Distinct evolutionary trajectories following loss of RNA interference in *Cryptococcus neoformans*

Jun Huang[1], Connor J. Larmore[1], Shelby J. Priest[1], Ziyan Xu[1], Fred S. Dietrich[1], Vikas Yadav[1], Paul M. Magwene[2], Sheng Sun[1], and Joseph Heitman[1]*

[1]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA
[2]Department of Biology, Duke University, Durham, NC 27710, USA.

* Corresponding author: Joseph Heitman
Email: heitm001@duke.edu

Author Contributions: J. Huang., S.J.P., V.Y., and J. Heitman designed research; J. Huang., C.J.L., S.J.P., Z.X., and S.S. performed research; J. Huang., P.M.M., C.J.L., and V.Y. analyzed data; P.M.M. and F.S.D. contributed new analytic pipelines; S.S. and J. Heitman supervised the study; J. Huang., C.J.L., S.J.P., V.Y., P.M.M., S.S., and J. Heitman wrote the paper.

Competing Interest Statement: The authors declare no competing interest.

Classification: Biological Sciences, Genetics
Keywords: Antifungal drug resistance, Hypermutation, KDZ transposon.

This PDF file includes:
Main Text
Figures 1 to 6
Tables 1 to 2

Abstract

While increased mutation rates typically have negative consequences in multicellular organisms, hypermutation can be advantageous for microbes adapting to the environment. Previously, we identified two hypermutator *Cryptococcus neoformans* clinical isolates that rapidly develop drug resistance due to transposition of a retrotransposon, Cnl1. Cnl1-mediated hypermutation is caused by a nonsense mutation in the gene encoding a novel RNAi component, Znf3, combined with a tremendous transposon burden. To elucidate adaptative mechanisms following RNAi loss, two bioinformatic pipelines were developed to identify RNAi loss-of-function mutations in a collection of 387 sequenced *C. neoformans* isolates. Remarkably, several RNAi-loss isolates were identified that are not hypermutators and have not accumulated transposons. To test if these RNAi loss-of-function mutations can cause hypermutation, the mutations were introduced into a non-hypermutator strain with a high transposon burden, which resulted in a hypermutator phenotype. To further investigate if RNAi-loss isolates can become hypermutators, *in vitro* passaging was performed. Although no hypermutators were found in two *C. neoformans* RNAi-loss strains after short-term passage, hypermutation was observed in a passaged *C. deneoformans* strain with increased transposon burden. Additionally, when an RNAi-loss isolate was crossed with an isolate containing a high Cnl1 burden, F1 hypermutator progeny were identified with distinct mutational spectra. In addition to Cnl1 transpositions, insertions of a novel gigantic DNA transposon KDZ1 (~11 kb), contributed to hypermutation in the progeny. Our results suggest that RNAi loss is relatively common (7/387, ~1.8%) and enables distinct evolutionary trajectories: hypermutation following transposon accumulation or survival without hypermutation.

Significance Statement

There is a dearth of antifungal drugs available to treat *Cryptococcus neoformans*, a human fungal pathogen of global impact. Resistance to current antifungal therapies has been observed. We previously identified natural hypermutators with a loss-of-function mutation in the RNAi machinery and transposon expansion. Here, we identified several novel natural isolates with RNAi defects, none of which display a hypermutator phenotype or have undergone transposon expansion. Furthermore, we demonstrate that these isolates can lie on a pathway to hypermutation following introduction of a transposon burden. In addition, a novel DNA transposon class was discovered that contributes to antifungal drug resistance. These findings highlight the importance of transposons in driving rapid adaptation in the absence of RNAi and reveal distinct evolutionary trajectories following RNAi loss, a relatively common event in *C. neoformans*.

Main Text

Introduction

Stochastic mutations and genomic rearrangements provide genetic variation enabling a microbial population to survive and proliferate in changing environments, driving rapid evolution. However, uncontrolled mutation and rearrangement can be deleterious, leading to gene malfunction and disorganized genome structure. To safeguard the genome, eukaryotic microbes have therefore evolved multiple mechanisms promoting genome stability. Relative to wildtype, microbes with defects in DNA repair, mitochondrial activity, or repressive chromatin modifications have been associated with higher mutation rates, which can lead to distinct adaptability in response to environmental stress (1–6).

RNA interference (RNAi) also serves to promote genome stability (7, 8). The RNAi machinery was first discovered in the nematode *Caenorhabditis elegans*, in which injection of double-stranded RNA led to specific silencing of the corresponding gene (9). This mechanism plays a critical role in modulating gene expression, defending against viral infection, and maintaining genome integrity (8, 10, 11). As such, RNAi is involved in the formation of heterochromatin, a condensed chromatin structure involved in gene silencing (12). Together with heterochromatin and other cellular mechanisms, RNAi suppresses transposable elements through post-transcriptional or transcriptional silencing within the genome (12, 13).

The RNAi pathway is conserved across a wide range of eukaryotic organisms, but has been independently lost in several fungal species, including the model yeast *Saccharomyces cerevisiae*, the corn smut pathogen *Ustilago maydis*, and the skin commensal/pathogen *Malassezia* species (14–17). As a consequence, several mycoviruses have been found in these RNAi-loss species, which provide benefits by, in some cases, producing toxins (18–20). In addition, shortened centromere length has been correlated with the loss of full-length retrotransposons in the RNAi-deficient *Cryptococcus deuterogattii* species compared to other pathogenic *Cryptococcus* species, highlighting the importance of RNAi in controlling transposons and maintaining genome integrity (21). RNAi was previously thought to be non-functional in the human

fungal pathogen *Candida albicans*. However, a recent study found that RNAi is intact in the vast majority of clinical isolates, emphasizing the limitations of utilizing a single reference strain (22).

Cryptococcus neoformans, a significant human fungal pathogen in the *Cryptococcus* pathogenic species complex, causes the majority of cryptococcal meningoencephalitis and poses a significant challenge to global human health (23, 24). Population analysis has revealed four distinct *C. neoformans* lineages: VNI, VNII, VNBI, and VNBII (25). Cryptococcal meningoencephalitis is associated with ~112,000 annual deaths and accounts for ~19% of all HIV/AIDS-related mortality worldwide (26, 27). Accordingly, *C. neoformans* was classified as a critical priority pathogen in the fungal priority pathogens list published by the World Health Organization (WHO) in 2022 (28, 29).

Unfortunately, options for treating cryptococcal infections are limited. The combination of amphotericin B and 5-fluorocytosine (5-FC) remains the gold standard treatment strategy with improved survival, albeit with side effects (30, 31). However, 5-FC monotherapy is untenable due to rapid development of resistance (32). Similarly, *C. neoformans* resistance to azoles (e.g., fluconazole) that are commonly used for maintenance therapies frequently develops via aneuploidy or mutations (33, 34).

Several natural and laboratory hypermutator isolates have been identified in the *Cryptococcus* pathogenic species complex, with increased mutation rates mediated by a variety of mechanisms. *C. neoformans* and *C. deuterogattii* isolates with mutations in components of the DNA mismatch repair pathway (e.g., *MSH2*) showed elevated mutation rates in homopolymeric tracts, following selection for antifungal drug resistance as well as rapid phenotypic switching (4, 35, 36). Additionally, disruption or mutation in *C. neoformans* of Rad51, a core component of homologous recombination, and in *C. deneoformans* of DNA polymerase delta subunit (*POL3*) are associated with an elevated mutation rate. (2, 37, 38). Furthermore, transposition of the T1 DNA transposon and the TCN12 retrotransposon into the 5-Fluoroorotic Acid (5-FOA) drug targets (*URA3* and *URA5*) underlies the vast majority of 5-FOA resistant colonies isolated from mice infected with the *C. deneoformans* strain XL280α (39).

Recently, we identified two natural hypermutators, Bt65 and Bt81 (40). These strains display a significantly higher mutation rate compared to H99, the type strain for *C. neoformans*, when selected for resistance to a combination of the antifungal drugs rapamycin and FK506 (40). Interestingly, the frequent movement of a non-LTR retrotransposon (i.e., Cnl1) into the gene encoding the FK506/rapamycin target FKBP12 is responsible for most of the resistance. Both hypermutator isolates harbor a nonsense mutation in the gene encoding a novel RNAi component, Znf3, and have accumulated a tremendous Cnl1 transposon burden (40). Further evidence supports the model that the loss of RNAi and high Cnl1 burden are together responsible for hypermutation and antifungal drug resistance (40, 41).

In this study, we sought to determine if other natural *C. neoformans* isolates have lost RNAi and display a similar hypermutator phenotype. Taking advantage of publicly available whole-genome sequencing data for 387 *C. neoformans* natural isolates, we successfully identified five novel RNAi-deficient strains from clinical and environmental sources. Surprisingly, unlike Bt65 and Bt81, none of these RNAi-loss strains display a hypermutator phenotype or contain a Cnl1 transposon burden. We demonstrated that these novel RNAi-deficient strains may lie on a pathway to hypermutation following transposon accumulation. Notably, we discovered several novel gigantic DNA transposons from the KDZ (Kyakuja, Dileera, and Zisupton) class with diverse regulatory mechanisms, contributing to the hypermutator phenotype of RNAi-deficient strains in *C. neoformans* and *C. deneoformans*. Taken together, our results suggest natural isolates with RNAi defects are relatively common and that they may exist at an evolutionary fork in the road, posed to either evolve into hypermutators as transposons accumulate, or persist without such changes despite the loss of RNAi.

Results

RNAi loss-of-function mutations occur frequently in *C. neoformans* natural isolates

To broadly analyze genetic variation among natural isolates, two pipelines were developed that utilize short-read sequencing data to identify mutations by mapping reads to a fully assembled and well-annotated reference genome. The first pipeline, Neostop, specifically identifies nonsense mutations, while the other, GenesComp, characterizes all genetic variants, including those occurring within the coding sequences (e.g., synonymous/nonsynonymous substitution, frameshift insertion/deletion, and nonsense mutation), as well as at intron splice sites. Both pipelines were applied to analyze the 387 *C. neoformans* natural isolates in the Strain Diversity Collection using publicly available whole genome sequencing data and the genome assembly and annotation of type strain H99 as reference (25, 42). We focused on genes encoding known components required for the RNAi pathway (SI Appendix, Table S1) (40, 43–47) and on variants predicted to cause loss-of-function (LOF) mutations, namely nonsense, frameshift, and splice-site mutations due to their high probability of causing truncated or dysfunctional gene products.

In total, we identified LOF mutations in five genes required for RNAi in 11 natural isolates (Table 1). Specifically, Neostop and GenesComp detected mutations in six and eight isolates, respectively, with three isolates detected by both pipelines. Most of the isolates had an LOF mutation in only one of the genes analyzed, with the exception of strain LP-RSA2296 which had LOF mutations in two RNAi-related genes *ZNF3* and *AGO1*. The mutations detected included the well-known nonsense mutation in the first exon of *ZNF3* in two hypermutator strains, Bt65 and Bt81, thus, providing a proof-of-principle for this approach (Fig. 1A) (40). All but one of the identified mutations were validated by PCR and Sanger sequencing; strain Bt84 was confirmed to be a false positive and excluded from further analysis (Table 1, SI Appendix, Dataset S1 and Fig. S1). Additionally, the splice-site mutation identified in the *RDE4* gene of strain Bt210 was validated by reverse transcription PCR (RT-PCR) and sequencing of the amplicon encompassing the intron 3/exon 4 boundary, confirming the predicted single nucleotide splicing error at the 3'-end of the

third intron, leading to a single nucleotide frame-shift insertion resulting in the introduction of a premature stop codon one amino acid into exon 4 (Fig. 1B and SI Appendix, Fig. S1).

Furthermore, the effects of the identified putative RNAi LOF mutations on small RNA production were analyzed by small RNA-sequencing (sRNA-seq), with H99 wild-type and H99 *rdp1*Δ strains serving as RNAi-proficient and -deficient controls, respectively (Fig. 1C and 1D). Our data showed that five of the strains tested (Bt210, LP-RSA2296, A2-102-5, D17-1, and NRHc5028.ENR.STOR (hereafter referred to as NRHc5028)) had sRNA profiles comparable to the RNAi-deficient control H99 *rdp1*Δ, that is, there was neither enrichment of sRNA species in the size range 21-24 nucleotides nor enrichment of 5'-uracil in the small RNA reads. This corresponds well with the size and feature of canonical small inferencing RNAs (siRNAs) utilized in RNAi-mediated silencing, thus confirming that these strains are indeed RNAi-deficient. On the contrary, strains Bt52, Bt152, and Bt208 had small RNA profiles similar to the RNAi-proficient control H99, with enrichment of both 21-24 nt species as well as a preference of 5'-uracil among these small RNAs, suggesting that while these strains have mutations in the *QIP1* gene, they are still able to produce small RNA species characteristic of RNAi-proficient strains (Fig. 1C and 1D). Interestingly, these three strains all have the same start-loss mutation in exon 1, which is predicted to result in a 15-amino-acid truncation of the Qip1 protein. We hypothesize that this modestly truncated Qip1 protein is still functional (Table 1). Taken together, these results indicate that five of the natural isolates identified by the bioinformatic pipelines, strains Bt210, LP-RSA2296, A2-102-5, D17-1, and NRHc5028, lack siRNAs necessary for RNAi silencing.

Hypermutation requires both RNAi-deficiency and elevated transposon burden

We have previously identified and characterized two clinical isolates (Bt65 and Bt81) with a nonsense mutation in the known RNAi component Znf3, which display a hypermutator phenotype when selected for rapamycin and FK506 resistance due to the increased mobilization of the retrotransposon Cnl1 into *FRR1*, which encodes FKBP12, the shared target of these drugs (40). To determine if the newly identified RNAi-deficient isolates are also hypermutators, fluctuation

assays were performed to quantify their mutation rates on medium containing rapamycin and FK506 (R+F) (Fig. 2 and SI Appendix, Dataset S2). None of these isolates showed significantly higher mutation rates on R+F YPD medium than the reference strain H99. They also had mutation rates comparable to H99 on YNB medium containing 5-FOA. Only the *msh2*Δ positive control strain, which lacks a core component of the DNA mismatch repair pathway, exhibited a significantly higher mutation rate than H99 on either media (Fig. 2 and SI Appendix, Dataset S3).

Given the important role of Cnl1 in causing R+F resistance, we hypothesized that although the five novel isolates had lost functional RNAi, they might not have undergone Cnl1 amplification as observed in Bt65 and Bt81 and, therefore, did not display a hypermutator phenotype. To test this hypothesis, publicly available short-read whole-genome sequence data were analyzed to assess the copy number of full-length Cnl1 (See Materials and methods) in these isolates. Compared to 178 or 102 copies predicted in Bt65 and Bt81, respectively, the estimated Cnl1 copy number in these isolates ranged from 0 copies (Bt210, A2-102-5, and NRHc5028) to 1 or 4 copies (D17-1 and LP-RSA2296) (Table 2). To validate these short-read based estimations, nanopore long-read sequencing was performed to obtain telomere-to-telomere assemblies for Bt210 and LP-RSA2296. Relative to Bt65 and Bt81, fewer full-length copies of Cnl1 (0 copies in Bt210 and 1 copy in LP-RSA2296) were confirmed by high-quality genome assemblies. These results provide evidence that being RNAi-deficient alone is likely not sufficient for hypermutation, and a relatively high transposon burden is likely also required.

Given the observed correlation between Cnl1 burden and hypermutator phenotype, we hypothesized that introduction of the novel RNAi LOF mutations identified here into a strain with a high Cnl1 burden would result in a hypermutator phenotype. In our previous study, restoration of functional RNAi against Cnl1 was successful by repairing the *znf3* nonsense mutation (to result in isolate Bt65+*ZNF3*) (40). Bt65+*ZNF3* contains a Cnl1 burden like that of Bt65 but is not a hypermutator due to its restored RNAi pathway, making it an ideal strain for testing this hypothesis. The *rde4* splice-site mutation from Bt210 and the *znf3* nonsense mutation from LP-RSA2296 (which differs from the original *znf3* nonsense mutation in Bt65, SI Appendix, Fig. S1A) were introduced

into the corresponding gene in Bt65+*ZNF3* via CRISPR-mediated allele exchange. One *rde4* transformant and two independent *znf3* transformants were successfully generated (SI Appendix, Fig. S2). In addition to the desired mutation, Sanger sequencing and Plasmidsaurus amplicon sequencing confirmed no additional mutations were found in these two loci. The *rde4* splicing error was also confirmed in Bt65+ZNF3 *rde4* by RT-PCR and Sanger sequencing (SI Appendix, Fig. S2). sRNA-seq further confirmed that like Bt65+*ZNF3 ago1*Δ, the canonical 21-24 nt sRNA peak with 5'-uracil identity was absent in these genetically modified *rde4* and *znf3* mutants, indicating defects in siRNA-mediated RNAi (Fig. 3A and SI Appendix, Fig. S3). To determine whether the introduced mutations confer a hypermutator phenotype in Bt65+*ZNF3*, mutation rates of these re-engineered RNAi-deficient strains were quantified by fluctuation analysis. In alignment with our prediction, when selected for R+F resistance, all three edited strains displayed significantly higher mutation rates than the recipient strain *Bt65+ZNF3* and negative control H99 (Fig. 3B and SI Appendix, Dataset S4).

To determine whether Cnl1 insertion is the main genetic change causing resistance, the FKBP12-encoding gene *FRR1* from spontaneous R+F resistant colonies derived from each strain was PCR amplified and sequenced. Consistent with our previous study, Cnl1 insertions were frequently observed in *FRR1* from the resistant colonies derived from the RNAi-deficient strain Bt65 (13/17, ~76%), but not the repaired RNAi-proficient strain Bt65+*ZNF3* (0/18, 0%) or H99 (0/9, 0%). In the three re-engineered RNAi-deficient strains, a high portion of R+F resistance was caused by insertion of Cnl1 into *FRR1* (ranging from ~37.5% to 69%), with the remaining resistant colonies having other types of genetic changes such as single nucleotide polymorphisms (SNPs), microINDELs (insertions or deletions $\leq$ 50 bp), or deletions (>50 bp deletion) (Fig. 3C and SI Appendix, Dataset S5). Interestingly, relative to the R+F sensitive progenitor, we also observed some resistant isolates from both Bt65+*ZNF3* (3/18, ~17%) and Bt65+*ZNF3 znf3-2* (2/16, 12.5%) that do not have any mutation (i.e., No change) at the *FRR1* locus including 1,699 bp 5' and 1,268 bp 3' sequences, leaving the cause of R+F resistance in these isolates unknown (Fig. 3C and SI Appendix, Dataset S5).

Movement of a novel DNA transposon occurs in the presence of functional RNAi, giving rise to drug resistance

During the fluctuation assay of the re-engineered RNAi-deficient strains, we noticed that some R+F resistant colonies had *FRR1* alleles containing insertions obviously larger than the expected size of the full-length Cnl1 element (~3.5 kb) (SI Appendix, Fig. S4 and SI Appendix, Dataset S5). Amplicon sequencing revealed these unusually large insertions were identical (10,944 bp element) and observed in the Bt65-derived strains regardless of their RNAi status. The frequency of resistant colonies attributable to this 10,944 bp insertion in *FRR1* ranged from ~6% (1/17 in Bt65) to 25% (2/8 in Bt65+*ZNF3 rde4*) (Fig. 3C). Two hallmark signatures of transposon movement, target site duplication (TSD, 8 bp) and terminal inverted repeats (TIR, 236 bp), were detected (Fig. 3D), suggesting the presence of a large uncharacterized mobile element. Although this element was not identified as mobile in H99, BLASTn results suggest the presence of a similar copy (100% query coverage and 98.41% identity) in the H99 genome. Based on the annotation of the H99 reference genome, two non-coding RNAs and three coding genes are included in this region; among these CNAG_00128 is predicted to contain a KDZ (Kyakuja, Dileera, and Zisupton) transposase domain and a putative zinc chelating domain, CxC1, indicative of a KDZ DNA transposon (48) (Fig. 3D). CNAG_12018, and CNAG_12019 are predicted non-coding RNAs, of which CNAG_12019 overlaps with CNAG_00128 in the opposite orientation. The other two genes, CNAG_00127 and CNAG_07313, encode unknown protein products (Fig. 3D). Due to its mobility and canonical KDZ transposon domain structure, we named this novel DNA transposon KDZ1 (the **1**st recognized **KDZ** element) in *C. neoformans*, whose movements have not been previously documented in pathogenic *Cryptococcus* species. In all cases of KDZ1 insertion into *FRR1*, full-length copies were observed (SI Appendix, Dataset S5).

The observation of KDZ1 transpositions in the repaired RNAi-proficient strain Bt65+*ZNF3* suggests that KDZ1 movement might evade functional RNAi machinery. To investigate whether RNAi is involved in silencing KDZ1, we re-analyzed our published sRNA-seq dataset for the RNAi-

298 deficient strain Bt65, two independent RNAi-repaired strains (Bt65+*ZNF3* and Bt65+*ZNF3-2*), as
299 well as H99 and its derived RNAi-deficient mutants H99 *rdp1*Δ and H99 *znf3*Δ (40). The sRNA
300 reads were first mapped against an array of Cnl1 elements located at the 3' end of the Bt65 Chr4.
301 Abundant 21 to 24 nt sRNA reads targeting Cnl1 in this region were present in the RNAi-repaired
302 strains Bt65+*ZNF3* and Bt65+*ZNF3-2*, while no such peaks were observed in the RNAi-deficient
303 Bt65 or other H99-derived strains lacking a full-length Cnl1 (Fig. 3E). In contrast, no obvious sRNA
304 targeting KDZ1 on the same Bt65 Chr4 was detected in Bt65+*ZNF3* or Bt65+*ZNF3*-2. The only
305 enrichment of sRNA reads was found in H99 around the overlapping region of CNAG_00128 and
306 the anti-sense non-coding RNA CNAG_12019, which also covers the KDZ transposase domain.
307 This enrichment is abolished in H99 *rdp1*Δ and H99 *znf3*Δ, indicating a possible role for these
308 sRNAs in RNAi silencing (Fig. 3E).

309 Our observation that the movement of a novel KDZ DNA transposon could lead to
310 antifungal resistance in *C. neoformans* prompted us to conduct further detailed investigation into
311 the diversity of KDZ transposons. Specifically, we first conducted a Hidden Markov Model (HMM)
312 search using the KDZ transposase domain (PF18758) against the genomes of the reference strain
313 H99 and the hypermutator strain Bt65 and identified five and six homologous sequences with high-
314 confidence E-values, respectively (SI Appendix, Table S2). We then conducted phylogenetic
315 analysis of these KDZ domain-containing proteins, together with those from the known KDZ
316 transposons from the model mushroom *Coprinopsis cinerea* that belong to the clades of Kyakuja,
317 Dileera, and Zisupton (Fig. 4A). Our results suggest that all the KDZ domain-containing proteins in
318 H99 and Bt65 belong to the Dileera clade with strong bootstrap support. The only exception was
319 CNAG_07523, whose placement in between the Dileera and Kyakuja clades still needed to be
320 further resolved (Fig. 4A). It should be noted that for the vast majority of the KDZ domain-containing
321 proteins in H99 and Bt65, we did not detect in their vicinity any of the other features that are
322 characteristic of a complete KDZ transposon, such as the CxCn domain, as well as the TIRs and/or
323 TSDs flanking the element, suggesting they may not represent active mobile elements.
324 Nevertheless, we did identify two KDZ-containing proteins in the Bt65 genome (termed KDZ2 and

KDZ3, which are absent in the H99 genome) that are associated with the CxC1 domain, residing in large DNA segments (10,053 bp for KDZ2 and 9, 218 bp for KDZ3) with TSDs and TIRs on both ends (Fig. 4B). Sequence analyses suggest that the KDZ proteins in KDZ2 and KDZ3 are significantly different from each other, although they are evolutionarily more closely related to each other than they are to KDZ1 (Fig. 4A and 4B). Indeed, we found one KDZ2 insertion in *FRR1* from a Bt65-derived F1 progeny (SI Appendix, Fig. S5). Thus, it is reasonable to propose that there are three different types of full-length KDZ transposons in Bt65, all of which are mobile and can induce genetic changes through their transposition.

Transposon accumulation during experimental evolution leads to hypermutation in *Cryptococcus*

We identified a strong correlation between the hypermutator phenotype and TE burden (e.g., Cnl1 burden) in the absence of functional RNAi. We next tested if a transition from non-hypermutator to hypermutator could occur following *in vitro* passage, a process known to lead to accumulation of TEs in *C. deneoformans*, a sister species of *C. neoformans* (49). Specifically, we selected two natural isolates that are RNAi-deficient and have low TE burdens in their genomes, Bt210 and LP-RSA2296, and conducted serial passages of the two strains, together with the *C. neoformans* reference strain H99 as control, on YPD solid medium for 2 months at 30°C and 37°C (transferred as patches every two days). Fluctuation assays were then conducted to estimate mutation rates on YPD+R+F solid medium for strains after 28 passages (mutation accumulation 28, i.e., MA28) and 56 passages (i.e., MA56). None of the passaged strains showed increased mutation rates on YPD+R+F solid medium compared to the H99 reference strain or the B210 and LP-RSA2296 progenitor strains (Fig. 5A; SI Appendix, Dataset S6), even though an initial pilot assay of these strains by a spread plate method suggested that Bt210-MA28 at 30°C, as well as LP-RSA2296-MA28 at both 30°C and 37°C, might have had increased mutation rates compared to the control strains (SI Appendix, Fig. S6). The mutational spectra of R+F resistant colonies from Bt210, LP-RSA2296, as well as their passaged strains were further characterized. We did not

observe any TE insertions in *FRR1*, and instead SNPs, deletions, and MicroINDELs were observed (Fig. 5B and SI Appendix, Dataset S7). Additionally, we estimated the copy numbers of Cnl1 and KDZ1 in the passaged strains using Illumina whole-genome sequencing reads and found no detectable accumulation of Cnl1 or KDZ1 in the passaged strains relative to their progenitor strains (SI Appendix, Table S3)

We further tested this hypothesis in *Cryptococcus deneoformans*, a sister species of *C. neoformans*. In this case, we utilized a strain, 37-02, which was derived from the laboratory strain XL280α by serial *in vitro* passage (49). Previous studies have shown that compared to the parental strain XL280α, strain 37-02 has undergone expansion of the Cnl1 retrotransposon (with ~40 additional copies) as well as other transposons in its genome (49). We also analyzed XL280α *rdp1*Δ and generated 37-02 *rdp1*Δ deletion strains to study the effects of RNAi on the mutation rates in these two genetic backgrounds and confirmed that they are indeed RNAi-deficient based on sRNA-seq analysis (Fig. 5C and SI Appendix, Fig. S7). Using fluctuation assays for R+F resistance, the passaged strain 37-02 had significantly higher mutation rates compared to its progenitor strain XL280α, suggesting that the expansion of transposons in the genome likely transformed strain 37-02 into a hypermutator following *in vitro* passage (Fig. 5D, SI Appendix, Dataset S8). Strain XL280α *rdp1*Δ had a mutation rate that was significantly higher than XL280α and comparable to the passage strain 37-02, indicating that inactivation of the RNAi machinery in this genetic background has a similar effect on mutation rates compared to transposon copy number expansion (Fig. 5D). The two independent *rdp1*Δ deletion strains in the 37-02 strain background exhibited mutation rates considerably higher than the XL280α *rdp1*Δ strain, and significantly higher than that of the 37-02 passage strain, providing evidence that the impacts of RNAi loss and TE accumulation on mutation rate are additive (Fig. 5D).

We next analyzed the mutation spectra of the R+F resistant colonies from all the tested strains by PCR and sequence analysis of the *FRR1* gene. With the exception of the wild-type progenitor strain XL280α, where the majority of mutations in the *FRR1* gene were due to changes other than transposon insertions (e.g., SNPs, deletions, and MicroINDELs), insertions of various

379 transposable elements into the *FRR1* gene were responsible for the majority of R+F resistance in
380 all of the other four derived and genetically modified strains tested, including 37-02 (Fig. 5E). In
381 total, seven different types of transposons were detected via insertions into the *FRR1* gene.
382 Interestingly, one of the mobile elements is a 6,397 bp novel element capped by 8 bp target site
383 duplications on both ends; insertion of this element occurred in all strains tested except XL280α.
384 Further analyses showed that this element contains two open reading frames based on the
385 annotation of the *C. deneoformans* type strain JEC21 (CNA07680 and CNA07690). While
386 CNA07680 encodes an unknown protein, the protein encoded by CNA07690 contains a KDZ
387 transposase domain, a CxC4 domain, and a SWIM (SWI2/SNF2 and MuDR) domain, which fits the
388 canonical domain structure of the Zisupton transposon from the KDZ DNA transposon class (8,
389 12). We therefore named this novel DNA transposon KDZx1 (**KDZ** in **X**L280 #**1**) in *C. deneoformans*
390 (Fig. 4A and Fig. 5E). Additional analysis showed that this element is also present in the *C.*
391 *deneoformans* strain JEC21, although it should be noted that in both cases we did not detect
392 terminal inverted repeats that typically are located at the two ends of KDZ transposons (Fig. 5E, F
393 and SI Appendix, Dataset S9). Analysis of complete genome assemblies of strains XL280α and
394 37-02 showed that KDZx1 had not undergone copy number expansion during *in vitro* passage (SI
395 Appendix, Table S4).

396 Surprisingly, we did not detect any Cnl1 insertions in *FRR1* among the 105 R+F resistant
397 colonies analyzed in the *C. deneoformans* strains (Fig. 5E and SI Appendix, Dataset S9). We
398 hypothesize that different transposons might be regulated by diverse mechanisms (e.g., RNAi-
399 dependent vs. RNAi-independent). To test this hypothesis, we counted the coverage of sRNA reads
400 in multiple transposons and found that KDZx1 and the DNA transposon T3 are enriched with 21 to
401 24 nt sRNA reads near their transposase domain, while no such enrichment was observed for the
402 DNA transposon T1 (Fig. 5G). Consistent with this finding, there is significant evidence to suggest
403 that the transposition of KDZx1 (Chi-squared test, p-value=0.003892) and T3 (Chi-squared test, p-
404 value=0.02092) is different between RNAi-proficient and RNAi-deficient backgrounds (Fig. 5E and
405 SI Appendix, Dataset S9). However, there is no such correlation for the movement of T1 (Chi-

406 squared test, p-value= 0.1336) (Fig. 5E and SI Appendix, Dataset S9). Taken together, our results
407 show that both RNAi loss and/or transposon accumulation can lead to hypermutation in
408 *Cryptococcus* via uncontrolled transposition.
409
410 **Hypermutation caused by RNAi loss and transposition can occur following sexual**
411 **reproduction**
412 We have identified a new class of mobile elements in *C. neoformans*, the KDZ transposons,
413 whose transposition can give rise to antifungal resistance. It is known that transposons are
414 activated during meiosis, and genetic variation resulted from sexual reproduction and meiotic
415 recombination could have contributed to the observed variation in KDZ and Cnl1 distribution and
416 activity across isolates (44, 50). Thus, we sought to investigate how sexual reproduction could
417 influence: 1) the activity and regulation of transposons, and 2) the mutation rates of RNAi-deficient
418 progeny inheriting a transposon burden. To this end, we crossed two natural isolates, recovered
419 meiotic progeny, and studied the mutation rates, and their relationship to transposon (e.g., KDZ
420 and Cnl1 transposon) movements and regulation. Specially, we crossed strain Bt210 (*ZNF3 rde4*)
421 with strain Bt65 (*znf3 RDE4*). Bt210 is RNAi-deficient due to a mutation in the *RDE4* gene and a
422 non-hypermutator likely due to a lack of sufficient transposon burden in the genome. Bt65 is an
423 RNAi-deficient hypermutator because of a nonsense mutation in a key RNAi gene, *ZNF3*, and with
424 an elevated load of transposable elements in its genome (e.g., KDZ and Cnl1 elements). A total of
425 57 meiotic progeny were recovered, and segregation analyses of *ZNF3*, *RDE4*, and the mating
426 type locus confirmed that this progeny set represents a population of random meiotic products (SI
427 Appendix, Dataset S10). Whole-genome sequencing also confirmed meiotic recombination on a
428 genome-wide scale (SI Appendix, Fig. S8).
429 Among the tested F1 progeny (all 11 progeny with *rde4 ZNF3* and 2 progeny with *rde4*
430 *znf3*), three exhibited a hypermutator phenotype on YPD medium containing R+F. Two of these
431 hypermutators (P12 and P25) were *ZNF3 rde4* (same genotype as the Bt210 parent), while the
432 other one (P57) was *znf3 rde4*, a recombinant genotype between the Bt210 and Bt65 parents.

While all three F1 progeny had mutation rates that were significantly higher than the non-hypermutator parent Bt210, when compared to that of the hypermutator parent Bt65, the mutation rate of P57 was relatively lower while those of P12 and P25 were comparable to each other (albeit lower than Bt65 as well) (Fig. 6A and SI Appendix, Dataset 11). sRNA sequencing confirmed that the three hypermutator F1 progeny (P12, P25, and P57) are RNAi-deficient, lacking the 21 to 24 nt sRNA species and the 5'-uracil enrichment in sRNA reads (Fig. 6B and SI Appendix, Fig. S9).

The mutation spectra at the *FRR1* locus in R+F resistant colonies derived from control, parental, and both hypermutator and some non-hypermutator progeny were assessed (Fig. 6C). Insertion of the KDZ1 and Cnl1 transposons at the *FRR1* locus contributed to a considerable proportion of R+F resistant isolates in the hypermutator parent Bt65, as well as in the F1 progeny regardless of whether they were hypermutators or not. This is in stark contrast to the non-hypermutator parent Bt210 as well as H99 and LP-RSA2296, where the vast majority of the R+F resistant colonies had SNPs in the *FRR1* gene, and no transposon insertions were detected. Interestingly, while in Bt65, the majority of transposon insertions were attributable to the Cnl1 element with few KDZ1 insertions being observed (~5%, 1 of 21 independent colonies). A relatively elevated number of KDZ1 insertions were observed in three F1 progeny (P20, P32, and P12), of which P12 is a hypermutator with 13 of 18 (~72%) independent resistant colonies derived from this progeny contained KDZ1 insertions in *FRR1* (Fig. 6C and SI Appendix, Dataset 12). Importantly, in the two other hypermutator progeny (P25 and P57), the majority of R+F resistant colonies (50%, 10/20 in P25 and ~59% 10/17 in P57) were attributable to Cnl1 insertions into the *FRR1* gene. Additionally, there were resistant colonies derived from P25 (35%, 7/20) and P57 (~35%, 6/17) in which we could not detect any mutation at the *FRR1* locus (including 1,699 bp 5' and 1,268 bp 3' sequences). To test the correlation between the hypermutator phenotype and the inherited transposon burden in the F1 progeny, we estimated the copy numbers of Cnl1 and KDZ1 by mapping the Illumina short-reads. We found that collectively the three hypermutator F1 progeny inherited significantly more Cnl1 elements than the remaining non-hypermutator progeny (Mann-Whitney U Test, p-value= 0.03571) (SI Appendix, Table S5). Moreover, progeny P12, which had

460 the highest frequency of KDZ1 insertion among resistant colonies, also contained the highest
461 abundance of KDZ1 (7 copies), which is higher than any other tested progeny as well as the two
462 parents (SI Appendix, Table S5). Thus, our results demonstrate that sexual reproduction has the
463 potential to 1) liberate the KDZ mobile elements and 2) cause the inheritance of a high Cnl1 element
464 burden along with a loss of RNAi mutation, resulting in hypermutation.

465 To determine whether hypermutation can result in resistance to clinically relevant
466 antifungal drugs, mutation rates on medium containing 5-fluorocytosine (5-FC), a systematic
467 antifungal that is one of the first-line treatments for cryptococcal infection, were measured.
468 Fluctuation analysis suggested that none of the three F1 progeny as well as Bt210 and Bt65 display
469 a hypermutator phenotype when selected for 5-FC resistance (SI Appendix, Fig. S10). We further
470 analyzed the 5-FC resistant colonies isolated from P12, P25, P57, and the parents Bt210 and Bt65
471 by PCR amplifying three genes in which mutations are known to confer 5-FC resistance (*FUR1*,
472 *UXS1* and *FCY2*) (51). Sequencing results revealed that KDZ1 insertions underlie resistance in
473 four 5-FC resistant colonies derived from P12. Two insertions were observed in *FUR1*, and two
474 were observed in *UXS1*. Additionally, we detected a KDZ1 insertion in *UXS1* in one of the resistant
475 colonies derived from Bt65 (SI Appendix, Fig. S10 and SI Appendix, Dataset 13). These results
476 demonstrate that KDZ1 can mediate resistance to a clinically relevant antifungal drug.

Discussion

Two bioinformatic pipelines were developed to identify mutations in required RNAi components using short-read whole-genome sequence data for 387 *C. neoformans* natural isolates (25). In addition to two previously identified RNAi-loss hypermutators, Bt65 and Bt81, these pipelines successfully identified five novel RNAi-deficient isolates (Bt210, LP-RSA2296, A2-102-5, D17, and NRHc5028), as evidenced by sRNA-seq analysis. In contrast to Bt65 and Bt81 (from the Sub-Saharan African VNBII lineage), all five newly identified RNAi-loss isolates are from the global VNI lineage; among these, clinical isolates Bt210 and LP-RSA2296 belong to two different clades, while environmental isolates A2-102-5, D17-1, and clinical isolate NRHc5028, are all from a single clade. The identification of five additional non-VNB RNAi-deficient isolates in *C. neoformans* suggests that the loss of RNAi within this species may be more widespread than previously appreciated. To date, RNAi loss has been confirmed in seven natural isolates, which account for ~1.8% of the 387 isolates in the Strain Diversity Collection (25, 40). Given our limited understanding of the RNAi components in *C. neoformans* and the nature of our pipeline (from genotype to phenotype and only focused on LOF mutations), the full extent of RNAi loss in this species may still be underestimated. Interestingly, a recent study revealed the presence of an intact and functional RNAi pathway in the human fungal pathogen *C. albicans*, which was previously thought to be RNAi-deficient based on a single reference strain (SC5314) (22). Our results, along with this recent finding from *C. albicans,* provide robust examples of the importance of variation not only between species, but also within a species and highlight the limitations of making conclusions based on analysis of a single reference strain.

All five newly identified RNAi-deficient isolates failed to display hypermutator phenotypes when selected for R+F or 5-FOA resistance. Compared to the hypermutator isolates Bt65 and Bt81, which have approximately 100 copies of the non-LTR retrotransposon Cnl1, the five newly identified RNAi-deficient isolates contain significantly fewer copies of Cnl1 (0 to 4 copies). This association between the hypermutator phenotype and copy number of the Cnl1 retrotransposon is in line with our hypothesis that it is the combination of RNAi loss and a high transposon burden that causes

hypermutation (40). This hypothesis is also supported by the transition from non-hypermutator to hypermutator observed in our CRISPR re-engineered allele exchange strains. These findings elucidate two distinct evolutionary trajectories following RNAi loss: one pathway leading to accumulation of transposons and resulting in a hypermutator phenotype, and the other where the organism maintains or loses transposons and persists as a non-hypermutator (Fig. 6D). Consistent with our model, the overall repeat content is similar between the RNAi-deficient isolates Bt210 and LP-RSA2296, and the RNAi-proficient reference strain H99 (SI Appendix, Dataset 14). Additionally, the loss of RNAi in the *C. deuterogattii* type strain R265 is associated with a reduction of full-length retroelements at the centromeres (21, 41). Although we did not observe significantly increased mutation rates or transposon accumulation following passage of RNAi-loss strains Bt210 and LP-RSA2296, possibly attributable to the relatively short passage time and low initial transposon numbers (0 copies Cnl1 in Bt210 and 1 copy of Cnl1 in LP-RSA2296), it remains possible that the transposon burden could increase over time without RNAi and eventually lead to hypermutation.

Our results in *C. deneoformans* indicate that even with functional RNAi, transposon accumulation can lead to hypermutation, as seen by the increased mutation rate of the passaged strain 37-02 compared to its progenitor XL280α (49). Additionally, removing RNAi in 37-02 further increased the mutation rate, highlighting an additive effect between RNAi loss and transposon accumulation. This suggests an alternate evolutionary path to become a hypermutator in which transposons first accumulate followed by RNAi loss (Fig. 6D). Interestingly, drug-resistant colonies with various types of transposon insertions were found in both RNAi-proficient XL280α and 37-02, as well as in their RNAi-deficient deletion mutants. This observation, combined with distinct sRNA read enrichment, suggests the presence of RNAi-independent mechanisms controlling transposable elements and highlights a substantial transposon reservoir in *C. deneoformans*. Furthermore, unlike Bt65 and Bt81 (40), after screening more than 100 independent R+F resistant colonies from *C. deneoformans*, we did not find any Cnl1 insertions in the *FRR1* gene. These results suggest a distinct insertion preference of Cnl1 between *C. neoformans* and *C. deneoformans* and might also be due to their difference in Cnl1 copy number. Similarly, a previous

study found no evidence of Cnl1 moving into non-telomeric regions when selected for antifungal drug resistance in *C. deneoformans* (49).

During our efforts to characterize mutational spectra, we surprisingly identified a novel class of large KDZ DNA transposons in *Cryptococcus*, KDZ1, KDZ2, KDZ3 in *C. neoformans* and KDZx1 in *C. deneoformans*, whose movements had not been previously documented. We found evidence indicating KDZ1, KDZ2, as well as KDZx1 can mobilize into genes encoding drug targets, thereby giving rise to antifungal drug resistance. Given these observations, it is reasonable to hypothesize that KDZ3 may similarly insert into genes encoding antifungal drug targets. The KDZ DNA transposon class, which consists of three different clades: **K**yakuja, **D**ileera and **Z**isupton, was first identified in the mushrooms *Coprinopsis* and *Laccaria* and shown to be associated with the expansion of TET/JBP dioxygenases in basidiomycetes (48). Our discovery that diverse KDZ transposons are present in *Cryptococcus*, as well as their transposition leading to drug resistance and elevated mutation rates, echoes the growing appreciation of the diversity of transposable elements, and the effects of their movements on the host phenotype and genome integrity and stability in a variety of organisms (39, 52, 53). The average size of our identified KDZ transposons is ~9.1 kb (ranging from ~6.4 kb for KDZx1 to ~10.9 kb for KDZ1), which is larger than most known transposons in *Cryptococcus*, such as the ~3.5 kb full-length Cnl1. In addition to the coding sequence containing the KDZ transposase domain along with its N-terminal zinc chelating domain (CxCn domain), these transposons also harbor extra open reading frames encoding proteins without predicted functional domains. Other recently characterized giant mobile elements (e.g., *Starship*) are known to carry cargo sequences contributing to genomic variability within various fungal species (54–56), and these additional sequences may function in a similar fashion.

Interestingly, we also found additional annotated proteins in *Cryptococcus* genomes that contain the KDZ transposase domain, although they lack other features of KDZ transposons, such as the CxCn domain, and we did not detect sequences indicative of the flanking TIRs and TSDs in their vicinity. In a recent study, one such protein, CNAG_04102, was implicated to be associated with virulence variation in *C. neoformans* based on a genome-wide association study (57). These

KDZ-transposase-containing proteins could represent relics of inactivated KDZ transposons, or they could be unique types of KDZ transposons that are yet to be characterized. Further studies are required to gain a better understanding of the characteristics and biology of this novel class of KDZ mobile elements.

In the *C. neoformans* reference strain H99, a non-coding RNA (CNAG_12019) was also annotated inside the KDZ1 transposon, and it is on the anti-sense strand of the region corresponding to the KDZ transposase domain. This suggests a possibility where double-stranded RNA is formed to target the KDZ transposase and initiate RNAi silencing of KDZ1. However, our observations of distinct activities between Cnl1 and KDZ1 in the RNAi-repaired strain, Bt65+*ZNF3*, suggest a distinct regulatory mechanism for KDZ1. Thus, while RNAi plays an important role in regulating Cnl1 (40), a different or additional mechanism might be involved in regulating KDZ1. Similarly, studies in *C. deneoformans* showed that while both KDZx1 and T3 DNA transposons (47) appear to be regulated by RNAi, the T1 DNA transposon remains active despite functional RNAi. These findings highlight the complexity of the regulation of transposable elements, as well as the consequent transposon-mediated antifungal drug resistance, and emphasize the need for further investigation into the regulatory mechanisms controlling various types of transposons in *Cryptococcus* and other species.

Three F1 hypermutator progeny inheriting the *rde4* splice mutation were identified from a genetic cross between Bt210 (*rde4*, RNAi-loss with limited transposons) and Bt65 (*RDE4*, RNAi-loss with enriched transposons). In addition to the Cnl1 insertions that predominated in two F1 progeny (P25 and P57), we found that KDZ1 transposition underlies most of the R+F resistance observed in one of the progeny (P12). Movement of Cnl1 and KDZ1 was also observed in several non-hypermutator F1 progeny, indicating the active status of these transposons. These findings suggest that hypermutation can evolve in two steps: an initial loss of RNAi by mutation followed by sexual reproduction with an isolate with a high transposon burden. F1 progeny from such a cross could display a hypermutator phenotype due to the combined effects of RNAi loss and inheritance of a transposon burden.

Taken together, our results demonstrate that rampant insertion of transposons, including the novel KDZ and Cnl1 mobile elements, conferred a hypermutator phenotype in the RNAi-deficient progeny produced through sexual reproduction, indicating that RNAi loss provides a pathway to hypermutation. The unique transposition profiles of the KDZ elements among these RNAi-deficient progeny suggest the presence of multifaceted and diverse regulatory mechanisms against different transposable elements, which could result from evolutionary pressure to suppress the propagation of selfish mobile elements and defend host genome integrity. Further research on the transgressive KDZ1 movement following sexual reproduction will illuminate the regulatory mechanism(s) controlling KDZ1. Notably, only a limited number of genetic loci and insertion sizes can be captured by the fluctuation analysis and PCR-based genotyping strategy applied in this study. Due to these experimental limitations, the activity and the role of the gigantic KDZ transposons within the genome are likely underappreciated.

Collectively, our findings revealed relatively frequent RNAi loss in *C. neoformans* natural isolates, and many of these may lie on a pathway that could lead to hypermutator evolution mediated by various types of transposons (Fig. 6D). With the discovery of KDZ transposons in *Cryptococcus*, our study sheds light on a darker side of the fungal genome, which could play a significant role in adaptation within fungal populations.

Materials and methods

Strains and growth

The *C. neoformans* and *C. deneoformans* strains analyzed in this study are listed in SI Appendix, Table S6. Strains were stored at -80°C and supplemented with 15% glycerol for long-term storage. Strains were inoculated on YPD solid medium at 30°C for 3 days and maintained at 4°C no more than two weeks. Fresh cells were streaked out from glycerol stocks as needed.

Pipeline constructions

Construction of reference-based genome assemblies for Pipeline 1 (Neostop). Raw Illumina sequencing reads generated by Desjardins et al. 2017 were downloaded from the NIH Sequence Read Archive (BioProject ID PRJNA382844) (25). For each of the 387 BioSamples (corresponding strains of interest) associated with the BioProject, a reference-based genome assembly was generated by aligning paired-end sequence data to the genome of the *C. neoformans* reference strain H99 (FungiDB R53) in Pipeline 1. In cases where there were multiple sequencing runs for a given BioSample, the sequencing run containing the largest number of paired-end reads was used. To produce reference-based genome assemblies, reads were aligned to the reference genome using BWA (v0.7.12), variants called using FreeBayes (v1.1), and strain-specific consensus assemblies generated by instantiating the called variants onto the reference genome (58, 59) The read alignment, variant calling, and consensus assembly were carried out using the Snippy (https://github.com/tseemann/snippy) pipeline tool. Following construction of consensus assemblies, genome feature annotation was "lifted over" from the H99 reference genome to each strain-specific genome using the software tool Liftoff (v1.6.3) (60). The `-polish` option of Liftoff was employed to re-align exons in cases where the lift-over procedure resulted in start/stop codon loss or introduced an in-frame stop codon. Based on the polished lift-over annotation, the AGAT GTF/GFF Toolkit software (https://github.com/NBISweden/AGAT) was used to predict protein sequences for all annotated genes in each strain-specific assembly using the `agat_sp_extract_sequences.pl` script. Where multiple protein isoforms are annotated in the

reference genome, predictions were generated for each isoform. For each RNAi-related gene of interest (SI Appendix, Table S1), the predicted amino acid sequences of each of the 387 *C. neoformans* strains was compared to the H99 reference genome (FungiDB R53) with Pipeline 1. Candidate loss-of-function alleles were classified as those encoding proteins whose predicted length is < 90% of the model protein length estimated from the entire set of strains.

A second, independent program to identify mutations of interest in *C. neoformans* sequencing data was constructed. Polymorphisms were independently identified with the GenesComp program (Pipeline 2), which is available through GitHub (https://github.com/derf0/genescomp). This is a species-specific program for finding polymorphisms. GenesComp attempts to classify polymorphisms according to their likelihood to alter function of the gene product. This includes polymorphisms that may increase the amount of protein, such as changes from rarely used codons to more frequently used codons, as well as changes in introns from poor splice site or branch site sequences to more advantages sequences. For *C. neoformans*, this analysis was carried out with the *C. neoformans* strain H99 (42) as the reference, and intron branch site, splice site, and permissible length rules determined for this strain, as well as codon usage (GenBank: CP003820 through CP003834). Gain of stop codons (nonsense mutations) and loss of stop codons ("resense" mutations) were detected, along with insertions/deletions, in-frame and out-of-frame, and synonymous and non-synonymous polymorphisms. Non-synonymous polymorphisms were categorized as either conservative or non-conservative. Polymorphisms were also identified in UTR regions and in RNA-coding genes such as those encoding tRNAs. The program uses the genomic sequence and annotation of the reference genome, so the accuracy of the predictions is dependent on the accuracy of the annotations and reference strain genome assembly. The program is written in Python3, though uses several external programs: EMBOSS package version 6 of DNA/RNA/protein programs, Blastx NCBI stand-alone blast package, search Smith-Waterman alignment, and tRNAscan-SE (61–64). For robustness, other analysis is done in the code rather than using external applications.

656 **Small RNA isolation, sequencing, and analysis**

657 Small RNA isolation and analysis were performed as described previously (40, 65). Briefly,
658 strains were cultured in 8 mL YPD liquid medium overnight at 30°C roller drum with 70 rpm.
659 Overnight cultures were pelleted and frozen at -80°C for 1 h. Cells were then lyophilized overnight.
660 Small RNA was isolated from ~70 mg of the lyophilized cell using the mirVana™ miRNA Isolation
661 Kit (Thermo Fisher) following manufacturer's instructions. Small RNA sequencing libraries were
662 prepared using QIASeq miRNA Seq Auto Kit (Qiagen) and sequenced on the Illumina NextSeq 500
663 (1x75 bp) or Novaseq X Plus (2x150 bp) platforms at the Duke University Sequencing and Genomic
664 Technologies Core facility. Sequencing adaptors were removal by Cutadapt (v1.18) (66), followed
665 by discarding untrimmed reads or reads shorted than 14 nt or longer than 40 nt. The remaining
666 reads were mapped to appropriate reference genome with Bowtie (v1.3.1) (67) that allowing
667 multiple alignments (Bt65 and XL280α) or up to five alignments (H99), but no mismatches. Custom
668 perl script was used for counting read size distribution and 5'-nucleotide identities from 14 to 28 nt
669 (https://github.com/timdahlmann/smallRNA) (68). The sorted and indexed BAM file, generated after
670 bowtie mapping was used as input for counting and visualizing the normalized read coverage
671 (RPM, Read Per Million) in the coordinates of specific transposon element using sRNA_Viewer
672 (https://github.com/MikeAxtell/sRNA_Viewer). All sRNA analyses were performed in at least two
673 biological replicates, except for Bt52, Bt152, Bt208, Bt210, and LP-RSA2296 (one biological
674 replicate).

675

676 **Fluctuation assay and mutational spectra analysis**

677 Fluctuation assays were performed as described previously (40). Briefly, 10 to 20
678 independent cultures of each strain were grown in 5 mL YPD for 12 to 18 h in a roller drum set to
679 70 rpm and at 30°C. The cultures were then centrifuged at 3000xg for 5 min before being
680 resuspended in sterile water. The resuspended cultures were diluted 100,000x, and 100 µL of each
681 were plated on YPD to assess the total number of cells in each culture. 100 µL of the undiluted
682 cultures were plated on YPD medium containing 100 ng/mL rapamycin + 1 µg/mL FK506 (R+F) or

YNB containing 1 mg/mL 5-Fluoroorotic Acid (5-FOA). For assessing hypermutation on YNB medium containing 100 µg/mL 5-fluorocytosine (5-FC), the resuspended cultures were diluted 100x, and 100 µL were plated. Colonies were counted following incubation for 3 to 4 days on YPD at 30°C and 6 to 12 days on drug plates (30°C for 5-FOA and 5-FC, and 37°C for R+F) using an eCount colony counter and ImageJ. Mutation rates and 95% confidence intervals for each strain were calculated using the FluCalc program (https://flucalc.ase.tufts.edu/) (69). As a strict standard, no overlap between respective 95% confidence intervals indicates a significant difference (70).

To ensure the identified mutant events were independent, only one resistant colony per drug plate was purified, as described previously (40). DNA extracted from independent drug-resistant colonies was used for genotyping PCR of target genes (e.g., *FRR1*, *FUR1*, *UXS1*, and *FCY2*). *ACTIN* was amplified as a DNA quality and loading control. LaTaq was utilized for amplifying extra-long PCR products (NEB). The PCR products were sequenced using Sanger or Plasmidsaurus amplicon sequencing (https://www.plasmidsaurus.com/), allowing for the detection of sequence variations.

Nanopore long-read sequencing and *de novo* genome assembly

High molecular-weight DNA was extracted by CTAB method as described previously (71). DNA size was estimated by pulsed-field gel electrophoresis and quantified by Qubit high sensitivity assay kit. Nanopore Sequencing library was prepared by using native barcoding kit SQK-NBD114.24 following manufacturer's protocol. Up to 2 samples were barcoded and sequencing in R10.4.1 flow cells (FLO-MIN114) for 72h. After sequencing, raw Pod5 files were trimmed, basecalled and demultiplexed with Dorado (v0.5.3, https://github.com/nanoporetech/dorado). Canu (v2.2) (72) was used for *de novo* genome assembly following sequence correction with medaka (v1.11.3, https://github.com/nanoporetech/medaka) and up to 5 rounds of Pilon (v1.24) (73) polish with Illumina WGS reads. Naming and orientation of each contigs were adjusted based on the H99 reference genome (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000149245.1/). Unplaced contigs with >50% duplicated with 14 core chromosomes or bacterial DNA were removed

710 from the assemblies. Telomere repeats (5' (TAACCCC)n '3) were found on both ends of the 14
711 chromosomes of Bt210 and LP-RSA2296 in the assemblies. The repeat content and annotation
712 analysis were performed using RepeatMasker (v4.0.7) with Dfam (v3.3) and
713 RepBaseRepeatMaskerEdition-20181026 libraries that was supplemented with RepBase EMBL
714 database (v26.04). All the repeats identified were used for the comparison between three strains,
715 H99, Bt210 and LP-RSA2296.
716
717 **CRISPR-mediated allele exchange and gene deletion**
718 Guide RNAs (gRNAs) for allele exchange assays were designed by using EuPaGDT
719 website (http://grna.ctegd.uga.edu/) and constructed as previously described (74, 75). Donor DNA
720 containing the sequence of interests was generated by mutating the endogenous sequences via
721 Q5 Site-Directed Mutagenesis Kit (for causing the *rde4* splice-site mutation or the *znf3* non-sense
722 mutation in Bt65+*ZNF3*) (NEB). Safe Haven 1 (SH1) locus was used for integration in
723 transformants (76).
724 Strains were transformed with 500 ng *C. neoformans* codon optimized Cas9 (75), 2 µg
725 allele exchange donor DNA and 300 ng gRNA against the mutation of interest as well as 300 ng
726 gRNA targeting SH1 and 2 µg AscI digested pSDMA25 plasmid (76), via a transient CRISPR-Cas9
727 coupled with electroporation (TRACE) system as previously described (77). Potential transformants
728 were selected on YPD+NAT, DNA extracted from resistant colonies were used as template for
729 amplifying a region outside the allele swap donor for avoiding the condition of ectopic insertion. In
730 addition to the desired mutations, no additional mutations were found by amplicon sequencing
731 (Plasmidsaurus). The transformants with expected sequence alternation were further validated for
732 the presence of Cas9, gRNA cassette and the status for SH1 by genotyping PCR.
733 To delete the *RDP1* gene in *C. deneoformans* 37-02, 1,069 bp upstream and 1,031 bp
734 downstream flanking sequences of *RDP1* were fused to *NAT* marker amplifying from pAI3 plasmid
735 by split-marker method. Two independent gRNAs were designed for improving knockout efficiency.
736 37-02 was co-transformed with Cas9, dual gRNAs, and the *NAT* split marker by the TRACE

737 protocol (77). Internal, 5' junction and 3' junction PCR were utilized for confirming the successful
738 gene deletion.
739
740 ***In vitro* passage assay and preliminary hypermutator screening**
741 Five single colonies of Bt210, LP-RSA2296, and H99 were passaged as patches on YPD
742 medium at 30°C and 37°C for 28 and 56 passages, respectively, every 48 hours. Two independent
743 cultures per strain were cultured overnight at 30°C and resuspended in 2 mL sterilizer water.
744 Homogenized cultures after vortex were swab on a quadrant of the YPD medium containing 100
745 ng/mL rapamycin + 1 µg/mL FK506 (R+F).
746
747 **Phylogenetic analysis**
748 The sequences of identified KDZ-transposase-domain containing proteins from
749 *Coprinopsis cinerea* were obtained from NCBI GenBank (https://www.ncbi.nlm.nih.gov/genbank/)
750 (48). The maximum-likelihood phylogeny was constructed with MEGA X (v10.0.1) with 1000
751 bootstrap replicates (78).
752
753 **Mating and spore dissection**
754 A cross between Bt210 and Bt65 was performed as previously described (79). A total of
755 57 F1 progeny were isolated via random spore microdissection from the MS plate, which were
756 plated with a mixed cell suspension containing both Bt210 and Bt65 (79).
757
758 **Transposon copy number estimates from assembled genomes using BLAT**
759 To estimate transposon copy number from assembled genomes, we used BLAT (80) to
760 search for regions homologous to KDZ1, KDZ2, KDZ3, and Cnl1 in the assembled *C. neoformans*
761 genomes given in SI Appendix, Dataset S15. The sequences of each transposon were taken from
762 the Bt65 genome where the KDZ transposons were first identified. BLAT searches were run with a
763 minimum identity of 0.9. The BLAT accessory tools pslReps was subsequently used to find the best

alignments for each query sequence with a minimum coverage of 0.95. The number of such best alignments per genome of interest was used as the BLAT-based estimate of transposon copy number.

Transposon copy number estimates via relative read depth analysis

To estimate transposon copy number from Illumina whole-genome short-read data we carried out the following analysis steps for each strain of interest: 1) Illumina short reads were mapped to the genome of the reference strain H99 using minimap2 (81), and the median genome-wide read depth was estimated using the tool Mosdepth (82); 2) The same reads were mapped to a "transposon genome" consisting of the concatenated sequences of KDZ1, KDZ2, KDZ3, and Cnl1; 3) For the transposon mapped reads, the cumulative read depth distribution output of Mosdepth was used to estimate the read-depth at which a given query sequence had a coverage of at least 0.95; 4) The read depth at coverage 0.95 (step 3) was divided by the genome-wide median read depth (step 2) to give a normalized read-depth estimate.

To provide a benchmark of the performance of our relative read depth estimates of transposon copy number, we used linear regression to fit copy number estimated via BLAT from whole genome assemblies to the read-depth estimated copy number (SI Appendix, Fig. S11). We find strong agreement between the two estimators of copy number, with R^2 values of the linear models between 0.95 and 0.98 depending on whether estimates for Cnl1 copy number in Bt65 are included or excluded.

Data Availability

The *de novo* assemblies of Bt210 and LP-RSA2296 have been deposited with accession number PRJNA1138746 in the NCBI Bioproject database. The raw sequence reads for Nanopore sequencing, sRNA-seq and Illumina whole-genome sequencing have also been deposited under the same BioProject accession number. sRNA data presented in Fig. 3E were collected previously and accessed through NCBI Bioproject accession PRJNA749953. The novel KDZ transposons

identified in this work have been submitted to GenBank and are available as GenBank number PQ181658 (KDZ1), PQ181659 (KDZx1), PQ181660 (KDZ2), and PQ181661 (KDZ3).

Acknowledgments

We thank Anna Floyd-Averette for constant support, and all the members of the Heitman Lab for constructive suggestions. We also thank Dr. Asiya Gusa and Anna Mackey (Duke University) for generously sharing the passaged strain 37-02 and XL280α *rdp1*Δ and for discussions and communication of unpublished results on their independent discovery of the KDZ transposons, Dr. Ruiyun Zeng (North Carolina State University) for assistance in data visualization, and Dr. Tim A. Dahlmann and Dr. Minou Nowrousian (Ruhr-Universität Bochum) for suggestions on sRNA data analysis. This study was supported by NIH/NIAID R01 grants AI039115-27, AI050113-20, and AI133654-07. J. Heitman is co-director and fellow of the CIFAR program Fungal Kingdom: Threats & Opportunities. We thank the Madhani laboratory and NIH grant R01 AI100272 for the KN99α *msh2*Δ and *frr1*Δ deletion strains.

References

1. M. Habig, C. Lorrain, A. Feurtey, J. Komluski, E. H. Stukenbrock, Epigenetic modifications affect the rate of spontaneous mutations in a pathogenic fungus. *Nat Commun* **12**, 5869 (2021).
2. J. Huang, D. E. Cook, The contribution of DNA repair pathways to genome editing and evolution in filamentous pathogens. *FEMS Microbiology Reviews* **46**, fuac035 (2022).
3. M. Villalba De La Peña, P. A. M. Summanen, M. Liukkonen, I. Kronholm, Chromatin structure influences rate and spectrum of spontaneous mutations in *Neurospora crassa*. *Genome Res.* **33**, 599–611 (2023).
4. R. B. Billmyre, S. A. Clancey, J. Heitman, Natural mismatch repair mutations mediate phenotypic diversity and drug resistance in *Cryptococcus deuterogattii*. *eLife* **6**, e28802 (2017).
5. A. K. Rasmussen, Mitochondria-mediated nuclear mutator phenotype in *Saccharomyces cerevisiae*. *Nucleic Acids Research* **31**, 3909–3917 (2003).
6. K. J. Boyce, Mutators enhance adaptive micro-evolution in pathogenic microbes. *Microorganisms* **10**, 442 (2022).
7. M. I. Navarro-Mendoza, C. Pérez-Arques, J. Heitman, Heterochromatin and RNAi act independently to ensure genome stability in Mucorales human fungal pathogens. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2220475120 (2023).
8. R. B. Billmyre, S. Calo, M. Feretzaki, X. Wang, J. Heitman, RNAi function, diversity, and loss in the fungal kingdom. *Chromosome Res* **21**, 561–572 (2013).
9. A. Fire, *et al.*, Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
10. N. Buchon, C. Vaury, RNAi: a defensive RNA-silencing against viruses and transposable elements. *Heredity* **96**, 195–202 (2006).
11. D. Baulcombe, RNA silencing in plants. *Nature* **431**, 356–363 (2004).
12. R. Martienssen, D. Moazed, RNAi and heterochromatin assembly. *Cold Spring Harb Perspect Biol* **7**, a019323 (2015).
13. T. Sijen, R. H. A. Plasterk, Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi. *Nature* **426**, 310–314 (2003).
14. I. A. Drinnenberg, *et al.*, RNAi in budding yeast. *Science* **326**, 544–550 (2009).
15. 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).
16. J. D. Laurie, *et al.*, Genome comparison of barley and maize smut fungi reveals targeted loss of RNA silencing components and species-specific presence of transposable elements. *Plant Cell* **24**, 1733–1745 (2012).
17. 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).
18. M. J. Schmitt, F. Breinig, Yeast viral killer toxins: lethality and self-protection. *Nat Rev Microbiol* **4**, 212–221 (2006).
19. C. Park, N. Banerjee, Y. Koltin, J. A. Bruenn, The *Ustilago maydis* virally encoded KP1 killer toxin. *Molecular Microbiology* **20**, 957–963 (1996).
20. S. Applen Clancey, F. Ruchti, S. LeibundGut-Landmann, J. Heitman, G. Ianiri, A novel mycovirus evokes transcriptional rewiring in the fungus *Malassezia* and stimulates beta interferon production in macrophages. *mBio* **11**, e01534-20 (2020).

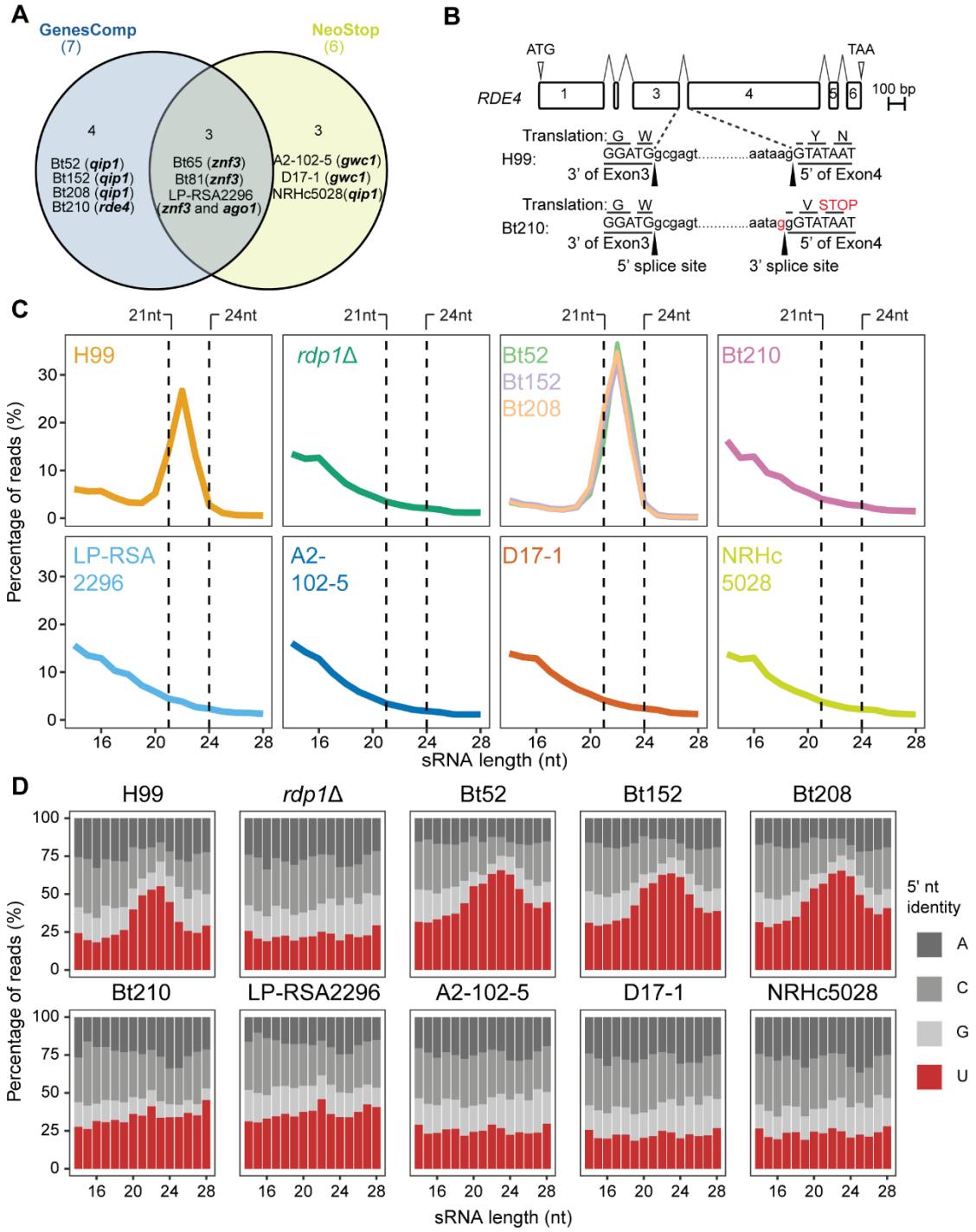
21. V. Yadav, *et al.*, RNAi is a critical determinant of centromere evolution in closely related fungi. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 3108–3113 (2018).
22. E. Iracane, *et al.*, Identification of an active RNAi pathway in *Candida albicans*. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2315926121 (2024).
23. S. Sun, M. A. Coelho, M. David-Palma, S. J. Priest, J. Heitman, The evolution of sexual reproduction and the mating-type Locus: links to pathogenesis of *Cryptococcus* human pathogenic fungi. *Annu. Rev. Genet.* **53**, 417–444 (2019).
24. G. D. Brown, *et al.*, The pathobiology of human fungal infections. *Nat Rev Microbiol* (2024). https://doi.org/10.1038/s41579-024-01062-w.
25. C. A. Desjardins, *et al.*, Population genomics and the evolution of virulence in the fungal pathogen *Cryptococcus neoformans*. *Genome Res.* **27**, 1207–1219 (2017).
26. R. Rajasingham, *et al.*, Global burden of disease of HIV-associated cryptococcal meningitis: an updated analysis. *The Lancet Infectious Diseases* **17**, 873–881 (2017).
27. R. Rajasingham, *et al.*, The global burden of HIV-associated cryptococcal infection in adults in 2020: a modelling analysis. *The Lancet Infectious Diseases* **22**, 1748–1755 (2022).
28. G. Casalini, A. Giacomelli, S. Antinori, The WHO fungal priority pathogens list: a crucial reappraisal to review the prioritisation. *The Lancet Microbe* **5**, 717–724 (2024).
29. M. C. Fisher, D. W. Denning, The WHO fungal priority pathogens list as a game-changer. *Nat Rev Microbiol* **21**, 211–212 (2023).
30. J. N. Day, *et al.*, Combination antifungal therapy for cryptococcal meningitis. *N Engl J Med* **368**, 1291–1302 (2013).
31. K. R. Iyer, N. M. Revie, C. Fu, N. Robbins, L. E. Cowen, Treatment strategies for cryptococcal infection: challenges, advances and future outlook. *Nat Rev Microbiol* **19**, 454–466 (2021).
32. Y. C. Chang, *et al.*, Moderate levels of 5-fluorocytosine cause the emergence of high frequency resistance in cryptococci. *Nat Commun* **12**, 3418 (2021).
33. R. C. May, N. R. H. Stone, D. L. Wiesner, T. Bicanic, K. Nielsen, *Cryptococcus*: from environmental saprophyte to global pathogen. *Nat Rev Microbiol* **14**, 106–117 (2016).
34. E. Sionov, Y. C. Chang, K. J. Kwon-Chung, Azole heteroresistance in *Cryptococcus neoformans*: emergence of resistant clones with chromosomal disomy in the mouse brain during fluconazole treatment. *Antimicrob Agents Chemother* **57**, 5127–5130 (2013).
35. K. J. Boyce, *et al.*, Mismatch repair of DNA replication errors contributes to microevolution in the pathogenic fungus *Cryptococcus neoformans*. *mBio* **8**, e00595-17 (2017).
36. J. Rhodes, *et al.*, A population genomics approach to assessing the genetic basis of within-host microevolution underlying recurrent Cryptococcal meningitis infection. *G3: Genes|Genomes|Genetics* **7**, 1165–1176 (2017).
37. K. J. Boyce, C. Cao, C. Xue, A. Idnurm, A spontaneous mutation in DNA polymerase POL3 during in vitro passaging causes a hypermutator phenotype in *Cryptococcus* species. *DNA Repair* **86**, 102751 (2020).
38. K.-W. Jung, J.-H. Jung, H.-Y. Park, Functional roles of homologous recombination and non-homologous end joining in DNA damage response and microevolution in *Cryptococcus neoformans*. *J. Fungi* **7**, 566 (2021).
39. A. Gusa, *et al.*, Transposon mobilization in the human fungal pathogen *Cryptococcus* is mutagenic during infection and promotes drug resistance in vitro. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 9973–9980 (2020).

40. S. J. Priest, *et al.*, Uncontrolled transposition following RNAi loss causes hypermutation and antifungal drug resistance in clinical isolates of *Cryptococcus neoformans*. *Nat Microbiol* **7**, 1239–1251 (2022).
41. M. Feretzaki, R. B. Billmyre, S. A. Clancey, X. Wang, J. Heitman, Gene network polymorphism illuminates loss and retention of novel RNAi silencing components in the *Cryptococcus* pathogenic species complex. *PLoS Genet* **12**, e1005868 (2016).
42. G. Janbon, *et al.*, Analysis of the genome and transcriptome of *Cryptococcus neoformans var. grubii* reveals complex RNA expression and microevolution leading to virulence attenuation. *PLoS Genet* **10**, e1004261 (2014).
43. P. A. Dumesic, *et al.*, Stalled spliceosomes are a signal for RNAi-mediated genome defense. *Cell* **152**, 957–968 (2013).
44. X. Wang, *et al.*, Sex-induced silencing defends the genome of *Cryptococcus neoformans* via RNAi. *Genes Dev.* **24**, 2566–2582 (2010).
45. M. Feretzaki, J. Heitman, Genetic circuits that govern bisexual and unisexual reproduction in *Cryptococcus neoformans*. *PLoS Genet* **9**, e1003688 (2013).
46. J. E. Burke, *et al.*, A non-dicer RNase III and four other novel factors required for RNAi-mediated transposon suppression in the human pathogenic yeast *Cryptococcus neoformans*. *G3: Genes|Genomes|Genetics* **9**, 2235–2244 (2019).
47. G. Janbon, *et al.*, Characterizing the role of RNA silencing components in *Cryptococcus neoformans*. *Fungal Genetics and Biology* **47**, 1070–1080 (2010).
48. L. M. Iyer, *et al.*, Lineage-specific expansions of TET/JBP genes and a new class of DNA transposons shape fungal genomic and epigenetic landscapes. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1676–1683 (2014).
49. A. Gusa, *et al.*, Genome-wide analysis of heat stress-stimulated transposon mobility in the human fungal pathogen *Cryptococcus deneoformans*. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2209831120 (2023).
50. X. Wang, S. Darwiche, J. Heitman, Sex-induced silencing operates during opposite-sex and unisexual reproduction in *Cryptococcus neoformans*. *Genetics* **193**, 1163–1174 (2013).
51. R. B. Billmyre, S. Applen Clancey, L. X. Li, T. L. Doering, J. Heitman, 5-fluorocytosine resistance is associated with hypermutation and alterations in capsule biosynthesis in *Cryptococcus*. *Nat Commun* **11**, 127 (2020).
52. M. F. Seidl, B. P. H. J. Thomma, Transposable elements direct the coevolution between plants and microbes. *Trends in Genetics* **33**, 842–851 (2017).
53. A. S. Urquhart, N. F. Chong, Y. Yang, A. Idnurm, A large transposable element mediates metal resistance in the fungus *Paecilomyces variotii*. *Current Biology* **32**, 937-950.e5 (2022).
54. E. Gluck-Thaler, *et al.*, Giant *Starship* elements mobilize accessory genes in fungal genomes. *Molecular Biology and Evolution* **39**, msac109 (2022).
55. E. Gluck-Thaler, A. A. Vogan, Systematic identification of cargo-mobilizing genetic elements reveals new dimensions of eukaryotic diversity. *Nucleic Acids Research* **52**, 5496–5513 (2024).
56. A. A. Vogan, *et al.*, The *Enterprise*, a massive transposon carrying *Spok* meiotic drive genes. *Genome Res.* **31**, 789–798 (2021).
57. P. Sephton-Clark, *et al.*, Genomic variation across a clinical *Cryptococcus* population linked to disease outcome. *mBio* **13**, e02626-22 (2022).
58. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

59. E. Garrison, G. Marth, Haplotype-based variant detection from short-read sequencing. [Preprint] (2012). Available at: https://arxiv.org/abs/1207.3907 [Accessed 10 July 2024].
60. A. Shumate, S. L. Salzberg, Liftoff: accurate mapping of gene annotations. *Bioinformatics* **37**, 1639–1643 (2021).
61. P. Rice, I. Longden, A. Bleasby, EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics* **16**, 276–277 (2000).
62. C. Camacho, *et al.*, BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
63. W. R. Pearson, D. J. Lipman, Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2444–2448 (1988).
64. P. P. Chan, T. M. Lowe, "tRNAscan-SE: searching for tRNA genes in genomic sequences" in *Gene Prediction*, Methods in Molecular Biology., M. Kollmar, Ed. (Springer New York, 2019), pp. 1–14.
65. V. Yadav, R. Mohan, S. Sun, J. Heitman, Calcineurin contributes to RNAi-mediated transgene silencing and small interfering RNA production in the human fungal pathogen *Cryptococcus neoformans*. *Genetics* **226**, iyae010 (2024).
66. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet j.* **17**, 10 (2011).
67. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).
68. T. A. Dahlmann, U. Kück, Dicer-dependent biogenesis of small RNAs and evidence for microRNA-like RNAs in the penicillin producing Fungus *Penicillium chrysogenum*. *PLoS ONE* **10**, e0125989 (2015).
69. E. A. Radchenko, R. J. McGinty, A. Y. Aksenova, A. J. Neil, S. M. Mirkin, "Quantitative analysis of the rates for repeat-mediated genome instability in a yeast experimental system" in *Genome Instability*, Methods in Molecular Biology., M. Muzi-Falconi, G. W. Brown, Eds. (Springer New York, 2018), pp. 421–438.
70. Q. Zheng, Methods for comparing mutation rates using fluctuation assay data. *Mutat Res-Fund Mol M* **777**, 20–22 (2015).
71. V. Yadav, S. Sun, M. A. Coelho, J. Heitman, Centromere scission drives chromosome shuffling and reproductive isolation. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 7917–7928 (2020).
72. S. Koren, *et al.*, Canu: scalable and accurate long-read assembly via adaptive *k* -mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).
73. B. J. Walker, *et al.*, Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLoS ONE* **9**, e112963 (2014).
74. D. Peng, R. Tarleton, EuPaGDT: a web tool tailored to design CRISPR guide RNAs for eukaryotic pathogens. *Microbial Genomics* **1** (2015).
75. M. Y. Huang, *et al.*, Short homology-directed repair using optimized Cas9 in the pathogen *Cryptococcus neoformans* enables rapid gene deletion and tagging. *Genetics* **220**, iyab180 (2022).
76. S. D. M. Arras, J. L. Chitty, K. L. Blake, B. L. Schulz, J. A. Fraser, A genomic safe haven for mutant complementation in *Cryptococcus neoformans*. *PLoS ONE* **10**, e0122916 (2015).
77. Y. Fan, X. Lin, Multiple applications of a transient CRISPR-Cas9 Coupled with Electroporation (TRACE) system in the *Cryptococcus neoformans* species complex. *Genetics* **208**, 1357–1372 (2018).

985 78. S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: molecular evolutionary genetics
986 analysis across computing platforms. *Molecular Biology and Evolution* **35**, 1547–1549
987 (2018).
988 79. S. Sun, S. J. Priest, J. Heitman, *Cryptococcus neoformans* mating and genetic crosses. *Curr.*
989 *Protoc. Microbiol.* **53**, e75 (2019).
990 80. W. J. Kent, BLAT —The BLAST -like alignment tool. *Genome Res.* **12**, 656–664 (2002).
991 81. H. Li, Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–
992 3100 (2018).
993 82. B. S. Pedersen, A. R. Quinlan, Mosdepth: quick coverage calculation for genomes and
994 exomes. *Bioinformatics* **34**, 867–868 (2018).
995

996 **Figures and Tables**



997

998 **Figure 1. Identification and confirmation of RNAi loss-of-function mutations in *C.***
999 ***neoformans* natural isolates.**

1000 (A) Venn diagram illustrating the 11 RNAi loss-of-function mutations initially predicted by the two
1001 independent pipelines, GenesComp and NeoStop, among *C. neoformans* natural isolates in the

Strain Diversity Collection. All of the 11 predicted mutations have been confirmed with Sanger sequencing. (B) *RDE4* gene structures in H99 and Bt210 are illustrated. The *RDE4* splice site mutation (A>G) found in Bt210 is highlighted in red. (C) Size distribution of sRNA reads from each of the indicated strain. Dashed vertical lines represent the 21 to 24 nt size range, corresponding to the RNAi-mediated sRNA population. (D) Proportion of 5'-nucleotide identity in sRNA reads. All the sRNA reads were mapped to the H99 genome.

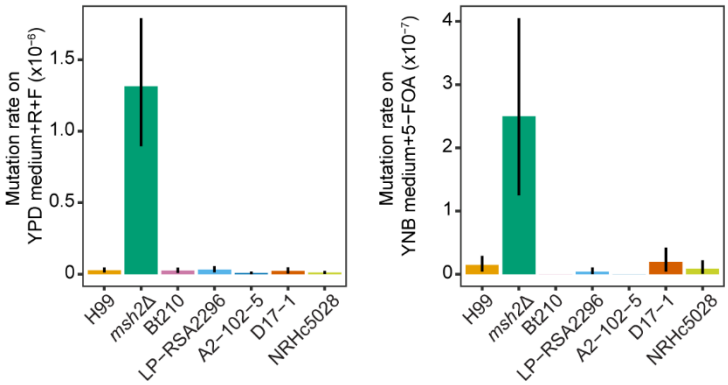


Figure 2. Five novel RNAi-loss isolates do not display a hypermutator phenotype. Mutation rates of five newly identified RNAi-loss isolates when grown on YPD medium containing rapamycin and FK506 (R+F) (left) or YNB medium containing 5-FOA (right), with H99 wildtype and *msh2*Δ strains serving as the negative and positive controls, respectively. The error bars represent 95% confidence intervals.

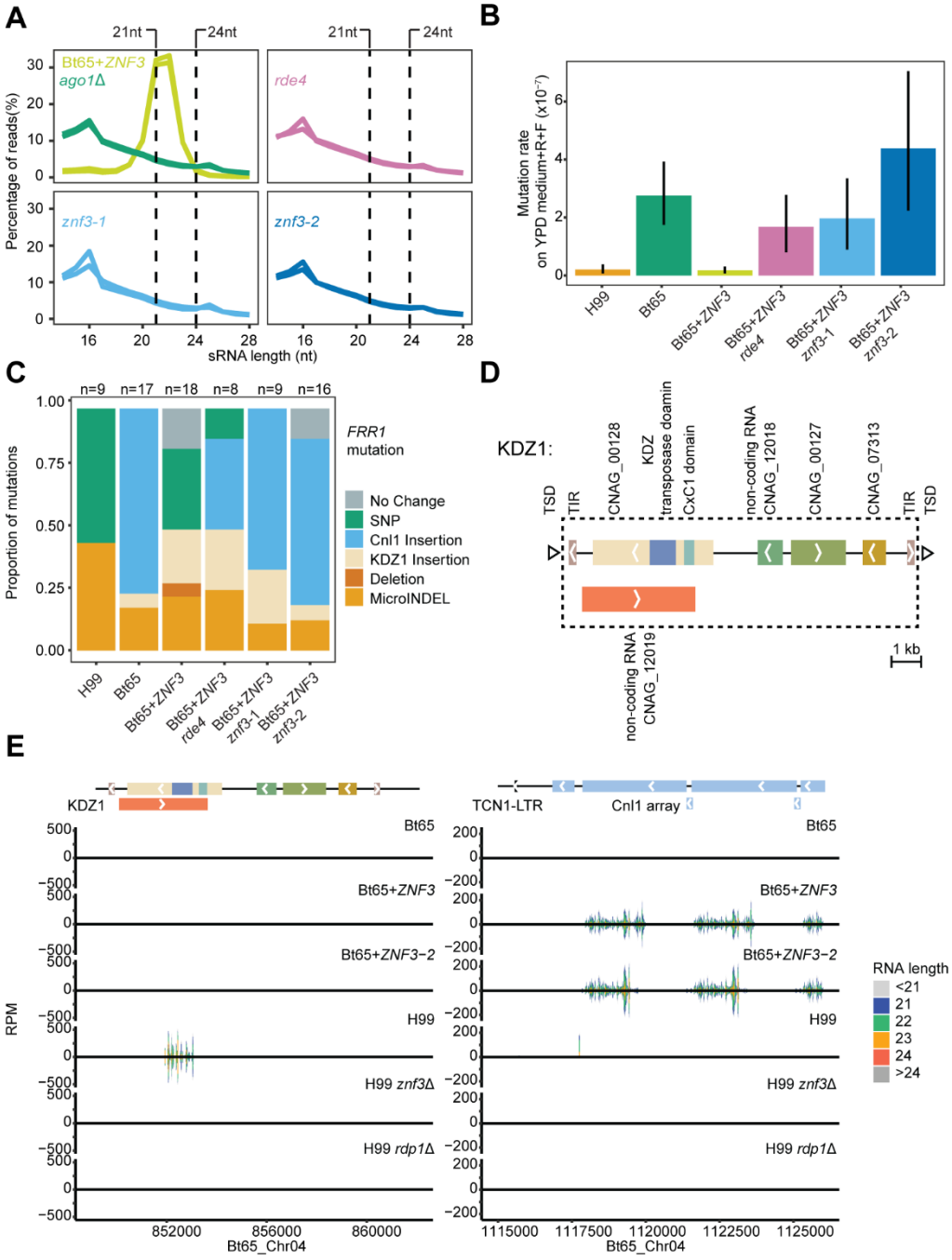


Figure 3. Engineered RNAi loss-of-function mutations result in transposition leading to hypermutation.

(A) Size distribution of sRNA reads from each of the respective strains (Bt65+*ZNF3*, Bt65+*ZNF3 ago1*Δ, Bt65+*ZNF3 rde4*, Bt65+*ZNF3 znf3-1*, and Bt65+*ZNF3 znf3-2*). Dashed vertical lines represent the 21 to 24 nt size range. Two biological replicates of each strain are plotted. All the sRNA reads were mapped to the Bt65 genome. (B) Mutation rates of Bt65+*ZNF3*, Bt65+*ZNF3 rde4* and two independent Bt65+*ZNF3 znf3* allele exchange strains when grown on YPD medium containing rapamycin and FK506, with H99 and Bt65 serving as negative and positive controls, respectively. The error bars represent 95% confidence intervals. (C) Mutational spectra in the *FRR1* gene of independent rapamycin+FK506 resistant colonies from each indicated strain. R+F resistant

colonies without mutations detected at the *FRR1* locus are labeled as 'No change'. (D) Diagram illustrating the structure of a novel gigantic DNA transposon in *C. neoformans*, KDZ1. Genes and non-coding RNAs are shown by solid blocks, with arrows indicating the orientation. Tags starting with "CNAG" correspond to the annotated number in the *C. neoformans* H99 reference genome. TIR: Terminal Inverted Repeat; TSD: Target Site Duplication (drawn out of scale). (E) Mapping of small RNA reads from Bt65, Bt65+*ZNF3*, Bt65+*ZNF3-2*, H99, H99 *znf3*Δ and H99 *rdp1*Δ aligning to a KDZ1 (left) and a Cnl1 array (right), both of which are located on chromosome 4 of Bt65. Shown on the y-axis are numbers of normalized reads (RPM, Reads Per Million on forward strand (positive value) or reverse strand (negative value)), representing normalized coverage counts at each position. Different lengths of small RNA reads are labeled as indicated by the color.

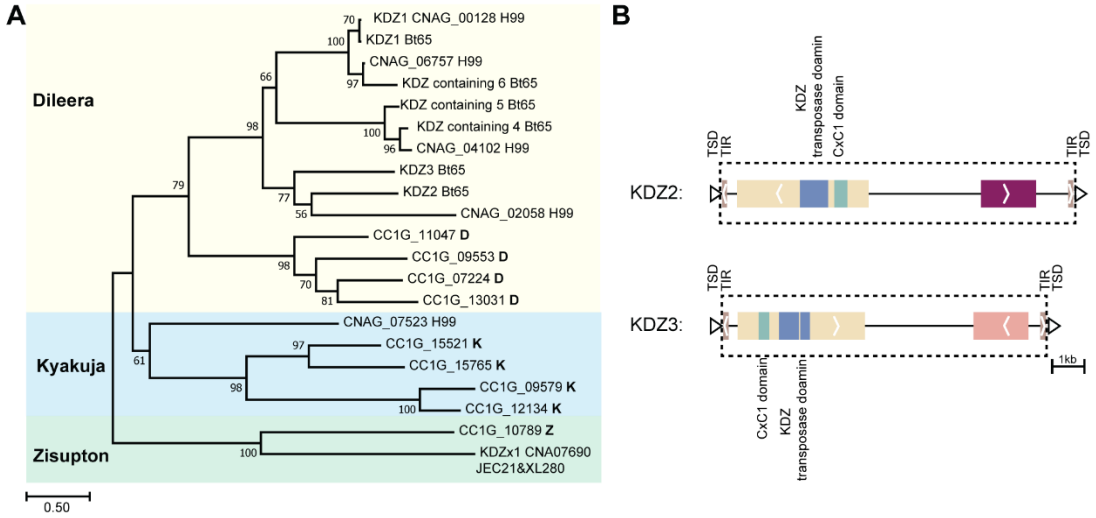


Figure 4. Evolutionary diversity of the giant KDZ DNA transposons in *Cryptococcus*. (A) Maximum-likelihood phylogeny of KDZ domain-containing proteins from Bt65 and H99, as well as representative KDZ domain-containing proteins that have been characterized from *Coprinopsis cinerea* (CC1G_). The tree was evaluated with 1000 bootstrap replicates. Shades of different colors highlight clades containing the *Coprinopsis cinerea* Dileera (D), Kyakuja (K), Zisupton (Z) proteins, respectively. (B) Diagram illustrating the structure of two novel gigantic DNA transposons in *C. neoformans*, KDZ2 and KDZ3. Genes are shown by solid blocks, with arrows inside indicating the orientation. TIR: Terminal Inverted Repeat; TSD: Target Site Duplication (drawn out of scale).

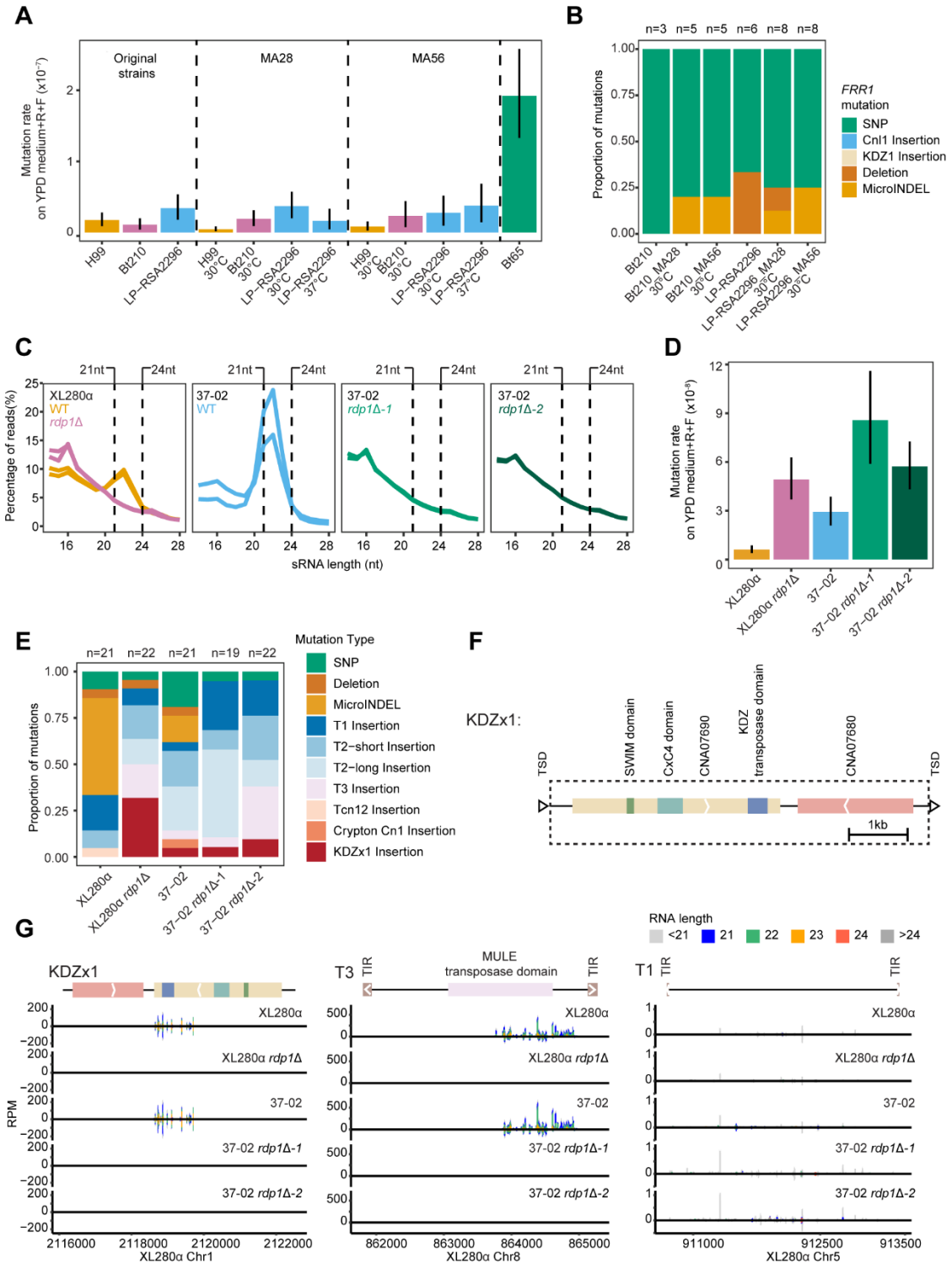


Figure 5. Dynamic outcomes of short-term experimental evolution.

(A) Mutation rates of the progenitor strains H99, Bt210, LP-RSA2296, and their derived strains after 28 (MA28) or 56 (MA56) *in vitro* passages, when grown on YPD medium containing rapamycin + FK506 (R+F), with the hypermutator Bt65 serving as positive control. The error bars represent 95%

confidence intervals. (B) Mutational spectra in the *FRR1* gene of the independent R+F resistant colonies from each indicated strain. (C) Size distribution of sRNA reads from each indicated strain. sRNA reads from two biologically independent replicates are plotted. Dashed vertical lines represent the 21 to 24 nt size range. All the sRNA reads were mapped to the XL280α genome. (D) Mutation rates of XL280α, XL280α *rdp1*Δ, 37-02, 37-02 *rdp1*Δ-*1*, and 37-02 *rdp1*Δ-*2* when grown on YPD medium containing R+F are shown. (E) Mutational spectra in the *FRR1* gene of independent R+F resistant colonies from each indicated *C. deneoformans* strain. (F) Diagram illustrating the structure of a novel DNA transposon in *C. deneoformans*, KDZx1. Solid colored blocks represent open reading frames, with arrows inside indicating the orientation. CNA07680 and CNA07690 correspond to two proteins in the *C. deneoformans* JEC21 reference genome. TIR: Terminal Inverted Repeat; TSD: Target Site Duplication (drawn out of scale). (G) Mapping of small RNA reads from XL280α, XL280α *rdp1*Δ, 37-02, 37-02 *rdp1*Δ-*1*, and 37-02 *rdp1*Δ-*2* to a KDZx1 (left), a DNA transposon T3 (middle, with MULE transposase domain and terminal inverted repeats), and a DNA transposon T1 (right, only known with terminal inverted repeats), which are located on different chromosomes of XL280α. Shown on the y-axis are normalized reads (RPM, Reads Per Million on forward strand (positive value) or reverse strand (negative value)), representing normalized coverage counts at each position. Different lengths of small RNA reads are labelled as indicated by the color.

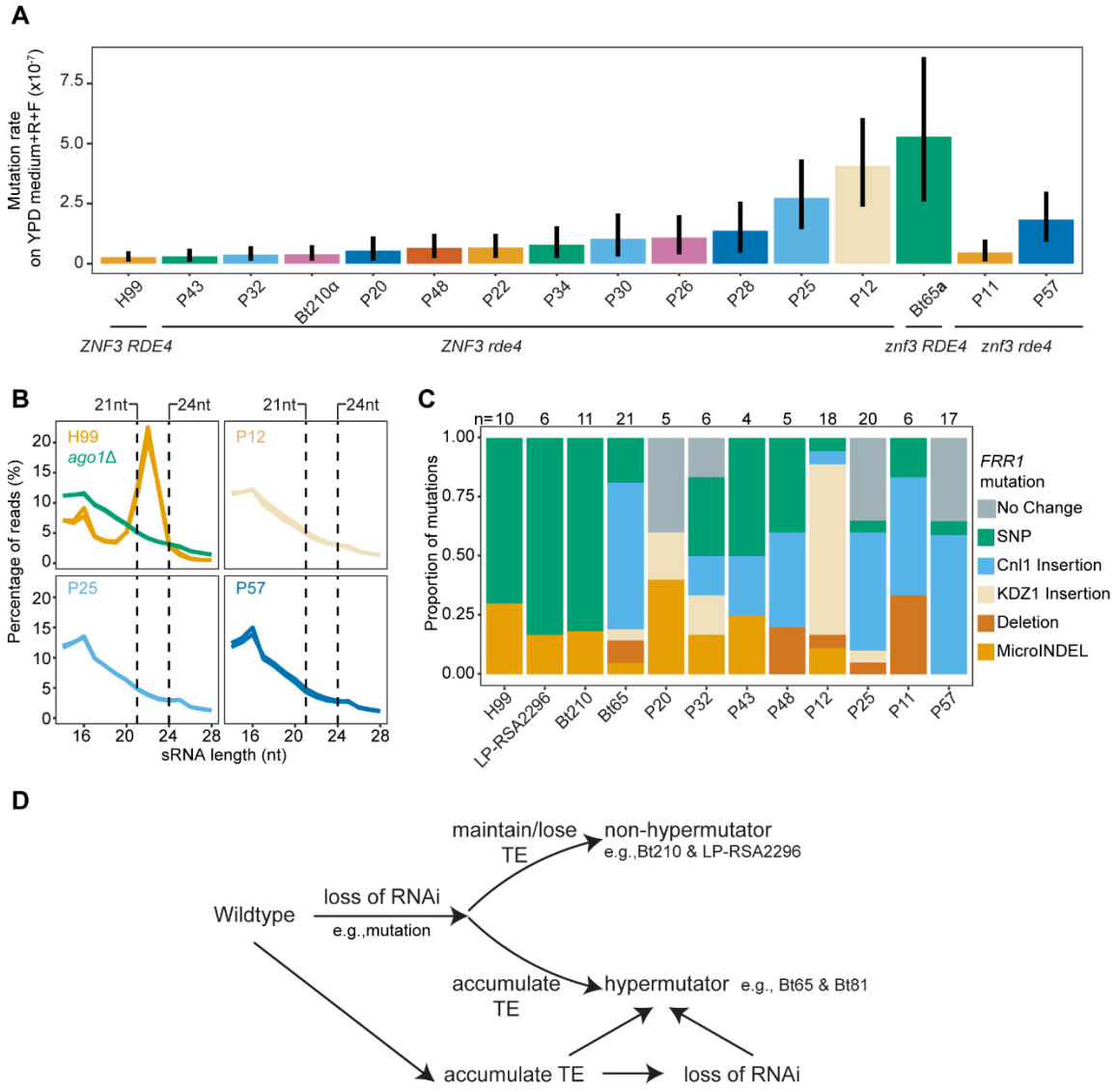


Figure 6. Transposition causes hypermutation after sexual reproduction.
(A) Mutation rates of 13 F1 progeny dissected from the cross between natural isolates Bt65**a** and Bt210α, with the two parents (Bt65**a** and Bt210α) and *C. neoformans* reference strain H99 serving as controls, when grown on YPD medium containing rapamycin + FK506 (R+F). The error bars represent 95% confidence intervals. Genotypes of *RDE4* and *ZNF3* are shown for each strain. (B) Size distribution of sRNA reads from each indicated strain with two independent biological replicates for each strain. Dashed vertical lines represent the 21 to 24 nt size range. All the sRNA reads were mapped to the Bt65 genome. (C) Mutational spectra in *FRR1* in independent rapamycin + FK506 (R+F) resistant colonies from each indicated strain. R+F resistant colonies without mutations detected at the *FRR1* locus are labelled as 'No Change'. (D) Current model for paths to hypermutation. After RNAi loss, isolates (e.g., Bt65 and Bt81) underwent transposon accumulation and evolved into hypermutators. In contrast, some other RNAi-deficient isolates (e.g., Bt210 and LP-RSA2296) might maintain or lose transposons and remain non-hypermutators. Additionally, transposon accumulation might occur without RNAi loss and cause a hypermutator phenotype with or without subsequent RNAi loss.

Table 1. Natural *C. neoformans* isolates with putative loss-of-function mutations in genes required for RNAi-mediated gene silencing.

Strain	Pipeline	Gene with mutation [1]	Mutation type	Putative mutation impact [2]
Bt84	GenesComp	*QIP1*	Nonsense; Exon 1	Protein truncated to 30 aa (4% WT length)
Bt65	NeoStop, GenesComp	*ZNF3*	Nonsense; Exon 1	Protein truncated to 96 aa (6% WT length)
Bt81	NeoStop, GenesComp	*ZNF3*	Nonsense; Exon 1	Protein truncated to 96 aa (6% WT length)
LP-RSA2296	NeoStop, GenesComp	*ZNF3*	Nonsense; Exon 1	Protein truncated to 176 aa (11% WT length)
Bt52	GenesComp	*QIP1*	Start-loss; Exon 1	Missing first 15 aa; truncated to 732 aa (98% WT length)
Bt152	GenesComp	*QIP1*	Start-loss; Exon 1	Missing first 15 aa; truncated to 732 aa (98% WT length)
Bt208	GenesComp	*QIP1*	Start-loss; Exon 1	Missing first 15 aa; truncated to 732 aa (98% WT length)
LP-RSA2296	NeoStop, GenesComp	*AGO1*	Frameshift; Exon 10	Deletion of 9 nt and 1 nt in Exon 10; truncated to 459 aa (50% WT length)
Bt210	GenesComp	*RDE4*	Splice-site; Intron 3	Splicing error in Intron 3 leads to frameshift; protein truncated to 207 aa (43% WT length)
A2-102-5	NeoStop	*GWC1*	Frameshift; Exon 1	Protein truncated to 47 aa (8% WT length)
D17-1	NeoStop	*GWC1*	Frameshift; Exon 1	Protein truncated to 47 aa (8% WT length)
NRHc5028. ENR.STOR	NeoStop	*QIP1*	Nonsense; Exon 4	Protein truncated to 250 aa (33% length)

[1]: *ZNF3*: CNAG_02700; *QIP1*: CNAG_01423; *AGO1:* CNAG_04609; *RDE4*: CNAG_01157; *GWC1*, CNAG_06486
[2]: WT: wild-type; aa: amino acids

Table 2. Full-length Cnl1 copy numbers.

Isolate	Reference-based assembly	Short-read-based estimation[1]	Hypermutator when selected on R+F?
H99	0	0	No
Bt65	108	178	Yes
Bt81	89	102	Yes
Bt210	0	0	No
LP-RSA2296	1	4	No
A2-102-5	NA	0	No
D17-1	NA	1	No
NRHc5028	NA	0	No

[1]: Whole-genome sequence data from Desjardins et al., 2017 were used for estimation (25).

NA: not applicable