Rampant transposition following RNAi loss causes hypermutation and antifungal drug resistance in clinical isolates of a human fungal pathogen

Shelby Priest (shelbyjpriest@gmail.com)
Duke University https://orcid.org/0000-0003-4577-167X
Vikas Yadav
Duke University Medical Center https://orcid.org/0000-0003-2650-9035
Cullen Roth
Duke University
Tim Dahlmann
Ruhr-Universität Bochum
Ulrich Kueck
Allgemeine und Molekulare Botanik, Ruhr-Universität Bochum
Paul Magwene
Duke University
Joseph Heitman
Duke Medical Center

Article

Keywords: Cryptococcus neoformans, antifungal drug resistance, non-LTR Cnl1 retrotransposon

Posted Date: August 20th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-804218/v1

Version of Record: A version of this preprint was published at Nature Microbiology on August 2nd, 2022.
See the published version at https://doi.org/10.1038/s41564-022-01183-z.

Abstract

Microorganisms survive and compete within their environmental niches and avoid evolutionary stagnation by stochastically acquiring mutations that enhance fitness. Although increased mutation rates are often deleterious in multicellular organisms, hypermutation can be beneficial for microbes in the context of strong selective pressures. To explore how hypermutation arises in nature and elucidate its consequences, we employed a collection of 387 sequenced clinical and environmental isolates of *Cryptococcus neoformans*. This fungal pathogen is responsible for ~ 15% of annual AIDS-related deaths and is associated with high mortality rates, attributable to a dearth of antifungal drugs and increasing drug resistance. Isolates were screened for the ability to rapidly acquire antifungal drug resistance, and two robust hypermutators were identified. Insertion of the non-LTR Cnl1 retrotransposon was found to be responsible for the majority of drug-resistant isolates. Long-read whole-genome sequencing revealed both hypermutator genomes have two unique features: 1) hundreds of Cnl1 copies organized in subtelomeric arrays on both ends of almost all chromosomes, and 2) a nonsense mutation in the first exon of *ZNF3*, a gene encoding an RNAi component involved in silencing transposons. Quantitative trait locus mapping identified a significant genetic locus associated with hypermutation that includes the mutant *znf3* allele, and CRISPR-mediated genome editing of the *znf3* single-base pair nonsense mutation abolished the hypermutation phenotype and restored siRNA production. In sum, hypermutation and drug resistance in these isolates results from loss of RNAi combined with subsequent accumulation of a large genomic burden of a novel transposable element in *C. neoformans*.

Introduction

Stochastic mutations and genomic rearrangements provide variation in populations for natural selection to act upon and enable evolution. However, genetic changes are a double-edged sword: too little variation can lead to evolutionary stagnation, while too much variation can lead to a lethal accumulation of deleterious mutations. Hypermutation, one extreme of this mutational spectrum, can lead to adaptation, disease, or eventual extinction if left unchecked.

Microbes are known to adopt highly mutable states that would normally be viewed as deleterious from the perspective of multicellular organisms. Studies have found that microorganisms with defects in pathways associated with maintaining genomic integrity, such as those involved in chromosome stability, DNA mismatch repair, DNA damage repair, and cell cycle checkpoints associated with recognizing DNA damage, accelerate adaptation to environmental stressors[1–3]. These defects can be beneficial in the short term, yet deleterious in the long term as mutations continue to accumulate. Defects in DNA mismatch repair resulting in increased mutation rates have been reported in fungi, including the model yeast *Saccharomyces cerevisiae*, the human pathogen *Candida glabrata*, in an outbreak strain of *Cryptococcus deuterogattii*, and in several clinical isolates of the model basidiomycete human fungal pathogen *Cryptococcus neoformans*[4–10]. Genomic stability in pathogenic *Cryptococcus* species is also significantly affected by karyotypic changes and transposable elements, both of which can mediate antifungal drug resistance[11–14].

Cryptococcus is an environmentally ubiquitous haploid basidiomycete and facultative human pathogen[15]. Approximately 95% of cryptococcal infections are attributable to the serotype A group, *C. neoformans* var. *grubii*, now known as *C. neoformans*, which is divided into four lineages: VNI, VNII, VNBI, and VNBII[16–18]. This species infects immunocompromised individuals and accounts for ~ 15% of HIV/AIDS-related deaths[19]. The threat of cryptococcal infections is exacerbated because the arsenal of antifungal drugs is limited. Amphotericin B, a fungicidal polyene, is often used in combination with 5-flucytosine (5-FC), an antimetabolite, to treat cryptococcal infections[20]. Unfortunately, amphotericin B and 5-FC have undesirable side effects, and 5-FC monotherapy frequently leads to resistance[21,22]. Fluconazole is used to treat asymptomatic patients with isolated cryptococcal antigenemia, those with disease limited to lung nodules or central nervous system infections after clearance of cerebrospinal fluid cultures, or for chronic maintenance therapy[20]. However, *C. neoformans* frequently develops resistance to fluconazole via aneuploidy or mutations in the sterol biosynthesis pathway, contributing to recurrent infections[11,23–25]. The limited number of drugs available to treat cryptococcosis, prevalence of resistance and recurrent infections, and difficulty in developing novel antifungal therapies combine to make *C. neoformans* drug resistance an important clinical problem.

Transposons in the C. *neoformans* H99 reference strain and the sister species *Cryptococcus deneoformans* JEC21 reference strain have been characterized[13,26,27]. The genomes of these species encode many retrotransposons with and without long-terminal repeats, known as LTR retrotransposons and non-LTR retrotransposons, respectively, which move via a copy-and-paste mechanism, allowing them to proliferate throughout the genome if unchecked. The most well-characterized *Cryptococcus* LTR-retrotransposons are Tcn1 through Tcn6, which are primarily located in centromeric regions[28]. The *C. deneoformans* JEC21 genome also encodes three types of DNA transposons (T1, T2, and T3), as well as ~ 25 copies of the non-LTR retrotransposon Cnl1 (*C. neoformans* LINE-1), which is thought to associate with telomeric repeat sequences[13,14,29]. In the *C. neoformans* H99 genome, there are no full-length copies of Cnl1, and DNA transposons are rare[13].

Studies have illustrated that transposon silencing in *Cryptococcus* is governed by RNAi through three primary lines of evidence: 1) siRNAs map predominantly to transposable elements, 2) RNAi mutants show increased transposon expression, and 3) spliceosomes stall on transposable element transcripts at an unusually high rate, triggering RNAi[13,30–34]. Other mechanisms thought to regulate *Cryptococcus* transposons include 5-methylcytosine DNA methylation[28,35,36] and heterochromatic marks[37]. Interestingly, the outbreak species *C. deuterogattii* is RNAi deficient because many genes encoding RNAi components are severely truncated or absent entirely[32]. This loss of RNAi has been shown to be associated with loss of all functional transposable elements, consequently shorter centromeres, and higher rates of intron retention[28,38].

Here, we identified two clinical, hypermutator *C. neoformans* isolates with significantly increased mutation rates on antifungal drug media. The majority of drug resistance in these two strains is mediated by insertions of the Cnl1 transposon into genes whose mutation confers resistance. Genetic

backcrossing, quantitative trait loci mapping, and CRISPR-mediated gene editing all confirmed that a nonsense mutation in the RNAi component *ZNF3*, resulting in loss of RNAi, is the cause of hypermutation in these two strains. Small RNA sequencing confirmed the role of Znf3 in silencing Cn1l, and whole-genome sequencing revealed both hypermutator genomes encode > 800 copies or fragments of Cnl1. This is the first time full-length copies of Cnl1 have been identified in *C. neoformans*, and the massive Cnl1 burden in these hypermutators is substantially higher than previously observed in any other *Cryptococcus* strain. Our results demonstrate the hypermutator phenotype described here is attributable to loss of RNAi, allowing rampant transposition of Cnl1. These transposition events lead to Cnl1 accumulation at subtelomeres and movement to novel genomic locations, which can result in drug resistance.

Results

Identification of two clinical, hypermutator *C. neoformans* isolates

To identify natural isolates of *C. neoformans* with increased mutation rates, we screened strains in the Strain Diversity Collection (SDC) for an increased ability to produce colonies resistant to various classes of antifungal drugs. The SDC contains 387 strains from all four *C. neoformans* lineages (VNI, VNII, VNBI and VNBII), including geographically diverse clinical and environmental isolates of both mating types. For each isolate in this collection, whole-genome sequencing (WGS) and phylogenetic relationships are available[18]. To screen isolates for increased mutation rates in a relatively high-throughput manner, strains were grown in liquid cultures, swabbed onto agar plates supplemented with either 5-fluorocytosine (5-FC) or a combination of FK506 and rapamycin (immunosuppressants that bind FKBP12 to form complexes that inhibit activity of calcineurin and TOR, respectively), and qualitatively evaluated for their ability to generate resistant colonies (Figure 1A)[39–42]. Strains that produced more spontaneously resistant colonies on average than the control strain H99 were categorized as hypermutator candidates. We screened 186 strains and identified 40 hypermutator candidates (Table S1). Interestingly, all but one of fourteen (93%) environmental isolates screened were identified as hypermutator candidates (as compared to only 16% (27/170) clinical isolates, *p*-value < 0.001, Fisher's exact test). Two previously identified hypermutator strains with mismatch repair defects, C23 and C45, were identified as hypermutator candidates as well[9].

We chose to focus on two clinical strains, Bt65 and Bt81, that produced the most rapamycin + FK506-resistant ($R+F^R$) colonies (Figure 1A). Bt65 and Bt81 are both VNBII *MAT***a** strains isolated from different HIV-positive individuals in Botswana[18,43]. To quantify the mutation rates of Bt65 and Bt81, we performed fluctuation assays on YPD + rapamycin + FK506, YNB + 5-fluoroorotic acid (5-FOA), and YNB + 5-FC media. Both Bt65 and Bt81 produced significantly higher mutation rates on YPD + rapamycin + FK506 compared to H99 as well as eleven of the most closely phylogenetically related strains (Figure 1B). On 5-FC, only the strain NRHc5014.ENR and the KN99α *msh2*Δ positive control had significantly higher

mutation rates compared to H99 (Figure S1A); on 5-FOA medium (Figure S1B), only KN99α *msh2*Δ produced a significantly higher mutation rate.

A recent study illustrated how incubation at an elevated temperature of 37°C results in increased mutation rates due to transposon mobilization in the closely related species *C. deneoformans*[14]. To determine if elevated temperature contributed to hypermutation in Bt65 and Bt81, we concurrently grew these strains as well as wild-type H99 and *msh2*Δ, *ago1*Δ, and *rdp1*Δ genetic deletion mutants in the H99 genetic background overnight at 30°C and 37°C, and then performed a fluctuation analysis on YPD + rapamycin + FK506 medium. Fluctuation analysis revealed H99, Bt65, and Bt81 had lower mutation rates when grown overnight at 37°C compared to growth at 30°C (Figure S2). Interestingly, all of the genetic deletion mutants showed increased mutation rates after growth at 37°C (Figure S2). These results suggest that unlike *C. deneoformans*, growth at 37°C reduces mutation rates in wild-type *C. neoformans* strains and does not contribute to or exacerbate hypermutation in Bt65 and Bt81.

Characterization of mutation spectra in *C. neoformans* hypermutator strains

After quantifying the mutation rates of Bt65 and Bt81 strains on media with various antifungal drugs, we investigated the types of mutations conferring resistance to the combination of rapamycin and FK506. PCR amplification of the *FRR1* gene (which encodes FKBP12, the shared target of rapamycin and FK506, and the only gene in which mutations confer resistance to both FK506 and rapamycin) followed by gel electrophoresis revealed the expected wild-type PCR product size (~1.2 kb) for all but two (35/37) H99 R+F^R colonies; the remaining two produced PCR products smaller than expected, indicative of deletions (Figure S3A). In contrast, large insertions of various sizes were observed in the majority of the Bt65 and Bt81 R+F^R colonies (62/77 and 36/37 independent colonies, respectively) (Figure 1C, S3B, C). Only one resistant colony derived from a non-hypermutator strain, Bt84, (1/10 independent colonies) had an insertion in *FRR1*. No insertions in *FRR1* were observed in any of the other closely related or control strains.

We subsequently sequenced *FRR1* in H99, Bt65, Bt81, and Bt84 R+F^R colonies to determine the genetic changes responsible for the larger PCR products (Figure 1C). In 37 R+F^R colonies of the H99 control strain, SNPs in *FRR1* were largely responsible for resistance (57%, 21/37 colonies), while resistance in the remaining colonies was attributable to small insertions/deletions (microINDELs; 38%, 14/37) or large deletions (5%, 2/37). Conversely, in the hypermutator isolate Bt65, insertions of the non-LTR retrotransposon Cnl1 were responsible for the majority of rapamycin + FK506 resistance (77.4%, 24/31). Rapamycin + FK506 resistance in the remaining Bt65 colonies was either due to SNPs (19.4%, 6/31) or microINDELs (3.2%, 1/31). In all sequenced PCR products from R+F^R colonies of Bt81, Cnl1 insertions were responsible for resistance (27/27 colonies). Cnl1 insertions in Bt65 and Bt81 ranged from 54 bp to ~3500 bp, and this range in transposon sizes is a common characteristic of non-LTR retrotransposons. The single insertion observed in *FRR1* in the R+F^R colony of Bt84 had no homology with any annotated

Cryptococcus transposons but was identified as a repetitive element by RepeatMasker and shared minor homology with a Copia-58 BG-I transposable element.

The 5-FC- and 5-FOA-resistant colonies of Bt65, Bt81, and H99 were similarly characterized to determine the sources of resistance to antifungal drugs with different mechanisms of action. Resistance to 5-FOA is conferred by mutations in the *URA3* or *URA5* genes of the uracil biosynthesis pathway[44,45]. Among the subset of H99, Bt65, and Bt81 5-FOA^R colonies sequenced, mutations were only identified in *URA5*. In almost all colonies, resistance was conferred by SNPs or INDELs, and only one Cnl1 insertion event was identified in a Bt81 5- FOA^R colony (Figure S3D). We also PCR amplified genes in which mutations are known to confer resistance to 5-FC, including *FUR1* and *UXS1*[46]. Only two of the 5-FC^R isolates analyzed from H99, Bt65, and Bt81 produced a PCR product larger than expected for all assessed genes, and subsequent sequencing revealed a Cnl1 insertion in *FUR1* in two independent 5-FC^R Bt81 isolates (Figure S3E).

Analysis of the Cnl1 insertions observed to confer resistance to rapamycin + FK506, 5-FC, and 5-FOA revealed Cnl1 preferentially inserts at guanine- and cytosine-rich regions of target genes, a known property of this element[27]. Target site duplication sequences flanking Cnl1 insertions were not present in many instances, but when present, ranged from 1 to 12 bp in length. Cnl1 insertions ranged greatly in size, from 25-bp fragments to full-length Cnl1 copies (3,494 bp). The smallest Cnl1 insertion (25 bp) was followed immediately by a 59-bp deletion in *FRR1*. Cnl insertions in the *FRR1* gene were observed in both orientations (5’ to 3’ and 3’ to 5’ relative to *FRR1* transcription). Of the 51 characterized Cnl1 *FRR1* insertions, 27 were in the 5’ UTR, and all but one were in the same orientation as *FRR1* transcription, 23 were in exons (7 oriented 5’ to 3’, 16 oriented 3’ to 5’), and one insertion was in an intron of *FRR1* in the 3’ to 5’ orientation, potentially disrupting splicing or transcription.

QTL mapping identifies loci that significantly contribute to the hypermutator phenotype

To determine the genetic cause of the hypermutator phenotype and rampant transposition in Bt65 and Bt81 and to determine the genetic consequences of this phenotype, we conducted quantitative trait locus (QTL) mapping. For this purpose, a total of 165 basidiospores were dissected from a genetic cross between Bt65 *MAT***a** and an H99 *crg1*Δ *MAT*α mutant with an enhanced mating phenotype, and 47 F_1 progeny germinated (28%).

Twenty-eight Bt65**a** x H99α F_1 progeny were selected for fluctuation analysis and whole-genome sequencing. Aligning the paired-read Illumina sequencing data from the 28 F_1 progeny identified 215,411 bi-allelic SNPs that were utilized for QTL mapping. For 24 of the segregants as well as for the Bt65 and H99 *crg1*Δ parental strains, the mutation rate on rapamycin + FK506 medium served as the phenotype for association tests (Figure 2A). Across the 14 chromosomes and bi-allelic SNP sites, two QTL with large effect (heritability = 64%) were identified at approximately 919-1,120 kb on Chromosome 3 and 987-1,193

kb on Chromosome 11 (Figure 2B and S4). Analysis of these QTLs revealed that the SNPs in each QTL were co-segregating and that they shared the same distributions of phenotype scores (Figure S4 and S5). The borders of the QTL spanning Chromosomes 3 and 11 were determined by calculating 95% confidence intervals and examining recombination break points along each chromosome. Interestingly, these two QTLs span the chromosomal translocation between Chromosomes 3 and 11 that is unique to H99 (Figure S4, S5, S6). We thus treated these QTLs as the same QTL for subsequent analysis.

Within the QTL there are a total of 108 and 85 genes along Chromosome 3 and Chromosome 11, respectively, and for 82 and 77 of these genes (respectively), the published annotation and SNP data was used to characterize differences in predicted protein sequence and expected protein lengths between the H99 and Bt65 parental strains (Figure S4 and Table S2).

Among these, 71 and 60 genes along Chromosome 3 and Chromosome 11, respectively, have at least one predicted nonsynonymous change in protein sequence, seven of which harbor a predicted nonsense (i.e. stop-gain) or stop-loss mutation. One of these genes is *ZNF3* (CNAG_02700), which encodes a C2H2 type zinc finger protein with three zinc finger domains. Znf3 was previously identified as an RNAi silencing component that localizes to P-bodies and whose mutation results in increased expression of transposable elements[32,33]. *ZNF3* is located on Chromosome 3 and has a SNP – C to T – within the first exon in the Bt65 genetic background, which is predicted to cause a nonsense mutation, severely truncating Znf3 from 1,561 amino acids to only 96 amino acids (Figure 2C). In addition, this nonsense mutation may also result in nonsense-mediated mRNA decay of the mutant *znf3* mRNA. Based on the publicly available whole-genome sequencing of all isolates in the SDC, the *znf3* nonsense mutation in exon 1 is unique to Bt65 and Bt81 and not present in any other strain. Another gene of known function within the QTL encodes a long-chain acyl-CoA synthetase (CNAG_01836, Chromosome 11) and a SNP – G to A – within the last exon of this gene is predicted to cause an early nonsense mutation in the Bt65 background (Figure S7). Given the dramatic difference in the predicted protein length of *ZNF3* between the H99 and Bt65 parental alleles (relative to other genes in this QTL with predicted stop-loss or nonsense mutations), and previous studies demonstrating the role of Znf3 in RNAi and transposon silencing, we hypothesized *ZNF3* could be the quantitative trait gene (QTG) and the SNP leading to the predicted stop gain in the first exon could be the quantitative trait nucleotide (QTN) underlying the hypermutation phenotype[32,33].

Few Bt81 F_1 progeny display a hypermutator phenotype

In addition to generating and analyzing Bt65**a** x H99α *crg1*Δ F_1 progeny, 42 F_1 progeny were derived from a genetic cross between the other hypermutator strain, Bt81**a,** and H99α *crg1*Δ. The *ZNF3* alleles of all 42 Bt81 F_1 progeny were sequenced to determine whether they had inherited the non-functional *znf3* allele from Bt81 or the functional *ZNF3* allele from H99. Of the 42 F_1 progeny, only four

inherited the mutant *znf3* allele from Bt81, a significantly lower number than would be expected based on Mendelian inheritance patterns (chi-square test, *p*-value < 0.01). The four progeny with non-functional Bt81 *znf3* alleles had the highest mutation rates of 18 F_1 progeny that were analyzed (Figure S8). The mutation rates for three of the four *znf3* progeny, however, were not significantly higher than the Bt81 F_1 progeny with functional the *ZNF3* allele and were not as high as would be expected based on the results from the Bt65**a** x H99α F_1 progeny (Figure 2A).

Cnl1 elements are organized into subtelomeric arrays in hypermutator genomes

For all strains in the SDC, including Bt65 and Bt81, only short-read WGS data was available[18]. Because of the known difficulties in assembling repetitive elements, such as Cnl1, with short-read sequencing data, we conducted long-read WGS with the Oxford Nanopore Technologies MinION sequencing platform to generate more complete assemblies for Bt65, Bt81, and two of the most closely phylogenetically related non-hypermutator strains, Bt89 and Bt133. With the long-read sequencing data, we were able to assemble chromosome-level genomes for all four strains. In the assemblies, we observed the known chromosomal translocation between Chromosomes 3 and 11 unique to H99[47] and identified a translocation between Chromosomes 1 and 13 unique to Bt65 and Bt81 (Figure S6). These two gross chromosomal rearrangements explain the relatively low germination frequency (28%) of Bt65**a** x H99 F_1 progeny because each translocation should decrease germination by ~50%.

Analysis of the genomes of Bt65 and Bt81 revealed large arrays of the Cnl1 transposon at all but one end of each of the 14 linear chromosomes (27/28 subtelomeric regions in Bt65 and 28/28 in Bt81) (Figure 3A, 3B). The assembled Cnl1 arrays (defined as ≥ 2 Cnl1 copies) in Bt65 and Bt81 range from 5 kb to 80 kb in length. These highly repetitive arrays made it difficult and, in some instances, impossible to confidently assemble telomeric repeat sequences at the ends of each Bt65 and Bt81 chromosome. Using manual telomere extension via read mapping, we were able to identify telomere repeats at only 20 chromosome ends in Bt65 and 13 in Bt81. In contrast to Bt65 and Bt81, genome assemblies for Bt89 and Bt133 were assembled with telomere repeats on all 28 chromosome ends without any manual extension (Figure S6). In these assemblies, some telomeres had no copies of Cnl1 while others had Cnl1 arrays up to 30 kb in length (Figure S9).

Further analysis revealed the Bt65 genome harbors at least 414 fragments of Cnl1, including 105 full-length copies, while the Bt81 genome appears to encode even more Cnl1 elements, with at least 449 fragments, including 147 full-length copies (Table 1). It is important to note that due to incomplete ends for most chromosomes, it is likely Bt81 and Bt65 encode additional copies of Cnl1 that were not assembled. The presence of long Cnl1 arrays in Bt65 and Bt81 was surprising because the *C. neoformans* H99 reference strain encodes only 22 fragments of Cnl1 and no full-length copies, and therefore, *C. neoformans* was not thought to harbor functional Cnl1 elements (Figure 3C) (Table 1).

Apart from the subtelomeres, retrotransposons in *Cryptococcus* are also enriched at centromeres, specifically the LTR retrotransposons Tcn1-Tcn6[26,28,47]. The changes in Cnl1 transposon content in Bt65 and Bt81 along with a previous study establishing a link between loss of RNAi and centromere length[28] motivated us to characterize the centromeres in Bt65, Bt81, Bt89, and Bt133. Identification of centromeres in these four isolates revealed shorter centromeres on average compared to H99: 40.3 kb in Bt65, 40 kb in Bt81, 35.6 kb in Bt89, and 41.2 kb in Bt133, compared to an average length of 47.3 kb in H99. However, this difference did not reach statistical significance (ANOVA, *p*-value = 0.153, Figure S10 and Table S3). Many centromeres in the assessed VNBII isolates had undergone numerous rearrangements and several inversions relative to one another (Figure S11). Centromeric alterations have also previously been observed in *C. neoformans* genetic deletion mutants lacking the canonical RNAi components Ago1 and Rdp1[28]. Combined, these analyses suggest that while Cnl1 is more abundant in the hypermutators Bt65 and Bt81 as well as the non-hypermutators Bt89 and Bt133, other retrotransposons are not substantially increased in number compared to H99 (Figure 3 and S9).

Table 1. Cnl1 burden in H99, hypermutator strains, related non-hypermutator strains, and six Bt65 x H99 *crg1*Δα F_1 progeny based on Nanopore sequencing data.

Strain	Hypermutator Status	Total Cnl1 burden (>50 bp)	Full-length Cnl1 copies (>99% in length)
H99	Non-hypermutator	22	0
Bt65	Hypermutator	414	105
Bt81	Hypermutator	449	147
Bt89	Non-hypermutator	261	24
Bt133	Non-hypermutator	246	23
Progeny 2	Hypermutator	212	30
Progeny 8	Hypermutator	172	40
Progeny 14	Non-hypermutator	296	68
Progeny 18	Non-hypermutator	321	88
Progeny 20	Non-hypermutator	425	136
Progeny 34	Hypermutator	187	41

Characterization of H99 *crg1*Δ x Bt65 F_1 progeny genomes reveals invasion of Cnl1 elements into naïve telomeres

Expression of transposable elements, including Cnl1, has been shown to be upregulated during sexual reproduction in *C. neoformans*[30,32,34]. To investigate how increased expression of Cnl1 during mating

impacts the genome, six of the 28 progeny utilized for QTL mapping were also selected for long-read whole-genome sequencing on the nanopore MinION sequencing platform: three hypermutator progeny that inherited the Bt65 *znf3* allele (P2, P8, and P34) and three non-hypermutator progeny that inherited the H99 *ZNF3* allele (P14, P18, and P20). Nanopore sequencing identified recombination points across the genomes of the progeny, providing information on which genomic regions were inherited from either parent and confirming these were F_1 genetic recombinants (Figure 4).

Surprisingly, the genome assemblies for the three progeny with the functional *ZNF3* allele inherited from H99 appear to encode more full-length Cnl1 elements and fragments compared to the three *znf3* progeny. However, of the three *znf3* progeny, telomeric repeat sequences were only identified at the end of two chromosomes (out of a possible 84 telomeric ends across the three progeny, 2/84 or 2%): on one arm of Chromosome 3 in P8 and one end of Chromosome 12 in P34 (Figure S12). This is in contrast to the 31 telomeres that were accurately assembled across the three *ZNF3* progeny (31/84 or 37%) (Figure S12). The smaller number of telomeres identified in the *znf3* progeny assemblies suggest there might be more Cnl1 elements that were not accurately included, similar to the assemblies for the hypermutators Bt65 and Bt81. Therefore, the Cnl1 burden quantified and presented in Table 1 might not accurately capture the entire Cnl1 burden of these strains.

A previous study found *ZNF3* to be a haploinsufficient gene because no progeny isolated from a *ZNF3* x *znf3*Δ cross showed evidence of sex-induced RNAi-mediated silencing[32]. This haploinsufficiency allowed us to analyze Cnl1 dynamics both hypermutator and non-hypermutator progeny and identified two additional phenomena. First, nearly all Cnl1 arrays in the progeny show signs of expansion and contraction relative to the Bt65 parental genome, suggesting these elements are either highly mobile during sexual reproduction, undergoing high levels of recombination, or both. Additionally, combined analysis of subtelomeric region inheritance patterns and Cnl1 arrays revealed Cnl1 elements are capable of invading naïve subtelomeric regions inherited from the H99 parent, i.e. regions that previously had no Cnl1 elements/fragments (Figure 4, Figure S12). In the three *znf3* progeny, 65% (28/43) of the naïve telomeric regions inherited from H99 acquired Cnl1 copies and arrays in many cases. In the three *ZNF3* progeny, 81% (35/44) of naïve telomeric regions inherited from H99 now had Cnl1 elements. Overall, both *ZNF3* and *znf3* F_1 progeny inherited roughly equivalent numbers of telomeric regions from either parent and Cnl1 invaded a majority of the naïve H99 telomeres.

Complementation of the nonsense mutation in *ZNF3* significantly lowers the mutation rate and restores production of siRNAs

All of the available evidence thus far suggested the nonsense mutation in *ZNF3* unique to Bt65 and Bt81 is responsible for the hypermutation phenotype, possibly due to compromised RNAi silencing of Cnl1 elements. To test this hypothesis, we used CRISPR-mediated gene editing to restore the functional *ZNF3* allele in Bt65. Gene editing was achieved with the transient CRISPR-Cas9 coupled with

electroporation (TRACE) system and the utilization of a functional *ZNF3* allele from a closely related strain, Bt133, such that only the SNP responsible for the nonsense mutation would be changed to the wild-type nucleotide (found in H99 and all SDC isolates except Bt65 and Bt81)[48]. Following transformation and selection, we identified two Bt65 transformants that had successfully integrated a single copy of the Bt133 *ZNF3* allele at the endogenous *znf3* locus, Bt65+*ZNF3-1* and Bt65+*ZNF3-2*. These two independent Bt65+*ZNF3* transformants were subjected to fluctuation analysis to determine if changing the single SNP responsible for the *znf3* nonsense mutation restored the mutation rate to a wild-type level. On YPD + rapamycin + FK506 medium, both transformants had significantly lower mutation rates than Bt65, similar to those observed in H99 and Bt133 (Figure 5A).

We next sequenced the sRNA repertoires of the Bt65+*ZNF3* isolates as well as Bt65, Bt81, and an H99 *znf3*Δ mutant; H99 and an H99 *rdp1*Δ mutant served as positive and negative controls, respectively. Analysis of the size distribution of sRNAs aligning to the *C. neoformans* genome showed that Bt65, Bt81, and H99 *znf3*Δ displayed profiles similar to that of the canonical RNAi mutant H99 *rdp1*Δ because they lacked the characteristic siRNA peak at 21-24 nt (Figure 5B). sRNA analysis also revealed that *ZNF3* complementation in Bt65 restored the 21-24 nt sRNA peak. We also characterized the 5’ nucleotide identity of sRNAs of different lengths and found only H99 and the two Bt65+*ZNF3* transformants had a peak of 21-24 nt sRNAs with a predominance for a 5’ U, another characteristic of siRNAs produced by the RNAi pathway (Figure 5C).

To determine how complementation of *ZNF3*, and thus restoration of RNAi, specifically impacted silencing of Cnl1, we quantified sRNAs aligning to Cnl1 elements across the Bt65 genome (Table S4). Relative to Bt65, the normalized expression of sRNAs corresponding to Cnl1 were increased 5.6-fold in H99 (a strain with no full-length Cnl1 elements and only 22 Cnl1 fragments), 10.6-fold in Bt65+*ZNF3-1*, and 12.8-fold in Bt65+*ZNF3-2* (Table S4). To demonstrate the marked difference in Cnl1 sRNAs in the two Bt65+*ZNF3* transformants compared to Bt65, sRNAs were plotted along a telomeric Cnl1 array on Chromosome 1 in Bt65 (Figure 5D). These results suggest that changing the single nucleotide responsible for the nonsense mutation in *ZNF3* back to the wild-type nucleotide found in closely related strains as well as H99 was able to successfully restore RNAi in Bt65. This result also provides evidence that all other RNAi components have been largely maintained in a functional form despite the loss of RNAi activity in Bt65.

Discussion

Transposable element mobilization can alter gene expression, gene function, and even genomic stability. In this study, we identified two clinical, hypermutator isolates of *C. neoformans*, Bt65 and Bt81, with a massive accumulation of the non-LTR retrotransposon Cnl1 at all but one subtelomeric loci of all fourteen chromosomes. Furthermore, the Cnl1 element was capable of inserting in non-subtelomeric regions of the genome, resulting in resistance to diverse classes of antifungal drugs. These findings were unprecedented as *C. neoformans* is thought to be an RNAi-proficient species[13,49]. Additionally, *C. neoformans* isolates were also thought to lack full-length copies of Cnl1, and only the sister species *C.*

deneoformans has been shown to harbor full-length Cnl1 elements capable of mobilization[13,27]. These findings highlight the importance of intraspecific diversity at both the genotypic and phenotypic levels.

Following isolation, phenotyping, and genotyping of the Bt81**a** x H99α F_1 progeny, a significantly smaller number of progeny inherited the Bt81 *znf3* allele than expected. These results were surprising giving the approximately 1:1 inheritance pattern of the *ZNF3* allele from either parent in the Bt65**a** x H99α F_1 progeny. Based on long-read whole-genome sequencing, Bt81 likely has a substantially higher burden of Cnl1 than Bt65. Previous studies have shown that expression of transposons, including Cnl1, is significantly upregulated during sexual reproduction in RNAi mutants, such as H99 *znf3*Δ[30,32–34]. The higher Cnl1 burden in Bt81 combined with RNAi-deficiency and transposon upregulation during mating could lead to an increased frequency of deleterious Cnl1 insertions in progeny lacking *znf3*, and thus a biased inheritance pattern of *ZNF3*. It is also possible that a higher Cnl1 burden favored the selection of a suppressor mutation in the Bt81 F_1 progeny. Additionally, *znf3*Δ mutants are the only *C. neoformans* RNAi mutants studied thus far in which progeny from unilateral genetic crosses (i.e. crosses in which only one parent lacks *ZNF3*) exhibit complete loss of RNAi-mediated silencing[32], and thus loss of *ZNF3* results in haploinsufficiency. Overall, the unequal *znf3* inheritance patterns in Bt81 F_1 progeny indicate a sufficiently high burden of Cnl1 elements may be deleterious during sexual reproduction.

In Bt65, QTL mapping and genetic complementation demonstrated the hypermutator phenotype was caused by a single SNP in the RNAi gene *ZNF3*. Changing the *znf3* nonsense mutation to the nucleotide found in the laboratory reference strain as well as phylogenetically closely related strains lowered the mutation rate to a wild-type level and restored RNAi function, including expression of siRNAs corresponding to Cnl1, thus confirming the role of Znf3 in Cnl1 regulation. Although *ZNF3* complementation restored the mutation rate and siRNA production, the identified QTL which spans Chromosomes 3 and 11 accounted for only 64% of the hypermutator phenotype. This may suggest the existence of additional contributing loci of small effect, such as a mutation in another RNAi component that results in a partial loss of function, although the mapping population used here is under powered to detect such small-effect loci. It is also possible that variation in Cnl1 burden across the F_1 progeny accounts for the remainder of the genetic loci that contribute to the hypermutator phenotype. However, because short-read sequencing data was utilized in variant calling, these Cnl1 loci may have been difficult to map in the F_1 progeny or failed to pass quality criteria and were subsequently filtered out during preprocessing of the genetic data used in QTL analysis.

Despite having the wild-type nucleotide in the first exon of *ZNF3*, where Bt65 and Bt81 have a nonsense mutation, both Bt89 and Bt133 (two of the most closely related strains) have a substantial accumulation of subtelomeric Cnl1 arrays relative to the laboratory reference strain H99, which has none. The considerable Cnl1 burden in Bt89 and Bt133 indicates that they may descend from an ancestral strain that had lost RNAi function, but which was subsequently restored, such as through a genetic cross in which they inherited a wild-type *ZNF3* allele. The bias against *znf3* inheritance in Bt81 F_1 progeny and the lower but still impressive Cnl1 burden in Bt89 and Bt133 potentially illustrate a natural example of how *C.*

neoformans genomes have struck a balance in their mutational capability, switching between high mutational capacities during times of RNAi loss, and genomic stability when RNAi is restored, all via one SNP in only one key RNAi gene. The expansion and contraction of Cnl1 arrays as well as the ability of Cnl1 to invade naïve subtelomeric regions inherited from the H99 parent in the Bt65 F_1 progeny genomes was also exceptional. The observed Cnl1 subtelomeric dynamics mirror those observed for the MoTeR transposons of the fungal plant pathogen *Magnaporthe oryzae*, which were also recently shown to localize to highly dynamic subtelomeric arrays[50]. It is possible the Cnl1 subtelomeric arrays identified here could eventually overcome the requirement for telomerase, as in the example of *Drosophila* telomeres, in which the functions of telomerase have been supplanted by a telomeric retrotransposon[51,52].

The finding that only a single SNP rendered the RNAi pathway non-functional in Bt65, and that no additional obvious mutations had occurred in other genes that might disrupt RNAi function, suggest Bt65 represents an intermediate evolutionary stage and could illustrate the natural consequences of relatively recent RNAi loss. Further analysis of the consequences of RNAi loss in Bt65 (and potentially Bt81) through experimental evolution or gene regulation analyses could shed light on the short-term consequences of RNAi loss at the genomic and phenotypic levels. Bt65 could thus serve as an interesting intermediate comparator between RNAi-proficient *C. neoformans* isolates and the closely related RNAi-deficient species *C. deuterogattii*[7,32]. Instances of relatively recent loss of RNAi have also been observed in a natural *Caenorhabditis elegans* isolate, which has a large deletion in a RIG-I homolog required for RNAi and was shown to be infected with an RNA virus[53,54]. Unlike the identified *C. elegans* virus-infected strain and several other RNAi-deficient fungal species, such as *Saccharomyces cerevisiae*, *Ustilago maydis*, and several *Malassezia* species, we were unable to identify a dsRNA virus in either of the two *C. neoformans* hypermutator strains identified here (see Materials and Methods)[49,55]. It is possible though, that the hypermutators harbor other types of mycoviruses (ssRNA, for example) that we were unable to detect through our approaches or that the mycovirus was cured by common microbiological isolation practices[55].

The identification of this hypermutator phenotype in natural *C. neoformans* clinical isolates has important implications for antifungal drug resistance and potentially other adaptive consequences. Here, we showed Cnl1 insertion could confer resistance to diverse classes of antifungal drugs, including the clinically relevant antifungal 5-FC. Insertion of Cnl1 into other genes, particularly those involved in the sterol biosynthesis pathway, could confer resistance to amphotericin B and fluconazole, the only other antifungal drugs effective for *C. neoformans* treatment[23,56]. This mechanism of drug resistance also has interesting implications for a novel antifungal approach that utilizes dsRNA to initiate RNAi silencing in fungal plant pathogens[57]. The effects of Cnl1 insertion at non-coding loci, such as promoters and 3' untranslated regions, could also impact overall genomic stability or alter gene expression to have important phenotypic implications for virulence, similar to the effects of the *Ac*/*Ds* elements of maize, the first transposable elements discovered[58]. Alterations in gene expression might also confer resistance to drugs for which resistance cannot be gained through loss of function mutations. Even if full resistance

isn't acquired, altered gene expression could contribute to antifungal drug tolerance, like the tolerance observed in *Candida albicans*, which contributes to persistent infections in immunocompetent patients[59,60].

At this stage it is difficult to know how Cnl1 transposition and accumulation may be selected for or against over time. The subtelomeric arrays identified in Bt65 and Bt81 may undergo cycles of amplification and recombination-mediated contraction allowing them to exploit Cnl1 mutagenesis when under stress, similar to the retrotransposons replication cycles observed in some plants[61,62]. Maintaining an RNAi-deficient background could also be adaptive in the context of viral infection, as has been shown in yeast harboring the killer virus, which outcompete neighboring uninfected strains, and in mice harboring latent herpesvirus, which are protected from the bacterial pathogens *Listeria monocytogenes* and *Yersinia pestis*[63,64]. Conversely, the mutational impact of Cnl1 mobilization could be highly deleterious over the long term and therefore may not represent a massive contributor to the rise of drug resistance. Natural selection could either select for reversion to a functional RNAi-pathway through mutation of the *znf3* nonsense mutation or instead preserve loss of RNAi and eliminate all full-length transposable elements, as has been observed in *C. deuterogattii*[28]. Future research on the potential for Cnl1 insertion to mediate resistance to amphotericin B and fluconazole, and the impact of hypermutation due to Cnl1 mobilization on *in vivo* drug resistance, adaptive potential, and genomic stability over time will be of great interest.

Materials And Methods

Strains and growth

The *C. neoformans* strains described in this study are listed in Table S5. Strains were stored at -80°C in liquid yeast extract peptone dextrose (YPD) supplemented with 15% glycerol. Strains were inoculated on YPD agar plates, grown for three days at 30°C, and maintained at 4°C. Due to the hypermutator phenotypes associated with several of the strains in this study, strains were not maintained on YPD agar plates for routine use for more than two weeks; fresh cells from frozen glycerol stocks were inoculated to YPD agar plates as needed.

Screening for hypermutator candidates

Assays for the emergence of resistance (papillation assays) were conducted as previously described[65]. In brief, ten independent overnights per strain were grown overnight at standard laboratory conditions in 5mL liquid YPD medium. Cultures were then spun down, washed, and concentrated in 2 mL dH_2O. Each culture was swabbed to a quadrant of either YPD + 100 ng/mL rapamycin + 1 µg/mL FK506 agar medium or YNB + 100 µg/mL 5-fluorocytosine agar medium. YPD + rapamycin + FK506 plates and YNB + 5-fluorocytosine plates were incubated for up to seven days at 37°C and 30°C, respectively. Fisher's

exact probability test was used to determine if the associations environmental isolates and the hypermutator phenotype was statistically significant using the VassarStats online software (http://vassarstats.net).

Fluctuation assays

Fluctuation assays were conducted as previously described[65]. Briefly, ten independent overnights of each strain were grown overnight in 5 mL liquid YPD medium at 30°C. Cultures were washed three times and resuspended in dH_2O. Cells were then plated to the appropriate medium (100 µL 10^{-5} cells on YPD, 100 µL 10^{-2} cells on YNB + 5-FC, and 100 µL undiluted cells on YPD + rapamycin + FK506 and YNB + 5-FOA). Mutation rates and 95% confidence intervals were calculated using the FluCalc program[66]. For the increased temperature fluctuation analysis, strains were grown overnight at either 30°C or 37°C before use in fluctuation assays, as indicated. YPD + rapamycin + FK506 plates were incubated and 37°C; all other media was incubated at 30°C. Mutation rates and confidence intervals for all fluctuation assays in this study are provided in Table S6.

Characterizing mutation spectra

Following selection on antifungal drug media, resistant colonies were streak purified to YPD medium. Genomic DNA was isolated from the purified colonies, and genes in which mutations are known to cause resistance to the corresponding antifungal drug were PCR amplified (*URA5* and *URA3* for 5-FOA-resistant colonies[44,45], *FRR1* for rapamycin+FK506-resistant colonies[41,42], and *FUR1*, and *UXS1* for 5-FC-resistant colonies[46]). Oligonucleotides used for all PCR reactions in this study are listed in Table S7. PCR products were subjected to gel electrophoresis, products of interest were extracted from agarose gels using a QIAgen gel extraction kit and sequenced through classical Sanger sequencing conducted at Genewiz. Sequenced mutations, including transposon insertions, were characterized with both Sequencher software and the Clustal Omega Multiple Sequence Alignment program[67]. Identified transposon insertion sequences in *FRR1*, *URA5*, and *FUR1* are listed in Table S8.

Illumina sequencing

Single colonies from strains for whole-genome Illumina sequencing were inoculated in 50 mL of liquid YPD medium and grown overnight at 30°C, shaking. Cells were collected and lyophilized as previously described[65], and high molecular weight DNA was isolated following the CTAB protocol as previously described[68]. Strains were barcoded and sequencing libraries were generated with the Kapa

HyperPlus library kit for 300bp inserts, pooled, and sequenced using paired-end, 2 x 150bp reads on an Illumina HiSeq 4000 platform at the Duke University Sequencing and Genomic Technologies Core facility.

Generation of F_1 progeny

Bt65**a** x H99α *crg1*Δ and Bt81**a** x H99α *crg1*Δ F_1 progeny were generated by genetically crossing either Bt65 or Bt81 with H99 *crg1*Δ on Murashige Skoog (MS) medium (Sigma) following Basic Protocol 1 as described in Sun et al. 2019[69]. Basidiospores were randomly isolated through microdissection after three weeks of incubation on MS following Basic Protocol 2 as described in Sun et al. 2019[69].

Nanopore sequencing and genome assemblies

The DNA samples for nanopore sequencing were isolated and purified using the CTAB DNA preparation protocol described previously[70]. The size estimation of the obtained DNA was done using PFGE electrophoresis and quality was determined using NanoDrop. Once the high-quality DNA was obtained, sequencing was performed using the MinION device with the MinKNOW interface. During sequencing, Bt65, Bt89, and Bt133 were multiplexed together whereas six of the Bt65**a** x H99α progeny were multiplexed for a second sequencing run. For multiplexing, samples were barcoded using EXP-NBD103/EXP-NBD104 kits and libraries were made using SQK-LSK109 kit as per the manufacturer's instructions. The libraries generated were sequenced on R9.4.1 flow cell and reads were obtained in .fast5 format. These reads were then converted to fastq format using Guppy_basecaller (v 4.2.2_linux64). The reads were de-multiplexed using qcat (https://github.com/nanoporetech/qcat) or Guppy_barcoder (part of Guppy_basecaller) with barcode trimming option during processing. Bt81 nanopore sequencing was done as a standalone sample using an R9 flow cell (FLO-MN106) and basecalling was performed during the run itself.

The sequences obtained for each sample were then assembled via Canu (v2.0 or v2.1.1) to obtain contig-level genome assemblies. For the assembly, only >2 kb long reads were used for the Bt65**a** x H99α F_1 progeny and Bt81, whereas >5 kb were used for Bt65, Bt89, and Bt133 genomes. Contigs were then assigned chromosome numbers based on their synteny with the reference genome, H99. The numbering of chromosomes involved in translocations was assigned based on the respective syntenic centromere. Some of the chromosomes were not fully assembled and were broken into multiple contigs (Chr 1, Chr 2 for Bt65, Chr 2, Chr11, Chr14 for Bt89, and Chr 2, Chr 5 for Bt133). For such cases, the respective contigs were joined artificially and then processed by read-mapping to obtain complete collinear chromosomes. Specifically, the contigs were stitched together in orientation as determined based on their synteny. Corrected reads obtained from Canu were then mapped to the respective genomes and duplicated or missing regions from the junction were identified. The chromosome sequence was then corrected accordingly by inserting/correcting/deleting sequences and a full-length chromosome sequence was

obtained. Complete resolution of junctions was obtained for Bt65, Bt89, and Bt133 genomes by this approach. However, some of the Bt65 F_1 progeny chromosomes could not be resolved, probably due to hybrid origin of sequencing reads, and were left with gaps as such.

Once chromosome level genome assemblies were obtained for the Bt65, Bt81, Bt89, and Bt133 genomes, the genome sequences were further processed to improve telomeric and subtelomeric regions. For this purpose, the corrected reads obtained from Canu were mapped back to the respective chromosome-level genomes using minimap2 v2.14. The obtained bam files were then analyzed manually by IGV and consensus or, in a few cases, individual reads (up to 30 kb) representing extra sequence beyond an assembled chromosome were extracted as sam files. These consensus extra sequences were then added onto the chromosome sequences to obtain longer chromosomes. In some cases, read mapping also resulted in the identification of incorrect sequence assembly at subtelomeric regions, and in those cases, the sequence was trimmed until a consensus sequence was observed at the end of the chromosome. Once these corrections were made, the genome assemblies were polished via one round of nanopolish and five rounds of pilon, except for the Bt81 genome, for which only 5X pilon polishing was performed. As a result of these corrections and polishing, final assemblies were obtained for each of the four isolates and are described in the study. For the Bt65 F_1 progeny genome assemblies, the subtelomeric extension/curation was not performed, but they were polished using both nanopolish and 5X pilon.

Centromere, telomere, and Cnl1 mapping

Centromeres in Bt65, Bt81, Bt89, and Bt133 were defined based on their synteny with the reference H99 genome (genome assembly ASM301198v1) The final polished genomes were used and centromere locations were identified by BLASTn analysis using H99 centromere-flanking genes as query sequences. Once the centromere locations were defined, Tcn1-Tcn6 locations within those regions were mapped by BLASTn analysis. For the representation, only BLAST hits longer than 400 bp were mapped. For the overlapping BLAST hits with multiple Tcn elements, the longest and best BLAST result was used, and the rest of the matches were discarded from further analysis. All the hits were then visualized using Geneious Prime and maps were exported as .svg files, which were then processed using Adobe Illustrator.

For the Cnl1 mapping at the subtelomeres, the longest *CNL1* insertion sequence from the Bt65 genome was used as the query sequence and BLASTn was performed against each genome. BLAST hits longer than 50 bp were mapped to the respective genomes and visualized using Geneious Prime where the hits were color-coded based on their lengths. The zoomed views for these maps were then exported as .svg files, processed using Adobe Illustrator, and combined with centromere Tcn mapping analysis to generate final figures.

RepeatMasker was used to annotate all transposons in the *de novo* genome assembly of Bt65. For this purpose, RepeatMasker (v4.0.7) with Dfam (v3.3) and RepBaseRepeatMaskerEdition-20181026

libraries was used, supplemented with RepBase EMBL database (v26.04)[71–73]. The "-species fungi" option was used to identify all repeats in the genome and provided additional support for the manual Tcn and Cnl1 mapping.

Synteny maps

Synteny comparisons between the genomes were performed using SyMAP v4.2 with the H99 genome as the reference (genome assembly ASM301198v1). The synteny comparison was conducted using default parameters and synteny block maps were exported as .svg files. The maps were processed using Adobe Illustrator for visualization. The phylogenetic relationship as depicted in Figure S6 was drawn based on the earlier representation[18]. The telomere and centromere locations were marked manually based on the presence of the telomere repeat sequence and Tcn mapping, respectively.

For the centromere comparisons, all centromere sequences along with Tcn annotations were converted into GenBank format. The files were then used for synteny comparison via EasyFig v2.2.3. The maps were exported as .svg files which were processed in Adobe Illustrator.

Recombination maps for Bt65 x H99 F_1 progeny

Six of the Bt65**a** x H99α F_1 progeny were sequenced with on the nanopore MinION sequencing platform and their genomes were assembled and polished using the methods described above. Once their genomes were assembled, recombination maps were generated by mapping the Illumina sequence data from the parental strains to each of the progeny genomes. For this purpose, both H99 and Bt65 Illumina reads were used from published datasets (SRR642222 and SRR647805 for H99; SRR836876, SRR836877, SRR836878, SRR836880, SRR836884, and SRR836885 for Bt65). Reads from all runs were merged to obtain a single file for both H99 and Bt65. The reads were then mapped to the progeny genomes using Geneious Prime default mapper with three iterations. Variants with 90X coverage and at least 90% variant frequency were called from these mapped files. These variants along with coverage analysis were then used to identify recombination sites and generate recombination maps. Cnl1 mapping for each of progeny genome was performed as described above. The location of *ZNF3* in each genome was identified by BLASTn analysis using H99 *ZNF3* (CNAG_02700) as the query sequence.

Genetic variant calling and segregant filtering

Whole-genome sequencing data of 28 F_1 progeny from the Bt65**a** x H99α *crg1*Δ cross were aligned via BWA (v0.7.12-r1039)[74] to an H99 reference genome (downloaded from FungiDB [http://fungidb.org/fungidb/] on April 15th, 2020; FungiDB-46_CneoformansH99_Genome.fasta) and

genetic variants between Bt65 and H99 were called using SAMtools (v0.1.19-96b5f22941)[75] and FreeBayes (v1.2.0)[76]. Approximately 300,000 raw genetic variants were identified across the segregants. The genotypic correlation between progeny, the read coverage per genetic variant, and the ratio of reads suggesting the H99 vs. Bt65 allele per variant were monitored across the genome to identify clones, progeny with aneuploid genomes, and heterozygotic diploids (respectively). Two pairs of clones were identified (Supplementary Table S9) and one segregant from each pair was retained for analysis. F_1 progeny 25 was identified as a heterozygotic diploid (Supplementary Figure S13) and removed from initial analysis. Instances of aneuploidy (and partial duplications) are observed along Chromosomes 3, 4, 11, and 13 within six segregants from this cross and for initial filtering and analysis, those with heterozygotic aneuploidy were removed from analysis (Supplementary Table S9).

Genetic variant filtering

After removing clones and samples with aneuploidy or diploidy, raw genetic variants were filtered by limiting sites to bi-allelic SNPs, called across all the progeny (100% call rate), with greater than 10X read coverage (and a maximum of 200X), a minor allele frequency of 5%, and a quality score greater than 4 (and less than 5.4). These filtering criteria were selected after examining the bivariate relationships between allele frequency, read depth, and quality scores per chromosome (Supplementary Figure S14A and 14B). Genetic variant sites were also removed if within one kb of the centromere along a given chromosome[28]. After filtering, a total of 215,411 bi-allelic SNPs were retained for further analysis. The median distance between contiguous SNP sites is 45 bp, and less than 0.01% of neighboring sites had a distance larger than two kb. The allele frequencies across the genome ranged between 25 and 75% of segregant with the Bt65 allele, except for a large portion of Chromosome 13, between 0 to 500 kb, where over 88% of segregants inherited the Bt65 allele (Figure S14C). With these data, a Poisson regression (methods described in Roth *et al*. 2018[77]) was used to relate the average number of crossovers across F_1 progeny as a function of chromosome size. With this model, the estimated genome-wide, physical-to-genetic distance in this cross is 8.14 kb/cM.

QTL mapping

For use in association tests, across 24 F_1 progeny and the two parental strains, the 215,411 bi-allelic SNPs were collapsed into 1,237 unique haplogroups made up of genetic variants sites that co-segregated within the segregant genomes, such that, between any two haplogroups, at least one segregant contains a change in allele (i.e. a recombination event between the Bt65**a** and H99α *crg1*Δ genomes). Collapsing genetic variants into haplogroups reduces the number of repeated tests in association mapping and computational costs[78]. Across the 1,237 haplogroups, the mutation rate on rapamycin + FK506 medium was used as the phenotype and tested for association using a Kruskal-Wallis H-test. The $-\log_{10}$ (*p*-value)

from this test across haplogroups was monitored to identify QTL. Significance thresholds were established via 10,000 permutations with an α = 0.01 as described in Churchill and Doerge (1994)[79], and 95% confidence intervals for the QTL locations were generated as described in Visscher *et al.* (1996)[80]. The heritability at the peak of identified QTL was estimated using linear regression with the model: $M = \mu + \beta I + e$, were M is the mutation rate x 10^7, e is an error term, μ is the mean mutation rate x 10^7, I is an indicator variable for the allele at the QTL peak – coded as 0 if from H99α *crg1*Δ or 1 if from Bt65**a** – and β is the effect of having the H99α *crg1*Δ vs. the Bt65**a** allele at the QTL. The variation explained (R^2) from this model was used as an estimate of the heritability.

Gene annotation and SNP effect prediction

For genes within the identified QTL spanning Chromosomes 3 and 11, the alleles between H99 and Bt65 were imputed using filtered SNP data (described above). The published H99 reference strain annotation (downloaded from FungiDB [http://fungidb.org/fungidb/] on April 15th, 2020; FungiDB-46_CneoformansH99.gff) was used to predict changes in protein sequence between the H99 and Bt65 parental backgrounds.

CRISPR-mediated genetic complementation of *ZNF3* in Bt65

To change the single nucleotide responsible for the nonsense mutation in the first exon of *ZNF3*, a thymine (base 976004 of H99 Chromosome 3 (CNA3 assembly, accession GCA_000149245.3)), to the wild-type cytosine found in H99 and other phylogenetically closely related strains, the Transient CRISPR-Cas9 Coupled with Electroporation (TRACE) system was used[48]. Briefly, the gene encoding Cas9 was PCR amplified from plasmid pXL1-CAS9-HYG. The *SH1-NEO* construct encoding *NEO* (G418 resistance) targeted to a safe haven locus (SH1) was amplified from plasmid pSDMA57[81]. For the sgRNA expression construct, the U6 promoter was amplified from XL280α gDNA, and the sgRNA scaffold was amplified from plasmid pYF515[82]. Overlap PCR was used to generate the sgRNA construct with the U6 promoter and sequences targeting either SH1 or the nonsense mutation in *ZNF3*. A 2,171bp region was amplified from Bt133, containing the wild-type C nucleotide in *ZNF3* exon 1 and no other mutations relative to Bt65 for integration at the *ZNF3* endogenous locus in Bt65 (1,197bp upstream of the *ZNF3* start codon to 971bp after the start codon). The *SH1-NEO* construct was linearized with the AscI restriction enzyme (NEB). 2μg of the Bt133 *ZNF3* recombination template, 2μg of the *SH1-NEO* linearized construct, 250ng of the *ZNF3* gRNA, 250ng of the SH1 gRNA, and 1.5μg of Cas9 DNA were transformed simultaneously into Bt65 via electroporation using a BIO-RAD Gene Pulser.

Transformants were selected on YPD + G418 agar plates. Successful transformants were identified through restriction enzyme digest with BtsI-v2 (NEB), which cleaves the first exon of Bt65 *znf3* at the nonsense mutation but does not cleave the first exon of *ZNF3* in Bt133 (or H99) (Figure S15A). PCR was

also used to ensure that no transformants had integrated copies of the gene encoding Cas9 or the gRNA constructs and that only a single Bt133 *ZNF3* allele had integrated correctly at the endogenous *znf3* locus (Figure S15B-F). Sanger sequencing was used to further confirm correct replacement of the Bt65 *znf3* allele including the nonsense mutation with the Bt133 *ZNF3* allele. No identified Bt65+*ZNF3* transformants also had a stably integrated copy of the *NEO* gene at the safe haven locus.

sRNA isolation and sequencing

C. neoformans cells were grown overnight in 50 mL YPD medium at standard laboratory conditions. Following culture, cells were spun down, supernatant was removed, and cells were frozen at -80°C for one hour. Cells were then freeze dried with a Labconco Freezone 4.5 lyophilizer overnight. 70 mg of lyophilized material was used for sRNA isolation following the mirVana miRNA Isolation Kit manufacturer's instructions. sRNA was quantified with a Qubit 3 Fluorometer and quality was verified with an Agilent Bioanalyzer using an Agilent Small RNA Kit. sRNA libraries were prepared with a QiaPrep miRNA Library Prep Kit and 1 x 75 bp reads were sequenced on the Illumina NextSeq 500 System at the Duke University Sequencing and Genomic Technologies Core facility.

sRNA data processing

Initial quality control of the small RNA libraries was performed with FastQC 0.11.9[83] followed by the removal of QIAseq library adapters (5': GTTCAGAGTTCTACAGTCCGACGATC; 3': AACTGTAGGCACCATCAAT) with cutadapt 2.8[84]. All untrimmed reads or reads smaller than 14 nt were discarded. The surviving trimmed reads were mapped with bowtie v1.2.3[85] to the *C. neoformans* Bt65 genome, allowing multiple alignments but no mismatches. The resulting SAM files were converted into BAM file format with SAMtools 1.9[75] and feature read counts of transposable elements were calculated with BEDTools[86] using the `intersect -wa´ option and the annotations of transposable elements, which were identified with RepeatMasker using the repbase database for *C. neoformans*[71,72]. Normalization of the read counts to reads per million (RPM) was performed, allowing the comparison of the libraries. Furthermore, the read depth on both DNA strands was analyzed with SAMtools and custom made perl scripts were used to calculate the read size distribution and 5'-nucleotide preference of the small RNA reads as previously described[87,88].

Double-stranded RNA enrichment

For dsRNA enrichment, *Cryptococcus* cells were grown overnight in 50mL liquid YPD medium at 30°C. RNA was extracted, and dsRNA was enriched as previously described[55]. dsRNA enrichment in H99,

Bt65, and Bt81 did not reveal the presence of any large dsRNA segments (Figure S16).

Data and software availability

Genetic variant filtering, QTL mapping, and SNP effect prediction was conducted in python (anaconda 3.7.3) via custom scripts; hosted at: https://github.com/magwenelab/Hypermutator_QTL . All sequencing data is available under BioProject PRJNA749953.

Acknowledgements

We thank and acknowledge Blake Billmyre for initial project guidance, Shelly Clancey for instruction in conducting fluctuation assays and dsRNA enrichment protocols, Josh Granek for preliminary analyses of hypermutator genomes, Zanetta Chang for assistance in sRNA isolation, Kayla Sylvester for assistance with screening of SDC isolates, and the laboratory of Chris Holley at Duke University for the use of their Nanodrop and BioAnalyzer equipment for preliminary sRNA analysis. We thank Mark Farman and Mostafa Rahnama for stimulating discussion on the impacts of transposons on telomere dynamics. We thank Kevin Zhu for assistance with figure generation. We also thank Sheng Sun, Blake Billmyre, Andy Alspaugh, Sue Jinks-Robertson, Asiya Gusa, and Kayla Sylvester for critical reading and comments on the manuscript. This work was funded by NIH/NIAID F31 Fellowship 1F31AI143136-02A1 awarded to S.J.P. and NIH/NIAID R37 MERIT award AI39115-23, R01 grant AI50113-16, and R01 grant AI33654-04 awarded to J.H. J.H. is co-Director and Fellow of the CIFAR program Fungal Kingdom: Threats & Opportunities. We also thank the Madhani Laboratory and NIH grant R01 AI100272 for the KN99α *msh2*Δ deletion strain. T.A.D. and U.K. are funded by the German Research Foundation (DFG) (Bonn Bad-Godesberg, Germany) (KU 517/15-1).

References

1. Sniegowski, P. D., Gerrish, P. J. & Lenski, R. E. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* **387**, 703–705 (1997).

2. Barrick, J. *et al.* Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243–1247 (2009).

3. Tenaillon, O. *et al.* Tempo and mode of genome evolution in a 50,000-generation experiment. *Nature* **536**, 165–170 (2016).

4. Harfe, B. D. & Jinks-Robertson, S. Mismatch repair proteins and mitotic genome stability. *Mutat. Res.* **451**, 151–167 (2000).

5. Fisher, K. J., Buskirk, S. W., Vignogna, R. C., Marad, D. A. & Lang, G. I. Adaptive genome duplication affects patterns of molecular evolution in *Saccharomyces cerevisiae*. *PLoS Genet.* **14**, e1007396 (2018).

6. Billmyre, R. B., Croll, D. & Li, W. Highly recombinant VGII *Cryptococcus gattii* population develops clonal outbreak clusters through both sexual macroevolution and asexual microevolution. *mBio* **5**, e01494-14 (2014).

7. Billmyre, R. B., Clancey, S. A. & Heitman, J. Natural mismatch repair mutations mediate phenotypic diversity and drug resistance in *Cryptococcus deuterogattii*. *eLife* **6**, e28802 (2017).

8. Healey, K. R. *et al.* Prevalent mutator genotype identified in fungal pathogen *Candida glabrata* promotes multi-drug resistance. *Nat. Commun.* **7**, 11128 (2016).

9. Boyce, K. J. *et al.* Mismatch repair of DNA replication errors contributes to microevolution in the pathogenic fungus *Cryptococcus neoformans*. *mBio* **8**, e00595-17 (2017).

10. Steenwyk, J. L. *et al.* Extensive loss of cell-cycle and DNA repair genes in an ancient lineage of bipolar budding yeasts. *PLoS Biol.* **17**, e3000255 (2019).

11. Sionov, E., Lee, H., Chang, Y. C. & Kwon-Chung, K. J. *Cryptococcus neoformans* overcomes stress of azole drugs by formation of disomy in specific multiple chromosomes. *PLoS Pathog.* **6**, e1000848 (2010).

12. Kwon-Chung, K. J. & Chang, Y. C. Aneuploidy and drug resistance in pathogenic fungi. *PLoS Pathog.* **8**, 8–11 (2012).

13. Janbon, G. *et al.* Characterizing the role of RNA silencing components in *Cryptococcus neoformans*. *Fungal Genet. Biol.* **47**, 1070–1080 (2010).

14. Gusa, A. *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).

15. Idnurm, A. *et al.* Deciphering the model pathogenic fungus *Cryptococcus neoformans*. *Nat. Rev. Microbiol.* **3**, 753–764 (2005).

16. Chayakulkeeree, M. & Perfect, J. R. Cryptococcosis. *Infect. Dis. Clin. North Am.* **20**, 507–544 (2006).

17. Hagen, F. *et al.* Recognition of seven species in the *Cryptococcus gattii/ Cryptococcus neoformans* species complex. *Fungal Genet. Biol.* **78**, 16–48 (2015).

18. Desjardins, C. A. *et al.* Population genomics and the evolution of virulence in the fungal pathogen *Cryptococcus neoformans*. *Genome Res.* **27**, 1207–1219 (2017).

19. Rajasingham, R. *et al.* Global burden of disease of HIV-associated cryptococcal meningitis: an updated analysis. *Lancet Infect. Dis.* **17**, 873–881 (2017).

20. Brouwer, A. *et al.* Combination antifungal therapies for HIV-associated cryptococcal meningitis: a randomised trial. *Lancet* **363**, 1764–1767 (2004).

21. Laniado-Laborín, R. & Cabrales-Vargas, M. N. Amphotericin B: side effects and toxicity. *Rev. Iberoam. Micol.* **26**, 223–227 (2009).

22. Vandeputte, P. *et al.* Molecular mechanisms of resistance to 5-fluorocytosine in laboratory mutants of *Candida glabrata*. *Mycopathologia* **171**, 11–21 (2011).

23. Perfect, J. R. & Cox, G. M. Drug resistance in *Cryptococcus neoformans*. *Drug Resist. Updat.* **2**, 259–269 (1999).

24. Nosanchuk, J. D., Cleare, W., Franzot, S. P. & Casadevall, A. Amphotericin B and fluconazole affect cellular charge, macrophage phagocytosis, and cellular morphology of *Cryptococcus neoformans* at subinhibitory concentrations. *Antimicrob. Agents Chemother.* **43**, 233–239 (1999).

25. Posteraro, B. *et al.* Identification and characterization of a *Cryptococcus neoformans* ATP binding cassette (ABC) transporter-encoding gene, *CnAFR1*, involved in the resistance to fluconazole. *Mol. Microbiol.* **47**, 357–371 (2003).

26. Loftus, B. J. *et al.* The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. *Science* **307**, 1321–1324 (2005).

27. Goodwin, T. J. D. & Poulter, R. T. M. The diversity of retrotransposons in the yeast *Cryptococcus neoformans*. *Yeast* **18**, 865–880 (2001).

28. Yadav, V. *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).

29. Cruz, M. C. *et al.* Rapamycin antifungal action is mediated via conserved complexes with FKBP12 and TOR kinase homologs in *Cryptococcus neoformans*. *Mol. Cell. Biol.* **19**, 4101–4112 (1999).

30. Wang, X. *et al.* Sex-induced silencing defends the genome of *Cryptococcus neoformans* via RNAi. *Genes Dev.* **24**, 2566–2582 (2010).

31. Dumesic, P. A. *et al.* Stalled spliceosomes are a signal for RNAi-mediated genome defense. *Cell* **152**, 957–968 (2013).

32. Feretzaki, M., Billmyre, R. B., Clancey, S. A., Wang, X. & Heitman, J. Gene network polymorphism illuminates loss and retention of novel RNAi silencing components in the *Cryptococcus* pathogenic species complex. *PLoS Genet.* **12**, e1005868 (2016).

33. Feretzaki, M. & Heitman, J. Genetic circuits that govern bisexual and unisexual reproduction in *Cryptococcus neoformans*. *PLoS Genet.* **9**, e1003688 (2013).

34. Wang, X., Darwiche, S. & Heitman, J. Sex-induced silencing operates during opposite-sex and unisexual reproduction in *Cryptococcus neoformans*. *Genetics* **193**, 1163–1174 (2013).

35. Huff, J. T. & Zilberman, D. Dnmt1-independent CG methylation contributes to nucleosome positioning in diverse eukaryotes. *Cell* **156**, 1286–1297 (2014).

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

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

38. Ferrareze, P. A. G. *et al.* Application of an optimized annotation pipeline to the *Cryptococcus deuterogattii* genome reveals dynamic primary metabolic gene clusters and genomic impact of RNAi loss. *G3* **11**, jkaa070 (2021).

39. Breuder, T., Hemenway, C. S., Movva, N. R., Cardenas, M. E. & Heitman, J. Calcineurin is essential in cyclosporin A- and FK506-sensitive yeast strains. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 5372–5376 (1994).

40. Parent, S. A. *et al.* Calcineurin-dependent growth of an FK506- and CsA-hypersensitive mutant of *Saccharomyces cerevisiae*. *J. Gen. Microbiol.* **139**, 2973–2984 (1993).

41. Heitman, J., Movva, N. R., Hiestand, P. C. & Hall, M. N. FK 506-binding protein proline rotamase is a target for the immunosuppressive agent FK 506 in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 1948–1952 (1991).

42. Heitman, J., Movva, N. R. & Hall, M. N. Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. *Science* **253**, 905–909 (1991).

43. Litvintseva, A. P., Thakur, R., Vilgalys, R. & Mitchell, T. G. Multilocus sequence typing reveals three genetic subpopulations of *Cryptococcus neoformans* var. *grubii* (serotype A), including a unique population in Botswana. *Genetics* **172**, 2223–2238 (2006).

44. Kwon-Chung, K. J., Varma, A., Edman, J. C. & Bennett, J. Selection of *ura5* and *ura3* mutants from the two varieties of *Cryptococcus neoformans* on 5-fluoroorotic acid medium. *J. Med. Vet. Mycol.* **30**, 61–69 (1992).

45. Edman, J. C. & Kwon-Chung, K. J. Isolation of the *URA5* gene from *Cryptococcus neoformans* var. *neoformans* and its use as a selective marker for transformation. *Mol. Cell. Biol.* **10**, 4538–4544 (1990).

46. Billmyre, R. B., Clancey, S. A., Li, L. X., Doering, T. L. & Heitman, J. 5-fluorocytosine resistance is associated with hypermutation and alterations in capsule biosynthesis in *Cryptococcus*. *Nat. Commun.* **11**, 127 (2020).

47. Janbon, G. *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).

48. Fan, Y. & Lin, X. Multiple applications of a Transient CRISPR-Cas9 Coupled with Electroporation (TRACE) system in the *Cryptococcus neoformans* species complex. *Genetics* **208**, 1357–1372 (2018).

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

50. Rahnama, M. *et al.* Transposon-mediated telomere destabilization: a driver of genome evolution in the blast fungus. *Nucleic Acids Res.* **48**, 7197–7217 (2020).

51. Mason, J. A. & Biessmann, H. The unusual telomeres of *Drosophila*. *Trends Genet.* **11**, 58–62 (1995).

52. Pardue, M. Lou & DeBaryshe, P. G. Retrotransposons provide an evolutionarily robust non-telomerase mechanism to maintain telomeres. *Annu. Rev. Genet.* **37**, 485–511 (2003).

53. Ashe, A. *et al.* A deletion polymorphism in the *Caenorhabditis elegans* RIG-I homolog disables viral RNA dicing and antiviral immunity. *eLife* **2**, e00994 (2013).

54. Félix, M.-A. *et al.* Natural and experimental infection of *Caenorhabditis* nematodes by novel viruses related to nodaviruses. *PLoS Biol.* **9**, e1000586 (2011).

55. Clancey, S. A., Ruchti, F., Leibundgut-Landmann, S., Heitman, J. & Ianiri, G. A novel mycovirus evokes transcriptional rewiring in the fungus *Malassezia* and stimulates beta interferon production in macrophages. *mBio* **11**, e01534-20 (2020).

56. Kanafani, Z. A. & Perfect, J. R. Resistance to antifungal agents: mechanisms and clinical impact. *Antimicrob. Resist.* **46**, 120–128 (2008).

57. Weiberg, A. *et al.* Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science* **342**, 118–123 (2013).

58. McClintock, B. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. U. S. A.* **36**, 344–355 (1950).

59. Rosenberg, A. *et al.* Antifungal tolerance is a subpopulation effect distinct from resistance and is associated with persistent candidemia. *Nat. Commun.* **9**, 2470 (2018).

60. Berman, J. & Krysan, D. J. Drug resistance and tolerance in fungi. *Nat. Rev. Microbiol.* **18**, 319–331 (2020).

61. Devos, K. M., Brown, J. K. M. & Bennetzen, J. L. Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*. *Genome Res.* **12**, 1075–1079 (2002).

62. Vitte, C. & Panaud, O. LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model. *Cytogenet. Genome Res.* **110**, 91–107 (2005).

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

64. Barton, E. S. *et al.* Herpesvirus latency confers symbiotic protection from bacterial infection. *Nature* **447**, 326–329 (2007).

65. Priest, S. J. *et al.* Factors enforcing the species boundary between the human pathogens *Cryptococcus neoformans* and *Cryptococcus deneoformans*. *PLoS Genet.* **17**, e1008871 (2021).

66. Radchenko, E. A., McGinty, R. J., Aksenova, A. Y., Neil, A. J. & Mirkin, S. M. Quantitative analysis of the rates for repeat-mediated genome instability in a yeast experimental system. *Methods Mol. Biol.* **1672**, 421–438 (2018).

67. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).

68. Pitkin, J. W., Panaccione, D. G. & Walton, J. D. A putative cyclic peptide efflux pump encoded by the *TOXA* gene of the plant-pathogenic fungus *Cochliobolus carbonurn*. *Microbiology* **142**, 1557–1565 (1996).

69. Sun, S., Priest, S. J. & Heitman, J. *Cryptococcus neoformans* mating and genetic crosses. *Curr. Protoc. Microbiol.* **53**, e75 (2019).

70. Yadav, V., Sun, S., Coelho, M. A. & Heitman, J. Centromere scission drives chromosome shuffling and reproductive isolation. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 7917–7928 (2020).

71. Bao, W., Kojima, K. K. & Kohany, O. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* **6**, 11 (2015).

72. Smit, A., Hubley, R. & Green, P. RepeatMasker Open-4.0. *http://www.repeatmasker.org* (2013).

73. Storer, J., Hubley, R., Rosen, J., Wheeler, T. J. & Smit, A. F. The Dfam community resource of transposable element families, sequence models, and genome annotations. *Mob. DNA* **12**, 2 (2021).

74. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

75. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

76. Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. *arXiv* 1207.3907 [q-bio.GN] (2012).

77. Roth, C., Sun, S., Billmyre, R. B., Heitman, J. & Magwene, P. M. A high-resolution map of meiotic recombination in *Cryptococcus deneoformans* demonstrates decreased recombination in unisexual reproduction. *Genetics* **209**, 567–578 (2018).

78. Xu, S. Genetic mapping and genomic selection using recombination breakpoint data. *Genetics* **195**, 1103–1115 (2013).

79. Churchill, G. A. & Doerge, R. W. Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963–971 (1994).

80. Visscher, P. M., Thompson, R. & Haley, C. S. Confidence intervals in QTL mapping by bootstrapping. *Genetics* **143**, 1013–1020 (1996).

81. Arras, S. D. M., Chitty, J. L., Blake, K. L., Schulz, B. L. & Fraser, J. A. A genomic safe haven for mutant complementation in *Cryptococcus neoformans*. *PLoS One* **10**, e0122916 (2015).

82. Fang, Y., Cui, L., Gu, B., Arredondo, F. & Tyler, B. M. Efficient genome editing in the oomycete *Phytophthora sojae* using CRISPR/Cas9. *Curr. Protoc. Microbiol.* **44**, 21A.1.1-21A.1.26 (2017).

83. Andrews, S. FastQC: a quality control tool for high throughput sequence data. *https://www.bioinformatics.babraham.ac.uk/projects/fastqc/* (2010).

84. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).

85. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).

86. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

87. Dahlmann, T. A. & Kück, U. Dicer-dependent biogenesis of small RNAs and evidence for microRNA-like RNAs in the penicillin producing fungus *Penicillium chrysogenum*. *PLoS One* **10**, e0125989 (2015).

88. Ianiri, G. *et al.* Mating-type-specific ribosomal proteins control aspects of sexual reproduction in *Cryptococcus neoformans*. *Genetics* **214**, 635–649 (2020).

Figures

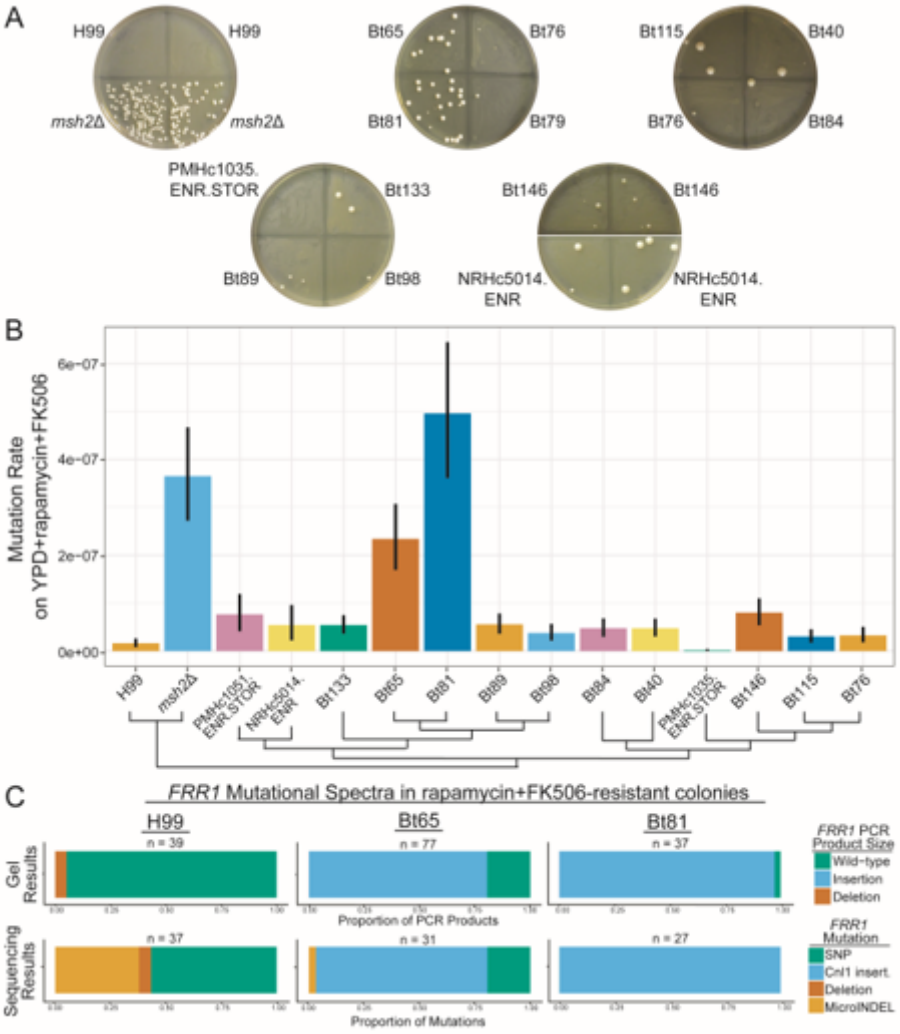

Figure 1

Hypermutation in Bt65 and Bt81 is driven primarily by the insertion of Cnl1 into FRR1. (A) Generation of spontaneously resistant colonies on YPD + rapamycin + FK506 medium was utilized to identify hypermutator candidates; pictures of representative plates are shown. Strains include the phylogenetically closely related strains involved in fluctuation assay in B as well as positive (msh2Δ) and negative (H99) controls. (B) Mutation rates of closely related VNBII strains and controls on YPD + rapamycin + FK506. Bars represent the mutation rate and error bars represent 95% confidence intervals; mutation rates represent the number of mutations per cell per generation. Schematic depicts the phylogenetic relationships of all strains included in fluctuation analyses based on Desjardins et al. 201718. Mutational spectra in FRR1 in YPD + rapamycin + FK506-resistant colonies of H99, Bt65, and Bt81 as characterized by (C) gel electrophoresis and Sanger sequencing of FRR1 PCR products. MicroINDELs are defined as insertions or deletions < 50 bp. All mutations are relative to the appropriate rapamycin + FK506-sensitive parental strain.

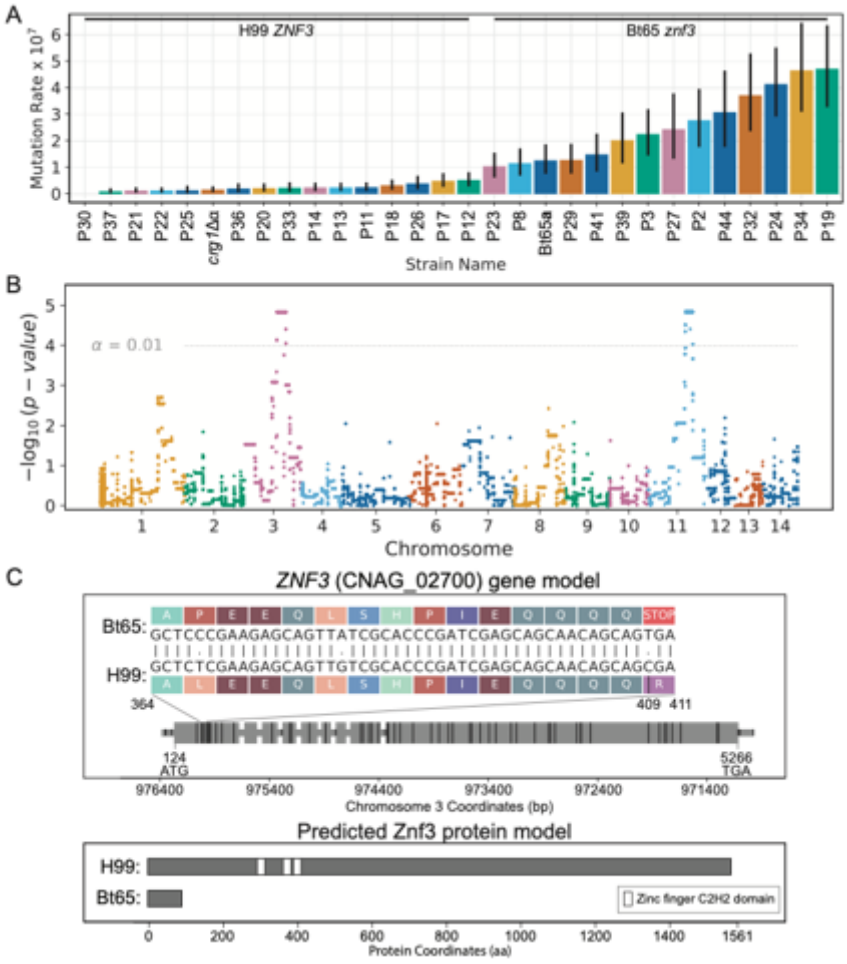

Figure 2

QTL analysis of hypermutator phenotype. (A) Quantification of mutation rates on YPD + rapamycin + FK506 medium – sorted smallest to largest, left to right – for F1 progeny and parental strains, H99 crg1Δ and Bt65. Inheritance of the Bt65 znf3 allele or H99 crg1Δ ZNF3 allele in F1 progeny is indicated above mutation rates. Colored bar plots and vertical black lines depict the mean mutation rate and associated 95% confidence intervals (CI) per segregant. Mutation rates represent the number of mutations per cell per generation. (B) Manhattan plot showing the strength in association (y-axis) between bi-allelic SNPs and hypermutator phenotype, across the 14 chromosomes (x-axis). Colors separate SNPs across chromosomes. The permutation-based significance threshold (α = 0.01) is depicted with a horizontal dashed line. (C) Predicted ZNF3 gene and Znf3 protein models in H99 and Bt65. A grey horizontal bar depicts the gene body in the upper panel, and larger grey rectangles represent exons; the gene is depicted 5’ to 3’ and is 5417 nt in length. The locations of SNPs differing between Bt65 and H99 are shown by vertical black rungs along the gene model. Amino acids specified by mRNA codons in the indicated region of ZNF3 exon 1 (nucleotides 364 to 411) are shown for H99 and Bt65 to illustrate the effect of the C to T mutation (nucleotide 409) predicted to cause a nonsense mutation in Bt65. The bottom panel depicts the predicted impact of the nonsense mutation on the Znf3 protein in Bt65. White rectangles along the protein schematic depict the three C2H2-type zinc finger domains of Znf3.

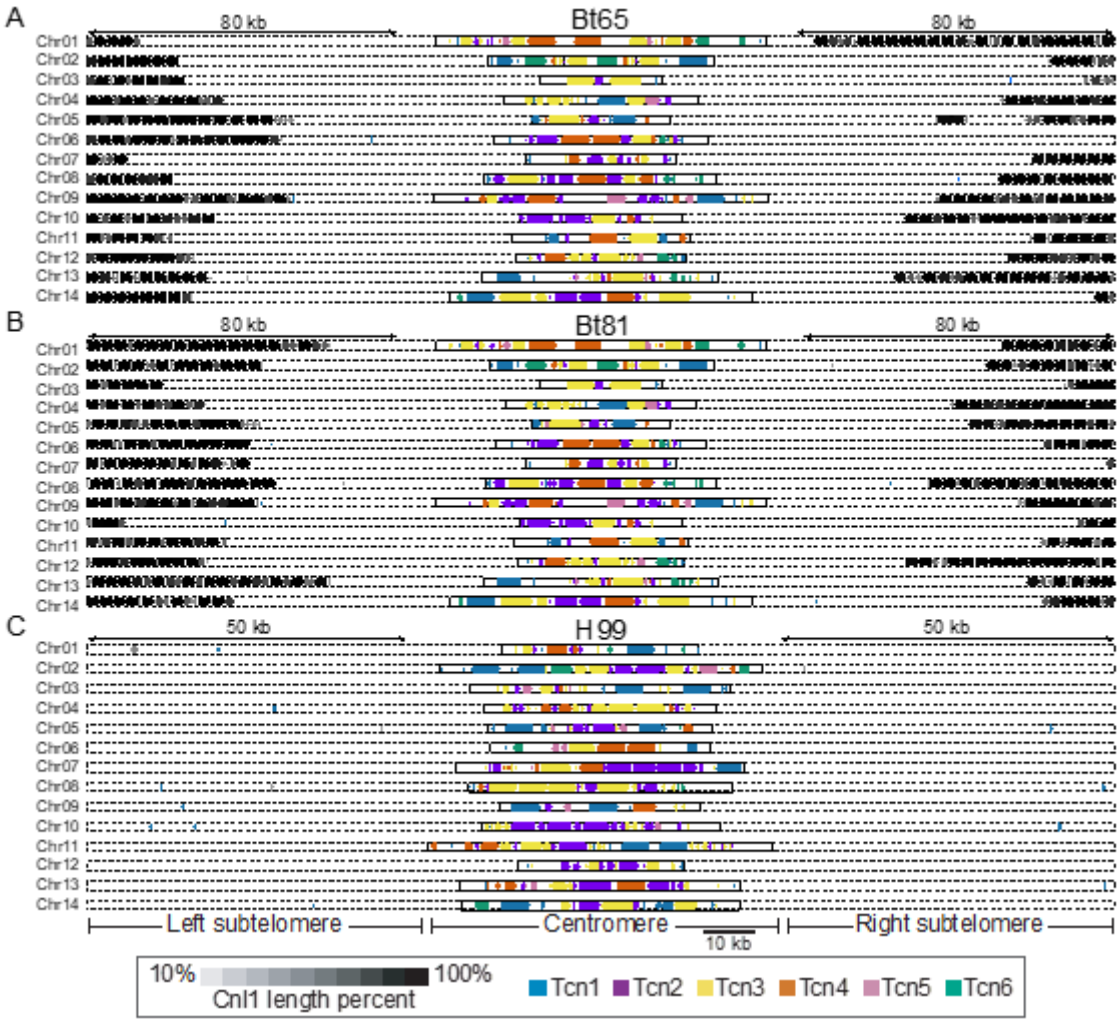

Figure 3

Retrotransposon content in the genomes of H99, Bt65, and Bt81. Distributions of the Tcn1 through Tcn6 LTR-retrotransposons and the Cnl1 non-LTR retrotransposon in subtelomeric and centromeric regions of (A) Bt65, (B) Bt81, and (C) H99 genomes depicted in Figure S6. In Bt65 and Bt81, 80 kb of subtelomeric regions are displayed, and 50 kb subtelomeric regions are displayed for H99 to show the full distribution of subtelomeric Cnl1 elements. Subtelomeric arrays of Cnl1 are depicted at the end of each chromosome in Bt65 and Bt81, while only 7 Cnl1 elements are localized subtelomerically in H99. Shading corresponds to fragments of the Cnl1 elements, and gene arrowheads indicate the direction of transcription for all retrotransposons.

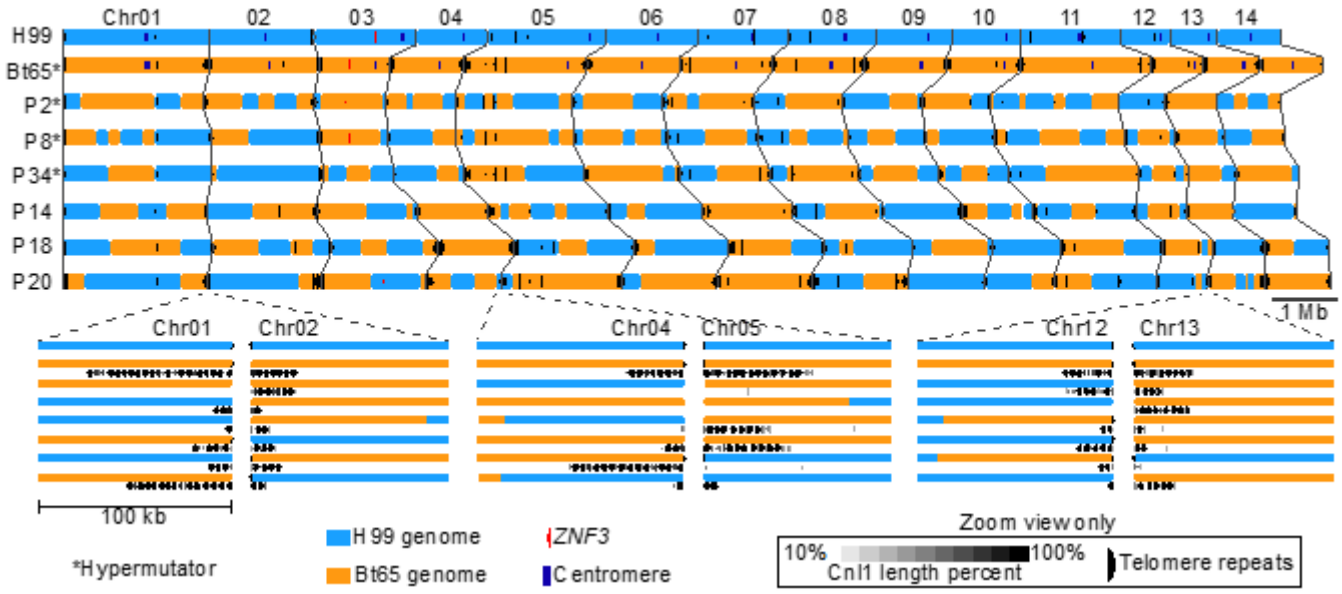

Figure 4

Genetic recombination sites and Cnl1 distribution in Bt65 x H99 F1 progeny. Recombination sites along each of the 14 chromosomes for the six Bt65a x H99α F1 progeny for which long-read whole-genome sequencing was conducted. Genomic loci depicted in blue were inherited from the H99 parent, and orange genomic loci were inherited from the Bt65 parent. Cnl1 elements throughout the F1 progeny and parental genomes are indicated by black arrowheads in the upper panel. Centromeres are indicated by dark blue boxes in only the parental genomes. Hypermutator F1 progeny are indicated with asterisks, and the ZNF3 locus is indicated in each strain with a red arrowhead. Regions enlarged below illustrate Cnl1 subtelomeric arrays on several chromosomes and depict examples of Cnl1 array expansion (e.g. Chromosome 4, P18), contraction (e.g. Chromosome 1, P14), and invasion of naïve H99 subtelomeres (e.g. Chromosome 1, P8). Telomeric repeat sequences are indicated by black half circles only in the enlarged panels.

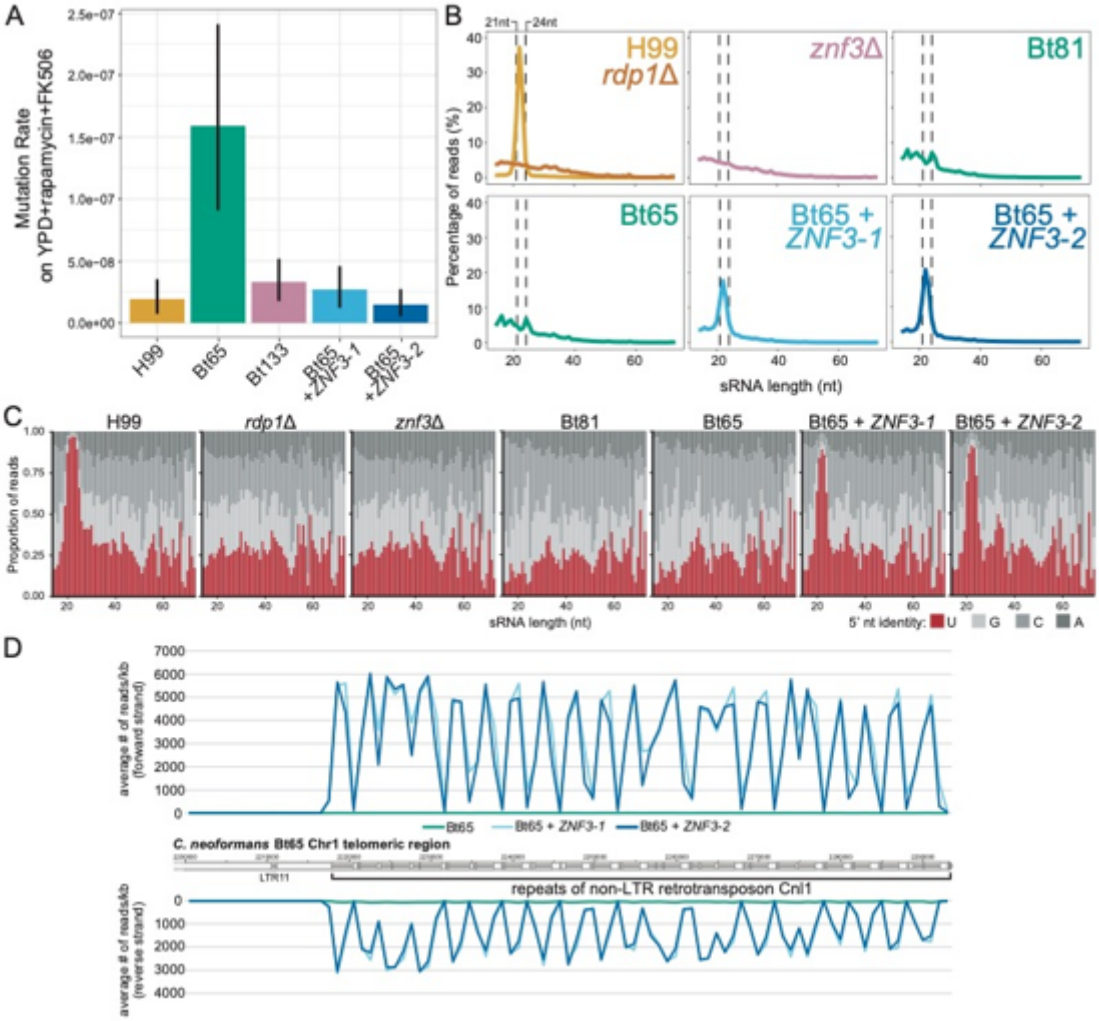

Figure 5

ZNF3 complementation in Bt65 significantly reduces mutation rates and restores siRNA production. (A) Mutation rates of H99, Bt65, Bt133, and the two independent ZNF3 complementation mutants, Bt65 + ZNF3-1 and Bt65 + ZNF3-2, on YPD + rapamycin + FK506 medium. Bars represent mutation rate (number of mutations per cell per generation) and error bars represent 95% confidence intervals. (B) Size distributions of sRNA reads from each indicated strain. Dashed vertical lines indicate the 21 to 24 nucleotide size range, the characteristic sizes of siRNAs produced by the RNAi pathway. (C) Proportion of sRNA reads (y-axis) with the indicated 5’ nucleotide identity (color of stacked bar) at each sRNA read size (x-axis). siRNAs produced by the RNAi pathway characteristically have a 5’ uracil nucleotide. (D) Quantification of sense and antisense sRNAs from Bt65, Bt65 + ZNF3-1, and Bt65 + ZNF3-2 aligning to an array of subtelomeric Cnl1 elements on Chromosome 1 of Bt65. Transposable elements along the chromosome are indicated by dark grey boxes, while intergenic regions are light grey.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1StrainsScreened.xlsx
- TableS2QTLvariants.xlsx
- TableS3CENlengthsandstats.docx
- TableS4sRNAdata.xlsx
- TableS5Strains.docx
- TableS6FluctuationData.xlsx
- TableS7Oligonucleotides.docx
- TableS8SangerSequencing.xlsx
- TableS9ProgenyPloidy.docx
- Supplement.docx