Natural mismatch repair mutations mediate phenotypic diversity and drug resistance in *Cryptococcus deuterogattii*

R. Blake Billmyre, Shelly Applen Clancey, and Joseph Heitman*

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

*Corresponding author:
Room 322 CARL Building
Box 3546 Research Drive
Department of Molecular Genetics and Microbiology
Duke University Medical Center, Durham, NC 27710, USA
Email: heitm001@duke.edu
Phone: (919) 684-2824
Fax: (919) 684-2790

Abstract

Pathogenic microbes confront a constant evolutionary conflict between the pressure to maintain genome stability and the need to adapt to mounting external stresses. Prokaryotes often respond with elevated mutation rates, but to date little evidence exists of stable eukaryotic hypermutators in nature. Whole genome resequencing of the human fungal pathogen *Cryptococcus deuterogattii* identified an outbreak lineage characterized by a nonsense mutation in *MSH2*. This defect in mismatch repair results in a moderate mutation rate increase in typical genes, and a larger increase in genes containing homopolymer runs. This allows facile inactivation of genes with coding homopolymer runs including *FRR1*, which encodes the target of the immunosuppresive antifungal drugs FK506 and rapamycin. Our study identifies a eukaryotic hypermutator lineage spread over two continents and suggests that pathogenic eukaryotic microbes may experience similar selection pressures on mutation rate as bacterial pathogens, particularly during long periods of clonal growth or while expanding into new environments.

Introduction

Mutation is the raw material of evolution. As a result, all organisms must strike a balance between allowing enough random mutations for selection to act upon, and the fact that most of these mutations are likely to be deleterious and must be purged from the population. This is particularly true for pathogenic microbes that take part in Red Queen conflicts with their hosts and require a continuous supply of mutations to maintain competitive fitness. In this case, increased pressure may exist to maintain an elevated mutation rate to increase the rate of adaptation. Increases in mutation rate may also serve to accelerate adaptation when microbes are introduced into new environments or encounter novel stresses, such as antimicrobial therapy.

Adaptive variation in mutation rate is common in bacteria. One example is the Long Term Evolution Experiment (LTEE), where defects in DNA repair emerged and the resulting hypermutator phenotype swept the population in six out of twelve *E. coli* lines (1). Mutator alleles likely emerge frequently but are typically purged from the population because individual mutations are more likely to be deleterious than adaptive; as a result hypermutators generally produce less fit offspring than non-hypermutators. However, occasionally a sufficiently beneficial mutation is potentiated by the presence of the hypermutator and the hypermutator allele is able to hitchhike to higher frequency within the population. This is more likely to occur in populations at local evolutionary minima, where many different large-effect adaptive mutations are possible. For example, hypermutator phenotypes are common in isolates of *Pseudomonas aeruginosa* growing within the lungs of patients with Cystic Fibrosis, where environmental conditions are constantly changing (2). This changing environment causes an ongoing need to adapt, and increases the likelihood of hypermutators emerging by increasing the target size for adaptive variation. However, once beneficial mutations have fixed, and organisms have adapted to their new environment, antimutator suppressor alleles can emerge to reduce the

mutation rate or defective mutator alleles can even be replaced by functional alleles through horizontal gene transfer (3, 4). This balance is likely critical to evolution in microbes as they face diverse stresses and environmental changes, and as a result over 1% of natural bacterial isolates display hypermutator phenotypes (5). Even more extreme cases also exist, like in *Mycoplasma pneumoniae* and *M. genitalium*, where the mismatch repair machinery has been completely lost (6).

In fact, many mutator phenotypes are the result of defects in mismatch repair, commonly referred to as MMR. In bacteria, MMR requires the MutHLS system, where MutS binds to mismatches and recruits MutL, which subsequently activates MutH to excise the mispair (7–10) . In eukaryotes, both the MutS and MutL families have expanded, yielding MutS homolog (MSH) (11) and MutL homolog (MLH) families (12). Heterodimers of various MSH family members participate in mismatch repair, but Msh2 in particular is a core component of the majority of mismatch repair pathways in yeast (13). As a result, defects in MMR, and *MSH2* mutations in particular, result in elevated rates of simple mismatches, but also dramatically elevated rates of repeat tract instability (12), which are associated with hereditary colon cancer in humans. MMR also plays a role in rejecting heteroduplexes of divergent sequences during both mitosis (14) and meiosis (15), meaning that loss of the complex can lower species boundaries and allow increased recombination between divergent chromosomes (15, 16).

In contrast to bacteria, substantially less is known about hypermutators in natural populations of eukaryotes. Evolutionary theory predicts that the presence of sex abrogates selection for hypermutator strains by eliminating genetic linkage between the mutator allele and the associated beneficial mutations it hitchhikes upon to high frequency (17). Studies of allele incompatibility within the model fungus *Saccharomyces cerevisiae* have supported this traditional paradigm. Incompatible alleles of *PMS1* and *MLH1* exist within the population, such that a cross of two strains with wildtype mutation frequency can give rise to progeny with elevated mutation rates (18, 19). However, the homozygous

incompatible arrangement was not found among any of the original 65 strains examined, and was only identified within one clinical isolate of 1,010 global diverse yeast isolates (20) and four clinical strains out of 93 sequenced in a separate study (21). This frequency is far below that of the individual alleles in the population, and in each case observed suppressors had arisen to restore a wildtype mutation rate despite the incompatibility. In contrast, eukaryotic mutators can adapt and thrive in clonally expanding populations such as in human tumor environments (22). Recent studies of *Candida glabrata* have begun to challenge this dogma in eukaryotic microorganisms as well, with over 50% of clinical *C. glabrata* isolates harboring loss of function mutations in *MSH2*, a critical component of mismatch repair (23). In previous work we also identified a clade of candidate hypermutators within the Pacific Northwest *Cryptococcus deuterogattii* outbreak that contain a coding single base deletion within the critical mismatch repair component *MSH2* (24).

Cryptococcus deuterogattii, previously known as *Cryptococcus gattii* VGII (25), is a basidiomycete human fungal pathogen. Unlike the other species of the *Cryptococcus* pathogenic species complex that infect immunocompromised hosts, *C. deuterogattii* is characterized by its ability to cause infection in otherwise healthy hosts (26) and by loss of the RNAi pathway (27, 28). *C. deuterogattii* is responsible for an ongoing outbreak in the Pacific Northwest region of the United States and Canada, which began in the late 1990's (29). This outbreak was originally analyzed using Multi-locus sequence typing (MLST) and shown to be comprised of three clonal expansions, denoted VGIIa, VGIIb, and VGIIc (30, 31). These three subtypes differ in terms of virulence and total number of cases, with VGIIa, responsible for the majority of infections. More recently, whole genome sequencing studies of these subpopulations showed that the clonal subtypes appeared to have different proximal geographic origins, and most interestingly, that VGIIa has three very closely related isolates, here referred to as VGIIa-like, that were considerably less virulent and differed in isolation location or

date from the outbreak (24). We identified a candidate potential large effect mutation shared by these three diminished virulence isolates that resulted in a predicted non-functional Msh2. Here we show that these VGIIa-like isolates exhibit a *bona fide* hypermutator phenotype. Furthermore, we show that homopolymer runs are particularly unstable, and that these runs are common within the coding regions of genes in the *Cryptococcus* genome. We predict that this results in dramatic phenotypic diversity from inactivation and possibly also activation or reactivation of genes. Finally, we show that the mutant *msh2* allele is not directly responsible for the decrease in virulence of the VGIIa-like isolates, nor does it appear to have directly played a role in the evolution of virulence in the VGIIa clonal outbreak. Rather, it appears to represent a parallel route of adaptation to a new environment in a successful lineage isolated from two continents over two decades and from both the environment and patient samples. This work suggests that hypermutator phenotypes are not limited to prokaryotes, but may represent a frequent avenue to evolutionary change and phenotypic diversity in eukaryotic microbes as well.

Results

We previously identified a sublineage of strains highly related to the VGIIa outbreak in the Pacific Northwest denoted VGIIa-like, in part because they differed either in isolation location or time from the typical outbreak isolates (24). The VGIIa-like sublineage is composed of three isolates: NIH444, a clinical isolate from Seattle in 1975, well before the start of the outbreak; ICB107, a clinical isolate from Brazil in 1981; and CBS7750, an environmental isolate from California in 1990. All three isolates are characterized by diminished virulence relative to the VGIIa clonal outbreak strains (30–32). The analysis here includes additional genomes for a South American outgroup to both the VGIIa-like and VGIIa groups (33) (Figure 1A) as well as alignment of these genomes to the improved *C.*

deuterogattii genome assembly (34). This allowed determination of the ancestral state of the multiple alleles differentiating the VGIIa and VGIIa-like groups. We also generated a second independent genome sequence for our original NIH444 isolate (NIH444(1) and NIH444(2)), as well as an NIH444 isolate acquired from a different lab (NIH444(v)). We identified a total of 18 mutations that are shared by all of the VGIIa-like isolates sequenced: 13 SNPs and 5 INDELs (Table 1). Five SNPs are noncoding while the remaining eight coding SNPs are characterized by a high rate of nonsynonymous mutations (7/8 coding mutations). In addition we confirmed that one of the INDELs in this branch is a single base deletion in a coding exon of *MSH2 (24)*. This deletion is unique to the VGIIa-like group.

We confirmed both the presence of this mutation and showed that it results in a hypermutator phenotype. Sanger sequencing verified that this single base deletion is not an Illumina sequencing artifact (Figure 1B). This single base deletion occurs within an early exon of *MSH2* and results in a nonsense mutation, with both an out of frame message and an early stop codon present in the predicted transcript (Figure 1C).

MSH2 defect results in hypermutator phenotype

Because Msh2 is a critical component of the mismatch repair complex, we next tested whether the VGIIa-like strains displayed an increased mutation rate. We utilized a standard fluctuation assay approach to determine the rate of 5-FOA resistance, as previously applied in *Cryptococcus* (35). The resistance rate was increased in the hypermutator lineage between 4.4-fold at the minimum (CBS7750) and 6.6-fold at the maximum (ICB107) compared to the type strain VGIIa R265 (Figure 2A). The majority of 5-FOA resistance in *Cryptococcus neoformans* is acquired through mutation of *URA5* (36); therefore, the *URA5* locus was PCR amplified and sequenced in independently derived 5-FOA resistant isolates to determine the genetic basis of resistance. Mutations were characterized as either

substitutions, or indels of either single or multiple bases (Figure 2B). The majority of resistant isolates were the result of substitutions, and the mutation profile was similar between the hypermutator VGIIa-like strains and the non-hypermutator VGIIa strain EJB17 (EJB17 vs NIH444: P=1.00, EJB17 vs CBS7750 P= 0.60, EJB17 vs ICB107 p= 1.00, Fisher's Exact Test).

MSH2 mutants accumulate indels in homopolymer runs

Next, the whole genome data was analyzed to ascertain whether the similarity in mutation spectrum observed in the *URA5* locus was recapitulated at multiple loci or if the *URA5* locus had unique properties. To do this, we focused on two relatively long branches within the *C. deuterogattii* phylogeny: the branch separating the outgroup VGIIb strain R272 from the VGIIa clonal cluster (58,375 variants) and the private ICB107 variants, which represented the longest available hypermutator branch (763 variants) (Figure 8). The frequency of individual SNPs was examined first. While a high rate of transitions as compared to transversions was maintained in the hypermutator branch, there was a slight reduction in the frequency of A->G and T->C mutations relative to C->T and G->A mutations (Figure 3A). However, much more striking was a dramatic increase in indels within homopolymer runs (Figure 3B). While shifts in homopolymer runs in the R272 branch account for only 0.7% of the variants, in ICB107 they account for 42.9%, exceeding even the proportion of SNPs.

By looking further at the context of the homopolymer run mutations within the ICB107 branch, it is apparent that longer base runs appear to be less stable than shorter runs, as approximately 17% of the mutations occur within the context of a nine base homopolymer run (Figure 3C). This peak is likely a function of both the decreased stability of longer runs and the low quantity of longer homopolymer runs in the genome. Many of these longer runs occur within intergenic regions in the genome, but a substantial portion of the coding genes in *C. deuterogattii* contain at least one longer coding

homopolymer run (Figure 3C). The longest homopolymer run in the *URA5* locus contains only four bases, which is predicted to be both relatively stable and among the coding genes with relatively shorter homopolymer runs (Figure 3C). In contrast, the *FRR1* gene that encodes the FKBP12 homolog responsible for the antifungal action of both FK506 and rapamycin has a run of seven cytosines (7C) within its coding region. This places it both within the range of commonly mutated homopolymers in the ICB107 branch, and in the upper half of homopolymer containing genes in *C. deuterogattii*, suggesting that *FRR1* is an appropriate locus to test the effect of mismatch repair mutants on homopolymers containing genes (Figure 3C).

Genes containing homopolymer runs are highly unstable in *msh2* mutants

The *FRR1* locus was utilized to further assess whether VGIIa-like strains were hypermutators at more than one locus and to test the hypothesis that homopolymer runs are particularly unstable in these isolates. The protein product of *FRR1*, FKBP12, binds to either FK506 or rapamycin to form a protein-drug complex that inhibits calcineurin or TOR, respectively. By selecting with both drugs at the same time at the non-permissive temperature of 37°C, loss of function mutations in *FRR1* are selected as the only single step mutation conferring resistance to both drugs. Here we utilized a semi-quantitative assay to identify gross differences in mutation rate whereby independent colonies were grown in liquid culture and then swabbed onto quadrants of a selective plate. On FK506/rapamycin media, the wildtype strains produced a small number of resistant colonies (Figure 4A). In contrast, the NIH444 hypermutator strain produced prolific FK506/rapamycin resistant colonies. Surprisingly, the CBS7750 hypermutator was completely resistant to both drugs. Upon examining the *FRR1* locus in the whole genome sequence data we discovered that there was already a single base deletion within the 7C homopolymer run in this strain. This may represent an unselected drug resistance phenotype

potentiated by the *msh2* defect. We attempted to use this assay for the ICB107 strain as well but were unsuccessful, likely as a result of its inability to grow at the higher growth temperature necessary for this assay that is based upon calcineurin-dependent growth at 37°C (Figure 7B).

Next a standard fluctuation assay approach was employed to quantify the mutation rate difference between the VGIIa and VGIIa-like hypermutators. There was a greater than 120-fold increase in mutation rate to FK506/rapamycin resistance in NIH444 compared to R265, in contrast with the maximum of 6.6-fold increase observed for 5-FOA resistance (Figure 4C). As before, we selected independent resistant colonies, PCR amplified, and sequenced the *FRR1* gene to determine the mechanism of resistance. All resistance was explained by mutations within *FRR1* as expected, but while R265 still primarily acquired resistance through substitutions, NIH444 now almost exclusively underwent single base additions or single base deletions, all within the 7C homopolymer run (R265 vs NIH444, p=0.0003, Fisher's Exact Test). These indels resulted in nonsense mutations and resistance to both FK506 and rapamycin.

Hypermutation phenotype is linked to the *msh2* defect

Once we established a *bona fide* hypermutator phenotype in the VGIIa-like strains, we sought to verify that this phenotype was linked to the *msh2* del131 allele. We crossed the NIH444 strain to a very closely related R265**a** congenic strain recently generated by serial backcrossing (37). *C. deuterogattii* matings are less fertile than *C. neoformans*, but we successfully dissected 14 viable spores from this cross. We typed these strains for the *MSH2* allele by sequencing and observed 1:1 segregation of the *MSH2* and the *msh2* del131 alleles. The spores were then tested for rate of FK506/rapamycin resistance using the same semi-quantitative swabbing assay described above (Figure 5A). While five of the *msh2* del131 strains displayed the predicted hypermutator phenotype, surprisingly two of the spores

(#4 and 5) appeared wildtype despite inheritance of a defective copy of *MSH2*.

Mating in *Cryptococcus* has previously been demonstrated to produce phenotypic plasticity through generation of aneuploid progeny at a high rate (38). These aneuploids often display defects in growth at high temperature; thus, we tested whether the high growth temperature of 37°C employed in our FK506/rapamycin resistance assay may have masked the ability of these meiotic progeny to develop drug resistance. Congruent with this hypothesis, both spore products #4 and #5 demonstrated growth defects at 37°C and 39°C (Figure 5C). Whole genome sequencing confirmed that both segregants carried an extra copy of one chromosome, suggesting that a 1N+1 aneuploidy was causing temperature sensitivity (Figure 5B). In fact, all of the progeny from this mating had an unusually high rate of aneuploidy with 8 of the 14 progeny aneuploid for at least one scaffold/chromosome (Table 2). By passaging all of the spores on YPD at 37°C we were able to restore the ability to grow at high temperatures in all of the progeny after four to nine passages after which all no longer demonstrated a high temperature growth defect (Figure 5C). As a result the passaged derivatives of isolates #4 and #5 now demonstrated a hypermutator phenotype as expected (Figure 5D) and all 14 progeny produced the results predicted through linkage of the *msh2* defect to the hypermutator phenotype (Supplemental Figure 1). Further, whole genome typing of the 14 meiotic progeny demonstrated that sufficient recombination had occurred to establish linkage. All 14 progeny demonstrated unique SNP profiles and only one SNP demonstrated inheritance congruent with the hypermutator phenotype: a SNP on scaffold 3 that was linked to the *msh2* del131 indel (~175 kb) (Figure 5E).

As a final verification of phenotypic linkage, we constructed two independent deletions of the *MSH2* gene in the R265 VGIIa background via biolistic transformation and homologous recombination replacing the *MSH2* ORF with a neomycin resistance cassette. A complete deletion of *MSH2* resulted in the same hypermutator phenotype as the *msh2* del131 allele, providing further evidence that the

hypermutator phenotype was linked to the loss of function in *MSH2* (Figure 5F). A fluctuation assay was used to determine the mutation rate for resistance to FK506 and rapamycin and demonstrated that the null mutants did not have an elevated mutation rate in comparison to the NIH444 strain (Figure 5G). This suggests that suppressors have not arisen on the NIH444 background to moderate the effects of the *msh2* del131 allele.

Hypermutation enables phenotypic diversity including reversion of mutations

During growth of the *msh2*Δ*::NEO* strains a spontaneous *ade2* mutant was fortuitously identified based on its classic red pigmentation (Figure 6A). We verified that this was the result of a defect in adenine biosynthesis as this strain grows on YNB supplemented with adenine, but not YNB media (Figure 6B). As expected, sequencing of the *ADE2* locus revealed that the ade$^-$ phenotype was linked to a single base mutation that results in a His->Arg amino acid change (Figure 6C).

The red pigment produced by *ade2* mutants is a toxic intermediate in adenine synthesis that accumulates in the vacuoles of mutant cells. As a result, these mutants have a growth defect, and suppressor mutations can readily be isolated that eliminate production of this toxic intermediate and now produce white colonies (Figure 6C). We isolated two red and two white derivative colonies and sequenced the *ADE2* locus. While the red colonies retained the causative substitution, one of the white colonies retained it and the other had undergone reversion at this site back to the functional nucleotide (Figure 6C). We confirmed that the revertant isolate was no longer an adenine auxotroph and that the second white isolate was still auxotrophic, despite the lack of pigmentation (Figure 6D). This is likely the result of a second mutation upstream in the adenine biosynthetic pathway, resulting in further inactivation and loss of the ability to produce the toxic red intermediate. We tested a number of independent white reverted colonies for the ability to grow on media lacking adenine and observed that

direct reversion to functional adenine biosynthesis was more common than additional inactivating mutations, with 37/42 (88%) white colonies demonstrating adenine prototrophy (Figure 6E). Taken together, these results suggest that hypermutation allows both frequent inactivation of pathways, and also the means to either adapt to the consequences or to simply directly revert the original mutation. In the context of antifungal drug action or activity, this phenotypic switching could be highly important.

Hypermutation does not directly affect virulence

We previously demonstrated that the VGIIa-like mutants exhibit diminished virulence in a nasal instillation model of virulence (30, 31). Here, we tested whether loss of Msh2 function was directly responsible for the decrease in virulence but observed no difference in virulence between R265 and two independent *de novo* deletions of *msh2* (Figure 7A). This suggested that the defects in murine virulence in the VGIIa-like isolates could be the result of multiple independent deleterious mutations unique to each lineage. We next tested the ability to grow at higher temperature, a critical virulence factor, in the VGIIa-like isolates. All three isolates showed defects in high temperature growth, varying from minor defects in NIH444, the isolate with the shortest branch to R265, to moderate defects in CBS7750, the isolate with an intermediate length branch, and finally severe defects in ICB107, the isolate with the longest branch affected by the hypermutator (Figure 7B). This suggests that instead of an immediate change in virulence, defects in mismatch repair may instead result in loss of virulence over time as mutations accumulate in critical pathways. Thus, rather than the virulence difference being explained by a single change or set of changes shared by all three VGIIa-like strains, instead they may all be avirulent for their own unique reasons.

Hypermutation is deleterious in rich conditions, but advantageous under stress

Our data suggests that over time a high mutation rate comes at a cost to organisms. We next addressed the early stages of hypermutator emergence and the direct effects of Msh2 inactivation to address if temporary benefits may be conferred. To this end, independent *msh2 de novo* deletions were employed in competition experiments with wildtype R265. Briefly, liquid YPD cultures were mixed in 50:50 ratios of wildtype R265 and a neomycin resistant *msh2* mutant or a random insertion of the neomycin resistance cassette as a control. After 48 hours incubation, co-cultures were spread onto YPD plates and then colonies were picked and restruck to media containing neomycin to determine the percentage of colonies derived from each original strain. If mutations are neutral, the expectation is that the 50:50 ratio will be maintained. Instead, growth defects were observed for the *msh2* mutants at either 30°C or 37°C, suggesting that the *msh2* mutation is deleterious under rich growth conditions (Figure 7C). Notably, in three of the 24 replicates at 30°C or 37°C the *msh2* mutant was able to grow moderately better than the wildtype parent strain. This may indicate that in these individual experimental replicates the hypermutator allowed a beneficial mutation and hitchhiked to higher frequency. However, under highly stressful conditions, such as exposure to the drugs FK506/rapamycin at 37°C, the hypermutator was highly advantageous. In multiple replicates the hypermutator rapidly acquired resistance and overtook the entire co-culture. This suggests that mutator alleles can be advantageous when a population faces an evolutionary landscape where large effect beneficial mutations exist. In contrast, mutators are deleterious in landscapes where few large effect mutations can provide advantages. Antifungal treatment and transitions from environment to host are likely to provide opportunities for adaptive mutation and favor mutator alleles.

Hypermutator lineage is derived and hypermutation did not cause the VGIIa outbreak

Historically, *C. deuterogattii* has been thought of as a tropical and subtropical pathogen. As a

result, the origin of the VGIIa outbreak in the Pacific Northwest, a non-tropical environment, was surprising. In addition, 1) the South American origin of the ICB107 strain (within the proposed cradle of the *C. deuterogattii* species (32)), 2) the isolation of NIH444 in Seattle near the outbreak origin, and 3) the diminished virulence of the VGIIa-like subclade (30, 31), all combined to suggest that the VGIIa-like group might have been the immediate precursors to the clonal VGIIa outbreak cluster (24, 33). Identification of the hypermutator phenotype further suggested that the defect in Msh2, carried by the older VGIIa-like group, may have played a role in adaptation to the climate of the Pacific Northwest and may have even potentiated the increase in virulence. To test this, a maximum parsimony phylogeny of the VGIIa-like and VGIIa strains was constructed, including additional sequences for a related South American VGII clade not included in our previous analysis (33) and using the VGIIb R272 strain as an outgroup (Figure 8). As described above, the *msh2* mutation can allow restoration of function mutations at the exact mutation site that are indistinguishable from the original sequence (Figure 6). Likewise, mating could also reintroduce a functional *MSH2* allele. For these reasons, we tested for the possible presence and impact of the past *msh2* del131 mutator allele throughout the phylogeny by examining the imputed mutation spectrum.

As discussed above (Figure 3), branches with defects in mismatch repair show an increased frequency of shifts in homopolymer runs. Increases in homopolymer run shifts were observed only on the most proximal branch ancestral to the VGIIa-like mutator lineage. This suggests that the mutator phenotype is congruent with the presence of the allele throughout the VGIIa-like phylogeny, and that the VGIIa group did not experience a transient period of *msh2* mediated hypermutation, followed by repair or mating-mediated replacement. Instead, the VGIIa-like lineage may represent a unique pathway to adaptation distinct from that followed by the VGIIa group that resulted in diminished rather than enhanced virulence.

Discussion

In this study we have identified and characterized a successful lineage of eukaryotic hypermutators. Elevated mutations rates are a common adaptive mechanism in bacteria, but are typically thought of as transient states that allow beneficial mutations in the short term but are selected against in the long term (3). Bacteria solve this problem through horizontal gene transfer of genes from the mismatch repair pathway, allowing the initial beneficial mutation to be separated from the deleterious mutator allele (4). In contrast, few cases have been identified in natural isolates of eukaryotes, suggesting that variation in mutation rate may play a less substantial role than in bacteria. Two recent studies have identified exceptions to this rule. In *S. cerevisiae*, incompatible alleles of *PMS1* and *MLH1* result in elevated mutation rates if present in combination (18, 19). However, in the rare cases where the incompatible arrangement is found in nature, additional suppressors have arisen to restore wildtype mutation rate, suggesting that hypermutator phenotypes may not be tolerated over long periods of time (20, 21). In addition, a recent study of *Candida glabrata* has demonstrated that a substantial proportion of clinical isolates (>50%) carry nonsynonymous mutations in the *MSH2* gene, some causing elevated mutation rates similar to the null phenotype and others with intermediate changes in mutation rate (23). The authors correlate the mismatch repair defects with multi-drug resistance. A second study with a different sample cohort confirmed the presence of *MSH2* mutations, but concluded that drug resistance was better correlated with drug exposure than with mismatch repair defects (39).

In contrast with the studies in *S. cerevisiae* and *C. glabrata*, here we identified a group of viable hypermutator strains resulting from a nonsense mutation in *MSH2*. These strains were isolated over a period of fifteen years from two different continents and include both clinical and environmental

strains. In addition, the *msh2* del131 allele carried by these strains results in a complete loss of function, rather than simply reducing the efficiency of mismatch repair. This is distinct from the mutators identified in *C. glabrata*, and suggests that the VGIIa-like lineage is a successful and relatively long-lived hypermutator lineage, capable of both disseminating over a large area and persisting in the environment. *Cryptococcus deuterogattii* may also represent an intermediate between mutator-rich *C. glabrata* and mutator-poor *S. cerevisiae*. *C. deuterogatttii* is not an obligate pathogen and can cycle between an environmental lifestyle and an infectious lifestyle as an "accidental pathogen" (26). Elevated mutation rates may be selected for by Red Queen interactions with a host, meaning that pathogens like *C. glabrata* that grow only in association with their hosts would experience higher mutation rates than facultative pathogens like *C. deuterogattii*, or effectively nonpathogenic yeasts like *S. cerevisiae*. In fact, a putative recurrent infection of *Cryptococcus neoformans* was recently demonstrated to contain nonsense mutations in *MSH2*, *MSH5*, and *RAD5*, predicted to result in a mutator phenotype (40).

An alternate hypothesis is that ploidy may play a role in the success of mutator strains. Both *C. deuterogattii* and *C. glabrata* exist primarily as haploids in nature, while *S. cerevisiae* is predominantly a diploid. Diploid strains are buffered from the effects of loss of function mutations like those observed in homopolymer runs in this study, which may reduce the supply of large effect beneficial mutations for mutator alleles to hitchhike upon in diploids. However, past work in *S. cerevisiae* suggests that diploids adapt more quickly in *msh2* mutant backgrounds, rather than less, suggesting that ploidy may play the opposite role, at least in *S. cerevisiae* (41).

Evolutionary theory predicts that the admixture provided by sex in a population nullifies the ability of mutator alleles to hitchhike to high frequency (17). Unlike obligately sexual animals, or asexual bacteria, fungi can reproduce both sexually and asexually, with the frequency of sex varying

399 substantially between different species. We previously described *C. deuterogattii* as a species
400 characterized by long periods of mitotic clonal expansion and only intermittent sexual crosses (24). A
401 contibuting factor is likely the highly biased mating type distribution in *C. deuterogattii*, as the vast
402 majority of *C. deuterogattii* isolates are *MAT*α, with only a handful of *MAT***a** isolates described globally
403 (31). In *Cryptococcus deneoformans*, which shares a similar biased mating type distribution, this has
404 resulted in the development of a unisexual α-α sexual cycle, dispensing with the obligate need for a
405 *MAT***a** partner (42). This unisexual cycle can result in both *de novo* variation, but also recombination
406 and admixture at similar levels to that observed in typical bisexual crosses (38, 42, 43). However, no
407 laboratory unisexual cycle has been observed in *C. deuterogattii* to this point. Consequently, as strains
408 are introduced to new locales, they may need to survive and adapt via mitotic growth without sexual
409 crosses for long periods of time. This could elevate linkage between mutator alleles and the beneficial
410 mutations they elicit, but also eliminate the ability to separate those beneficial mutations from
411 additional deleterious alleles.

412 We also showed that hypermutators can allow both inactivating mutations in genes, and
413 reversion of those mutations to the wildtype. Mismatch repair-defective mutator strains are
414 characterized by particularly high rates of slippage within homopolymer runs, and these occur at run
415 lengths that are common within fungal genomes. This may indicate that eukaryotes may harbor
416 contingency loci featuring homopolymer runs, like those observed in bacteria (44). Fungal contingency
417 loci could be critical for responses to antifungal agents, both in the environment and the clinical setting.
418 In addition, variation could be important in host-pathogen interactions as well. For example, in *S.*
419 *cerevisiae*, tandem repeats are enriched within cell surface genes, which could enable alteration of
420 antigenic diversity during pathogenic interactions (45). While these loci are unstable even in wildtype
421 populations, defects in the MMR pathway could enhance this instability and result in increased

422 diversity. In *C. neoformans*, one mutator strain has been previously described in a lab-passaged strain
423 notable for frequent mutations in the RAM pathway, which enables a dimorphic transition between
424 pseudohyphal and yeast phase growth, although the molecular basis of the elevated mutation rate is
425 unknown (35). This transition is highly important for survival in the face of environmental threats such
426 as amoeba, but compromises pathways responsible for high temperature growth, suggesting that
427 oscillation between RAM+ and RAM- states may allow populations to survive as conditions change.
428 In addition, we provided evidence that the VGIIa-like lineage was not a subvirulent progenitor
429 of the Pacific Northwest outbreak, in contrast with previous hypotheses from two groups (24, 33).
430 Rather, the VGIIa-like lineage may represent either an alternative pathway of adaptation or perhaps a
431 preview of the evolutionary trajectory of the outbreak. The ability to grow at high temperatures is a
432 critical component of virulence for fungal pathogens. However, many pathogenic fungi are thought to
433 be “accidental” pathogens where virulence factors are selected for by other environmental factors. In
434 this model, thermotolerance could be selected by high ambient temperatures in a non-host environment
435 (46). Moving from a tropical/sub-tropical environment in South America to the more temperate climate
436 of the Pacific Northwest may relieve this high temperature selection. Consequently, the loss of ability
437 to grow under thermal stress observed in the VGIIa-like strains may be adaptive for the majority of
438 their growth conditions in the environment, but also render them relatively avirulent in mammals. This
439 could suggest that as the VGIIa outbreak strains continue to adapt to the Pacific Northwest, they will
440 gradually lose virulence potential, making the outbreak self-limiting over time. Alternatively, the
441 defects could simply reflect a decrease in viability caused by a long period of growth with an elevated
442 mutation rate, resulting in mutational meltdown. The most severe temperature defect was observed in
443 ICB107, a strain isolated from a patient in Brazil and also the strain with the largest number of
444 mutations separating it from the VGIIa group, which would support the second hypothesis. These

445 strains may have been able to persist in the Pacific Northwest in spite of the high mutation load because
446 of a bottleneck effect during their introduction.

447 Our original hypothesis was that the common variants that differentiated the VGIIa group from
448 the progenitor of the VGIIa-like mutator group would explain the decline in virulence observed in the
449 VGIIa-like strains (24, 31). However, the *msh2* mutation, the only predicted large effect mutation, has
450 no obvious immediate effect on virulence in a murine inhalation model. This is interesting, at least in
451 part, because *msh2* mutants of *C. neoformans* were previously reported to exhibit a modest growth
452 advantage in mouse lungs in pooled signature tagged mutant experiments (47). This suggests that the
453 mutator strains are better able to adapt to growth in the mouse lung, but not in a way that correlates
454 directly with virulence in a dissemination and survival model of virulence. In addition, our results
455 suggest that loss of virulence in the VGIIa-like strains may be a consequence of independent private
456 mutations, i.e. each VGIIa-like strain is avirulent for its own reason. In previous work, CBS7750 and
457 ICB107 have the most substantial virulence defects, while NIH444 has a more modest defect (31). This
458 is congruent with the fact that NIH444 has the shortest divergence from the VGIIa clonal group, while
459 CBS7750 and ICB107 have much longer branches, and is also congruent with the high temperature
460 growth defects we observed.

461 Finally, we demonstrated that mutation spectrum analysis could be utilized within a phylogeny
462 to infer the presence of a hypermutator allele through an increase in the proportion of indels within
463 homopolymer runs. In bacteria, the hypermutator state is often transient and wildtype mismatch repair
464 is often restored via horizontal gene transfer (4). This paradigm could also operate in fungi, with sexual
465 recombination replacing direct DNA transfer. However, these events are difficult to detect if the
466 hypermutator allele is purged from the population by selection after it is separated from its linked
467 beneficial allele. The widespread availability of whole genome sequencing could now allow changes in

mutation spectrum to be used to detect episodic hypermutation throughout entire populations of microbes. We suspect that these episodes are common throughout the evolution of eukaryotic microbes and may be even more common among pathogenic microbes, reflecting their natural history as well as the result of Red Queen host-pathogen conflicts.

Materials and Methods

Strains and media

Strains used in this study are listed in table S1. Strains were routinely grown on YPD media at 30°C and maintained in permanent glycerol stocks at -80°C. Strains marked with neomycin resistance were grown on YPD supplemented with G418.

Genetic crosses and spore dissection

To conduct crosses and isolate spores, NIH444 was cocultured with the R265**a** congenic strain (37) on solid V8 pH =5.0 medium in the dark for eight weeks. Basidiospores were isolated using a microdissection microscope equipped with a 25-µm needle (Cora Styles Needles 'N Blocks, Dissection Needle Kit) as previously described (48).

Virulence assays

Approximately eight week old A/J mice were anesthetized with phenobarbital via intraperitoneal injection and then were infected with $5x10^4$ cells from strains R265, RBB17, and RBB18 by intranasal inhalation. Ten animals per group were infected. Mice were monitored daily for signs of cryptococcal infection and sacrificed when exhibiting signs of clinical distress. The animal study was conducted in the Division of Laboratory Animal Resources (DLAR) facilities at Duke University Medical Center (DUMC). All of the animal work was performed according to the guidelines of NIH and Duke University Institutional Animal Care and Use Committee (IACUC) under protocol number A245-13-09.

494 **Gene disruption and strain construction**

495 Deletions of the *MSH2* gene were constructed using a standard overlap PCR approach as
496 previously described (49). Briefly, 1 kb flanking regions of genomic DNA were amplified from both
497 the 5' and 3' regions flanking the *MSH2* open reading frame in R265 and the selectable marker for *NEO*
498 resistance was amplified from plasmid pJAF1. An overlap PCR was carried out to generate a full-
499 length deletion cassette and R265 was transformed using biolistic transformation. Gene replacement
500 was confirmed using in-gene, 5' junction, 3' junction, and spanning PCRs. Two independent
501 transformations of independent overnight cultures of R265 were carried out to isolate independent
502 mutants.

503 The R265 *NEO* resistance marked strain was constructed by transforming XbaI digested pJAF1
504 plasmid into wild-type R265 by biolistics. Transformants were selected for on YPD containing
505 neomycin. Transformants were checked for tandem insertions using primers JOHE40500/JOHE40501
506 and a strain with only a single insertion was chosen and employed for the competition assay.

507

508 **Fluctuation assays**

509 Fluctuation assays were performed on either synthetic medium containing 5-FOA (1 g/L) at
510 30°C or YPD supplemented with rapamycin (1 µg/mL) and FK506 (1 µg/mL) at 37°C. For each strain
511 tested, ten independent 5 mL YPD liquid cultures were grown overnight at 30°C. Cultures were then
512 split and either spread directly on selective media or diluted and spread on solid YPD to determine
513 colony forming units. Resistant colonies and total colonies were counted and mutation rate was
514 calculated using the Maximum Likelihood method as implemented via the FALCOR calculator (50).
515

Co-culture competition assays

Competition assays were carried out by growing independent liquid cultures overnight in 5 mL YPD. Cultures were then counted using a hemocytometer and 500,000 cells each of a tester neomycin resistant strain and a wildtype R265 culture were mixed in a 5 mL liquid YPD culture. Co-cultures were grown for 48 hours and then spread onto a YPD plate, such that individual colonies could be isolated. All colonies were picked (up to 60 and at least 5) and restruck to neomycin media to determine the proportion of colony forming units derived from each of the original strains in the competition.

Genome sequencing and assembly

DNA was isolated with the CTAB isolation protocol as previously described (51). Library construction and genome sequencing were carried out at the University of North Carolina Next Generation Sequencing Facility. Paired-end libraries with approximately 300-base inserts were constructed. A number of genome sequences were previously available (24, 33). The remaining were generated here and are deposited on the SRA under project accession no. PRJNA387047. All sequences were mapped to the V2 R265 reference genome (34). Alignment was performed using the short read component of the BWA aligner (52). Further processing was carried out using the Genome Analysis Toolkit (GATK) version 3.1-1 (53), including SAMtools (54) and Picard. SNPs and indels were called with UnifiedGenotyper from GATK, using the haploid setting. GATK's VariantAnnotator was used to define the homopolymer context of indels. VCFtools was used for filtering and extracting private variants (55). SnpEff was employed to determine the predicted impact of mutations (56).

Inference of phylogeny

Maximum parsimony phylogenies were constructed using MEGA6 (57) to analyze SNP matrices extracted from VCF format using a custom Perl script. SNPs and indels were placed on branches where they were predicted to change states using VCFtools (55) to extract the variants supporting each node.

Acknowledgements

Supported by NIH/NIAID R37 grant AI39115-19 and R01 grant AI50113-13 to JH. We thank Sheng Sun, Shelby Priest, Paul Magwene, and Tom Petes for helpful discussion and suggestions during the writing of this manuscript.

549

550 1. **Tenaillon O**, **Barrick JE**, **Ribeck N**, **Deatherage DE**, **Blanchard JL**, **Dasgupta A**, **Wu GC**,
551 **Wielgoss S**, **Cruveiller S**, **Médigue C**, **Schneider D**, **Lenski RE**. 2016. Tempo and mode of
552 genome evolution in a 50,000-generation experiment. Nature **536**:165–170.

553 2. **Oliver A**, **Canton R**, **Campo P**, **Baquero F**, **Blazquez J**. 2000. High frequency of
554 hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. Science **288**:1251–
555 1254.

556 3. **Wielgoss S**, **Barrick JE**, **Tenaillon O**, **Wiser MJ**, **Dittmar WJ**, **Cruveiller S**, **Chane-Woon-**
557 **Ming B**, **Médigue C**, **Lenski RE**, **Schneider D**. 2013. Mutation rate dynamics in a bacterial
558 population reflect tension between adaptation and genetic load. Proc Natl Acad Sci **110**:222–
559 227.

560 4. **Denamur E**, **Lecointre G**, **Darlu P**, **Tenaillon O**, **Acquaviva C**, **Sayada C**, **Sunjevaric I**,
561 **Rothstein R**, **Elion J**, **Taddei F**, **Radman M**, **Matic I**. 2000. Evolutionary implications of the
562 frequent horizontal transfer of mismatch repair genes. Cell **103**:711–721.

563 5. **LeClerc JE**, **Li B**, **Payne WL**, **Cebula TA**. 1996. High mutation frequencies among
564 *Escherichia coli* and *Salmonella pathogens*. Science **274**:1208–1211.

565 6. **Himmelreich R**, **Plagens H**, **Hilbert H**, **Reiner B**, **Herrmann R**. 1997. Comparative analysis
566 of the genomes of the bacteria *Mycoplasma pneumoniae* and *Mycoplasma genitalium*. Nucleic
567 Acids Res **25**:701–712.

568 7. **Su SS**, **Modrich P**. 1986. *Escherichia coli* mutS-encoded protein binds to mismatched DNA
569 base pairs. Proc Natl Acad Sci **83**:5057–5061.

570 8. **Au KG**, **Welsh K**, **Modrich P**. 1992. Initiation of methyl-directed mismatch repair. J Biol Chem
571 **267**:12142–12148.

572 9. **Grilley M**, **Welsh KM**, **Su SS**, **Modrich P**. 1989. Isolation and characterization of the
573 *Escherichia coli* mutL gene product. J Biol Chem **264**:1000–1004.

574 10. **Cox EC**, **Degnen GE**, **Scheppe ML**. 1972. Mutator gene studies in *Escherichia coli*: the *mutS*
575 gene. Genetics **72**:551 LP-567.

576 11. **Reenan RA**, **Kolodner RD**. 1992. Isolation and characterization of two *Saccharomyces*
577 *cerevisiae* genes encoding homologs of the bacterial HexA and MutS mismatch repair proteins.
578 Genetics **132**:963–73.

579 12. **Strand M**, **Prolla TA**, **Liskay RM**, **Petes TD**. 1993. Destabilization of tracts of simple
580 repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature **365**:274–6.

581 13. **Harfe BD**, **Jinks-Robertson S**. 2000. DNA mismatch repair and genetic instability. Annu Rev
582 Genet **34**:359–399.

583 14. **Datta A**, **Hendrix M**, **Lipsitch M**, **Jinks-Robertson S**. 1997. Dual roles for DNA sequence
584 identity and the mismatch repair system in the regulation of mitotic crossing-over in yeast. Proc
585 Natl Acad Sci U S A **94**:9757–9762.

586 15. **Alani E**, **Reenan RA**, **Kolodner RD**. 1994. Interaction between mismatch repair and genetic
587 recombination in *Saccharomyces cerevisiae*. Genetics **137**:19–39.

588 16. **Rayssiguier C**, **Thaler DS**, **Radman M**. 1989. The barrier to recombination between
589 *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. Nature
590 **342**:396–401.

591 17. **Tenaillon O**, **Le Nagard H**, **Godelle B**, **Taddei F**. 2000. Mutators and sex in bacteria: conflict
592 between adaptive strategies. Proc Natl Acad Sci U S A **97**:10465–70.

593 18. **Heck JA**, **Argueso JL**, **Gemici Z**, **Reeves RG**, **Bernard A**, **Aquadro CF**, **Alani E**. 2006.
594 Negative epistasis between natural variants of the *Saccharomyces cerevisiae MLH1* and *PMS1*
595 genes results in a defect in mismatch repair. Proc Natl Acad Sci U S A **103**:3256–3261.

596 19. **Bui DT**, **Dine E**, **Anderson JB**, **Aquadro CF**, **Alani EE**. 2015. A genetic incompatibility
597 accelerates adaptation in yeast. PLOS Genet **11**:e1005407.

598 20. **Bui DT**, **Friedrich A**, **Al-Sweel N**, **Liti G**, **Schacherer J**, **Aquadro CF**, **Alani E**. 2017.
599 Mismatch repair incompatibilities in diverse yeast populations. Genetics.

600 21. **Skelly DA**, **Magwene PM**, **Meeks B**, **Murphy HA**. 2017. Known mutator alleles do not
601 markedly increase mutation rate in clinical *Saccharomyces cerevisiae* strains. Proceedings Biol
602 Sci **284**.

603 22. **Bielas JH**, **Loeb KR**, **Rubin BP**, **True LD**, **Loeb LA**. 2006. Human cancers express a mutator
604 phenotype. Proc Natl Acad Sci U S A **103**:18238–18242.

605 23. **Healey KR**, **Zhao Y**, **Perez WB**, **Lockhart SR**, **Sobel JD**, **Farmakiotis D**, **Kontoyiannis DP**,
606 **Sanglard D**, **Taj-Aldeen SJ**, **Alexander BD**, **Jimenez-Ortigosa C**, **Shor E**, **Perlin DS**. 2016.
607 Prevalent mutator genotype identified in fungal pathogen *Candida glabrata* promotes multi-drug
608 resistance. Nat Commun **7**:11128.

609 24. **Billmyre RB**, **Croll D**, **Li W**, **Mieczkowski P**, **Carter DA**, **Cuomo CA**, **Kronstad JW**,
610 **Heitman J**. 2014. Highly recombinant VGII *Cryptococcus gattii* population develops clonal
611 outbreak clusters through both sexual macroevolution and asexual microevolution. MBio
612 **5**:e01494-14.

613 25. **Hagen F**, **Khayhan K**, **Theelen B**, **Kolecka A**, **Polacheck I**, **Sionov E**, **Falk R**, **Parnmen S**,
614 **Lumbsch HT**, **Boekhout T**. 2015. Recognition of seven species in the *Cryptococcus*
615 *gattii*/*Cryptococcus neoformans* species complex. Fungal Genet Biol **78**:16–48.

616 26. **Springer DJ**, **Phadke S**, **Billmyre B**, **Heitman J**. 2012. *Cryptococcus gattii*, no longer an
617 accidental pathogen? Curr Fungal Infect Rep **6**:245–256.

618 27. **Feretzaki M**, **Billmyre RB**, **Clancey SA**, **Wang X**, **Heitman J**. 2016. Gene network
619 polymorphism illuminates loss and retention of novel RNAi silencing components in the
620 *Cryptococcus* pathogenic species complex. PLoS Genet **12**:e1005868.

621 28. **D'Souza CA**, **Kronstad JW**, **Taylor G**, **Warren R**, **Yuen M**, **Hu G**, **Jung WH**, **Sham A**, **Kidd**
622 **SE**, **Tangen K**, **Lee N**, **Zeilmaker T**, **Sawkins J**, **McVicker G**, **Shah S**, **Gnerre S**, **Griggs A**,
623 **Zeng Q**, **Bartlett K**, **Li W**, **Wang X**, **Heitman J**, **Stajich JE**, **Fraser JA**, **Meyer W**, **Carter D**,
624 **Schein J**, **Krzywinski M**, **Kwon-Chung KJ**, **Varma A**, **Wang J**, **Brunham R**, **Fyfe M**,
625 **Ouellette BFF**, **Siddiqui A**, **Marra M**, **Jones S**, **Holt R**, **Birren BW**, **Galagan JE**, **Cuomo**
626 **CA**. 2011. Genome variation in *Cryptococcus gattii*, an emerging pathogen of
627 immunocompetent hosts. MBio **2**:e00342-10.

628 29. **Kidd SE**, **Hagen F**, **Tscharke RL**, **Huynh M**, **Bartlett KH**, **Fyfe M**, **Macdougall L**, **Boekhout**
629 **T**, **Kwon-Chung KJ**, **Meyer W**. 2004. A rare genotype of *Cryptococcus gattii* caused the
630 cryptococcosis outbreak on Vancouver Island (British Columbia, Canada). Proc Natl Acad Sci U
631 S A **101**:17258–63.

632 30. **Fraser JA**, **Giles SS**, **Wenink EC**, **Geunes-Boyer SG**, **Wright JR**, **Diezmann S**, **Allen A**,
633 **Stajich JE**, **Dietrich FS**, **Perfect JR**, **Heitman J**. 2005. Same-sex mating and the origin of the
634 Vancouver Island *Cryptococcus gattii* outbreak. Nature **437**:1360–4.

635 31. **Byrnes EJ**, **Li W**, **Lewit Y**, **Ma H**, **Voelz K**, **Ren P**, **Carter DA**, **Chaturvedi V**, **Bildfell RJ**,
636 **May RC**, **Heitman J**. 2010. Emergence and pathogenicity of highly virulent *Cryptococcus*
637 *gattii* genotypes in the northwest United States. PLoS Pathog **6**:e1000850.

638 32. **Hagen F**, **Ceresini PC**, **Polacheck I**, **Ma H**, **van Nieuwerburgh F**, **Gabaldón T**, **Kagan S**,
639 **Pursall ER**, **Hoogveld HL**, **van Iersel LJJ**, **Klau GW**, **Kelk SM**, **Stougie L**, **Bartlett KH**,
640 **Voelz K**, **Pryszcz LP**, **Castañeda E**, **Lazera M**, **Meyer W**, **Deforce D**, **Meis JF**, **May RC**,
641 **Klaassen CHW**, **Boekhout T**. 2013. Ancient dispersal of the human fungal pathogen
642 *Cryptococcus gattii* from the Amazon rainforest. PLoS One **8**:e71148.

643 33. **Engelthaler DM**, **Hicks ND**, **Gillece JD**, **Roe CC**, **Schupp JM**, **Driebe EM**, **Gilgado F**,
644 **Carriconde F**, **Trilles L**, **Firacative C**, **Ngamskulrungroj P**, **Castañeda E**, **Lazera MDS**,
645 **Melhem MSC**, **Pérez-Bercoff A**, **Huttley G**, **Sorrell TC**, **Voelz K**, **May RC**, **Fisher MC**,
646 **Thompson GR**, **Lockhart SR**, **Keim P**, **Meyer W**. 2014. *Cryptococcus gattii* in North

American Pacific Northwest: Whole-population genome analysis provides insights into species evolution and dispersal. MBio **5**:e01464-14-.

34. **Farrer RA**, **Desjardins CA**, **Sakthikumar S**, **Gujja S**, **Saif S**, **Zeng Q**, **Chen Y**, **Voelz K**, **Heitman J**, **May RC**, **Fisher MC**, **Cuomo CA**. 2015. Genome evolution and innovation across the four major lineages of *Cryptococcus gattii*. MBio **6**:e00868-15.

35. **Magditch DA**, **Liu T-B**, **Xue C**, **Idnurm A**. 2012. DNA mutations mediate microevolution between host-adapted forms of the pathogenic fungus *Cryptococcus neoformans*. PLoS Pathog **8**:e1002936.

36. **Kwon-Chung KJ**, **Varma A**, **Edman JC**, **Bennett JE**. 1992. 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.

37. **Zhu P**, **Zhai B**, **Lin X**, **Idnurm A**. 2013. Congenic strains for genetic analysis of virulence traits in *Cryptococcus gattii*. Infect Immun **81**:2616–25.

38. **Ni M**, **Feretzaki M**, **Li W**, **Floyd-Averette A**, **Mieczkowski P**, **Dietrich FS**, **Heitman J**. 2013. Unisexual and heterosexual meiotic reproduction generate aneuploidy and phenotypic diversity de novo in the yeast *Cryptococcus neoformans*. PLoS Biol **11**:e1001653.

39. **Dellière S**, **Healey K**, **Gits-Muselli M**, **Carrara B**, **Barbaro A**, **Guigue N**, **Lecefel C**, **Touratier S**, **Desnos-Ollivier M**, **Perlin DS**, **Bretagne S**, **Alanio A**. 2016. Fluconazole and echinocandin resistance of *Candida glabrata* correlates better with antifungal drug exposure rather than with *MSH2* mutator genotype in a French cohort of patients harboring low rates of resistance. Front Microbiol **7**:2038.

40. **Rhodes J**, **Beale MA**, **Vanhove M**, **Jarvis JN**, **Kannambath S**, **Simpson JA**, **Ryan A**, **Meintjes G**, **Harrison TS**, **Fisher MC**, **Bicanic T**. 2017. A population genomics approach to assessing the genetic basis of within-host microevolution underlying recurrent cryptococcal meningitis infection. G3 (Bethesda) **7**:1165–1176.

41. **Thompson DA**, **Desai MM**, **Murray AW**. 2006. Ploidy controls the success of mutators and nature of mutations during budding yeast evolution. Curr Biol **16**:1581–1590.

42. **Lin X**, **Hull CM**, **Heitman J**. 2005. Sexual reproduction between partners of the same mating type in *Cryptococcus neoformans*. Nature **434**:1017–1021.

43. **Sun S**, **Billmyre RB**, **Mieczkowski PA**, **Heitman J**. 2014. Unisexual reproduction drives meiotic recombination and phenotypic and karyotypic plasticity in *Cryptococcus neoformans*. PLOS Genet **10**:e1004849.

44. **Moxon R**, **Bayliss C**, **Hood D**. 2006. Bacterial contingency loci: the role of simple sequence

680 DNA repeats in bacterial adaptation. Annu Rev Genet **40**:307–333.

681 45. **Verstrepen KJ**, **Jansen A**, **Lewitter F**, **Fink GR**. 2005. Intragenic tandem repeats generate
682 functional variability. Nat Genet **37**:986–90.

683 46. **Casadevall A**, **Steenbergen JN**, **Nosanchuk JD**. 2003. "Ready made" virulence and "dual use"
684 virulence factors in pathogenic environmental fungi — the *Cryptococcus neoformans* paradigm.
685 Curr Opin Microbiol **6**:332–337.

686 47. **Liu OW**, **Chun CD**, **Chow ED**, **Chen C**, **Madhani HD**, **Noble SM**. 2008. Systematic genetic
687 analysis of virulence in the human fungal pathogen *Cryptococcus neoformans*. Cell **135**:174–88.

688 48. **Hsueh Y-P**, **Idnurm A**, **Heitman J**. 2006. Recombination hotspots flank the *Cryptococcus*
689 mating-type locus: implications for the evolution of a fungal sex chromosome. PLoS Genet
690 **2**:e184.

691 49. **Davidson RC**, **Cruz MC**, **Sia RA**, **Allen B**, **Alspaugh JA**, **Heitman J**. 2000. Gene disruption
692 by biolistic transformation in serotype D strains of *Cryptococcus neoformans*. Fungal Genet Biol
693 **29**:38–48.

694 50. **Hall BM**, **Ma C-X**, **Liang P**, **Singh KK**. 2009. Fluctuation AnaLysis CalculatOR: a web tool
695 for the determination of mutation rate using Luria–Delbrück fluctuation analysis. Bioinformatics
696 **25**:1564–1565.

697 51. **Pitkin JW**, **Panaccione DG**, **Walton JD**. 1996. A putative cyclic peptide efflux pump encoded
698 by the *TOXA* gene of the plant-pathogenic fungus *Cochliobolus carbonum*. Microbiology
699 **142**:1557–1565.

700 52. **Li H**, **Durbin R**. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform.
701 Bioinformatics **25**:1754–60.

702 53. **McKenna A**, **Hanna M**, **Banks E**, **Sivachenko A**, **Cibulskis K**, **Kernytsky A**, **Garimella K**,
703 **Altshuler D**, **Gabriel S**, **Daly M**, **DePristo MA**. 2010. The Genome Analysis Toolkit: A
704 MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res
705 **20**:1297–303.

706 54. **Li H**, **Handsaker B**, **Wysoker A**, **Fennell T**, **Ruan J**, **Homer N**, **Marth G**, **Abecasis G**,
707 **Durbin R**. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**:2078–
708 9.

709 55. **Danecek P**, **Auton A**, **Abecasis G**, **Albers CA**, **Banks E**, **DePristo MA**, **Handsaker RE**,
710 **Lunter G**, **Marth GT**, **Sherry ST**, **McVean G**, **Durbin R**. 2011. The variant call format and
711 VCFtools. Bioinformatics **27**:2156–8.

712 56. **Cingolani P**, **Platts A**, **Wang LL**, **Coon M**, **Nguyen T**, **Wang L**, **Land SJ**, **Lu X**, **Ruden DM**.
713 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms,
714 SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly (Austin)
715 **6**:80–92.

716 57. **Tamura K**, **Stecher G**, **Peterson D**, **Filipski A**, **Kumar S**. 2013. MEGA6: Molecular
717 Evolutionary Genetics Analysis version 6.0. Mol Biol Evol **30**:2725–2729.

718 58. **Springer DJ**, **Ren P**, **Raina R**, **Dong Y**, **Behr MJ**, **McEwen BF**, **Bowser SS**, **Samsonoff WA**,
719 **Chaturvedi S**, **Chaturvedi V**. 2010. Extracellular fibrils of pathogenic yeast *Cryptococcus*
720 *gattii* are important for ecological niche, murine virulence and human neutrophil interactions.
721 PLoS One **5**:e10978.

722

Figure legends

Figure 1. VGIIa-like isolates harbor a frameshift resulting in a predicted nonfunctional protein.

A) The VGIIa-like clade is part of the clonal radiation that includes the Pacific Northwest Outbreak. All three isolates are characterized by a frameshift in the *MSH2* gene at position 131. B) Sanger sequencing of the *MSH2* gene confirms that NIH444 and CBS7750 have both undergone deletion of a single T within the coding region of *MSH2*, while the VGIIa R265 strain and the outgroup VGIIb strain have not. C) This deletion results in a frameshift beginning in the first functional domain of *MSH2* and an early premature stop. This truncated protein is predicted to be non-functional.

Figure 2. *msh2* mutants are hypermutators.

A) A fluctuation assay for resistance to 5-FOA was carried out for both wildtype VGIIa strains (R265 and EJB17) and hypermutator VGIIa-like strains (NIH444, EJB17, and ICB107), as well as an outgroup VGIIb (R272) strain. Resistance rates were normalized to the rate observed in R265. The VGIIa-like strains demonstrate an increase in mutation rate. Data shown are the mean of ten replicates with 95% confidence interval. B) The molecular basis of resistance was determined for 5-FOA resistance at the *URA5* locus. All isolates tested demonstrated predominantly substitutions as the molecular basis of resistance.

Figure 3. Hypermutators are characterized by high homopolymer instability.

A) Percentage of transitions and transversions among SNPs private to R272, an outgroup, and ICB107, a hypermutator, are shown. Transitions are more common in both strains, while one type of transition is modestly reduced in the hypermutator. B) All private mutations within R272 and ICB107 were characterized either as SNPs, indels, or homopolymer indels. The cutoff used to distinguish a

homopolymer indel was longer than 4 bases. ICB107 had a substantial increase in homopolymer run shifts. C) Frequency of indels within all homopolymers longer than 1 base from the private ICB107 mutations are shown in blue. All genes from the *C. deuterogattii* transcriptome were grouped by frequency of the longest homopolymer run within the coding region of that gene and shown in red. *URA5* only has a run of 4 bases, while *FRR1*, which encodes FKBP12, the target of FK506 and rapamycin, has a run of 7 bases.

Figure 4. A homopolymer run within the *FRR1* gene allows rapid inactivation and resistance to FK506 and rapamycin.

A) Swab assays of the VGIIa and VGIIa-like strains were conducted to determine whether VGIIa-like stains develop resistance to FK506 and rapamycin at 37°C at an elevated rate compared to VGIIa. NIH444 demonstrated a large increase in resistance rate, while CBS7750 was completely resistant. B) The resistance in CBS7750 is attributable to a single base deletion within the coding 7C run in the *FRR1* gene that has been fixed in this strain. C) Fluctuation assays for NIH444 and R265 show that the hypermutator conferred a greater than 100-fold increase in mutation rate compared to the wildtype VGIIa strain. Data shown are the mean of ten replicates with 95% confidence intervals and are normalized to the rate in R265. D) Analysis of the molecular basis of resistance shows that substitutions are still the predominant mechanism for resistance in R265, but in the hypermutator strain single base additions and single base deletions within the homopolymer run are responsible for the vast majority of resistance.

Figure 5. The hypermutator phenotype is linked to the frameshift in *MSH2*.

A) Progeny from a cross between NIH444 and R265 show co-segregation of the *msh2* del131 allele

with the hypermutator phenotype in all except two cases. B) Whole genome depth of coverage plots show that scaffolds five and six are aneuploid in spores #4 and #5, respectively. C) Temperature sensitivity assay for the incongruent spores before and after serial passaging at 37°C. D) After passage at 37°C, the non-congruent spores now properly demonstrate linkage with the *msh2* del131 allele. E) Whole genome sequencing of NIH444 x R265**a** progeny, with variant regions indicated as reference in blue or alternate in red. The boxed region indicates a SNP in close linkage to the *msh2* del131 allele. F) Two independent *de novo* deletions of *MSH2* via biolistic transformation demonstrate elevated mutation rates in an FK506/rapamycin swabbing assay at 37°C. G) Fluctuation assays for R265, NIH444, and two independent deletions of *msh2* in R265 show that the hypermutator phenotype is recapitulated in the null mutants. Data shown are the mean of ten replicates with 95% confidence intervals and are normalized to the rate observed in R265.

Figure 6. Hypermutation allows inactivation and reactivation of adenine biosynthetic pathway. A) A spontaneous red colony was isolated from the *de novo msh2* deletion (RBB18). B) This colony (RBB22) demonstrated adenine auxotrophy, suggesting that it was an *ade2* mutant. C) Sequencing of the *ADE2* locus confirmed that the original colony was an *ade2* mutant. In addition, two red (RBB25 and RBB26) and two white derivatives (RBB23 and RBB24) were tested. One white derivative had reverted the original mutation (RBB23), while the second had eliminated production of the red intermediate but had not reverted the original *ade2* mutation (RBB24). D) One revertant colony (RBB23) demonstrated adenine prototrophy, while the other (RBB24) remained an auxotroph despite losing the red pigmentation. E) An assay to test direct reversion frequency versus secondary mutation to eliminate the red toxic intermediate demonstrated that the most common mutations were direct reversions and restoration of adenine prototrophy.

Figure 7. Hypermutation does not have immediate virulence defects but may potentiate long term deficits.

A) Virulence tested in the murine inhalation model was not strongly affected by deletion of the *MSH2* gene. B) All three VGIIa-like strains demonstrated defects in high temperature growth via a spot dilution assay in comparison with the VGIIa R265 strain and the VGIIb R272 strain as an outgroup. Strains with longer branches exhibited larger high temperature growth defects. C) Competition experiments between a tester strain with the neomycin resistance marker and the wildtype R265 strain. (Strain used: SEC501, RBB17, RBB18). Original cultures were mixed in a 1:1 ratio and then grown overnight in liquid YPD. Both hypermutators showed a modest growth defect at 30°C and 37°C but a dramatic growth advantage in the high stress FK506/rapamycin 37°C condition. Boxplots show minimum, first quartile, median, third quartile, and maximum values. Points represent the results from six individual replicates summarized by the box plot. The *NEO* vs WT competition is shown in gray, while the two *msh2*Δ::*NEO* competitions are shown in dark and light blue.

Figure 8. The VGIIa-like hypermutator is derived and not ancestral to the Pacific Northwest outbreak.

A maximum parsimony phylogeny of the VGIIa group with the VGIIb R272 genome as an outgroup demonstrates that the VGIIa-like group is a branch parallel to the VGIIa group. To test for the presence of a defect in MMR throughout the tree, the mutation spectrum was examined on each branch. High rates of homopolymer run shifts were observed throughout the VGIIa-like group, but no evidence was apparent at branch A). Instead it appears that the hypermutator first arose on branch B).

Supplemental Figure 1. Passaged spore progeny demonstrate linkage between *msh2*del131 allele and mutator phenotype.

Four independent overnight cultures of spore progeny passaged at 37°C to select against aneuploidies that confer temperature sensitivity were struck on YPD media containing FK506 and rapamycin at 37°C. All fourteen passaged spore progeny demonstrated behavior predicted by the state of the *MSH2* locus for the majority of the four samples tested. Spores #4 and #5 and controls are depicted in Figure 5.

Supplemental Figure 2. Reversion of *ade2* mutants primarily occurs through repair of *ADE2*.

42 white colonies were selected as independent revertants from a red *ade2* mutant (RBB22). Colonies were struck to both YPD and YNB to test for auxotrophy indicative of either direct reversion or a second site mutation upstream of *ade2* in the adenine biosynthetic pathway.

828 Supplemental Table 1: Strains used in this study.

Strain	**Genotype**	**Source/Reference**
R265	VGIIa	(29)
EJB17	VGIIa	(31)
R272	VGIIb	(30)
NIH444	VGIIa-like, *msh2*del131	(30)
NIH444(v)	VGIIa-like, *msh2*del131	(58)
CBS7750	VGIIa-like, *msh2*del131	(30)
ICB107	VGIIa-like, *msh2*del131	(30)
RBB17	R265, *msh2*Δ::*NEO*	This study
RBB18	R265, *msh2*Δ::*NEO*	This study
RBB22	RBB18, *msh2*Δ::*NEO*, *ade2*	This study
RBB23	RBB22, *msh2*Δ::*NEO*, white, *ADE2*	This study
RBB24	RBB22, *msh2*Δ::*NEO*, white, *ade2*	This study
RBB25	RBB22, *msh2*Δ::*NEO*, red, *ade2*	This study
RBB26	RBB22, *msh2*Δ::*NEO*, red, *ade2*	This study
ECt1	XL280 *ade2*::*NEO*	This study
R265**a**	R265, *MAT***a**	(37)
SEC016	NIH444xR265**a** spore #1, *msh2*del131	This study
SEC017	NIH444xR265**a** spore #2	This study
SEC018	NIH444xR265**a** spore #3, *MAT***a**, *msh2*del131	This study
SEC019	NIH444xR265**a** spore #4, *MAT***a**, *msh2*del131	This study
SEC020	NIH444xR265**a** spore #5, *msh2*del131	This study
SEC021	NIH444xR265**a** spore #6, *MAT***a**, *msh2*del131	This study
SEC022	NIH444xR265**a** spore #7	This study
SEC023	NIH444xR265**a** spore #8	This study
SEC024	NIH444xR265**a** spore #9, *MAT***a**	This study
SEC025	NIH444xR265**a** spore #10, *MAT***a**	This study
SEC026	NIH444xR265**a** spore #11	This study
SEC027	NIH444xR265**a** spore #12, *MAT***a**	This study
SEC028	NIH444xR265**a** spore #13, *msh2*del131	This study

SEC029	NIH444xR265**a** spore #14, *MAT***a**, *msh2*del131	This study
SEC501	R265, *NEO*	This study
SEC559	SEC016, passaged 4 times at 37°C	This study
SEC560	SEC017, passaged 4 times at 37°C	This study
SEC561	SEC018, passaged 4 times at 37°C	This study
SEC562	SEC019, passaged 4 times at 37°C	This study
SEC563	SEC020, passaged 4 times at 37°C	This study
SEC564	SEC021, passaged 4 times at 37°C	This study
SEC565	SEC022, passaged 4 times at 37°C	This study
SEC566	SEC023, passaged 4 times at 37°C	This study
SEC567	SEC024, passaged 4 times at 37°C	This study
SEC568	SEC025, passaged 4 times at 37°C	This study
SEC569	SEC026, passaged 4 times at 37°C	This study
SEC570	SEC027, passaged 4 times at 37°C	This study
SEC571	SEC028, passaged 4 times at 37°C	This study
SEC572	SEC029, passaged 4 times at 37°C	This study
SEC573	R265**a**, passaged 4 times at 37°C	This study
SEC574	NIH444, passaged 4 times at 37°C	This study
SEC575	SEC562, passaged 5 additional times at 37°C	This study

829

830

831 Supplemental Table 2. Oligonucleotides used in this study.

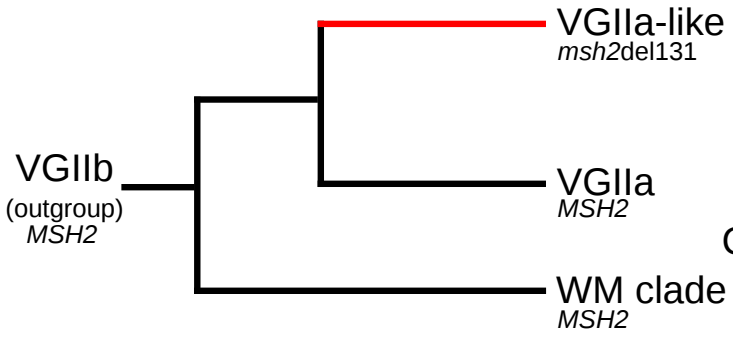
Primer	Sequence (5' to 3')	Comment
M13F	GTAAAACGACGGCCAG	universal oligo
M13R	CAGGAAACAGCTATGAC	universal oligo
JOHE41021/BB265	GAGCGTATTTCGAAGCAGG	*MSH2* 5' F deletion
JOHE41022/BB266	CATCTGGCCATAGTGACGC	*MSH2* 5' F nested deletion
JOHE41023/BB267	CTGGCCGTCGTTTTACAAAGACGAACTTCATTATGAGC	*MSH2* 5' R deletion
JOHE41024/BB268	GTCATAGCTGTTTCCTGCCTTGACAAGTTTGATCGC	*MSH2* 3' F deletion
JOHE41025/BB269	CGACATTTGATGAACCTTCACC	*MSH2* 3' R deletion
JOHE41026/BB270	GTCATATACCCGGCACACTTCG	*MSH2* 3' R nested deletion
JOHE41027/BB271	CCGAGAAAGCAGAAGTGACC	*MSH2* F sequencing
JOHE41245/BB275	CCAAGCAGATGCGTATCG	*MSH2* R sequencing
JOHE42336/BB280	CCTTCCAACAGGCCAAAGTGG	*ADE2* 5' F for sequencing
JOHE42337/BB281	GGTAATTTGTGCCTGACTGG	*ADE2* 5' F nested for sequencing
JOHE42338/BB282	GCGAAGTCCAGCCAAGTCC	*ADE2* F ~400 for sequencing
JOHE42339/BB283	GGCTGAGAAGGCAGTCGG	*ADE2* F ~800 for sequencing
JOHE42340/BB284	GGATTGATCTCATTGCACCTCC	*ADE2* F ~1300 for sequencing
JOHE42341/BB285	CCGACATCAAGCCAACAGG	*ADE2* 3' R for sequencing
JOHE42342/BB286	CGTGCTGCAGATGCTGG	*ADE2* 3' R nested for sequencing
JOHE42343/BB287	CCACGCCGTCTAGTACACTGG	*ADE2* R ~1600 for sequencing
JOHE42344/BB288	GCAGCACCAGAAACAGTGAGAGC	*ADE2* R ~1100 for sequencing
JOHE42345/BB289	CCCAACCTTCAGCATAAAGAGG	*ADE2* R ~600 for sequencing
JOHE26938/BB14	GTCTTCCCAAGCCCTCGACTC	*URA5* 5' F for sequencing
JOHE26939/BB15	CCGGTGAGCCATATCGCAGC	*URA5* 5' F nested for sequencing
JOHE26941/BB17	CCTGTACTTCCTGACCTCTCG	*URA5* 3' R for sequencing
JOHE26942/BB18	CCCACTTTCCGGAGCCTTCC	*URA5* 3' R nested for sequencing
JOHE40363/BB257	GTGTTTGGACGAGCAGTCGG	*FRR1* 3' R for sequencing
JOHE40364/BB258	GCAGCAATGCAATCCTGG	*FRR1* 3' R nested for sequencing
JOHE40365/BB259	GGTACAGGGCGTTGGACC	*FRR1* 5' F for sequencing
JOHE40366/BB260	CGACCTGCAATAGTTTCCC	*FRR1* 5' F nested for sequencing

JOHE40500/SEC129	GGGATGACAGGAGATCCTGC	NEO 3' R for insert confirmation
JOHE40501/SEC130	GCAACAATCCATCCGTGCTGG	NEO 5' F for insert confirmation

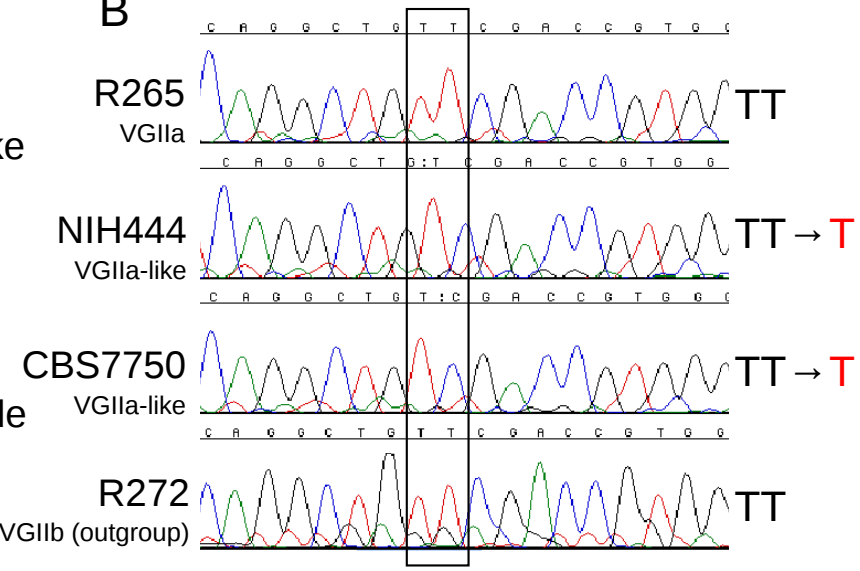
832

Figure 1

A



B



C

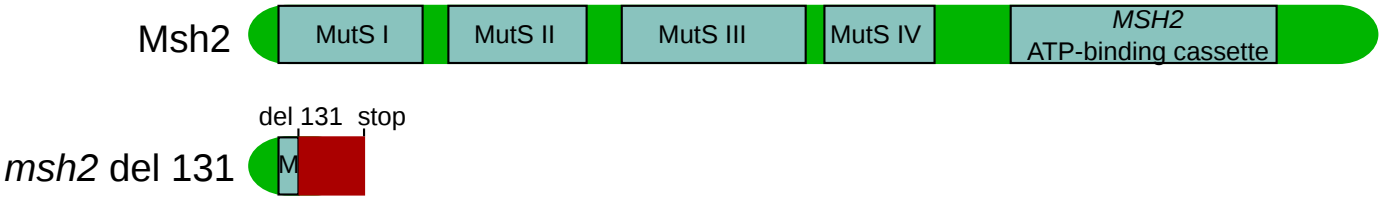


Figure 2

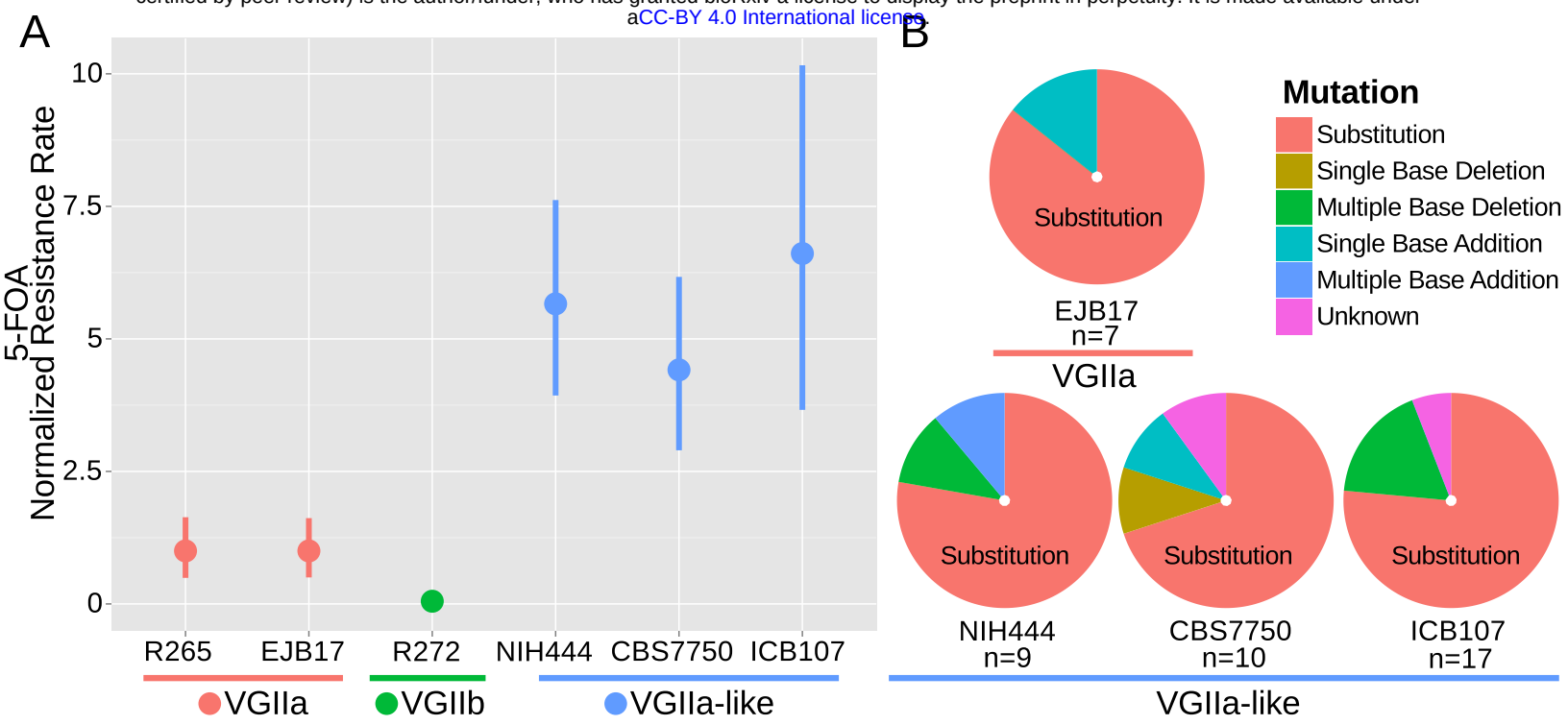


Figure 3

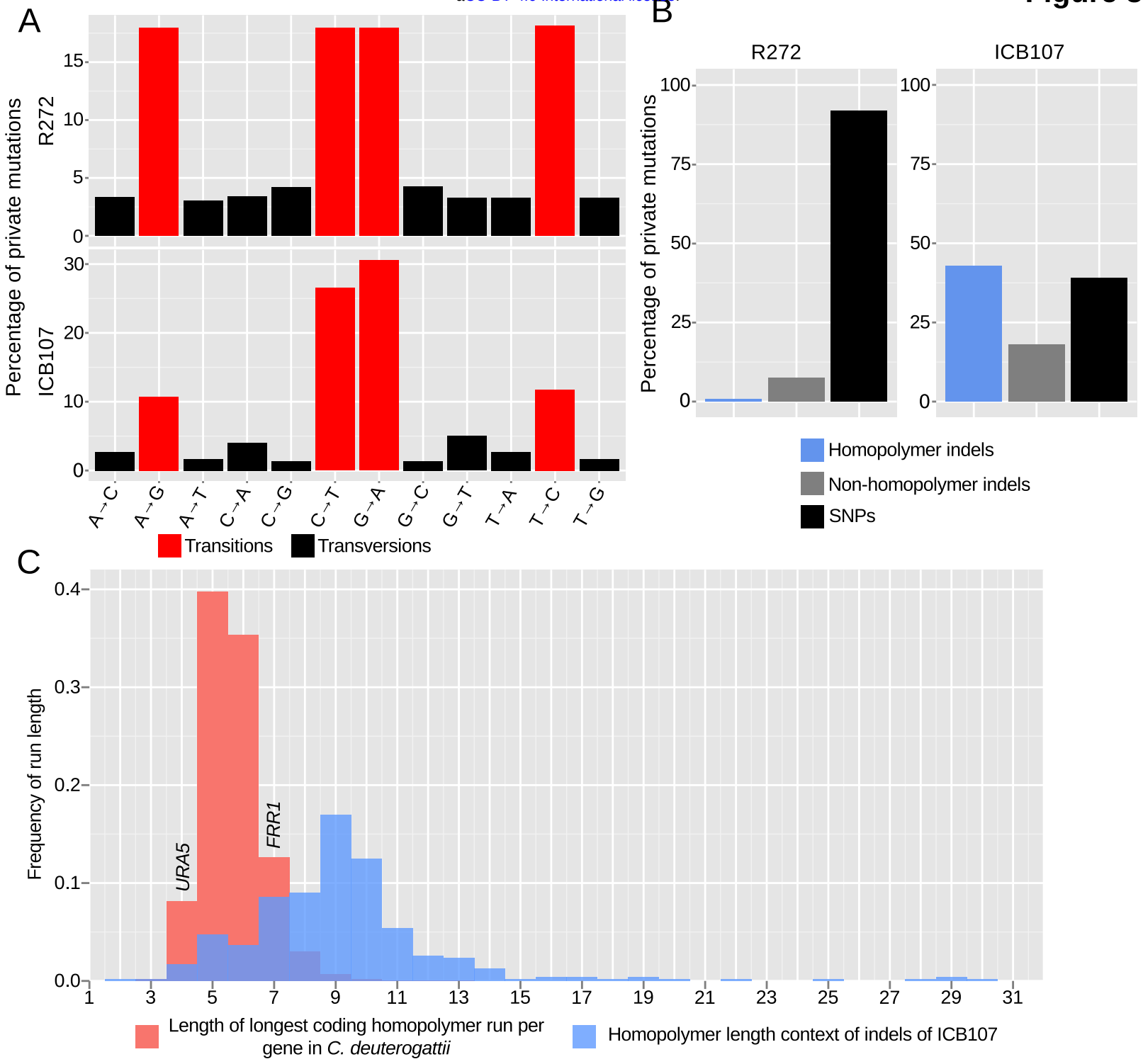
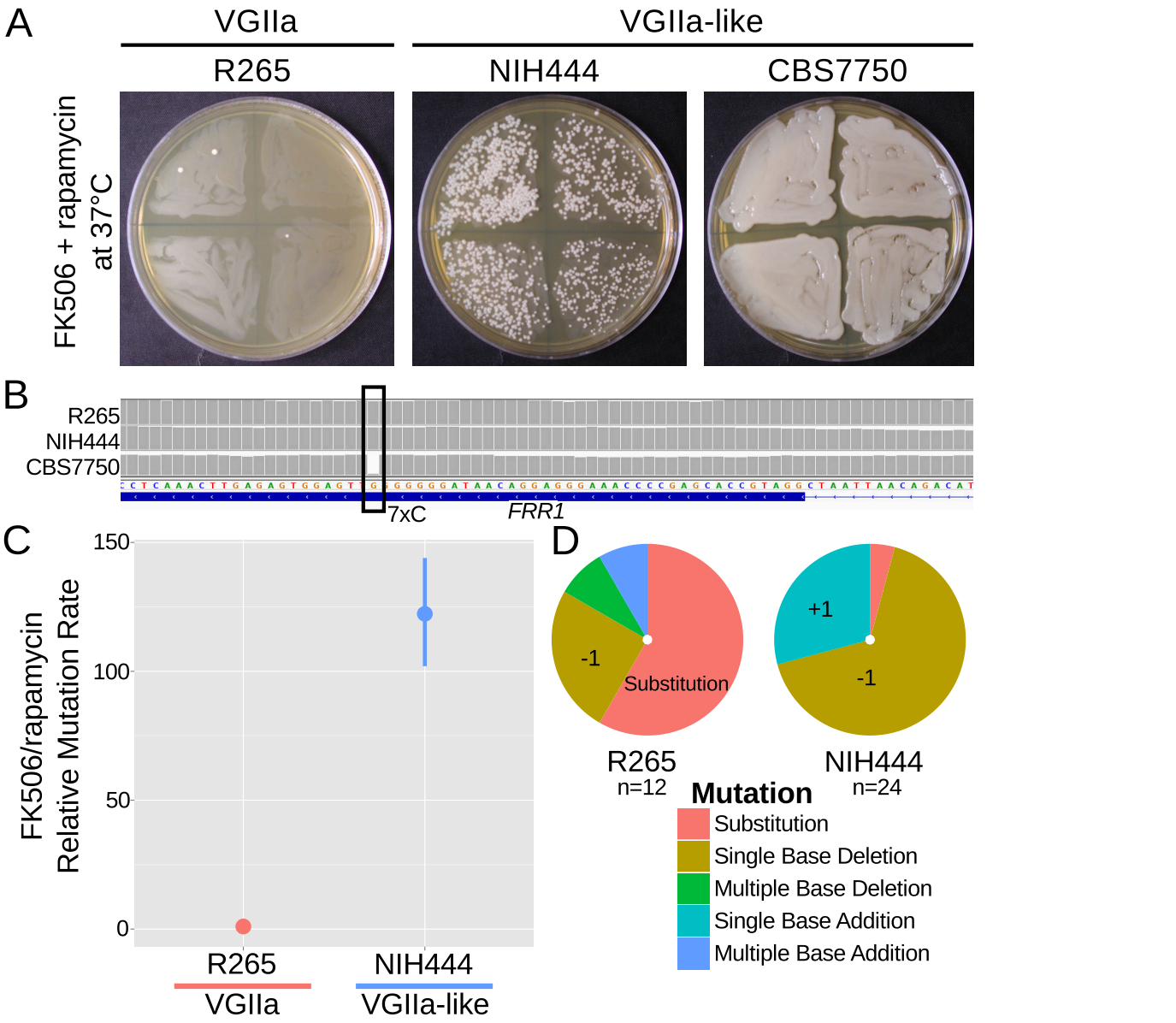


Figure 4



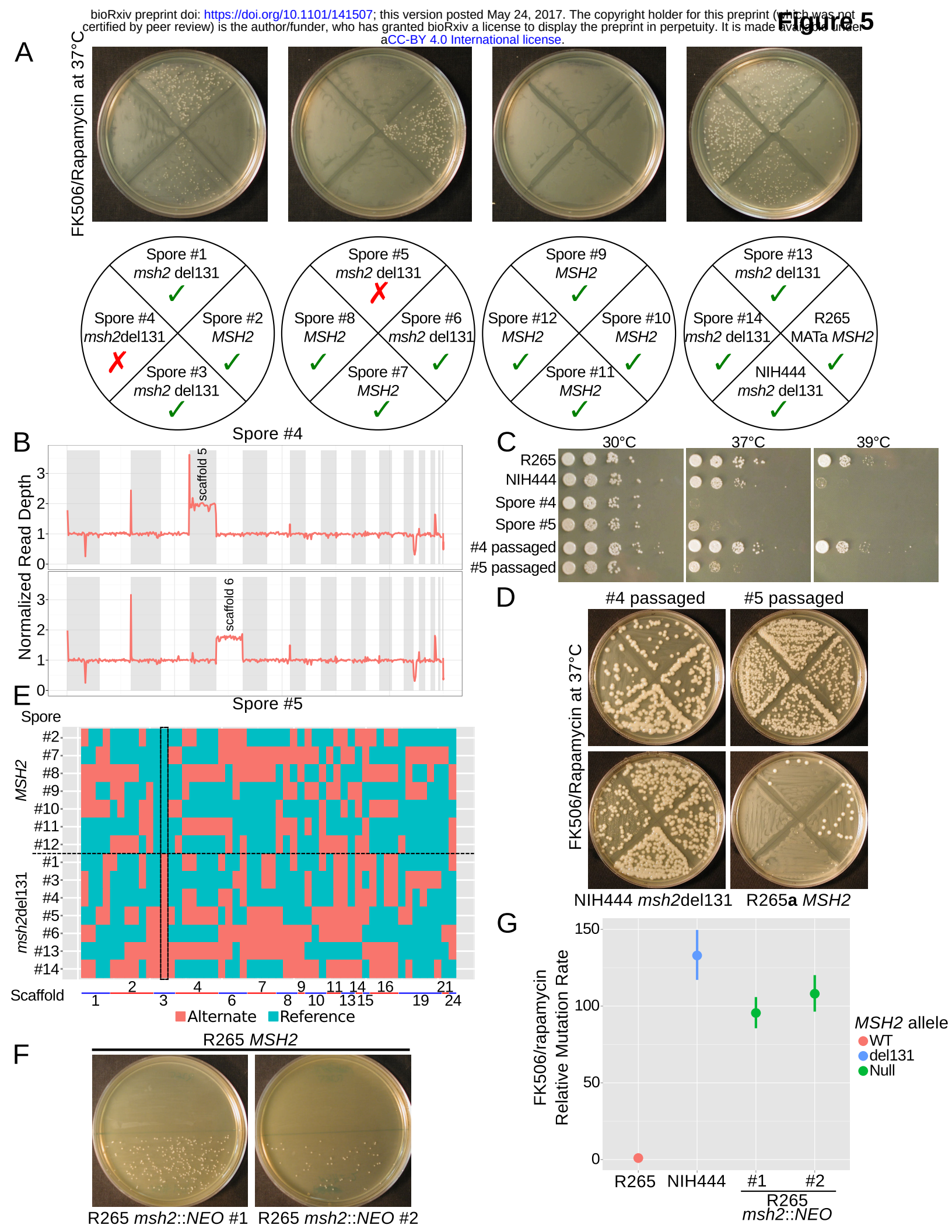


Figure 6

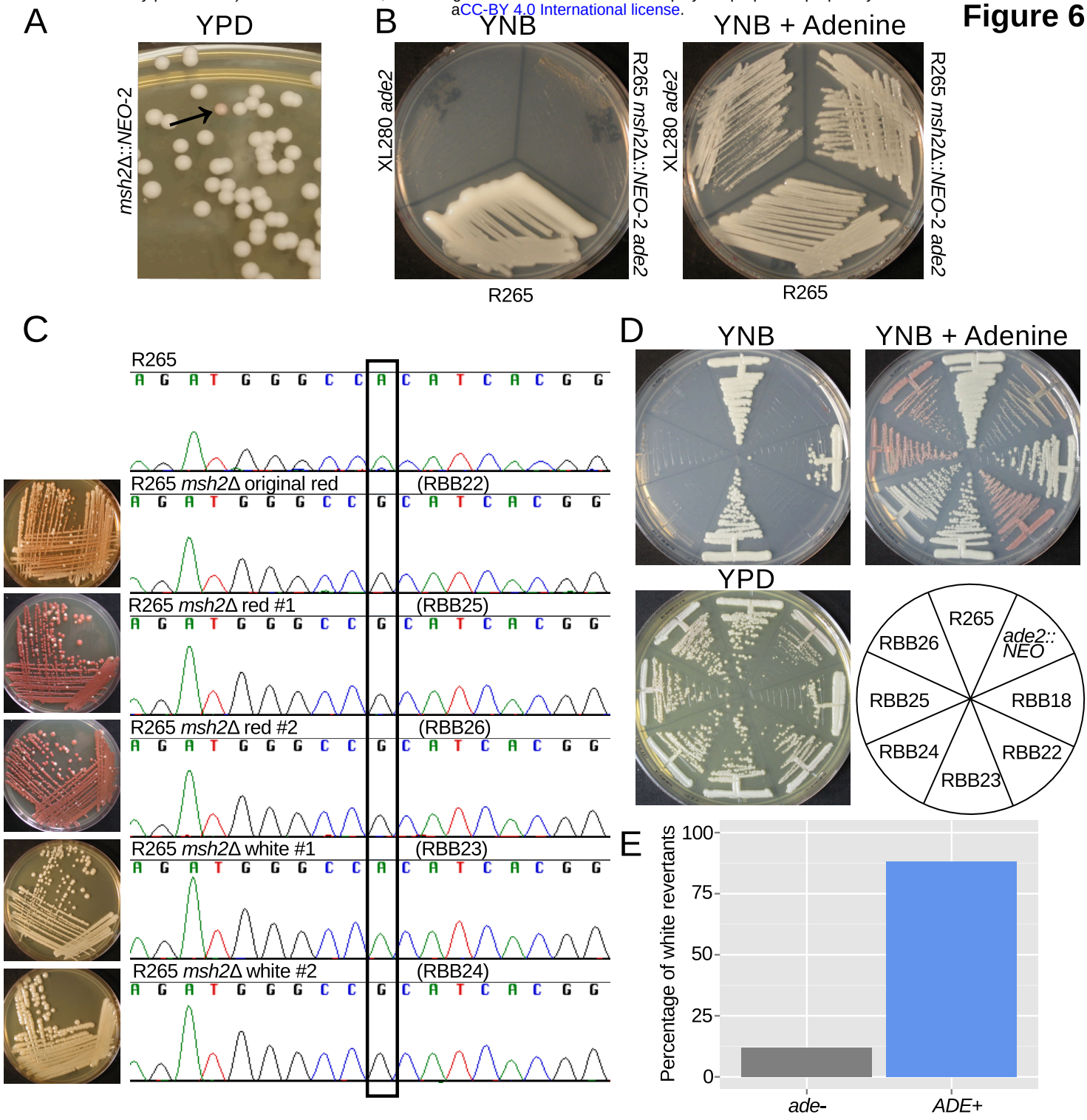


Figure 7

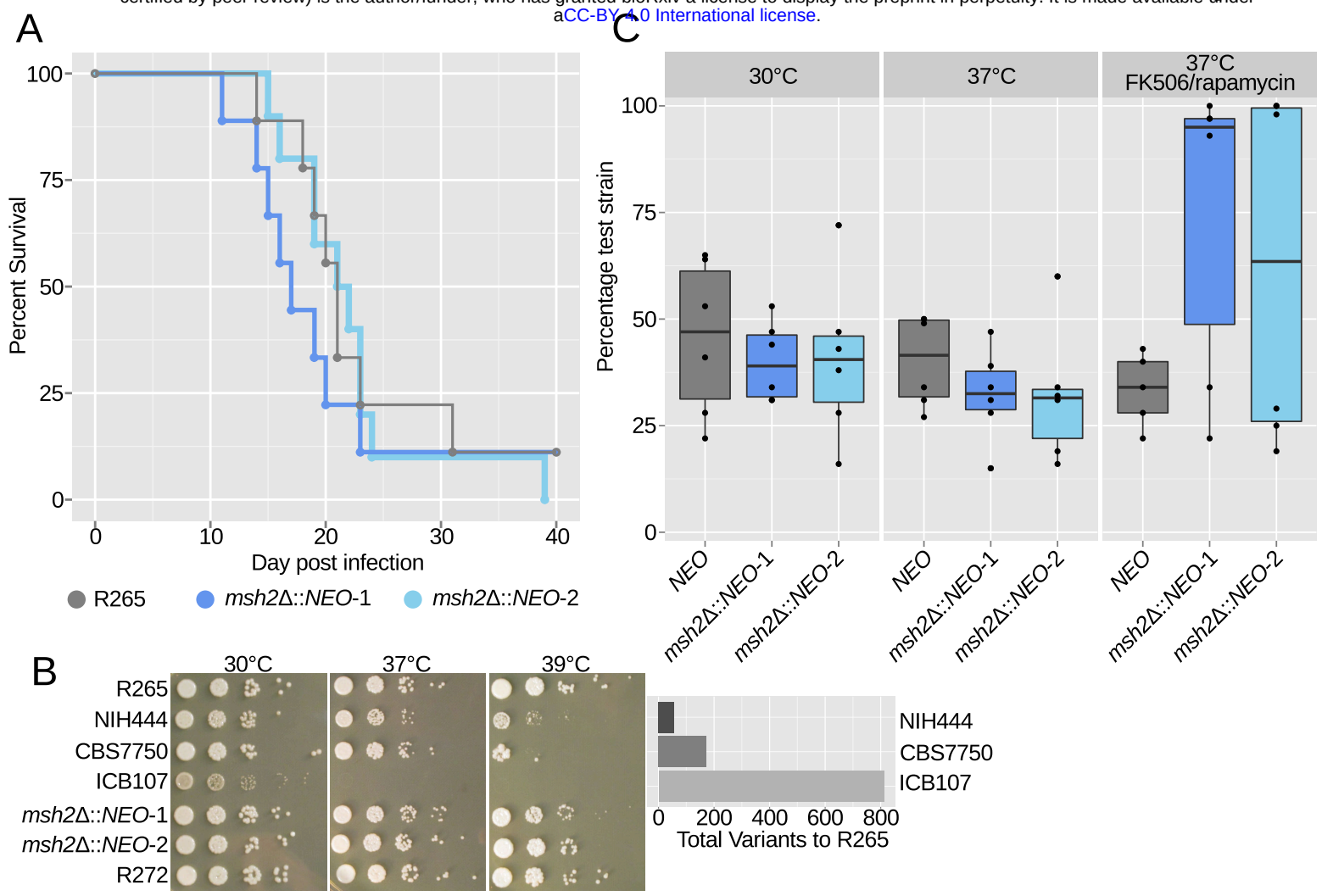
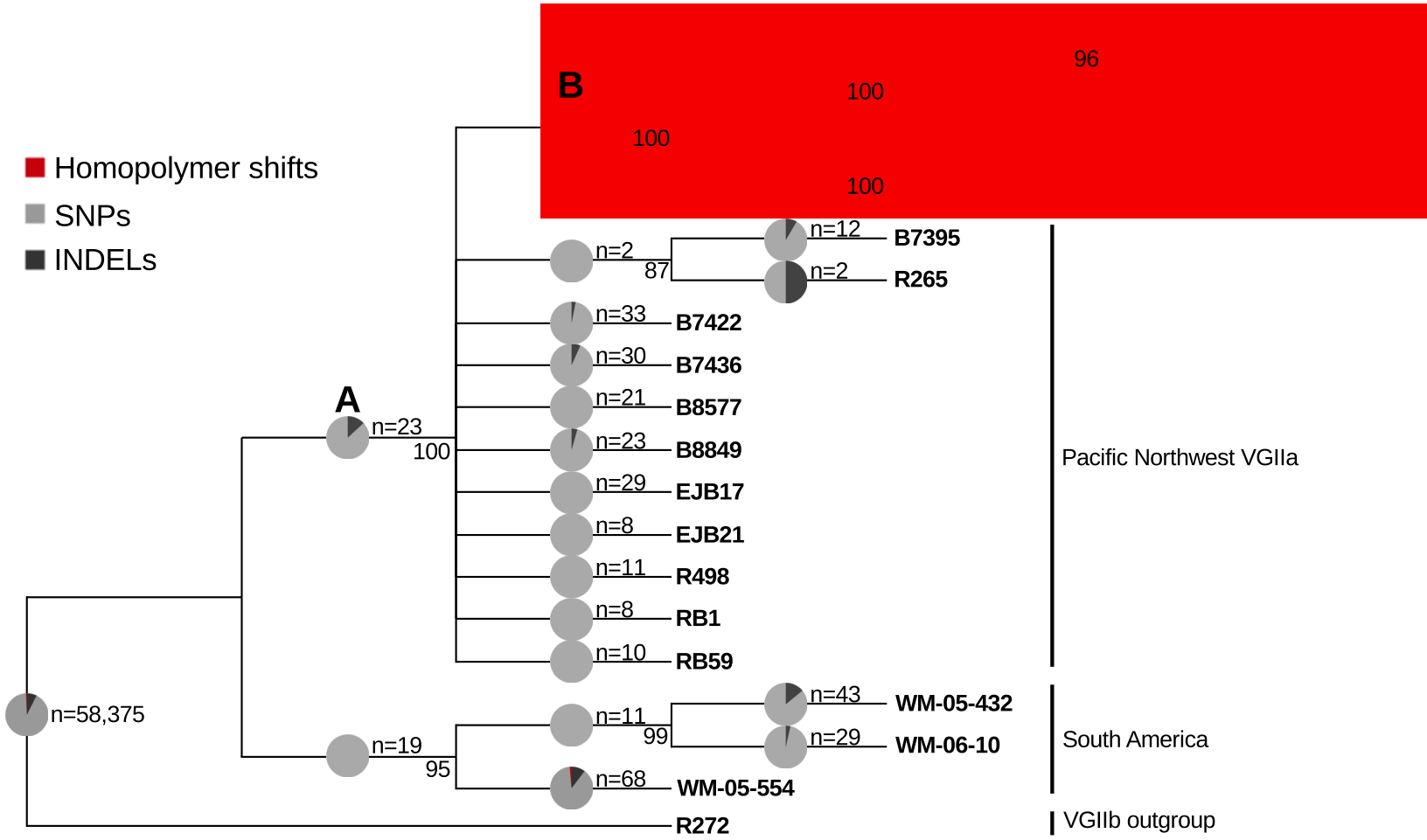


Figure 8



1 Table 1. Mutations shared by VGIIa-like strains relative to VGIIa

Mutation	Location	Predicted Effect	Gene	Predicted Function
SNP	SC2.1:1493731	Missense (Ser->Pro)	CNBG_0538	t-RNA ligase
SNP	SC2.2:591810	Splice site acceptor	CNBG_0776	HAD hydrolase
SNP	SC2.2:965997	Missense (Gly->Asp)	CNBG_0907	Clp protease subunit
SNP	SC2.4:899106	Missense (Leu->Val)	CNBG_1894	RhoGEF
SNP	SC2.6:160627	Missense (Ala->Val)	CNBG_2540	Ubiquitin-like
SNP	SC2.6:410824	Intergenic	-	
SNP	SC2.6:1157157	Intergenic	-	
SNP	SC2.7:290975	Missense (His->Tyr)	CNBG_3034	BRCT domain-containing
SNP	SC2.8:594146	Synonymous	CNBG_3547	snRNP
SNP	SC2.10:94373	Intronic	CNBG_3750	Dcp2
SNP	SC2.13:413395	Missense (Leu->Ser)	CNBG_4876	esterase
SNP	SC2.15:96883	Intergenic	-	
SNP	SC2.16:143098	Missense (Pro->Leu)	CNBG_5437	Protein kinase Nrc2
INDEL	SC2.1:15401-15402	Intronic	CNBG_0005	sugar metabolism
INDEL	SC2.3:240763-240764	Nonsense	CNBG_1661	Msh2
INDEL	SC2.9:241948-241949	Intronic	CNBG_4065	SAM methyltransferase
INDEL	SC2.14:162875-162876	Intergenic	-	
INDEL	SC2.15:299154-299155	Intergenic	-	

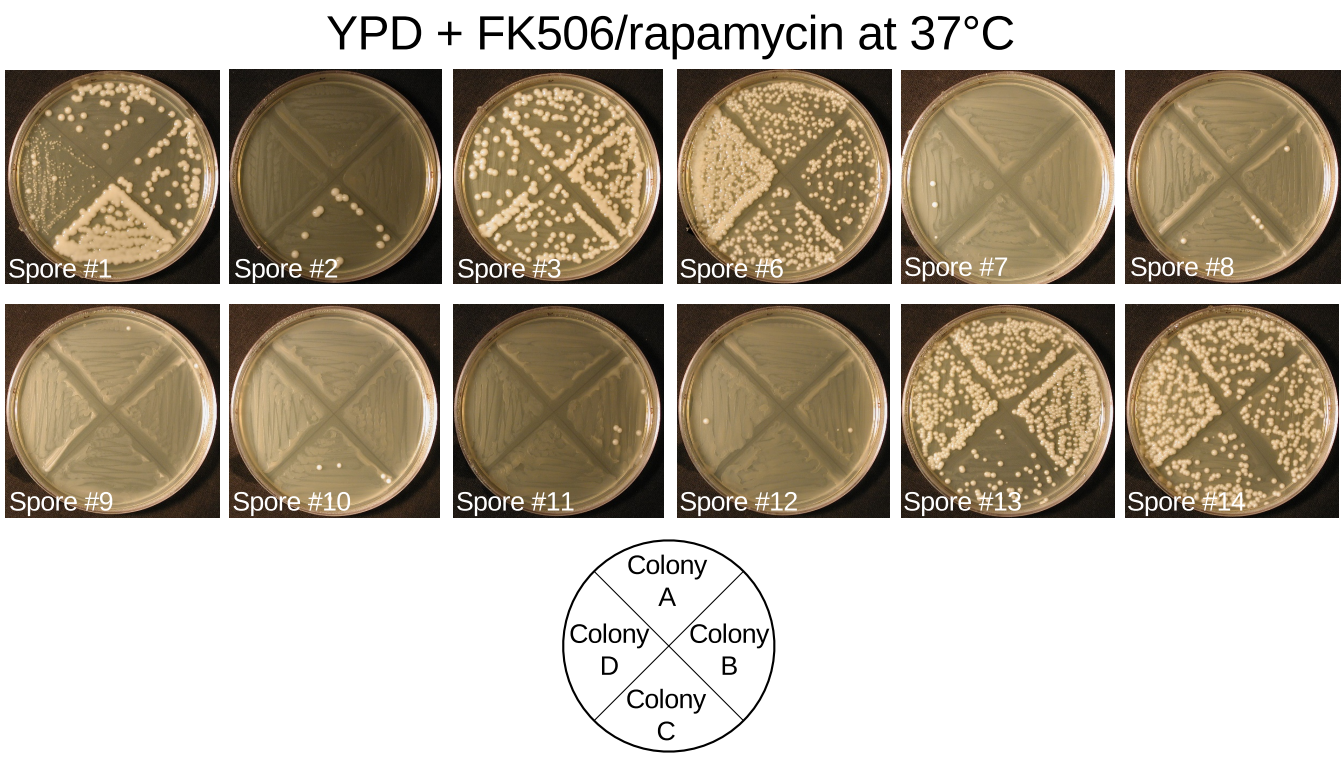
2

3

Table 2. Aneuploidy in NIH444 x R265**a** cross

Spore #	Aneuploid scaffold(s)	*MSH2* allele
1	-	*msh2*del131
2	19,21,26	WT
3	-	*msh2*del131
4	5	*msh2*del131
5	6	*msh2*del131
6	-	*msh2*del131
7	6	WT
8	-	WT
9	6	WT
10	6	WT
11	6	WT
12	-	WT
13	-	*msh2*del131
14	19,21,26	*msh2*del131

Supplemental Figure 1



Supplemental Figure 2