Heterochromatin and RNAi act independently to ensure genome stability in Mucorales human fungal pathogens

María Isabel Navarro-Mendoza[1,2,*] , Carlos Pérez-Arques[1,2,*], and Joseph Heitman[1,*]

[1]Department of Molecular Genetics and Microbiology, Duke University School of Medicine, NC 27710, USA

[2]Equally-contributing authors

*To whom correspondence may be addressed

Corresponding authors: Joseph Heitman (heitm001@duke.edu), María Isabel Navarro-Mendoza (maribel.navarro@duke.edu), and Carlos Pérez-Arques (carlos.parq@duke.edu). 322 Research Drive, CARL Building, Box 3546. Durham, NC 27710. USA. Telephone: 919-684-2824

Classification: Biological Sciences – Genetics

Keywords: RNA interference (RNAi), H3K9-methylation, heterochromatin, epigenetics, transposable elements, Mucorales, fungal pathogens.

Abstract

Chromatin modifications play a fundamental role in controlling transcription and genome stability and yet despite their importance, are poorly understood in early-diverging fungi. We present the first comprehensive study of histone-lysine and DNA methyltransferases across the Mucoromycota, uncovering H3K9-methylated heterochromatin as the major chromatin modification repressing transcription in these fungi, which lack both Polycomb silencing and cytosine methylation. Although RNA interference (RNAi) facilitates the formation of heterochromatin in many eukaryotic organisms, we show that RNAi is not required to maintain either genomic or centromeric heterochromatin in *Mucor*. H3K9-methylation and RNAi act independently to control centromeric regions, suggesting a functional sub-specialization. Whereas the H3K9 methyltransferase Clr4 and heterochromatin formation are essential for cell viability, RNAi is dispensable for viability yet acts as the main epigenetic, regulatory force repressing transposition of GremLINE1 elements. Mutations inactivating canonical RNAi lead to rampant transposition and insertional inactivation of targets resulting in antimicrobial drug resistance. This fine-tuned, Rdrp2-dependent RNAi activity is critical for genome stability, restricting GremLINE1 transposable elements to the centromeres where they occupy long heterochromatic islands. Taken together, our results suggest that RNAi and heterochromatin formation are independent genome defense and regulatory mechanisms in the Mucorales, contributing to a paradigm shift from the co-transcriptional gene silencing observed in fission yeasts to models in which heterochromatin and RNAi operate independently in early-diverging basal fungi.

Significance

The Mucorales are a group of basal fungi known to cause mucormycosis, a fungal infection with high mortality rates and for which treatment options are limited. Understanding key regulatory mechanisms that drive adaptation to the host, influence virulence traits, and control genome stability will provide insights to overcome the challenges posed by these fungal pathogens. Our research establishes constitutive heterochromatin and RNAi as two central epigenetic, regulatory forces that act independently to repress transposable elements in the Mucorales. Modulating these regulatory mechanisms may lead to reduced transposition and subsequently to less frequent drug target disruption that could be exploited to combat antimicrobial drug resistance.

Introduction

Fungi are a diverse clade of eukaryotic organisms that have successfully adapted to niches across the planet (1). To do so, they have evolved ecological strategies to survive and compete in diverse environments (2). In some cases, these result in pathogenic interactions and host disease, causing economic and human losses (3). Adaptation to hosts is a consequence of different molecular mechanisms modulating gene expression, and understanding this provides insights to lessen the burden associated with fungal diseases (4, 5). Many of these mechanisms rely on epigenetic modifications that do not alter the DNA sequence but rather control chromatin accessibility by producing distinct chromatin states. The repressed chromatin state, known as heterochromatin, is frequently associated with low gene density and highly repetitive DNA content, and heterochromatic regions are marked by methylation of histone H3 lysine 9 (H3K9me) and/or lysine 27 (H3K27me), considered as constitutive or facultative heterochromatin, respectively. Methylation at these sites is catalyzed by histone lysine methyltransferases (or KMT enzymes) (6). In addition, other epigenetic modifications may contribute to heterochromatin formation, including cytosine methylation at the carbon-5 position (5mC) catalyzed by DNA methyltransferases (DNMT) (7). These “writer” proteins add the methylation mark, as opposed to “eraser” demethylases that remove it; and specific “reader” proteins, which recognize epigenetic marks for signaling, and recruit “writers” to their targets for reinforcement (8, 9).

In the fission yeast *Schizosaccharomyces pombe*, pericentric heterochromatin formation by co-transcriptional gene silencing (CTGS) serves as a paradigmatic example. RNA polymerase II-based transcription from repetitive sequences triggers RNA interference (RNAi), generating small non-coding RNAs (sRNAs) that guide Argonaute and associated RNAi-induced transcriptional silencing (RITS) complex components to complementary nascent transcripts (10–12). Next, the RITS complex recruits the “writer” methyltransferase Clr4/KMT1 to catalyze H3K9me (13), and methylation is reinforced, maintained, and extended by several

chromodomain-containing “readers”: Chp1 from the RITS complex (14), the heterochromatin protein 1 (HP1) proteins Chp2 and Swi6 (15, 16), and Clr4 (17). Beyond this paradigm, heterochromatin formation has been characterized in ascomycetous and basidiomycetous fungi revealing new insights into 5mC DNA methylation and Polycomb silencing through KMT6-mediated H3K27me (18–24). Interestingly, *S. pombe* lacks both DNA methylation and H3K27me and the necessary “writer” proteins to produce these, highlighting the importance of alternative fungal models and developing techniques to study them. The fact that the Dikarya – Ascomycota and Basidiomycota– comprises only two out of the numerous fungal phyla suggests other regulatory, epigenetic mechanisms remain to be discovered in the early-diverging fungi, which the latest classifications estimate may comprise as many as seventeen phyla (25, 26).

The Mucorales are a group of early-diverging fungal pathogens known to cause devastating infections called mucormycoses. These infections cause significant morbidity and mortality (27), and mechanisms leading to antifungal resistance and pathogenesis are poorly understood. The fungus *Mucor lusitanicus* serves a model for this group of pathogens. Molecular techniques have been developed and applied to characterize RNA interference (RNAi) as a post-transcriptional gene silencing (PTGS) mechanism (28). Double-stranded RNA (dsRNA) produced by RNA-dependent RNA polymerase (RdRP) (29, 30) is cleaved into sRNAs by Dicer ribonucleases (30–32). sRNAs are loaded into the RNAi-induced silencing complex (RISC) where Argonaute (Ago) directs degradation of the complementary messenger RNA by base-pairing homology and ribonuclease activity (33, 34). The canonical RNAi machinery competes with an alternative RNAi degradative pathway relying on RdRP and the novel ribonuclease R3B2 to silence its own set of target transcripts (35, 36). These RNAi pathways modulate distinct biological processes including pathogenesis, regulation of centromeric transposons, and epigenetic antimicrobial drug resistance (36–41).

RNAi-based epimutations were discovered for the first time in two species of *Mucor,* causing drug resistance by sRNA silencing of the *fkbA* gene encoding the FK506 target, FKBP12 (37). Later, epimutations were found to silence other drug targets, e.g., uracil biosynthetic genes targeted by pyrimidine analogs (5-FOA) (42). These epimutations result in unstable, transient drug resistances that revert to sensitivity after several mitotic growth cycles in the absence of drug. During host infection, epimutations persist and are induced in specific organs, suggesting a role in pathogenesis (38). Epimutations that promote drug resistance and host colonization pose a concern both for public health and agriculture (23, 37, 43–45), as they could elude detection due to their transient, unstable nature. Therefore, it is critical to elucidate the repertoire of epigenetic modifications these pathogens deploy, and how these pathways operate and interact to modify gene expression in response to environmental cues.

The diversity of RNAi mechanisms in *Mucor* suggests distinct roles and functional specialization that may involve recruiting other epigenetic modifications as seen in several other eukaryotes (46–48). Here, we define the main forces driving heterochromatin formation in the Mucorales, characterize the role of RNAi and other epigenetic modifications in establishing and maintaining altered chromatin states, and determining the impact of epigenetic regulation on centromere identity and genome stability.

Results

Clr4-mediated histone-lysine 9 methylation is an essential regulatory mechanism in Mucorales

The Mucorales express a diverse repertoire of RNAi components that are involved in distinct repressive, regulatory pathways relying on sRNA-directed degradation of complementary mRNAs that results in PTGS (10, 46, 47). These sRNAs also serve as initiators of epigenetic modifications leading to transcriptional repression in other eukaryotic organisms. While many of

these modifications have been characterized in the Dikarya, their conservation and role in early-diverging fungi was unexplored. Therefore, the presence of epigenetic modifications was first surveyed in early-diverging fungal genomes, focusing on the Mucorales. As in most eukaryotes, Mucoralean histone proteins H3 and H4 are highly conserved (40). The *M. lusitanicus* genome encodes four histone H3 and three histone H4 copies (*SI Appendix*, Fig. S1A-B), and the KMT-lysine substrates and neighboring regions are also conserved. Thus, these proteins could harbor epigenetic modifications catalyzed by KMT enzymes or other histone-lysine modifiers. To identify RNAi components, and KMT and DNMT homologs in the Mucorales, 86 fungal proteomes were analyzed including at least one representative of most fungal phyla, focusing on the Mucoromycota phylum and major fungal pathogens (*SI Appendix*, Dataset S1). Out of these proteomes, 43 were selected as representatives to capture the presence or absence of different epigenetic modifying enzymes (Fig. 1). We identified homologs of the three major RNAi enzymes –RdRP, Dicer, and Ago– across all fungal phyla and particularly, in every Mucoromycota species, though not in some isolated Dikarya species or clades, e.g., *Saccharomyces cerevisiae*, *Candida auris, Ustilago maydis*, and *Malassezia pachydermatis*. These losses are well documented (49–51) and validate the approach. In *S. pombe*, CTGS leading to constitutive heterochromatin initiates in regions producing complementary sRNA that are loaded into the RITS complex by the Argonaute siRNA chaperones (ARC), Arb1 and Arb2. These two proteins are mostly conserved in the Mucorales, ascomycetes that have retained RNAi, and some early-diverging fungi. However, RITS-exclusive components –Chp1 and Tas3– are only found in *S. pombe*, indicating that any possible interplay among RNAi and heterochromatin formation might require novel components.

Next, KMT homologs and accessory proteins in the methyltransferase complexes were identified. KMT enzymes are often composed of a SET domain which is involved in methyltransferase activity, and a variety of other motifs that mediate binding to specific lysine

targets and accessory proteins. Clr4/KMT1 (also known as DIM-5, Su(var)3-9, and SUV39H), KMT6 (SET-7, E(z), and EZH2), and KMT5 (Set9, SET-10, Hmt4-20, and SUV420H) catalyze methylation of H3K9, H3K27, and H4K20, respectively (*SI Appendix*, Fig. S1A-B), resulting in transcriptional repression. In contrast, KMT2 (Set1, SET-1, and SETD1), KMT3 (Set2, SET-2, and SET2D), and KMT4 (Dot1, DOT-1/HLM-2, Gpp, and DOT1L) methylate H3K4, H3K36, and H3K79, respectively, and mediate transcriptional activation. Clr4 and associated proteins, Rik1 and the ubiquitin ligase Cul4, are highly conserved in fungi, particularly in the Mucorales (Fig 1). In *S. pombe*, the Clr4 complex (CLRC) has been shown to interact with Ago1 through Stc1, which connects both pathways to establish H3K9me via RNAi (50). Interestingly, Stc1 homologs are present in the Mucoromycota despite the lack of other RITS components. The heterochromatin protein 1 (HP1) complex (Swi6/Chp2 homologs) is also conserved among the Mucoromycota, and all early-diverging fungi that have retained RNAi. Surprisingly, KMT6 was lost in every Mucoromycota species, yet is conserved in some other early-diverging fungal phyla. It was also lost in some Ascomycota and many Basidiomycota, suggesting that H3K27me and Polycomb silencing was lost in many independent events during fungal evolution. KMT enzymes involved in transcription-activating methylation (KMT2, KMT3, and KMT4) are conserved across fungal evolution, except for the H3K79 methyltransferase KMT4 in *S. pombe.*

Lastly, we searched for DNMT enzymes, specifically those capable of producing transcription-repressive 5mC DNA methylation (Fig. 1). Fungal specific *de novo* methyltransferase RID-1 is exclusively conserved in the Pezizomycotina, while DIM-2-like homologs were identified in early-diverging Endogonales and Glomeromycota. Traditionally, DIM-2 and RID-1 were considered the main *de novo* DNA methyltransferases operating in fungi, and recent phylogenetic aligns them with a monophyletic clade that includes DNMT1 from animals and plants (52). While DNMT1 homologs were identified in some Mucorales and other early-diverging fungi, it was notably absent in the pathogenic *Mucor* and *Rhizopus* species.

Overall, most Mucoromycota species have lost KMT enzymes and associated proteins involved in H3K27me3-mediated Polycomb silencing through facultative heterochromatin, as well as DNMT proteins catalyzing DNA 5mC. Because Clr4/KMT1 is the only transcription-repressive chromatin-modifying enzyme present in every Mucoralean species analyzed, we focused on the *M. lusitanicus clr4* gene. First, the previous *clr4* gene model prediction was curated and reannotated using our RNA sequencing transcriptomic data. This revealed three additional exons at the 5'-end of the reannotated gene, encoding a predicted chromodomain essential for H3K9me binding activity (*SI Appendix*, Fig. S2A). The presence of a chromodomain in the *M. lusitanicus* Clr4 protein prompted us to inspect every Clr4 homolog in species with available transcriptomic data and alternative gene models. We found many N-terminally truncated, misannotated proteins and uncovered as many chromodomains in the early-diverging fungi (*SI Appendix*, Fig. S2B, Dataset S2 Clr4/KMT1). A phylogenetic analysis across the fungal kingdom revealed that most fungal, specifically early-diverging, Clr4 enzymes harbor an N-terminal chromodomain, suggesting six major and independent losses of this domain in the last common ancestors of the Basidiomycota, Saccharomycotina and Pezizomycotina, Phycomycetaceae, Umbelopsidales, Olpidiomycota, and Blastocladiomycota (Fig. 1, *SI Appendix*, Fig. S2B).

To gain a better understanding of Clr4 function in Mucorales, we attempted to delete *clr4* through CRISPR-Cas9 and homologous recombination in *M. lusitanicus*. Mucorales have non-septate hyphae and multinucleated asexual spores. As such, genetic transformation often results in initially heterokaryotic mutants, producing spores that harbor both wild-type and mutant nuclei. Mutant allele segregation by sexual reproduction is not possible because *M. lusitanicus* is not able to undergo a complete sexual cycle under laboratory conditions. To render these mutants homokaryotic, i.e., harboring mutant nuclei exclusively, transformants must be passaged through several vegetative growth cycles in a medium that positively selects the mutant allele. Surprisingly, every transformation experiment aiming at *clr4* deletion resulted in heterokaryotic

mutants that were unable to achieve complete mutant homokaryosis, and after twenty vegetative passages they still harbored wild-type nuclei (*SI Appendix*, Fig. S3). We tested if Clr4 activity could be essential for viability by genotyping the asexual progeny of a heterokaryotic mutant, reasoning an equal rate of wild-type or mutant homokaryons in the spore population is expected for a non-deleterious mutation. To do so, 40 spores from a heterokaryotic mutant were dissected and plated on non-selective medium, and after incubation, 30 spores were able to germinate. Next, we analyzed the germinated spores and found either wild-type homokaryotic or heterokaryotic colonies (with wild-type and *clr4*Δ nuclei), yet no *clr4*Δ homokaryons were observed (*SI Appendix*, Fig. S3). These results indicate that homokaryotic *clr4*Δ spores are not able to survive, implying Clr4 is essential and suggesting that a functional H3K9me-mediated heterochromatin formation mechanism is required for cell viability in *Mucor*.

Regions targeted by H3K9 methylation are transcriptional deserts enriched with repeats

Chromatin immunoprecipitation followed by Illumina sequencing (ChIP-seq) using α-H3K9 di- (H3K9me2) or tri-methylated (H3K9me3) antibodies was performed to characterize H3K9me in *Mucor*. Reads were mapped to the publicly available *M. lusitanicus* MU402 v1.0 genome (https://mycocosm.jgi.doe.gov/Muccir1_3/Muccir1_3.home.html), and H3K9me2 and -me3 enriched peaks were identified across the genome. Genome-wide observations (Fig. 2A and *SI Appendix*, Fig. S4) revealed a striking overlapping pattern between H3K9me2 and -me3, which was confirmed by linear correlation measurements of every enriched peak detected (Fig. 2B).

To determine if H3K9me-mediated heterochromatin results in transcriptional silencing, ChIP-seq against RNA polymerase II (RNA pol II) and ribosomal RNA-depleted strand-specific RNA sequencing (RNA-seq) were performed. H3K9me peaks emerge as heterochromatin islands in transcriptional deserts, which are devoid of mRNA and RNA pol II occupancy (Fig. 2A and *SI Appendix*, Fig. S4). Interestingly, genes are absent from most of these heterochromatic regions, suggesting H3K9me is targeting other genomic elements. A *de novo* repeat identification analysis

was conducted across the genome of *M. lusitanicus*, resulting in an estimate of 22.3% genome-wide repeated DNA content, 53.9% gene sequences, and the remaining percent intergenic sequences. An in-depth, base pair-resolution analysis confirmed the genome-wide observations, showing that ~ 19.7% of the genome is bound by H3K9me chromatin which is mainly directed at repeats (15.3%, Fig. 2C). In contrast, only 1.2% of the genome is targeted by H3K9me across coding sequences, suggesting that H3K9me is mostly controlling repeats rather than regulating gene expression.

Because previous studies have shown that *M. lusitanicus* Dicer-dependent, canonical sRNAs are able to target repeated and transposon sequences (30, 36) and most of the H3K9me islands are localized to repeats, the sRNA distribution across the whole genome was analyzed by conducting sRNA-enriched RNA purification and sequencing (sRNA-seq). The sRNA content analyzed exhibits key features of canonical sRNAs: 1) a discrete length ranging from 21 to 24 nt and 2) a strong preference for uracil at the 5'-end (Fig. 2D). These canonical sRNAs coincide extensively with H3K9me at silenced repeated regions (Fig. 2A and Fig. S4), indicating a dual regulation by post-transcriptional and transcriptional silencing mechanisms. In many eukaryotic genomes, repeated regions represent copies of transposable elements (TE) whose transposition gave rise to fine-tuned regulatory mechanisms that shaped their genome evolution (53).

To determine the extent of transposable elements regulated by either or both H3K9me and RNAi, we conducted the first comprehensive TE identification analysis in the *Mucor* species complex. Briefly, open reading frames (ORFs) were predicted in every repeat type identified in our previous analysis, and searched for protein domains frequently associated with TE activity. The presence and relative position of these domains within each sequence was used to manually curate TE predictions, identifying six different TE families: DNA or Class II elements Tc1/Mariner, PIF/Harbinger, and RC/Helitron; and RNA or Class I elements non-LTR GremLINE1 and other LINE, and LTR Ty3/Gypsy (Fig. 2E, *SI Appendix*, Table S1, Dataset S3). All of these families

comprised several distinct elements that were classified into subfamilies according to sequence similarity, particularly the previously identified GremLINE1 retrotransposons (40). After generating this curated TE library, the whole genome and annotated TE sequences were screened using restrictive parameters, ensuring only highly similar and full-length elements were included. Then, TE distribution across the genome was examined to detect conspicuous clustering patterns. Only the GremLINE1 elements occupy a specific genomic region, colonizing the pericentric regions of *M. lusitanicus*, whereas the remaining elements have a cosmopolitan distribution across the genome (Fig. 2F).

Taken together, these findings reveal an intimate overlap between H3K9me2 and -me3 islands primarily forming in repeated regions subject to transcriptional silencing. Interestingly, most of these genomic regions are also targeted by sRNAs, indicating a possible interconnection between RNAi and heterochromatin formation pathways to control transposition that was explored further.

RNAi is dispensable for heterochromatin maintenance at transposable elements

RNAi relies on sRNAs to induce post-transcriptional silencing of complementary transcripts but in addition, these sRNAs may participate in CTGS by interacting with other repressive epigenetic mechanisms to initiate and maintain heterochromatin formation (10, 46–48). To explore the interaction between the principal epigenetic repressive mechanisms in *Mucor*, RNAi and H3K9me levels were compared in the newly identified TEs. Class I and II transposons are traditionally classified according to their transposition intermediates: RNA or DNA, respectively. An RNA intermediate could be more readily targeted by RNAi and consequently, the analysis was focused on RNA transposons. H3K9me binds to most copies of GremLINE1 and Gypsy, and several copies of LINE in the wild-type *M. lusitanicus* strain (Fig. 3). In these elements, H3K9me is deposited across the entire sequence, with the highest enrichment at the start and end positions. As expected, sRNAs are abundantly produced from these regions but interestingly,

they target more TE copies than H3K9me. For instance, some LINE subfamilies generate sRNAs, but H3K9me is completely absent (Fig.3, e.g., LINE subfamily highlighted in green). Nevertheless, none of these elements are being actively transcribed, not even those lacking H3K9me, as evidenced by the transcriptomic and RNA pol II occupancy data (Fig. 3). Because several TE copies are differentially targeted by H3K9me or sRNAs, these mechanisms may be acting independently to regulate these elements.

To investigate this hypothesis, an extensive mutant collection comprised of every key RNAi enzyme identified in *M. lusitanicus* was studied. To confirm their involvement in RNAi, sRNA content was sequenced and compared to previous studies (30, 36) (*SI Appendix*, Fig. S5A). These results recapitulate that sRNAs are almost depleted in *dcl1*Δ *dcl2*Δ, a double mutant lacking any Dicer activity; and in a*go1*Δ, lacking the main Argonaute enzyme, confirming that both activities are essential for sRNA biogenesis and canonical RNAi (*SI Appendix*, Fig. S5B). These canonical sRNAs are mostly produced from exons and repeated sequences (*SI Appendix*, Fig. S5C), suggesting gene and transposon regulation are their primary targets. Therefore, we focused our analysis on canonical (*dcl1*Δ *dcl2*Δ, *ago1*Δ, and *rdrp2*Δ) and alternative (*rdrp1*Δ and *r3b2*Δ) RNAi pathway mutants (Fig. 3). Other RNAi paralog mutants (*dcl1*Δ, *dcl2*Δ, *ago2*Δ, *ago3*Δ and, *rdrp3*Δ) are covered in supplementary material because their function is not essential for *Mucor* RNAi (*SI Appendix*, Fig. S6). ChIP and RNA-seq analyses were conducted to identify H3K9me2 and -me3 enrichment, as well as transcription across the genome.

Strikingly, not a single RNAi-defective mutant showed a significant decrease in H3K9me3 at any of these sequences (Fig. 3, *SI Appendix*, Fig. S6), and no differences in H3K9me2 and -me3 content were observed in any region (*SI Appendix*, Fig. S7). These findings indicate that RNAi is neither sufficient nor necessary to maintain histone methylation across transposable elements. Despite not showing alterations in H3K9me, both mRNA expression and sRNA accumulation were substantially affected in the mutants. The production of sRNAs targeting most

of these regions decreased dramatically in *dcl1*Δ *dcl2*Δ, and to a lesser extent in *ago1*Δ, as expected of Dcl- and Ago-dependent, canonical sRNAs. More importantly, the absence of Dicer or Argonaute activity increased transcript levels in many TE despite being embedded in H3K9me chromatin, especially but not restricted to GremLINE1 elements. Conversely, some TE subfamilies showed a marked decrease in canonical sRNA production in RNAi-defective mutants but remained transcriptionally silent or inactive, suggesting H3K9me and RNAi may operate independently to repress TE expression. Interestingly, RNA pol II occupancy remained stable in all of the RNAi mutants, indicating that increased transcript accumulation was a result of lacking PTGS instead of increased transcription. These results confirm that *Mucor* RNAi pathways act exclusively via PTGS.

S. pombe Rdp1 and the RNA-directed RNA polymerase complex (RDRC) are essential to initiate co-transcriptional gene silencing at pericentric repeats (54). In contrast, *M. lusitanicus* encodes three subspecialized RdRP enzymes (29, 55), suggesting more complex regulatory dynamics. To address this, sRNA production and transcript expression across TE sequences were studied in the RdRP and alternative RNAi pathway mutants. Surprisingly, obvious differences were observed in centromeric GremLINE1 retrotransposons compared to other elements (Fig. 3). Rdrp1 and R3B2 –the main components of the alternative pathway– are important to produce sRNAs from LINE and Gypsy retrotransposons and to regulate their expression, suggesting an interaction between canonical and alternative RNAi pathways to control these elements. In contrast, post-transcriptional silencing of GremLINE1 elements is completely dependent on the canonical pathway because sRNA and transcript levels remain unchanged in mutants of the alternative RNAi pathway. These differences in TE regulation suggest that centromeric GremLINE1s are controlled by a unique RNAi-based mechanism, possibly as a consequence of their role in centromere identity.

To confirm the extent of GremLINE1 unique regulation, the remaining DNA transposable elements (Tc1/Mariner, PIF/Harbinger, and RC/Helitron) were examined (*SI Appendix*, Fig. S8A). Most copies of DNA transposons show H3K9me across the coding sequence. Despite relying on a DNA intermediate for transposition, sRNAs are produced from Mariner and Harbinger copies as well as several Helitron subfamilies through canonical and alternative RNAi pathways. Only Mariner and Helitron subfamilies 1 (Fig.S8A, green in both Mariner and Helitron) depend on RNAi to repress their expression as *dcl1*Δ *dcl2*Δ and *ago1*Δ exhibit increase transcript levels, similar to Gypsy and LINE RNA transposons. On the other hand, the remaining DNA element copies are transcriptionally silenced, possibly as a result of H3K9me. Due to the presence of full-length elements and the necessary protein products for autonomous transposition, we reasoned that transcripts encoding transposase protein domains might be targeted by RNAi. Indeed, sRNAs are specifically directed to the transposase-encoding transcript, suggesting RNAi is silencing transposase transcripts to prevent DNA transposition events as a genome defense mechanism (*SI Appendix*, Fig. S8B). Intriguingly, H3K9me3 is more broadly distributed, extending to neighboring structural repeats, reinforcing that histone methylation is maintained in regions not targeted by RNAi and that both epigenetic mechanisms may recognize and target repeats differently.

RNAi prevents genome instability by repressing GremLINE1 transposition to non-centromeric regions

GremLINE1 elements represent a novel LINE L1-like subfamily of RNA TEs exclusively located in pericentric regions and they are conserved in Mucoromycota species that lost CENP-A, suggesting an important role in centromere-kinetochore dynamics (40). We investigated the regulatory mechanisms controlling GremLINE1 expression, analyzing H3K9me and sRNA accumulation at the centromeres. H3K9me delimits the pericentric regions, extending further from the GremLINE1s and their complementary sRNAs into other repeated sequences, but sparing

neighboring genes (Fig. 4A). Notably, H3K9me is absent at the kinetochore binding region, possibly to maintain an open chromatin state that allows kinetochore binding. Additionally, RNAi exerts fine-tuned, post-transcriptional silencing of GremLINE1 elements (Fig. 3) independently of Rdrp1 –the main RdRP associated with silencing initiation in *Mucor* (29).

Because all other TEs rely on Rdrp1 to generate sRNAs, we determined which RdRPs were triggering RNAi to target the GremLINE1s by searching for differences in natural antisense transcription and thus, dsRNA, across the GremLINE1 retrotransposons. Many GremLINE1 copies exhibited a significant decrease of antisense transcripts in *rdrp2*Δ compared to the wildtype and *rdrp1*Δ (Fig. 4B, FDR ≤ 0.05), indicating that Rdrp2 may generate dsRNA from these elements and induce RNAi. Rdrp2 is the only RdRP that harbors an RNA-recognition motif in its N-terminus (*Appendix SI*, Dataset S2 RdRP), and its activity is essential to maintain silencing through an sRNA-amplifying positive feedback loop (29). We propose that both activities, generating antisense transcription resulting in dsRNA and amplifying silencing, are important to initiate and maintain PTGS directed at the GremLINE1s, and could explain the slight increase of GremLINE1 transcript levels in *rdrp2*Δ (Fig. 3).

Our data confirms that GremLINE1 regulation is unique compared to other repeated elements outside of the centromeres, therefore we hypothesized that loss of RNAi could incite their transposition and lead to genome instability. To test this hypothesis, an artificial evolution experiment was conducted by culturing RNAi-defective and wild-type strains for twenty asexual passages in minimal medium (Fig. 4C). We estimated GremLINE1 copy number variation across these strains as the fold change compared to a single-copy gene by qPCR, resulting in a significant increase in *dcl1*Δ *dcl2*Δ and *ago1*Δ –key components of PTGS (Fig. 4D). Interestingly, transcript levels seem to correlate with the increase in copy number in each given mutant, particularly in *dcl1*Δ *dcl2*Δ and *ago1*Δ, but also in *dcl2*Δ and *rdrp2*Δ though their increase in copy numbers followed a trend that was not statistically significant (Fig. 3 and Fig. 4D).

To further characterize these transposition events, the strains were cultured under positive selection exerted by FK506, an antifungal drug that binds FKBP12, generating a protein-drug complex that inhibits calcineurin and enforces yeast-like growth. Mycelium-growing colonies were selected to capture GremLINE1s in the FKBP12-encoding locus *fkbA* and thus cause FKBP12 loss-of-function and FK506 resistance (Fig. 4C). After passage on the selective medium, 36 spontaneous FK506-resistant colonies were isolated: 10 in *dcl1*Δ *dcl2*Δ, 5 in *ago1*Δ, 1 in *rdrp2*Δ, 10 in *rdrp1*Δ, and 10 in the wild-type strains. Mutation types in each of these isolates were characterized by analyzing *fkbA* fragment size and whole-locus Sanger-sequencing, classifying these as insertions, point mutations or deletions, or epimutations when no genetic alteration was found (*SI Appendix*, Fig. S9, Table S2). Six insertions disrupting the *fkbA* locus were detected in only *dcl1*Δ *dcl2*Δ and *ago1*Δ resistant isolates. Four independent insertion events were fully characterized: one full-length and one truncated GremLINE1 copies in each of the *dcl1*Δ *dcl2*Δ and *ago1*Δ mutants (Fig. 4E, *SI Appendix,* Fig S9, and Table S3). In addition, two double GremLINE1 insertions were detected in *dcl1*Δ *dcl2*Δ, confirmed by PCR with a specific *fkbA* and GremLINE1 primer pair. Every insertion analyzed revealed exclusively a GremLINE1 copy, and the remaining resistant isolates in other mutant or wild-type strains did not harbor any insertions. The presence of A/T-5'-end target site duplications and truncated insertions reflect the LINE L1 canonical transposition mechanism, involving element-encoded endonuclease and reverse transcriptase activities (Fig. 2). Briefly, the endonuclease nicks genomic DNA in AT-rich regions generating an exposed 3'-OH end that primes a reverse transcription reaction of the A-rich 3'-end of the element (Fig. 5). Abrupt termination of the reverse transcription step frequently leads to incomplete insertions, explaining truncated copies (56).

Discussion

Here in this study, we have identified essential chromatin modifying enzymes across the fungal tree of life, focusing on the diverse yet understudied early-diverging fungi. Our findings

imply that transcription-activating epigenetic mechanisms are ancestral and likely essential in fungi, present even in ascomycetaceous yeasts that lost many other chromatin modifying enzymes (57), which highlights the differences in transcriptional ground states between eukaryotic and prokaryotic organisms (58, 59). On the other hand, several transcriptional repressive chromatin-modifying enzymes are recursively lost in many fungi, except Clr4/KMT1 and its associated proteins from the CLRC that are generally conserved in all fungal lineages. The vast majority of early-diverging fungal Clr4 proteins exhibit an N-terminal chromodomain, including those encoded in Mucoralean genomes. When present, the chromodomain H3K9me-binding activity is essential for heterochromatin maintenance and spreading to neighboring regions beyond the initiation sites in fungi (17) and animals (60). The presence of this chromodomain in fungal Clr4 enzymes is surprising because traditionally, *Schizosaccharomyces* spp. Clr4 was considered the only fungal Clr4 protein with a chromodomain, possibly due to a heavy bias towards Dikarya fungal models (9). Most Clr4 methyltransferases in animals (Su(var)3-9 or SUVH39) harbor an N-terminal chromodomain (61), whereas Clr4 enzymes in plants (SUVH) lack this domain (62). This convoluted Clr4/KMT1 evolutionary history of recurrent domain gain or loss led to the hypothesis that the fungal ancestor lacked the chromodomain (9, 63), yet our results indicate that the N-terminal chromodomain-containing variant is the ancestral Clr4 of animals and fungi.

Our research focuses on the human fungal pathogens Mucorales, and we have noted losses of certain chromatin modifying enzymes that must have important implications in their epigenetic regulatory landscape. For instance, the loss of KMT6 indicates the absence of H3K27me and Polycomb-mediated facultative heterochromatin. This is not the only relevant transcription-repressive modification lost in these fungi, because the majority of Mucorales also lack DNA 5mC methyltransferases. These mechanisms have been reported as essential for epimutations governing effector expression during host colonization in several plant pathogens

(23, 43). Interestingly, *S. pombe* lost both Polycomb silencing and 5mC but still develops epimutations based on H3K9me heterochromatin –triggered by sRNAs– to acquire drug resistance (44). We reasoned that fungi may exploit any available epigenetic mechanism to modulate gene expression and better adapt to their current environment and therefore, emphasize how critical it is to characterize epigenetic modifications and their interactions that could contribute to epimutations in fungal pathogens.

Our results suggest that Clr4/KMT1, the CLRC, and HP1 complexes are the major chromatin modifying mechanisms involved in transcriptional silencing through heterochromatin in the Mucorales. Indeed, Clr4 is essential in *M. lusitanicus*, indicating that H3K9me repression controls critical biological processes. H3K9 is methylated sequentially from mono- to trimethylation, and these differentially methylated states define distinct domains of silent chromatin in many organisms, including mammals (64), nematodes (65), and plants (66). All of these organisms either lost the chromodomain necessary for Clr4 reading activity and/or possess a set of subspecialized Clr4 variants mediating the different H3K9 methylation states. In *Mucor*, H3K9me2 and -me3 enriched regions overlap and show a robust correlation across the whole genome and particularly, at repeated sequences. Similarly, wild-type *S. pombe* cells exhibit a strong correlation between H3K9me2 and -me3 in pericentric, telomeric, and the *mat* locus regions. *S. pombe* also shows distinct methylation states but these are only evident upon altering either the Clr4 chromodomain, its SET domain (67), or the bromodomain-containing protein Abo1 (68). Interestingly, the genomes of both *S. pombe* and the Mucorales encode a single Clr4 protein with an N-terminal chromodomain. We hypothesize that *M. lusitanicus* Clr4 is the main histone-lysine methyltransferase driving both H3K9me2 and -me3 owing to its essentiality, and suspect that either or both marks have a critical role in cell viability.

An extensive collection of literature endorses the evolution of epigenetic modifications as a defensive mechanism against invasive nucleic acids, particularly transposable elements and

repeated DNA (43, 44 and references therein). Indeed, distinct epigenetic mechanisms for heterochromatin formation have evolved in many fungi to control transposition, e.g., CTGS in *S. pombe* and H3K9me and 5mC in *C. neoformans* (18, 19) and in several sordariomycetes where the process is initiated by repeat-induced point mutations (RIP) (70, 71). Interestingly in these organisms, sRNAs are also directed to RIPed sequences but the interaction between RNAi and heterochromatin to regulate repeats remains poorly understood in filamentous fungi, and has only been thoroughly established in the fission yeast.

Mucor deploys an extensive array of RNAi components involved in two competing RNAi pathways. The presence of both active RNAi and H3K9me prompted us to explore interactive dynamics between both mechanisms in this fungus. Our findings indicate that RNAi loss in *Mucor* does not impair heterochromatin maintenance although sRNAs and H3K9me frequently co-occur in the same regions. This is intriguing because RNAi components have been shown to participate in transcriptional or chromatin-based silencing in other eukaryotic organisms. For example, loss of Ago activity results in defective H3K9me in plant (46), animal (48), and ciliate species (47). Also, in the fission yeast, Ago, Dicer, and RdRP defective mutants are unable to establish or maintain H3K9me across pericentric regions and the mating-type locus (10, 13). Once heterochromatin is established, the RDRC and RITS complexes are directed to H3K9-methylated regions by the chromodomain-containing factor Chp1, generating more sRNAs to sustain a co-dependent feedback loop between RNAi and H3K9me (72). In contrast, our findings reveal RNAi is dispensable for maintaining H3K9me across the genome of *M. lusitanicus*, and particularly at pericentric TEs. This is reinforced by our lethality analysis (*SI Appendix*, Fig. S3), showing that Clr4 activity is essential for cell viability, in contrast to canonical or alternative RNAi pathway mutants which are viable.

Importantly, we identified regions targeted by canonical sRNAs that are devoid of H3K9me and conversely, H3K9-methylated transposable elements that do not trigger RNAi.

RNAi-independent epigenetic mechanisms have been extensively characterized before in the filamentous fungal model *N. crassa*. Cogoni and Macino showed that H3K9me deposited at RIP-targeted transposon relics is not affected in RNAi mutants, although these relics do accumulate canonical sRNAs (73). Later, another study revealed that the RNAi machinery is also dispensable for DNA 5mC co-occurring at these transposon relics (74). Taken together, these findings suggest that RNAi and H3K9me arose as independent regulatory mechanisms that later evolved into co-dependent relationships in some organisms, probably due to their shared targets.

Overall, our work reinforces that epigenetic regulation in fungi is complex and diverse, reflecting a rich evolutionary history of loss and retention of different epigenetic mechanisms. Besides loss of both DNA 5mC and H3K27me-mediated facultative heterochromatin formation machineries, Mucoralean genomes were shaped by loss of two critical kinetochore proteins, CENP-A and CENP-C, resulting in a novel type of mosaic centromere that harbors hybrid features of point and regional centromeric sequences (40). Mosaic pericentromeres are colonized by large, inverted repeats comprising GremLINE1 retrotransposons, and our work has uncovered an exclusive, dual regulation by both H3K9me and canonical RNAi that sets them apart from the remaining transposable elements resident in the genome. H3K9-methylated pericentric regions, spanning up to 70 kilobases, are the largest heterochromatin islands in the genome (Fig. 2A and 4A and *SI Appendix,* Fig. S4) but surprisingly, these rely on RNAi and PTGS instead of TGS to repress GremLINE1 expression (Fig. 5). This suggests an alternative regulatory role for H3K9me besides transcriptional silencing and this transcriptional role may have an involvement in centromere identity, possibly recruiting the remaining kinetochore proteins in the absence of CENP-A and explaining Clr4 essentiality. Because H3K9me-based heterochromatin by itself has been shown to recruit the kinetochore machinery in *S. pombe* (75, 76), we propose a similar scenario in which heterochromatin-bound GremLINE1 retrotransposons could act as a beacon for the inner kinetochore complex CENP-T-W-S-X (77, 78) or a novel, alternative protein to

CENP-A. *Mucor* centrochromatin harbors every element to support this model: large, inverted repeats that trigger H3K9me and a chromodomain-containing Clr4 that can maintain heterochromatin formation even in the absence of RNAi activity. However, there is robust RNAi activity directed against GremLINE1s and though it is dispensable for maintaining H3K9me across these regions, we hypothesized that its role is to restrain the repeats within the bounds of the pericentric regions. The absence of GremLINE1 elements beyond the pericentric regions together with their exceptional, Rdrp2-dependent post-transcriptional regulation strongly supports this hypothesis. Indeed, we have shown that lack of canonical RNAi activity results in GremLINE1 retrotransposition to non-centromeric regions, with multiple implications for *Mucor* biology. First and foremost, GremLINE1s are strictly controlled by canonical Dicer and Ago RNAi activities. Second, this control is essential to preserve genome stability by preventing transposition to other genomic regions. Third, our results indicate that GremLINE1 retrotransposition may disrupt drug targets to confer antifungal drug resistance, confirming TE insertion as a source of genomic variability that favors adaptation (Fig. 5). The antagonistic interaction between canonical and alternative RNAi pathways in *Mucor* (36, 55) may fine-tune transposition and act as a fail-safe mechanism, typically ensuring proper GremLINE1 repression and genome stability but enabling transposition when beneficial to overcome adverse stressful conditions.

Materials and methods

Fungal strains, culturing, and transformation

Every strain used in this work derives from the natural isolate *Mucor lusitanicus* CBS277.49. The parental strain of all the mutant strains is the leucine and uridine double auxotroph MU402 (Ura-, Leu-). This strain was used for *clr4* deletion mutant transformation following a CRISPR-Cas9 procedure as previously described (28, 79). Briefly, we designed two guide RNAs (gRNA) to generate breaks up- and downstream of the reannotated *clr4* locus (gRNA_clr4_5': CTCCTGGTGACTGGTGAAAGTGG, and gRNA_clr4_3':

GGGTTTTCATTGGCCGTGTCTCC), and a linear DNA construct containing the selectable marker *pyrG* flanked by 1-kb upstream and downstream regions of the gRNA cleavage sites (*SI Appendix*, Table S4). Ribonucleoprotein complexes with Cas9 and gRNAs were assembled in vitro following the supplier instructions (Alt-R™ CRISPR Custom Guide RNAs and Cas9 enzyme, Integrated DNA Technologies). The DNA cassette and the ribonucleoprotein complexes were electroporated into *Mucor* protoplasts. After transformation, colonies were plated on selective minimal medium with casamino acids (MMC) at pH of 3.2 to positively select prototrophic colonies. DNA deletion allele integration and heterokaryosis was assessed by PCR using primers that generate discriminatory amplicons (*SI Appendix*, Table S4), as described previously (28). At least 10 vegetative passages were conducted in an attempt to achieve homokaryotic mutants, consisting of collecting spores from a single asexual sporangium and subsequent plating in MMC medium. Lethality of the *clr4*Δ allele was assessed by dissecting spores from a heterokaryotic mutant onto rich, non-selective YPD medium, and analyzing the *clr4* locus as described above. RNAi mutants analyzed in this study include single Dicer mutant *dcl1*Δ (strain MU406), *dcl2*Δ (MU410), and double mutant *dcl1*Δ/*dcl2*Δ (MU411) (31, 32), Argonaute mutants *ago1*Δ, *ago2*Δ, and *ago3*Δ (MU413, MU416 and MU414, respectively) (34), RdRP mutants *rdrp1*Δ, *rdrp2*Δ, and *rdrp3*Δ (MU419, MU420 and MU439, respectively) (29, 55), and the alternative ribonuclease R3B2 mutant *r3b2*Δ (MU412) (35). Media were supplemented with uridine (200 mg/L) or leucine (20 mg/L) when needed to supplement auxotrophic requirements and cultures were grown at room temperature unless otherwise stated.

Ortholog search and phylogenomic tree inference

Experimentally curated protein sequences were used as queries in a PSI-BLAST v2.12.0 against the publicly available proteomes of 85 fungal species (*SI Appendix*, Dataset S1). Then, BLASTp searches were conducted to retrieve those matches that displayed a positive reciprocal BLAST hit. Protein domains were predicted on those hits using InterProScan v5.59-91.0 and

visually examined for manual curation (*SI Appendix*, Dataset S2). BUSCO v5.4.3 identified 389 single-copy fungal orthologs from fungi_odb10 database that were present in ≥ 90 % of the species. These single-copy gene protein sequences were aligned using MAFFT v7.475, alignments trimmed by TrimAl v1.4.rev15, and used to infer a phylogenomic species tree using IQ-TREE v2.2.0.3.

Repeats and transposable element prediction

Repeated DNA sequences were predicted using RepeatModeler2 v2.0.3. Then, EMBOSS v6.6.0.0 predicted ORFs across these repeats, and InterProScan identified protein domains. Repeat sequences were manually curated to retain those encoding protein domains frequently associated with TE activity and were annotated according to presence and relative position of those domains. Both raw structural repeat and curated TE libraries were used by RepeatMasker v4.1.3 to produce annotation files.

Chromatin immunoprecipitation and RNA isolation and sequencing

2.5×10^5 spore/mL YPD cultures were grown for 16 hours at 26 ºC and 250 rpm in duplicates. The same culture was utilized to obtain ChIP and RNA samples to ensure minimal experimental differences. ChIP-grade antibodies α-H3K9me2 (ab1220, Abcam), α-H3K9me3 (ab8898, Abcam), α-RNA pol II (39497, Active Motif) were utilized to immunoprecipitate chromatin-bound DNA, following a previously-established procedure (40). Libraries were prepared using Roche KAPA HyperPrep Kits and sequenced with Illumina NovaSeq 6000 sequencing system for 100-bp paired-end reads. Total RNA was purified with a QIAGEN miRNeasy Mini Kit and RNA samples were divided into small and long RNA preparations. rRNA-depleted RNA libraries (long RNA) were prepared using Illumina Stranded Total RNA Prep with Ribo-Zero Gold rRNA Removal Kit and *M. lusitanicus* rDNA specific probes, and cDNA sequenced in a NovaSeq 6000 sequencing system for 150-bp paired-end reads. Lastly, sRNA

libraries were amplified using QIAseq miRNA library kit, and sequenced in an Illumina NextSeq 500 High-Output sequencing system to obtain 75-bp single-reads.

Sequencing data analyses

FASTQ dataset quality was assessed by FASTQC v0.11.9 and reads processed by TrimGalore! V0.6.7 to remove adapters, low quality reads and adjust minimal length for subsequent analyses. Reads were aligned to the *M. lusitanicus* MU402 genome (https://mycocosm.jgi.doe.gov/Muccir1_3/Muccir1_3.info.html) employing BWA-MEM v.0.7.17 for ChIP DNA, STAR v.2.7.10a for long RNA, and ShortStack v3.8.5 for sRNA reads. ChIP-enrichment was determined as IP/Input DNA ratio of counts per million (CPM) reads. Broad ChIP-enriched regions were identified by MACS2 v2.2.7.1. Long RNA data was quantified by featureCounts v2.0.1 from the Subread package, modifying strandness to the forward or reverse strand as needed to generate sense or antisense counts, respectively. Log_2 CPM values were determined using the limma package v3.50.3, normalizing by the trimmed mean of M-values (TMM) method. Pairwise comparisons and log_2 fold change values were determined and tested for significance (FDR test corrected by the Benjamini-Hochberg). sRNA aligned reads were deduplicated by UMI-tools v1.1.2. Reads aligning to tRNA or rRNA sequences were removed, and the remaining genome-aligning reads were quantified by ShortStack and normalized to reads per million (RPM). Coverage files for ChIP DNA and RNA alignments were generated by Deeptools2 v3.5.1.

Artificial evolution experiment and isolation of spontaneous FK506 resistant isolates

RNAi mutant and parental MU402 strains were cultured during 20 vegetative passages by collecting spores from single colonies and plating them in MMC at pH of 4.5. After 20 passages, spores were collected, counted, and adjusted to the same concentration to plate them on YPD supplemented with FK506 (1 mg/L). After 3-7 days incubation, colonies growing mycelial sectors

were selected and plated in YPD supplemented with FK506 and rapamycin (0.1 mg/L) to confirm resistance was caused by FKBP12 loss-of-function. Genomic DNA was purified using MasterPure™ Complete DNA and RNA Purification Kit (Lucigen). The whole *fkbA* locus from the FK506 resistant isolates was amplified using primers JOHE52223 and JOHE52224 and LA Taq DNA polymerase for long-range PCR with GC Buffer I (Takara). PCR conditions were optimized for long fragments as follows: 1 min 94ºC, 30 cycles: 94ºC 5s, 55ºC 30s, 72ºC 12min; and 72ºC 10 min. PCR fragments were Sanger-sequenced using primers listed in *SI Appendix*, Table S4. A primer walking strategy was designed to sequence the full-length insertions. Their full-sequence was used as query for BLASTp searches against a nucleotide database containing every TE element previously identified, ensuring alignment length covered ≥ 97 % of the query length (-qcov_hsp_perc 97), i.e., the full insertion sequence except for possible mismatches at either 5'- or 3'-end.

Quantitative PCR

Genomic DNA was purified using MasterPure™ Complete DNA and RNA Purification Kit (Lucigen). Quantitative PCR (qPCR) was performed using diluted genomic DNA from each sample, SYBR Green PCR Master Mix (Applied Biosystems™) and primers that specifically amplified a TE subfamily or the single-copy gene *vma1* (*SI Appendix*, Table S4). qPCR reactions were performed in triplicate using a QuantStudio™ 3 Real-Time PCR System. TE Ct values were normalized to *vma1* using the $2^{-\Delta CT}$ method to estimate copy numbers.

Acknowledgements and funding sources

We thank Drs. Minou Nowrousian, Patricia Peterson, and Sheng Sun for critical reading and insightful suggestions on the manuscript. Also, thanks to our laboratory manager Anna Floyd Averette for her constant support. We extend our gratitude to Drs. Rosa M. Ruiz-Vázquez, Santiago Torres-Martínez, and Francisco E. Nicolás for providing fungal strains. We appreciate

Drs. Kaustuv Sanyal, Victoriano Garre, and their teams for their input and discussion sessions on fungal epigenetics and centromere biology. Particularly, we thank Dr. Shweta Panchal for her counsel on H3K9me2 antibodies for chromatin immunoprecipitation. We are also grateful to Drs. Zanetta Chang and Silvia Calo for their recommendations on isolating FKBP12 loss-of-function insertional mutants in *M. lusitanicus*. We commend the Duke Computer Cluster for their computing equipment and Duke's Sequencing and Genomic Technologies Core Facility for their assistance, especially its Associate Director Dr. Devjanee Swain Lenz for her sage advice and expertise. Also, special thanks to Dr. Fred Dietrich for his invaluable computational resources and dedication. This study was supported by NIH/National Institute of Allergy and Infectious Diseases R01 Grant R01 AI170543 awarded to J.H. J.H. is Co-Director and Fellow of the CIFAR program Fungal Kingdom: Threats & Opportunities.

References

1. L. Tedersoo, *et al.*, Global diversity and geography of soil fungi. *Science* **346**, 1256688–10 (2014).

2. M. A. Naranjo-Ortiz, T. Gabaldón, Fungal evolution: diversity, taxonomy and phylogeny of the Fungi. *Biol. Rev.* **94**, 2101–2137 (2019).

3. M. C. Fisher, N. A. R. Gow, S. J. Gurr, Tackling emerging fungal threats to animal health, food security and ecosystem resilience. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20160332 (2016).

4. A. Rokas, Evolution of the human pathogenic lifestyle in fungi. *Nat. Microbiol.* **7**, 607–619 (2022).

5. H. C. van der Does, M. Rep, Adaptation to the host environment by plant-pathogenic fungi. *Annu. Rev. Phytopathol.* **55**, 427–450 (2017).

6. C. Martin, Y. Zhang, The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* **6**, 838–849 (2005).

7. T. Kouzarides, Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).

8. A. Buscaino, Chromatin-mediated regulation of genome plasticity in human fungal pathogens. *Genes* **10**, 855 (2019).

9. M. Freitag, Histone methylation by SET domain proteins in Fungi. *Annu. Rev. Microbiol.* **71**, 413–439 (2017).

10. T. A. Volpe, *et al.*, Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**, 1833–1837 (2002).

11. A. Verdel, RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**, 672–676 (2004).

12. K. Noma, *et al.*, RITS acts in cis to promote RNA interference–mediated transcriptional and post-transcriptional silencing. *Nat. Genet.* **36**, 1174–1180 (2004).

13. J. Nakayama, J. C. Rice, B. D. Strahl, C. D. Allis, S. I. S. Grewal, Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292**, 110–113 (2001).

14. T. Schalch, *et al.*, High-affinity binding of Chp1 chromodomain to K9 methylated histone H3 is required to establish centromeric heterochromatin. *Mol. Cell* **34**, 36–46 (2009).

15. A. J. Bannister, *et al.*, Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124 (2001).

16. H. P. Cam, *et al.*, Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nat. Genet.* **37**, 809–819 (2005).

17. K. Zhang, K. Mosch, W. Fischle, S. I. S. Grewal, Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nat. Struct. Mol. Biol.* **15**, 381–388 (2008).

18. P. A. Dumesic, *et al.*, Product binding enforces the genomic specificity of a yeast Polycomb repressive complex. *Cell* **160**, 204–218 (2015).

19. S. Catania, *et al.*, Evolutionary persistence of DNA methylation for millions of years after ancient loss of a de novo methyltransferase. *Cell* **180**, 263–277 (2020).

20. K. Jamieson, M. R. Rountree, Z. A. Lewis, J. E. Stajich, E. U. Selker, Regional control of histone H3 lysine 27 methylation in *Neurospora*. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 6027–6032 (2013).

21. L. R. Connolly, K. M. Smith, M. Freitag, The *Fusarium graminearum* histone H3 K27 methyltransferase KMT6 regulates development and expression of secondary metabolite gene clusters. *PLoS Genet.* **9**, e1003916 (2013).

22. D. E. Cook, H. M. Kramer, D. E. Torres, M. F. Seidl, B. P. H. J. Thomma, A unique chromatin profile defines adaptive genomic regions in a fungal plant pathogen. *eLife* **9**, e62208 (2020).

23. W. Zhang, J. Huang, D. E. Cook, Histone modification dynamics at H3K27 are associated with altered transcription of in planta induced genes in *Magnaporthe oryzae*. *PLoS Genet.* **17**, e1009376 (2021).

24. M. Kamei, *et al.*, IMITATION SWITCH is required for normal chromatin structure and gene repression in PRC2 target domains. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2010003118 (2021).

25. K. Voigt, *et al.*, Early-diverging fungal phyla: taxonomy, species concept, ecology,

distribution, anthropogenic impact, and novel phylogenetic proposals. *Fungal Divers.* **109**, 59–98 (2021).

26. J. F. H. Strassert, M. T. Monaghan, Phylogenomic insights into the early diversification of fungi. *Curr. Biol.* **32**, 3628-3635.e3 (2022).

27. O. A. Cornely, *et al.*, Global guideline for the diagnosis and management of mucormycosis: an initiative of the European Confederation of Medical Mycology in cooperation with the Mycoses Study Group Education and Research Consortium. *Lancet Infect. Dis.* **19**, e405–e421 (2019).

28. F. E. Nicolás, *et al.*, "Molecular tools for carotenogenesis analysis in the mucoral *Mucor circinelloides*" in *Microbial Carotenoids. Methods in Molecular Biology*, C. Barreiro, J. L. Barredo, Eds. (Humana Press, New York, NY, 2018).

29. S. Calo, F. E. Nicolás, A. Vila, S. Torres-Martínez, R. M. Ruiz-Vázquez, Two distinct RNA-dependent RNA polymerases are required for initiation and amplification of RNA silencing in the basal fungus *Mucor circinelloides*. *Mol. Microbiol.* **83**, 379–394 (2012).

30. F. E. Nicolás, *et al.*, 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 (2010).

31. F. E. Nicolás, J. P. de Haro, S. Torres-Martínez, R. M. Ruiz-Vázquez, Mutants defective in a *Mucor circinelloides* dicer-like gene are not compromised in siRNA silencing but display developmental defects. *Fungal Genet. Biol.* **44**, 504–516 (2007).

32. J. P. De Haro, *et al.*, A single dicer gene is required for efficient gene silencing associated with two classes of small antisense RNAs in *Mucor circinelloides*. *Eukaryot. Cell* **8**, 1486–1497 (2009).

33. F. E. Nicolás, S. Torres-Martínez, R. M. Ruiz-Vázquez, Two classes of small antisense RNAs in fungal RNA silencing triggered by non-integrative transgenes. *EMBO J.* **22**, 3983–3991 (2003).

34. M. Cervantes, *et al.*, A single argonaute gene participates in exogenous and endogenous RNAi and controls cellular functions in the basal fungus *Mucor circinelloides*. *PLoS One* **8**, e69283 (2013).

35. T. A. Trieu, *et al.*, A non-canonical RNA silencing pathway promotes mRNA degradation in basal fungi. *PLoS Genet.* **11**, e1005168 (2015).

36. C. Pérez-Arques, *et al.*, A non-canonical RNAi pathway controls virulence and genome stability in Mucorales. *PLoS Genet.* **16**, e1008611 (2020).

37. S. Calo, *et al.*, Antifungal drug resistance evoked via RNAi-dependent epimutations. *Nature* **513**, 555–558 (2014).

38. Z. Chang, J. Heitman, Drug-resistant epimutants exhibit organ-specific stability and induction during murine infections caused by the human fungal pathogen *Mucor circinelloides*. *mBio* **10**, e02579-19 (2019).

39. F. E. Nicolás, *et al.*, The RNAi machinery controls distinct responses to environmental signals in the basal fungus *Mucor circinelloides*. *BMC Genomics* **16**, 237 (2015).

40. M. I. Navarro-Mendoza, *et al.*, Early diverging fungus *Mucor circinelloides* lacks centromeric histone CENP-A and displays a mosaic of point and regional centromeres. *Curr. Biol.* **29**, 3791-3802.e6 (2019).

41. C. Pérez-Arques, *et al.*, The RNAi mechanism regulates a new exonuclease gene involved in the virulence of Mucorales. *Int. J. Mol. Sci.* **22**, 2282 (2021).

42. Z. Chang, R. B. Billmyre, S. C. Lee, J. Heitman, Broad antifungal resistance mediated by

RNAi-dependent epimutation in the basal human fungal pathogen Mucor circinelloides. *PLoS Genet.* **15**, e1007957 (2019).

43. H. M. Kramer, M. F. Seidl, B. P. H. J. Thomma, D. E. Cook, Local rather than global H3K27me3 dynamics are associated with differential gene expression in *Verticillium dahliae*. *mBio* **13**, e03566-21 (2022).

44. S. Torres-Garcia, *et al.*, Epigenetic gene silencing by heterochromatin primes fungal resistance. *Nature* **585**, 453–458 (2020).

45. N. Robbins, M. D. Leach, L. E. Cowen, Lysine deacetylases Hda1 and Rpd3 regulate Hsp90 function thereby governing fungal drug resistance. *Cell Rep.* **2**, 878–888 (2012).

46. D. Zilberman, X. Cao, S. E. Jacobsen, ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* **299**, 716–719 (2003).

47. Y. Liu, K. Mochizuki, M. A. Gorovsky, Histone H3 lysine 9 methylation is required for DNA elimination in developing macronuclei in *Tetrahymena*. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 1679–1684 (2004).

48. M. Pal-Bhadra, *et al.*, Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. *Science* **303**, 669–672 (2004).

49. I. A. Drinnenberg, *et al.*, RNAi in budding yeast. *Science* **326**, 544–550 (2009).

50. J. D. Laurie, R. Linning, G. Bakkeren, Hallmarks of RNA silencing are found in the smut fungus Ustilago hordei but not in its close relative *Ustilago maydis*. *Curr. Genet.* **53**, 49–58 (2008).

51. G. Wu, *et al.*, Genus-wide comparative genomics of *Malassezia* delineates its phylogeny, physiology, and niche adaptation on human skin. *PLoS Genet.* **11**, e1005614 (2015).

52. A. J. Bewick, *et al.*, Diversity of cytosine methylation across the fungal tree of life. *Nat. Ecol. Evol.* **3**, 479–490 (2019).

53. N. V Fedoroff, Transposable elements, epigenetics, and genome evolution. *Science* **338**, 758–67 (2012).

54. M. R. Motamedi, *et al.*, Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* **119**, 789–802 (2004).

55. S. Calo, *et al.*, A non-canonical RNA degradation pathway suppresses RNAi-dependent epimutations in the human fungal pathogen Mucor circinelloides. *PLoS Genet.* **13**, e1006686 (2017).

56. Q. Feng, J. V Moran, H. H. Kazazian, J. D. Boeke, Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87**, 905–916 (1996).

57. R. J. Price, E. Weindling, J. Berman, A. Buscaino, Chromatin profiling of the repetitive and nonrepetitive genomes of the human fungal pathogen *Candida albicans*. *mBio* **10**, e01376-19 (2019).

58. J. T. Kadonaga, The transformation of the DNA template in RNA polymerase II transcription: a historical perspective. *Nat. Struct. Mol. Biol.* **26**, 766–770 (2019).

59. K. Struhl, Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. *Cell* **98**, 1–4 (1999).

60. A. Shirai, *et al.*, Impact of nucleic acid and methylated H3K9 binding activities of Suv39h1 on its heterochromatin assembly. *eLife* **6**, e25317 (2017).

61. V. Krauss, A. Fassl, P. Fiebig, I. Patties, H. Sass, The evolution of the histone methyltransferase gene Su(var)3-9 in metazoans includes a fusion with and a re-fission from a functionally unrelated gene. *BMC Evol. Biol.* **6**, 18 (2006).

62. L. O. Baumbusch, *et al.*, The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic Acids Res.* **29**, 4319–4333 (2001).

63. C. S. Veerappan, Z. Avramova, E. N. Moriyama, Evolution of SET-domain protein families in the unicellular and multicellular Ascomycota fungi. *BMC Evol. Biol.* **8**, 190 (2008).

64. J. C. Rice, *et al.*, Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol. Cell* **12**, 1591–1598 (2003).

65. B. D. Towbin, *et al.*, Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell* **150**, 934–947 (2012).

66. K. Naumann, *et al.*, Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in *Arabidopsis*. *EMBO J.* **24**, 1418–1429 (2005).

67. G. Jih, *et al.*, Unique roles for histone H3K9me states in RNAi and heritable silencing of transcription. *Nature* **547**, 463–467 (2017).

68. W. Dong, *et al.*, Abo1 is required for the H3K9me2 to H3K9me3 transition in heterochromatin. *Sci. Rep.* **10**, 6055 (2020).

69. R. Castanera, *et al.*, Transposable elements versus the fungal genome: impact on whole-genome architecture and transcriptional profiles. *PLoS Genet.* **12**, e1006108 (2016).

70. M. Freitag, R. L. Williams, G. O. Kothe, E. U. Selker, A cytosine methyltransferase homologue is essential for repeat-induced point mutation in *Neurospora crassa*. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 8802–8807 (2002).

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

72. T. Sugiyama, H. Cam, A. Verdel, D. Moazed, S. I. S. Grewal, RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 152–157 (2005).

73. A. Chicas, RNAi-dependent and RNAi-independent mechanisms contribute to the silencing of RIPed sequences in *Neurospora crassa*. *Nucleic Acids Res.* **32**, 4237–4243 (2004).

74. Z. A. Lewis, *et al.*, Relics of repeat-induced point mutation direct heterochromatin formation in *Neurospora crassa*. *Genome Res.* **19**, 427–437 (2009).

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

76. A. Kagansky, *et al.*, Synthetic heterochromatin bypasses RNAi and centromeric repeats to establish functional centromeres. *Science* **324**, 1716–1719 (2009).

77. N. Cortes-Silva, *et al.*, CenH3-independent kinetochore assembly in Lepidoptera requires CCAN, including CENP-T. *Curr. Biol.* **30**, 561-572.e10 (2020).

78. Y. Takenoshita, M. Hara, T. Fukagawa, Recruitment of two Ndc80 complexes via the CENP-T pathway is sufficient for kinetochore functions. *Nat. Commun.* **13**, 851 (2022).

79. C. Lax, *et al.*, Stable and reproducible homologous recombination enables CRISPR-based engineering in the fungus *Rhizopus microsporus*. *Cell Reports Methods* **1**, 100124 (2021).

Figure legends

Fig 1. Distribution of RNA interference components and epigenetic regulatory methyltransferases in Fungi. The matrix displays presence (blue) or absence (gray) of different epigenetic regulatory complexes and their corresponding proteins (left) across the fungal kingdom (top). Predicted N-terminal chromodomain in Clr4 methyltransferase is depicted (black dot). Methyltransferases involved in either transcriptional activation (green) or repression (orange), as well as RNAi components involved in post- or co-transcriptional silencing (PTGS or CTGS respectively, also orange) are shown. The fungal phylogenomic tree contains 43 species selected as representatives from each major color-coded phylum and a robust sampling of Mucoromycota and other fungal pathogens, showing evolutionary distance as substitutions per amino acid (solid black lines).

Fig 2. H3K9me heterochromatin and RNAi are coincident at transcriptionally silent, repeated sequences. (A) H3K9me2 (light purple) and -me3 (dark purple) and RNA polymerase II (RNA pol II, gold) enriched distribution is shown across scaffold 9. Enrichment coverage is computed as the ratio of immunoprecipitated (IP) DNA/Input DNA reads, previously normalized as $\log_2$ counts per million (CPM). Transcripts (mRNA, green) and sRNA (red) coverages correspond to $\log_2$ CPM. Genes (light red) and repeated sequences (light blue) including pericentric regions (bright blue) are displayed, as well as GC content coverage as a percentage. **(B)** Dot plot and Pearson's correlation coefficient (R, dark line) of H3K9me2 (x-axis) and -me3 (y-axis) average $\log_2$ CPM enrichment values are plotted as 10-kb regions representing the whole genome. Overlapping dots are color-coded as a density scale. **(C)** Proportion of gene, repeated, and intergenic sequences are shown color-coded as percentage of total base pairs in the genome (inner pie chart). The proportion of H3K9-methylated base pairs in the same categories is highlighted (darker shade, outer pie chart) from the remaining non-methylated base pairs (lighter shade, outer pie chart). **(D)** The percentage of small RNA aligned reads according to length and

5'-end nucleotide distribution (color-coded) is depicted. **(E)** Schematic view of the six most abundant TE families identified in the genome, showing total length (nt) of the elements (solid black lines), open reading frames (gray rectangular arrows), and predicted protein domains (color-coded rectangles) abbreviated as follows: Zinc finger, CCHC-type (ZnF-CCHC); Endonuclease/exonuclease/phosphatase superfamily (APE); Reverse transcriptase domain (RT); Reverse transcriptase zinc-binding domain (RT-ZnB); Retrotransposon GAG domain (GAG); Aspartic peptidase domain superfamily (PR); Integrase zinc-binding domain and catalytic core (INT); Chromodomain (CD); Homeobox-like domain superfamily (HOX); Transposase, Tc1-like (Tc1-Tpn) and DDE domain (DDE); Harbinger transposase-derived nuclease domain (Harbinger-Tpn); Helitron helicase-like domain (Helitron); DNA helicase Pif1-like (Pif1). **(F)** Genomic distribution of the six most abundant TE families shown as color-coded rectangular arrows to indicate transcription direction (element length is scaled to the rectangle width, not including the arrow point). Centromeric (*CEN*) and telomeric (*) regions are designated in the plot. The first eleven out of twenty-four scaffolds (> 94% of the genome) containing nine centromeres and their length (Mb) are shown for reference.

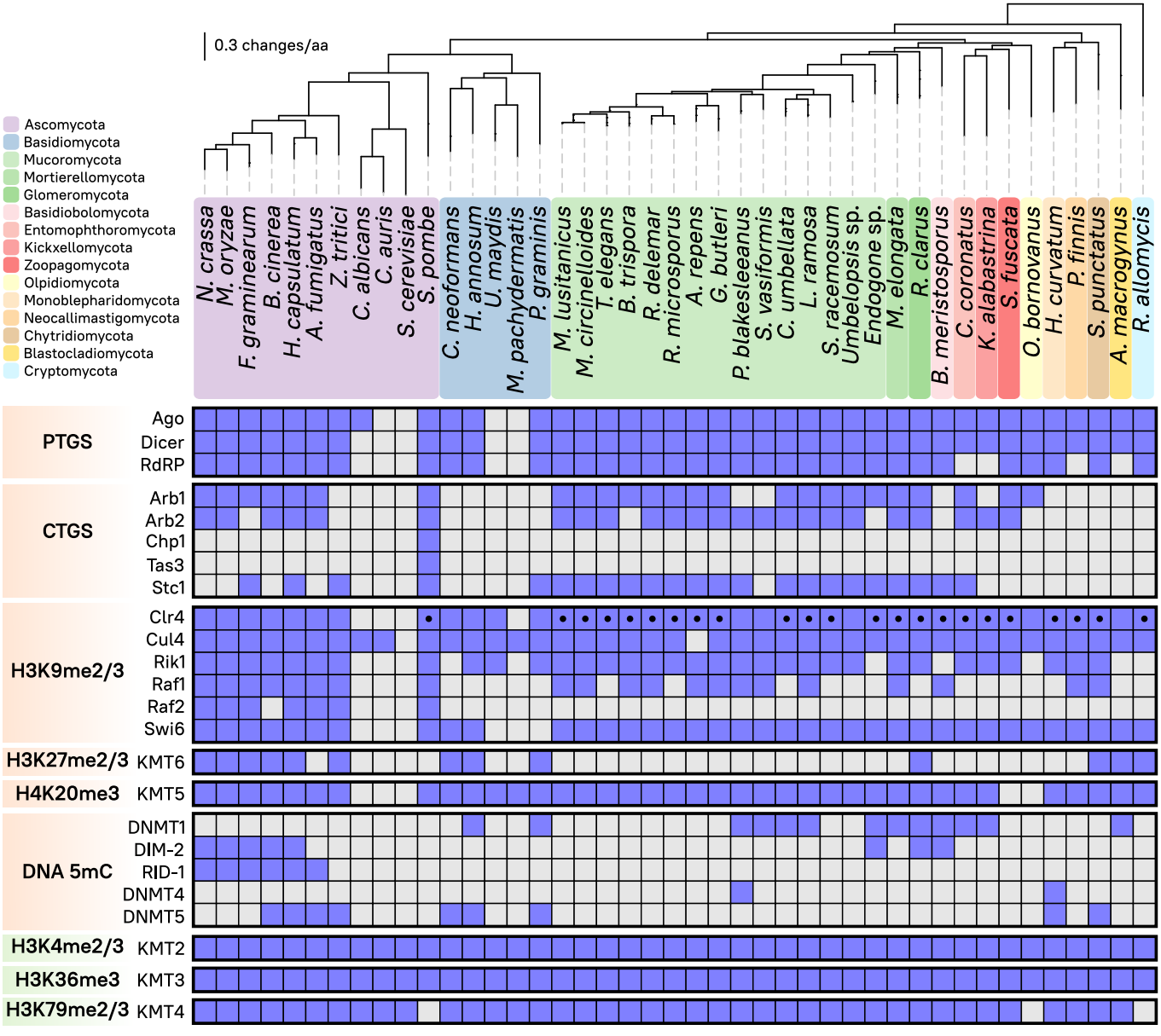
Fig 3. Dual epigenetic regulation –H3K9me-based heterochromatin and RNAi– targets RNA transposons. All copies of the GremLINE1 (n=58), other LINE (n=79), and Gypsy (n=92) retrotransposons identified in the genome are shown. TE copies from the same family are clustered according to their subfamily, H3K9me3, sRNA, and mRNA abundance in the wild-type strain. Heatmaps depicting H3K9me3, sRNA, mRNA, and RNA polymerase II abundance in wild-type and RNAi deficient strains are plotted (see plot name above). Each row represents a single TE copy, which are aligned across all heatmaps to compare their values that read from left to right as follows. First, a clustered dendrogram indicates each element subfamily (color-coded). Second, H3K9me3 enrichment values are shown as the immunoprecipitation (IP)/input ratio of normalized $\log_2$ CPM (CPM, yellow-to-red legend) from start to end of each copy (divided in 250

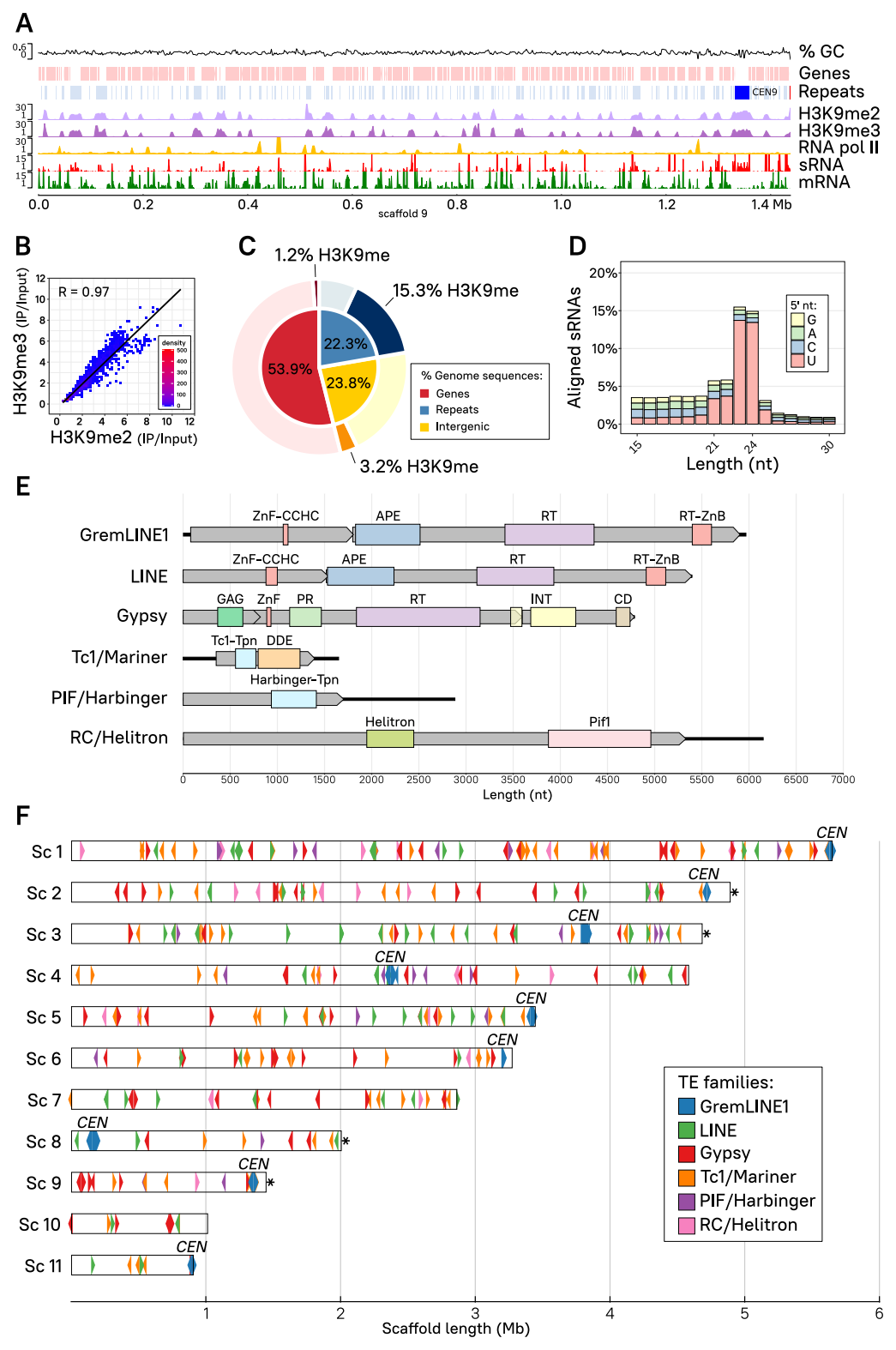
bins) and 50 kb upstream and downstream; as well as the mean value of all of the copies in each family shown in a profile plot (on top, bold red line shows the average and faded red background shows the standard deviation). Third, sRNA values normalized to $\log_2$ reads per million (RPM, yellow-to-red legend) are presented. Fourth, sense RNA values normalized to $\log_2$ CPM (yellow-to-blue legend) are shown. Lastly, RNA pol II occupancy values are depicted as IP/Input ratio of normalized $\log_2$ CPM (yellow-to-red legend).

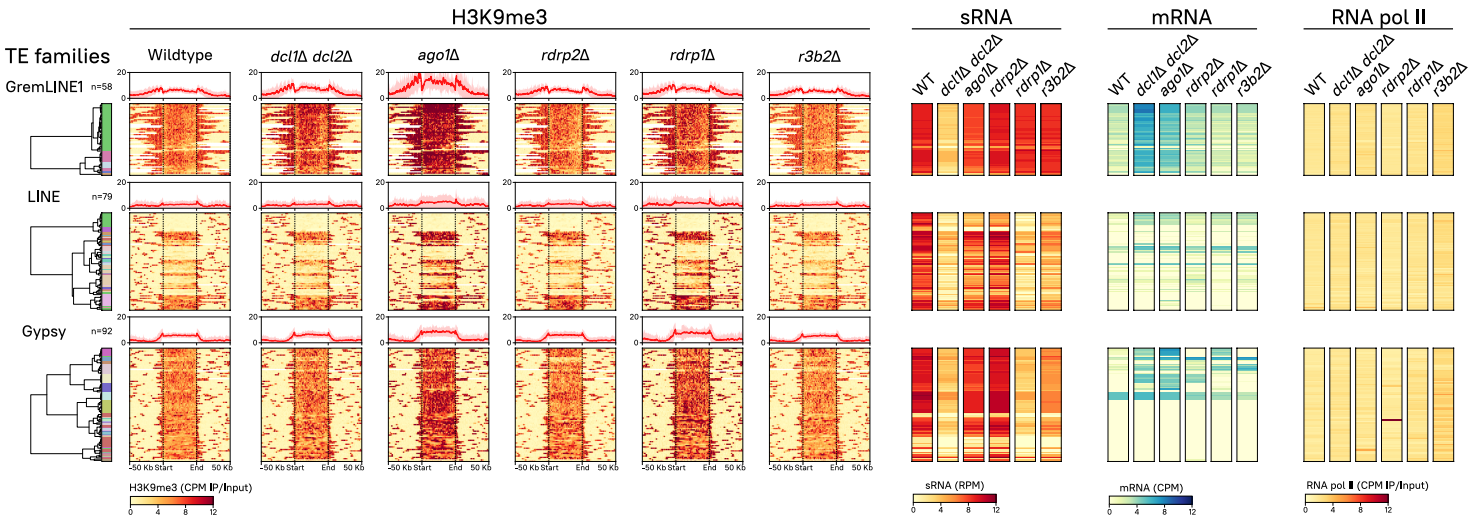
Fig 4. H3K9me defines pericentric boundaries while canonical RNAi inhibits GremLINE1 transposition to non-centromeric regions. (A) Genomic plot of *CEN4* and its pericentric regions displaying H3K9me3 enrichment (purple), and sRNA (red) and transcript levels (mRNA, green) in a wild-type (WT) and canonical RNAi-deficient strains (double *dcl1*Δ *dcl2*Δ and *ago1*Δ). Read coverage is normalized to $\log_2$ CPM, and ChIP enrichment is shown as IP/Input DNA ratio. GC proportion (black line) is computed across the whole sequence. Genes (light red), repeated sequences (structural repeats in light blue and TE in turquoise), as well as the centromeric sequence (bright blue) are shown for reference. **(B)** Heatmap of significant antisense transcription differences across GremLINE1 copies (FDR ≤ 0.05) in the wild-type strain compared to either *rdrp1*Δ or *rdrp2*Δ (color-coded values shown as $\log_2$ fold change of CPM). GremLINE1 subfamilies (color-coded) and identification are shown. **(C)** Artificial evolution experiment in RNAi mutants. Briefly, the strains were cultured through 20 vegetative growth passages, and analyzed for GremLINE1 copy numbers **(D)** or FK506 spontaneous resistance **(E)**. **(D)** GremLINE1 copies in different RNAi mutants after 20 vegetative passages are shown. Copy numbers are estimated as the relative fold ($2^{-\Delta CT}$) between GremLINE1 (multiple copies) and *vma1* (single copy) sequence amplification in triplicate quantitative PCR reactions. Significantly distinct values are indicated with different letters and were calculated by one-way ANOVA and Tukey's HSD tests. **(E)** GremLINE1 transposition into the *fkbA* locus was positively selected on FK506-containing medium. Each insertion was named after their genetic background, the TE inserted, and a systematic number;

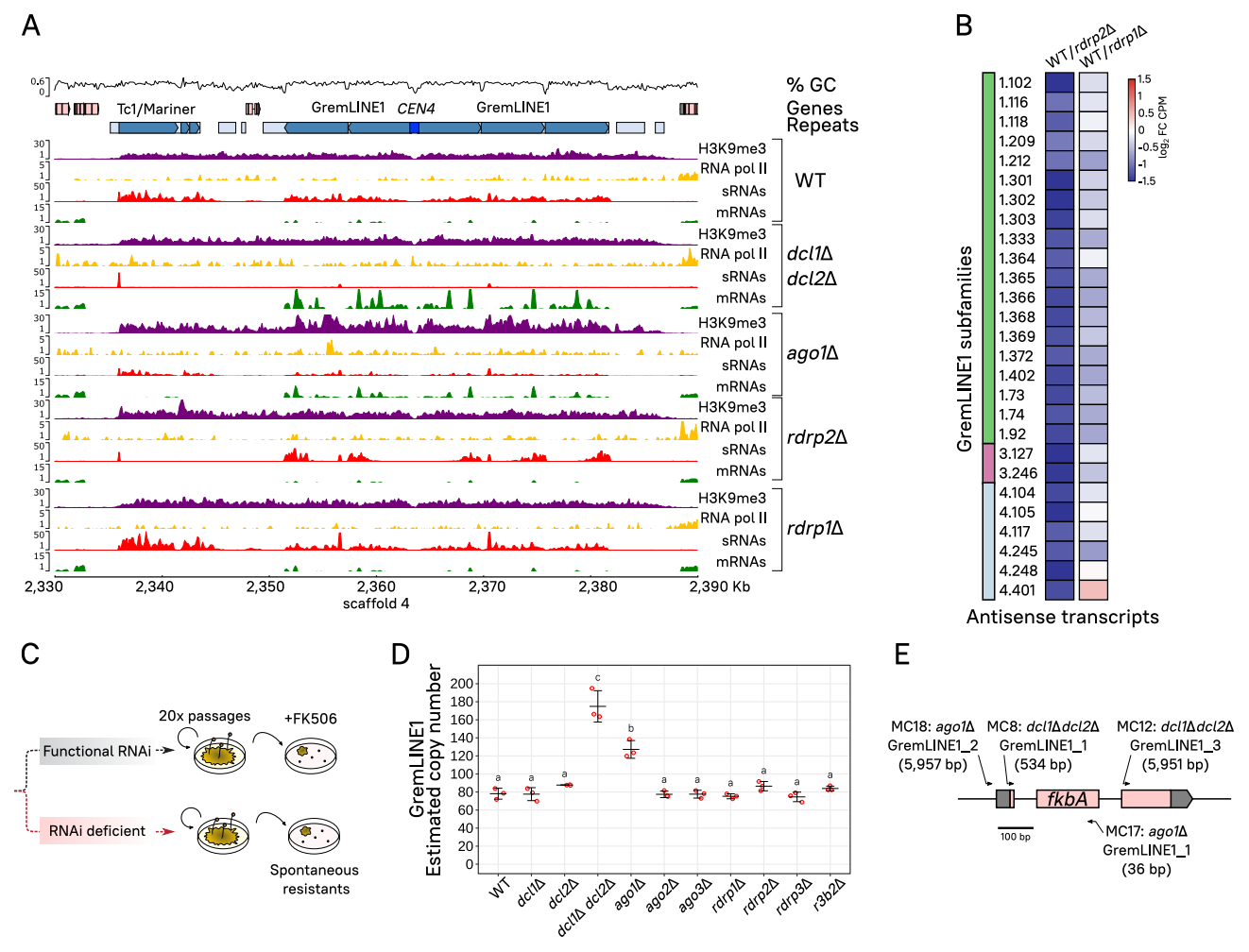
insertion length (in bp) is also shown after the name. Insertion sites are indicated by arrows, whose direction follows the element orientation with respect to *fkbA*, and whose width reflects the scaled length of the target site duplication marking the insertion site.

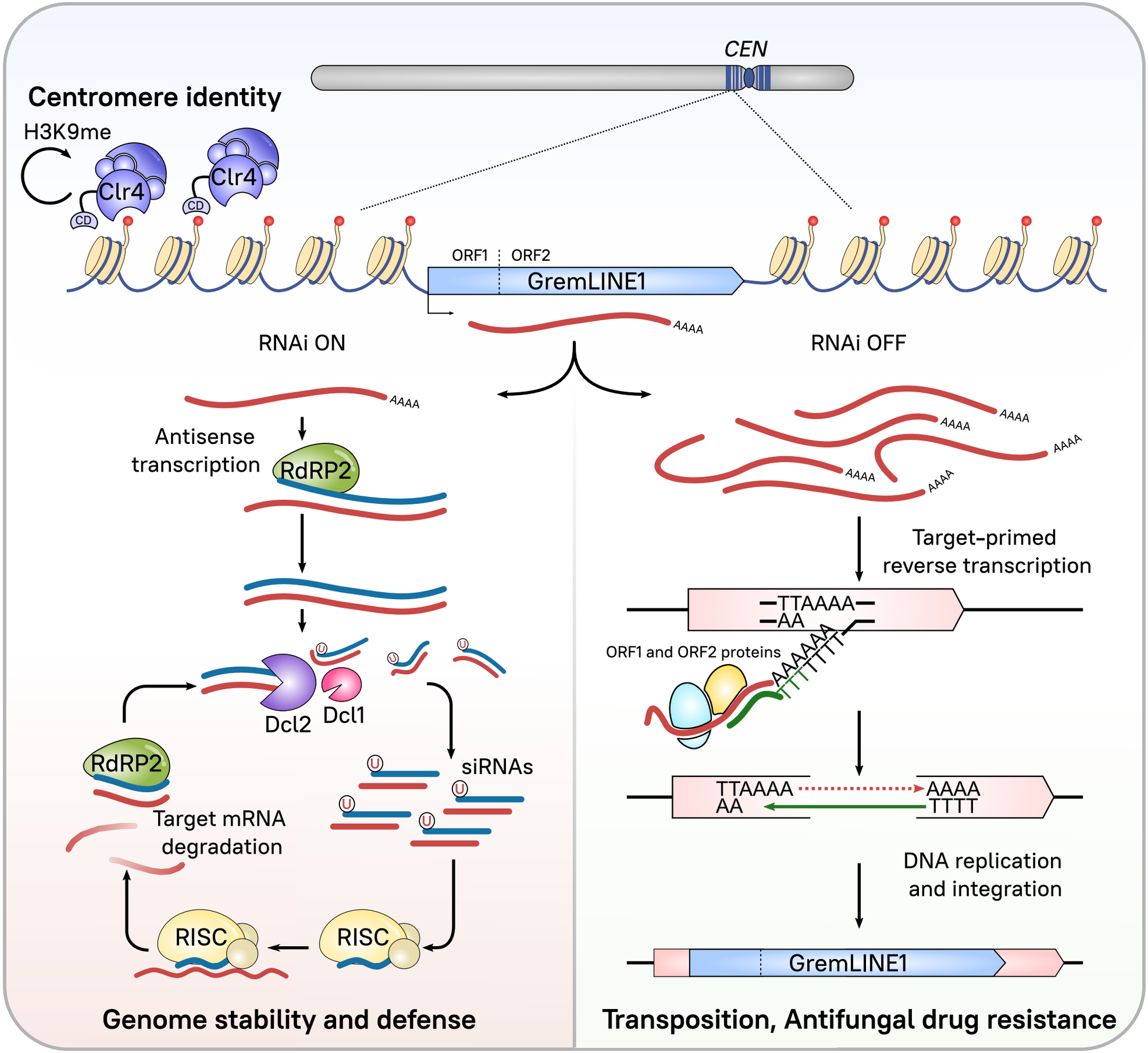
Fig. 5. Current model for centromeric GremLINE1 dual regulation by H3K9me-based heterochromatin and RNAi. GremLINE1 copies are retrotransposons exclusively located in the pericentric regions of Mucoromycota species, forming long H3K9me heterochromatic islands that define centromere identity. Clr4 methyltransferase catalyzes H3K9me, and its chromodomain allows for H3K9me reading activity that could be involved in maintaining and spreading the mark in an RNAi-independent process. GremLINE1 elements display two ORFs encoding the enzymes necessary for their transposition, and their expression is usually post-transcriptionally repressed by RNAi in a process that involves Rdrp2 antisense transcription, Dcl2 nuclease activity that can be partially rescued by Dcl1, and RISC targeted degradation of GremLINE1 transcripts. In addition, Rdrp2 activity could reamplify the silencing signal by generating dsRNA from these degraded transcript fragments. This PTGS mechanism ensures genome stability by preventing GremLINE1 transposition and consequently, transposition is increased in RNAi-deficient strains. Lack of RNAi leads to increased GremLINE1 expression and consequently, to target-primed reverse transcription integration events in non-centromeric regions. These transposition events have been shown to affect drug targets, promoting drug resistance.


0.3 changes/aa
Ascomycota
Basidiomycota
Mucoromycota
Mortierellomycota
Glomeromycota
Basidiobolomycota
Entomophthoromycota
Kickxellomycota
Zoopagomycota
Olpidiomycota
Monoblepharidomycota
Neocallimastigomycota
Chytridiomycota
Blastocladiomycota
Cryptomycota
N. crassa
M. oryzae
F. graminearum
B. cinerea
H. capsulatum
A. fumigatus
Z. tritici
C. albicans
C. auris
S. cerevisiae
S. pombe
C. neoformans
H. annosum
U. maydis
M. pachydermatis
P. graminis
M. lusitanicus
M. circinelloides
T. elegans
B. trispora
R. delemar
R. microsporus
A. repens
G. butleri
P. blakesleeanus
S. vasiformis
C. umbellata
L. ramosa
S. racemosum
Umbelopsis sp.
Endogone sp.
M. elongata
R. clarus
B. meristosporus
C. coronatus
K. alabastrina
S. fuscata
O. bornovanus
H. curvatum
P. finnis
S. punctatus
A. macrogynus
R. allomycis
PTGS
Ago
Dicer
RdRP
CTGS
Arb1
Arb2
Chp1
Tas3
Stc1
H3K9me2/3
Clr4
Cul4
Rik1
Raf1
Raf2
Swi6
H3K27me2/3
KMT6
H4K20me3
KMT5
DNA 5mC
DNMT1
DIM-2
RID-1
DNMT4
DNMT5
H3K4me2/3
KMT2
H3K36me3
KMT3
H3K79me2/3
KMT4


A
% GC
Genes
Repeats
CEN9
H3K9me2
H3K9me3
RNA pol II
sRNA
mRNA
scaffold 9
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4 Mb
B
R = 0.97
density
H3K9me3 (IP/Input)
H3K9me2 (IP/Input)
C
1.2% H3K9me
15.3% H3K9me
22.3%
53.9%
23.8%
3.2% H3K9me
% Genome sequences:
Genes
Repeats
Intergenic
D
Aligned sRNAs
20%
15%
10%
5%
0%
5' nt:
G
A
C
U
Length (nt)
E
GremLINE1
ZnF-CCHC
APE
RT
RT-ZnB
LINE
ZnF-CCHC
APE
RT
RT-ZnB
Gypsy
GAG
ZnF
PR
RT
INT
CD
Tc1/Mariner
Tc1-Tpn
DDE
PIF/Harbinger
Harbinger-Tpn
RC/Helitron
Helitron
Pif1
Length (nt)
F
Sc 1
Sc 2
Sc 3
Sc 4
Sc 5
Sc 6
Sc 7
Sc 8
Sc 9
Sc 10
Sc 11
CEN
TE families:
GremLINE1
LINE
Gypsy
Tc1/Mariner
PIF/Harbinger
RC/Helitron
Scaffold length (Mb)


H3K9me3
sRNA
mRNA
RNA pol II
TE families
Wildtype
dcl1Δ dcl2Δ
ago1Δ
rdrp2Δ
rdrp1Δ
r3b2Δ
GremLINE1 n=58
LINE n=79
Gypsy n=92
-50 Kb Start End 50 Kb
H3K9me3 (CPM IP/Input)
sRNA (RPM)
mRNA (CPM)
RNA pol II (CPM IP/Input)


A
% GC
Genes
Repeats
Tc1/Mariner
GremLINE1
CEN4
GremLINE1
H3K9me3
RNA pol II
sRNAs
mRNAs
WT
dcl1Δ
dcl2Δ
ago1Δ
rdrp2Δ
rdrp1Δ
2,330
2,340
2,350
2,360
2,370
2,380
2,390 Kb
scaffold 4
B
WT/rdrp2Δ
WT/rdrp1Δ
GremLINE1 subfamilies
1.102
1.116
1.118
1.209
1.212
1.301
1.302
1.303
1.333
1.364
1.365
1.366
1.368
1.369
1.372
1.402
1.73
1.74
1.92
3.127
3.246
4.104
4.105
4.117
4.245
4.248
4.401
log2 FC CPM
Antisense transcripts
C
20x passages
+FK506
Functional RNAi
RNAi deficient
Spontaneous resistants
D
GremLINE1 Estimated copy number
WT
dcl1Δ
dcl2Δ
dcl1Δ dcl2Δ
ago1Δ
ago2Δ
ago3Δ
rdrp1Δ
rdrp2Δ
rdrp3Δ
r3b2Δ
E
MC18: ago1Δ GremLINE1_2 (5,957 bp)
MC8: dcl1Δdcl2Δ GremLINE1_1 (534 bp)
MC12: dcl1Δdcl2Δ GremLINE1_3 (5,951 bp)
fkbA
100 bp
MC17: ago1Δ GremLINE1_1 (36 bp)


CEN
Centromere identity
H3K9me
Clr4
CD
Clr4
CD
ORF1
ORF2
GremLINE1
AAAA
RNAi ON
RNAi OFF
AAAA
Antisense transcription
RdRP2
Dcl2
Dcl1
U
siRNAs
RdRP2
Target mRNA degradation
RISC
RISC
Genome stability and defense
AAAA
AAAA
AAAA
AAAA
AAAA
Target-primed reverse transcription
TTAAAA
AA
ORF1 and ORF2 proteins
AAAAAA
TTTTTT
TTAAAA
AA
AAAA
TTTT
DNA replication and integration
GremLINE1
Transposition, Antifungal drug resistance