith et al., 2012). **B**: Bright field and florescence microscopic picture of CenH3 in *Z. tritici* show multiple foci inside the nucleus.

An intensive screen with the fluorescence microscope showed that the number of foci in every cell differs and that the distribution of the foci is random and unorganized. To date, it is not clear what

causes the formation of multiple foci inside the nucleus of *Z. tritici*.

Mapping of the ChIP-seq of CenH3

ChIP experiments targeting the GFP-tag of the CenH3 were conducted. Prior to the ChIP, the DNA was fragmented by Mnase, and the ChIPed fragments were enriched by PCR. DNA fragments were finally sequenced using Illumina sequencing (Read length: 50 bp with an average amount of reads: 8.9 x 10^6 reads per sample). The reads were mapped using Bowtie or BWA and Tophat to the genome of *Z. tritici* IPO323 (table 2.21) (Langmead, 2010) & (Kim et al., 2013) & (Goodwin et al., 2011). Read mapping led to the identification of the centromere specific sequences of all chromosomes in *Z. tritici*. (Figure 2.2.2).

Characterization of centromeres

As shown in Figure 2.2.1, the centromeres are generally acrocentrically located close to the ends of the chromosomes. Based on the position of the centromeres, it is not possible to distinguish between the essential and dispensable chromosomes.

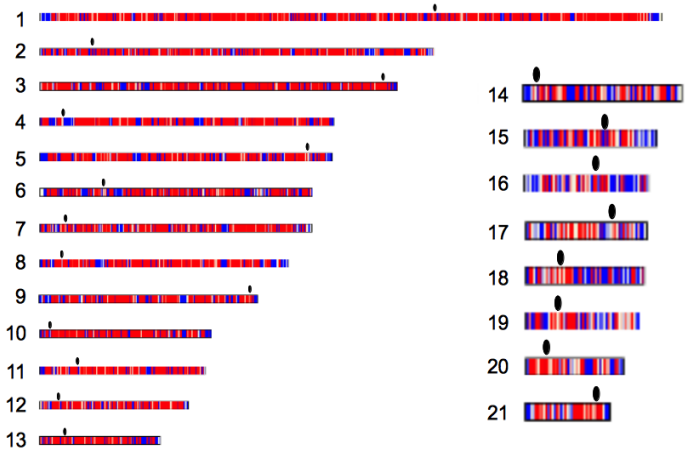

Figure 2.2.2: Centromeric positions on chromosomes.

The essential chromosomes (1-13) are displayed on scale, the dispensable chromosomes (14-21) are shown twice of scale. The blue blocks are repeat rich sequences and in red unique sequences are displayed. The centromeres are shown as black dots (Figure is modified from: (Goodwin et al., 2011)).

In comparison to other characterized fungal centromeres, the centromeres of *Z. tritici* are short, on average 9 kb (Figure 2.2.3 and Table 2.2.1). The largest centromere (14 kB) is found on chromosome 8. The smallest is only 5 kB and is located on chromosome 13. Hence, there is thereby no apparent correlation between chromosome size or centromere lengths.

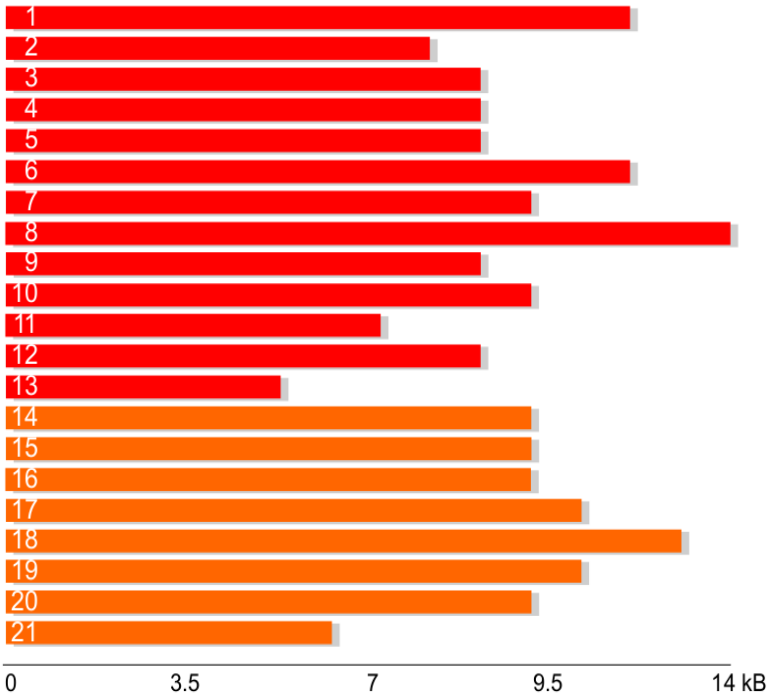

Figure 2.2.3: Length of the centromeres of *Z. tritici* IPO323.

The essential chromosomes (1-13) are displayed in red and the dispensable chromosomes (14-21) are displayed in orange. The x-axis shows the length of the centromeres (in kilo bases).

Compared to the genome-wide average AT content of 48.3%, the centromeres of *Z. tritici* are enriched with AT-sequences (average 52.5%) (Table 2.2.1). However, the centromeres are not located in the longest AT-rich region of the chromosome, as observed in several other fungal genomes. Rather, they are found in shorter AT-rich regions as shown in figure 2.2.5 and figure 2.2.6.

Table 2.2.1: Centromeric positions of *Z. tritici* IPO323.

Summary data for the 21 centromeres in the genome of IPO323 as identified by ChIP-seq.

		Chromosome length (mb) / AT content	Chromosomal position of centromere	Length of the centromere (kb)	AT content of the centromere	Repetitive DNA
Essential chromosomes	**Chr1**	6.09 / 46.9 %	Chr_1:3839000..3850999	12	55.7 %	27.26 %
	Chr2	3.86 / 47.6 %	Chr_2:513000..520999	8	55.2 %	62.36 %
	Chr3	3.51 / 47.4 %	Chr_3:3348000..3356999	9	49.9 %	0 %
	Chr4	2.88 / 47.8 %	Chr_4:217500..226499	9	49.8 %	19.11 %
	Chr5	2.68 / 48.0 %	Chr_5:2606000..2614999	9	57.8 %	22.03 %
	Chr6	2.68 / 48.6 %	Chr_6:625000..636999	12	53.6 %	41.97 %
	Chr7	2.67 / 47.2 %	Chr_7:257000..266999	10	48.4 %	2.81 %
	Chr8	2.44 / 48.3 %	Chr_8:214000..227999	14	52.9 %	38.42 %
	Chr9	2.14 / 48.5 %	Chr_9:2067000..2075999	9	53.0 %	0.47 %
	Chr10	1.68 / 47.5 %	Chr_10:99000..108999	10	52.7 %	21.94 %
	Chr11	1.62 / 47.2 %	Chr_11:366000..372999	7	62.9 %	71.77 %
	Chr12	1.46 / 47.7 %	Chr_12:179000..187999	9	48.7 %	11.47 %
	Chr13	1.19 / 48.0 %	Chr_13:237000..241999	5	47.7 %	0.56 %
Dispensable chromosomes	**Chr14**	0.77 / 51.5 %	Chr_14:60000..69999	10	46.5 %	52.78 %
	Chr15	0.64 / 49.0 %	Chr_15:384000..393999	10	51.3 %	25.61 %
	Chr16	0.61 / 48.5 %	Chr_16:332000..341999	10	58.6 %	42.01 %
	Chr17	0.58 / 48.0 %	Chr_17:407000..417999	11	57.3 %	50.85 %
	Chr18	0.57 / 50.8 %	Chr_18:159000..171999	13	51.9 %	46.62 %
	Chr19	0.55 / 48.7 %	Chr_19:148000..158999	11	54.9 %	2.48 %
	Chr20	0.47 / 48.5 %	Chr_20:95000..104999	10	46.0 %	13.21 %
	Chr21	0.41 / 48.1%	Chr_21:340000..346099	6	47.7 %	1.87 %

To confirm if the sequences were correctly assembled, the lengths and positions of all the centromeres were confirmed by PCR, using centromere specific primers designed to amplify fragments inside and outside of the centromeres (Table 2.2.2 and Figure 2.2.4). For all centromeres, three PCR reactions were performed. The first PCR was designed to amplify the whole centromere. In most of the PCR reactions this failed due to size limitations of the PCR reaction. The second and third PCR reactions amplify each half of the centromere. Between the two products there is a short overlap. Here, the PCR reactions for centromere 9 and 18 were unsuccessful. The PCR reactions for centromere 9 were repeated and the expected product was correctly amplified (data not shown). All other PCR reactions showed products of the predicted size, which led to the conclusion that the centromeres were correctly assembled. Besides the predicted band, in some PCR reactions multiple bands were amplified. This is might be due to repetitive DNA inside the centromeres and the enrichment for AT nucleotides.

Table 2.2.2: PCR reactions to confirm centromere assembly

PCR #	Centromere	Primer 1	Primer 2	Product (kb)	PCR #	Centromere	Primer 1	Primer 2	Product (kb)
1	Cen1	1192	1193	10.1	33	Cen11	1237	1236	1.9
2	Cen1	1192	1195	4.1	34	Cen12	1239	1240	8.7
3	Cen1	1194	1193	6.2	35	Cen12	1239	1242	4.7
4	Cen2	1197	1198	7.7	36	Cen12	1241	1240	4.1
5	Cen2	1197	1200	4.9	37	Cen13	1243	1244	4.8
6	Cen2	1199	1198	2.9	38	Cen13	1243	1246	2.0
7	Cen3	1201	1202	8.9	39	Cen13	1245	1244	2.9
8	Cen3	1201	1204	4.4	40	Cen14	1248	1249	9.5
9	Cen3	1203	1202	4.2	41	Cen14	1248	1251	4.3
10	Cen4	1205	1206	7.8	42	Cen14	1250	1249	5.4
11	Cen4	1205	1208	4.9	43	Cen15	1253	1254	9.8
12	Cen4	1207	1206	3.0	44	Cen15	1253	1256	5.2
13	Cen5	1209	1210	8.6	45	Cen15	1255	1254	4.8
14	Cen5	1209	1212	5.1	46	Cen16	1257	1258	9.3
15	Cen5	1211	1210	3.7	47	Cen16	1257	1260	4.9
16	Cen6	1214	1215	11.6	48	Cen16	1259	1258	4.5
17	Cen6	1214	1217	5.3	49	Cen17	1261	1262	10.8
16	Cen6	1214	1215	11.6	50	Cen17	1261	1264	5.1
19	Cen7	1218	1219	9.9	51	Cen17	1263	1262	5.9
20	Cen7	1218	1221	5.6	52	Cen18	1265	1266	12.1
21	Cen7	1220	1219	4.4	53	Cen18	1265	1268	5.9
22	Cen8	1222	1223	13.0	54	Cen18	1267	1266	6.2
23	Cen8	1222	1225	6.3	55	Cen19	1269	1270	9.8
24	Cen8	1224	1223	6.9	56	Cen19	1269	1272	4.2
25	Cen9	1227	1228	8.3	57	Cen19	1271	1270	5.8
26	Cen9	1227	1230	4.6	58	Cen20	1273	1274	9.3
27	Cen9	1229	1228	3.8	59	Cen20	1273	1276	3.8
28	Cen10	1231	1232	9.1	60	Cen20	1275	1274	5.6
29	Cen10	1231	1234	4.0	61	Cen21	1277	1278	5.4
30	Cen10	1233	1232	5.3	62	Cen21	1277	1280	2.7
31	Cen11	1235	1236	6.7	63	Cen21	1279	1278	2.8
32	Cen11	1235	1238	4.9					

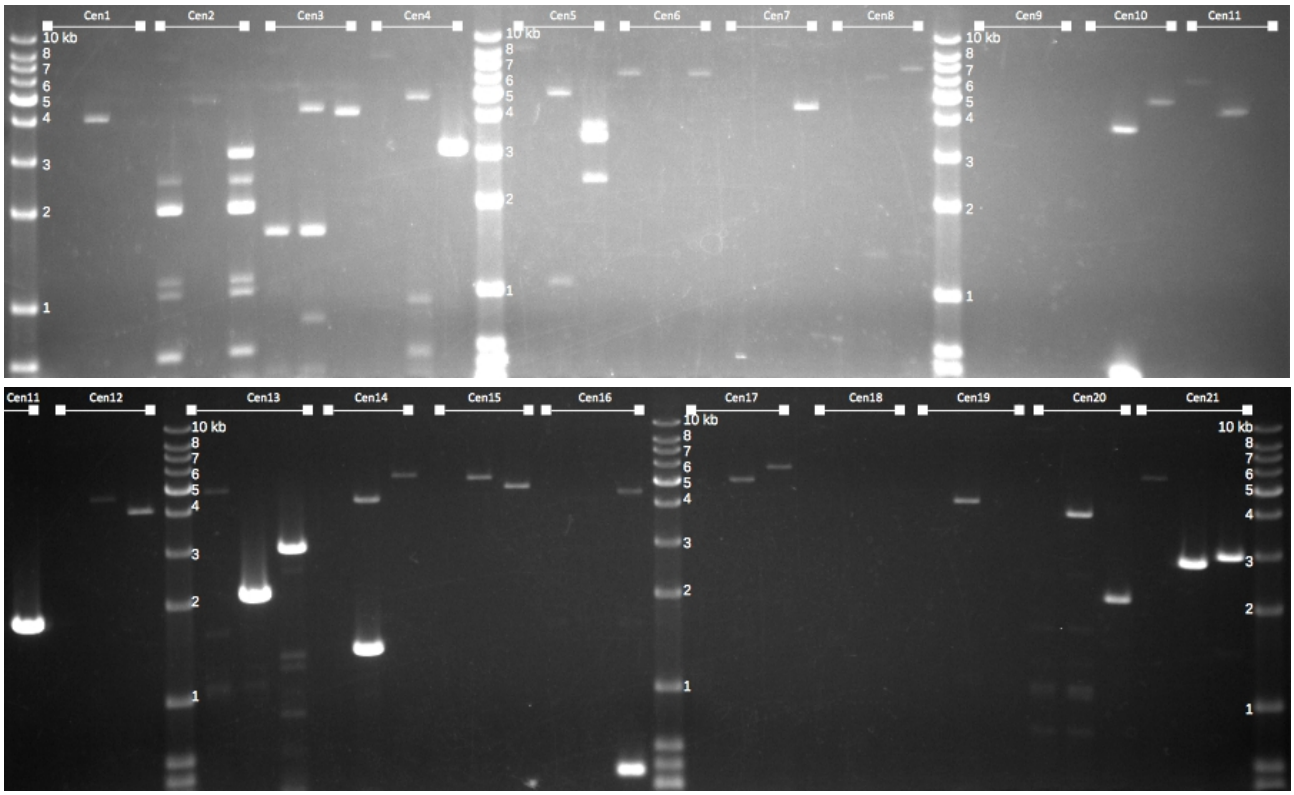

Figure 2.2.4: Conformation of the length the centromeres on gDNA by PCR analyzes.

In total three PCRs for each centromere are performed on gDNA. The 1st PCR overlaps the whole centromere, the 2nd PCR amplifies the left part of the centromere and the 3th PCR amplifies the right part of the centromere. There is a short overlap between the products of the second and third PCR.

For all chromosomes, the centromeric sequences lack any conserved structural domains and motives, but contain several repetitive sequences and paralogous sequences from other centromeres. Some of the centromeres lack completely homology to other centromeres, others have homologous sequences to one or more centromeres. The homologous sequences are all located in repetitive sequences (See the table in the supplementary data). From the essential chromosomes, the centromere of chromosome 3, 9, 10 and 13 have unique sequences without homology to any other loci, and of the dispensable chromosomes this is the case for chromosome 14, 15, 19, 20 and 21. Centromere 3 is the only centromere which has no repetitive DNA.

Based on the homology search, there is no clear trend in relation to the length of the centromere. The lack of conserved sequence motives and the observation that some centromeres consist entirely of unique sequences led to the conclusion that the centromeres of *Z. tritici* are not defined according to their sequences.

The centromeres of the essential and dispensable chromosomes show the same pattern of histone modifications, as analyzed in this study (see chapter 2.1). H3K4me2, H3K9me3 and H3K27me3 are absent in the centromeric regions. These modifications were characterized by ChIP experiments

with antibodies against the specific histone modifications. Subsequently, the associated DNA was processed and sequenced, like the centromere reads. As is shown in figure 2.2.5 and 2.2.6, H3K4me2 is absent inside the centromere, but present at the surrounding regions of the centromere. H3K9me3 is completely lacking both at the centromeric region and also in the surrounding regions. Also H3K27me3 is absent at the centromeres and surrounding regions of the essential chromosomes. In general, the dispensable chromosomes are enriched with H3K9me3 and H3K27me3, nevertheless the centromeric regions lack any enrichment of both histone modifications (See also chapter 2.1).

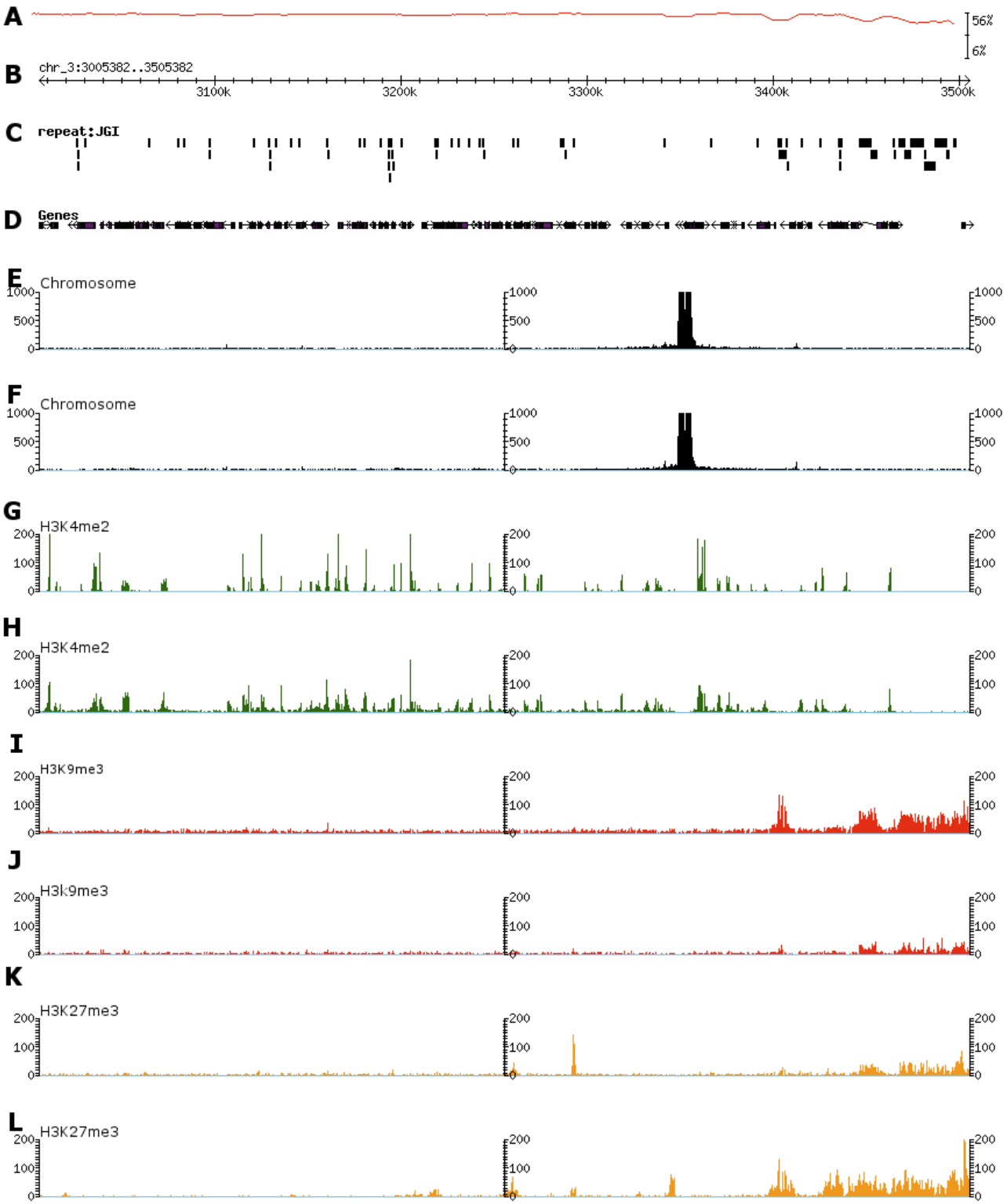

Figure 2.2.5: Centromeric position of the essential chromosome 3.

The figure shows 500kb of the centromeric and surrounding region of chromosome 3. Panel A shows the GC enrichment, the ruler shows the length of the chromosome and indicates the position of the specific reads. Panel C shows the amount of repeats and panel D shows the gene predications based on the JGI annotation. Panel E & F show the coverage of centromeric reads and panel G & H show the coverage H3K4me2 reads. Panel I & J show the coverage of H3K9me3 reads and panel K & L show the coverage of H3K27me3 reads.

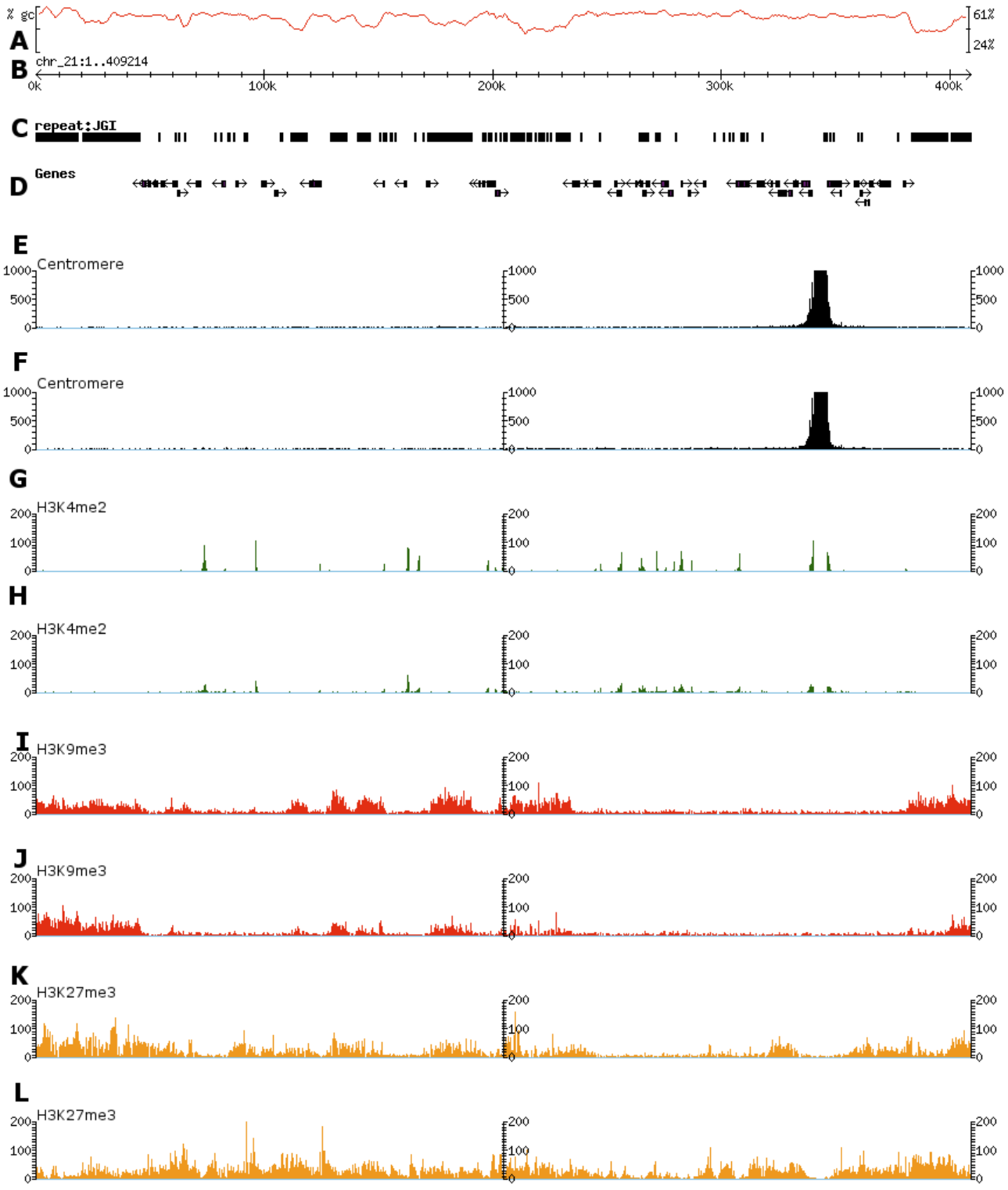

Figure 2.2.6: Centromeric position of the dispensable chromosome 21.

The figure shows chromosome 21. Panel A shows the GC enrichment, the ruler shows the length of the chromosome and indicates the position of the specific reads. Panel C shows the amount of repeats and panel D shows the gene predications based on the JGI annotation. Panel E & F show the coverage of centromeric reads and panel G & H show the coverage H3K4me2 reads. Panel I & J show the coverage of H3K9me3 reads and panel K & L show the coverage of H3K27me3 reads.

According to the JGI annotation, the centromeres of *Z. tritici* IPO323 contain sequences that encodes for 29 genes (Goodwin et al., 2011). Interestingly, 19 genes are expressed in axenic culture, or at four days post infection of *Triticum aestivum* (Kellner et al., 2014). Of the 19 genes, four are paralogous genes. Nine genes have a RPKM value higher than 5.0 in at least one of the two conditions which were tested. From the nine genes with a RPKM value higher than 5.0, three genes are paralogous. Table 2.2.3 shows the 19 genes with a RPKM higher than 1.0 in one of the two conditions tested.

Table 2.2.3: Gene expression of genes located in the centromeric regions.

The mean gene expression of 19 genes located in the centromeric regions of *Z. tritici* IPO323 are shown.

Centromere	Centromere position	Gene ID	Mean position of gene (bp)	Gene length (bp)	KOG Group	Mean RPKM at axenic culture	Mean RPKM at 4dpi	Paralogous genes
1	Chr_1:3839000..3850999	26410	3839964.5	761	NA	4.5	5.0	
1	Chr_1:3839000..3850999	31911	3841420	1250	NA	2.2	3.2	
1	Chr_1:3839000..3850999	89702	3843050	1284	NA	1.3	2.4	
3	Chr_3:3348000..3356999	70541	3351430.5	1057	NA	8.0	1785.7	1
3	Chr_3:3348000..3356999	92187	3354850.5	1797	NA	1.3	1.7	2
4	Chr_4:217500..226499	108931	223114.5	1417	NA	1.4	0.2	
5	Chr_5:2606000..2614999	72635	2613897.5	1631	METABOLISM	114.9	18.2	
7	Chr_7:257000..266999	45154	257182	1878	METABOLISM	3.5	9.1	1
7	Chr_7:257000..266999	100649	264078.5	503	NA	1.9	2.3	
7	Chr_7:257000..266999	105049	266459.5	781	CELLULAR PROCESSES AND SIGNALING	7342.9	7626.8	
8	Chr_8:214000..227999	81715	217848	550	NA	2453.3	2495.6	5
8	Chr_8:214000..227999	94908	222399	1340	NA	4.8	3.1	
8	Chr_8:214000..227999	94910	227745	722	NA	2.1	6.0	
10	Chr_10:99000..108999	105826	107790	850	NA	81.1	50.7	
12	Chr_12:179000..187999	96902	185635	2026	NA	0.5	1.3	
12	Chr_12:179000..187999	106260	179662.5	773	NA	37.8	5.7	
13	Chr_13:237000..241999	111601	241388	1744	NA	1.1	6.2	
20	Chr_20:95000..104999	98014	95240.5	533	NA	1.6	0.0	
20	Chr_20:95000..104999	98016	102461	826	NA	1.4	1.0	

Analysis of the ChIP-seq data obtained from the *Z. tritici* IPO323 GFP-CenH3 strain revealed that the *Z. tritici* IPO323 strain used in our laboratory lacks chromosome 18. This was confirmed with several PCR reactions wherein DNA from *Z. tritici* IPO323 and *Z. tritici* IPO323 ΔKu70 was used as a template (data not shown). The PCR reactions with the template of *Z. tritici* IPO323 failed in the amplification of chromosome 18 specific products. The same PCR reactions performed with *Z. tritici* IPO323 ΔKu70 however resulted in the predicted amplification products. The centromere of chromosome 18 was obtained from the *Z. tritici* IPO323 ΔKu70 (Figure 2.2.7). All other centromeres show the same position and sequence in these two lineages as expected.

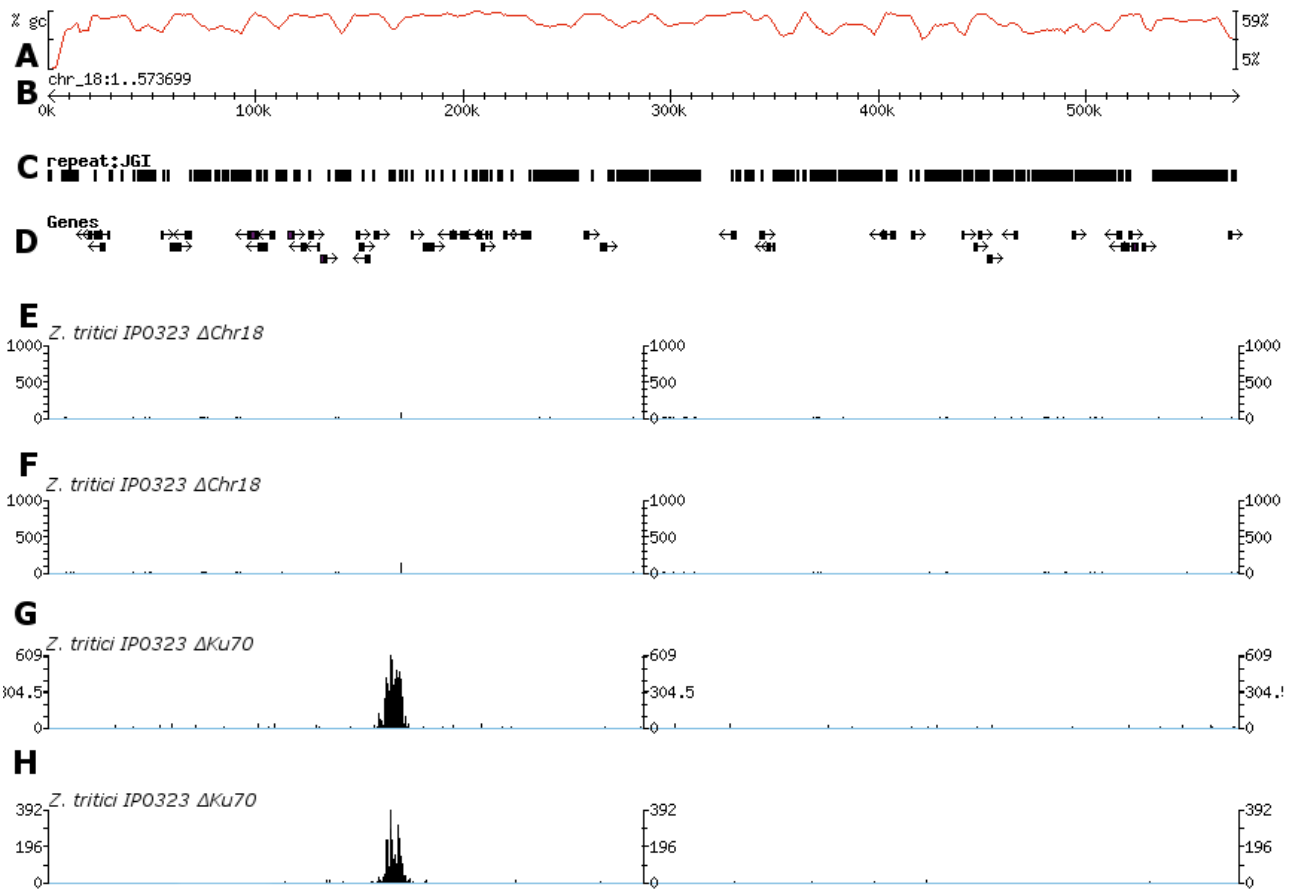

Figure 2.2.7: ChIP-seq reads from chromosome 18 of *Z. tritici* IPO323 ΔChr18 and *Z. tritici* IPO323 ΔKu70.
Panel A shows the GC enrichment, the ruler shows the length of the chromosome and indicates the position of the specific reads. Panel C shows the amount of repeats and panel D shows the gene predications based on the JGI annotation. Panel E & F show the coverage of centromeric reads of the *Z. tritici* IPO323 ΔChr18 strain. Panel G & H show the coverage of centromeric reads of the *Z. tritici* IPO323 ΔKu70 strain.

In addition to the CenH3 protein, one additional centromeric protein was tagged with GFP in *Z. tritici* IPO323. This was the Centromere B protein (CenB or Cenp-B in human) (Masumoto et al., 2004). Cenp-B was the only centromere specific protein in humans that binds to a DNA motive. The motive, the so-called B-box, has a length of 17 bp. In humans, Cenp-B recruits the CenH3 and contribute to the centromere assembly (Masumoto et al., 2004). Cenp-B is lacking at neocentromeres and at the centromere of the Y-chromosome (Masumoto et al., 2004). It was here hypothesized that the centromeres of essential and dispensable chromosomes could be differently assembled. The frequent rearrangement of dispensable chromosomes could cause a higher frequency of neocentromeres on these chromosomes (Croll et al., 2013) & (Mehrabi et al., 2007). To analyze this hypothesis, the CenB was tagged with GFP and subsequently ChIP-seq of the DNA associated to the CenB could be performed.

The CenB protein was C-terminal tagged with a GFP, and the correct integration of the CenB-GFP at the endogenous locus was confirmed with Southern blot analyses. The cells of the CenB-GFP strain were screened, with the fluorescence microscope, and this showed that the CenB protein of *Z. tritici* was located inside the nucleus. Typical foci, observed in the GFP-CenH3 strain, are however absent (figure 2.2.8). It is possible that CenB is a nuclear protein, that is not involved in the centromere assembly of *Z. tritici.* Therefore, ChIP-seq experiments with the CenB were not performed here.

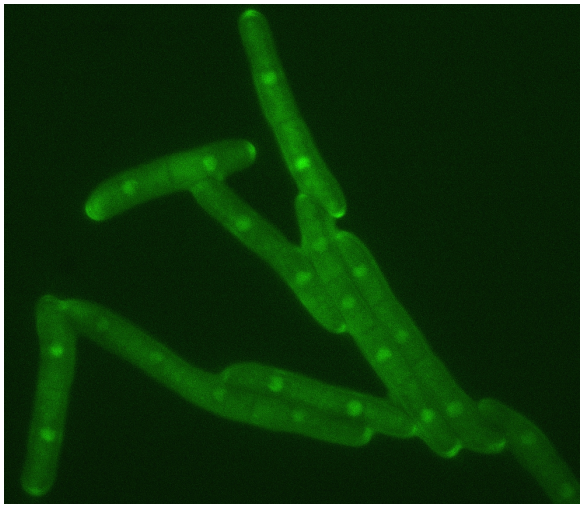

Figure 2.2.8: Localization of the CenB protein in *Z. tritici IPO323.*

A fluorescence microscopic picture shows that the CenB protein is localized inside the nucleus.

Centromeric behavior after meiosis

To study the behavior of the centromeres after meiosis, the GFP-CenH3 construct was introduced in four additional *Z. tritici* isolates that represent parents and progenies from a crossing experiment (table 2.1.2) (Wittenberg et al., 2009). ChIP-seq was used to identify the centromeres in these strains. Unlike the genome of *Z. tritici* IPO323*,* the genomes of IPO95052, IPO94269 and progeny strain 02133 are incompletely assembled and consist of many contigs (Table 2.1.2) (Kema, unpublished). The genome of the progeny strain 01151 has not been sequenced so far and the ChIP-seq reads of 01151 were therefore mapped to the genomes of both parents (*Z. tritici* IPO323 and *Z. tritici* IPO95052*)* (Table 2.2.4). After mapping of the reads, the centromeric sequences were extracted as consensus sequences. The statistics of the mapping are shown in table 2.2.5 and 2.2.7.

Table 2.2.4: Statistics of the ChIP-seq after mapping the reads to the genomes.

Summary data from ChIP-seq analyses, Illumina sequencing and Bowtie mapping of reads using *Z. tritici* IPO323 and four additional strains are shown. For 01511 the same reads were mapped to genomes of IPO323 (1) and IPO95052 (2).

Strain	Replicate	Antibody	Number of reads (x10 ^{6})	Number of mapped reads (x10 ^{6})	% mapped
Z. tritici IPO323 GFP-CenH3	1	GFP	14.64	12.74	87.0
Z. tritici IPO323 GFP-CenH3	2	GFP	15.90	13.22	83.1
Z. tritici IPO323 ΔKU70 GFP-CenH3	1	GFP	2.41	1.88	78.0
Z. tritici IPO323 ΔKU70 GFP-CenH3	2	GFP	0.69	0.57	83.4
Z. tritici 01151 GFP-CenH3	1 (1)	GFP	11.03	8.28	75.1
Z. tritici 01151 GFP-CenH3	2 (1)	GFP	5.52	4.20	76.2
Z. tritici 01151 GFP-CenH3	1 (2)	GFP	11.03	7.07	64.1
Z. tritici 01151 GFP-CenH3	2 (2)	GFP	5.52	3.82	69.2
Z. tritici 02133 GFP-CenH3	1	GFP	4.17	2.98	71.3
Z. tritici 02133 GFP-CenH3	2	GFP	4.79	3.45	72.1
Z. tritici IPO94269 GFP-CenH3	1	GFP	17.39	13.45	77.4
Z. tritici IPO94269 GFP-CenH3	2	GFP	25.18	21.19	84.2
Z. tritici IPO95052 GFP-CenH3	1	GFP	2.72	0.69	25.3
Z. tritici IPO95052 GFP-CenH3	2	GFP	3.23	0.16	5.0

It was possible to describe the centromeres of all strains. However the lower read outcome from IPO95052 resulted in lesser centromere coverage. The lower read amount of one of the *Z. tritici* IPO323 ΔKu70 strains is most likely due to the preparation method for this strain. Instead of MNase digestion, sonication was used to share the DNA.

Table 2.2.5: Centromeres of 01151 mapped to the parental genomes.

Summary of read mapping to 21 reference chromosomes of the parental genomes. NA: non applicable

		01151					
		Mapped to IPO323			Mapped to IPO95052		
	Chr	Length (kb)	AT (%)	N (%)	Length (kb)	AT (%)	N (%)
Essential chromosomes	**1**	14	57.7	0.0	11	51.4	4.6
	2	9	54.2	0.0	4	45.5	0.0
	3	9	50.0	0.0	9	49.3	1.1
	4	8	52.2	0.0	8	48.5	0.0
	5	9	56.4	0.4	10	42.7	1.0
	6	21	52.6	0.3	8	50.1	0.0
	7	10	48.7	0.0	9	48.9	0.0
	8	9	47.3	0.3	10	48.7	2.1
	9	9	53.1	0.0	9	51.1	2.6
	10	8	52.2	0.0	7	52.4	1.1
	11	17	51.6	0.2	7	49.5	3.7
	12	8	48.4	0.0	9	47.7	0.0
	13	9	45.9	0.2	8	46.3	0.0
Dispensable chromosomes	**14**	14	45.5	0.0	15	45.4	0.7
	15	6	47.5	0.1	9	49.9	1.1
	16	12	56.9	0.0	8	48.8	0.7
	17	14	48.8	0.1	14	57.1	1.0
	18	NA	NA	NA	NA	NA	NA
	19	10	54.2	0.0	12	47.9	0.4
	20	10	45.9	0.0	11	46.0	0.0
	21	6	47.9	0.0	6	47.8	0.0

To analyze from which parent the centromeres were inherited, alignments for the progeny and parental centromeres were created (Table 2.2.5) (Figure 2.2.9). As is shown in Figure 2.2.9, for centromere 3, there is a clear distinction between the two groups of centromeres. Alignments for other centromeres show the same clear distinct pattern, except for centromere 2, 5, 15, 18 and 21 of 01151 and centromere 2 and 21 of 02133 were the inherency could not be assessed based on the alignments (data not shown).

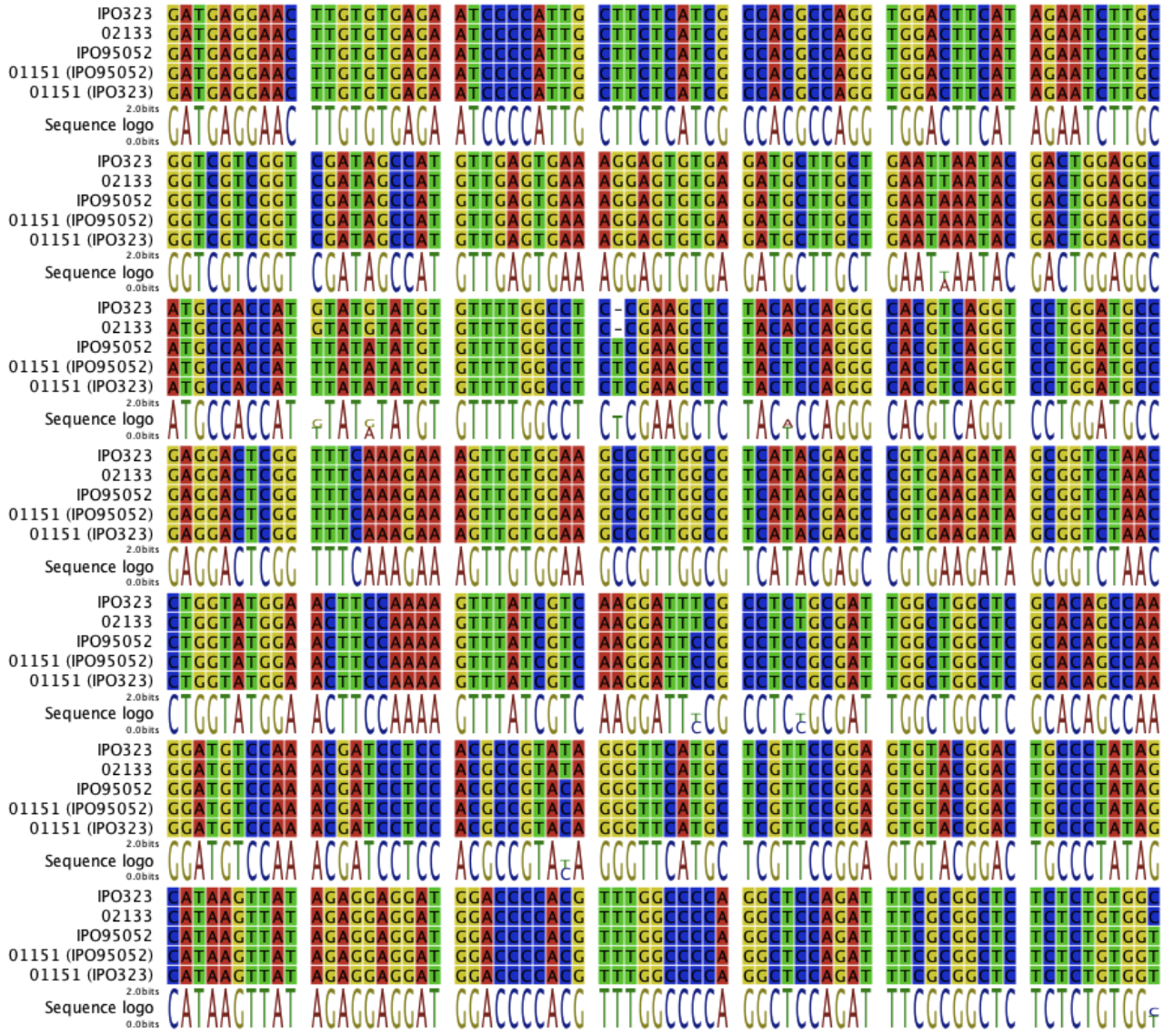

Figure 2.2.9: Detail from an alignment with centromere 3 from progeny and parental strains.

A short region of the alignment of centromere 3 is shown. The sequence logo shows the nucleotide conservation. IPO323 and progeny strain 02133 share the same centromeric sequence. While centromere 3 from progeny strain 01151 is derived from IPO95052.

Table 2.2.6: Centromeres are inherent from one parent.

For each centromere it is shown from with parental centromere they derive. NA: non applicable. Absent: Chromosome is not present in the genome of the progeny strain.

	Chr	01151	02133
Essential chromosomes	**1**	IPO323	IPO95052
	2	NA	NA
	3	IPO95052	IPO323
	4	IPO95052	IPO95052
	5	NA	IPO95052
	6	IPO323	IPO323
	7	IPO95052	IPO95052
	8	IPO95052	IPO95052
	9	IPO323	IPO323
	10	IPO323	IPO323
	11	IPO323	IPO323
	12	IPO323	IPO323
	13	IPO95052	IPO95052
Dispensable chromosomes	**14**	IPO323	IPO323
	15	NA	Absent
	16	IPO323	IPO95052
	17	IPO95052	IPO323
	18	NA	Absent
	19	IPO95052	IPO95052
	20	NA	IPO323
	21	IPO323	NA

For a few centromeres the inheritance was not clear due to problems in the assembly of the centromeric reads. For centromere 2 this is most likely due to the high amount of repetitive DNA (Cen2 of IPO323: 62.36 %). For 01151 the centromeres were mapped to both parental genomes. For centromere 2, 5, 15, 18 and 20 this resulted in centromeres which were biased for the parental genome and as result of this both centromeres differ. Based on the inheritance of the centromeres (Table 2.2.6), the centromeres of progeny strain 01151 mapped to correct parental genome were chosen (table 2.2.5) and used in table 2.2.7.

Table 2.2.7: Centromeres of four additionally *Z. tritici* strains to study centromeric behavior after meiosis. Summary of read mapping to 21 reference chromosomes. NA: non applicable

		IPO323		01151			IPO94269			02133			IPO95052		
	Chr	Length (kb)	AT (%)	Length (kb)	AT (%)	N (%)	Length (kb)	AT (%)	N (%)	Length (kb)	AT (%)	N (%)	Length (kb)	AT (%)	N (%)
Essential chromosomes	**1**	12	55.7	14	57.7	NA	5	48.1	NA	10	54.9	2.5	17	50.0	3.6
	2	8	55.2	NA	NA	NA	9	48.0	0.2	10	47.6	0.7	NA	NA	NA
	3	9	49.9	9	49.3	1.1	9	49.7	NA	9	49.4	0.4	10	48.3	2.3
	4	9	49.8	8	48.5	NA	9	47.9	NA	6	64.8	4.6	NA	NA	NA
	5	9	57.8	NA	NA	NA	10	50.5	NA	8	46.5	0.4	12	47.6	1.0
	6	12	53.6	21	52.6	0.3	4	47.8	NA	7	52.0	4.4	15	47.7	1.6
	7	10	48.4	9	48.9	NA	10	45.8	0.1	11	47.9	0.4	NA	NA	NA
	8	14	52.9	10	48.7	2.1	11	49.4	NA	8	50.2	2.2	8	48.4	3.5
	9	9	53.0	9	53.1	NA	7	54.4	NA	8	46.3	1.7	8	49.3	3.3
	10	10	52.7	8	52.2	NA	14	49.4	NA	7	49.9	2.6	8	51.1	1.5
	11	7	62.9	17	51.6	0.2	7	51.3	0.3	7	49.7	2.4	11	46.8	4.5
	12	9	48.7	8	48.4	NA	9	47.9	NA	8	47.7	2.1	7	47.6	1.1
	13	5	47.7	8	46.3	NA	7	46.6	NA	6	46.8	NA	NA	NA	NA
Dispensable chromosomes	**14**	10	46.5	14	45.5	NA	14	45.7	NA	12	46.6	1.8	13	45.7	3.8
	15	10	51.3	NA	NA	NA	14	47.3	NA	Absent	NA	NA	8	52.6	4.1
	16	10	58.6	12	56.9	NA	11	49.4	NA	7	50.5	0.6	13	49.0	1.5
	17	11	57.3	14	57.1	1.0	7	47.3	NA	14	57.5	0.8	8	54.1	3.5
	18	13	51.9	NA	NA	NA	NA	NA	NA	Absent	NA	NA	NA	NA	NA
	19	11	54.9	12	47.9	0.4	11	49.5	NA	10	47.3	NA	13	52.8	2.8
	20	10	46.0	NA	NA	NA	Absent	NA	NA	8	46.3	0.1	NA	NA	NA
	21	6	47.7	6	47.9	NA	12	46.3	NA	Absent	NA	NA	NA	NA	NA

The overall length of the centromeres is not conserved and varies between the five *Z. tritici* strains. Centromere 3 has a minor length difference of 1kb and centromere 2, 7, 9, 12 and 19 have lengths which range within 2 kb. Centromere 4 and 13 have length differences between the strains of 3 kb. The other centromeres have a larger length variation between all the strains. However, some chromosomes show extreme variation in the length of the centromeres, especially the centromeres of progeny strain 01151. The centromere of chromosomes 6 and 11 have a large gap inside the centromere. The sequences which surround the gap show enrichment for the GFP-reads and as result of this both surrounding regions and the gap are considered as centromere. The centromere of chromosome 4 of 01151 is located at the end of the scaffold, as also seen in centromeres 1, 6 and 17 of IPO94269. The variation in length could be due to assembly problems.

Table 2.2.8 shows the difference in length and AT-content between the progeny centromere and the parental centromere from which the progeny centromere was inherited.

Table 2.2.8: Difference in length of the centromere and the AT percentage between the progeny and parental centromeres.

Summary of differences between progeny and parental centromeres (Table 2.2.6). NA: non applicable. Absent: the chromosome is not present in the strain.

	Chr	01151		02133	
		Difference in Length (kb)	Difference in AT (%)	Difference in Length (kb)	Difference in AT (%)
Essential chromosomes	**1**	1.6	2	7.4	4.9
	2	NA	NA	NA	NA
	3	0.4	1	0.2	0.5
	4	NA	NA	NA	NA
	5	NA	NA	3.2	1.1
	6	8.7	1	4.9	1.6
	7	NA	NA	NA	NA
	8	1.6	0.3	0	1.8
	9	0.1	0.1	0.8	6.7
	10	2.5	0.5	3	2.8
	11	9.7	11.3	0.2	13.2
	12	0.7	0.3	1.4	1
	13	NA	NA	NA	NA
Dispensable chromosomes	**14**	3.9	1	1.9	0.2
	15	NA	NA	Absent	Absent
	16	1.8	1.7	5.4	1.5
	17	5.6	3	2.7	0.2
	18	NA	NA	Absent	Absent
	19	1.2	4.9	3.1	5.5
	20	NA	NA	2.3	0.3
	21	0.3	0.2	Absent	Absent

Alignments of sequences of the centromere and the surrounding region were created with Muscle (Edgar, 2004). The aligned sequences are approximately 25 kb and the centromere is located in the centre of each sequence. Figure 2.2.10 show the alignments of the centromeres of chromosome 3, 16 and 17. In general sequences of the centromeres and surrounding regions are conserved.

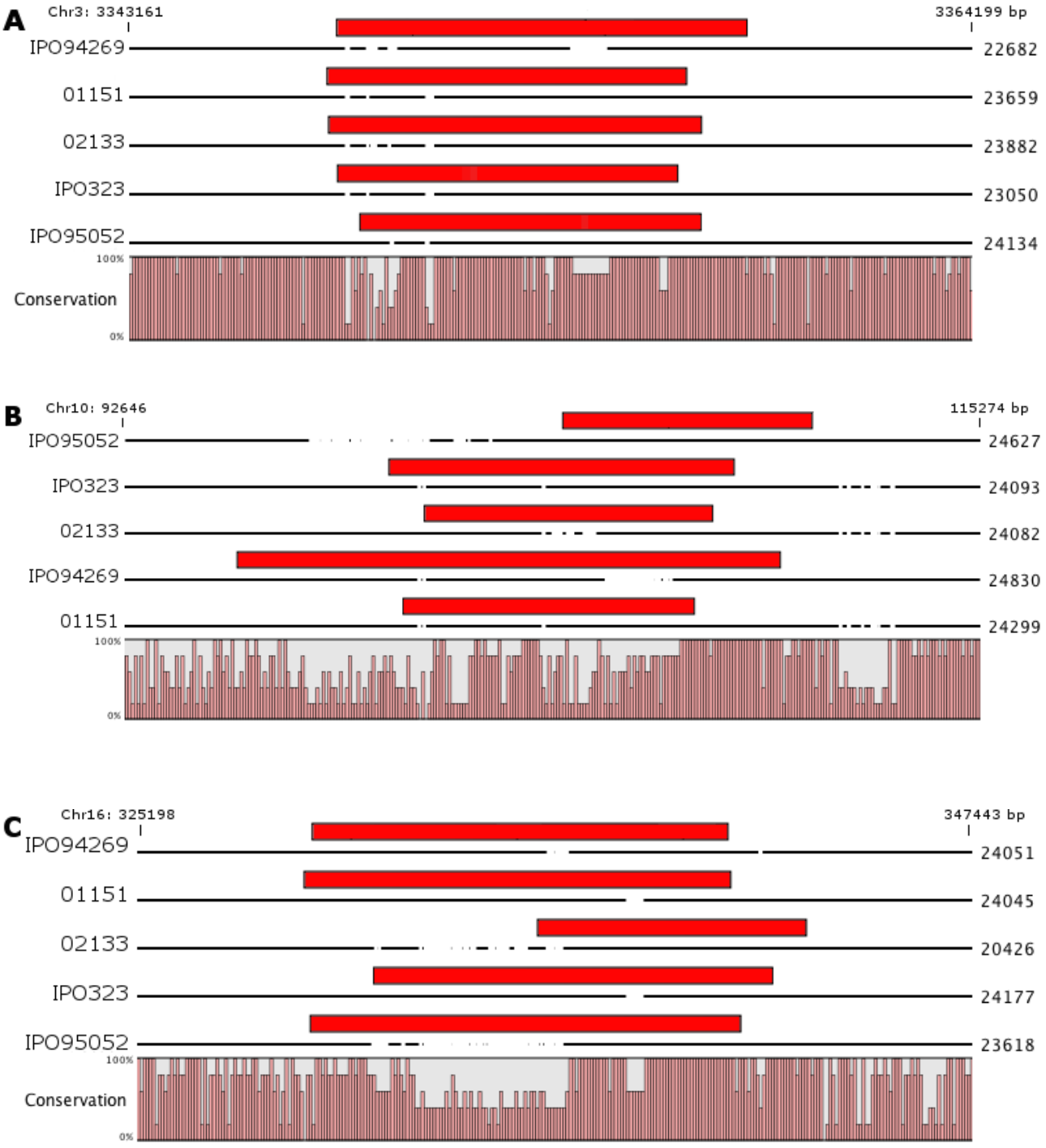

Figure 2.2.10: Alignment of the centromere and surrounding regions of chromosome 3, 10 and 16.

The centromere alignments were created with Muscle. The centromere is shown with a red rectangle and the corresponding strain number is shown before each line. The length of the sequence is displayed at the end of each sequence and the ruler on top of each alignment shows the chromosomal position of IPO323. The conservation plot displays similarity at each position of the alignment. The centromere of progeny strain 01151 was obtained from direct mapping to the genome IPO323. **A)** Centromere and surrounding regions of chromosome 3. **B)** Centromere and surrounding regions of chromosome 10. **C)** Centromere and surrounding regions of chromosome 16.

Neocentromere formation

Neocentromere formation was also investigated in this study. Since frequent structural changes affect the dispensable chromosomes, a hypothesis was that new centromeres are formed in rearranged chromosomal fragments without existing centromeres.

To induce neocentromere formation, an experiment was designed to delete the centromeres in two chromosomes. To delete a centromere, the sequence of the centromere should be replaced with the antibiotic Nourseotricin by homologous recombination. To study neocentromere formation of both essential and dispensable chromosomes, the centromeres of chromosome 13 and 14 were the focus of this study. All of the experiments were performed with the GFP-CenH3 *Z. tritici* IPO323 strain, allowing the use of a GFP antibody agains the CenH3 in a ChIP-seq experiments. This would to subsequently allow us to identify the new centromeres.

For the deletion of the centromere of chromosome 13, a total of 106 mutants were screened. All isolates however still contained the centromere of the chromosome 13. We conclude that the deletion of the centromere of chromosome 13 was lethal to the fungus in this experiment. For the deletion of the centromere of chromosome 14 a total of 745 mutants were screened. Here PCR analysis suggested that 23 isolates had lost the centromere of chromosome 14. Southern blot analysis however confirmed the correct deletion of the centromere of chromosome 14 for only one isolate. However in the second Southern blot, the wild-type bands were still visible (Figure 2.2.11).

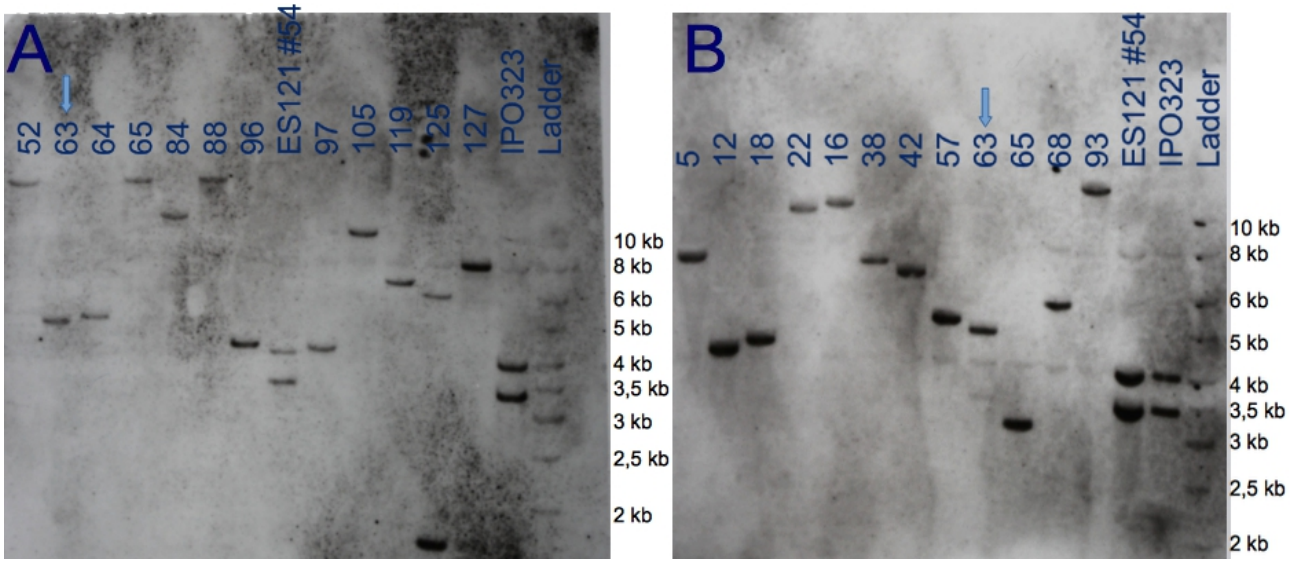

Figure 2.2.11: Southern blot analyses of centromere 14 mutants.

DNA was digested with EcoRV and the probe based on the up-stream and down-steam region of the centromere deletion construct. **A**. The wild-type (*Z. tritici* IPO323 and *Z. tritici* IPO323 GPF-CenH3 (ES121#54)) shows the expected band of 3.4 and 5.0 kb while the strain #63 (marked with the arrow) has a 5.0 kb band that corresponds to the expected size by centromere deletion. **B.** Blot shows additional mutant strains. Sample #63 (marked with the arrow) was included.

To confirm if chromosome 14 is still present after the deletion of the centromere, primer pairs spending eight regions along chromosome 14 were designed. Both wild-type controls (*Z. tritici* IPO323 ΔChr18 and *Z. tritici* IPO323 ΔChr18 GFP-CenH3) and the candidate strain 63 showed the same pattern, indicating that chromosome 14 is still present in the strain with the Δcentromere 14. The same PCR reactions were also performed on 22 transformed strains which showed deviating bands in the Southern blot analysis. Here, the PCR experiments show that the strains have lost the whole chromosome 14 (table 2.2.9). The strains which have lost chromosome 14, however, still contain the resistance against Nourseotricin.

Table 2.2.9: Results from a PCR screen for presence/absence of chromosome 14.

The positive amplification (√) of DNA from chromosome 14 indicates the presence of the chromosome in the strains. *Z. tritici IPO323* and *Z. tritici* IPO323-GFP-CenH3 (ZT121#54) (used as positive control). Isolate 63 is the by Southern blot confirmed Δ centromere 14 isolate.

	295x296	1580x1581	1582x1583	286x285	1584x1585	1586x1587	291x292	159x158
52								
63	√	√	√	√		√	√	
64				√				
65				√				
84				√				
88				√				
96				√				
121#54	√	√	√	√	√	√		√
97				√				
105				√				
119				√				
125				√				
106								
IPO323	√	√	√	√	√	√		√
5				√				
12				√				
18				√				
22				√				
26				√				
38				√				
42				√				
57				√				
65				√				
68								
93				√				

ChIP experiments with the candidate strain 63 were performed. Currently the samples are in sequencing.

Centromeres of closely related species

The centromeres of the two closely related species *Z. pseudotritici* and *Z. ardabiliae* are also investigated here to compare centromere evolution across species boundaries (Stukenbrock et al., 2011). To identify the centromeres of these two species, *Z. pseudotritici* (Zp13) and *Z. ardabiliae (*Za17) were transformed with the GFP-CenH3 construct designed for *Z. tritici*. The correct integration of the GFP-CenH3 construct was confirmed by PCR. Western blot analyses confirmed the correct translation of the integrated GFP-CenH3.

As in *Z. tritici* multiple foci inside the nucleus were observed by fluorescence microscopy (Figure 2.2.12).

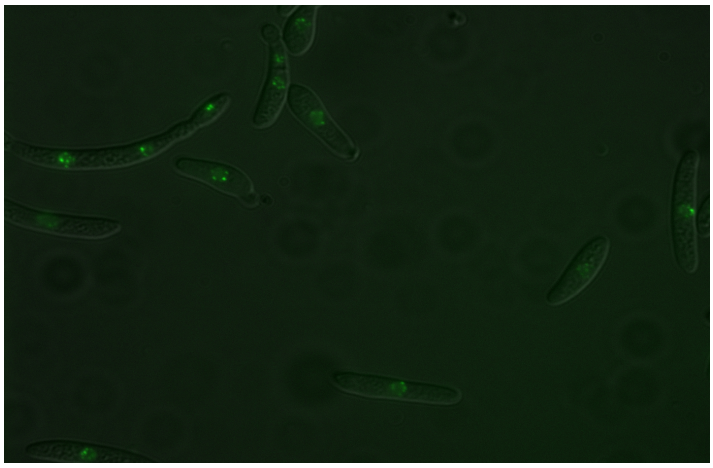

Figure 2.2.12: Localization of the GFP-CenH3 in *Z. ardabiliae.*

Overlay from a bright field and florescence microscopic picture of CenH3 in *Z. ardabiliae* shows multiple foci inside the nucleus. *Z. pseudotritici* shows the same pattern (data not shown).

ChIP experiments were performed and currently the sequencing of the ChIPed samples is in progress.

To confirm if the centromeres are conserved among *Z. pseudotritici* (Zp13) and *Z. ardabiliae* (Za17), a BlastN search with the centromeres of *Z. tritici* IPO323 as query was performed (Stukenbrock et al., 2011) The results are summarized in table 2.2.10. The overall sequence identity of *Z. pseudotritici* (Zp13) to *Z. tritici* IPO323 is 93-94% and the overall sequence identity of *Z. ardabiliae* (Za17) to *Z. tritici* IPO323 is 90-91% (Stukenbrock et al., 2011). On average the sequence identity of the centromeres of *Z. pseudotritici* (Zp13) to the centromeres of *Z. tritici* IPO323 is higher than to the sequence identity of the centromeres of *Z. ardabiliae* (Za17). Nevertheless, almost all centromeric sequences of *Z. pseudotritici* (Zp13) and *Z. ardabiliae* (Za17) have a lower sequence identity to *Z. tritici* IPO323 than the average genome sequence identity.

Table 2.2.10: BlastN search in the *Z. pseudotritici* (Zp13) and *Z. ardabiliae* (Za17) genomes using the centromeres of *Z. tritici* IPO323 as query.

Shown is the conservation of centromeres based on length of BlastN hit (> 1kb) and % sequence identity to *Z. tritici* IPO323. When more than one hit, the total homology is shown and the average % sequence identity are shown. The number of homologous regions are shown between the brackets. See for the isolate numbers in table 5.3.2.

Cen	*Z. pseudotritici* (Zp13)	*Z. ardabiliae* (Za17)
	Hit length (bp) / sequence %identity	Hit length (bp) / sequence %identity
1	5255 / 91.96% (2)	5472 / 75.97% (2)
2	2303 / 88.45% (2)	NA
3	6265 / 90.25% (1)	5425 / 80.84% (2)
4	NA	2479 / 74.97% (2)
5	1702 / 91.74% (1)	3528 / 85.13% (2)
6	NA	4639 / 85.52% (1)
7	NA	2183 / 73.34% (2)
8	1191 / 87.37 (1)	1980 / 84.78% (2)
9	NA	3715 / 77.30% (3)
10	1157 / 89,62 % (1)	1035 / 71.14% (1)
11	NA	NA
12	NA	NA
13	1308 / 91.67% (1)	1202 / 90.46% (1)
14	NA	NA
15	NA	1216 / 71.35% (1)
16	NA	NA
17	2857 / 87.09% (2)	NA
18	1361 / 89.31% (1)	3341 / 90.75% (1)
19	NA	NA
20	NA	2582 / 75.47% (2)

Chapter 2.3: Functional and phylogenetic characterization of proteins involved in the RNAi-pathways

To understand the importance of RNAi for fungal growth and pathogenicity, experiments to investigate the functional role of genes encoding proteins of the RNAi machinery were here performed.

Identification of RNAi components in *Zymoseptoria tritici*

One Dicer-like (Dcl) and four Argonaute (Ago) encoding genes were identified by a blast search in the JGI database using homologs (both protein and nucleotide) of Dicer (-like) and Argonaute genes of *Neurospora crassa* (Dcl-1/Sms-3 & Dcl-2 and Qde-2/Sms-1), *Drosophila melanogaster* and *Caenorhabditis elegans* as input (Goodwin et al., 2011) & (Galagan et al., 2003) & (Adams, 2000) & (Hillier et al., 2005). Furthermore, keyword searches for the functional domains of Dicer and Argonaute proteins: Piwi, PAZ, RNAIII and helicase were performed (Ghildiyal and Zamore, 2009). The results of the Blast and domain searches are listed in table 2.3.1.

Table 2.3.1: Genes involved in the RNAi-pathway of *Z. tritici* IPO323.

Gene	Gene-ID	Chromosomal position	Description	Functional domains
Dcl	47983	chr_9:1660747-1665410	Dicer-like	Helicase, Duf283, 2x RNAIII domain, Double-stranded RNA binding motif
Ago1	90232	chr_1:5667619-5671029	Argonaute	Duf1785, PAZ, Piwi
Ago2	38035	chr_3:2630540-2633932	Argonaute	Duf1785, PAZ, Piwi
Ago3	10621	chr_11:940973-943399	Argonaute	Duf1785, PAZ, Piwi
Ago4	25632	chr_1:4995449-4996504	Argonaute	Duf1785, Piwi

In addition to the Dcl and Ago genes, several other proteins, like the RNA-dependent RNA polymerase (RdRP) and Hen1, also involved in the RNAi-pathway, were identified (Ghildiyal and Zamore, 2009). These proteins were not studied further here, but add support to the presence of RNAi in *Z. tritici*.

Expression of the Dcl and Ago genes in *Z. tritici* IPO323

The wheat (*Triticum aestivum*) cultivar Obelisk was infected with the *Z. tritici* reference isolate IPO323 and RNA was isolated at 4, 11, 16 and 28 days post infection (dpi) from two independent replicates. For each replicate, RNA was isolated from 3 leaves. The infection process developed as expected and the symptoms were screened macroscopically. Subsequently, cDNA was synthesized to determine relative expression. cDNA isolated from axenic culture was used as a control to compare gene regulation during infection.

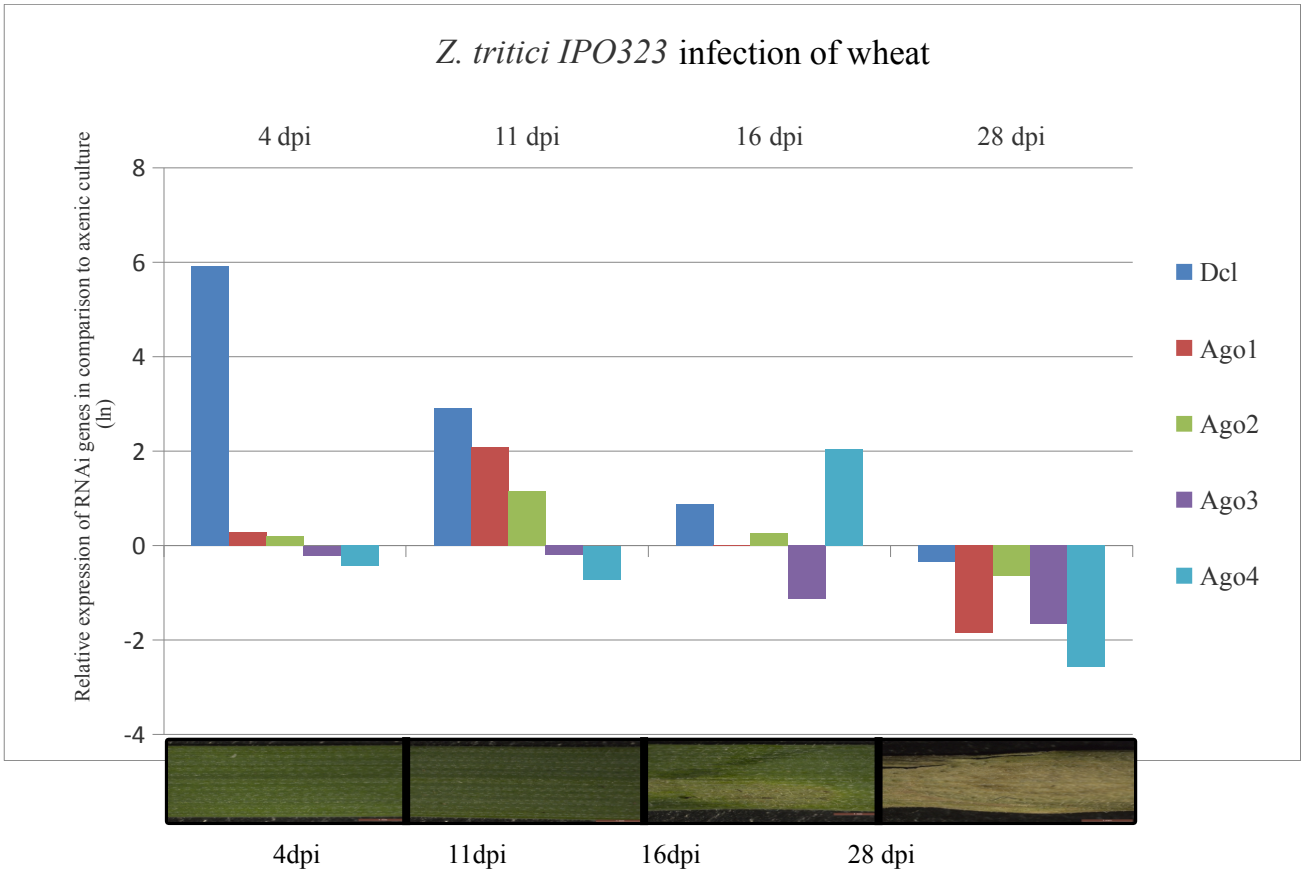

Figure 2.3.1: Expression of RNAi related genes during infection of wheat.

Relative expression of Dcl and Ago genes of *Z. tritici IPO323* during an infection of *T. aestivum* in comparison to gene expression on axenic culture. (The values are show as natural logarithm). Bellow the graph, *T. aestivum* leaves infected with *Z. tritici IPO323* isolated at four time-points during the host infection.

As shown in figure 2.3.1 *Z. tritici IPO323* first grows symptomless, during biotrophic growth and shows the first symptoms of fungal infection shortly after 11 dpi. At 16 dpi, chlorosis, necrosis and pycnidia become visible. At 28 dpi, the leaf only has death tissue and the pycnidia are matured and visible as dark black spots. The fungus grows saprotrophically in the dead plant tissue.

In figure 2.3.1, the relative expression (as natural logarithm) of the genes in comparison to the gene expression on axenic culture is shown. Compared to the GAPDH expression, none of the RNAi genes are up-regulated in axenic culture. The Dcl gene is significantly up-regulated during early infection and the expression is decreased during late infection. Ago1 and Ago2 are up-regulated at 4 dpi and the expression is considerably increased at 11 dpi. This is in contrast to Ago3 which, compared to GAPDH, is down regulated during all five time points. Ago4 is also down regulated during all stages of infection, except at 16 dpi were it is highly up-regulated. At 28 dpi during saprotrophic growth all genes are down regulated, consistent with expression in the axenic culture.

The role of Dcl and the Ago genes during plant infection of *Z. tritici*.

To study the role of RNAi during plant infection, single mutants of the Dcl, Ago1, Ago2, Ago3 and Ago4 were created by replacing the ORF with a hygromycin resistance cassette. Mutants were confirmed by Southern blot analysis and the deletion loci were sequenced to confirm the correct replacement of each gene with a hygromycin resistance cassette. To assay the virulence of the mutants *dcl, ago1, ago2* and *ago3*, 40 leaves of the wheat (*T. aestivum*) cultivar Obelisk were infected double blind (see material and methods) with either a mutant or wild type. A mock infection with water was used as control (12 leaves). 28 days post infection, the leaves were screened for necrosis levels and pycnidia abundance. The resulting phenotypes are schematically shown in figure 2.3.2.

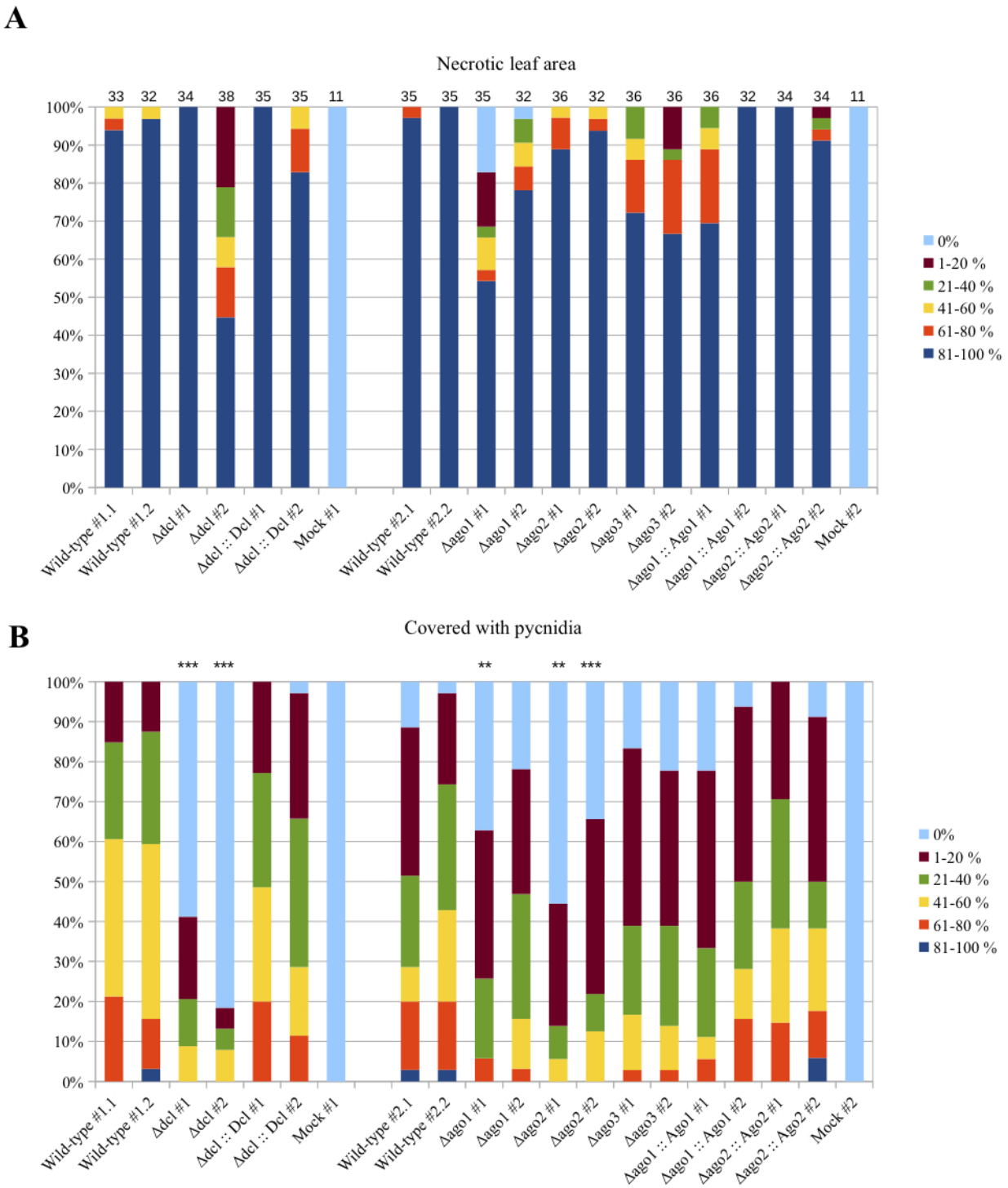

Figure 2.3.2: Phenotypes from virulence assay of Dcl mutants and wild-type strain of *Z. tritici*.

Results 28 days post infection with *Z. tritici IPO323 on T. aestivum* leaves. The values are shown in six categories, where 0% means that symptoms are lacking and 100% corresponds to a complete coverage of the leaf area. The experiment was performed with two replicates. **A**: Shows the percentage of necrosis of the leaf area. The amount of leaves is displayed above the each isolate. **B**: Shows the amount of pycnidia in the same marked area. Significant levels are shown in table 2.3.2.

To verify if the mutant phenotypes deviate significantly from the wild-type a student T-test for the amount pycnidia was performed (Table 2.3.2).

Table 2.3.2: P-values of a Students T-test.

Dicer	Between wild-types	Δ*dcl* #1	Δ*dcl* #2	Δ*dcl* :: Dcl #1	Δ*dcl* :: Dcl #2
Wild-type 1	0.819957	1.47E-006	5.55E-016	0.741274	0.741274
Wild-type 2		2.76E-006	2.86E-015	0.906450	0.906450

Argonaute	Between wild-types	Δ*ago1* #1	Δ*ago1* #2	Δ*ago2* #1	Δ*ago2* #2	Δ*ago3* #1	Δ*ago3* #2	Δ*ago1* :: Ago1 #1	Δ*ago1* :: Ago1 #2	Δ*ago2* :: Ago2 #1	Δ*ago2* :: Ago2 #2
Wild-type 1	0.1657	2.25E-006	0.004330	6.8E-006	8.99E-009	0.000303	0.000664	0.097313	0.000165	0.574565	0.309989
Wild-type 2		0.002214	0.204853	0.004175	5.53E-005	0.051849	0.082183	0.830716	0.033726	0.358442	0.751896

All mutants were significantly reduced in virulence. However one replicate of *ago1, ago2* and *ago3* were less avirulent when compared to wild-type 2. Clearly, more replicates are required to further document the observed phenotypes of the mutants. For the restoration of the wild-type, we inserted the wild-type gene in the endogenous locus.
Wild-type could be restored except of one Δ*ago1* :: Ago1 strain which was reduced in virulence similar to the mutant. The other replicate was however pathogenic at a level similar to the wild-type.

As displayed in figure 2.3.2, an infection with the wild type (IPO323) results in leaves which are almost completely necrotic and have an average pycnidia formation covering more than 30% of the infected region. All the mutants show necrosis levels similar to the wild type. However, the pycnidia formation is significantly reduced in the mutants.

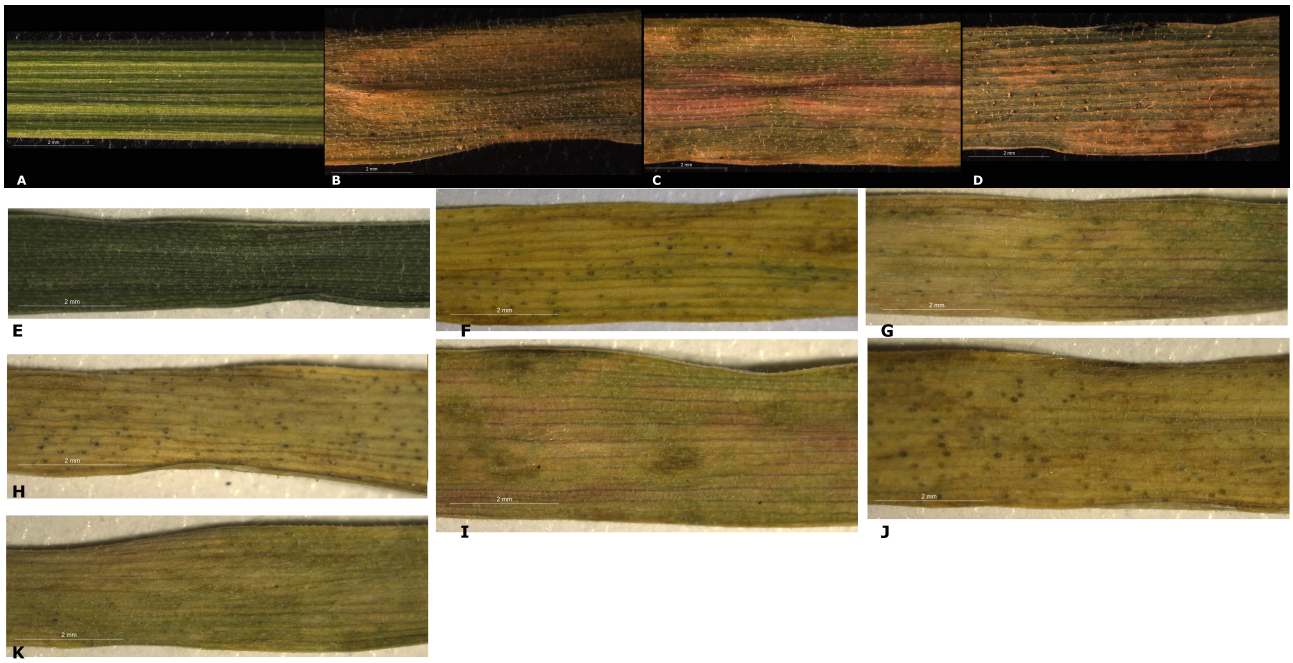

Figure 2.3.3: Macroscopic phenotypes of wild-type, *dcl, ago1, ago2* and *ago3* infection.

Macroscopic pictures of leaves at 28 dpi infected with: A) Water control, B) Wild-type, C) Δ*dcl*, D) Δ*dcl* :: Dcl, E) Water control, F) Wild-type, G) Δ*ago1*, H) Δ*ago1* :: Ago1, I) Δ*ago2*, J) Δ*ago2* :: Ago2, K) Δ*ago3*.

Sequence conservation of RNAi genes in the *Zymoseptoria* clade.

To assess the extend of sequence conservation in the RNAi genes, sequences of the five genes were obtained from different *Z. tritici* isolates and compared with the *Z. tritici* reference strain IPO323. The genomes of 14 additional *Z. tritici* isolates were sequenced by Illumina sequencing (J. Grandaubert and E. Stukenbrock, unpublished) & (G. Kema, unpublished). Unlike the genome of IPO323, these genomes have not been fully assembled and consist of scaffolds. To identify the homologs of the RNAi related genes, a Blast search, at the nucleotide level, using the IPO323 genes as template was performed. The results are summarized in table 2.3.3.

Table 2.3.3: Homologs of RNAi-genes in *Z. tritici* isolates

Gene conservation based on BlastN results when using homologous genes of IPO323 as query. √: present in genome; -: absent in genome of the respective isolate. Information about the isolate numbers are listed in table 5.3.2 in the supplement material.

	IPO323	Zt02	Zt05	Zt07	Zt10	Zt11	Zt128	Zt129	Zt130	Zt148	Zt150	Zt151	Zt153	Zt154	Zt155
Dcl	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Ago1	√	√	-	√	-	-	√	√	√	-	√	-	-	-	-
Ago2	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Ago3	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Ago4	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√

The Dcl sequences of all strains are conserved at the nucleotide level. A sequence %identity of minimal 97.62% and the full length (4.66 kb) sequences were found. Ago1 has a percentage of sequence identity of at least 99.82% in seven strains. The full length of Ago1 of IPO323 was found in the seven strains (3.41 kb). However Ago1 is absent in eight of the *Z. tritici* isolates. Remarkably, the flanking region of Ago1 are still conserved and present in all 15 isolates. To confirm that Ago1 is absent in the eight isolates, and to eliminate possible errors in the sequence assembly, primers in the conserved flanks were designed to conduct primer-walking and sequencing in the two isolates Zt10 and Zt11. This resulted in the identification of a highly distinct allele of Ago1 in both Zt-10 and Zt-11. A BlastN search with the Ago1 allele of Zt10 as query, resulted in the identification of Ago1 Zt10 homologs in the genomes of the eight strains were the Ago1 allele of IPO323 was not found. The eight newly identified Ago1 Zt10 alleles are conserved among themselves with the full length of 3.41 kb and with a percentage sequence identity of 99.7%.

Ago2 has a sequence identity of at least 85 % among the *Z. tritici* isolates and is conserved for the full length of the ORF, except for the Ago2 of Zt05, Zt11, Zt148, Zt151, Zt154 and Zt155 which have two gaps of 91 and 185 bp in the gene. The 91 bp region is a region without a functional domain. However the 185 bp gap span a Piwi domain. The sequence has so far not been resequenced in the isolates and it cannot be excluded that the gaps originate from assembly errors.

In all, 15 genomes Ago3 was identified with a sequence identity of 75.92 % and a full length ORF of 2.87 kb.

Ago4 has a sequence identity of 98.6% between Zt02, Zt05, Zt07, Zt10, Zt11, Zt150 and Zt155. For the other isolates, the full length ORF could not be found. As mentioned above this may be due to problems in the assembly of the sequences.

Based on the BlastN searches, the other RNAi genes are conserved at the nucleotide level. To confirm this, alignments and nucleotide diversity plots were created for the Dcl and Ago1-4. The results are shown in figures 2.3.4.

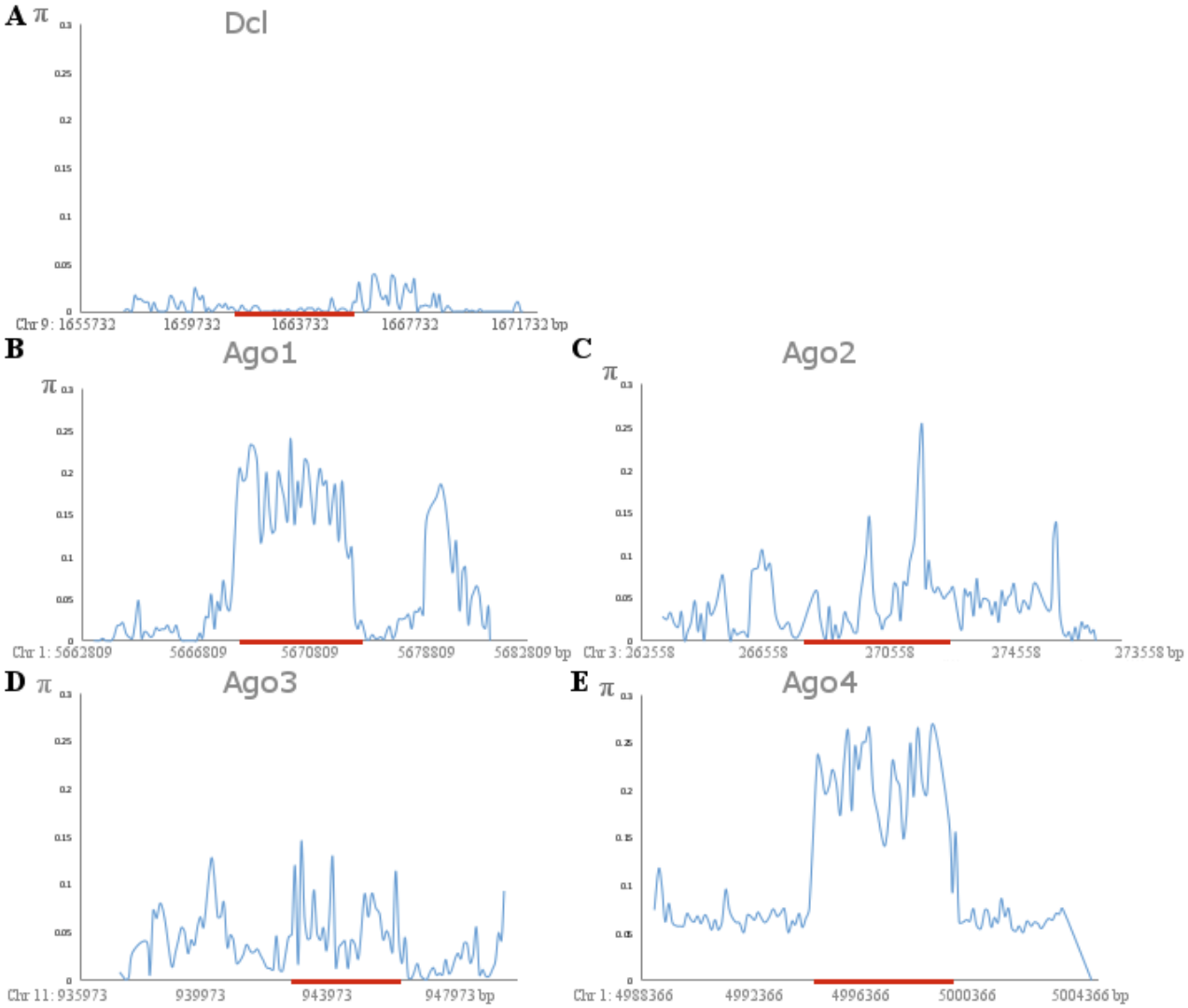

Figure 2.3.4: Nuclear diversity plot and alignment of Dcl, Ago1-4 and 5 kb surrounding regions.

The alignments are based on the sequences of 15 *Z. tritici* isolates with 5 kb surrounding regions up and downstream of the ORF. The x-axis shows the chromosome coordinates of IPO323 and the ORF of IPO323 is marked in red. The diagram shows variation of nucleotide diversity (π) in windows of 100 bp with a step size of 100. A) Dcl, B) Ago1, C) Ago2, D) Ago3 and E) Ago4.

Both the ORF and the surrounding regions of Dcl are conserved in all 15 isolates (Figure 2.3.4.A). The alignment of Ago1 shows a highly variable region between isolates. The nucleotide diversity measured as π was plotted across the ORF and the up and downstream flanking regions (Figure 2.3.4.B). The conservation and nucleotide diversity plots confirm the results of the BlastN search and show that the Ago1 has two alleles which are equally distributed through the 15 isolates tested. All the ORFs of Ago2 shows similar levels of nucleotide diversity compared to Ago1 (Figure 2.3.4.C). However, the surrounding regions of Ago2 are less conserved than the surrounding region of both Dcl and Ago1. The allelic difference of the ORF is considerable lower than for Ago1.

The ORF of Ago3 shows similar levels of nucleotide diversity as the surrounding regions (Figure 2.3.4.D). This is in contrast to the nucleotide diversity plot were both the flanks and the ORF of Ago3 shows similar levels of nucleotide diversity.

Ago4, shows a similar extraordinary pattern as Ago1. The surrounding regions are conserved while the ORF of Ago4 shows two highly diverged allele-groups (Figure 2.3.4.E).

Functional characterization of distinct Ago1 alleles in two Iranian *Z. tritici* isolates

To understand the functional implications of sequence divergence of Ago1, we set out to investigate the role of Ago1 in two Iranian *Z. tritici* isolates, Zt10 and Zt11 (Stukenbrock et al., 2011).

The wheat (*T. aestivum*) cultivar Obelisk was infected with the *Z. tritici* isolates Zt10 and Zt11 and RNA was isolated at 4, 11, 16 and 28 dpi from two independent replicates. For each replicate, RNA was isolated from 2 leaves. The infection process developed as expected and the symptoms were screened macroscopically. cDNA was synthesized to determine relative expression of the RNAi genes. cDNA isolated from axenic culture was used as a control to compare gene regulation during infection.

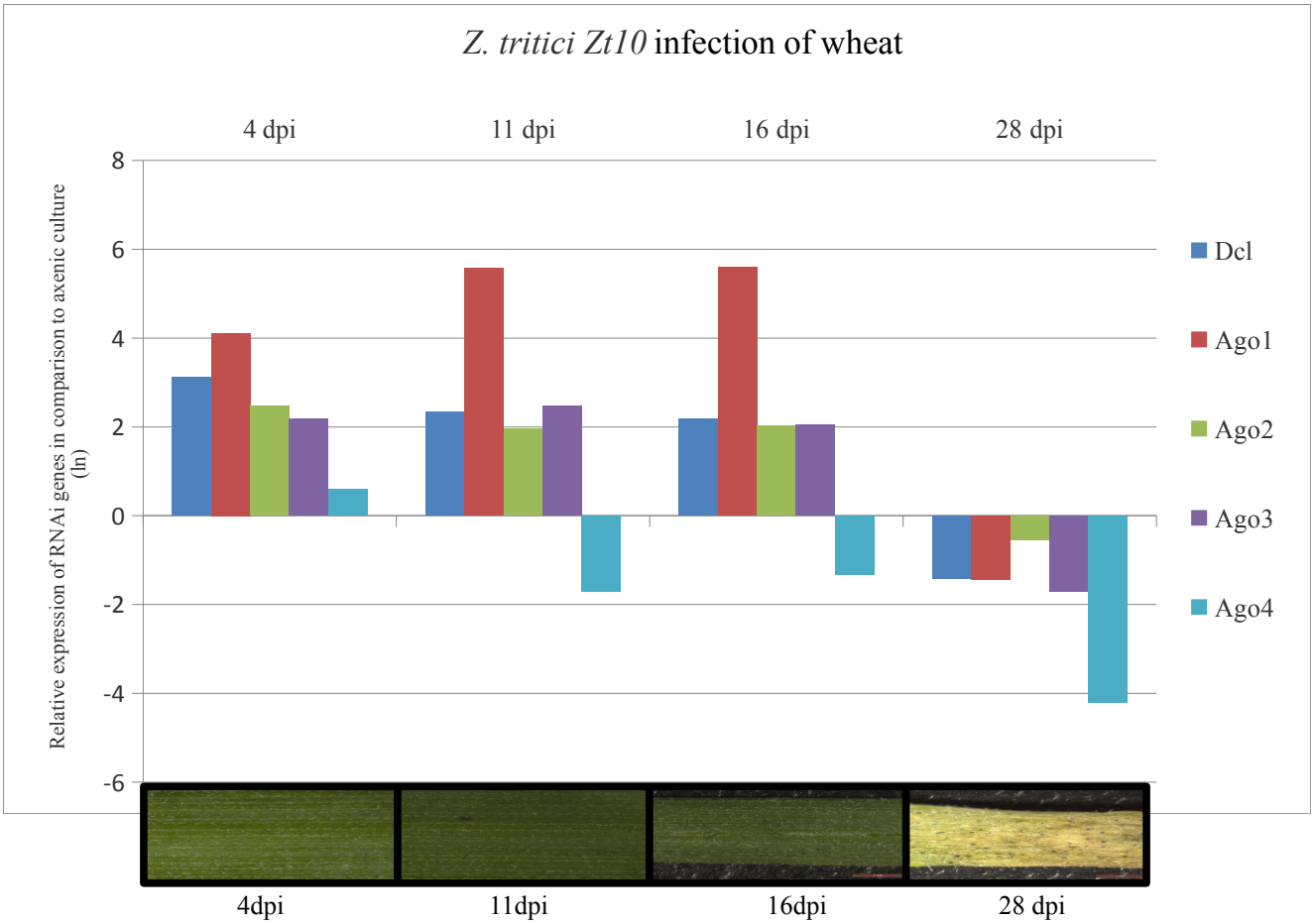

Figure 2.3.5: Relative expression of Dicer-like and Argonaute genes of *Z. tritici* Zt10 during an infection of *T. aestivum* in comparison to gene expression on axenic culture. (The values are show as natural logarithm). Below the graph: *T.*

aestivum leaves infected with *Z. tritici* Zt10 isolated at 4 time-points during the life-cycle.

As for the RNAi genes in the reference isolate IPO323, all genes were down regulated in axenic media. Dcl, Ago1, Ago2 and Ago3 are up-regulated at 4, 11 and 16 dpi and later down regulated at 28 dpi (Figure 2.3.5). However Ago4 is only up-regulated at 4 dpi. The symptoms of Zt10 appear between 11 and 16 dpi which is consistent with its hemibiotrophic lifestyle switch. At 16 dpi, chlorosis is induced and the pycnidia are visible. At 28 dpi, the leaf has only death tissue and the pycnidia are mature (Figure 2.3.5).

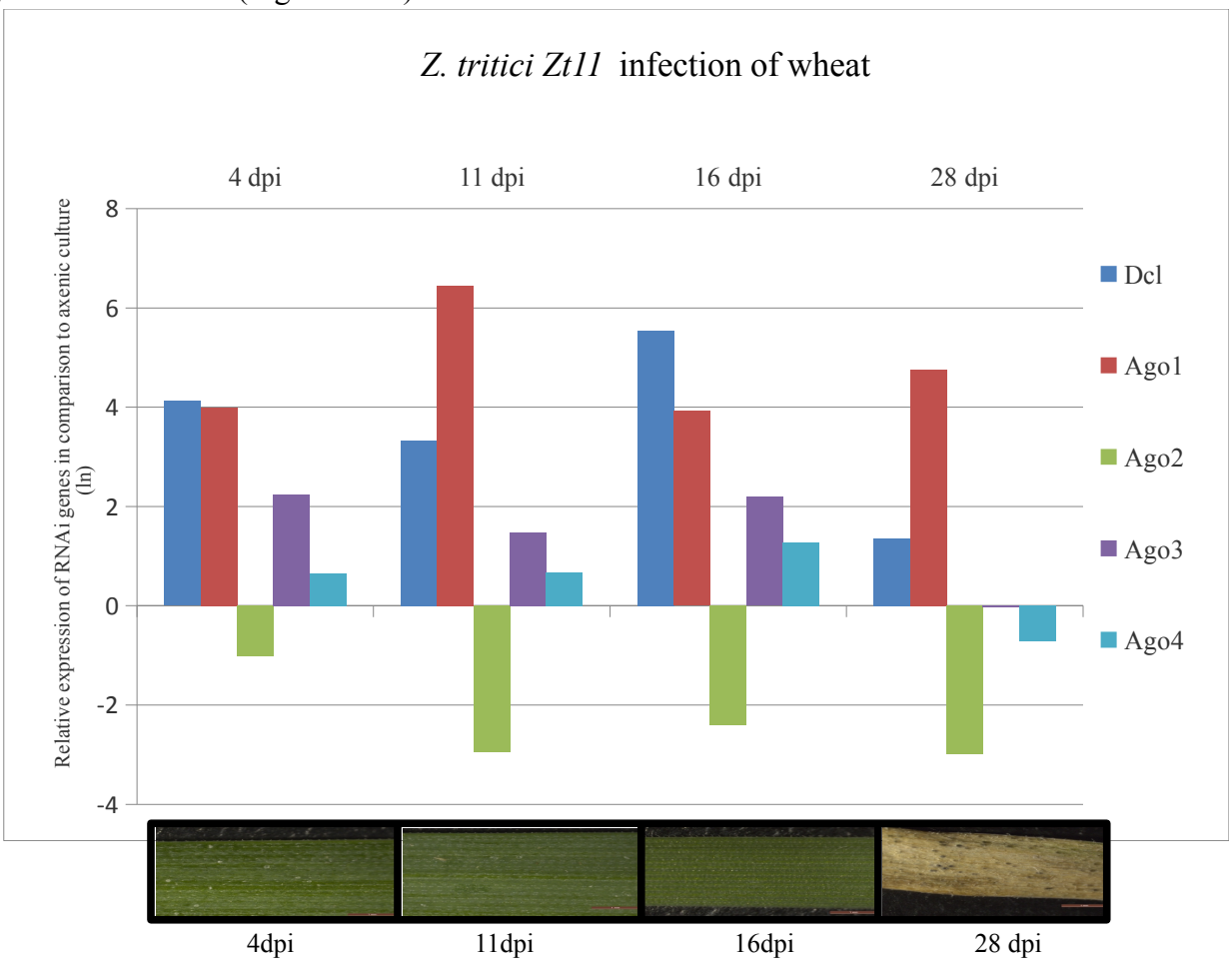

Figure 2.3.6: Relative expression of Dicer-like and Argonaute genes of *Z. tritici* Zt11 during an infection of *T. aestivum* in comparison to gene expression on axenic culture. (The values are show as natural logarithm). Bellow the graph: *T. aestivum* leaves infected with *Z. tritici* Zt11 isolated at 4 time-points during the life-cycle.

Figure 2.3.6 shows that the Dcl and the Ago1, Ago3 and Ago4 genes are up-regulated during the plant infection of isolate Zt11. Ago2 is however not up-regulated during the whole infection process. *Z. tritici* Zt11 growths symptomless until 16dpi. However, the delay in the life-cycle mature pycnidia and necrosis are visible at 28dpi.

Several genomes of the closely related species *Z. pseudotritici*, *Z. ardabiliae* and *Z. passerinii* have been sequenced (Stukenbrock et al., 2011). A BlastN search was conducted to identify homologs of the RNAi related genes in these species (table 2.3.4).

Table 2.3.4: Homologs of *Z. tritici* RNAi genes in *Z. pseudotritici*, *Z. ardabiliae* and *Z. passerinii* isolates.
Genes present in sister species of *Z. tritici* as identified by BlastN searches using the *Z. tritici* reference strain IPO323 or *Z. tritici* Zt-10 as input. √: present in the genome; -: absent in the genome; √/-: partial present, partial absent in the genome. Zp: *Z. pseudotritici,* Za: *Z. ardabiliae* and Zpa: *Z. passerinii.*

	Zp12	Zp13	Zp14	Zp15	Zp16	Za17	Za18	Za19	Za20	Zpa21
Dcl	√	√	√	√	√	√	√	√	√	√
Ago1	√	√	√	√	√	-	-	-	-	√
Ago2	√	√	√	√	√	√/-	√/-	√	√	√
Ago3	√	√	√	√	√	√	√	√	√	√
Ago4	√/-	√/-	√/-	√/-	√/-	√/-	√/-	√/-	√/-	√/-

For all isolates the whole ORF was identified based on the BlastN. The Dcl gene has at least a sequence identity of 94.32% to the *Z. pseudotritici* and *Z. ardabiliae* homologs. As expected, the Dcl of *Z. passerinii* has less similarity (84.16% identity). The Ago1 gene of *Z. pseudotritici* and *Z. ardabiliae* group with the Ago1 allele from IPO323. The Ago1 allele from Zt10 and Zt11 show a considerably larger distance to the other *Zymoseptoria* sequences.

The Ago2 has a sequence identity of 91.62% and 82.65% for *Z. pseudotritici* and *Z. ardabiliae* respectively. The whole ORF is present inside the genomes. Ago2 of *Z. passerinii* has a sequence identity of 79.30%.

Ago3 has a sequence identity of 92.17%, 87.77% and 83.24% for *Z. pseudotritici, Z. ardabiliae* and *Z. passerinii* respectively. The whole ORF is present inside the genomes. Ago4 is not conserved in *Z. pseudotritici, Z. ardabiliae* and *Z. passerinii*. All Ago4 ORFs have two gaps and the three fragments have a sequence identity of 64.93%, 71.53% and 88.74%.

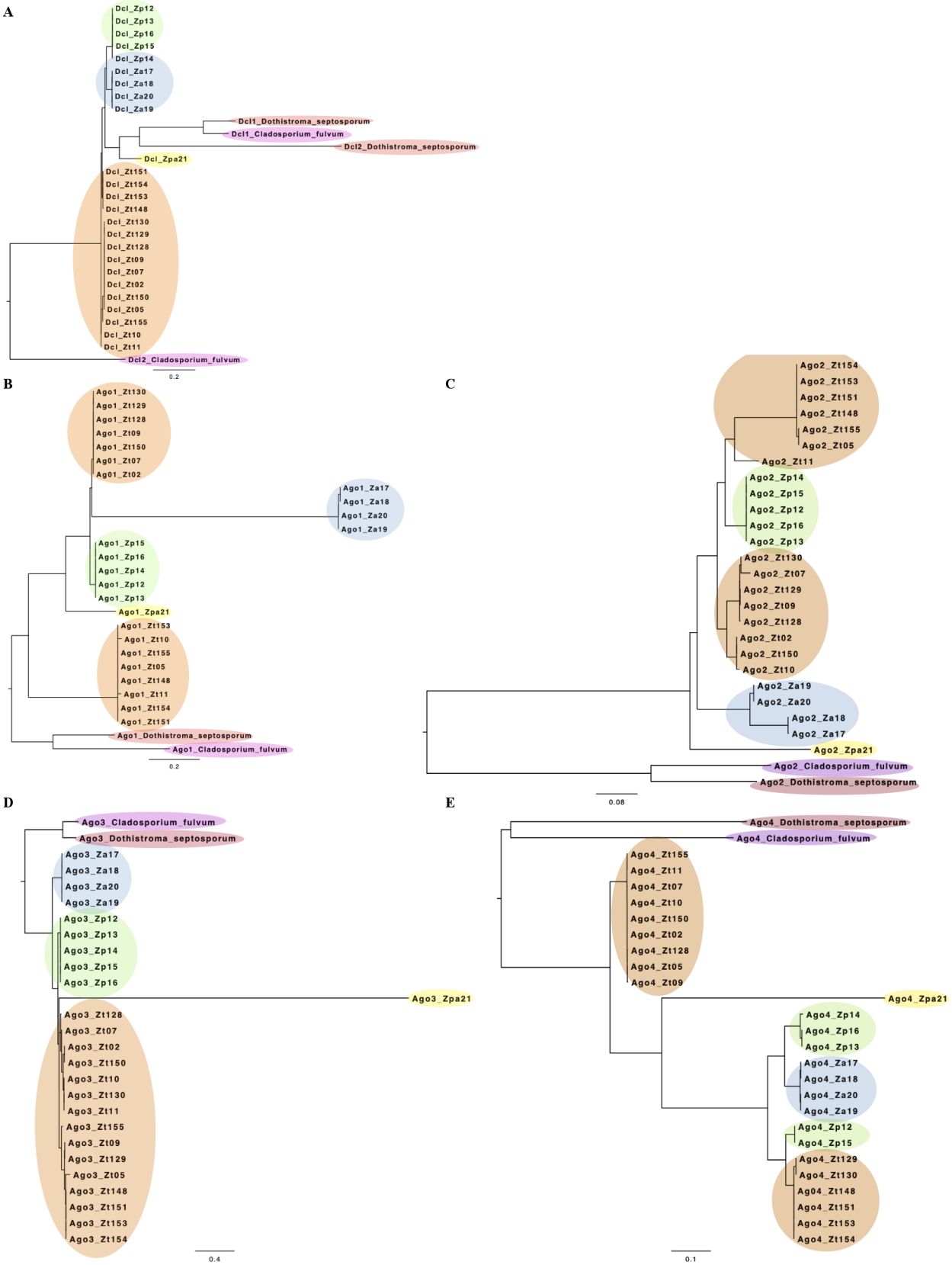

Figure 2.3.7: Phylogenic relationship of RNAi genes in *Z. tritici*, *Z. pseudotritici*, *Z. ardabiliae* and *Z. passerinii.* Nucleotide alignments (Muscle) of the ORFs of Dcl and Ago1-4 using all 15 *Z. tritici* isolates were created. All ORF contained start and stop codons. The phylogenetic trees were conducted using a maximum likelihood as implemented in

PhyML using a GTR substitution model. A) Dcl, B) Ago1, C) Ago2, D) Ago3 and E) Ago4.

Figure 2.3.7 shows the phylogenetic relationship of the Ago genes of the *Zymoseptoria* species. *Dothistroma septosporum* and *Cladosporium fulvum* were included as out-groups (de Wit et al., 2012). Ago1 shows three distinct clades in the phylogenetic tree. Interestingly, there is a discrepancy in the clustering of the isolates of the Ago1 allele and the Ago4 allele which suggest that the genes are not physically linked.

Identification of a Argonaute (pseudo)gene

With the BlastN search for homologous sequence of Ago3, some of the *Z. tritici* isolates contained an additional hit (J. Grandaubert and E. Stukenbrock, unpublished) & (G. Kema unpublished). This hit led to the identification of an additional putative Argonaute gene, here named Ago5, with a length of 2.82 kb and a sequence identity of 94.35% to the Ago3 of IPO323. Suggesting that Ago5 is derived from Ago3. Ago5 sequences lack the start codon (start with ATA) instead of ATG. This led to the conclusion that Ago5 is a pseudogene. The C/G content of the ORF is 50.5% which is lower than for the Dcl and other Argonaute genes. This could suggest that Ago5 underwent repeat induced point (RIP) mutations, where the C::G is replaced by A::T. Ago5 is present in 8 of the 15 *Z. tritici* genomes and located on a scaffold with a length of 18.5 to 68 kB. The scaffold is absent in the remaining seven *Z. tritici* genomes. The genomic locus is also absent in the genomes of *Z. pseudotritici*, *Z. ardabiliae* and *Z. passerinii* as shown by Blast searches (Stukenbrock et al., 2011). The nuclear diversity plot of figure 2.3.8 shows that the both Ago5 and the surrounding regions are conserved in all eight *Z. tritici* genomes.

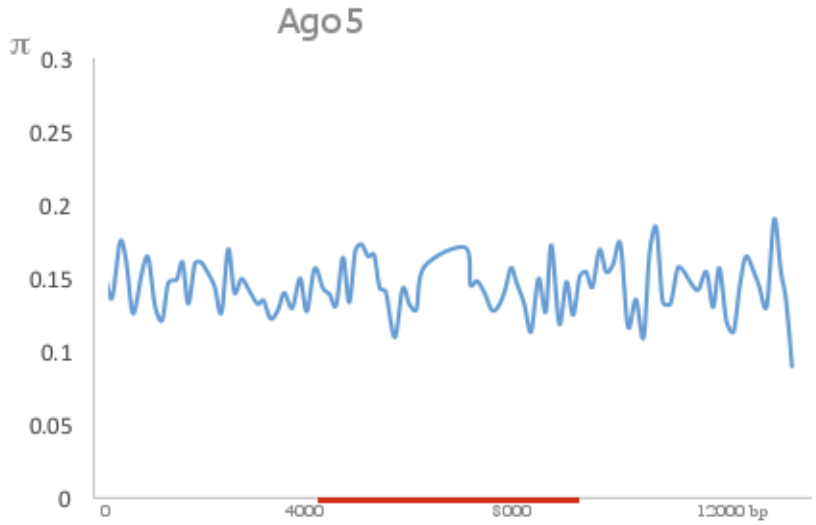

Figure 2.3.8: Nuclear diversity plot and alignment of Ago5 and 5 kb surrounding regions.

The diagram shows variation of nucleotide diversity (π) in windows of 100 bp with a step size of 100. Ago5 of Zt05 is marked with in red.

Chapter 3: Discussion

Chapter 3.1: Histone modifications of essential and dispensable chromosomes in *Z. tritici*

Here, we describe histone modifications for fungal dispensable chromosomes for the first time. Compared to the essential chromosomes of *Z. tritici* the dispensable chromosomes are highly enriched for heterochromatin, in contrast the essential chromosomes are mainly euchromatic. This correlates with the transcription patterns (Goodwin et al., 2011) & (Kellner et al., 2014). The dispensable chromosomes only code for 6% of the genes and show a significantly reduced level of gene transcription in comparison to the essential chromosomes (Goodwin et al., 2011) & (Kellner et al., 2014).

Goodwin et al showed that the dispensable chromosomes have a higher AT-content (49.1% instead of 47.7% compared to the essential chromosomes (Goodwin et al., 2011). Apart from the AT enrichment, the dispensable chromosomes have twice the amount of repetitive elements compared to the essential chromosomes (Goodwin et al., 2011). A hypothesis is that heterochromatin silence transposons and therefore is enriched at the dispensable chromosomes (Peng and Karpen, 2008).

The low number of genes on the dispensable chromosomes could be the result of heterochromatin spreading into coding regions. Due to the spreading of heterochromatin, genes could be silenced and over time loss their function (Grewal and Jia, 2007).

Several regions of the essential chromosomes are also enriched with heterochromatic histone modifications, lack predicted genes, have a high content of repetitive sequences and low levels of gene transcription. Grewal et al hypothesized that repetitive sequences are masked by heterochromatin to prevent recombination, and rearrangements with other repetitive regions in the genome (Grewal and Jia, 2007). A hypothesis is that the heterochromatic islands in *Z. tritici* are associated with chromosomes breakage, which could lead to variability of karyotypes of distinct isolates. Not only can the dispensable chromosomes be lost after meiosis, also the length of both dispensable and essential chromosomes varies between isolates (McDonald and Martinez, 1991) & (Mehrabi et al., 2007). These structural differences may be mediated by the presence of heterochromatin.

Recently, Croll et al reported that breakage-fusion-bridge cycles could result in the formation of a dispensable chromosome (Croll et al., 2013). On chromosome 14 so-called alignment breakpoints were identified. One of these alignment breakpoints overlaps exactly with a heterochromatic island identified here. The variation in length of the chromosomes could be explained by rearrangement of segments by sequence homology of the flanking heterochromatic islands. Apart from the described

heterochromatic islands, there are many more regions in the *Z. tritici* IPO323 genome which are enriched with H3K9me3. These regions are smaller than 10 kb and not included in table 2.1.1. Further analyses of gene content, sequence composition and transcription will elucidate the role of these local modifications.

Also the B-chromosomes of *Zea mays* are enriched with heterochromatin and lack coding genes (Marschner et al., 2007). Nevertheless one third of the B-chromosome is euchromatic (Lamb et al., 2005). This is in contrast to the dispensable chromosomes of *Z. tritici* which are enriched with H3K9me3 and H3K27me3 and only have short stretches of sequences which are enriched with H3K4me2. H3K27me3 is a histone modification which is also associated with secondary clusters and metabolites. In general, genes located in regions enriched with H3K27me3 are not transcribed and the histone modifications allows them to switch from an active state to a silent state. Connolly et al have shown that the deletion of the H3K27 methyltransferase gene Kmt6 in *Fusarium graminearum,* led to the transcription of secondary metabolites which were silenced in the wild-type under standard conditions (Connolly et al., 2013). Soyer et al silenced the H3K9 methyltransferase gene Dim-5 and heterochromatic protein 1 (HP-1) in *Leptosphaeria maculans*, resulting in the transcription of 400 genes, which were silenced in the wild-type (Soyer et al., 2014). Both studies show that gene transcription is regulated by chromatin changes. Currently, studies to characterize the Kmt6 and Dim-5 in *Z. tritici* are in progress. It is hypothesized that also in *Z. tritici* several genes, silenced in the wild-type, will be expressed in the mutant strains due to a change in the chromatin stage.

Until today, an overall overview of the gene transcription during the complete lifecycle of *Z. tritici* is lacking. Kellner et al showed that a whole region of chromosome 7 lacks gene transcription while growing in axenic culture and at 4 dpi in two different hosts (Kellner et al., 2014). Based on the histone modification, it can not be ruled out that some genes located in the region of chromosome 7 show are transcribed at later stages of the lifecycle. RNAseq or qPCRs experiments of infected plants could be designed to test this hypothesis.

All the telomeres of *Z. tritici* are enriched with H3K27me3, consistent with findings in other fungi like *N. crassa* (Jamieson et al., 2013). The H3K27me3 enriched region of chromosome 7 is most likely also a result of chromosomal rearrangements. The region of chromosome 7 could originate from a dispensable chromosome that has fused to chromosome 7. However the histone modifications differ, the dispensable chromosomes are mainly enriched with H3K9me3, which is absent at the facultative heterochromatic region. Both the dispensable chromosomes and the facultative heterochromatic region of chromosome 7 show a drastic decrease in gene transcription.

The region of chromosome 7 lacks similarity to other parts of the IPO322 genome and based on this we concluded that this region did not originate from a duplication event. The AT-content of the whole region is 47.0 % which is similar to the overall AT content of *Z. tritici* IPO323. The AT-content of the dispensable chromosomes are slightly higher (49.1%) (Goodwin et al., 2011). For this particular region it is possible that two essential chromosome fused resulting in a new chromosome arm of chromosome 7 of IPO323.

Mehrabi et al showed, by pulsed field gel electrophoresis (PFGE), that three *Z. tritici* strains show chromosome length polymorphisms. Based on the PFGE, a chromosome with the same length as chromosome 7 is absent in the other two strains (Mehrabi et al., 2007). Since this analysis is based on a PFGE gel, it is difficult to prove that a specific band of the other *Z. tritici* isolates indeed match to corresponding chromosome of *Z. tritici IPO323*. The alignment of figure 2.1.8 and PCR analyses have shown that the facultative heterochromatic region still is present in the genome of the strains which were used by Mehrabi et al. However, they might be rearranged and be part of other chromosomes. To verify this hypothesis, PFGE and Southern blots could be applied to locate the region in the genome.

Chapter 3.2: Characterization of centromeres in *Zymoseptoria tritici*

A previous study of the centromeres in *Z. tritici* used native anti-bodies against the *Z. tritici* CenH3 protein (Dhillon and Goodwin, unpublished). However, this approach was unsuccessful in mapping centromeres. In this study, both an N-terminal GFP construct and a C-terminal GFP construct were created. In contrast to the N-terminal GFP construct, the C-terminal GFP construct did not result in a correct replacement of the CenH3 and did not show multiple foci inside the nucleus. This indicate that the CenH3 protein of *Z. tritici* may have an essential functional domain on the C-terminal end of the protein. Fluorescence microscopic screening of several GFP-CenH3 *Z. tritici* isolates showed that there are multiple foci visible inside the nucleus, that the number of foci differs and that the distribution of the foci in each cell is random and unorganized. However the microscopic screening was preformed with a fluorescence microscope and not with a confocal microscope. Using a confocal microscope could improve the visualization of GFP signal since it would allow to visualize multiple layers inside the fungal cell.

The multiple foci could originate from structural difference between essential and dispensable chromosomes. A hypothesis is that the essential and dispensable chromosomes are ordered differentially inside the nucleus. Chromosome conformation capture technology (3C, HiC, etc) experiments have shown that the active transcription of genes occurs in special compartments of the nucleus (de Wit and de Laat, 2012). Knowing that the dispensable chromosomes are silenced, it is possible that the non-transcribed chromosomes are located differentially in the nucleus compared to the essential chromosomes (Kellner et al., 2014).

The “telomere to telomere” sequenced genome of *Z. tritici* IPO323 allowed complete mapping of the centromeric reads (Goodwin et al., 2011). Similar studies based on ChIP-seq in other fungal species, failed to completely map centromere reads due to incomplete assemblies of genomes. Centromeres are often found in AT rich, repetitive DNA which is more difficult to assemble. We have here obtained sufficiently high coverage of centromeric sequences to describe centromeres in *Z. tritici.*

Centromeres in *Z. tritici* are located close to the chromosome ends. The location of the centromeres of *Z. tritici* IPO323 thereby differ from other centromeres described in fungal species (Smith et al., 2012). Centromeres described in other fungal species (*N. crassa* and *F. graminearum*) are located in the longest repeat rich sequences of the chromosomes and are often enriched with AT nucleotides (Smith et al., 2012) & (Smith et al., 2011). Although the centromere of chromosome 6 of *C. albicans* is located ~50 kb from the telomere (Ketel et al., 2009).

The centromeres of *Z. tritici* are relatively small (~9 kb) in comparison to centromeres identified in other ascomycete fungi such as *N. crassa* (175-300 kb) and *F. graminearum* (~40 kb) (Smith et al., 2012) & (Smith et al., 2011). Based on the size and the lack of structural domains, the centromeres of *Z. tritici* are similar to the centromeres of *C. albicans* (Sanyal et al., 2004). We ensured correct assembly and mapping by PCR amplification of multiple loci with each predicted centromere of IPO323.

The centromeres of the essential and dispensable chromosomes are similar in structure, organization and length. Based on centromeres the chromosomes cannot be distinguished and the centromeric organization of the dispensable chromosomes cannot cause the instability of the dispensable chromosomes.

In general the centromeric sequences of *Z. tritici* are enriched in AT nucleotides (mean 52%). The overall AT-content of the genome of IPO323 is 48%. The centromeres of chromosome 3, 4, 7, 12, 13, 14, 20 and 21 are not enriched in AT nucleotides (mean 48.09%). Of these eight chromosomes, the centromeres of chromosome 13, 14, 20 and 21 have even a lower AT content than the overall genome (Goodwin et al., 2011). This low centromeric AT-content is rare and an unusual feature of centromeric sequences (Smith et al., 2012).

The centromeres of *Z. tritici* are directly surrounded by H3K4me2 histone modifications, which is however a typical feature of centromeres as also is found at the centromeres of *S. pombe* and *N. crassa* (Ishii et al., 2008) & (Smith et al., 2011). Where the centromeres are surrounded by genes and H3K4me2 histone modifications. Interestingly, also the centromeres of the dispensable chromosomes, have H3K4me2 enrichment at the surrounding regions of the centromeres. Otherwise these chromosomes entirely lack enrichment for H3K4me2 (Chapter 2.1).

The centromeres of *Z. tritici* IPO323 are not enriched for H3K9me3 or H3K27me3 this is in contrast to the centromeric regions in *N. crassa*. In *N. crassa* the centromeres are enriched with H3K9me3 (Jamieson et al., 2013). It is however possible that the centromeres are enriched with other heterochromatic histone modification that were not studied here, for example H3K27me3, H4K20me2, H3K36me2 or H3K9me2.

Based on the genome annotation of the JGI and RNA-seq data, 29 genes are encoded inside the centromeres (Goodwin et al., 2011) & (Kellner et al., 2014). Nine genes have an RPKM value higher than 5.0 and of these nine genes, two are partially located inside centromeres. Five additional genes are located 2 kb from the centromeric boarder. The remaining two genes are paralogous genes located in the middle of the centromere. None of the 29 genes are located inside a repetitive regions, although several are surrounded by repetitive DNA. Based on the gene expression data the essential

regions of the centromeres might be smaller than the region which is covered by CenH3. From the nine genes inside the centromeres, three are assigned to a KOG category (2 for metabolism and one for cellular processes and signaling). The regions which were predicted to contain genes inside the centromeres are not enriched with H3K4me2, which is in contrast to the expression data of these genes. To verify the gene expression qPCRs could further verify transcription.

In the centromeres of *N. crassa,* three putative genes were identified (Smith et al., 2011). In contrast to the genes located in the centromeres of *Z. tritici* all three genes were not expressed at all tested conditions. The three genes of *N. crassa* are not enriched with H3K4me2 and were found to be associated with two other centromeric proteins (Cen-C and Cen-T) (Smith et al., 2011). These results together suggest that the three genes are pseudogenes. In *C. albicans* genes are absent in the centromeric regions (Ketel et al., 2009). Like *Z. tritici* the centromeres of *C. albicans* are surrounded with genes and these genes are transcribed at similar levels as the genome wide transcription level (Ketel et al., 2009).

Neocentromere formation in *S. pombe* occurs in regions with genes which are low expressed during normal growth, but induced at nitrogen starvation (Ishii et al., 2008). After neocentromere formation the gene expression remains low after nitrogen removal, suggesting that the CenH3 silenced the expression of the genes. In *S. pombe* the replacement of the centromeric sequence, with non centromeric sequence did not disturbed the centromere formation (Castillo et al., 2007). Interestingly, the introduction of naked centromere DNA, including the central core and outer repeats resulted in the formation of heterochromatin and a functional centromere (Folco et al., 2008). This suggest that centromere formation in *S. pombe* depends on heterochromatin.

Centromere 3 is the only centromere without any repetitive regions. All other centromeres have at least one (Cen 9, 13 and 18) or more repetitive elements. Centromere 20, 6 and 8 have the highest number of repetitive sequences 13, 11 and 10 respectively. Interestingly, one single repetitive domain from centromere 18 is 6 kb covers 46.62% of the centromere.

All centromeres of the *S. pombe, Schizosaccharomyces octosporous* and *Schizosaccharomyces japonicus* are enriched with repetitive sequences (Rhind et al., 2011). The repeats in the centromeres of S. *pombe* and *S. octosporous* are similar and the centromeres of *S. japonicus* are only centromeres enriched with transposons in the pericentric regions (Rhind et al., 2011).

The assemblies of the *Z. tritici* isolates IPO94269, IPO95052, 02133 and 01151 consist of many contigs (table 2.1.2). Therefore most centromeres are represented by more than one contig. The exact centromeric position of the additional *Z. tritici* isolates was consequently not shown. The centromeric sequences of parental and progeny strains were however compared. The alignment of

centromeres contain gaps, most likely due to the repetitive elements.

I n *N. crassa* the centromeres contain relics of transposable elements (Smith et al., 2011). *C. albicans* lacks repetitive DNA and transposons in the centromeres and all of the eight centromeres have unique DNA sequence thus also conserved domains are lacking (Mishra et al., 2007). This suggest that the centromeres are also sequence independent in *C. albicans*. Mishra et al compared centromeric sequences of 15 *C. albicans* strains. These strains have diverged ~1-3 million years ago. From the 15 strains a subgroup of six strains were selected and the nucleotide diversity in theses centromeric sequences (Cen1) is 5 to 15-fold higher than in the coding genes. Southern blot analyses for centromere 6 showed an identical pattern. Both the PCR and the Southern blot analyses showed that the centromere sequences are identical for centromere 1 and 6 and the authors suggested that this also applies for the other centromeres. ChIP-PCR analyses showed that all eight centromeres of the six *C. albicans* strains could be amplified with the same primer set and all have the same length. This suggest that the centromeric sequences were not enriched with transposons, in contrast to the whole genome level, where extensive chromosome rearrangements had occurred.

Centromeres of *C. albicans* have also been compared with the closely related species *C. dubliniensis* (Last common ancestor: 20 million years ago) (Padmanabhan et al., 2008). The centromeres of *C. albicans* and *C. dubliniensis* vary in sequences, this is in contrast to the nucleotide diversity at the genome wide level. The genomes of the two *Candida* species are otherwise conserved. As for the with-in species comparison, it was also found that the centromeric positions between *C. albicans* and *C. dubliniensis* are conserved between the species. As for *C. albicans* the centromeric sequence of the *Z. tritici* isolates are conserved. Bensasson et al showed also that the centromeric regions (point centromeres) of *Saccharomyces paradoxus* show a higher rate of nucleotides substitutions than the genome average (Bensasson et al., 2008).

In human the centromere of the X-chromosome consists of a pericentric regions and a central core (Schueler et al., 2001). The central core is an array from twelve 171 bp repeat (α-satellite sequences) and this repetitive sequence binds to the kinetochore proteins. The pericentric regions exist of monomeric units and is not involved in the kinetochore assembly. Comparative genomics between human and great apes showed that the nucleotide substitution rate was higher inside the central core in comparison to the pericentric regions of the X-chromosome (Rudd et al., 2009).

Besides the centromeric sequences, also CenH3 shows a high rate of positive selection within *Drosophila melanogaster* and *Drosophila simulans (Malik and Henikoff, 2001)*. Especially the highly variable N-terminal tail and the histone fold domain evolved under positive selection.

The nucleotide diversity and the positive selection of the CenH3 protein have given rice to the

"centromere-drive" model (Henikoff et al., 2001). According to this model, the centromeric sequence expand resulting in longer centromeric sequence. This improves the attraction of microtubles. This is followed by a second step, that involved the mutation of centromere binding proteins. Positive selection of the centromere binding proteins restores the genetic stability, that occurred after the expansion of the centromeric sequence. The model suggest that centromeres are a dynamic interaction between centromeric regions and centromere binding proteins. However the interaction between the centromere binding proteins and the centromere is independent from a specific centromeric sequence and dependents on the ability of the centromeric regions to bind with centromeric proteins to assemble a kinetochore.

In human CenB, is important for the centromere formation. CenB associates B-box sequences (with (α-satellite sequences). In human, the introduction of naked DNA can lead to the formation of functional centromeres, although 30 kb of α-satellite sequences and B-boxes are required (Harrington et al., 1997). However CenB might not be involved in centromere assembly in fungi (Smith et al., 2012). Zaratiegui et al, have shown that CenB of *S. pombe* is involved in the silencing of transposons and has a role in DNA replication (Zaratiegui et al., 2011). Interestingly, the CenB mutant of *S. pombe* shows double-strand breaks in the three chromosomes of the fungus. Also *C. albicans* lacked centromere formation when naked DNA containing α-satellite sequences was introduced (Baum et al., 2006). In *Z. tritici* the CenB was also tagged with GFP. Fluorescence microscopic analyses confirmed that the CenB of *Z. tritici* is localized inside the nucleus. However, the GFP-CenH3 multiple foci were not present and instead the GFP signal was spread in the whole nucleus. We speculate that CenB is not involved in the centromeric assembly of *Z. tritici*. Based on the function in *S. pombe,* it will be interesting to study the mutant of CenB in *Z. tritici.*

In summary the *Z. tritici* centromeres are short and not AT-rich. They are similar to the centromeres of *C. albicans* especially in terms of their organization and length of the centromeres.

We have aimed to describe the formation of new centromeres, and set out to delete centromeres on essential and dispensable chromosomes. The strain which lost the centromere of chromosome 14 during transformation, still has the wild-type centromere and also neocentromere formation. The wild-type pattern is still visible in one of the Southern blots. Compared to the band which is specific for the centromere deletion, the wild-type bands are weak.

After the attempt to delete the centromere of chromosome 14, several transformed strains were obtained which had lost chromosome 14. This was shown by eight PCR reactions. Further pulsed-field gel electrophoresis will confirm this. The loss of chromosome 14 suggests that the centromere is essential for the chromosome stability. Interestingly, the isolates which have lost chromosome 14

are still resistant to the antibiotic used for the deletion of the centromere. Resistance against the antibiotic suggests that the antibiotic resistance is integrated on another chromosome. Ectopic integrations of the antibiotic resistance could be the reason why the isolates which have lost chromosome 14 still show resistance to the antibiotic.

The cell morphology of *Z. tritici* IPO323 ΔChr14&18 is similar to the cell shapes of *Z. tritici* IPO323 ΔChr18. This observation is agreement with previous claims that a dispensable chromosomes can be lost without any effect on fitness. Comparative inplanta experiments with Z. *tritici* IPO323*, Z. tritici* IPO323 ΔChr18 and *Z. tritici* IPO323 ΔChr14&18 will allow us to study if chromosome 14 is essential for the host-fungus interaction.

Besides the deletion of the centromere of chromosome 14, experiments to induce neocentromere formation on chromosome 13 were performed. The deletion of the centromere of chromosome 13 was here unsuccessful. This result is in agreement with the idea that the genes on the essential chromosomes code for primary functions and that the loss of an essential chromosome is lethal for the fungus.

The formation of neocentromeres in fungi is rare. neocentromere formation has not been described in filamentous fungi (Smith et al., 2012). Ishii et al showed that the neocentromere formation has a mean frequency of 8.7×10^{-4} in *S. pombe* (Ishii et al., 2008). Thankur et al screened an average of colonies of 2800 mutants to study the neocentromere formation on 3 chromosomes (Thakur and Sanyal, 2013). Screening these numbers of mutants after a *Z. tritici* transformation is impossible due to technical limitations of the transformation protocol and the lack of a macroscopic marker to screen the mutants. Also *C. albicans* has a diploid genome, which makes it possible to delete the centromere first on one chromosome and subsequently on the other chromosome (Ketel et al., 2009) & (Thakur and Sanyal, 2013).

As in *Z. tritici* the introduction of GFP-CenH3 in *Z. pseudotritici* (Zp13) and *Z. ardabiliae* (Za17) led to the observation of multiple foci in the nucleus. This suggest that this particular phenomenon is conserved among *Zymoseptoria* species.

Like the additional genomes of *Z. tritici* (IPO95052, IPO94269 and progeny strain 02133) the genomes of *Z. pseudotritici* (Zp13) and *Z. ardabiliae* (Za17) are assembled into many scaffolds, preventing the complete assembly of centromeric sequences. Therefore the centromeres were only partial identified inside the genomes of *Z. pseudotritici* and *Z. ardabiliae*. As consequence of the variation in the number of dispensable chromosomes in the *Zymoseptoria* species, the presence or absence of the centromeres varies as-well. The higher variation in sequence diversity, correlates with the results of Padmanabhan et al and Bensasson et al (Padmanabhan et al., 2008) &

(Bensasson et al., 2008).

Currently, the sequencing of the, by ChIP obtained, centromeric reads of *Z. pseudotritici* (Zp13) and *Z. ardabiliae* (Za17) are in progress. After mapping of the centromere derived reads to the specific genomes, the exact location and nucleotide sequence of the centromeres of *Z. pseudotritici* (Zp13) and *Z. ardabiliae* (Za17) are known and this allows us to compare the centromeres of three species of the *Zymoseptoria* clade.

Chapter 3.3: Functional and phylogenetic characterization of proteins involved in the RNAi-pathways

Very little is known about the role of RNAi related genes in pathogenic fungi (Chang et al., 2012) & (Katiyar-Agarwal and Jin, 2010) & (Billmyre et al., 2013). In this study, the roles of Dicer and Argonaute genes in the plant pathogenic fungi *Z. tritici* were addressed. It was shown that the RNAi system plays a role in virulence of the pathogen.

Most studied fungal genomes contain two dicer (-like) proteins, with distinct function during the life-cycle. Most of the fungi have a Dcl specific for vegetive growth and one Dcl involved in meiosis (Hu et al., 2013). In the rice blast pathogen *Magnaporthe oryzae* it was shown that only the *dcl2*, and not the *dcl1* mutant has a disruption in the siRNA production (Kadotani et al., 2004). During quelling in *N. crassa* it was shown that the two Dcls are redundant to each other and that only the double mutant shows a defect in quelling (Catalanotto et al., 2004). The number of Argonaute encoding genes in fungi varies. *N. crassa* has two, but some basidiomycetes have up to eight (Hu et al., 2013). For most fungi, experimental evidence for the function of the Argonaute is lacking and the number of Argonautes is based on genome annotations. *Z. tritici* has only one Dcl coding gene and four Argonaute genes (Table 2.31) and both the Dcl and Ago genes have all the functional domains.

Expression studies showed that the five RNAi genes are not up-regulated in axenic culture, suggesting that the RNAi mechanism is not involved in basic growth in axenic culture. The cell morphology, in axenic culture, of the *dcl* and *ago1-4* single mutants are identical to the wild-type. However, the expression of the dicer and Argonaute genes were also tested in-planta and showed a distinct expression pattern. At the end of the life-cycle (28dpi) of *Z. tritici,* all genes were down-regulated again in the three strains studied here. The lack of expression in axenic media and during saprothrophic growth at 28dpi, together with the up-regulation of the dicer-like and Argonaute genes at 4, 11 and 16 dpi suggest that RNAi pathways play a role in the switching of life-style (from biotrophic to necrotrophic growth) of *Z. tritici*. Also, the similarity of the Dcl up-regulation in the three isolates of *Z. tritici* shortly before the pycnidia formation leads to the conclusion that the RNAi-pathways play a role in the virulence of *Z. tritici*. To conclude the results of the expression studies, the gene expression suggest that the RNAi-pathways of *Z. tritici* are linked to the host-fungus interaction and are not involved in basic growth on axenic culture or during saprotrophic growth in planta.

The expression pattern of the dicer gene is similar to what was observed in *Phytophthora infestans*

(Tsutsumi et al., 2011). Also, in *P. infestans* the Dcl gene is only weakly expressed during axenic growth and expression is increased two to three fold at 24 hpi. At 48 and 72 hpi, the expression levels are similar to axenic growth. Like in *Z. tritici,* the five Argonaute genes of *P. infestans* showed different expression levels during infection. At 24 hpi, the relative increase of expression of Ago5 compared to the axenic conditions was five to seven fold. This is in contrast to the expression levels of Ago1-2, which showed a two fold increase in axenic media and which were not up-regulated during infection.

In 2012, Vetukuri et al, showed that 77% of the small-RNAs in *P. infestans* derive from genes coding for Crinkler (CRN) effector genes, and have a length of 21 nt (Vetukuri et al., 2012). This is in contrast to the small-RNAs which map to RxLR effectors. These are mainly 22, 25 and 26 nt in length. From a comparative analysis of small-RNAs in a virulent and non-virulent strain of *P. infestans*, it was clear that small-RNAs play a role in pathogenicity: Two RxLR effector genes are enriched for small-RNAs in a strain with reduced pathogenicity, this is in contrast to a pathogenic strain, where the same two genes lack enrichment of the same small-RNAs. Unfortunately, the authors did not show the expression levels of the two RxLR effector genes in the absence of either a dicer or an Argonaute encoding gene. This experiment could however have further documented the importances of small-RNAs in pathogenicity.

The trigger that switch the metabolism of *Z. tritici* from biotrophic to necrotrophic growth is unknown. A hypotheses is that the fungus needs the biotrophic stage to produce a certain amount of fungal biomass to reproduce in necrotic tissue. All mutants show similar necrosis levels as the wild-type. However, it can not be excluded that the necrosis occurred through other factors than the infection of *Z. tritici*. The appearance of necrosis due to secondary effects could prevent the biotrophic spread of *Z. tritici.* The lower amount of pycnidia of the mutants could thereby reflect a lower amount of intercellular hyphea. To confirm this hypotheses, the biomass of the wild-type and mutants should be assayed.

In fungal pathogen *Sclerotinia sclerotiorum* computational, and experimental evidence for two potential miRNAs were shown by Northern blot (Zhou et al., 2012). Pre- and pri-RNAs of the predicted miRNAs were however not confirmed by Northern blot. The miRNAs of *S. sclerotiorum* are 20-24 nt in length and originate predominately from antisense exonic regions of protein coding genes. For all eight miRNA tested the highest expression value was observed 72 hpi. Expression levels were lower at 96 and 120 hpi. Interestingly, the expression of the Dicers and Argonaute encoding genes were consistent during the three time-points. The authors suggest that the change in expression of the miRNAs is due to changes in gene transcriptions. The authors did not show the

expression levels of the RNA-dependent RNA polymerase, which could have shed some more light on the roles of small-RNA transcription.

The fungal plant pathogen *Magnaporthe oryzae* has two dicer proteins. Only Dcl2 is essential for siRNA production in axenic culture (Kadotani et al., 2004). In axenic culture the *dcl2* mutants has a slightly longer cell cycle at 22°C and 30°C, when compared to the wild-type and the *dcl*1 mutant. Interestingly, the two mutants show the same growth curves as the wild-type when grown at 20°C. Tissue specific small-RNAs were observed in *M. oryzae.* Nunes et al showed that small-RNAs isolated from apperessoria predominately have a length of 28-35 nt and originate from the tRNA loci (Nunes et al., 2011). This is in contrast to mycelium growth. Here, the small-RNA mainly have a length of 18-23 nt and are derived from transposable elements. Raman et al show showed that small-RNAs, of *M. oryzae,* isolated from in planta experiments, predominately have a length of 26 nt and the small-RNAs obtained from mycelial mainly have a length of 25 nt (Raman et al., 2013). Raman et al also showed the relevance of small-RNAs during several stress conditions and plant infection. A stress assay resulted in the identification of 98 genes which were down-regulated and in planta 12 down-regulated genes. To verify if the down-regulation of the genes was due to post-transcriptional regulation, the gene expression of dicer-1, dicer-2 mutants were compared to the wild-type expression levels. Interestingly, the gene expression levels were increased 1.7 fold in the *dcl1* mutant. This shows that, the *dcl1* is involved in stress response of *M. oryzae*.

Small-RNAs are important in fungi during stress responses, gene regulation, and during plant infection. Interestingly, the *dcl* and *ago1-3* single mutants of *Z. tritici* show macroscopical the same phenotype as the wild-type when induced to several types of cell-wall stress. Additional stress factors, like DNA damage, could be tested to further study the role of RNAi in *Z. tritici* during stress (Lee et al., 2009).

As shown, most studies are based on axenic growth and the results presented in this thesis are one of the first which test the role of RNAi during plant infection. Recently Weiberg et al showed that small-RNAs of *Botrytis cinerea* hijack the RNAi machinery of the host (Weiberg et al., 2013). The double *dcl1* and *dcl2* mutant of *B. cinerea* lacks the ability to produce small-RNAs and compared to the wild-type this mutant shows reduced virulence. In *Z. tritici* the *dcl* and *ago1-3* single mutants have necrosis levels comparable to the wild-type. This suggests that the same mechanism which transports small-RNAs to the host in the *B. cinerea* system does not exist in *Z. tritici.*

In this study, we also analyzed sequence evolution of RNAi related genes using phylogenetic approaches. The Dcl and Argonautes genes from a total of 15 *Z. tritici* isolates were compared. The isolates have been collected from leaves of bread wheat (*T. aestivum*) and durum wheat (*T.*

turgidum ssp. *Durum*) growing in the Netherlands, Iran, Denmark, Germany and Algeria. The phylogenetic tree (figure 2.3.7) shows that Ago1 and Ago4 are split in two distinct clades. Ago2 and Ago3 are more conserved among the isolates. The nuclear diversity plot of the Argonaute genes shows that the surrounding regions of all the isolates are conserved. But the coding parts of Ago1 and Ago4 are highly diverged.

Based on the conserved surrounding region of all the Argonautes genes, the introduction of a new Argonaute gene by horizontal gene transfer is unlikely. If horizontal gene transfer had occurred, the transferred Argonaute gene would most likely be integrated in another part of the genome and would not replace an already existing Argonaute gene. All Argonaute genes show signs of diversification, which suggest that a horizontal gene transfer event had occurred a long time ago, leaving time for new mutations to occur. Differentiation of two allele-groups is also present in the closely related species *Z. pseudotritici* and *Z. ardabiliae.* This suggests, that the divergence of the allele-groups is ancient and occurred before differentiation of the *Zymoseptoria* genus. The dicer gene is, however, highly conserved between all the 15 *Z. tritici* isolates. This is most likely due to the fact that *Z. tritici* only has one gene encoding a dicer protein.

In 2012, Bernhardt and co-workers published an article where they described a high evolutionary rate of RNAi related genes in the mosquito *Aedes aegypti* (Bernhardt et al., 2012). All the genes of the exo-siRNA and the miRNA pathway in *Ae. aegypti,* showed signatures of positive selection. A similar finding was shown in three *Drosophila* species (*D. melanogaster, D. simulans* and *D. yakuba*) (Obbard et al., 2006). Like in *Ae. aegypti*, the genes involved in the exo-siRNA pathway were found to be under positive selection and among the top 3% of fastest evolving genes in *Drosophila*. This is in contrast to the genes involved in the miRNA pathway which show a much lower rate of evolutionary changes. It was hypothesized that the differences in selection of the two pathways is due to the function of the exo-siRNA pathway. The authors' hypothesis was that this pathway is under selective pressure due to the infection with flaviviruses. These viruses are insect specific. The role of viral infection in *Z. tritici* is not known. Several fungi, like *Saccharomyces cerevisiae* and *Ustilago maydis* have lost the complete RNAi pathways due to infection of the Killer virus (Drinnenberg et al., 2011). It is unlikely that an infection with the Killer virus has led to divergence of the genes in the RNAi pathways of *Z. tritici.* Nevertheless, an infection with another virus could have led to the diversification of the Argonaute genes.

Within several species it is shown that a specific class of small-RNAs are associated to a particular Argonaute protein (Mallory and Vaucheret, 2010). However, other species have only one highly abundant class of small-RNAs (Braun et al., 2010) & (Wang et al., 2010). It is possible that the

RNAi pathways of *Z. tritici* similarly are associated with specific classes of small-RNAs. A specific binding of a small-RNA class by each Argonaute could also lead to the diversification of two allelic groups of Argonaute genes. To verify this hypothesis, Argonaute genes could be tagged with a protein tag allowing the sequencing of Argonaute associated small-RNAs.

Goodwin et al predict that the dispensable chromosomes of *Z. tritici* are enriched with miRNAs (Goodwin et al., 2011). This is based on the sequencing of size selected RNA and computational predications. Small-RNAs isolated from a Page gel do not exclude the presence of degradation products of messenger RNAs. Better results would have been obtained by sequencing of small-RNAs obtained from immunoprecipitation of small-RNAs associated to Argonaute proteins (Gu et al., 2011) & (Lee et al., 2010). As the expression data here shows, most of the RNAi related genes are up-regulated during the switch from biotrophic to necrotrophic growth. This suggests that the small-RNAs are important during the switching of the two stages and isolating of small-RNAs at several time-points during a plant infection would give more information about the biological relevance than the small-RNA expression during axenic growth.

Chapter 4: General conclusions

Chapter 4: General conclusions

The work in this thesis have characterized epigenetic features of *Z. tritici* for the first time. The differentiation of the essential and dispensable chromosomes is not due to specific centromeres, but rather due to distinct histone modifications. These histone modifications lead to strong differences in transcriptional patterns of the two types of chromosomes. The essential chromosomes are euchromatic and show high levels of gene transcription (Goodwin et al., 2011) & (Kellner et al., 2014). This is in contrast to the dispensable chromosomes which are heterochromatic and have low levels of gene transcription (Goodwin et al., 2011) & (Kellner et al., 2014). In addition to the heterochromatic dispensable chromosomes, a large region (780 kb) of the essential chromosome 7 is non-transribed (Kellner et al., 2014). This region is enriched with H3K27me3. It was also shown here that other essential chromosomes contain long AT-rich regions (12-105 kb), which are not centromeric. These regions lack genes and consist of repetitive DNA. These AT-rich regions are enriched for H3K9me3. Interestingly, one of the AT-rich regions overlaps with a synteny breakpoint described by Croll et al (Croll et al., 2013). Due to the H3K9me3 histone modification DNA repeats and transposons appear to be silenced underlining the importance of histone modifications in maintaining genome stability.

Compared to other fungal centromeres, the centromeres of *Z. tritici* are short (up to 14 kb) and lack conserved domains (Smith et al., 2012). They have a higher AT-content than the genome average (52.5% compared to 48.3%), but as mentioned above, are not located in the longest AT-rich regions of the chromosomes. Overall, the centromeres of *Z. tritici* have similarities with the centromeres of *C. albicans* in terms of size and organization (Sanyal et al., 2004). The importance of the centromeres in chromosome stability was shown by the deletion of the centromere of the dispensable chromosome 14. This resulted in several strains in which the entire chromosome 14 was lost. Also the centromeres of progeny and parental strains were identified and characterized here to study the stability of centromeric sequences during meiosis. Here, the centromeres of the progeny strains were compared to the parental centromeres. Overall, the nucleotide diversity inside the centromeres was higher rate than the genome average. The same conclusions could be drawn after comparing the centromeric sequences of two closely related species *Z. pseudotritici* and *Z. ardabiliae*. Similar findings were shown in *C. albicans* (Padmanabhan et al., 2008). The assembly of centromeric DNA using illumine reads may however bias the analyses and comparison of nucleotide diversity.

It was also investigated here if another component of epigenetics in *Z. tritici*, RNA interference is

involved in the regulation of virulence in this pathogen. Mutants where the Dicer and three Argonaute genes were replaced by an antibiotic, were created. The virulence of the mutants was reduced documenting the importance of RNAi in the hemibiotrophic life-style and the development of pycnidia of *Z. tritici*. This further suggests that small-RNAs are involved in the host-pathogen interaction. A peculiar pattern of gene evolution was found for the Ago1 gene. Within the *Z. tritici* population two highly diverged allele groups exist. The functional divergence of these allele groups is being investigated.

Overall, it is here shown that epigenetic mechanisms plays an important role in genome stability and the in virulence of *Z. tritici*.

Chapter 5: Material and Methods

5.1 Chemicals, Enzymes, Buffers and Solutions

5.1.1 Chemicals and enzymes

All chemicals used in the described research are from the analytical grade and purchased from the companies Sigma-Aldrich, Invitrogen, Fermentas, Roche, Thermo scientific, Merck and Roth.
Restriction enzymes were purchased from New England Biolabs (NEB). The DNA polymerases phusion (Thermo scientific) and Taq (NEB), were used for standard PCR reactions. The geneRuler 1 kb DNA Ladder of Fermentas (#SM0312) was used as size indicator for gel electrophoresis analysis.

5.1.2 Kits, buffers and solutions

The following kits were used with protocols recommended by the suppliers: TopoTA Cloning Kit (Invitrogen) for direct cloning of PCR products, Wizard SV Gel and PCR clean-up system (Promega) for DNA extraction from gels and PCR purification and QIAquick plasmid Purification Kit (QIAGEN) for plasmid isolation and purification.
Standard buffers and solutions were prepared according to Sambrook *et al*. (1989). If required, buffers and solutions were autoclaved for 5 min at 121°C or filter-sterilized.

5.2 Media

5.2.1 Media for *Escherichia coli* and *Agrobacterium tumefaciens* cultivation

dYT: 20 g Trypton, 10 g Yeast-Extract, 5 g NaCl, ddH2O was added to 1 liter followed by autoclaving at 121°C for 5 min. (Sambrook et al., 1989).

dYT-agar: 20 g Trypton, 10 g Yeast-Extract, 5 g NaCl, 20 g Bacto Agar (Difco), ddH2O was added to 1 liter followed by autoclaving at 121°C for 5 min.

5.2.2 Media for *Zymoseptoria tritici* cultivation

Yeast Malt Sucrose (YMS): 4 g Yeast Extract, 4 g Malt Extract, 4 g Sucrose, ddH2O was added to 1 liter followed by autoclaving at 121°C for 5 min.

YMS-agar: 4 g Yeast Extract, 4 g Malt Extract, 4 g Sucrose, 20 g Agar, ddH2O was added to 1 liter followed by autoclaving at 121°C for 5 min.

Induction medium (After Bowler et al, 2010): $K_2HPO_4*3H_2O$ (10 mM), KH_2PO_4 (10 mM), NaCl (2.5 mM), $MgSO_4*7H_2O$ (2.0 mM), $FeSO_4*7H_2O$ (9 µM), MES buffer (40 mM), $CaCl2*2H_2O$ (0.7

mM), Glycerol (0.5%), Glucose (10mM). Final pH 5.6-5.8. Carbenicillin (100 µg/ml), Rifampicin (50 µg/ml), Kanamycin (40 µg/ml), Acetosyringone (200 µM).

5.2.3 Cell density

The absorption of cultures with *Z. tritici* cells and *A. tumefaciens* were measured at a 600 nm (OD600) with a NovaspecII photo spectrometer. *Z. tritici* has $1x10^7$ cells/ml at an OD600 of 1.0.

5.3 Strains, plasmids and oligonucleotides

5.3.1 *Escherichia coli* and *Agrobacterium tumefaciens* strains

E. coli strain TOP10 (Invitrogen) was used as host strain for plasmid constructions and amplifications.

A. tumefaciens strain AGL1 was used for *Zymoseptoria* transformations.

5.3.2 *Zymoseptoria tritici*, *Zymoseptoria pseudotritici* and *Zymoseptoria ardabiliae* strains

Table 5.3.2: *Zymoseptoria sp.* strains used in this study. Hyg) Hygromycin, Gen) Geneticin, Nat) Nourseotricin

ID	Full name	Transformed strain	Reference	Transformed plasmid	Antibiotic
ZT-01	ZTDk09 U42	Wild-type	Unpublished		
ZT-02	ZTDk09 F24	Wild-type	Unpublished		
ZT-03	ZTDk09 U21	Wild-type	Unpublished		
ZT-04	ZTDk09 U33	Wild-type	Unpublished		
ZT-05	ZTDk09 U34	Wild-type	Unpublished		
ZT-06	ZTDk09 U35	Wild-type	Unpublished		
ZT-07	ZTDk09 U43	Wild-type	Unpublished		
ZT-08	ZTDk09 U52	Wild-type	Unpublished		
ZT-09	IPO323	Wild-type	Goodwin (2011)		
ZT-10	ZT STIR01 A26b	Wild-type	Stukenbrock (2011)		
ZT-11	ZT STIR01 A48b	Wild-type	Stukenbrock (2011)		
ZP-12	Zp STIR04 3.11.1	Wild-type	Stukenbrock (2011)		
ZP-13	Zp STIR04 2.2.1	Wild-type	Stukenbrock (2011)		
ZP-14	Zp STIR04 4.3.1.	Wild-type	Stukenbrock (2011)		
ZP-15	Zp STIR04 5.3	Wild-type	Stukenbrock (2011)		
ZP-16	Zp STIR04 5.9.1	Wild-type	Stukenbrock (2011)		
ZA-17	Za STIR04 1.1.1	Wild-type	Stukenbrock (2011)		
ZA-18	Za STIR04 1.1.2	Wild-type	Stukenbrock (2011)		
ZA-19	Za STIR04 3.3.2	Wild-type	Stukenbrock (2011)		
ZA-20	Za STIR04 3.13.1	Wild-type	Stukenbrock (2011)		
SP-21	*S. passerinii* P63	Wild-type	Stukenbrock (2011)		
ZT-84	IPO323 ΔKu70	IPO323	Bowler (2010)		
ZT-89	IPO323 ΔZT90232	ZT-09 IPO323	This study	pES34 ΔZT90232	Hyg
ZT-108	IPO323 ΔZT38035	ZT-09 IPO323	This study	pES33 ΔZT38035	Hyg
ZT-117	IPO323 ΔZT47983	ZT-09 IPO323	This study	pES36 ΔZT47983	Hyg
ZT-118	IPO323 ΔKu70 GFP-CenH3	ZT-84 IPO323 ΔKu70	This study	pES42 CenH3 N-terminal-GFP	Hyg, Gen
ZT-121	IPO323 GFP-CenH3	ZT-09 IPO323	This study	pES42 CenH3 N-terminal-GFP	Hyg
ZT-127	IPO01151	Wild-type	Wittenberg (2009)		
ZT-128	IPO94269	Wild-type	Wittenberg (2009)		
ZT-129	IPO02133	Wild-type	Wittenberg (2009)		
ZT-130	IPO95052	Wild-type	Wittenberg (2009)		
ZT-135	IPO323 CENB-GFP	ZT-09 IPO323	This study	pES73 CENB-GFP	Hyg
ZT-136	IPO01151 GFP-CenH3	ZT-127 IPO01511	This study	pES42 CenH3 N-terminal-GFP	Hyg
ZT-137	IPO94269 GFP-CenH3	ZT-128 IPO94269	This study	pES42 CenH3 N-terminal-GFP	Hyg
ZT-138	IPO02133 GFP-CenH3	ZT-129 IPO02133	This study	pES42 CenH3 N-terminal-GFP	Hyg
ZT-139	IPO95052 GFP-CenH3	ZT-130 IPO95052	This study	pES42 CenH3 N-terminal-GFP	Hyg
ZT-140	IPO323 ΔKu70 CENB-GFP	ZT-84 IPO323 ΔKu70	This study	pES73 CENB-GFP	Hyg, Gen
ZT148	ZT Ger 13.1.1.2	Wild-type	Unpublished		
ZT150	ZT Ger 13.2.1.1	Wild-type	Unpublished		
ZT151	ZT Ger 13.3.1.1	Wild-type	Unpublished		
ZT153	ZT Ger 13.5.1.1	Wild-type	Unpublished		
ZT154	ZT Ger 13.6.1.1	Wild-type	Unpublished		
ZT155	ZT Fr 13.1.1.1	Wild-type	Unpublished		
ZT193	IPO323 ΔZT90030	ZT-09 IPO323	This study	pES38 ΔZT90030	Hyg
ZT194	IPO323 47983 comp	ZT-117 IPO323	This study	pES77 47983 comp	Gen
ZT195	IPO323 CenB C-terminal GFP	ZT-09 IPO323	This study	pES73 CenB with C-terminal GFP	Hyg
ZT196	IPO323 ΔCentromere 14	ZT-121 IPO323	This study	pES106 ΔCentromere 14	Hyg, Gen, Nat
ZT197	IPO323 Comp 38035 Myg	ZT-108 IPO323	This study	pES115 38035_Myg	Gen

ZT198	IPO323 90030 GFP	ZT-09 IPO323	This study	pES116 90030_GFP	Gen
ZT199	IPO323 10621 Myg	ZT-09 IPO323	This study	pES117 10621_Myg	Gen
ZT200	IPO323 90030 Myg	ZT-09 IPO323	This study	pES118 90030_Myg	Gen
ZT201	IPO323 38035 GFP	ZT-108 IPO323	This study	pES131 38035 GFP	Gen
ZT202	IPO323 Δ ZT10621	ZT-09 IPO323	This study	pES134 ZT10621 deletion	Hyg
ZT203	IPO323 90323 GFP	ZT-09 IPO323	This study	pES135 90323_GFP	Gen
ZT204	ZT STIR01 A26b ΔZT90323	ZT-10	This study	pES137 ZT90323 deletion in ZT10	Hyg
ZT205	ZT STIR01 A26b CompZT-09 ZT90323	ZT-10	This study	pES138 ZT90323 complementation in ZT10	Gen
ZT206	IPO323 CompZT-10 ZT90323	ZT-09 IPO323	This study	pES139 ZT09 ZT90323 complementation in ZT09	Gen

5.3.3 Oligonucleotides used in this study

Table 5.3.3: Oligonucleotides used in this study

Oligo name	Experiment	Sequence 5'-3'
ES447_Dicer_38035_Chr3R	QPCR for 38035	ACAACGGTTCCAGGCAGA
ES448_Dicer_38035_Chr3F	QPCR for 38035	ACACGATTCTACCCCACCC
ES451_Dicer_10621_Chr11R	QPCR for 10621	GGTGAATGGAGACCGGAGTG
ES452_Dicer_10621_Chr11F	QPCR for 10621	TGATGAACGAGGCCGGTGT
ES453_Dicer_90232_Chr1R	QPCR for 90232	GCTGAGTGGGTAGAAACGGG
ES454_Dicer_90232_Chr1F	QPCR for 90232	CGCCGCACAAACGATGAGA
ES457_Dicer_25632_Chr1R	QPCR for 25632	GCACAGTATGAGACGCCTTT
ES458_Dicer_25632_Chr1F	QPCR for 25632	GCAGAATGGAATGGCCCT
oES632_u2_EcoR1_38035	pES33	GTACCGAATTCATCCTGCCAGCTTATCCG
oES633_u3_38035	pES33	CCTTCAATATCAAAGCTTGTCAGTCAACACCTCCG
oES634_k1_38035	pES33	AGGTGTTGACTGACAAGCTTTGATATTGAAGGAGC
oES635_k2_38035	pES33	CTCTACCATGGCCGAGGCTAGCAGATCTCTATTC
oES636_d1_38035	pES33	AGAGATCTGCTAGCCTCGGCCATGGTAGAGAAC
oES637_d2_Apa1_38035	pES33	AATAGGGCCCAGCTTCTTCGCCTTCCAG
oES640_u2_EcoRI_90232	pES34	GCACTGAATTCTTCCCAGGCTCGACATAG
oES641_u3_90232	pES34	CGAGCATGCAAGCAGTTC
oES642_k2_90232	pES34	GATGCACGGTCGGAGGGCTAGCAGATCTCTATTC
oES643_k1_90232	pES34	CTGCTTGCATGCTCGGCTTTGATATTGAAGGAGC
oES644_d1_90232	pES34	AGAGATCTGCTAGCCCTCCGACCGTGCATCAAG
oES645_d2_Apa1_90232	pES34	GATAAGGGCCCGTGACGTTCGTGCATTC
ES661_K1_47983	pES36	TGCGCTGTCAATGGAGCTTTGATATTGAAGGAGC
ES662_K2_47983	pES36	GGTCGGAGATGTCGTGGCTAGCAGATCTCTATTC
ES663_U2_EcoR1_47983	pES36	ACTGCAGAATTCCTTCATGCCGGCGGAAAG
ES664_U3_47983	pES36	CCTTCAATATCAAAGCTCCATTGACAGCGCACCC
ES665_D1_47983	pES36	TAGAGATCTGCTAGCCACGACATCTCCGACCACG
oES666_D2_SBF1_47983	pES36	ATCGTCTCCTGCAGGTGCGATCACGCCAAGAAC
ES707_K1_CenH3	pES42	TCTTGGAATTGCACATAACATGGTGAGCAAGGG
ES708_K2_CenH3	pES42	GCCTCCTCCGCCGCCTCCCTTGTACAGCTCGTCCAT
ES709_D1_CenH3	pES42	GGAGGCGGCGGAGGAGGCATGGCGCGAACAAAGCAA
ES710_D2_CenH3	pES42	GGCGCGCCCCAGGCGGTTAAGGTGTC
ES711_Pr2_CenH3	pES42	TTGCTCACCATGTTATGTGCAATTCCAAGATTC
ES712_Pr1_CenH3	pES42	CGAGGAAGTTTGTGTGCGGGAGGCTCAGTCGTC
ES713_Hyg1_CenH3	pES42	ATGTGTTTGAAGGAGCTAGTCGACAGAAGATGATA
ES714_Hyg2_CenH3	pES42	ACACAAACTTCCTCGCTCTATTCCTTTGCCCTCGG
ES715_U3_CenH3	pES42	TCTTCTGTCGACTAGGCTCCTTCAAACACATTC
ES716_U2_CenH3	pES42	GGGCCCCCAAGATCCGGCCGTTTC
oES763_47983_Fq	QPCR for 47983	CTCGACCTTCATCCAGAACC
oES764_47983_Rq	QPCR for 47983	TGTACTCCAAGGCATGTTCG
oES786_99044_Fq	QPCR for GAPDH	TGTCGGCAAGGTCATTCCAG
oES787_99044_Rq	QPCR for GAPDH	ATGCGGCAGGTCAAGTCAAC
oES805_Rev-RT-primer	QPCR for 90232 (ZT10 and ZT10)	TAAGGCGGAGGAGTTC
oES930_Fwd_UF_CenB	ES73	TCGATGGGATCCGCCCACAAATCACTGGTC
oES931_Rev_UF_CenB	ES73	GCCTCCGCCGCCTCCGCCGGCGTTGCTTCCCGAAGG
oES932_Fwd_GFP_CenB	ES73	GGCGGAGGCGGCGGAGGCATGGTGAGCAAGGGCGAG
oES933_Rev_GFP_CenB	ES73	TTGCGCGGCTAAGACCTATTCCTTTGCCCTCGG
oES934_Fwd_DF_CenB	ES73	AGGGCAAAGGAATAGGTCTTAGCCGCGCAAATC
oES935_Rev_DF_CenB	ES73	TCAGCAGGATCCATGCGCTTCCGGGATATG
oES1015_UF_Fwd	pES77	CCTGCAGGTGAGACATCGGTGCATCGAGAC
oES1016_UF_Rev	pES77	CCTTCAATATCAGGTTGCCGGAGATGTCGTCGAAAC

oES1017_Gen_Fwd	pES77	TCGACGACATCTCCGGCAACCTGATATTGAAGGAGC
oES1018_Gen_Rev	pES77	CGATCTCGTGGTCGGACCTCAGAAGAACTCGTCAAG
oES1019_DF_Fwd	pES77	ACGAGTTCTTCTGAGGTCCGACCACGAGATCGGATG
oES1020_DF_Rev	pES77	TGCGAATTCTGTGCCAACATGCTCTTC
ES1192_ZtCen1F1	ZtCen1F1	TTCGCGGCTATCTTCGTACT
ES1193_ZtCen1R1	ZtCen1R1	CTACCTTACTGCCCGCAGAG
ES1194_ZtCen1F2	ZtCen1F2	AACCTCGAAGCGGTAGGATT
ES1195_ZtCen1R2	ZtCen1R2	CTAGCCTTACGGGCCTTCTT
ES1196_ZtCen1F3	ZtCen1F3	GCCGAAGGAAGGGCTTAGAT
ES1197_ZtCen2F1	ZtCen2F1	GGTCGTACCCAGAGCATTGT
ES1198_ZtCen2R1	ZtCen2R1	CTAGCTCTCCGGACCTCCTT
ES1199_ZtCen2F2	ZtCen2F2	ACATGCGGTCCAATTTAAGC
ES1200_ZtCen2R2	ZtCen2R2	CCAAAGGACCTGGCTAACAA
ES1201_ZtCen3F1	ZtCen3F1	CGCTTAGTTAGCCTGCTGCT
ES1202_ZtCen3R1	ZtCen3R1	AGCTCTCGTTTCTGGGAACA
ES1203_ZtCen3F2	ZtCen3F2	AACACCTTCGTCCAAACAGG
ES1204_ZtCen3R2	ZtCen3R2	CGGATCTTGAAGGACACCAT
ES1205_ZtCen4F1	ZtCen4F1	CTGATTTCGCCTCCAGACTC
ES1206_ZtCen4R1	ZtCen4R1	GCCTGGCCAAAAAGACTAAA
ES1207_ZtCen4F2	ZtCen4F2	ATCAGCAGTCCAGACCGAGT
ES1208_ZtCen4R2	ZtCen4R2	TAGTGTTAGAGGGGCGGAGA
ES1209_ZtCen5F1	ZtCen5F1	TGCTACGAAGTGCAACAAGG
ES1210_ZtCen5R1	ZtCen5R1	GTGGGCTAGTCTCGTTGAGG
ES1211_ZtCen5F2	ZtCen5F2	GGGTCGCGCTAGACATACAT
ES1212_ZtCen5R2	ZtCen5R2	ATGCCCTTCGAGTCGATAAC
ES1213_ZtCen5F3	ZtCen5F3	GGGCATAAAACAGAATTCCTC
ES1214_ZtCen6F1	ZtCen6F1	TAAGGGGAAGGCAGCTAACA
ES1215_ZtCen6R1	ZtCen6R1	AACGATCCGAGAGCGTCTTA
ES1216_ZtCen6F2	ZtCen6F2	TAAGGGGAAGGCAGCTAACA
ES1217_ZtCen6R2	ZtCen6R2	GTCGGTTTTGGATCGCTTTA
ES1218_ZtCen7F1	ZtCen7F1	GGATTTCAGGGTGTCGAGAA
ES1219_ZtCen7R1	ZtCen7R1	TCGATGTTGTGATTCGCATT
ES1220_ZtCen7F2	ZtCen7F2	AAAGCCATCGAGCAAAAGAA
ES1221_ZtCen7R2	ZtCen7R2	ATGGCAGGAGAGGTATGTCG
ES1222_ZtCen8F1	ZtCen8F1	CGTCCTCGAAGGAGTCGTAG
ES1223_ZtCen8R1	ZtCen8R1	TACCATGTGCAATCGGTGTT
ES1224_ZtCen8F2	ZtCen8F2	TATCGAGGGTAGGCAGAGGA
ES1225_ZtCen8R2	ZtCen8R2	GCCGACCTAGCACAACTAGC
ES1226_ZtCen8F3	ZtCen8F3	GCCGAAGGACTTCAAAACAG
ES1227_ZtCen9F1	ZtCen9F1	AGAGATCGTCCGTGACAGGT
ES1228_ZtCen9R1	ZtCen9R1	CCTGTCGCCTCTTCCACTAC
ES1229_ZtCen9F2	ZtCen9F2	CGCCAAACCTAATTTGCACT
ES1230_ZtCen9R2	ZtCen9R2	CGTCGCTTCCAGAGGTGTAT
ES1231_ZtCen10F1	ZtCen10F1	TTCGAGCGTCGACATAGAAT
ES1232_ZtCen10R1	ZtCen10R1	CTTTTGTCCGGTCTCTCCTG
ES1233_ZtCen10F2	ZtCen10F2	TCTAGGACGCCAAGATTGCT
ES1234_ZtCen10R2	ZtCen10R2	CAAGGTCGTACGTGCTCTGA
ES1235_ZtCen11F1	ZtCen11F1	CCTTCGAGTGAAGGCTGAAC
ES1236_ZtCen11R1	ZtCen11R1	TGCAAGCTTGTAAGCGAATC
ES1237_ZtCen11F2	ZtCen11F2	GCTCTTATAACTTTACTACCCGGATCT
ES1238_ZtCen11R2	ZtCen11R2	CGACGTCTGCTTCCGACTAC
ES1239_ZtCen12F1	ZtCen12F1	CACGAGAAACGCGAGATACA
ES1240_ZtCen12R1	ZtCen12R1	GGTGTGATGAGGCTTTTGGT
ES1241_ZtCen12F2	ZtCen12F2	GCTCGACAAATGCTCGGTAT
ES1242_ZtCen12R2	ZtCen12R2	TGCAGTTCCGAACAAGCTAA

ES1243_ZtCen13F1	ZtCen13F1	ATCCGACCATCTCCACTCTG
ES1244_ZtCen13R1	ZtCen13R1	CTGCTGTGAGTGGTGAGGAA
ES1245_ZtCen13F2	ZtCen13F2	TCTTACTACCGCGGTCCAGT
ES1246_ZtCen13R2	ZtCen13R2	CACAGCAAGGCTGCACTTAG
ES1247_ZtCen13F3	ZtCen13F3	CTCCTTCCAGGCCTACCTCT
ES1248_ZtCen14F1	ZtCen14F1	CTTTGGCGAGGAACTTCTTG
ES1249_ZtCen14R1	ZtCen14R1	ATTCCAAATCATCCCCTTCC
ES1250_ZtCen14F2	ZtCen14F2	TCGCGTGCTTCGTAGTAATG
ES1251_ZtCen14R2	ZtCen14R2	AACCGACCGACTCTCACATC
ES1252_ZtCen14F3	ZtCen14F3	CCTCACCAAGGAGGTTACCA
ES1253_ZtCen15F1	ZtCen15F1	GCGATCACGATGTTCAAATG
ES1254_ZtCen15R1	ZtCen15R1	GGCGTGGAAACGACACTTAT
ES1255_ZtCen15F2	ZtCen15F2	TAAGCGCGGTCTAGTTGGAT
ES1256_ZtCen15R2	ZtCen15R2	TTTCGCACCGATAGAAGGTT
ES1257_ZtCen16F1	ZtCen16F1	GCCAGGCGCAATTGTTATAC
ES1258_ZtCen16R1	ZtCen16R1	TGTCGACAAAGGACAACGAG
ES1259_ZtCen16F2	ZtCen16F2	CAAAAAGCAGTCGAGCGTTA
ES1260_ZtCen16R2	ZtCen16R2	CGCCTTCTGTGCGATAAAGT
ES1261_ZtCen17F1	ZtCen17F1	GCCGATGACTACACCTTCGT
ES1262_ZtCen17R1	ZtCen17R2	GGCTTAAGGGGGAGCTAGAA
ES1263_ZtCen17F2	ZtCen17F2	AGACAGCCCTACATGCTCGT
ES1264_ZtCen17R2	ZtCen17R2	ACTGCCAGTTCCGTGAAGTC
ES1265_ZtCen18F1	ZtCen18F1	TGAGGTGTACCCCCTACGAG
ES1266_ZtCen18R1	ZtCen18R1	TGCTTTCTTCCAATGCTGTG
ES1267_ZtCen18F2	ZtCen18F2	GGGTGCGGATTGACAGTTAT
ES1268_ZtCen18R2	ZtCen18R2	TGTCTAATCCAGACGCAACG
ES1269_ZtCen19F1	ZtCen19F1	CACGGCACCATCACTGTATC
ES1270_ZtCen19R1	ZtCen19R1	CTGAGATTGTTGACGCTGGA
ES1271_ZtCen19F2	ZtCen19F2	GCATAGACCTTTGCCTCCAG
ES1272_ZtCen19R2	ZtCen19R2	TTAAGTTGCGGGTGTGACAG
ES1273_ZtCen20F1	ZtCen20F1	AAAGAGAGAGTCCGCGTCTG
ES1274_ZtCen20R1	ZtCen20R1	TATGGGCACGAACCAAGAGT
ES1275_ZtCen20F2	ZtCen20F2	AGTTGTTCACGCAGCACTTG
ES1276_ZtCen20R2	ZtCen20R2	ATGCGGTAAACACGGAGAAC
ES1277_ZtCen21F1	ZtCen21F1	GTCCTGAACGGACTGGTGTT
ES1278_ZtCen21R1	ZtCen21R1	GCGACGCAAAAGCATATGTA
ES1279_ZtCen21F2	ZtCen21F2	CTGCTCTGCTGCTTGTTCTG
ES1280_ZtCen21R2	ZtCen21R2	CCAACTTGGAGGATCTGGAA
ES1281_ZtCen21F3	ZtCen21F3	TCCTCGGAAACGTTACAACC
ES1291_14DUpRev	pES106	CATGCCCTGCCCCTAGAATGGCTTTCAACACGAAGGCT
ES1292_14DNatFwd	pES106	TTCGTGTTGAAAGCCATTCTAGGGGCAGGGCATGCTCA
ES1293_14DNatRev	pES106	TGTGGGTGTTCAGGAAGTCCGCTTTGATATTGAAGGAGCATTT
ES1294_14DDoFwd	pES106	TCCTTCAATATCAAAGCGGACTTCCTGAACACCCACACCAT
ES1299_13DUpRev	pES107	ATGCCCTGCCCCTAGCGACGGTAAAGTGGAGGTGTCG
ES1300_13DNatFwd	pES107	CCTCCACTTTACCGTCGCTAGGGGCAGGGCATGCTC
ES1301_13DNatRev	pES107	CCGCTTTGATATTGAAGGAGCATTTTTTGGGCTTGGCTGG
ES1302_13DDoFwd	pES107	GCTCCTTCAATATCAAAGCGGCATCGACCTTCTCCAAGGAAAC
ES1303_13DDoRev	pES107	TCTCGAGGTTTAAACGGGGCCTCCTGAATCCTGTTCACCA
ES1304_13DBaFwd	pES107	GTGAACAGGATTCAGGAGGCCCCGTTTAAACCTCGAGAGATC
oES1348_ES10_qPCR	QPCR for 90232 (ZT10 and ZT10)	TCAACACCGCTCTCAG
ES1373_bbFwd106	pES106	TCGTCATCGCCAAAGTACAGAAGCTTATGGCCGTATCCGCAATGTG
ES1374_inRev106	pES106	CACATTGCGGATACGGCCATAAGCTTCTGTACTTTGGCGATGACGA
ES1375_inFwd106	pES106	GAACTGACAGAACCGCAACGAAGCTTCCATCTCTGTCGGCCAGTAT
ES1376_bbRev106	pES106	ATACTGGCCGACAGAGATGGAAGCTTCGTTGCGGTTCTGTCAGTTC
ES1377_bbFwd107	pES106	GTGAACAGGATTCAGGAGGCGAATTCATGGCCGTATCCGCAATGTG
ES1378_inRev107	pES106	CACATTGCGGATACGGCCATGAATTCGCCTCCTGAATCCTGTTCAC
ES1379_inFwd107	pES106	GAACTGACAGAACCGCAACGGAATTCCAGGAGAAGGCGAGTGTAGG
ES1380_bbRev107	pES106	CCTACACTCGCCTTCTCCTGGAATTCCGTTGCGGTTCTGTCAGTTC
oES1505_115_Gen_Fwd	pES115	GAAGTTGAGGTTGGGTAACGGGACCACGTACCTGATATTG

oES1506_115_Gen_Rev	pES115	GTGGTTCTCTACCATGGCCGTCGATCGCTCAGAAGAACTC
oES1507_115_Gene_Fwd_Myc	pES115	AGAAGCTGATCTCAGAGGAGGACCTGTCAGGAAACCCAAGAGGCCG
oES1508_115_Gene_Rev	pES115	CAATATCAGGTACGTGGTCCCGTTACCCAACCTCAACTTC
oES1509_115_UF_Fwd	pES115	TACGAATTCTTAATTAAGATCATATCCTGCCAGCTTATCC
oES1510_115_UF_Rev_myc	pES115	TCCTCTGAGATCAGCTTCTGCTCCATTTCGAGAGCTCGAAGAGAAA
oES1511_130_Gene_Fwd_GFP	pES115	TGTACAAGGGAGGCGGCGGAGGAGGCTCAGGAAACCCAAGAGGCCG
oES1512_130_UF_Rev_GFP	pES115	AACAGCTCCTCGCCCTTGCTCACCATTTCGAGAGCTCGAAGAGAAA
oES1513_130_GFP_Rev	pES115	GGGTTTCCTGAGCCTCCTCCGCCGCCTCCCTTGTACAGCTCGTCCA
oES1514_130_GFP_Fwd	pES115	TTTCTCTTCGAGCTCTCGAAATGGTGAGCAAGGGCGAGGAGCTGTT
oES1515_117_DF_Fwd	pES117	GAGTTCTTCTGAGCGATCGAAACGAGGCATGACGGATGAG
oES1516_117_DF_Rev	pES117	TCGAGGGTACCGAGCTCGATTCTCCGTCCTCTGGCAACAC
oES1517_117_Gen_Fwd	pES117	GACTTCATGATGGTGGGAACGGACCACGTACCTGATATTG
oES1518_117_Gen_Rev	pES117	CTCATCCGTCATGCCTCGTTTCGATCGCTCAGAAGAACTC
oES1519_117_Gene_Fwd_Myc	pES117	GAAGCTGATCTCAGAGGAGGACCTGGCTGGTGCAGCCAAGAACCG
oES1520_117_Gene_Rev	pES117	CAATATCAGGTACGTGGTCCGTTCCCACCATCATGAAGTC
oES1521_117_UF_Fwd	pES117	TACGAATTCTTAATTAAGATCTCTTCGCGATGAATGAGAG
oES1522_117_UF_rev_Myc	pES117	CCTCTGAGATCAGCTTCTGCTCCATGTTGACAGTGCGGTTGGTGG
oES1523_131_Gene_GFP_Fwd	pES131	AGGGAGGCGGCGGAGGAGGCGCTGGTGCAGCCAAGAACCG
oES1524_131_GFP_Rev	pES131	GCCTCCTCCGCCGCCTCCCTTGTACAGCTCGTCCATGCCG
oES1525_131_GFP_Fwd	pES131	CCACCAACCGCACTGTCAACATGGTGAGCAAGGGCGAGGA
oES1526_131_UF_GFP_Rev	pES131	TCCTCGCCCTTGCTCACCATGTTGACAGTGCGGTTGGTGG
oES1527_u3_10621	pES134	GCTCCTTCAATATCAAAGCGAGGATGTTAGATGAGCTGC
oES1528_k1_10621	pES134	GCAGCTCATCTAACATCCTCGCTTTGATATTGAAGGAGC
oES1529_k2_10621	pES134	CGTCATGCCTCGTTGCTCGTTGGCTAGCAGATCTCTATTC
oES1530_d1_10621	pES134	GAATAGAGATCTGCTAGCCAACGAGCAACGAGGCATGACG
oES1531_u1_10621	pES134	GATTGCTTGCTGGATCAC
oES1532_d3_10621	pES134	ATCGTCGCCATAATCCTC
oES1533_u2_90232	pES135	TACGAATTCTTAATTAAGATGTATGCGGACAGCGCTCTGG
oES1534_u3_90232	pES135	CCTCGCCCTTGCTCACCATGGCGGCGTGGTTGAAGTCTG
oES1535_gfp-f_90232	pES135	CAGACTTCAACCACGCCGCCATGGTGAGCAAGGGCGAGG
oES1536_gfp-r_90232	pES135	CACCAAGTACCTGCTTTAGGGCCTCCTCCGCCGCCTCCC
oES1537_orf-f_90232	pES135	GGGAGGCGGCGGAGGAGGCCCTAAAGCAGGTACTTGGTG
oES1538_orf-r_90232	pES135	CTTGACGAGTTCTTCTGAGTCGGACCATGTTGATGATGC
oES1539_k1_90232	pES135	GCATCATCAACATGGTCCGACTCAGAAGAACTCGTCAAG
oES1540_k2_90232	pES135	GACAGAAACAGTGCGGTGTCCAACCTGATATTGAAGGAGC
oES1541_d1_90232	pES135	GCTCCTTCAATATCAGGTTGGACACCGCACTGTTTCTGTC
oES1542_d2_90232	pES135	TCGAGGGTACCGAGCTCGATTGCGTGACGTTCGTGCATTC
oES1543_u3_90232_myc	pES135	TAGGGAGGTCCTCTTCCGAGATGAGCTTCTGCTCCATGGCGGCGTGGTTGAAGTCTG
oES1544_orf-f_90232_myc	pES135	CGCCATGGAGCAGAAGCTCATCTCGGAAGAGGACCTCCCTAAAGCAGGTACTTGGTC
oES1572_u3_D90232	pES138	GCTCCTTCAATATCAAAGCGACTGCGTGATCGACTTGGC
oES1573_k1_D90232	pES138	GCCAAGTCGATCACGCAGTCGCTTTGATATTGAAGGAGC
oES1574_k2_D90232	pES138	GTGATGAGTCGTGAAGATCGGGCTAGCAGATCTCTATTC
oES1575_d1_D90232	pES138	GAATAGAGATCTGCTAGCCCGATCTTCACGACTCATCAC
oES1576_u3	pES137	CAAGTACCTGCTTTAGGCATGACTGCGTGATCGACTTGGC
oES1577_orf-f	pES137	GCCAAGTCGATCACGCAGTCATGCCTAAAGCAGGTACTTG
oES1578_k2	pES137	CGAACAGCAGCTGTGAAGTGGCAACCTGATATTGAAGGAGC
oES1579_d1	pES137	GCTCCTTCAATATCAGGTTGCCACTTCACAGCTGCTGTTCG

oES1590_u3	pES139	CTTGACCATGACTGCGTGATGGCGTGGTTGAAGTCTGGTG
oES1591_orf-f	pES139	CACCAGACTTCAACCACGCCATCACGCAGTCATGGTCAAG
oES1592_orf-r	pES139	TCTTGACGAGTTCTTCTGAGCATGTTTCGCCGAACAGCAG
oES1593_k1	pES139	CTGCTGTTCGGCGAAACATGCTCAGAAGAACTCGTCAAGA
oES1594_k2	pES139	CCTGTTGAGCGGTCACAGTTCAACCTGATATTGAAGGAGC
oES1595_d1	pES139	GCTCCTTCAATATCAGGTTGAACTGTGACCGCTCAACAGG
oES1612_UF_Rev	pES116	CTCGCCCTTGCTCACCATAGTTCGTTCGTTCGTGTTGG
oES1613_UF_Fwd	pES116	TACGAATTCTTAATTAAGATCGCGATGGAGATGGGAATAC
oES1614_GFP_Rev	pES116	CATGCGTTGGTCGAGCCTCCGCCGCCTCCGCCCTTGTACAGCTCGTCCAT
oES1615_GFP_Fwd	pES116	CCAACACGAACGAACGAACTATGGTGAGCAAGGGCGAG
oES1616_G418_Rev	pES116	GAATTGGCCTGTCAAAGCTCTCGATCGCTCAGAAGAACTC
oES1617_G418_Fwd	pES116	GAGTTGGAGTTGCAGTTGGTGGACCACGTACCTGATATTG
oES1618_Gene_Rev	pES116	CAATATCAGGTACGTGGTCCACCAACTGCAACTCCAACTC
oES1619_Gene_Fwd	pES116	GAGCTGTACAAGGGCGGAGGCGGCGGAGGCTCGACCAACGCATGGACTCA
oES1620_DF_Rev	pES116	TCGAGGGTACCGAGCTCGATGGCGATAGCACAGCAATACC
oES1621_DF_Fwd	pES116	GAGTTCTTCTGAGCGATCGAGAGCTTTGACAGGCCAATTC
oES1622_UF_Rev	pES118	CAGGTCCTCCTCTGAGATCAGCTTCTGCTCCATGGTTGCTGAGTTCGTTCG
oES1623_Gene_Fwd	pES118	ATGGAGCAGAAGCTGATCTCAGAGGAGGACCTGTCGACCAACGCATGGACTC
oES1624_UF_Fwd	pES132	TACGAATTCTTAATTAAGATCTCGCTATCAAGCCCATGAC
oES1625_UF_Rev	pES132	TCCTCGCCCTTGCTCACCATGGACGGCGGATTTGAAGGCA
oES1626_GFP_Fwd	pES132	TGCCTTCAAATCCGCCGTCCATGGTGAGCAAGGGCGAGGA
oES1627_GFP_Rev	pES132	GCGCAGAGCCTCCTCCGCCGCCTCCCTTGTACAGCTCGTCCATGC
oES1628_Gene_Fwd	pES132	GTACAAGGGAGGCGGCGGAGGAGGCTCTGCGCCATCAGATCTTCT
oES1629_Gene_Rev	pES132	CAATATCAGGTACGTGGTCCCGGAGATGTCGTCGAAACAC
oES1630_G418_Fwd	pES132	GTGTTTCGACGACATCTCCGGGACCACGTACCTGATATTG
oES1631_G418_Rev	pES132	TAATCATCCGATCTCGTGGTTCGATCGCTCAGAAGAACTC
oES1632_DF_Fwd	pES132	GAGTTCTTCTGAGCGATCGAACCACGAGATCGGATGATTA
oES1633_DF_Rev	pES132	TCGAGGGTACCGAGCTCGATGGCGGAGAATCTCCCTAAAC

5.3.4 *Agrobacterium tumefaciens* plasmids to create *Zymoseptoria sp.* mutants

pES17 D0893pNOVpGpda SDHB_H267Y tTrpC Hyg (Bowler et al, 2010).

Plasmids contains multiple restriction sides and was used as a vector for ligations. Ampicillin and Kanamycin were used for selection in E. coli. The plasmid contains the Hygromycin-resistance cassette under the control of the pTrpC promoter and was used as template for the amplification of the Hygromycin cassette.

pES21 pNOV pGpda SDHB H267Y tTrpC Hyg (Bowler et al, 2010).

Plasmids contains multiple restriction sides and was used as a vector for ligations. Ampicillin and Kanamycin were used for selection in E. coli. The plasmid contains the Hygromycin-resistance cassette under the control of the pTrpC promoter and was used as template for the amplification of the Hygromycin cassette.

pES26 pZErOR_2-GFP-lox (Honda et al, 2009).

Contains Hygromycin-resistance and C-terimal GFP with 6x G-linker.

pES27 CenH3 N-terminal GFP (This study).

Contains CenH3 with 1kb surrounding regions. The CenH3 is N-terminal tagged with a 6x G-linker and GFP obtained from pES 26. The selection antibiotic Hygromycin-resistance is adjacent to the CenH3. pES21 was used as vector.

pES33 ZT38035 deletion (This study).

Primers 632, 633, 636, 637, 634 and 635 were used to amplify the fragments.
Vector pES21 was digested with ApaI and EcoRI. In the Plasmid contains 1 kb surrounding regions of 38035 and the ORF of 38035 is replaced by the selection antibiotic is Hygromycin.

pES34 ZT90232 deletion (This study).

Primers 640, 641, 644, 645, 643 and 642 were used to amplify the fragments.
Vector pES21 was digested with ApaI and EcoRI. In the Plasmid contains 1 kb surrounding regions of 90323 and the ORF of 90323 is replaced by the selection antibiotic is Hygromycin.

pES36 ZT47983 deletion (This study).

Primers 663, 664, 661, 662, 665 and 666 were used to amplify the fragments.
Vector pES17 was digested with EcoRI and SBF-I. In the Plasmid contains 1 kb surrounding regions of 47983 and the ORF of 47983 is replaced by the selection antibiotic is Hygromycin.

pES38 ZT90030 deletion (This study).

Primers 1465, 668, 671, 672, 1465 and 1466 were used to amplify the fragments.
Vector pES17 was digested with ApaI. In the Plasmid contains 1 kb surrounding regions of 90030 and the ORF of 90030 is replaced by the selection antibiotic is Hygromycin.

pES42 CenH3_N-terminal-GFP (This study).

Primers 715, 713, 714, 716, 711, 707, 708, 709 and 710 were used to amplify the fragments.
Vector pES21 was digested with Bsp120I and ASCI. In the Plasmid contains 1 kb surrounding regions of CenH3. CenH3 is N-terminal tagged with a 6x G-linker and GFP, the selection antibiotic is Hygromycin

pES61 modified mcs of pES1 (Happel, unpublished).

Plasmids contains multiple restriction sides and was used as a vector for ligations. Kanamycin was used for selection in E. coli. The plasmid contains the Hygromycin-resistance cassette under the control of the pTrpC promoter and was used as template for the amplification of the Hygromycin cassette.

pES73 CenB with C-terminal GFP antibiotic (This study).

Primers 930, 931, 934, 935, 932 and 933 were used to amplify the fragments.

Vector pES61 was digested with BamHI. Plasmid contains 1 kb surrounding regions of CenB. Cenb is C-terminal tagged with a 6x G-linker and GFP, the selection antibiotic is Hygromycin.

pES77 ZT47983 complementation (This study).

Primers 1023, 1016, 1019, 1020, 1017 and 1018 were used to amplify the fragments.

Vector pES61 was digested with EcoRI and SBF-I. Plasmid contains the ORF of 47983 and 1 kb surrounding regions, the selection antibiotic is Geneticin.

pES106 Centromere 14 deletion (This study).

Primers 1375, 1291, 1294, 1374, 1292, 1293, 1373 and 1376 were used to amplify the fragments.

Vector pES61 was digested with HindIII. Plasmid contains 1 kb surrounding regions of centromere 14 and centromere 14 is replaced with the selection antibiotic is Nourseotricin.

pES107 Centromere 13 deletion (This study).

Primers 1379, 1299, 1304, 1378, 1300, 1301, 1377 and 1380 were used to amplify the fragments.

Vector pES61 was digested with EcoRI. Plasmid contains 1 kb surrounding regions of centromere 13 and centromere 13 is replaced with the selection antibiotic is Nourseotricin.

pES115 ZT38035_Myg (This study).

Primers 1509, 1510, 1503, 1504, 1507, 1508, 1505 and 1506 were used to amplify the fragments.

Vector pES61 was digested with EcoRV. Plasmid contains the ORF of ZT38035 tagged with Myg and 1 kb surrounding regions, the selection antibiotic is Geneticin.

pES116 ZT90030_GFP (This study).

Primers 1613, 1612, 1620, 1621, 1619,1618, 1617, 1616, 1615 and 1619 were used to amplify the fragments. Vector pES61 was digested with EcoRV. Plasmid contains the ORF of ZT90030 tagged with 6 x G-linker and GFP and 1 kb surrounding regions of the ORF, the selection antibiotic is

Geneticin.

pES117 ZT10621_Myg (This study).

Primers 1521, 1522, 1515, 1516, 1519, 1520, 1517 and 1518 were used to amplify the fragments. Vector pES61 was digested with EcoRV. Plasmid contains the ORF of ZT10621 tagged with Myg and 1 kb surrounding regions, the selection antibiotic is Geneticin.

pES118 ZT90030_Myg (This study).

Primers 1613, 1622, 1620, 1621, 1623, 1618, 1617 and 1616 were used to amplify the fragments. Vector pES61 was digested with EcoRV. Plasmid contains the ORF of ZT90030 tagged with Myg and 1 kb surrounding regions, the selection antibiotic is Geneticin.

pES131 ZT38035 GFP (This study).

Primers 1509, 1512, 1514, 1513, 1511, 1508, 1505, 1506, 1504 and 1503 were used to amplify the fragments. Vector pES61 was digested with EcoRV. Plasmid contains the ORF of ZT38035 tagged with 6 x G-linker and GFP and 1 kb surrounding regions, the selection antibiotic is Geneticin.

PES131 ZT47983 GFP (This study).

Primers 1521, 1526, 1523, 1520, 1516, 1515, 1517, 1518, 1525 and 1524 were used to amplify the fragments. Vector pES61 was digested with EcoRV. Plasmid contains the ORF of ZT47983 N-terminal tagged 6x G-linker and GFP and 1 kb surrounding regions, the selection antibiotic is Geneticin.

pES 134 ZT10621 deletion (This study).

Primers 1521, 1527, 1528, 1529, 1530 and 1516 were used to amplify the fragments.
Vector pES61 was digested with EcoRV. Plasmid contains 1 kb surrounding regions of ZT10621 and the ORF is replaced with the selection antibiotic is Hygromycin.

pES135 ZT90323_GFP (This study).

Primers 1533, 1534, 1535, 1536, 1537, 1538, 1539 and 1540 were used to amplify the fragments. Vector pES61 was digested with EcoRV. Plasmid contains the ORF of ZT90323 N-terminal tagged 6x G-linker and GFP and 1 kb surrounding regions, the selection antibiotic is Geneticin.

pES 137 ZT90323 deletion in ZT10 (This study).

Primers 1533, 1572, 1573, 1574, 1575 and 1542 were used to amplify the fragments.
Vector pES61 was digested with EcoRV. Plasmid contains 1 kb surrounding regions of ZT90323 and the ORF is replaced with the selection antibiotic is Hygromycin.

pES 138 ZT90323 complementation in ZT10 (This study).
Primers 1533, 1576, 1577, 1538, 1539, 1578, 1579 and 1542 were used to amplify the fragments.
Vector pES61 was digested with EcoRV. Plasmid contains the ORF of 90323 of ZT09 and 1 kb surrounding regions, the selection antibiotic is Geneticin.

pES 139 ZT09 ZT90323 complementation in ZT09 (This study).
Primers 1533, 1590, 1591, 1592, 1593, 1594, 1595 and 1542 were used to amplify the fragments.
Vector pES61 was digested with EcoRV. Plasmid contains the ORF of 90323 of ZT10 and 1 kb surrounding regions, the selection antibiotic is Geneticin.

5.4 Microbiological methods

5.4.1 *Escherichia coli* methods

E. coli was grown overnight in liquid dYT-medium or on dYT-agar plates at 37°C. For liquid dYT-medium 200 rpm shaking was applied. Antibiotics were added to the media with the following final concentrations: Ampicillin (100 µg/ml) or Kanamycin (50 µg/ml). For TopoTA cloning x-Gal ß (60 µl/plate) was added.

5.4.1.2 TopoTA cloning

The TopoTA cloning was performed according the TopoTA Manual (Invitrogen).

5.4.2 *Agrobacterium tumefaciens* methods

5.4.2.1 Electro-transformation of *Agrobacterium tumefaciens*

Approximately 100 ng of plasmid-DNA was added to 50 µl of electrocompetent *Agrobacterium tumefaciens* cells and was incubated for one minute on ice. After incubation the suspension of cells was transferred to a, pre-chilled, electroporation cuvette (with 1 mm distance between the electrodes) and electroporated twice with the following conditions: 1,8 kV, 200 Ohm, 25 µF. Directly after electrophoresis, 950 µl pre-warmed dYT-medium was added to the cell suspension and the content of the electrophoresis cuvette was transferred to a test tube. After incubating for 2 hours at 28°C, while shaking, 200 µl of the transformed *A. tumefaciens* cells were plated on half of

a dYT-agar plate (Carbenicillin (100 µg/ml), Rifampicin (50 µg/ml) and Kanamycin (40 µg/ml) and streaked out twice to obtain single colonies. After 2 days of incubating at 28°C, 4 single colonies where inoculated in 2.5 ml of liquid dYT medium (Carbenicillin (100 µg/ml), Rifampicin (50 µg/ml) and Kanamycin (40 µg/ml)) and incubated for approximately 24 hours at 28°C while shaking at 200 rpm.

5.4.3 *Zymoseptoria tritici* methods

5.4.3.1 Cultivation of *Zymoseptoria tritici*

Z. tritici was grown for 4-6 days on YMS-agar at 18°C. Or for 3 days in YMS liquid medium, at 18°C, at 140 or 200 rpm until the cultures had reached OD600 of 0.6-0.8.

For a mutant strain, the media was supplemented with Hygromycin (150 µg/ml), Geneticin (150 µg/ml) or Nourseotricin (35 µg/ml).

5.4.3.2 Transformation of *Zymoseptoria tritici* by *Agrobacterium tumefaciens* mediated transformation

Transformed *A. tumefacien*s cells were diluted with induction medium to an OD600 of 0.15, in a total volume of 5ml. After the dilution the culture was incubated for 4 hours at 28°C at 200 rpm until an OD600 of ~0.3 was reached.

Freshly grown *Z. tritici* cells were isolated from a YMS-agar plate and re-suspend in 1 ml of sterile water to obtain a dilution of 10,000 cells per µl. To transform *Z. tritici*, 100 µl of the obtained *Z. tritici* dilution and 100 µl of the *A. tumefaciens* culture were mixed and the mixture was plated onto cellulose nitrate discs which were placed on top of the induction medium agar and incubated at 18°C for 2 to 3 days. Hereafter, each cellulose nitrate disc was transferred to an YMS-agar plate containing a selection antibiotic and Cefotaxime (200 µg/ml) and incubated at 18°C for 15-20 days. When single colonies appeared, the colonies were streaked out and incubated for a week at 18°C.

5.4.3.3 Planta experiments with *Zymoseptoria tritici*

5.4.3.3.1 Plant material

Triticum aestivum (e.g. Obelisk) was pre-germinated in a phytochamber at 22°C with a 16-h light period. After four days the seedlings were transferred to pots (4 seedlings per pot with a size of 10 x 10 cm) with non-sterile soil and grown for an additional 7 days at 22°C with a 16-h light period. Prior to the infection, a distinct area of the second leaf (10-15cm) was marked.

5.4.3.3.2 *Zymoseptoria tritici* cultures

Prior to the infection, the spore density of the inoculation culture was adjust to $1x10^7$ cells/ml in non-sterile tap water and Tween20 to a concentration of ca 0.1% was added.

5.4.3.3.3 Infection of *Triticum aestivum* with *Zymoseptoria tritici*

To infect *T. aestivum*, the *Z. tritici* spore solution was painted onto the second leaf by using a thin sterile paint-brush. The infected *T. aestivum* plants were incubated for 48-h under high humidity, obtained by packing the plants and pots into plastic bags with sufficient water in the bottom (0.5-1 liter). Hereafter the infected *T. aestivum* plants were grown at 22°C and 75% humidity 16-h light period for another 26 days. To avoid cross-contamination each pot was covered with a small plastic bag.

28 days post infection, the leaves were inspected for disease symptoms. The symptoms were evaluated by screening the fungal caused necrosis and the amount of pycnidia. The obtained values were indexed using six categories representing the percentages of necrotic leaf area and the percentage of leaf area covered by pycnidia (0, 1-20, 21-40, 41-60, 61-80, 81-100).

5.5 Standard microbiological and biochemical methods

5.5.1 Nucleotide isolation and Southern blotting

5.5.1.1 Polymerase chain reaction (PCR)

5.5.1.1.1 Polymerase chain reaction with NEB polymerase

According to the instructions of the Taq DNA polymerase with thermoPol buffer (NEB M0267L).

10X Standard Taq Reaction Buffer	1x
10 mM dNTPs	200 µM each
Primer A	0.2 µM
Primer B	0.2 µM
Template DNA	>1000 ng
Taq DNA Polymerase	1.25 U / 50 µl

PCR program: 95°C/30 sec – [95°C/15-30 sec – 45-68°C/15-60 sec – 68°C/1 min /1 kb] x 35 – 68°C/5 min

5.5.1.1.2 Polymerase chain reaction (PCR) with phusion high-fidelity DNA polymerase

According to the instructions of the Phusion high-fidelity DNA polymerase (Thermo scientific F-530S).

5x Phusion HF Buffer	1x

10 mM dNTPs	1x 200 µM each
Primer A	0.5 µM
Primer B	0.5 µM
Template DNA	50-250 ng
DMSO	(3 %)
Phusion DNA Polymerase	0.02U/µl

PCR program: 98°C/30 sec – [98°C/5-10 sec - Tm + 3°C/10-30 sec - 72°C/15-30 sec /1 kb] x 35 – 72°C/5 min

5.5.1.1.3 Overlap PCR to ligate PCR products

Up to five fragments with an overlap of 20 nt were used to create a construct consisting of all fragments. To perform the PCR, the protocol for the Phusion polymerase was used with modifications: Fragments of 0.1 pmol were used as template. The first five PCR cycles were without primers and after the adding of the primers. These were added for the sixth cycle and the PCR was continued for an additionally 30 cycles. The obtained PCR product was gel-purified for T4 DNA ligation (5.5.1.2.1)

5.5.1.2 Ligation of DNA fragments

5.5.1.2.1 Ligation with T4 DNA ligase

According to the instructions of T4 DNA ligase (Roche, Mannheim) DNA fragments were ligated. In short, a linearized plasmid and up to five DNA fragments were incubated in a 1:5 molecular proportion in a total volume of 20 µl with 1U T4 DNA ligase at 16°C overnight. After incubation, the reaction was heat deactivated at 65°C for 15 minutes.

5.5.1.2.2 Gibson Assembly Cloning

According to instructions of Gibson Assembly Cloning protocol (NEB) constructs and vectors were produced.

In short, the following reaction was performed on ice:

	2–3 fragment Assembly	4–6 fragment Assembly
Total Amount of fragments	0.02–0.5 pmols x µl	0.2–1 pmols x µl
gibson Assembly Master Mix (2x)	10 µl	10 µl
Deionized H2o	10-x µl	10-x µl
Total Volume	20 µl	20 µl

The samples were incubated in a thermocycler at 50°C for 60 minutes. The circular products were used with the protocol for chemical transformation of *E. coli* (5.4.1.2 TopoTA cloning).

5.5.1.3 Plasmid preparation with *Escherichia coli*

5.5.1.3.1 Plasmid DNA isolation QIA Mini-Prep

Plasmid DNA isolation by QIA Mini-Prep was performed according to the protocol of the QIAprep Spin Miniprep Kit.

5.5.1.3.2 Plasmid DNA isolation by heat incubation (Sambrook et al, 1989)

A colony was inoculated in 2 ml dYT and a selection antibiotic was added. After overnight incubation, at 37°C, the culture was transferred to a 2 ml Eppendorf tube and after centrifuging for 2 min. at 13,000 rpm the supernatant was discard. The obtained pellet was dissolved in 200 µl STET buffer (0.1 M NaCl, 10 mM Tris/HCl (pH 8,0), 1 mM EDTA, 5% Triton-X-100) and Lysozyme (1,0 %). This was followed by incubation for 60 sec at 95°C and the suspension was centrifuged at 13,000rpm for 10min. After the centrifugation, the pellet was discard and to the supernatant 20 µl of 3M NaOAc and 500µl 2-Propanol was added and the suspension was mixed by inversion, followed by another centrifugation at 13,000rpm for 10min. After discarding the supernatant, the pellet was washed by adding of 500µl 70% Ethanol, incubated for a few minutes, and centrifuged for 5 min and the supernatant was discarded. The purified plasmid DNA was dissolved in 100 µl TE.

5.5.1.4 Genomic DNA isolation from *Zymoseptoria tritici*

Cells of a densely grown *Z. tritici* culture (OD600 ~1.0) were harvested by centrifugation at 13,000 rpm, for 2 min and the supernatant was discarded. To the pellet, one spoon of glass beads to the cell pellet, 500 µl Phenol/Chloroform (1:1) and 500 µl of *Zymoseptoria* lysis buffer (10% Triton X 100, 1% SDS, 100 mM NaCl, 10 mM TrisHCl pH 8.0, 10 mM EDTA pH 8.0) were added. To disrupt the cell-wall, the cell suspension was placed on a Vibrax for 30 min and followed by centrifugation at 13,000 rpm for 15 min at room temperature. The aqueous (upper) phase was transferred to a 1.5ml Eppendorf tube containing 1 ml 100% Ethanol and the suspension was mixed by inverting, followed by centrifugation at 13,000 rpm for 5 min at room temperature. The supernatant was discarded and the pellet was re-suspended in 30-50 µl TE.

5.5.1.5 Quick DNA extraction for pre-selection (Modified after Zwiers and De Waard 2001).

This protocol is only suitable for a pre-selection of *Z. tritici* transformants for Southern Blot

analysis.
Due to technical limitations of the protocol, only products smaller than 750 bp can be amplified by PCR.

A freshly grown *Z. tritici* colony was transferred to an Eppendorf tube and diluted in 50 µl of TE buffer. To disrupt the cell-wall, the suspension was microwaved for 2 minutes at 900 watt. This was followed by vortexing the suspension for 5 min and additional 2 minutes microwaving at 900 watt. To obtain a pellet, the suspension was centrifuged at 13,000 rpm for 2 minutes and the supernatant used directly for PCR reaction.

PCR- Mix for 1 reaction:

- 7.5 µl Quick DNA
- 3.0 µl 10 x HF buffer
- 0.45 µl DMSO
- 0.15 µl dNTPs (10mM)
- 0.15 µl Primer (100pmol/µl)
- 0.15 µl Primer (100pmol/µl)
- 0.15 µl Phusion Polymerase
- 3.45 µl H_2O

Total volume: 15 µl

PCR program:

98°C – 30 sec.		
98°C – 8 sec.	|	
xx°C – 20 sec.	|	35 cycles
72°C – x sec. (30 sec/1 kb)	|	
72°C – 5 min.		

5.5.1.6 Sequencing of plasmids

All sequence reactions were done by Eurofins Genomics, Ebersberg. Purified plasmids, with a concentration of 50-100 ng/µl in 15 µl, were submitted and sequenced, with Sanger sequencing, with primers (10 pmol/µl). Sequences were inspected manually using the CLC-Genomic workbench.

5.5.1.7 Southern Analyses

10-25 µg of DNA was digested overnight with an enzyme of choice (or a combination of enzymes) in a total volume of 25 µl. Samples were separated on a 0.8% TAE agarose gel and run at 80-100 V for 3-4 hours. After electrophoresis, the gel was documented and the ladder marked by scoring the bands with a knife. Prior to the blotting, the gels were incubated in 0.25 M HCl for 20 minutes while shaking. The HCl was replaced with 0.4 M NaOH and incubated for 20 minutes while shaking. The DNA was transferred by capillary transfer to a Nylon membrane (Hybond N+, GE

healthcare, Munich, Germany) in 0.4 M NaOH overnight. After blotting, the DNA was UV cross-linked to the membrane. Prior to the hybridization, the membrane was pre-hybridized at 65°C for at least 30 minutes in hybridization buffer (For 1 liter: 500 ml 1M NaPO4, pH7.0; 350 ml 20% SDS, fill-up with water)

Dig labeled probes were generated according to the protocol of the manufacture (PCR DIG labeling Mix protocol, Roche, Mannheim, Germany). The gel purified DIG-probe was added to 15 ml hybridization buffer and boiled in a water bath at 100°C for at least 10 minutes. After discarding the pre-hybridization buffer, the probe was added and the membrane was hybridized at 65°C (overnight). After hybridization, the membrane was washed twice for 15 minutes at 65°C with few ml of southern wash buffer (For 1 liter: 100 ml 1M NaPO4, pH7.0 (>0,1M);50 ml 20% SDS (>0,1M) fill-up with water) and subsequently washed in DIG wash buffer at 25°C for 5 minutes. After discarding the DIG wash buffer, the membrane was incubated in 20ml DIG buffer 2 (For 100ml: 1g skimmed milk powder; 1.16g Maleic acid; 3 ml 5M NaCl and fill-up with water) for 30 minutes at 25°C. Subsequently the membrane was transferred to 15ml antibody solution in a 1:10,000 dilution in DIG buffer 2 and incubated for 30 minutes at 25°C. Followed by washing the membrane twice for 15 minutes at 25°C with DIG wash buffer (997 ml DIG buffer 1 and 3 ml Tween-20 (0.3%)) and equilibration of the membrane for 5 minutes in DIG buffer 3 (For 1 liter: 100 ml 1M Tris-HCl (pH 9.5); 20 ml 5M NaCl; 50 ml 1M MgCl2 and fill-up with water). Finally the membrane was incubated for 5 min with 4ml CDP Star (1:100 dilution of CDP Star (Roche Cat no 11 685 627 001) for 15 minutes at 37°C and exposed to an X-ray film.

5.5.1.8 RNA isolation of a *Zymoseptoria tritici* culture

Axenic cultures of *Z. tritici* were densely grown (OD600 ~1.0) and harvested by centrifugation at 13,000 rpm, for 2 min and the supernatant was discarded. Infected plant material was isolated at a specific time-point (4, 11, 16 and 28 dpi). To disrupt the cell-wall the samples (e.g. supernatant or infected plant leaves) were grounded thoroughly in liquid nitrogen and the obtained material was transferred to a pre-chilled 1.5 ml Eppendorf tube. To the grounded material, 0.75 ml TRIzol LS reagent (Invitrogen) and 200 µl Chloroform was added. The suspension was mixed by inversion and centrifuged at 12,000 rpm for 15 min at 4°C. The water (upper) phase was transferred into a new Eppendorf tube and 500 µl Isopropyl alcohol was added. After incubating at room temperature for 10 min, the suspension was centrifuged at 12,000 rpm for 15 min at 4°C and the supernatant was discarded. The obtained pellet was washed with 1 ml 75% EtOH followed by centrifugation at 12,000 rpm for 5 min at 4°C and the supernatant was again discarded. The pellet was briefly dried and dissolved in RNAse free water (20-50µl). The obtained RNA was stored at -80°C.

5.5.1.9 cDNA synthesis

Prior to the cDNA synthesis, the samples were treated with DNAseI.

To 1 µg total RNA, 0.9 µl 10x DNAseI buffer and 1 µl DNAseI (DNAseI Ambion AM2222) were added. The total volume was finalized to 9 µl with RNAse free DEPC water. The reaction was incubated at 37°C for 30 min. To terminate the reaction, 1µl 50mM EDTA was added and the DNAseI was heat inactivated at 75°C for 15 min.

To the DNAseI treated samples, 1µl Oligo dT primer, 4µl 5x reaction buffer, 1µl RNAse Inhibitor, 2µl dNTPs mix and 2µl Reverse transcriptase were added. The reaction was incubated for 60 min at 37°C and subsequently the reaction was terminated for 5 min at 70°C. All synthesized cDNA was tested for the presence of genomic DNA by two PCR reactions with intron spanning primers:

- oES68 & oES69 for wheat actin (571bp for gDNA and 277bp for cDNA)
- oES1026 & oES1027 for Zt Glyceraldehyde 3-phosphate dehydrogenase (491bp for gDNA and 431bp for cDNA)

5.5.1.10 Quantitative PCR

To 4 µl gDNA free cDNA, 12 µl Sybr Green (BioRad, cat: 170-8882), 0.4 µl Forward primer, 0.4 µl Reverse primer and 7.2 µl water were added. To amplify products, a program with an annealing temperature of 59°C, elongation time of 10 seconds and 40 cycles was used. The qPCR amplifications were performed in a BioRad CFX-connect real-time system.

5.5.1.11 Pulsed Field Gel Electrophoresis (PFGE) of *Zymoseptoria sp*.

A densely grown culture was centrifuged at 4000 rpm for 10 min at 4°C. The supernatant was discarded. The obtained pellet was washed with 1.2 M Sorbitol, followed by centrifugation as above and supernatant was discarded. The pellet was resuspended in 5 ml Sorbitol 1.2M/Novozyme (6mg/ml) and incubated for 1 hour. The protoplast formation was verified by microscopic analysis. The protoplasts were resuspended in 25 ml Sorbitol 1.2M and centrifuged for 10 min at 4000 rpm and the supernatant was discarded. From here on all steps were performed on ice. 10 ml cold STC (30 ml Sorbitol 2M, 0.5 ml Tris-HCl 1M pH 7.5, 5 ml CaCl2 1M and filled to 50 ml with water) was added to the sample and centrifuged for 10 min at 4000 rpm and the supernatant was discarded. The washing with STC and centrifugation were repeated for three times. Subsequently, the pellet was resuspended in 1 ml cold STC and an equal volume of 2.2% low melting point agarose in TBE 0.5x was added to the protoplasts. To create the PFGE plugs, the suspension of protoplasts and agarose was mixed gently and directly filled into a well of a plug mold. Followed by solidification

for 2 hours at 4°C. Subsequently the plugs were removed from the plug molds and incubated in 10 ml lysis buffer (0.2 g Sarkosyl, 18 ml EDTA 0.5M, 20 mg Proteinase K and 2 ml water) for 48 hours at 50°C, while replacing lysis buffer once after 24 hours. Finally, the plugs were rinsed once in EDTA 0.5M and stored in EDTA 0.5M at 4°C.

The samples were loaded on a 1% low melting agarose (TBE). To separate the dispensable chromosomes the following conditions were applied: switching time: 60–120s, running time: 48 h, Volts/cm: 4.1, angle: 120° and 14°C.

5.5.2 Protein preparation

5.5.1 Histone isolation (Acid extraction method)

A *Z. tritici* culture was grown to an OD600 of 1.0 in liquid YMS media and a pellet was obtained by centrifugation. The pellet was quick frozen and grounded in liquid N_2. To the pellet 20 ml of ice cold Histone extraction buffer 1 (0.3 M Sucrose, 40 mM $NaHSO_3$, 25 mM Tris pH 7.4, 10 mM $MgSO_4$, 0.5 mM EDTA and 0.5% NP40 (Igepal) and protein inhibitors 1 mM PMSF, 1 µg/ml pepstatin and 1 µg/ml leupeptin) were added to the grounded mycelia was suspended by pipetting on ice. Subsequently the suspension was centrifuged at 8000 rpm for 10 minutes at 4°C. The supernatant was discarded and the pellet was washed with 20 ml Histone extraction buffer 1 and the suspension was centrifuged at 8000 rpm for 10 minutes at 4°C. The pellet was resuspended in 1 ml of CW buffer (1 mM beta-mercaptoethanol, 150 mM NaCl and 10 mM Tris pH 8.0) and including 1 mM PMSF, 1 µg/ml pepstatin and 1 µg/ml leupeptin) and an equal volume of 0.4 M $H2SO_4$ was added. After overnight incubation at 4°C, while shaking, the sample was split into four equal volumes, separated in four tubes. A pellet was obtained by centrifugation at 12.000 rpm for 10 minutes at 4°C. The supernatant was transferred to a new tube and mixed with 1/4 volume of 100% TCA and incubated on ice for 1 hour. To obtain a pellet, the suspension was centrifuged at 12.000 rpm for 30 minutes at 4°C. Subsequently the pellet was washed twice with ice-cold acetone at room temperature and centrifuged at 12.000 rpm for 5 minutes. After drying the pellet at room temperature, the pellet was resuspended in a total volume of 50 µl of 20 mM Tris pH 8.0 and after centrifugation at 12.000 rpm for 5 minutes the pellets of the four tubes were pooled.

5.5.2.2 Western blotting

A Bradford protein assay was used to measure the protein concentration for each sample. A protein concentration of 10 µg was used for each sample. For each gel, a ladder was included (Protein STP Small, Kaleidoscope, polypeptide standard Cat#161-0325). After denaturation of the samples at 100°C, for 5 minutes, the samples were loaded on a 15% acrylamide gel and separate for 90

minutes. After the electrophoresis, the gel was either stained in chromasie stain for 40 minutes while shaking and subsequently de-stained in de-stain fast for 45 minutes while shaking. For Western blotting, the samples were transferred from the gel to a PVDF membrane according to the instructions of the Mini Trans-blot Electrophoretic transfer cell (Bio-Rad). After transferring the samples to the membrane, the membrane was incubated face-down for 30 minutes at room temperature in 10 ml of TBS + 0.05% Tween + 3% milk, while shaking. After the blocking reaction, 1-2 µl of the primary antibody was added and the membrane was incubated overnight at 4°C. Hereafter the membrane was washed five times with water for 10 minutes and 1 µl of secondary antibody in 10 ml of TBS + 0.05% Tween and 0.5% milk was added and incubated for 30 minutes to one hour at room temperature, while shaking. Again the membrane was washed five times with water for 10 minutes. Finally the membrane was developed according to the instructions of chemiluminescent hyglow quickspray.

Table 5.1: Antibodies for Western blot

Antibody	Host	Cat no.	Supplier
GFP	Mouse	9996	Santa Cruz biotechnology
H3K4me3: Histone H3 trimethyl Lys 4	Rabbit	39159	Active motif
Anti-MgCenH3	Rabbit		
Secondary antibody			
Goat anti rabbit	Goat	31460	Hrp Pierce
Goat anti mouse	Goat		Invitrogen

5.5.2.3 Chromatin Immunoprecipation (ChIP)

To a densely grown *Z. tritici* culture of 40 ml, 2 ml of 20% Formaldehyde was added directly to the flask and incubated for 15 minutes at room temperature while shaking (100 rpm). The formaldehyde was quenched by adding 2 ml 2.5 M glycine and the pellet was obtained by centrifugation at 2000 rpm for 5 min. The pellet was washed with 1x PBS and the suspension was centrifuged for 1 min at 2000 rpm to obtain a pellet and the supernatant was discarded. The obtained pellet was ground in liquid Nitrogen. Ice cold chromatin buffer (50 mM Hepes-NaOH pH 7.5, 20 mM NaCl, 1 mM Na-EDTA pH 8.0, 1% Triton X-100 and 0.1% DOC) was added to the pellet in a ratio of 5 µl chromatin buffer to 1 mg pellet. Subsequently, protein inhibitors were added to the suspension: 0.001 M PMSF, 1x Leupeptin, 1x E-64 and 0.1x Pepstatin. Subsequently $CaCl_2$ to an end concentration of 2 mM was added and the suspension was mixed by inversion. Followed by adding of 5 µl of Micrococcal nuclease (Mnase; #M02479; Lot nr: 0091211; NEB), the reaction was incubated for 10 min in a 37°C water bath and mixed every 2 minutes. To stop the reaction, 4 µl of 0.5 M Na-EGTA pH 8.0 was added and the samples were placed on ice. The final NaCl

concentration was brought to 140 mM, mixed and the centrifuged 4.000 rpm for 5 min at 4 °C. The supernatant was transferred to a new tube and the sample was separated in two fractions, one fraction of 800 µl and two fractions of 100 µl. The 800 µl fraction was used for the actual ChIP experiment and the two fractions of 100 µl were used to verify the DNA quality. In short the verification of the DNA: 1 volume TES and 1 µl proteinase-K were added and the sample was de-crosslinked for 6-16 hours at 65°C (overnight). One volume of water and 1.9 µl of 20 mg/mL RNAseA were added and the sample was incubated for 2 hours at 50 °C. To precipitate the DNA, one volume of phenol/chloroform was added and the sample was centrifuged 5 min at 14.000 rpm. The aqueous layer was extracted and 250 µl chloroform was added and the sample was centrifuged 5 min at 14.000 rpm. To the obtained aqueous layer 1 µl of glycogen, 25 µl of Na-acetate pH5.2 and 865 µl of 100% EtOH were added and the sample was precipitated overnight at -20 °C. Followed by centrifugation for 10 min at 14.000 rpm at 4°C. The obtained pellet was washed with 1000 µl of 70% EtOH and after centrifugation at 14.000 rpm for 10 minutes at 4 °C a pellet was obtained. The pellet was dried and resuspended in 30 µl of TE, 1 µl was ran on a 2% TAE gel and the remaining volume was stored at -20°C.

The ChIP experiment used the 800 µl fraction and 20 µl of protein A beads were added and the lysate was pre-cleared by incubating at 4°C on a rotator for 1-3 hours. Hereafter the sample was centrifuged at 5000 rpm for 1 min and 250 µl of the pre-cleared supernatant was transferred into a new tube. The desired antibody was added to the sample and incubated overnight at 4 °C on a rotator. Hereafter 20 µl of the equilibrated protein-A agarose beads are added and incubated for 1-2 hours at 4°C on a rotator to allow antibody binding. This was followed by centrifugation for 1 min at 5000 rpm at 4°C to pelletize the beads. The supernatant was then discarded. To the pellet, 1 ml cold ChIP lysis buffer (50 mM HEPES pH 7.4, 140 mM NaCl, 1 mM EDTA,1 % Triton-X-100 and 0.01% DOC (Na-Deoxycholate) was added and incubated for 5 min at 4°C on while shaking. Hereafter the sample was centrifuged for 1 min at 5000 rpm at 4 °C and the supernatant was then discarded. This was followed by adding again 1 ml cold ChIP lysis buffer, incubation and centrifugation. Subsequently the pellet was washed with cold ChIP lysis buffer + 0.5 M NaCl, incubated and centrifuged as described with the first wash step with the ChIP lysis buffer. This was followed by washing of the pellet with cold LiCl wash buffer as described with the first wash step with the ChIP lysis buffer. Finally the pellet was washed as described with the first wash step before with cold TE. By adding of 62.5 µl of (pre-warmed) TES and incubating for 10 min at 65 °C the DNA was eluted from the agarose beads. During the incubation, the sample was vortexed several times. To obtain a supernatant the sample was centrifuged for 1 min at 5000 rpm and the supernatant was saved. To the sample again 62.5 µl of (pre-warmed) TES was added and incubated

and centrifuged as before. Both obtained supernatants were pooled.

The protein-DNA interactions were de-crosslinked by incubating the samples for 6-16 hours in a 65°C incubator. Followed by adding 250 µl of water (1:1) and 1.9 µl of 20 mg/ml RNAseA to the sample and the sample was incubated at 50°C for 2 hours. Subsequently, 9.4 µl of 20 mg/ml Proteinase K was added and the sample was incubated for 2 hours at 50°C.

To precipitate the DNA, 250 µl of phenol/chloroform (1:1) was added and the aqueous layer was extracted after centrifugation for 5 min at 14.000 rpm. This was followed by adding 250 µl of chloroform and the aqueous layer was extracted after centrifugation for 5 min at 14.000 rpm. 1 µl of glycogen, 25 µl of Na-acetate pH5.2 and 865 µl of 100% EtOH were added to the sample and the DNA was precipitated overnight at -20°C.

The sample was centrifuged for 10 min at 14.000 rpm at 4 °C, the obtained pellet was washed with 1000 µl of 70% EtOH and again centrifuged for 10 min at 14.000 rpm at 4 °C. The supernatant was discarded and the pellet dried and resuspended in 30 µl of 1/10 TE. The sample was stored at -20°C or used for library construction.

Before the library construction with TruSeq adapters, the AMPure XP beads were pre-heated on the bench for +/- 30 min. All the reactions displayed below are performed in PCR tubes.

Blunt DNA ends (end repair) were generated by adding of 5 µl of NEB T4 DNA ligase buffer, 2 µl of 10mM (each) dNTPS, 0.5 µl of End Repair Enzyme mix to 25 µl of sample, and fill-up to 50 µl with water (17.5 µl). (End Repair enzyme mix: 30 µl of 3 U/µl T4 DNA polymerase (NEB), 6 µl of 5 U/µl Klenow fragment (NEB) and 30 µl of 10 U/µl of T4 DNA Polynucleotide Kinase (NEB)). Subsequently the reaction was incubated at 20°C for 30 minutes. To purify the DNA, 50 µl of AMPure XP beads and 50 µl of 30% PEG8000, 1.25 M NaCl were added to the sample. Followed by incubation at room temp for 15 min. After placing the sample on a magnetic rack for 5 min the supernatant was removed and discarded. The samples stayed in the magnetic rack and 200 µl of freshly prepared 80% Ethanol was added. After an incubation for 30 seconds, the supernatant was removed and discarded. The washing with 80% Ethanol was repeated once and subsequently the AMPure XP beads were dried for 5 minutes. 18 µl of 1/10 TE was added and incubated at room temp for 2 min. The samples were placed in a magnetic rack for 5 min and 16.5 µl of the supernatant was transferred to a new PCR tube.

To the end-repaired DNA, 2 µl of 10x NEB buffer #2, 1 µl of 4 mM dATP and 0.5 µl of 5 U/µl Klenow 3’ to 5’ exo minus (NEB) were added and incubated at 37°C for 30 minutes to ligate an A-overhang to the DNA. To 20 µl of the ‘A’-tailed DNA, 25 µl of 2x Quick ligase buffer, 2.5 µl RNA TruSeq adapters (or 2.5 µl DNA TruSeq adapters diluted 1:100), 2.5 µl of ligase (400 U/µl NEB) were added. (2x Quick ligase buffer: 132 mM Tris-HCl pH 7.6, 20 mM $MgCl_2$, 2 mM DTT, 2 mM

ATP and 15% PEG6000). TrueSeq PCR primer cocktail: 25 µl of OMF3028, 25 µl of OMF3029 and 50 µl of water, OMF3028: AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGA. OMF3029: CAAGCAGAAGACGGCATACGAGAT). The reaction was incubated for 20 minutes at room temperature and subsequently 5 µl of 0.5 M EDTA, pH 8.0, 55 µl of AMPure XP beads and 55 µl of 30% PEG8000, 1.25 M NaCl were added and the sample was incubated at room temperature for 20 minutes. To purify the sample, the tubes were placed on a magnetic rack for 5 minutes and the supernatant was discarded. To the sample, 200 µl of freshly prepared 80% ethanol was added and the sample was incubated for 30 seconds. The washing with 80% ethanol was repeated once and subsequently the AMPure XP beads were dried for 5 minutes. The DNA was resolved in 15.5 µl of 1/10 TE and the sample was incubated at room temperature for 2 minutes and placed in a magnetic rack for 5 minutes. 14 µl of the supernatant was transferred to a new PCR tube.

To size select the library, the samples were separated on a 2.5% NuSieve agarose gel. The gel that contains material between 250 and 450 bp was sliced out and place in a 2 mL tube. Three volumes of Qiagen Buffer QG were added to the gel and subsequently the samples were incubated at room temperature until the gel slice was completely dissolved. 1 volume of isopropanol was added, followed by transferring of the sample to a Qiagen minElute column. After centrifugation for 30 seconds at 14.000 rpm, the liquid from the collection tube was discarded. 500 µl of Qiagen Buffer QG was added to the column, centrifuged for 1 min at 14.000 rpm and the liquid from the collection tube was discarded. 750 µl of Qiagen buffer PE was applied to the column, incubated for 2 minutes and centrifuged for 1 min at 14.000 rpm. The liquid from the collection tube was discarded and the sample was centrifuged for one additional minute at 14.000 rpm. The column was transferred to a new 1.5 mL tube and 20 µl of Qiagen buffer EB was added and after 5 minutes incubation, the sample was centrifuged for 1 min at 14.000 rpm. An additional 20 µl of Qiagen buffer EB was added and again the incubation and centrifugation steps were repeated.

To determine the minimum number of cycles needed for final amplification a test PCR was performed.

To 4.75 µl size selected DNA, 0.5 µl of TrueSeq PCR primer cocktail, 4.75 µl of water and 10 µl of 2x Phusion Flash master mix were added. The amplification was performed with the following program: 45 sec at 98°C, 13 cycles of (15 sec at 98°C, 30 sec at 63°C and 30 sec at 72°C), 1 min at 72°C.

After the test PCR, the libraries were amplified by PCR. To 19 µl size selected DNA, 1 µl of TrueSeq PCR primer cocktail and 20 µl of 2x Phusion Flash master mix were added. The amplification was performed with the following program: 45 sec at 98°C, 13 cycles of (15 sec at

98°C, 30 sec at 63°C and 30 sec at 72°C), 1 min at 72°C. To purify the samples, 40 µl of AMPure XP beads were added directly to the PCR tube and incubated at room temperature for 15 min. Subsequently, the PCR tubes were placed on a magnetic rack for 5 min and the supernatant was discarded. The samples were washed with 200 µl of freshly prepared 80% ethanol, incubated for 30 seconds and the supernatant was discarded. The washing step with 80% ethanol was repeated once. Subsequently, the AMPure XP beads were dried for 5 minutes and 12 µl of 1/10 TE was added to the sample. After incubating at room temperature for 2 min, the PCR tubes were placed in a magnetic rack for 5 minutes and 10 µl of the supernatant was transferred to a new 1.5 mL tube.
Finally, the concentration was determined by Nanodrop and 2 µl of the sample was ran on a 2% gel to verify the quality of the library.

Table 5.2: Antibodies used for ChIP experiments.

Antibodies for ChIP:	**Host/Clonality**	**Cat no.**	**Supplier**
Protein A Agarose beads		15918-014	Invitrogen
Rb pAB to GFP	Rabbit/Polyclonal	ab290	Abcam
H3K4me2	Rabbit/Polyclonal	07-030	Millipore
H3K9me3	Rabbit/Polyclonal	39161	Active Motif
H3K27me3	Rabbit/Polyclonal	39155	Active Motif

5.6 Computational methods

5.6.1 Plasmid design

All plasmids were design with the bioinformatic software: Clone manager of Scientific & Educational Software, USA (www.scied.com).

5.6.2 qPCR analysis

All qPCR reactions were analyzed with the Rest2009 software (Pfaffl et al., 2002). The reads were normalized to gene expression in axenic cultures and standard settings were applied.

5.6.3 ChIP-seq analysis and mapping of the reads

All grooming, trimming, filtering and masking steps where performed with tools from the Galaxy project (www.galaxyproject.org).
Grooming was performed according to the standard settings of the Galaxy project. After grooming, the reads were trimmed 4 bp from the left. Subsequently, reads with an overall quality score bellow 20 were discarded. In addition, reads which had less than 25% of nucleotides with a quality score

above 20 were also discard. Masking of nucleotides with a quality score bellow 20 was applied to all nucleotides which had a quality score less than 20.

Subsequently, the reads were mapped to the specific genomes with Bowtie2 and standard settings were applied (Langmead, 2010). After the Bowtie2 mapping the reads were converted in to Bam-files with Samtools (Li et al., 2009). The Bam-files were visualized with the CLC-Genomic Workbench (www.clcbio.com).

5.6.4 Alignments and Phylogenetic analyses

All alignments were created in the software packet Seaview with the Muscle alignment software (Gouy et al., 2010) & (Edgar, 2004). The phylogenetic trees were conducted using a maximum likelihood as implemented in PhyML using a GTR substitution model (Guindon et al., 2010).

5.6.5 Nucleotide diversity

The nucleotide diversity was calculated with the software: DNAsP (Librado and Rozas, 2009). The nucleotide diversity was shown with π. Π is the measurement of nucleotide diversity which accounts for the sample site.

Chapter 6: Supplementary data

Table 6.1: Repetitive sequences in the centromeric regions

Centromere	Start (bp)	End (bp)	Length	Strand	Motif or Repeat
Cen01	3165	3232	67	+	Target "Motif:(TTCAC)n" 1 69
Cen01	7827	7877	50	+	Target "Motif:(TA)n" 1 53
Cen01	8331	8793	462	-	Target "Motif:RLC_element4_ZPST04IR55" 4908 5397
Cen01	8800	8830	30	-	Target "Motif:RLC_element6_ZB163" 383 413
Cen01	8826	8879	53	+	Target "Motif:RLC_element6_ZB163" 473 530
Cen01	9391	12000	2609	-	Target "Motif:RYN_element1_ZTIPO323" 2036 4652
Cen02	1928	4511	2583	+	Target "Motif:NoCat_element1_ZTIPO323" 336 2929
Cen02	2442	4759	2317	+	Target "Motif:NoCat_element1_ZTIPO323" 468 2787
Cen02	6217	6306	89	-	Target "Motif:DTA_element7_ZTIPO323" 1 90
Cen04	8114	8382	268	+	Target "Motif:RLX-TRIM_element4_ZTIPO323" 400 668
Cen04	8162	8829	667	+	Target "Motif:RLX-TRIM_element4_ZTIPO323" 1 671
Cen04	8301	8980	679	+	Target "Motif:RLX-TRIM_element4_ZTIPO323" 1 674
Cen04	8893	8999	106	+	Target "Motif:RLX-TRIM_element4_ZTIPO323" 1 107
Cen05	1	1832	1831	+	Target "Motif:RYN_element1_ZTIPO323" 1165 2996
Cen05	5341	5405	64	+	Target "Motif:(ATATCG)n" 1 72
Cen05	5386	5407	21	+	Target "Motif:(AT)n" 1 22
Cen05	6611	6645	34	+	Target "Motif:(TA)n" 1 36
Cen05	8874	8907	33	+	Target "Motif:(TTCAAGC)n" 1 36
Cen06	1158	1222	64	+	Target "Motif:(TA)n" 1 69
Cen06	1888	6539	4651	+	Target "Motif:RYN_element1_ZTIPO323" 1 4652
Cen06	8661	8701	40	+	Target "Motif:(TCC)n" 1 41
Cen06	9011	9029	18	+	Target "Motif:(ACT)n" 1 19
Cen06	10129	10161	32	+	Target "Motif:(GCT)n" 1 33
Cen06	10253	10286	33	+	Target "Motif:(ACCTTA)n" 1 34
Cen06	10420	10455	35	+	Target "Motif:(GTTGCTA)n" 1 38
Cen06	10541	10560	19	+	Target "Motif:(CTC)n" 1 20
Cen06	10584	10671	87	+	Target "Motif:(TAAGT)n" 1 91
Cen06	11628	11653	25	+	Target "Motif:(ACTTT)n" 1 26
Cen06	11803	11835	32	+	Target "Motif:(AGG)n" 1 33
Cen07	6583	6622	39	-	Target "Motif:RLC_element4_ZPST04IR55" 5341 5384
Cen07	9175	9228	53	+	Target "Motif:(CATCATC)n" 1 55
Cen07	9175	9227	52	+	Target "Motif:(CATC)n" 1 57
Cen07	9759	9854	95	+	Target "Motif:(TTGGTG)n" 1 98
Cen07	9787	9829	42	+	Target "Motif:(GT)n" 1 43
Cen08	1	2491	2490	-	Target "Motif:RYN_element1_ZTIPO323" 1 2491
Cen08	2491	3547	1056	-	Target "Motif:NoCat_element1_ZTIPO323" 167 1224
Cen08	3548	3884	336	+	Target "Motif:NoCat_element10_ZTIPO323" 5 370
Cen08	3863	4099	236	+	Target "Motif:NoCat_element10_ZTIPO323" 405 715
Cen08	4100	4128	28	+	Target "Motif:(CTT)n" 1 29
Cen08	4129	4338	209	+	Target "Motif:NoCat_element10_ZTIPO323" 716 946
Cen08	4339	4518	179	-	Target "Motif:NoCat_element1_ZTIPO323" 1 180
Cen08	5433	5712	279	+	Target "Motif:NoCat_element1_ZTIPO323" 1721 2014
Cen08	7221	7729	508	-	Target "Motif:RLC_element4_ZPST04IR55" 4889 5397
Cen08	11183	11216	33	+	Target "Motif:(ACCTCC)n" 1 34
Cen09	817	859	42	+	Target "Motif:(ACGGCT)n" 1 44
Cen10	1	1926	1925	-	Target "Motif:DTX_element6_ZTIPO323" 1 1990
Cen10	6650	6919	269	-	Target "Motif:RLG_element15_ZPST04IR55" 5959 6233
Cen11	1045	2420	1375	-	Target "Motif:NoCat_element1_ZTIPO323" 1524 2849
Cen11	1100	1890	790	-	Target "Motif:NoCat_element1_ZTIPO323" 1357 2143
Cen11	1907	4766	2859	+	Target "Motif:NoCat_element1_ZTIPO323" 277 3138
Cen12	49	317	268	+	Target "Motif:DTH_element9_ZB163" 3306 3574

Cen12	3020	3087	67	+	Target "Motif:(TCC)n" 1 68
Cen12	3974	4041	67	-	Target "Motif:RLC_element4_ZPST04IR55" 5295 5364
Cen12	4122	4621	499	+	Target "Motif:RLC_element4_ZPST04IR55" 4896 5398
Cen12	5156	5189	33	+	Target "Motif:(GCTA)n" 1 34
Cen12	5730	5762	32	+	Target "Motif:(TCGTGC)n" 1 33
Cen12	6860	6908	48	+	Target "Motif:(TTGTGC)n" 1 49
Cen12	7087	7105	18	+	Target "Motif:(GGC)n" 1 19
Cen13	1060	1088	28	+	Target "Motif:(CCTCAT)n" 1 31
Cen14	5776	6254	478	-	Target "Motif:NoCat_element87_ZB163" 1541 2045
Cen14	6827	7002	175	-	Target "Motif:DTM_element1_ZTIPO323" 222 404
Cen14	7194	9761	2567	+	Target "Motif:RLC_element5_ZB163" 2190 4814
Cen14	7833	9761	1928	+	Target "Motif:RLC_element9_ZTIPO323" 1 1975
Cen14	9870	10000	130	-	Target "Motif:DTX-MITE_element5_ZTIPO323" 443 579
Cen15	185	2393	2208	+	Target "Motif:DTT_element1_ZTIPO323" 1 2229
Cen15	2856	2899	43	+	Target "Motif:(TCTTCG)n" 1 45
Cen15	4944	5192	248	-	Target "Motif:RLC_element6_ZB163" 1803 2066
Cen15	9319	9381	62	+	Target "Motif:(TCC)n" 1 63
Cen16	778	3915	3137	+	Target "Motif:NoCat_element1_ZTIPO323" 1 3138
Cen16	4494	4543	49	+	Target "Motif:RLX-TRIM_element4_ZTIPO323" 617 669
Cen16	4500	5124	624	+	Target "Motif:RLX-TRIM_element4_ZTIPO323" 40 670
Cen16	4893	5153	260	+	Target "Motif:RLX-TRIM_element4_ZTIPO323" 1 255
Cen16	5041	5172	131	+	Target "Motif:RLX-TRIM_element4_ZTIPO323" 1 131
Cen17	4071	4100	29	+	Target "Motif:(GGATA)n" 1 30
Cen17	5396	7914	2518	+	Target "Motif:RLC_element6_ZB163" 23 2629
Cen17	7917	10963	3046	-	Target "Motif:NoCat_element1_ZTIPO323" 1 3047
Cen18	4935	10995	6060	-	Target "Motif:RLC_element6_ZB163" 15 6109
Cen19	7387	7516	129	+	Target "Motif:RLC_element4_ZAST11IR611" 2644 2773
Cen19	7468	7527	59	+	Target "Motif:RLC_element4_ZAST11IR611" 2837 2896
Cen19	7600	7620	20	+	Target "Motif:(TGC)n" 1 21
Cen19	10499	10533	34	+	Target "Motif:(GAG)n" 1 35
Cen19	10953	10984	31	+	Target "Motif:(ACGAG)n" 1 32
Cen20	14	45	31	+	Target "Motif:(TTCGTC)n" 1 32
Cen20	268	316	48	+	Target "Motif:DTX_element5_ZTIPO323" 5403 5455
Cen20	4355	4546	191	+	Target "Motif:DTA_element2_ZTIPO323" 3644 3823
Cen20	4624	4674	50	+	Target "Motif:NoCat_element60_ZAST11IR611" 1835 1886
Cen20	4844	4978	134	+	Target "Motif:NoCat_element60_ZAST11IR611" 2077 2217
Cen20	6024	6079	55	+	Target "Motif:DTM_element1_ZTIPO323" 2233 2288
Cen20	6080	6318	238	+	Target "Motif:DTM_element6_ZTIPO323" 1086 1325
Cen20	6348	6401	53	+	Target "Motif:DTM_element6_ZTIPO323" 1419 1471
Cen20	6348	6393	45	+	Target "Motif:DTM_element1_ZTIPO323" 2731 2776
Cen20	6393	6439	46	+	Target "Motif:DTM_element6_ZTIPO323" 1505 1551
Cen20	6426	6651	225	+	Target "Motif:DTM_element1_ZTIPO323" 2911 3144
Cen20	7985	8109	124	+	Target "Motif:DTM_element1_ZTIPO323" 3424 3551
Cen20	7985	8066	81	+	Target "Motif:NoCat_element32_ZB163" 2262 2341
Cen21	5913	5962	49	-	Target "Motif:NoCat_element32_ZB163" 2311 2360
Cen21	6037	6100	63	-	Target "Motif:DTM_element1_ZTIPO323" 3220 3284

References

References

Adams, M.D. (2000). The Genome Sequence of Drosophila melanogaster. Science. *287*, 2185–2195.

Ahn, J., and Walton, J.D. (1996). Chromssomal Organization of TOx2, a Complex Locus Controlling Host-Selective Toxin Biosynthesis in Cochliobolus carbonum. Plant Cell *8*, 887–897.

Allshire, R.C., and Karpen, G.H. (2008). Epigenetic regulation of centromeric chromatin: old dogs, new tricks? Nat. Rev. Genet. *9*, 923–937.

Balesdent, M.-H., Fudal, I., Ollivier, B., Bally, P., Grandaubert, J., Eber, F., Chèvre, A.-M., Leflon, M., and Rouxel, T. (2013). The dispensable chromosome of Leptosphaeria maculans shelters an effector gene conferring avirulence towards Brassica rapa. New Phytol. *198*, 887–898.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823–837.

Bártová, E., Krejcí, J., Harnicarová, A., Galiová, G., and Kozubek, S. (2008). Histone modifications and nuclear architecture: a review. J. Histochem. Cytochem. *56*, 711–721.

Baum, M., Sanyal, K., Mishra, P.K., Thaler, N., and Carbon, J. (2006). Formation of functional centromeric chromatin is specified epigenetically in Candida albicans. Proc. Natl. Acad. Sci. U. S. A. *103*, 14877–14882.

Bensasson, D., Zarowiecki, M., Burt, A., and Koufopanou, V. (2008). Rapid evolution of yeast centromeres in the absence of drive. Genetics *178*, 2161–2167.

Bernhardt, S.A., Simmons, M.P., Olson, K.E., Beaty, B.J., Blair, C.D., and Black, W.C. (2012). Rapid intraspecific evolution of miRNA and siRNA genes in the mosquito Aedes aegypti. PLoS One *7*, e44198.

Bernstein, D.A., Vyas, V.K., Weinberg, D.E., Drinnenberg, I.A., Bartel, D.P., and Fink, G.R. (2011). Candida albicans Dicer (CaDcr1) is required for efficient ribosomal and spliceosomal RNA maturation. Proc. Natl. Acad. Sci. U. S. A. *109*, 523–528.

Billmyre, R.B., Calo, S., Feretzaki, M., Wang, X., and Heitman, J. (2013). RNAi function, diversity, and loss in the fungal kingdom. Chromosom. Res. *21*, 561–572.

Bowler, J., Scott, E., Tailor, R., Scalliet, G., Ray, J., and Csukai, M. (2010). New capabilities for Mycosphaerella graminicola Research. Mol. Plant Pathol. *11*, 691–704.

Braun, L., Cannella, D., Ortet, P., Barakat, M., Sautel, C.F., Kieffer, S., Garin, J., Bastien, O., Voinnet, O., and Hakimi, M.-A. (2010). A Complex Small RNA Repertoire Is Generated by a Plant/Fungal-Like Machinery and Effected by a Metazoan-Like Argonaute in the Single-Cell

Human Parasite Toxoplasma gondii. PLoS Pathog. *6*, e1000920.

Van der Burgt, A., Karimi Jashni, M., Bahkali, A.H., and de Wit, P.J.G.M. (2014). Pseudogenization in pathogenic fungi with different host plants and lifestyles might reflect their evolutionary past. Mol. Plant Pathol. *15*, 133–144.

Burrack, L.S., and Berman, J. (2012). Neocentromeres and epigenetically inherited features of centromeres. Chromosom. Res. *20*, 607–619.

Campos, E.I., and Reinberg, D. (2009). Histones: annotating chromatin. Annu. Rev. Genet. *43*, 559–599.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642–655.

Castillo, A.G., Mellone, B.G., Partridge, J.F., Richardson, W., Hamilton, G.L., Allshire, R.C., and Pidoux, A.L. (2007). Plasticity of fission yeast CENP-A chromatin driven by relative levels of histone H3 and H4. PLoS Genet. *3*, e121.

Catalanotto, C., Pallotta, M., Refalo, P., Sachs, M.S., Vayssie, L., Macino, G., and Cogoni, C. (2004). Redundancy of the Two Dicer Genes in Transgene-Induced Posttranscriptional Gene Silencing in Neurospora crassa. Mol. Cell. Biol. *24*, 2536–2545.

Chang, S.-S., Zhang, Z., and Liu, Y. (2012). RNA Interference Pathways in Fungi: Mechanisms and Functions. Annu. Rev. Microbiol. 305–323.

Clarke, L., and Carbon, J. (1980). Isolation of a yeast centromere and construction of functional small circular chromosomes. Nature *287*, 504–509.

Clarke, L., Amstutz, H., Fishel, B., and Carbon, J. (1986). Analysis of centromeric DNA in the fission yeast Schizosaccharomyces pombe. Proc. Natl. Acad. Sci. U. S. A. *83*, 8253–8257.

Coleman, J.J., Rounsley, S.D., Rodriguez-Carres, M., Kuo, A., Wasmann, C.C., Grimwood, J., Schmutz, J., Taga, M., White, G.J., Zhou, S., et al. (2009). The genome of Nectria haematococca: contribution of supernumerary chromosomes to gene expansion. PLoS Genet. *5*, e1000618.

Collas, P. (2010). The current state of chromatin immunoprecipitation. Mol. Biotechnol. *45*, 87–100.

Connolly, L.R., Smith, K.M., and Freitag, M. (2013). The Fusarium graminearum Histone H3 K27 Methyltransferase KMT6 Regulates Development and Expression of Secondary Metabolite Gene Clusters. PLoS Genet. *9*, e1003916.

Covert, S.F. (1998). Supernumerary chromosomes in filamentous fungi. Curr. Genet. *33*, 311–319.

Croll, D., and McDonald, B.A. (2012). The accessory genome as a cradle for adaptive evolution in pathogens. PLoS Pathog. *8*, e1002608.

Croll, D., Zala, M., and McDonald, B.A. (2013). Breakage-fusion-bridge Cycles and Large Insertions Contribute to the Rapid Evolution of Accessory Chromosomes in a Fungal Pathogen. PLoS Genet. *9*, e1003567.

Cumberledge, S., and Carbon, J. (1987). Saccharomyces cerevisiae. Genetics *117*, 203–212.

Dang, Y., Li, L., Guo, W., Xue, Z., and Liu, Y. (2013). Convergent Transcription Induces Dynamic DNA Methylation at disiRNA Loci. PLoS Genet. *9*, e1003761.

Dang, Y., Zhang, Z., and Liu, Y. (2014). Small RNA-Mediated Gene Silencing in Neurospora. In Fungal RNA Biology, A. Sesma, and T. von der Haar, eds. (Cham: Springer International Publishing), pp. 269–289.

Dhar, M.K., Friebe, B., Koul, A.K., and Gill, B.S. (2002). Origin of an apparent B chromosome by mutation, chromosome fragmentation and specific DNA sequence amplification. Chromosoma *111*, 332–340.

Dinant, C., and Luijsterburg, M.S. (2009). The emerging role of HP1 in the DNA damage response. Mol. Cell. Biol. *29*, 6335–6340.

Drinnenberg, I.A., Weinberg, D.E., Xie, K.T., Mower, J.P., Wolfe, K.H., Fink, G.R., and Bartel, D.P. (2009). RNAi in budding yeast. Science. *326*, 544–550.

Drinnenberg, I.A., Fink, G.R., and Bartel, D.P. (2011). Compatibility with killer explains the rise of RNAi-deficient fungi. Science. *333*, 1592.

Duncan, K.E., and Howard, R.J. (2000). Cytological analysis of wheat infection by the leaf blotch pathogen Mycosphaerella graminicola. Mycol. Res. *104*, 1074–1082.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.

Ellendorff, U., Fradin, E.F., de Jonge, R., and Thomma, B.P.H.J. (2009). RNA silencing is required for Arabidopsis defence against Verticillium wilt disease. J. Exp. Bot. *60*, 591–602.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806–811.

Folco, H.D., Pidoux, A.L., Urano, T., and Allshire, R.C. (2008). Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. Science. *319*, 94–97.

Freitag, M. (2014). Fungal Chromatin and Its Role in Regulation of Gene Expression. In Fungal Genomics, M. Nowrousian, ed. (Springer-Verlag Berlin Heidelberg),.

Galagan, J.E., and Selker, E.U. (2004). RIP: the evolutionary cost of genome defense. Trends Genet. *20*, 417–423.

Galagan, J.E., Calvo, S.E., Borkovich, K. a, Selker, E.U., Read, N.D., Jaffe, D., FitzHugh, W., Ma, L.-J., Smirnov, S., Purcell, S., et al. (2003). The genome sequence of the filamentous fungus Neurospora crassa. Nature *422*, 859–868.

Gaszner, M., and Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. Nat. Rev. Genet. *7*, 703–713.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. Nat. Rev. Genet. *10*, 94–108.

Goodwin, S.B., Waalwijk, C., and Kema, G.H.J. (2004). Genetics and genomics of Mycosphaerella graminicola A model for the dothideales. In Applied Mycology and Biotechnology, pp. 315–330.

Goodwin, S.B., Barek, S.B.M., Dhillon, B., Wittenberg, A.H.J., Crane, C.F., Hane, J.K., Foster, A.J., van der Lee, T.A.J., Grimwood, J., Aerts, A., et al. (2011). Finished Genome of the Fungal Wheat Pathogen Mycosphaerella graminicola Reveals Dispensome Structure, Chromosome Plasticity, and Stealth Pathogenesis. PLoS Genet. *7*, e1002070.

Gouy, M., Guindon, S., and Gascuel, O. (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. *27*, 221–224.

Grewal, S.I.S., and Jia, S. (2007). Heterochromatin revisited. Nat. Rev. Genet. *8*, 35–46.

Gu, W., Claycomb, J.M., Batista, P.J., Mello, C.C., and Conte, D. (2011). Cloning Argonaute-Associated Small RNAs from Caenorhabditis elegans. In Argonaute Proteins: Methods and Protocols, T.C. Hobman, and T.F. Duchaine, eds. (Totowa, NJ: Humana Press), pp. 251–280.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. *59*, 307–321.

Gullerova, M., and Proudfoot, N.J. (2008). Cohesin complex promotes transcriptional termination between convergent genes in S. pombe. Cell *132*, 983–995.

Gupta, O.P., Permar, V., Koundal, V., Singh, U.D., and Praveen, S. (2012). MicroRNA regulated defense responses in Triticum aestivum L. during Puccinia graminis f.sp. tritici infection. Mol. Biol. Rep. *39*, 817–824.

Halic, M., and Moazed, D. (2010). Dicer-independent primal RNAs trigger RNAi and heterochromatin formation. Cell *140*, 504–516.

Hammond, T.M., Xiao, H., Boone, E.C., Decker, L.M., Lee, S.A., Perdue, T.D., Pukkila, P.J., and Shiu, P.K.T. (2013). Novel proteins required for meiotic silencing by unpaired DNA and siRNA generation in Neurospora crassa. Genetics *194*, 91–100.

Han, F., Lamb, J.C., and Birchler, J.A. (2006). High frequency of centromere inactivation resulting in stable dicentric chromosomes of maize. Proc. Natl. Acad. Sci. U. S. A. *103*, 3238–3243.

Harrington, J.J., Van Bokkelen, G., Mays, R.W., Gustashaw, K., and Willard, H.F. (1997). Formation of de novo centromeres and constructution of first-generation human artificial

micochromosomes. Nat. Genet. *15*.

Henikoff, S., Ahmad, K., and Malik, H.S. (2001). The centromere paradox: stable inheritance with rapidly evolving DNA. Science. *293*, 1098–1102.

Hieter, P., Pridmore, D., Hegemann, J.H., Thomas, M., Davis, R.W., and Philippsent, P. (1985). Functional selection and analysis of Yeast Centromeric DNA. Cell *42*, 913–921.

Hillier, L.W., Coulson, A., Murray, J.I., Bao, Z., Sulston, J.E., and Waterston, R.H. (2005). Genomics in C. elegans: so many genes, such a little worm. Genome Res. *15*, 1651–1660.

Honda, S., Lewis, Z.A., Huarte, M., Cho, L.Y., David, L.L., Shi, Y., and Selker, E.U. (2010). The DMM complex prevents spreading of DNA methylation from transposons to nearby genes in Neurospora crassa. Genes Dev. *24*, 443–454.

Houben, A., and Carchilan, M. (2012). Plant B Chromosomes: What Makes Them Different? In Plant Cytogenetics, Plant Genetics and Genomics: Crops and Models 4, H.W. Bass, and J.A. Birchler, eds. (New York, NY: Springer New York), pp. 59–77.

Hu, Y., Stenlid, J., Elfstrand, M., and Olson, A. (2013). Evolution of RNA interference proteins dicer and argonaute in Basidiomycota. Mycologia *105*, 1489–1498.

Ishii, K., Ogiyama, Y., Chikashige, Y., Soejima, S., Masuda, F., Kakuma, T., Hiraoka, Y., and Takahashi, K. (2008). Heterochromatin integrity affects chromosome reorganization after centromere dysfunction. Science. *321*, 1088–1091.

Jamieson, K., Rountree, M.R., Lewis, Z.A., Stajich, J.E., and Selker, E.U. (2013). Regional control of histone H3 lysine 27 methylation in Neurospora. Proc. Natl. Acad. Sci. U. S. A. *55*, 1–6.

Jiang, N., Yang, Y., Janbon, G., Pan, J., and Zhu, X. (2012). Identification and functional demonstration of miRNAs in the fungus Cryptococcus neoformans. PLoS One *7*, e52734.

Jones, K. (1998). Robertsonian Fusion and Centric Fission in Karyotype Evolution of Higher Plants. Bot. Rev. *64*, 273–289.

Kadotani, N., Nakayashiki, H., Tosa, Y., and Mayama, S. (2004). One of the two Dicer-like proteins in the filamentous fungi Magnaporthe oryzae genome is responsible for hairpin RNA-triggered RNA silencing and related small interfering RNA accumulation. J. Biol. Chem. *279*, 44467–44474.

Katiyar-Agarwal, S., and Jin, H. (2010). Role of Small RNAs in Host-Microbe Interactions. Annu. Rev. Phytopathol. *48*, 225–246.

Kellner, R., Bhattacharyya, A., Poppe, S., Hsu, T.Y., Brem, R.B., and Stukenbrock, E.H. (2014). Expression profiling of the wheat pathogen Zymoseptoria tritici reveals genomic patterns of transcription and host-specific regulatory programs. Genome Biol. Evol.

Ketel, C., Wang, H.S.W., McClellan, M., Bouchonville, K., Selmecki, A., Lahav, T., Gerami-Nejad, M., and Berman, J. (2009). Neocentromeres form efficiently at multiple possible loci in Candida albicans. PLoS Genet. *5*, e1000400.

Ketting, R.F. (2011). The many faces of RNAi. Dev. Cell *20*, 148–161.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Lamb, J.C., Kato, A., and Birchler, J.A. (2005). Sequences associated with A chromosome centromeres are present throughout the maize B chromosome. Chromosoma *113*, 337–349.

Langmead, B. (2010). Aligning short sequencing reads with Bowtie. Curr Protoc Bioinforma. 1–24.

Lee, H.-C., Chang, S., Choudhary, S., Aalto, A.P., Maiti, M., Bamford, D.H., and Liu, Y. (2009). qiRNA is a new type of small interfering RNA induced by DNA damage. Nature *459*, 274–277.

Lee, H.-C., Li, L., Gu, W., Xue, Z., Crosthwaite, S.K., Pertsemlidis, A., Lewis, Z.A., Freitag, M., Selker, E.U., Mello, C.C., et al. (2010). Diverse pathways generate microRNA-like RNAs and dicer-independent small interfering RNAs in fungi. Mol. Cell *38*, 803–814.

Lefrançois, P., Auerbach, R.K., Yellman, C.M., Roeder, G.S., and Snyder, M. (2013). Centromere-Like Regions in the Budding Yeast Genome. PLoS Genet. *9*, e1003209.

Lewis, Z.A., Honda, S., Khlafallah, T.K., Jeffress, J.K., Freitag, M., Mohn, F., Schübeler, D., and Selker, E.U. (2009). Relics of repeat-induced point mutation direct heterochromatin formation in Neurospora crassa. Genome Res. *19*, 427–437.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics *25*, 1451–1452.

Liu, H., Cottrell, T.R., Pierini, L.M., Goldman, W.E., and Doering, T.L. (2002). RNA Interference in the Pathogenic Fungus Cryptococcus neoformans. Genetics *160*, 463–470.

Lu, S., Sun, Y.-H., Amerson, H., and Chiang, V.L. (2007). MicroRNAs in loblolly pine (Pinus taeda L.) and their association with fusiform rust gall development. Plant J. *51*, 1077–1098.

Luijsterburg, M.S., Dinant, C., Lans, H., Stap, J., Wiernasz, E., Lagerwerf, S., Warmerdam, D.O., Lindh, M., Brink, M.C., Dobrucki, J.W., et al. (2009). Heterochromatin protein 1 is recruited to various types of DNA damage. J. Cell Biol. *185*, 577–586.

Ma, L.-J., Van Der Does, H.C., Borkovich, K.A., Coleman, J.J., Daboussi, M.-J., Di Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., Henrissat, B., et al. (2010). Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. Nature *464*, 367–373.

Malik, H.S., and Henikoff, S. (2001). Adaptive evolution of Cid, a centromere-specific histone in Drosophila. Genetics *157*, 1293–1298.

Mallory, A., and Vaucheret, H. (2010). Form, function, and regulation of Argonaute proteins. Plant

Cell *22*, 3879–3889.

Malone, C.D., and Hannon, G.J. (2009). Small RNAs as guardians of the genome. Cell *136*, 656–668.

Marschner, S., Kumke, K., and Houben, A. (2007). B chromosomes of B. dichromosomatica show a reduced level of euchromatic histone H3 methylation marks. Chromosom. Res. *15*, 215–222.

Martienssen, R.A., Zaratiegui, M., and Goto, D.B. (2005). RNA interference and heterochromatin in the fission yeast Schizosaccharomyces pombe. Trends Genet. *21*, 450–456.

Masumoto, H., Nakano, M., and Ohzeki, J.-I. (2004). The role of CENP-B and alpha-satellite DNA: de novo assembly and epigenetic maintenance of human centromeres. Chromosom. Res. *12*, 543–556.

McDonald, B.A., and Martinez, J.P. (1991). Chromosome length polymorphisms in a Septoria tritici population. Curr. Genet. *19*, 265–271.

Mehrabi, R., Taga, M., and Kema, G.H.J. (2007). Electrophoretic and cytological karyotyping of the foliar wheat pathogen Mycosphaerella graminicola reveals many chromosomes with a large size range. Mycologia *99*, 868–876.

Mishra, P.K., Baum, M., and Carbon, J. (2007). Centromere size and position in Candida albicans are evolutionarily conserved independent of DNA sequence heterogeneity. Mol. Genet. Genomics *278*, 455–465.

Napoli, C., Lemieux, C., and Jorgensen, R. (1990). lntroduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes Ín trans. Plant Cell *2*, 279–289.

Ng, R., and Carbon, J. (1987). Mutational and In Vitro Protein-Binding Studies on Centromere DNA from Saccharomyces cerevisiae. Mol. Cell. Biol. *7*.

Nicolas, F.E., Moxon, S., de Haro, J.P., Calo, S., Grigoriev, I. V., Torres-Martínez, S., Moulton, V., Ruiz-Vázquez, R.M., and Dalmay, T. (2010). Endogenous short RNAs generated by Dicer 2 and RNA-dependent RNA polymerase 1 regulate mRNAs in the basal fungus Mucor circinelloides. Nucleic Acids Res. *38*, 5535–5541.

Nicolás, F.E., Torres-Martínez, S., and Ruiz-Vázquez, R.M. (2013). Loss and retention of RNA interference in fungi and parasites. PLoS Pathog. *9*, e1003089.

Nunes, C.C., Gowda, M., Sailsbery, J., Xue, M., Chen, F., Brown, D.E., Oh, Y., Mitchell, T.K., and Dean, R.A. (2011). Diverse and tissue-enriched small RNAs in the plant pathogenic fungus, Magnaporthe oryzae. BMC Genomics *12*, 288–308.

Obbard, D.J., Jiggins, F.M., Halligan, D.L., and Little, T.J. (2006). Natural selection drives extremely rapid evolution in antiviral RNAi genes. Curr. Biol. *16*, 580–585.

Padmanabhan, S., Thakur, J., Siddharthan, R., and Sanyal, K. (2008). Rapid evolution of Cse4p-rich centromeric DNA sequences in closely related pathogenic yeasts, Candida albicans and Candida

dubliniensis. Proc. Natl. Acad. Sci. U. S. A. *105*, 19797–19802.

Page, B.T., Wanous, M.K., and Birchler, J.A. (2001). Characterization of a Maize Chromosome 4 Centromeric Sequence: Evidence for an Evolutionary Relationship With the B Chromosome Centromere. Genetics *159*, 291–302.

Palmer, J.M., and Keller, N.P. (2010). Secondary metabolism in fungi: does chromosomal location matter? Curr. Opin. Microbiol. *13*, 431–436.

Peng, J.C., and Karpen, G.H. (2008). Epigenetic regulation of heterochromatic DNA stability. Curr. Opin. Genet. Dev. *18*, 204–211.

Pfaffl, M.W., Horgan, G.W., and Dempfle, L. (2002). Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res. *30*, e36.

Poláková, S., Blume, C., Zárate, J.A., Mentel, M., Jørck-Ramberg, D., Stenderup, J., and Piskur, J. (2009). Formation of new chromosomes as a virulence mechanism in yeast Candida glabrata. Proc. Natl. Acad. Sci. U. S. A. *106*, 2688–2693.

Ponomarenko, A., Goodwin, S.B., and Kema, G.H.J. (2011). Septoria tritici blotch (STB) of wheat. Plant Heal. Instr. 1–8.

Pumplin, N., and Voinnet, O. (2013). RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence. Nat. Rev. Microbiol. *11*, 745–760.

Raman, V., Simon, S.A., Romag, A., Demirci, F., Mathioni, S.M., Zhai, J., Meyers, B.C., and Donofrio, N.M. (2013). Physiological stressors and invasive plant infections alter the small RNA transcriptome of the rice blast fungus, Magnaporthe oryzae. BMC Genomics *14*, 326.

Rando, O.J., and Winston, F. (2012). Chromatin and transcription in yeast. Genetics *190*, 351–387.

Rhind, N., Chen, Z., Yassour, M., Thompson, D.A., Haas, B.J., Habib, N., Wapinski, I., Roy, S., Lin, M.F., Heiman, D.I., et al. (2011). Comparative Functional Genomics of the Fission Yeasts. Science. *930*, 930–936.

Riddle, N.C., and Birchler, J.A. (2003). Effects of reunited diverged regulatory hierarchies in allopolyploids and species hybrids. Trends Genet. *19*, 593–597.

Rodríguez-Paredes, M., and Esteller, M. (2011). Cancer epigenetics reaches mainstream oncology. Nat. Med. *17*, 330–339.

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

Rouxel, T., Grandaubert, J., Hane, J.K., Hoede, C., van de Wouw, A.P., Couloux, A., Dominguez, V., Anthouard, V., Bally, P., Bourras, S., et al. (2011). Effector diversification within compartments of the Leptosphaeria maculans genome affected by Repeat-Induced Point mutations. Nat. Commun. *2*, 202.

Roy, B., and Sanyal, K. (2011). Diversity in requirement of genetic and epigenetic factors for centromere function in fungi. Eukaryot. Cell *10*, 1384–1395.

Rudd, M.K., Endicott, R.M., Friedman, C., Walker, M., Young, J.M., Osoegawa, K., de Jong, P.J., Green, E.D., and Trask, B.J. (2009). Comparative sequence analysis of primate subtelomeres originating from a chromosome fission event. Genome Res. *19*, 33–41.

Sanyal, K., Baum, M., and Carbon, J. (2004). Centromeric DNA sequences in the pathogenic yeast Candida albicans are all different and unique. Proc. Natl. Acad. Sci. U. S. A. *101*, 11374–11379.

Schmidt, S.M., Houterman, P.M., Schreiver, I., Ma, L., Amyotte, S., Chellappan, B., Boeren, S., Takken, F.L.W., and Rep, M. (2013). MITEs in the promoters of effector genes allow prediction of novel virulence genes in Fusarium oxysporum. BMC Genomics *14*, 119.

Schübeler, D., MacAlpine, D.M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D.E., O'Neill, L.P., Turner, B.M., Delrow, J., et al. (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. Genes Dev. *18*, 1263–1271.

Schueler, M.G., Higgins, A.W., Rudd, M.K., Gustashaw, K., and Willard, H.F. (2001). Genomic and genetic definition of a functional human centromere. Science. *294*, 109–115.

Smith, K.M., Phatale, P.A., Sullivan, C.M., Pomraning, K.R., and Freitag, M. (2011). Heterochromatin is required for normal distribution of Neurospora CenH3. Mol. Cell. Biol. *31*, 2528–2542.

Smith, K.M., Galazka, J.M., Phatale, P.A., Connolly, L.R., and Freitag, M. (2012). Centromeres of filamentous fungi. Chromosom. Res. *20*, 635–656.

Soyer, J.L., El Ghalid, M., Glaser, N., Ollivier, B., Linglin, J., Grandaubert, J., Balesdent, M.-H., Connolly, L.R., Freitag, M., Rouxel, T., et al. (2014). Epigenetic control of effector gene expression in the plant pathogenic fungus Leptosphaeria maculans. PLoS Genet. *10*, e1004227.

Stimpson, K.M., and Sullivan, B.A. (2010). Epigenomics of centromere assembly and function. Curr. Opin. Cell Biol. *22*, 772–780.

Stukenbrock, E.H., Banke, S., Javan-Nikkhah, M., and McDonald, B.A. (2007). Origin and domestication of the fungal wheat pathogen Mycosphaerella graminicola via sympatric speciation. Mol. Biol. Evol. *24*, 398–411.

Stukenbrock, E.H., Bataillon, T., Dutheil, J.Y., Hansen, T.T., Li, R., Zala, M., McDonald, B.A., Jun, W., and Schierup, M.H. (2011). The making of a new pathogen: Insights from comparative population genomics of the domesticated wheat pathogen Mycosphaerella graminicola and its wild sister species. Genome Res. *21*, 2157–2166.

Thakur, J., and Sanyal, K. (2013). Efficient neocentromere formation is suppressed by gene conversion to maintain centromere function at native physical chromosomal loci in Candida albicans. Genome Res. 638–652.

Tsutsumi, A., Kawamata, T., Izumi, N., Seitz, H., and Tomari, Y. (2011). Recognition of the pre-miRNA structure by Drosophila Dicer-1. Nat. Struct. Mol. Biol. *18*, 1153–1158.

Vagnarelli, P., Ribeiro, S.A., and Earnshaw, W.C. (2008). Centromeres: old tales and new tools. FEBS Lett. *582*, 1950–1959.

VanEtten, H.D., Jorgensen, S., Enkerli, J., and Covert, S.F. (1998). Inducing the loss of conditionally dispensable chromosomes in Nectria haematococca during vegetative growth. Curr. Genet. *33*, 299–303.

Vetukuri, R.R., Avrova, A.O., Grenville-briggs, L.J., van West, P., Söderbom, F., Savenkov, E.I., Whisson, S.C., and Dixelius, C. (2011). Evidence for involvement of Dicer-like, Argonaute and histone deacetylase proteins in gene silencing in Phytophthora infestans. Mol. Plant Pathol. *12*, 772–785.

Vetukuri, R.R., Åsman, A.K.M., Tellgren-Roth, C., Jahan, S.N., Reimegård, J., Fogelqvist, J., Savenkov, E., Söderbom, F., Avrova, A.O., Whisson, S.C., et al. (2012). Evidence for Small RNAs Homologous to Effector-Encoding Genes and Transposable Elements in the Oomycete Phytophthora infestans. PLoS One *7*, e51399.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I.S., and Martienssen, R.A. (2002). Regulation of Heterochromatic Silencing and Histone H3 Lysine-9 Methylation by RNAi. Science. *297*, 1833–1837.

Volpe, T.A., Schramke, V., Hamilton, G.L., White, S.A., Teng, G., Martienssen, R.A., and Allshire, R.C. (2003). RNA interference is required for normal centromere function in fission yeast. Chromosom. Res. *11*, 137–146.

Wang, X., Hsueh, Y.-P., Li, W., Floyd, A., Skalsky, R., and Heitman, J. (2010). Sex-induced silencing defends the genome of Cryptococcus neoformans via RNAi. Genes Dev. *24*, 2566–2582.

Wang, X., Wang, P., Sun, S., Darwiche, S., Idnurm, A., and Heitman, J. (2012). Transgene induced co-suppression during vegetative growth in Cryptococcus neoformans. PLoS Genet. *8*, e1002885.

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M.Q., et al. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nat. Genet. *40*, 897–903.

Weiberg, A., Wang, M., Lin, F.-M., Zhao, H., Zhang, Z., Kaloshian, I., Huang, H.-D., and Jin, H. (2013). Fungal Small RNAs Suppress Plant Immunity by Hijacking Host RNA Interference Pathways. Science. *342*, 118–123.

Westwood, J.H., Roney, J.K., Khatibi, P.A., and Stromberg, V.K. (2009). RNA translocation between parasitic plants and their hosts. Pest Manag. Sci. *65*, 533–539.

Wilson, R.C., and Doudna, J.A. (2013). Molecular mechanisms of RNA interference. Annu. Rev. Biophys. *42*, 217–239.

De Wit, E., and de Laat, W. (2012). A decade of 3C technologies : insights into nuclear organization. Genes Dev. *21*, 11–24.

De Wit, P.J.G.M., van der Burgt, A., Ökmen, B., Stergiopoulos, I., Abd-Elsalam, K.A., Aerts, A.L., Bahkali, A.H., Beenen, H.G., Chettri, P., Cox, M.P., et al. (2012). The genomes of the fungal plant pathogens Cladosporium fulvum and Dothistroma septosporum reveal adaptation to different hosts and lifestyles but also signatures of common ancestry. PLoS Genet. *8*, e1003088.

Wittenberg, A.H.J., van Der Lee, T.A.J., Ben M'barek, S., Ware, S.B., Goodwin, S.B., Kilian, A., Visser, R.G.F., Kema, G.H.J., and Schouten, H.J. (2009). Meiosis drives extraordinary genome plasticity in the haploid fungal plant pathogen Mycosphaerella graminicola. PLoS One *4*, e5863.

Wood, V., Gwilliam, R., Rajandream, M., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S., et al. (2002). Genome sequence of Schizosaccharomyces pombe. Tanpakushitsu Kakusan Koso. *47*, 1215–1220.

Yin, W., Birchler, J. a., and Han, F. (2011). Maize centromeres: where sequence meets epigenetics. Front. Biol. (Beijing). *6*, 102–108.

Zaratiegui, M., Vaughn, M.W., Irvine, D. V., Goto, D., Watt, S., Bähler, J., Arcangioli, B., and Martienssen, R.A. (2011). CENP-B preserves genome integrity at replication forks paused by retrotransposon LTR. Nature *469*, 112–115.

Zhang, P., Li, W., Friebe, B., and Gill, B.S. (2008). The origin of a "zebra" chromosome in wheat suggests nonhomologous recombination as a novel mechanism for new chromosome evolution and step changes in chromosome number. Genetics *179*, 1169–1177.

Zhang, X., Zhao, H., Gao, S., Wang, W.-C., Katiyar-Agarwal, S., Huang, H.-D., Raikhel, N., and Jin, H. (2011). Arabidopsis Argonaute 2 regulates innate immunity via miRNA393(∗)-mediated silencing of a Golgi-localized SNARE gene, MEMB12. Mol. Cell *42*, 356–366.

Zhou, J., Fu, Y., Xie, J., Li, B., Jiang, D., Li, G., and Cheng, J. (2012). Identification of microRNA-like RNAs in a plant pathogenic fungus Sclerotinia sclerotiorum by high-throughput sequencing. Mol. Genet. Genomics *287*, 275–282.

Zofall, M., Fischer, T., Zhang, K., Zhou, M., Cui, B., Veenstra, T.D., and Grewal, S.I.S. (2009). Histone H2A.Z cooperates with RNAi and heterochromatin factors to suppress antisense RNAs. Nature *461*, 419–422.

Acknowledgements

Curriculum Vitae