Centromere scission drives chromosome shuffling and reproductive isolation

Vikas Yadav, Sheng Sun, Marco A. Coelho, and Joseph Heitman*

Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA

*Corresponding author

Joseph Heitman
Box 3546, 322 CARL Building, Research Drive
Department of Molecular Genetics and Microbiology
Duke University Medical Center
Durham, NC 27710
Phone: 919-684-2824
Fax: 919-684-2790
Email: heitm001@duke.edu

Running title: Chromosome shuffling mirrors species boundaries

Keywords: *Cryptococcus neoformans*; DSB repair; Chromosome translocation; Karyotype evolution.

Abstract

A fundamental characteristic of eukaryotic organisms is the generation of genetic variation via sexual reproduction. Conversely, significant large-scale genome variation could hamper sexual reproduction, causing reproductive isolation and promote speciation. The underlying processes behind large-scale genome rearrangements are not well understood and include chromosome translocations involving centromeres. Recent genomic studies in the *Cryptococcus* species complex revealed that chromosome translocations generated via centromere recombination have reshaped the genomes of different species. In this study, multiple DNA double-strand breaks (DSBs) were generated via the CRISPR/Cas9 system at centromere-specific retrotransposons in the human fungal pathogen *Cryptococcus neoformans*. The resulting DSBs were repaired in a complex manner, leading to the formation of multiple inter-chromosomal rearrangements and new telomeres. The newly generated strains harboring chromosome translocations exhibited normal vegetative growth but failed to undergo successful sexual reproduction with the parental wild-type strain. One of these strains failed to produce any spores, while another produced ~3% viable progeny. The germinated progeny exhibited aneuploidy for multiple chromosomes and showed improved fertility with both parents. All chromosome translocation events were accompanied without any detectable change in gene sequences and thus, suggest that chromosomal translocations alone may play an underappreciated role in the onset of reproductive isolation and speciation.

Introduction

Chromosomes are prone to undergo several rearrangement events, including fusion, fission, deletion, and segmental duplication. In some cases, one chromosome segment is exchanged with another to generate chromosomal translocations. Such exchanges between homologs are regularly observed during meiosis when homologous chromosomes exchange arms via meiotic recombination[1]. Chromosome rearrangements can also occur during mitosis, but in a less well-regulated manner, and sometimes as a result of disease conditions like cancer[2,3]. Additionally, rearrangements can also occur within a single chromosome. As a result, chromosome rearrangements during mitosis can cause mutations, gene disruption, copy number variations, as well as alter the expression of genes near the breakpoints[4]. Cancer cells show a high level of chromosome rearrangements as compared to healthy cells, and this contributes to critical pathological conditions observed in these cells, such as activation of oncogenes[5,6].

Chromosomal translocations are initiated by double-strand breaks (DSBs) in DNA[7]. Rearrangements involving a single DSB are mainly repaired by the invasion of the broken DNA molecule into a homologous DNA molecule, in a process termed homologous recombination (HR)[6,8]. This invasion can lead to the exchange of DNA between the two molecules of DNA, leading to reciprocal crossover or gene conversion[2]. These types of rearrangements occur during meiosis and are regulated to give rise to an error-free repaired sequence. Other types of chromosomal translocation involve two or more DNA DSB sites, which are then fused in an error-prone mechanism known as non-homologous end joining (NHEJ)[9]. The two sites can be present on either the same chromosome or different chromosomes. Repair of two DSBs present on the same chromosome can result in the deletion of the intervening sequence or inversion if the sequence is rejoined in reverse orientation[7]. On the other hand, the fusion of DSBs from two

different chromosomes can result in chromosomal translocation. In some instances, other repair pathways like microhomology-mediated end joining (MMEJ) or alternative End Joining (alt-EJ) also participate in the repair of DSB ends[9,10].

The occurrence of multiple DSBs in mitotically growing cells at the same time is rare but occurs at a relatively higher frequency in cancer cells. Such events could occur in a genome due to replication defects or exogenous factors such as ionizing radiation or chemotherapeutic agents[11]. Multiple DSBs also occur naturally during processes like V(D)J recombination[12]. The occurrence of multiple DSBs can also be induced in the micronuclei of cancer cells[13]. Micronuclei are small nuclei harboring one or a few chromosomes and are generated as a result of a mitotic failure. These small nuclei act as hotspots of chromosome fragmentation, where multiple DSB sites are almost simultaneously generated and subsequently rejoined in a random order, a process known as chromothripsis. Previous reports have suggested that the occurrence of multiple DSBs alters HR pathways leading to NHEJ or alt-EJ mediated repair[14,15]. Because both NHEJ and alt-EJ are error-prone, they lead to a significant increase in mutation at the repair junctions and also randomly join broken fragments.

Apart from generating mutations and gene disruptions, chromosome translocations can also result in reproductive isolation during meiosis and facilitate speciation[16]. The presence of multiple rearrangements between the two homologous chromosomes from the parents leads to failures in chromosome pairing during meiosis or crossovers that result in loss of essential genes[17]. Meiosis that is defective in this way will result in the production of progeny with abnormal genome content. In fungi, the parental nuclei fuse and undergo meiosis before sporulation that gives rise to progeny. Thus, defects in meiosis lead to the production of spores with abnormal or incomplete genetic compositions rendering them inviable. *Cryptococcus*

neoformans is a basidiomycete fungus that largely infects immunocompromised humans causing Cryptococcal meningoencephalitis[18-20]. *C. neoformans* harbors a 19 Mb genome with 14 chromosomes[21]. While most of the genome is devoid of repeat regions, centromeres in *C. neoformans* are rich in a set of retrotransposons that are shared across multiple centromeres[22,23]. *C. neoformans* also has a meiotic cycle making it a suitable model organism to study meiosis and sexual development[20,24]. It exhibits two mating types, α and **a**, defined by the *MAT*α and *MAT***a** alleles at the single mating-type locus[25,26].

In this study, we exploited the genomic features of *C. neoformans* to study the impact of chromosome translocations on reproductive isolation. First, retrotransposons present in centromeres were targeted with CRISPR, generating multiple DSBs simultaneously. Next, the presence of chromosome rearrangements was screened by Pulsed-Field Gel Electrophoresis (PFGE), and isolates with multiple chromosomal translocations were identified. The genomes of these strains were assembled based on long-read nanopore sequencing to characterize the chromosome rearrangements. Although the strains with new karyotypes did not exhibit growth defects as compared to the wild-type, the impact of chromosomal rearrangements had a profound effect on sexual reproduction. These findings demonstrate that *C. neoformans* can tolerate multiple chromosomal translocations, but that such large-scale changes can cause reproductive isolation, and promote incipient speciation.

Results

Simultaneous breaks at multiple centromeres lead to chromosome shuffling

Centromeres in *C. neoformans* were previously identified and shown to possess multiple retrotransposons named Tcn1-Tcn6[21,22]. These elements are present only in centromeres, albeit distributed randomly and non-uniformly across all 14 centromeres. We specifically targeted the Tcn2 element using a guide RNA (gRNA) that would cleave nine centromeres, generating a total of 18 DSBs (Figure 1A). We hypothesized that these DSBs would be repaired to generate chromosomal translocations (Figure 1B).

After the transformation of wild-type cells with three DNA fragments, each constitutively expressing a dominant selection marker (G418), Cas9, and the gRNA, colonies were obtained and screened via pulsed-field gel electrophoresis (PFGE) (Figure S1). We found karyotypic variation for at least one chromosome in three out of twelve colonies screened (Figure S2A). Of the three, two (VYD135 and VYD136) harbored more than three changes in chromosome banding pattern as compared to the wild-type strain *C. neoformans* H99 (Figure 1C and S2A). To characterize these chromosomal alterations, we sequenced the genomes of these two strains using Oxford Nanopore sequencing and were able to assemble them into 17 contigs each (Figure S2B).

A synteny block comparison of these genomes with that of the wild-type revealed the presence of multiple translocations. However, some of these contigs were broken with the centromere at one of the ends. Additionally, Illumina sequencing of these strains revealed a near euploid genome for both of these strains (Figure S2C) except for the smaller arm of chromosome 13 in VYD136 that was present in two copies. To resolve the status of incompletely assembled contigs, Southern blots of PFGE-separated chromosomes (chromoblots) followed by

hybridization with chromosome-specific probes were conducted for multiple chromosomes (Figure S3). Contigs validated to be arms of the same chromosome by this approach were fused manually with a 50-bp sequence gap to generate a chromosome-size scaffold. The presence of telomere repeat sequences at both the ends of these scaffolds suggests that these scaffolds represent full-length chromosomes. The integrity of each genome assembly was further verified by mapping the respective nanopore reads to the genome assembly. Thus, using nanopore and Illumina sequencing, and chromoblot analysis, we were able to assemble the genomes of isolates VYD135 and VYD136 to the chromosome level (Figure 1D). This analysis also revealed that the duplicated arm of chromosome 13 of isolate VYD136 exists as an isochromosome with two broken centromeres fused with each other. Overall, these results show that multiple breaks at centromeres can lead to karyotype shuffling in *C. neoformans* that is tolerated by the organism.

Centromere breaks generate new telomeres and increase the number of chromosomes

No species in the *Cryptococcus* species complex has been observed to harbor chromosomes smaller than 500 kb. However, a genome-level assembly of VYD135 and VYD136 revealed the presence of 2 to 3 chromosomes that are shorter than the shortest naturally occurring chromosome of wild-type H99 (chromosome 13 of 757 kb length). Bands of the expected size for these chromosomes were observed in PFGE, indicating that these chromosomes are *bona fide* and not a result of assembly error (Figure 2A). In addition to these small chromosomes, a few novel features were observed for the genomes of VYD135 and VYD136. Three of the new chromosomes had generated *de novo* telomere sequences on one of the ends, next to the broken centromeres (chromosomes 13 and 15 of VYD135; chromosome 13 of VYD136) (Figure 2B, and S4A). While in two cases, the telomere repeat sequences were

present next to the Tcn2 elements, in the third multiple copies of the Cas9 sequence were found to be inserted between the Tcn2 element and telomere repeat sequences (Chr15 of VYD135). Our analysis of these regions did not reveal any common motif suggesting that addition of telomere repeat sequences independently occurred on random sequences.

Centromere locations in these two genome assemblies were defined based on synteny with centromere flanking regions (Table S1). Specifically, BLASTn analysis with centromere-flanking ORF sequences as query sequences were used to identify the syntenic regions defining the boundary for centromeres. Next, Tcn elements (Tcn1-6) were mapped across centromeres of the new strains, VYD135 and VYD136. Surprisingly, foreign DNA elements, such as Cas9 and neomycin resistance gene sequence, were found to have integrated into multiple centromeres (Figure 2C and D). Both of these sequences were introduced as linear DNA molecules during the CRISPR transformation experiment. In some cases, Cas9 and the neomycin gene sequences were present in multiple copies and in a random order/orientation. Further analysis revealed that these elements were present at the junction where two parental centromeres fused with each other. This result suggests that these foreign sequences may have assisted in joining the broken centromeres to form hybrid centromeres.

A comparison of centromere length between H99, VYD135, and VYD136 revealed the presence of some significantly shorter and longer centromeres in both of these new strains (Figure S4B). However, the reduction in centromere length did not seem to confer a visible growth defect indicating no change in centromere function due to a decrease or increase of length (Figure S5A). These strains also did not show any difference when tested for various stress conditions like temperature, fluconazole, or DNA damaging agent like Phleomycin (Figure S5A). When grown with another strain expressing NAT resistant gene, to test competitive

fitness, both VYD135 and VYD136 did not show any significant difference as compared to wild-type H99 (Figure S5B). Additionally, these two strains do not exhibit any difference in virulence in the *Galleria* model of infection[27] (Figure S5C). Both VYD135 and VYD136 lead to lethal infection of *Galleria* with the same efficiency as the wild-type H99α and KN99**a** isolates. Taken together, these results suggest that the changes in chromosome structure do not grossly affect mitotic growth or infectivity of *C. neoformans*.

Chromosome shuffling is driven by the Cas9 induced breaks

Isolates VYD135 and VYD136 show inter-chromosomal rearrangements between seven and eight chromosomes, respectively. The rearrangements in the two strains are not identical, although they do involve the same set of chromosomes, suggesting that both of these strains underwent recombination via different routes (Figure 3A and B). Next, the breakpoints in each of the centromeres were defined by mapping nanopore reads to the wild-type genome. One caveat is that long reads from nanopore sequencing might not align precisely with the parental centromere sequence due to the loss of regions of the original sequence. Many reads were found that mapped to a single location next to the gRNA cleavage site in almost all of the centromeres that underwent recombination (Figure 3C and S6). In a few cases, the reads did not map to the cleavage site flanking sequences suggesting deletions occurred during recombination. This loss of sequence is prominent in centromeres with three or more cleavage sites (*CEN10* in Figure 3C). This mapping pattern suggests these non-essential, small fragments were lost, and their loss does not compromise centromere function. Interestingly, among the chromosomes that contain Tcn2 elements and could be targeted by Cas9 with the Tcn2-specific gRNA, chromosome 11 was not found to be involved in recombination in either of the two strains, even though *CEN11* is

predicted to be cleaved once. It is possible that the gRNA cleavage site prediction for *CEN11* could be a result of an incorrect genome assembly or sequence error. Overall, this analysis supports that recombination was initiated by the Cas9 DSBs. Also, that centromeres lacking Tcn2 elements were not involved in recombination reflects the specificity of Cas9 and the repair machinery.

Multiple types of repair machinery were involved in the DSB repair process

Next, the new centromere sequences were compared with the original centromeres in a pairwise fashion to understand the repair process (Figure 4 and S7). For this purpose, we utilized our newly generated assembly of the wild-type strain, which showed better coverage for centromere sequences (See Materials and Methods for details). This analysis provides evidence that both NHEJ and HR pathways participated in the repair of these broken ends. Insertion of the *CAS9* sequence and the *NEO* gene sequence conferring G418 resistance seems to be the result of DSB repair via NHEJ in all cases as there was no additional sequence added between the ends of the breaks. Additionally, *CEN12* of VYD136 results from a single fusion event between wild-type *CEN8* and *CEN14* after the DNA breaks (Figure S7). For *CEN14* of VYD135, a sequence of 2.7 kb aligned with both parental centromeres suggesting that the hybrid centromere formed as a result of homologous recombination. *CEN8* of VYD135 exhibited evidence for repair via multiple mechanisms, including NHEJ, microhomology, and invasion into multiple different centromeres (Figure 4). Because all of the sequences involved have a Tcn2 sequence at the ends, it is not possible to infer the order of these events.

A comparison of wild-type *CEN7* and VYD135 *CEN11* shows evidence of resection beyond the DSB sites. Resection was probably followed by strand invasion into broken pieces of

CEN7 that were released due to multiple DSBs, and a final fusion event with *CEN8* leading to the formation of hybrid centromeres. Similarly, *CEN13* of VYD136 seems to have arisen as a result of resection followed by invasion into multiple centromeres before adding telomere sequences at the end (Figure S7). In addition to these multiple recombination events at these junctions, we also observed inversion events for sequences that were released due to multiple DSBs in a single centromere (*CEN15* of VYD135). On the other hand, the inversion in VYD135 *CEN13* has signatures of invasion into another intact copy of *CEN13*, resulting in the inversion.

Apart from inter-chromosome recombination, we also observed intra-chromosomal recombination in *CEN12* of VYD135. Wild type *CEN12* has two cleavage sites separated by a 10 kb sequence. In VYD135, *CEN12* is smaller due to the absence of the 10 kb sequence, and two flanking sequences were joined with an overlap of 640 bp (Figure 4). This event also led to a reduction in centromere length for *CEN12*, shortening it significantly (21 kb versus 31 kb in the wild type). Combined, these results suggest that both HR and NHEJ processes repair DSBs at the centromeres. Given the high level of identity shared between Tcn2 elements present among multiple centromeres, other plausible routes to these rearrangements are possible.

Strains with chromosomal rearrangements fail to undergo normal meiosis

Chromosome shuffled strains, VYD135 and VYD136, exhibit multiple chromosomal translocations compared to the wild-type karyotype, as described above. We hypothesized that this would lead to incompatibility during meiosis and defects in producing viable spores. To study this, the two shuffled strains were crossed with the wild-type strain, KN99**a**, which is congenic with the parental strain H99α. During *Cryptococcus* sexual reproduction, cells of two opposite mating types fuse and then produce hyphae, basidia, and long chains of spores over a

two-week course of incubation[26]. In the basidium, the two parental haploid nuclei fuse and undergo meiosis producing four haploid nuclei. These meiotic products subsequently undergo repeated rounds of mitosis that produce four spore chains from the surface of the basidium by budding. A pairing defect during meiosis would compromise the segregation of chromosomes, giving rise to progeny with incomplete genomes.

After two weeks of mating, intact spore chains were observed by light microscopy in the H99α x KN99**a** wild-type cross. On the other hand, the VYD135α x KN99**a** cross formed basidia but showed almost no sporulation, whereas the VYD136α x KN99**a** cross was largely indistinguishable from wild-type and did not show a prominent sporulation defect (Figure 5A). Meiotic structures were further examined by scanning electron microscopy (SEM). Abundant spores were produced in the H99α x KN99**a** cross, whereas no spore chains were observed in the VYD135 x KN99**a** cross (Figure 5B). The VYD136 x KN99**a** cross showed defective sporulation, where some basidia formed fewer spores while others had none (Figure 5B). Next, the germination rate (equivalent to spore viability) of spores from the H99α x KN99**a** and VYD136 x KN99**a** crosses was assessed. The spore germination rate for the VYD136 x KN99**a** cross was only 3% compared to wild-type with 80 to 90% spore germination[28,29]. The lack of sporulation in the VYD135 x KN99**a** cross and the reduced spore germination rate for the VYD136 x KN99**a** cross provide evidence that these strains fail to undergo normal meiosis. The translocations in VYD135 involve seven chromosomes, and in VYD136 eight chromosomes are involved, yet these isolates show very different phenotypes during meiosis. We hypothesize that this observed difference in the sporulation phenotype may be attributable to different chromosome configurations observed in these two strains. Specifically, different configurations

could lead to differences in meiotic pairing and could cause more severe meiotic defects in one compared to the other. The failure to undergo meiosis will lead to no or defective sporulation.

VYD136 backcross progeny are aneuploid and exhibit improved sexual reproduction

Three viable F1 progeny were isolated from the cross of VYD136 with KN99**a**. These F1 progeny (P1, P2, P3) were characterized to further understand the impact of chromosomal translocations on meiosis by assessing mating type and ploidy (Figure 6A and B). Flow cytometry analysis revealed that two of the progeny (P1 and P2) are aneuploid. To investigate this further, the relative chromosome copy number for the three progeny was determined from read counts after Illumina and nanopore sequencing, revealing that they are aneuploid for multiple chromosomes (Figure 6C). P1 and P2 are both aneuploid for parts of chromosomes 5, 12, 13, and 14, whereas P2 is also aneuploid for the entire chromosome 6. The third progeny (P3) also showed aneuploidy but only for the shorter arms of chromosomes 12 and 13.

These three F1 progeny were backcrossed to both parents (VYD136 and KN99**a**) and the wild-type H99α. All three progeny mated with strains of opposite mating type, as expected. Surprisingly, spore dissections from these crosses revealed a much higher spore germination rate for the three progeny (11 to 58%) as compared to their parental cross (3% germination rate) (Figure 6A). P1 and P2 exhibited a much higher (>50%) germination rate for all crosses as compared to P3, which showed 11 and 21% germination rates. The segmental aneuploidy of these isolates may explain their higher germination rate compared to the parent VYD136.

The presence of copy number changes for only one arm of most chromosomes suggested these progeny may have a mixed karyotype. To address this, genomes for the three progeny were assembled using nanopore sequencing. Due to aneuploidy, these genomes were not assembled

completely and harbored multiple breaks (Figure S8A and B). However, based on their ploidy profiles, mapping of nanopore reads, as well as synteny comparison with both parents, their genome configurations were largely resolved (Figure 6D). The final karyotype shows that most chromosomes match either with the wild-type or the VYD136 chromosome profile (Figure 6D and S8A). Based on this, we propose that segmental aneuploidy may contribute to overcome the reproductive barrier that might otherwise arise due to multiple changes in chromosome configuration. Genetic transmission via an aneuploid intermediate may yield a transitional population that is more compatible with the ancestral population and slowly diverges to give rise to a new species. On the other hand, a new species might also arise because of a single event of multiple translocations occurring throughout the genome. This process is only one of several factors contributing to reproductive isolation and could play a greater or lesser role dependent upon the species distribution and mode of reproduction.

Discussion

Chromosome rearrangements shape the genome in multiple ways by affecting gene positions, causing mutations, and compromising recombination during meiosis. In this study, chromosomal rearrangements were generated by severing centromeres in a human fungal pathogen, *C. neoformans*. The centromeres in this species are rich in retrotransposons that are common among multiple centromeres[22]. CRISPR/Cas9 targeting of retrotransposons cleaved multiple centromeres at the same time. This approach does not affect gene-rich regions, thus avoiding the risks involved with the loss or mutation of genes. At the same time, this allowed us to study a) how multiple breaks in the genome are tolerated, b) how DSBs in heterochromatic regions such as centromeres are repaired, and c) how the resulting chromosomal rearrangements

impact reproductive isolation. Centromere-mediated chromosomal translocation events have been observed in the *Cryptococcus* species complex[21,30] (Figure S9). Genome comparisons between non-pathogenic *C. amylolentus* and pathogenic species, such as *C. neoformans*, show at least six centromere-mediated translocations[30]. Our findings suggest these translocations could have driven speciation in this species complex, in addition to other factors.

Recent studies have suggested that centromeres can undergo recombination, at least mitotic recombination, as compared to previous models in which centromeres were recalcitrant to recombination[22,31-36]. Our results further support that centromeres can undergo recombination when DSBs are generated in these regions. Each of the broken ends was processed and subjected to HR or NHEJ mediated repair and recombination. We also observed a case where a centromere, following the generation of DSBs, recombined, causing the loss of intermediate DNA sequences present between the two ends. In similar events, many fragments of centromere sequences were either lost or fused within other centromeres, altering the architecture of centromeres in this species (Figure 7). This result suggests that variation in centromere structure is not critical for centromere function, but might play some other role in genome organization. This conclusion is also supported by a previous study that showed a significant reduction in centromere length in *C. deuterogattii*, a closely related species of *C. neoformans*[22]. This variation in centromere architecture also raises questions about the requirement of longer centromeres in most organisms. According to previously proposed models, longer centromeres with more centromere repeats might bind more kinetochore proteins[37-39]. This would enable stronger kinetochore-microtubule interactions for longer centromeres and favor transmission of longer centromeres during meiosis. Our system can be harnessed to understand the dynamics of centromere transmission and test this hypothesis in future studies.

Interestingly, we also observed the addition of *de novo* telomere repeat sequences adjacent to the broken centromeres. While the mechanism underlying this process remains to be elucidated, the frequency of occurrence of *de novo* telomere addition seems to be high at ~10% (3/32 ends). Additionally, two of the three telomere repeat additions involved invasion or fusion with other sequences before the addition of repeat sequences. *De novo* telomere formation on broken DNA ends has been observed in many species and is commonly known as chromosome healing[40-42]. Chromosome healing occurs during mammalian development and has been observed in mouse embryonic stem cells as well as human germline cells[43-45]. It is also proposed to occur in cancer cells to stabilize the ends of broken chromosome fragments arising as a result of chromothripsis or telomere crisis[46,47]. This process maintains the genome content while increasing the number of chromosomes in the case of DSBs that could not be repaired. In our study, we did not observe a significant impact of telomere formation next to centromeres or the loss of centromere sequences on the growth of the chromosomal shuffled strains as compared to the parental strain. This could suggest that shorter centromere sequences are sufficient to propagate the genome content, and centromeres and telomeres do not influence each other's function.

Repair of a DSB site is a complex process and involves multiple repair machineries including HR and NHEJ. HR mainly takes place during S-phase, whereas NHEJ occurs throughout the cell cycle[48,49]. HR mainly leads to gene conversion and can also drive recombination between repetitive sequences leading to translocation. NHEJ is more error-prone and can join any two sites resulting in translocations. Thus, both of these processes can lead to translocations but can result in very different types of junctions[8,50]. Our analysis of junctions exhibited evidence for both HR and NHEJ. Previous reports suggested that the NHEJ pathway is

preferred over HR in the case of multiple DSBs[14,15]. Although the number is small, our results suggest that both of these pathways might take place at a similar rate in *C. neoformans*. However, the underlying regulatory mechanisms remain a subject for further investigation.

We also observed the inversion of sequences suggesting either strand invasion or simple fusion of a broken piece in reverse order by NHEJ. Our analysis also suggests the occurrence of multi-invasion repair (MIR), a subtype of HR, and micro-homology-mediated BIR (MMBIR), both of which can be favored for the repair of repeat sequences[6,51]. According to the MIR pathway, a single broken end might enter multiple target DNA molecules based on homology[52,53]. The DNA breaks in our experiments were made in the Tcn2 element, which is present in multiple copies across centromeres; thus, it is possible that some of the ends might have been repaired via MIR. Our results also suggest that breaks in repeat molecules are repaired either by NHEJ or HR, where other repeat sequences present in the genome may aid in the repair process. Similar observations have also been made previously in *Saccharomyces cerevisiae*, where DSBs were generated with gamma radiation, HO-endonuclease, or CRISPR[54-56]. In all of these studies, multiple copies of Ty elements were found to be involved in the generation of translocations. In the study that employed CRISPR, similar multiple translocation events were observed[55]. However, Ty elements in *S. cerevisiae* are distributed across the genome, unlike *C. neoformans*, where all Tcn2 elements are clustered in centromeres.

Multiple concurrent breaks in the genome are rare and occur mainly in pathogenic conditions such as cancer. A commonly observed outcome is chromothripsis, where one or several chromosomes are broken into multiple smaller pieces and rejoined randomly to shuffle the targeted chromosomal region[11,13]. This phenomenon is associated with the generation of multiple mutations, as well as the activation of oncogenes. Chromothripsis is initiated by the

generation of multiple breaks in a localized manner. The source of these breaks could be either internal factors such as replication associated breaks or external factors, including ionizing radiation or chemotherapy. These break sites are then repaired in a random manner. While this process is well-known, the mechanisms involved in this process are poorly understood[13,57]. In our experiments, we observed that similar events take place during DSBs repair (Step 3 in Figure 7). The connection between these processes needs to be further explored to establish whether our system can be extended to understand chromothripsis.

Our approach induced multiple simultaneous breaks in the genome, which were then repaired to generate chromosome shuffling. Centromeres are known to cluster in *C. neoformans* during mitosis[58,59], and this may have promoted their interaction during the DSB repair events that generated these alterations. This centromere clustering in *C. neoformans* also mimics the clustering of DSB sites observed during the process of chromothripsis. We posit our approach could provide answers to critical questions regarding chromothripsis. The MIR pathway has been implicated as one pathway contributing to chromothripsis[6,53]. In our study, we also found evidence that the MIR pathway contributes to chromosome shuffling, further suggesting similarities between chromothripsis and the events we observed. Notably, chromothripsis is mainly observed in chromosome arms, whereas centromeres were targeted in our studies[13]. However, some studies have indicated an association between chromothripsis break sites and the presence of transposon sequences[11,60]. In our approach, the breaks were also located within transposon sequences, which are part of the centromeres. Thus, understanding the factors governing this process in *C. neoformans* could also shed light on facets of chromothripsis.

Chromosome rearrangements have been implicated in speciation, acting via reducing fertility or gene flow in the progeny that inherit the translocation[61-64]. However, most models

proposing speciation in this manner also consider geographical isolation as other major criteria[16]. Other models consider chromosome rearrangement as an effect of speciation, which happens as a result of simultaneous rearrangement after the speciation. Studies in marsupials suggested a role of chromosome rearrangements involving centromeres in the process of speciation in this species complex[65-69]. These involved centromere-mediated translocations as well as differences in centromere lengths between two species. In our study, no genes were found to be affected or mutated because all rearrangements were generated via centromere recombination. Thus, a meiotic failure of rearrangement harboring strains suggests that chromosome translocations solely can drive reproductive isolation in a species.

Here, we find that the presence of multiple centromere-mediated chromosomal rearrangements dramatically reduces the efficiency of meiosis and generates post-zygotic reproductive isolation. A complete failure to undergo meiosis with the existing population could lead to loss of new chromosome configurations, as they will not be passed to the new generations. We found that the presence of a mix of the original karyotype along with the new karyotype enhanced meiotic success. Interestingly, all three viable progeny exhibited similar karyotypes suggesting that this chromosome configuration is probably selected for cell survival while also keeping most of the genome as haploid. Thus, aneuploidy could also shield the newly established chromosome rearrangements and allow it to persist and spread in the population, eventually leading to the fixation of new changes. This hypothesis is supported by our results in which the aneuploid F1 progeny exhibited a much higher spore germination rate with either parent as compared to the crosses between the haploid parents. Additionally, the less aneuploid progeny exhibited less successful meiosis, further suggesting the role of aneuploidy in this process. Based on this, we propose a model in which the rearranged chromosomes can be present

along with the parent chromosomes in an aneuploid intermediate (Figure 7). The aneuploid intermediate then allows the transmission of rearranged chromosomes to the next generation until a particular rearrangement becomes fixed in part of the population. Once fixed, the rearranged population will stably transmit itself and can give rise to a new subpopulation. The presence of geographical barriers, as well as possible advantageous selective roles for the new rearrangement, might further promote fixation of this newly acquired rearrangement.

Acknowledgments

We thank Valerie Lapham at North Carolina State University electron microscopy facility for assistance with scanning electron microscopy. We thank Tom Petes and Sue Jinks-Robertson for critical reading of this manuscript. This work was supported by NIH/NIAID R37 MERIT award AI39115-21 to JH, R01 grant AI50113-15 to JH, and R01 grant AI112595-04 awarded to JH, David Tobin, and Paul Magwene. We also thank the CIFAR program, Fungal Kingdom: Threats & Opportunities for their support.

Materials and Methods

Strains and media

C. neoformans wild-type strains H99α and KN99**a** served as the wild-type isogenic parental lineages for the experiments. Strains were grown in YPD media for all experiments at 30°C unless stated otherwise, and G418 was added at a final concentration of 200 µg/ml for selection of transformants. MS media was used for all the mating assays, which were performed as described previously[26]. Basidia specific spore dissections were performed after two weeks of mating, and the spore germination rate was scored after four days of dissection. All strains used in this study are listed in Table S2.

Phenotyping and growth competition assays

For cell growth and phenotyping assays, cultures were grown overnight, serially diluted (10-fold dilutions), and 3 µl of cell suspension was spotted on YPD or defined stress solid medium followed by incubation at 30°C and 37°C for two days. To make drug-containing media, fluconazole and phleomycin were added into YPD media at a concentration of 24 µg/ml and 1 µg/ml, respectively, and plates were made. The plates were imaged after two days of growth and are presented.

To test the competitive fitness, H99, VYD135, and VYD136 strains were separately mixed with CNVY156[59] at an equal ratio based on OD measurements. CNVY156 strain expresses NAT resistant marker and was used as the counter-selective strain to score competitive growth in co-culture experiment. After the strains were mixed, a part of the mix (1 ml) was immediately spread on YPD plates (0h sample), and cells were allowed to grow for 48 h at 30°C. The remaining mix (5 ml) was grown as co-culture at 30°C for 24 h, and then a portion of cells

from the mix was again spread on YPD plates (24h sample) and incubated for 48 h. From both initial and final samples, 100-150 randomly selected single colonies were replica patched on YPD and YPD+NAT media plates. After 24 h of growth, number of growing colonies from both YPD and YPD+NAT plates was counted. The growth fraction for H99, VYD135 and VYD136 strains was calculated by subtracting number of colonies on YPD+NAT plate from the total number of colonies on YPD plate for each co-culture. The percent growth was calculated and plotted using GraphPad with the total number of colonies on YPD representing 100%, the number of NAT resistant colonies representing NAT^+ population while remaining colonies were considered to be either H99, VYD135, and VYD136.

CRISPR transformations

CRISPR transformation experiments were performed as described previously, with minor modifications[70] (Figure S1). Briefly, *C. neoformans* cells were grown overnight in YPD media and then inoculated into 100 ml of fresh YPD media and grown until OD_{600} reached 0.8 to 1. The cells were then pelleted, washed, and resuspended in EB buffer and incubated for one hour. Next, the cells were pelleted and resuspended again in 50 µl of EB buffer. The cells were mixed with three DNA fragments expressing guide RNA (350 ng), Cas9 (500 ng), and selection marker (2 µg). The cell-DNA mix was subjected to electroporation using the Eppendorf multiporator, with the bacterial mode operating at V=2 kV with τ optimized at 5 ms. Fresh YPD media was immediately added into the transformed cells, and the cells were allowed to recover for two hours before spreading onto selection media (YPD+200 µg/ml G418). The transformants were recovered after 3 to 5 days after transformation. The transformants were further colony purified on YPD+G418 medium to obtain single colony stable transformants. The guide RNA coding

sequences used for Tcn2 and safe-haven locus targeting are TAAGTACTTCTGGGATGGTA and AGTGCTGTGGTGAAAGAGAT, respectively.

PFGE and chromoblot analysis

The PFGE plugs were prepared as described previously with minor modifications[71]. The PFGE was performed with 1% agarose gel at 3.6 V/cm and a switching frequency of 120 to 360 sec for 120 h at 14°C in 0.5X TBE. *S. cerevisiae* CHEF DNA marker (Bio-rad, Cat #1703605) served as markers for estimating the chromosome lengths in all PFGE experiments. To separate shorter chromosomes, the switching frequency was changed to 116 to 276 sec, while all other conditions were kept the same. Following electrophoresis, gels were stained with ethidium bromide (EtBr), and bands were observed by UV transillumination and photographed. Chromoblot analysis was performed as described previously[30]. Briefly, the DNA was hybridized to membranes and probed with chromosome arm regions from targeted chromosomes. The probed membranes were washed, and hybridization signals were observed with a phosphorimager.

Genome assembly and synteny comparison

The *C. neoformans* H99 genome was reassembled with Canu v1.7 based on previously published PacBio and Illumina data to obtain a better resolution of the centromeric regions (see Table S3 for details)[21,22,72]. The resulting draft assembly was improved by correcting errors using five rounds of Pilon (v1.22) polishing ('--fix all' setting) and the Illumina reads mapped to the respective assemblies by the use of BWA-MEM (v0.7.17-r1188)[73,74]. Centromere locations were mapped based on BLAST analysis with centromere flanking genes, and coordinates for these

new locations are provided in Table S1. Because some of the centromere lengths differed in the new assembly as compared to the previous one, we validated the new centromere lengths by mapping the Canu-corrected PacBio read using Minimap2, followed by visual inspection in IGV[75]. Because our work involved analysis of centromere sequences, we utilized this new improved assembly as the reference for all of our analyses. Once the locations were mapped, we oriented all of the chromosomes such that the longer arm (q arm) begins the chromosome, and the smaller arm (p arm) is the distal part of the chromosome.

De novo genome assemblies for the VYD135 and VYD136 isolates, and progeny (P1, P2, and P3) were generated with Canu using Oxford Nanopore reads > 2 kb (see Table S3 and S4 for details) followed by Pilon polishing as described above. When necessary, and after validating by chromoblot analysis, broken contigs were joined artificially with a 50 bp sequence gap to generate a full-length chromosome. Once completed, the chromosomes were numbered based on their length with the longest chromosome as the first. For progeny P1, P2, and P3, extra chromosome configuration was inferred to obtain the final karyotype based on their Illumina and nanopore sequencing reads mapping.

Synteny comparisons between the genomes were performed with SyMAP v4.2 and default parameters[76] (http://www.agcol.arizona.edu/software/symap/). The comparison block maps were exported as .svg files and were then processed using Adobe Illustrator and Adobe Photoshop for representation purposes.

Genomic DNA isolation for nanopore sequencing

The strains with chromosome alterations, as well as progeny from the VYD136α x KN99**a** cross, were subjected to nanopore and Illumina sequencing. For nanopore sequencing,

large molecular weight genomic DNA was prepared using the CTAB based lysis method. For this purpose, 50 ml of an overnight culture was pelleted, frozen at -80°C, and subjected to lyophilization. The lyophilized cell pellet was broken into a fine powder by vortexing with beads for 3 to 5 min at room temperature. 20 ml of CTAB extraction buffer (100 mM Tris-Cl, pH=7.5; 0.7 M NaCl; 10 mM EDTA, pH=8.0; 1% CTAB; 1% β-mercaptoethanol) was added, mixed, and incubated at 65°C for an hour with intermittent shaking after every 20 min. The mix was cooled on ice for 10 min, and the supernatant solution was decanted into a fresh tube. An equal volume of chloroform (~15 ml) was added to the tube and mixed gently for 5 to 10 min.

The mix was centrifuged at 3200 rpm for 10 min, and the supernatant was transferred to a fresh tube. An equal volume of isopropanol (~18 to 20 ml) was added into the supernatant and mixed gently until thread-like structures were visible and formed a clump. The mix was incubated at -20°C for an hour, and centrifuged at 3200 rpm for 10 min to pellet the DNA. The supernatant was discarded, and the pellet was washed with 70% ethanol. The pellet was air-dried and dissolved in 1ml of 1X TE buffer. RNase A was added into the resuspended DNA to a final concentration of 100 µg/ml and incubated at 37°C for 30 to 45 min. Sodium acetate solution was added into the mix to a final concentration of 0.5 M, and the solution was transferred to a 1.5 ml Eppendorf tube/s in the aliquots of 0.5-0.6 ml each. An equal volume of chloroform was added, mixed gently, and centrifuged at 13,000 rpm for 15 min. The supernatant was transferred to a fresh tube, followed by isopropanol precipitation. The DNA threads were removed with a pipette tip and transferred to a fresh tube followed by ethanol washing, air drying, and finally dissolved in 200 µl 1X TE buffer. The DNA quality was estimated with NanoDrop whereas DNA quantity was estimated with Qubit. The size estimation of DNA was carried by electrophoresis of DNA samples on PFGE. For this purpose, the PFGE was carried out at 6V/cm at a switching frequency

of 1 to 6 sec for 16 h at 14°C. Samples with most of the DNA ≥100 kb or larger were selected for nanopore sequencing.

Nanopore and Illumina sequencing

The DNA samples isolated as above were subjected to library preparation and sequencing, as recommended by the manufacturer's instructions. For nanopore sequencing, three or four different DNA samples were barcoded as per the manufacturer's instructions using SQK-LSK109 and EXP-NBD103/EXP-NBD104 kits. The DNA samples were then pooled together on a single R9 flow-cell (FLO-MN106), and sequenced by the MinION system. MinION sequencing and live-base calling were controlled using MinKNOW. DNA sequencing was performed at the default voltage for 48 hours. After sequencing, reads were de-multiplexed with qcat (https://github.com/nanoporetech/qcat) (parameters: --trim -k NBD103/NBD104 --epi2me). Each set of reads was then assembled to obtain genome assembly using Canu as described previously.

Illumina sequencing of the strains was performed at the Duke sequencing facility core (https://genome.duke.edu/), and the data was employed to error correct the genome assemblies for VYD135 and VYD136. The Illumina sequencing data were also mapped to the wild-type H99 genome assembly to estimate the ploidy of strains. Specifically, the Illumina reads were mapped to the H99 genome using Geneious default mapper or bow-tie2 mapper. The resulting BAM file was converted to a .tdf file, which was then visualized through IGV to estimate the ploidy based on read coverage for each chromosome. Sequence data generated in this study were submitted to NCBI with the BioProject accession number PRJNA577944.

To map the breakpoints, nanopore reads from both VYD135 and VYD136 were mapped to the wild-type genome assembly using minimap2. The mapped file (.bam) file was visualized in IGV, and centromere specific snapshots were exported for representation. Comparative analysis of the new centromere sequences with parental sequences was performed by both BLASTn analysis and pairwise sequence alignment. The *CAS9* and *NEO* gene sequences were identified in the new centromeres based on BLAST analysis with sequences of original plasmids.

Galleria mellonella killing assay

G. mellonella infection experiments were performed as described previously with some modifications[27]. *G. mellonella* caterpillars in the final instar larval stage were used to test the pathogenicity of *C. neoformans* strains, and healthy caterpillars were employed in all assays. 20 to 22 chosen caterpillars were infected with each strain. Four μl cell suspension (10^6 cells/ml) of a strain was injected into each caterpillar via the last left proleg. After injection, caterpillars were incubated in plastic containers at 37°C, and the number of dead caterpillars was scored daily. Caterpillars were considered dead when they exhibited black coloration of the body and displayed no movement in response to touch. The few caterpillars that became pupae during the experimental duration were removed from the analysis.

References

1 Hunter, N. Meiotic recombination: The essence of heredity. *Cold Spring Harb Perspect Biol* **7**, (2015).

2 Andersen, S. L. & Sekelsky, J. Meiotic versus mitotic recombination: Two different routes for double-strand break repair. *Bioessays* **32**, 1058-1066, (2010).

3 LaFave, M. C. & Sekelsky, J. Mitotic recombination: why? when? how? where? *PLoS Genet* **5**, e1000411, (2009).

4 Harewood, L. & Fraser, P. The impact of chromosomal rearrangements on regulation of gene expression. *Hum Mol Genet* **23**, R76-82, (2014).

5 Hasty, P. & Montagna, C. Chromosomal rearrangements in cancer: Detection and potential causal mechanisms. *Mol Cell Oncol* **1**, e29904, (2014).

6 Piazza, A. & Heyer, W. D. Homologous recombination and the formation of complex genomic rearrangements. *Trends Cell Biol* **29**, 135-149, (2019).

7 Lysák, M. A. & Schubert, I. in *Plant Genome Diversity Volume 2: Physical Structure, Behaviour and Evolution of Plant Genomes* (eds Johann Greilhuber, Jaroslav Dolezel, & Jonathan F. Wendel) 137-147 (Springer Vienna, 2013).

8 Mehta, A. & Haber, J. E. Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harb Perspect Biol* **6**, a016428, (2014).

9 Chang, H. H. Y., Pannunzio, N. R., Adachi, N. & Lieber, M. R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol* **18**, 495-506, (2017).

10 Lieber, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* **79**, 181-211, (2010).

11 Koltsova, A. S. *et al.* On the complexity of mechanisms and consequences of chromothripsis: An update. *Front Genet* **10**, 393, (2019).

12 Arya, R. & Bassing, C. H. V(D)J recombination exploits DNA damage responses to promote immunity. *Trends Genet* **33**, 479-489, (2017).

13 Ly, P. & Cleveland, D. W. Rebuilding chromosomes after catastrophe: Emerging mechanisms of chromothripsis. *Trends Cell Biol* **27**, 917-930, (2017).

14 Gudjonsson, T. *et al.* TRIP12 and UBR5 suppress spreading of chromatin ubiquitylation at damaged chromosomes. *Cell* **150**, 697-709, (2012).

15 Pellestor, F. Chromoanagenesis: cataclysms behind complex chromosomal rearrangements. *Mol Cytogenet* **12**, 6, (2019).

16 Rieseberg, L. H. Chromosomal rearrangements and speciation. *Trends Ecol Evol* **16**, 351-358, (2001).

17 Pellestor, F. *et al.* Complex chromosomal rearrangements: origin and meiotic behavior. *Hum Reprod Update* **17**, 476-494, (2011).

18 Lin, X. & Heitman, J. The biology of the *Cryptococcus neoformans* species complex. *Annu Rev Microbiol* **60**, 69-105, (2006).

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 Sun, S., Coelho, M. A., David-Palma, M., Priest, S. & Heitman, J. The evolution of sexual reproduction and the mating-type locus: Links to pathogenesis of *Cryptococcus* human pathogenic fungi. *Annu Rev Genet* **53**, (2019).

21 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).

22 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).

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

24 Kozubowski, L. & Heitman, J. Profiling a killer, the development of *Cryptococcus neoformans*. *FEMS Microbiol Rev* **36**, 78-94, (2012).

25 Lengeler, K. B. *et al.* Mating-type locus of *Cryptococcus neoformans*: a step in the evolution of sex chromosomes. *Eukaryot Cell* **1**, 704-718, (2002).

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

27 Mylonakis, E. *et al. Galleria mellonella* as a model system to study *Cryptococcus neoformans* pathogenesis. *Infect Immun* **73**, 3842-3850, (2005).

28 Botts, M. R., Giles, S. S., Gates, M. A., Kozel, T. R. & Hull, C. M. Isolation and characterization of *Cryptococcus neoformans* spores reveal a critical role for capsule biosynthesis genes in spore biogenesis. *Eukaryot Cell* **8**, 595-605, (2009).

29 Velagapudi, R., Hsueh, Y. P., Geunes-Boyer, S., Wright, J. R. & Heitman, J. Spores as infectious propagules of *Cryptococcus neoformans*. *Infect Immun* **77**, 4345-4355, (2009).

30 Sun, S. *et al.* Fungal genome and mating system transitions facilitated by chromosomal translocations involving intercentromeric recombination. *PLoS Biol* **15**, e2002527, (2017).

31 Jaco, I., Canela, A., Vera, E. & Blasco, M. A. Centromere mitotic recombination in mammalian cells. *J Cell Biol* **181**, 885-892, (2008).

32 Talbert, P. B. & Henikoff, S. Centromeres convert but don't cross. *PLoS Biol* **8**, e1000326, (2010).

33 Thakur, J. & Sanyal, K. Efficient neocentromere formation is suppressed by gene conversion to maintain centromere function at native physical chromosomal loci in *Candida albicans*. *Genome Res* **23**, 638-652, (2013).

34 Zafar, F. *et al.* Regulation of mitotic recombination between DNA repeats in centromeres. *Nucleic Acids Res* **45**, 11222-11235, (2017).

35 Liebman, S. W., Symington, L. S. & Petes, T. D. Mitotic recombination within the centromere of a yeast chromosome. *Science* **241**, 1074-1077, (1988).

36 Symington, L. S. & Petes, T. D. Meiotic recombination within the centromere of a yeast chromosome. *Cell* **52**, 237-240, (1988).

37 Dawe, R. K. & Henikoff, S. Centromeres put epigenetics in the driver's seat. *Trends Biochem Sci* **31**, 662-669, (2006).

38 Kursel, L. E. & Malik, H. S. The cellular mechanisms and consequences of centromere drive. *Curr Opin Cell Biol* **52**, 58-65, (2018).

39 Rosin, L. F. & Mellone, B. G. Centromeres drive a hard bargain. *Trends Genet* **33**, 101-117, (2017).

40 Kurzhals, R. L. *et al.* Chromosome healing Is promoted by the telomere cap component Hiphop in *Drosophila*. *Genetics* **207**, 949-959, (2017).

41 Melek, M. & Shippen, D. E. Chromosome healing: spontaneous and programmed *de novo* telomere formation by telomerase. *Bioessays* **18**, 301-308, (1996).

42 Pennaneach, V., Putnam, C. D. & Kolodner, R. D. Chromosome healing by *de novo* telomere addition in *Saccharomyces cerevisiae*. *Mol Microbiol* **59**, 1357-1368, (2006).

43 Gao, Q. *et al.* Telomerase-dependent and -independent chromosome healing in mouse embryonic stem cells. *DNA Repair (Amst)* **7**, 1233-1249, (2008).

44 Kulkarni, A., Zschenker, O., Reynolds, G., Miller, D. & Murnane, J. P. Effect of telomere proximity on telomere position effect, chromosome healing, and sensitivity to DNA double-strand breaks in a human tumor cell line. *Mol Cell Biol* **30**, 578-589, (2010).

45 Varley, H., Di, S., Scherer, S. W. & Royle, N. J. Characterization of terminal deletions at 7q32 and 22q13.3 healed by *de novo* telomere addition. *Am J Hum Genet* **67**, 610-622, (2000).

46 Maciejowski, J. & de Lange, T. Telomeres in cancer: tumour suppression and genome instability. *Nat Rev Mol Cell Biol* **18**, 175-186, (2017).

47 Murnane, J. P. Telomere loss as a mechanism for chromosome instability in human cancer. *Cancer Res* **70**, 4255-4259, (2010).

48 Hustedt, N. & Durocher, D. The control of DNA repair by the cell cycle. *Nat Cell Biol* **19**, 1-9, (2016).

49 Murray, J. M. & Carr, A. M. Integrating DNA damage repair with the cell cycle. *Curr Opin Cell Biol* **52**, 120-125, (2018).

50 Roukos, V. & Misteli, T. The biogenesis of chromosome translocations. *Nat Cell Biol* **16**, 293-300, (2014).

51 Sakofsky, C. J. & Malkova, A. Break induced replication in eukaryotes: mechanisms, functions, and consequences. *Crit Rev Biochem Mol Biol* **52**, 395-413, (2017).

52 Piazza, A. & Heyer, W. D. Multi-invasion-induced rearrangements as a pathway for physiological and pathological recombination. *Bioessays* **40**, e1700249, (2018).

53 Piazza, A., Wright, W. D. & Heyer, W. D. Multi-invasions are recombination byproducts that induce chromosomal rearrangements. *Cell* **170**, 760-773 e715, (2017).

54 Argueso, J. L. *et al.* Double-strand breaks associated with repetitive DNA can reshape the genome. *Proc Natl Acad Sci U S A* **105**, 11845-11850, (2008).

55 Fleiss, A. *et al.* Reshuffling yeast chromosomes with CRISPR/Cas9. *PLoS Genet* **15**, e1008332, (2019).

56 Muramoto, N. *et al.* Phenotypic diversification by enhanced genome restructuring after induction of multiple DNA double-strand breaks. *Nat Commun* **9**, 1995, (2018).

57 Pellestor, F. & Gatinois, V. Chromothripsis, a credible chromosomal mechanism in evolutionary process. *Chromosoma* **128**, 1-6, (2019).

58 Kozubowski, L. *et al*. Ordered kinetochore assembly in the human-pathogenic basidiomycetous yeast *Cryptococcus neoformans*. *mBio* **4**, e00614-13, (2013).

59 Yadav, V. & Sanyal, K. Sad1 spatiotemporally regulates kinetochore clustering to ensure high-fidelity chromosome segregation in the human fungal pathogen *Cryptococcus neoformans*. *mSphere* **3**, e00190-18, (2018).

60 Hancks, D. C. A role for retrotransposons in chromothripsis. *Methods Mol Biol* **1769**, 169-181, (2018).

61 Fischer, G., James, S. A., Roberts, I. N., Oliver, S. G. & Louis, E. J. Chromosomal evolution in *Saccharomyces*. *Nature* **405**, 451-454, (2000).

62 Livingstone, K. & Rieseberg, L. Chromosomal evolution and speciation: a recombination-based approach. *New Phytol* **161**, 107-112, (2004).

63 Noor, M. A., Grams, K. L., Bertucci, L. A. & Reiland, J. Chromosomal inversions and the reproductive isolation of species. *Proc Natl Acad Sci U S A* **98**, 12084-12088, (2001).

64 Brown, J. D. & O'Neill, R. J. Chromosomes, conflict, and epigenetics: chromosomal speciation revisited. *Annu Rev Genomics Hum Genet* **11**, 291-316, (2010).

65 Deakin, J. E. Chromosome evolution in marsupials. *Genes (Basel)* **9**, (2018).

66 Metcalfe, C. J. *et al*. Genomic instability within centromeres of interspecific marsupial hybrids. *Genetics* **177**, 2507-2517, (2007).

67 O'Neill, R. J., Eldridge, M. D. & Metcalfe, C. J. Centromere dynamics and chromosome evolution in marsupials. *J Hered* **95**, 375-381, (2004).

68 Potter, S. *et al.* Chromosomal speciation in the genomics era: Disentangling phylogenetic evolution of rock-wallabies. *Front Genet* **8**, 10, (2017).

69 Bulazel, K. V., Ferreri, G. C., Eldridge, M. D. & O'Neill, R. J. Species-specific shifts in centromere sequence composition are coincident with breakpoint reuse in karyotypically divergent lineages. *Genome Biol* **8**, R170, (2007).

70 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).

71 Findley, K. *et al.* Discovery of a modified tetrapolar sexual cycle in *Cryptococcus amylolentus* and the evolution of *MAT* in the *Cryptococcus* species complex. *PLoS Genet* **8**, e1002528, (2012).

72 Koren, S. *et al.* Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* **27**, 722-736, (2017).

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

74 Walker, B. J. *et al.* Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* **9**, e112963, (2014).

75 Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100, (2018).

76 Soderlund, C., Bomhoff, M. & Nelson, W. M. SyMAP v3.4: a turnkey synteny system with application to plant genomes. *Nucleic Acids Res* **39**, e68, (2011).

Figure legends

Figure 1. Centromere-specific DSBs mediated by CRISPR lead to chromosome rearrangements. (A) Centromere maps showing the distribution of retrotransposons (Tcn1-Tcn6) in the centromeres of wild-type strain H99 of *C. neoformans*. **(B)** An outline depicting the model for achieving multiple chromosome translocations in *C. neoformans*. **(C)** PFGE revealed many differences in the karyotype of VYD135 and VYD136 as compared with wild-type, H99. **(D)** Chromosome maps for VYD135 and VYD136 compared to the H99 genome revealed multiple chromosome translocations in these strains. Chromosomes are colored with H99 chromosomes as reference. 'q' represents the longer arm while 'p' represents the shorter arm according to the wild-type chromosome configuration.

Figure 2. Sequence analysis of new centromeres and telomeres. (A) PFGE EtBr staining showing bands for newly generated short chromosomes and new telocentric chromosomes. M represents *S. cerevisiae* chromosomes. **(B)** Chromosome map of one of the newly formed telomeres showing the presence of telomere sequence repeats next to Tcn2 elements present in the centromere. **(C-D)** Maps showing the distribution of retrotransposons, along with the integration of foreign sequences from CRISPR/Cas9 and the neomycin resistance gene (NEO^R), in the centromeres of VYD135 (C) and VYD136 (D). Numbers in brackets next to *CEN* numbers represent the wild-type *CEN* numbers that rearranged to form new centromeres.

Figure 3. Chromosome rearrangements are mediated via the gRNA cleavage site. (A-B) Simplified outline maps depicting the chromosomes that underwent translocation in VYD135 (A) and VYD136 (B). Black semicircles, telomeres; red semicircles, de novo telomeres, narrow

bands, centromeres; shaded box on chromosome 5, *MAT* locus. **(C)** Nanopore reads mapping to the wild-type H99 genome revealed the sites of DSB formation or repair junctions at centromeres. The reads either converge on a single site (*CEN5*, *CEN8*) or exhibit sequence gaps between sites (*CEN10*) marking the location of junctions. Red bars indicate the centromeres whereas the black vertical lines mark the site of gRNA cleavage. Cov, Coverage of nanopore reads; Map, Mapping of nanopore reads.

Figure 4. Synteny analysis of rearranged centromeres with wild-type H99 centromeres reveals complex rearrangements. A pair-wise comparison of newly generated and wild-type centromeres revealed that translocations are mediated by double-strand breaks (DSBs) generated via CRISPR. Centromere specific events are described in the individual panels. Grey shades represent direct synteny, while the cyan shade represents inversion events. In the cases that are shown in detail, the cross represents the evidence for homologous recombination, whereas the connecting grey lines represent joining events marking non-homologous end-joining. VYD135-*CEN7* was generated after artificial fusion of two contigs and hence was not analyzed in detail.

Figure 5. Chromosome shuffled isolates exhibit defects in sexual reproduction. (A) Light microscopy images showing hyphae, basidia, and spore chains in crosses between *MAT***a** wild-type strain KN99**a** and *MAT*α wild-type (H99) and rearranged strains (VYD135 and VYD136). Scale bar, 100 µm. **(B)** Scanning Electron Microscope (SEM) images depicting a complete or partial sporulation defect when wild-type KN99**a** was crossed with VYD135 and VYD136, respectively. Scale bar, 10 µm.

Figure 6. VYD136 progeny are aneuploid and exhibit mixed karyotypes. (A) Analysis of the mating-type locus and the mating efficiency of three progeny of VYD136 with either parent. The numbers in brackets represent spore germination rates of respective crosses. **(B)** Flow cytometry profiles of wild-type haploid, diploid, and three progeny of VYD136. **(C)** Nanopore sequencing data mapped to the wild-type H99 genome revealed aneuploidy for multiple chromosomes in the three VYD136 progeny. **(D)** Karyotypes for three progeny showing synteny as compared to the wild-type H99 genome. The red stars represent breaks that were fused later based on synteny and ploidy. The chromosomes shown with red bar on top were not assembled *de novo* but represent possible chromosomes configuration based on Illumina and nanopore sequencing analysis. Contigs 14 in VYD136.P3 could not be resolved into their chromosome configuration. 'q' represents longer arm while 'p' represents shorter arm according to the wild-type chromosome configuration.

Figure 7. A model proposing the evolution of reproductive isolation induced by centromere breaks. DSBs generated using CRISPR at centromeres (step 2) reshapes the karyotype following complex repair events. These complex events include the loss of centromere DNA, isochromosome formation, and *de novo* telomere formation (step 3), similar to what is observed during the process of chromothripsis. The new karyotype can generate a reproductive barrier with the parental isolate and lead to speciation. On the other hand, the strain with the rearranged karyotype could mate with wild-type isolate, albeit at low frequency, leading to aneuploid progeny (step 4), which can independently establish itself as a new species.

Figure 1

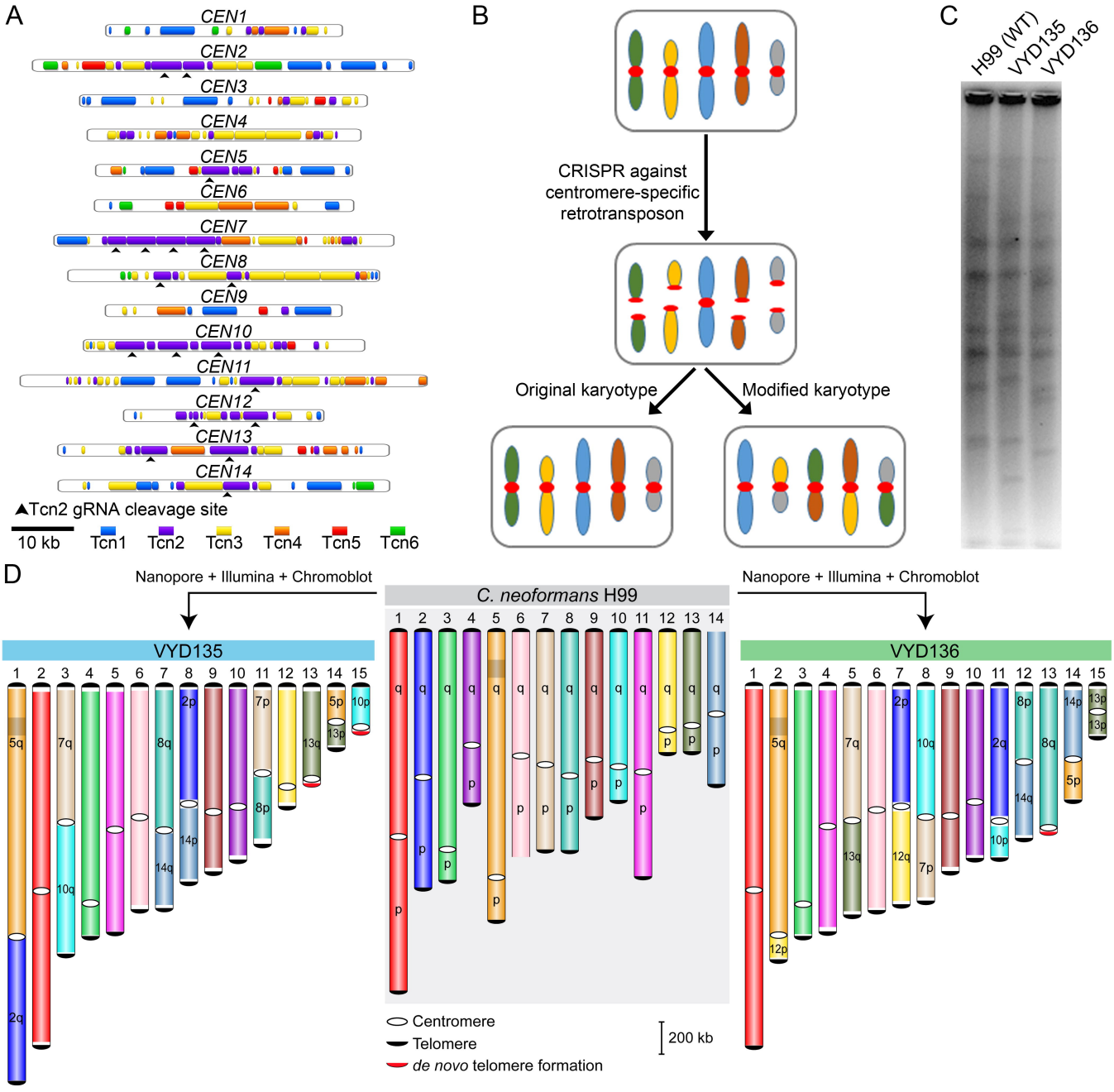
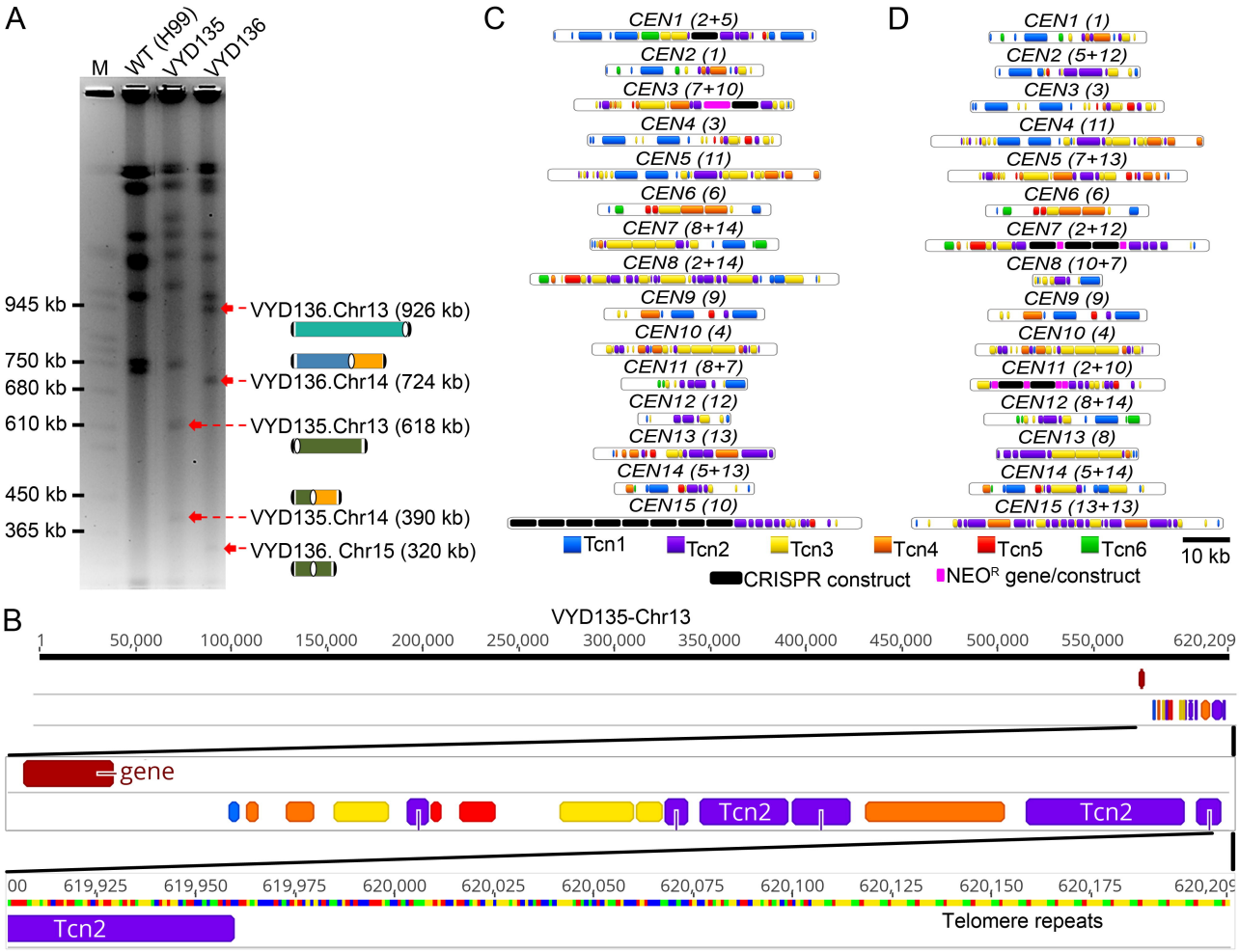

A
CEN1
CEN2
CEN3
CEN4
CEN5
CEN6
CEN7
CEN8
CEN9
CEN10
CEN11
CEN12
CEN13
CEN14
▲Tcn2 gRNA cleavage site
10 kb
Tcn1
Tcn2
Tcn3
Tcn4
Tcn5
Tcn6
B
CRISPR against centromere-specific retrotransposon
Original karyotype
Modified karyotype
C
H99 (WT)
VYD135
VYD136
D
Nanopore + Illumina + Chromoblot
C. neoformans H99
Nanopore + Illumina + Chromoblot
VYD135
VYD136
Centromere
Telomere
de novo telomere formation
200 kb

Figure 2


A
M
WT (H99)
VYD135
VYD136
945 kb
750 kb
680 kb
610 kb
450 kb
365 kb
VYD136.Chr13 (926 kb)
VYD136.Chr14 (724 kb)
VYD135.Chr13 (618 kb)
VYD135.Chr14 (390 kb)
VYD136. Chr15 (320 kb)
C
CEN1 (2+5)
CEN2 (1)
CEN3 (7+10)
CEN4 (3)
CEN5 (11)
CEN6 (6)
CEN7 (8+14)
CEN8 (2+14)
CEN9 (9)
CEN10 (4)
CEN11 (8+7)
CEN12 (12)
CEN13 (13)
CEN14 (5+13)
CEN15 (10)
D
CEN1 (1)
CEN2 (5+12)
CEN3 (3)
CEN4 (11)
CEN5 (7+13)
CEN6 (6)
CEN7 (2+12)
CEN8 (10+7)
CEN9 (9)
CEN10 (4)
CEN11 (2+10)
CEN12 (8+14)
CEN13 (8)
CEN14 (5+14)
CEN15 (13+13)
Tcn1
Tcn2
Tcn3
Tcn4
Tcn5
Tcn6
10 kb
CRISPR construct
NEO^R gene/construct
B
VYD135-Chr13
1
50,000
100,000
150,000
200,000
250,000
300,000
350,000
400,000
450,000
500,000
550,000
620,209
gene
Tcn2
Tcn2
619,925
619,950
619,975
620,000
620,025
620,050
620,075
620,100
620,125
620,150
620,175
620,209
Tcn2
Telomere repeats

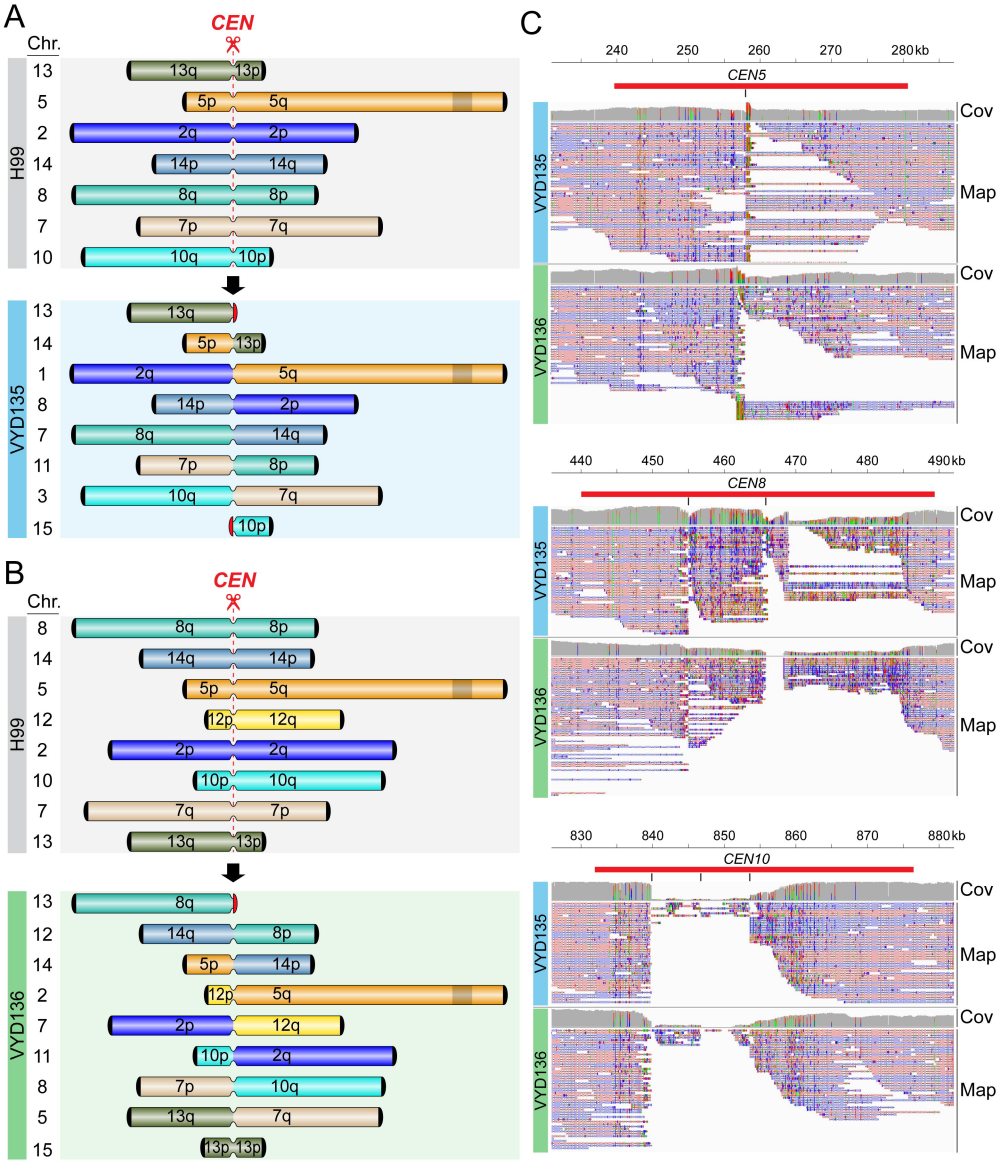

A
CEN
Chr.
H99
13
13q
13p
5
5p
5q
2
2q
2p
14
14p
14q
8
8q
8p
7
7p
7q
10
10q
10p
VYD135
13
13q
14
5p
13p
1
2q
5q
8
14p
2p
7
8q
14q
11
7p
8p
3
10q
7q
15
10p
B
CEN
Chr.
H99
8
8q
8p
14
14q
14p
5
5p
5q
12
12p
12q
2
2p
2q
10
10p
10q
7
7q
7p
13
13q
13p
VYD136
13
8q
12
14q
8p
14
5p
14p
2
12p
5q
7
2p
12q
11
10p
2q
8
7p
10q
5
13q
7q
15
13p
13p
C
240
250
260
270
280kb
CEN5
VYD135
VYD136
Cov
Map
440
450
460
470
480
490kb
CEN8
830
840
850
860
870
880kb
CEN10

Figure 4

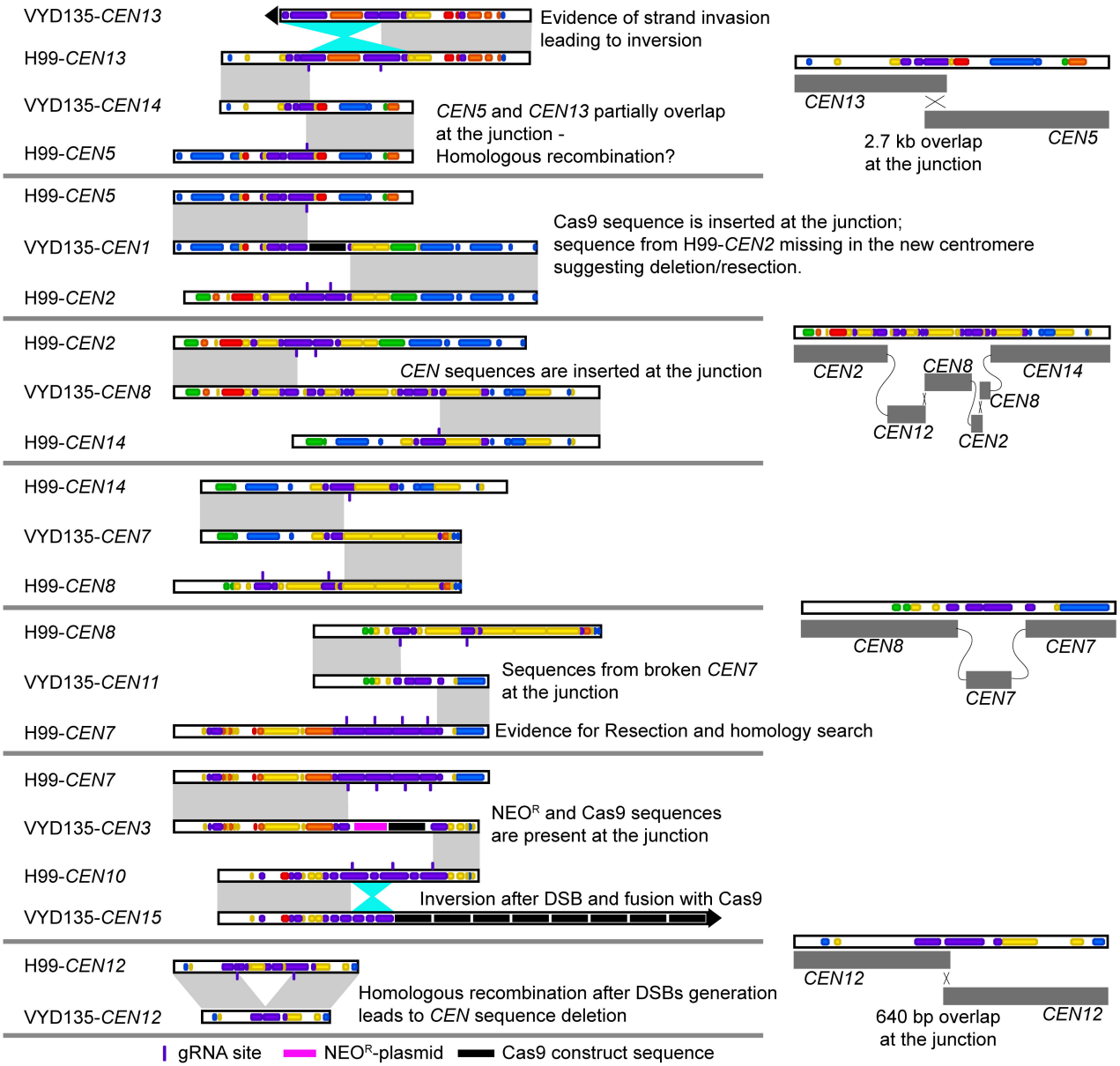

VYD135-CEN13
H99-CEN13
Evidence of strand invasion leading to inversion
VYD135-CEN14
H99-CEN5
CEN5 and CEN13 partially overlap at the junction - Homologous recombination?
CEN13
CEN5
2.7 kb overlap at the junction
H99-CEN5
VYD135-CEN1
H99-CEN2
Cas9 sequence is inserted at the junction; sequence from H99-CEN2 missing in the new centromere suggesting deletion/resection.
H99-CEN2
VYD135-CEN8
H99-CEN14
CEN sequences are inserted at the junction
CEN2
CEN8
CEN14
CEN8
CEN12
CEN2
H99-CEN14
VYD135-CEN7
H99-CEN8
H99-CEN8
VYD135-CEN11
H99-CEN7
Sequences from broken CEN7 at the junction
Evidence for Resection and homology search
CEN8
CEN7
CEN7
H99-CEN7
VYD135-CEN3
H99-CEN10
VYD135-CEN15
NEO^R and Cas9 sequences are present at the junction
Inversion after DSB and fusion with Cas9
H99-CEN12
VYD135-CEN12
Homologous recombination after DSBs generation leads to CEN sequence deletion
CEN12
CEN12
640 bp overlap at the junction
gRNA site
NEO^R-plasmid
Cas9 construct sequence

Figure 5

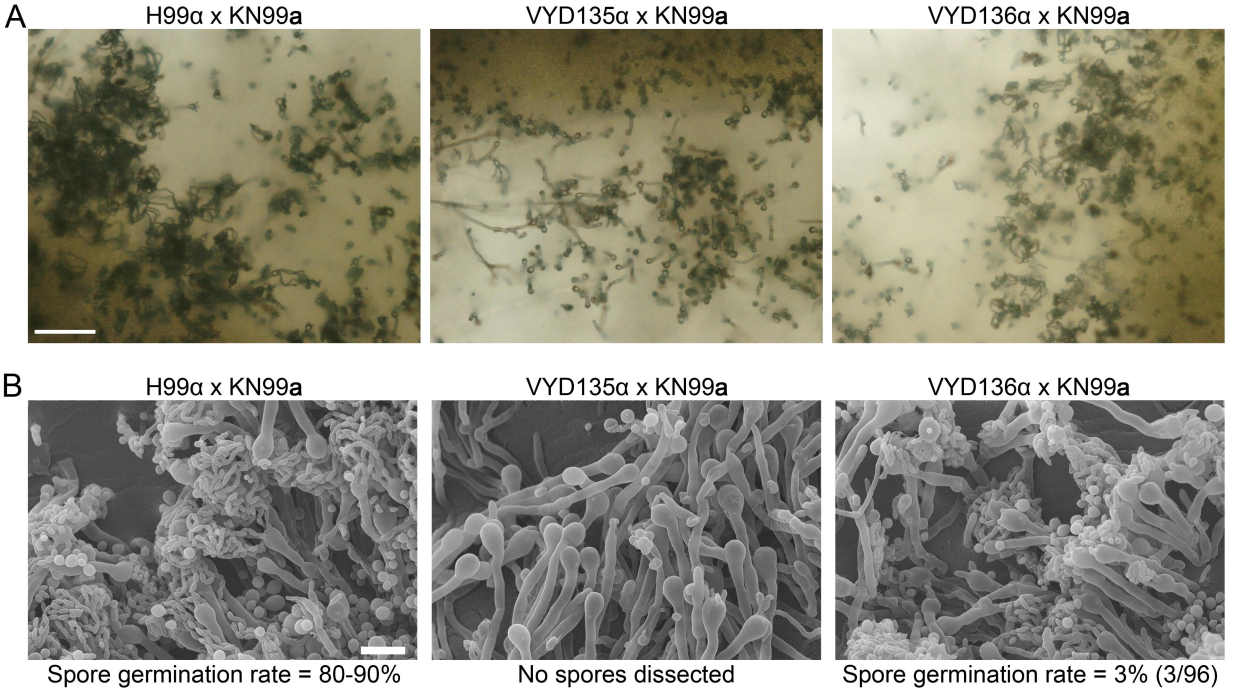
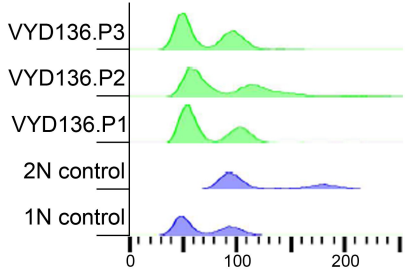

A
H99α x KN99a
VYD135α x KN99a
VYD136α x KN99a
B
H99α x KN99a
VYD135α x KN99a
VYD136α x KN99a
Spore germination rate = 80-90%
No spores dissected
Spore germination rate = 3% (3/96)

Figure 6

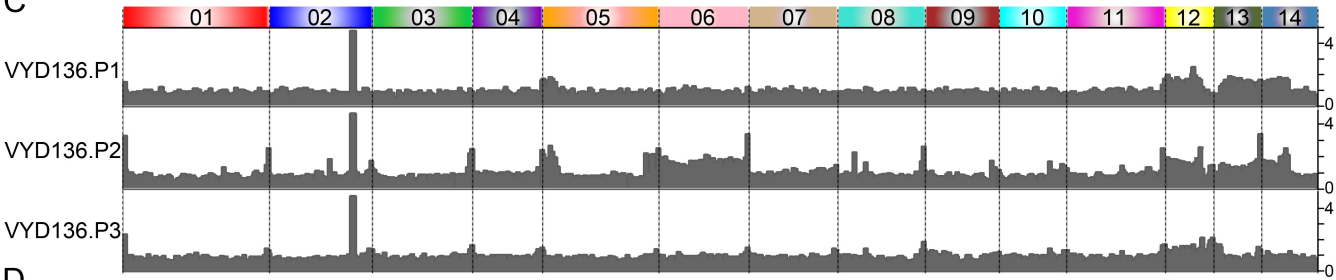
A

VYD136 progeny #	*MAT* locus (PCR)	Ploidy (FACS)	Crossed with (% spore germination rate) VYD136 α	H99 α	KN99 a	Self
VYD136.P1	α	1N+	✗	✗	✓ (54)	✗
VYD136.P2	a	1N+	✓ (58)	✓ (51)	✗	✗
VYD136.P3	a	1N	✓ (21)	✓ (11)	✗	✗

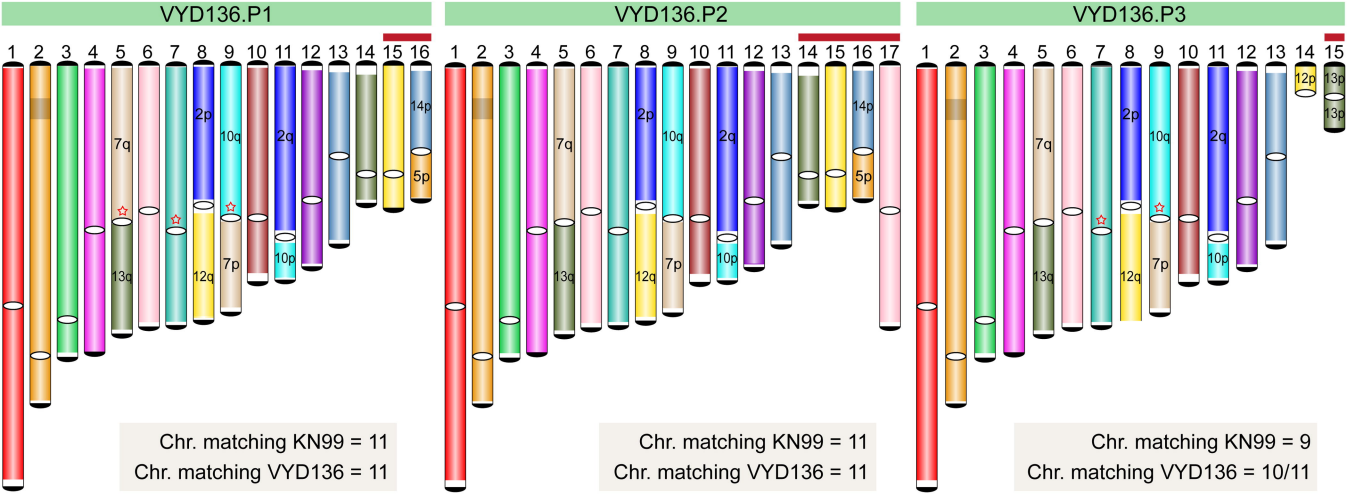
B



C



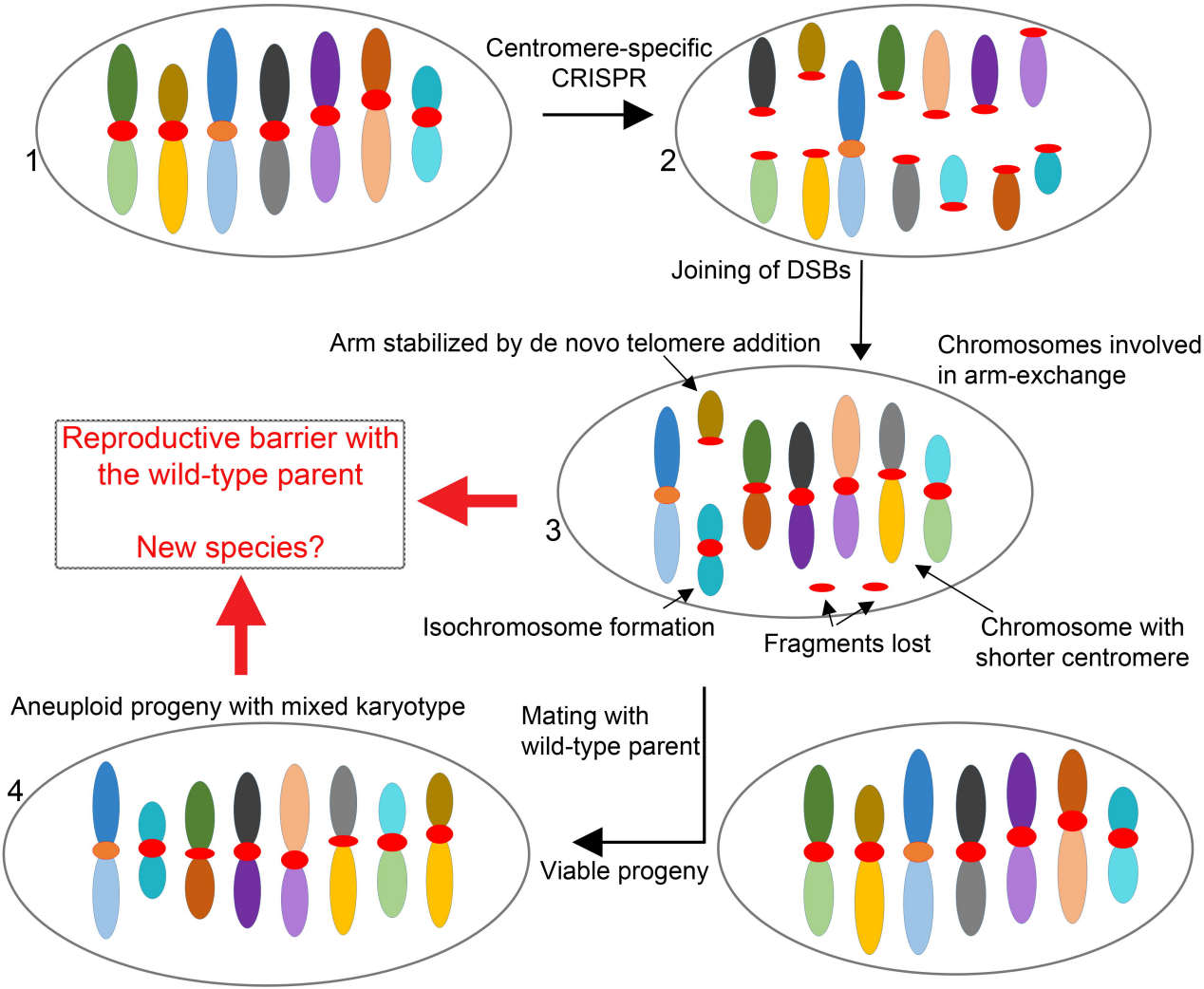
D



H99 Chromosomes:

Chr1 Chr2 Chr3 Chr4 Chr5 Chr6 Chr7 Chr8 Chr9 Chr10 Chr11 Chr12 Chr13 Chr14

Figure 7


1
Centromere-specific CRISPR
2
Joining of DSBs
Arm stabilized by de novo telomere addition
Chromosomes involved in arm-exchange
3
Reproductive barrier with the wild-type parent
New species?
Isochromosome formation
Fragments lost
Chromosome with shorter centromere
Aneuploid progeny with mixed karyotype
Mating with wild-type parent
4
Viable progeny