Tracing the evolution and genomic dynamics of mating-type loci in *Cryptococcus* pathogens and closely related species

Marco A. Coelho[1,§,*], Márcia David-Palma[1,§], Seonju Marincowitz[2], Janneke Aylward[2,3], Nam Q. Pham[2], Andrey M. Yurkov[4], Brenda D. Wingfield[2], Michael J. Wingfield[2], Sheng Sun[1], and Joseph Heitman[1*]

[1]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, North Carolina, USA

[2]Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa

[3]Department of Conservation Ecology and Entomology, Stellenbosch University, Stellenbosch, South Africa

[4]Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany

[§]These authors contributed equally as first authors.

*Corresponding authors

Emails: marco.dias.coelho@duke.edu and heitm001@duke.edu

Running title: Mating-type loci evolution in *Cryptococcus* and *Kwoniella*

Keywords: Mating-type locus, *MAT* genes, chromosomal rearrangements, human fungal pathogens.

Abstract

Sexual reproduction in basidiomycete fungi is governed by *MAT* loci (*P/R* and *HD*), which exhibit remarkable evolutionary plasticity, characterized by expansions, rearrangements, and gene losses often associated with mating system transitions. The sister genera *Cryptococcus* and *Kwoniella* provide a powerful framework for studying *MAT* loci evolution owing to their diverse reproductive strategies and distinct architectures, spanning bipolar and tetrapolar systems with either linked or unlinked *MAT* loci. Building on recent large-scale comparative genomic analyses, we generated additional chromosome-level assemblies uncovering distinct evolutionary trajectories shaping *MAT* loci organization. Contrasting with the small-scale expansions and gene acquisitions observed in *Kwoniella*, our analyses revealed independent expansions of the *P/R* locus in tetrapolar *Cryptococcus*, possibly driven by pheromone gene duplications. Notably, these expansions coincided with an enrichment of AT-rich codons and a pronounced GC-content reduction, likely associated with recombination suppression and relaxed codon usage selection. Diverse modes of *MAT* locus linkage were also identified, including three previously unrecognized transitions: one resulting in a pseudobipolar arrangement and two leading to bipolarity. All the three transitions involved translocations. In the pseudobipolar configuration, the *P/R* and *HD* loci remained on the same chromosome but genetically unlinked, whereas the bipolar transitions additionally featured rearrangements that fused the two loci into a nonrecombining region. Mating assays confirmed a sexual cycle in *C. decagattii*, demonstrating its ability to undergo mating and sporulation. Progeny analysis in *K. mangrovensis* revealed substantial ploidy variation and aneuploidy, likely stemming from haploid-diploid mating, yet evidence of recombination and loss of heterozygosity indicates that meiotic exchange occurs despite irregular chromosome segregation. Our findings underscore the importance of continued diversity sampling and provides further evidence for convergent evolution of fused *MAT* loci in basidiomycetes, offering new insights into the genetic and chromosomal changes driving reproductive transitions.

Introduction

Sexual reproduction is a cornerstone of eukaryotic biology, generating genetic variation through recombination and allele segregation, which fuels adaptation in changing environments [1-3]. The Fungal Kingdom is amongst the most ecologically diverse, with species thriving in nearly every environment, and fulfilling roles as decomposers, symbionts, commensals, as well as pathogens of plants, animals, and other fungi [4-6]. Their ecological versatility is mirrored in a wide range of dispersal stages and reproductive strategies, ranging from heterothallism, requiring genetically distinct individuals of opposite mating types, to homothallism, which allows self-fertilization by a single individual as well as outcrossing [7, 8].

In heterothallic basidiomycete fungi, sexual compatibility is typically governed by two mating-type (*MAT*) loci. The pheromone/receptor (*P/R*) locus encodes at least one mating pheromone and one G-protein-coupled pheromone receptor, enabling mate recognition and fusion. The *HD* locus encodes two homeodomain transcription factors (HD1/Sxi1 and HD2/Sxi2) that heterodimerize post-mating to control dikaryotic growth and sexual development [9]. Importantly, successful mating and offspring production can only occur between mating partners carrying different, compatible alleles of both *MAT* loci.

Most basidiomycetes exhibit a tetrapolar breeding system, where the *P/R* and *HD* loci are unlinked, located on separate chromosomes, and may harbor multiple alleles, allowing each meiotic event to generate up to four distinct mating types among haploid progeny [10]. Some lineages, however, have evolved a bipolar system, in which sexual compatibility is controlled by a single *MAT* locus. This transition can occur either through loss of function of one locus in mating-type determination, as observed in *Coprinellus disseminatus* [11] and other mushroom-forming *Agaricomycetes* [12-14], or via the physical linkage and subsequent fusion (i.e., genetic linkage) of the *P/R* and *HD* loci. In the latter case, translocations or chromosome fusion bring the two loci together onto the same chromosome, followed by additional rearrangements that suppress

recombination, resulting in a single, fused, nonrecombining *MAT* region that is often expanded, highly rearranged, and biallelic [15-18]. Additionally, some basidiomycetes exhibit a pseudobipolar configuration, where the two *MAT* loci are on the same chromosome but remain genetically unlinked because they are sufficiently distant to allow recombination, as seen in skin-commensal/pathogenic *Malassezia* species [19, 20]. Beyond heterothallism, several basidiomycetes reproduce without a partner. This strategy is broadly classified as homothallism but encompasses diverse genetic mechanisms, including primary homothallism, where a single individual carries all of the *MAT* alleles necessary for sexual development [21-24], and unisexual reproduction, where mating occurs between cells of the same mating type [25].

Evidence suggests that tetrapolarity is ancestral in basidiomycetes, with bipolarity via *P/R-HD* fusion arising independently multiple times [15, 16, 18, 26, 27]. The bipolar configuration, often observed in fungal species associated with plants or animals as commensals or pathogens, increases mating success among siblings from a single parental pair from 25% (tetrapolar) to 50% (bipolar), thereby facilitating inbreeding [9, 28, 29]. This may be particularly advantageous during host colonization, where spore dispersal and access to unrelated mating partners may be inherently more limited. Bipolar species with fused *MAT* loci are found across the three basidiomycete subphyla, including smut fungi of grasses [15, 30, 31], anther-smut *Microbotryum* infecting *Caryophyllaceae* [17, 27, 32, 33], and two *Tremellomycetes* lineages: *Trichosporonales* [18] and the human-pathogenic *Cryptococcus* species [16, 34, 35].

An important aspect of fused *MAT* loci is their tendency to incorporate additional genes involved in sexual development, thereby influencing key morphological processes (e.g., dikaryotic hyphal growth) and dispersal mechanisms (e.g., spore production) [36]. Additionally, recombination suppression often extends beyond the core mating-type genes, encompassing larger regions of the *MAT*-containing chromosome [37]. While this may preserve beneficial allele combinations, it also contributes to genomic degeneration, including gene loss, accumulation of

transposable elements (TEs), and extensive structural rearrangements. These features share similarities with sex chromosome evolution in other eukaryotes, highlighting potential convergent evolution in sexual reproduction regulation [34, 38].

The genus *Cryptococcus* includes both human-pathogenic species of critical importance and nonpathogenic saprobic members. The pathogenic clade consists of *Cryptococcus neoformans*, *C. deneoformans*, and six species within the *C. gattii* complex [39, 40], with *C. neoformans* ranking first on the World Health Organization's list of critical fungal pathogens [41, 42]. Nonpathogenic species include *C. wingfieldii*, *C. amylolentus*, *C. floricola*, *C. depauperatus*, and *C. luteus* [23, 43-46] along with newly identified lineages [47, 48], many associated with insects, particularly bark beetles [49]. In contrast, its sister genus, *Kwoniella*, consists solely of saprophytic species found in diverse environments, including soil, seawater, plant material, and insect frass [45, 47, 48, 50, 51].

Sexual reproduction is well-characterized in *Cryptococcus* and has been linked to pathogenicity in medically relevant species [52-54]. A hallmark of the *Cryptococcus* sexual cycle is the yeast-to-hypha transition, which occurs after mating between cells of opposite mating types (**a** and α), or during unisexual reproduction [25]. This morphological switch appears to aid in foraging for nutrients and mating partners [55, 56] and in enhancing survival against environmental stressors and microbial predators, thought to be major selective forces in *Cryptococcus* evolution [52, 57-61]. Eventually, the resulting hyphae differentiate into basidia, where diploid nuclei undergo meiosis, producing basidiospores that are more resilient to environmental stresses and significantly smaller than yeast cells. These spores facilitate alveolar deposition upon inhalation and subsequent dissemination within the host [62-64]. In contrast, sexual reproduction in *Kwoniella* has only been documented in *K. mangrovensis* and *K. heveanensis* [51, 65, 66]. These species do not form aerial basidia with spore chains, which likely makes them less efficient at environmental dispersal than *Cryptococcus*.

Recent surveys have identified several isolates that likely represent novel *Cryptococcus* and *Kwoniella* species [40, 47, 48, 67, 68]. However, these discoveries are often based on single strains, limiting direct mating tests and hindering the characterization of reproductive strategies and compatibility systems. In such cases, high-quality genome assemblies offer a powerful alternative, enabling comparative analyses of *MAT* loci to elucidate reproductive systems. While our recent work generated chromosome-level genome assemblies for several *Cryptococcus* and *Kwoniella* species, providing insights into their genomic architecture and evolution [48], this was restricted to single strains per species, precluding in-depth analyses of *MAT* locus variation.

Here, we consider a broader diversity of species than previously available and, where possible, strains of opposite mating types within the same species, allowing for a systematic characterization of *MAT* loci structure, their chromosomal organization, and key evolutionary transitions across the two genera. Our findings support that tetrapolarity is the ancestral state of both *Cryptococcus* and *Kwoniella*, with *P/R-HD* loci fusion (i.e., genetic linkage) evolving independently three times: (i) in the common ancestor of pathogenic *Cryptococcus* species, (ii) in a recently identified nonpathogenic *Cryptococcus* species, and (iii) in a single *Kwoniella* species. Additionally, our analyses revealed that the *P/R* locus itself is highly dynamic, sharing only a core set of four genes across both genera, and has undergone lineage-specific expansions, particularly in *Cryptococcus* where increased pheromone gene copy number appears to have driven substantial rearrangements. Experimental crosses also confirmed a previously unobserved sexual cycle in *C. decagattii* and provided the first direct evidence of recombination in *K. mangrovensis*, revealing meiotic exchange in progeny and non-haploid offspring with potential for self-fertility. Collectively, our findings contribute to the growing body of evidence that fused *MAT* loci have evolved multiple times independently in basidiomycetes, reinforcing the role of convergent evolution in shaping reproductive strategies and ecological adaptation.

Results

Chromosome-level assemblies and updated phylogenetic relationships of *Cryptococcus* and *Kwoniella*

To enable detailed analyses of *MAT* locus structure and evolution, we expanded our previous genomic dataset [48] by incorporating genome sequences for 12 *Cryptococcus* and 7 *Kwoniella* strains, including opposite mating-types where available. Among the *Cryptococcus* strains, four represent three currently undescribed species (*Cryptococcus* sp. 3, sp. 4, and sp. 5; **Fig 1A**), while the remaining eight are opposite mating types of pathogenic *Cryptococcus* species previously sequenced [48]. For *Kwoniella*, three strains belong to the recently described species *K. ovata*, *K. endophytica*, and *K. fici* [47, 69, 70], while the remaining four strains are different mating types of *K. mangrovensis*, *K. europaea*, *K. botswanensis*, and *K. heveanensis* [51, 65, 66]. All but three genomes were assembled using a combination of long-read (PacBio or Oxford Nanopore) and short-read (Illumina) sequencing, yielding chromosome-level genome assemblies. The exceptions, *K. europaea* PYCC6162, *K. botswanensis* CBS12717, and *K. heveanensis* BCC8398 were sequenced solely with Illumina, resulting in more fragmented assemblies. In total, the expanded dataset comprises 27 *Cryptococcus* strains (spanning 17 species) and 22 *Kwoniella* strains (spanning 18 species) (see **S1 Appendix** for additional details).

To establish evolutionary relationships across species and determine the placement of the newly sequenced taxa, we identified 3,086 single-copy orthologous genes shared across all *Cryptococcus* and *Kwoniella* species, as well as three outgroup species (*Tremella mesenterica*, *Saitozyma podzolica*, and *Bullera alba*) following a previously described approach [48]. Phylogenetic analyses employing both Maximum Likelihood (ML) and coalescent-based methods yielded largely congruent results, except for the variable placement of *Cryptococcus* clades B and C relative to clade A (**Fig 1A** and **S1 Fig**). This inconsistency, noted in earlier studies [23, 44, 46] remains unresolved with the current dataset. Despite this, our analyses resolved the phylogenetic

placement of the newly identified *Cryptococcus* species. Among them, *Cryptococcus* sp. 3, currently known from a single strain (CMW60451) isolated from a bark beetle (*Lanurgus* sp.) infesting twigs of the endangered conifer *Widdringtonia cedarbergensis* in the Cederberg Mountains of South Africa, was found to be more closely related to *C. depauperatus*, albeit positioned on a relatively long branch. Additionally, *Cryptococcus* sp. 4 [71] and *Cryptococcus* sp. 5 clustered with the previously identified *Cryptococcus* sp. 6 (OR918) [48, 72], together forming an early-branching clade within the genus. Among *Kwoniella*, *K. ovata* and *K. endophytica* were identified as sister species closely related to *K. dendrophila*, while *K. fici* grouped more closely with *Kwoniella* sp. 4. The new *Cryptococcus* and *Kwoniella* species will be formally described elsewhere.

Sexual reproduction, breeding systems, and *MAT* gene identification and organization across *Cryptococcus* and *Kwoniella*

Sexual cycles have been described for eight *Cryptococcus* species, seven of which are primarily heterothallic (*C. neoformans*, *C. deneoformans*, *C. gattii*, *C. bacillisporus*, *C. deuterogattii*, *C. amylolentus*, *C. floricola*), while one (*C. depauperatus*) is homothallic (**Fig 1A**) [23, 46, 73-77]. In *Kwoniella*, sexual cycles are known for only two species (*K. mangrovensis* and *K. heveanensis*), both heterothallic [44, 51, 65]. To identify sexual cycles in species where they remain unknown, we attempted to induce mating and detect sexual structures in species for which at least two strains of opposite mating types are available. Crosses were performed under conditions known to stimulate mating in *Cryptococcus* [78] and examined for the presence of hyphal growth or other sexual structures, including basidia and basidiospores.

In *C. decagattii*, sexual reproduction was confirmed under laboratory conditions upon crossing strains 7685027 (*MAT*α) and CBS11687 (*MAT***a**), resulting in the formation of sexual structures (**S2A Fig**). In contrast, crosses between the two *C. tetragattii* strains (IND107 and

CBS11718) did not produce observable sexual structures (**S2B Fig**). For *Cryptococcus* sp. 3, sister to *C. depauperatus*, we observed key differences: while *C. depauperatus* is homothallic and exhibits continuous hyphal growth throughout its lifecycle [23], *Cryptococcus* sp. 3 grew mostly as yeast and did not produce sexual structures alone, suggesting it is unlikely to be homothallic (**S2C Fig**). Finally, within clade D species, we tested *Cryptococcus* sp. 4, where two strains of opposite mating types were available. Although some hyphal growth was observed, it was inconsistent, and no discernible sexual structures were detected in the resulting mycelia even after prolonged incubation (**S2D Fig**). Given the absence of confirmed sexual reproduction in most of the species tested, we resorted to genomic data and BLAST-based searches to investigate *MAT* gene content and organization. This approach enabled us to determine the likely breeding system (heterothallic or homothallic) for species without defined sexual cycles and classify *MAT* loci configurations as tetrapolar, pseudobipolar, or bipolar across the two genera. The results of these analyses are summarized in **Fig 1B**.

Expanding on previous studies that identified the tetrapolar-to-bipolar transition in the ancestor of the *Cryptococcus* lineage containing human pathogens (clade A) [34, 77], our analysis confirmed that all species in this clade, including *C. decagattii* and *C. tetragattii*, have a bipolar mating configuration. This analysis resolves gaps stemming from the previous lack of genome assemblies for both mating types in these species. In contrast, most of the other non-pathogenic *Cryptococcus* and *Kwoniella* species appeared to be heterothallic and tetrapolar, consistent with the proposed ancestral state of basidiomycetes [9, 79, 80]. However, three exceptions were identified: *Cryptococcus* sp. 3 (clade C), *K. europaea* (clade E), and *K. fici* (clade H) (**Figs 1B** and **1E**). In *Cryptococcus* sp. 3 and *K. fici*, BLAST searches revealed that the key *MAT* genes (*HD* and the pheromone receptor *STE3*) are close together on the same chromosome, a hallmark of bipolar configurations (**Fig 1E**). In contrast, *K. europaea* exhibited a pseudobipolar configuration, with the *HD* and *STE3* genes residing on the same chromosome but separated by ~8.95 Mb (**Fig 1E**).

These findings prompted us to precisely determine the composition of each *MAT* loci across species and retrace the chromosomal structural changes underlying the configurations found in all bipolar species.

Delineation of the *MAT* loci structure in *Cryptococcus* and *Kwoniella* species

The structure and gene composition of the *MAT* locus has been previously determined for several *Cryptococcus* species within clades A, B, and C [16, 23, 34, 35, 46, 77, 81]. In contrast, analyses of *Kwoniella* species have been limited to *K. mangrovensis* and *K. heveanensis*, relying on fosmid library data in the absence of genome sequencing and focusing on a single mating type, leaving the precise structure and full gene composition of the *MAT* loci unresolved [65, 66]. To overcome these limitations and expand the analysis to newly included species, we conducted synteny analyses complemented by gene genealogies to comprehensively characterize *MAT* loci across *Cryptococcus* and *Kwoniella*.

Our analyses revealed that the *HD* locus in tetrapolar *Cryptococcus* and *Kwoniella* species is a small chromosomal region containing only the *HD1*/*SXI1* and *HD2*/*SXI2* genes, which are divergently transcribed from a shared promoter region (**S3 Fig**). Similar to other basidiomycetes, differences between mating types are primarily confined to the N-terminal regions of the encoded proteins. This conserved two-gene structure suggests strong selection to maintain the *SXI1*-*SXI2* pair as a single inheritable unit. Notably, this conservation persists in *Kwoniella* species, even though the *HD* locus is predominantly located in subtelomeric regions, which are inherently more prone to variability and rearrangements (**S3J–S3O Fig**).

In contrast to the *HD* locus, the *P/R* locus is notably more variable, spans a larger chromosomal region, and exhibits substantial variation across species and between mating types within the same species (**Figs 2** and **S4**). The *P/R* locus is significantly more expanded in tetrapolar *Cryptococcus* species, averaging 91.7 kb compared to 29.3 kb in *Kwoniella* species

(**Figs 2, S4 Fig**, and **S2 Appendix**; $P < 0.0001$, Mann–Whitney U Test). Within *Cryptococcus*, the *P/R* locus ranges from 84.5 kb estimated in *Cryptococcus* sp. 5 (*P/R a2*) to 101.5 kb in *C. floricola* (*P/R a1*), while in *Kwoniella*, it spans from 24.3 kb in *K. newhampshirensis* (*P/R a1*) to 34.4 kb in *K. heveanensis* (*P/R a2*) (**S2 Appendix**).

Evolution of the *P/R* locus in *Kwoniella*

Gene content analysis of the *Kwoniella P/R* locus revealed 9 genes shared across all species (**Figs 1C** and **2C-D,** and **S3 Appendix**). These include a single pheromone gene (*MF*α or *MF***a**) containing the canonical CAAX motif at its C-terminus, a pheromone receptor (*STE3*α or *STE3***a**), and two additional genes (*STE20* and *STE12*) with established roles in mating-related functions and known to reside within *MAT* in *Cryptococcus* pathogens [34, 82, 83]. Other genes identified include *DBP10* and *REI1*, both implicated in ribosome biogenesis in *Saccharomyces cerevisiae* [84, 85]; *ATG7*, an autophagy-related gene [86]; a C2 domain-containing protein (*C2prot*), whose ortholog in *S. pombe* (Ync13) coordinates exocytosis, endocytosis, and cell-wall integrity during cytokinesis [87]; and *NDC80*, encoding a component of the NDC80 kinetochore complex, essential for chromosome segregation and spindle checkpoint activity [88]. In *K. heveanensis* and other species of clade G, the *P/R* locus has expanded to include two additional genes: *BPS3*, encoding a protein of unknown function with BTP/PZ and CAP-Gly domains (also present in the *MAT* locus of *Cryptococcus* pathogens), and *UPB6,* a predicted deubiquitinating enzyme associated with the 26S proteasome [89]. Of these genes, three (*DBP10*, *REI1* and *NDC80*) are predicted to be essential in *S. cerevisiae* and in *C. neoformans*, based on evidence from previous studies [84, 85, 88, 90, 91].

Within a species, the *P/R a2* allele is generally larger than the *a1* allele (**Figs 2C-D**, **S4 Fig** and **S2 Appendix**), primarily due to the presence of an additional gene of unknown function. This size difference is also accompanied by truncation of the *STE12* gene within the *a2* allele, with

fragments of *STE12* flanking the *STE3* gene on both sides in several species, such as in *K. mangrovensis* and *K. pini* (**Figs 2C** and **S4**). Similarly, fragments of the 3' end of *NDC80* and *BSP3* were detected at the right edge of the *P/R a2* allele in certain clade E/F and clade G species, respectively, in addition to apparently intact copies of these genes (**Figs 2C** and **S4**). These fragments seem to be remnants of structural rearrangements, potentially associated with duplication tracks at the inversion breakpoints. To elucidate the origin of these truncations and structural differences, we reconstructed the likely rearrangements leading to the extant *P/R* configuration in *Kwoniella*. In clades E/F, three distinct inversions are predicted to have shaped the locus: the first inversion involved *STE3* and *STE12*, truncating *STE12* in the *a2* allele; the second relocated *DBP10* and *REI1* from the edge of the locus to its center; and the third moved *NDC80* from the edge to the middle of the locus (**S5A Fig**). The first two inversions appear to be conserved in clade G, suggesting they are likely ancestral in *Kwoniella*. In contrast, two additional rearrangements relocating *BSP3-NDC80* and *UBP6* to the center of the locus (**S5B Fig**) seem to be clade G-specific events. Collectively, this data suggests that the *P/R a2* allele has experienced more modifications compared to the *P/R a1* allele.

The *P/R* locus in *K. ovata* and *K. endophytica* present further complexity in that it contains *STE3***a** and *MF***a** alleles alongside truncated versions of *STE3*α, with *K. ovata* additionally harboring a truncated *MF*α gene (**Fig 3A**). This unique configuration suggests that the extant *P/R* locus in these species may have arisen through intra-*P/R* recombination, resulting in a mosaic assembly of loci elements derived from both mating types. Supporting this hypothesis, synteny analysis showed that the left side of the *P/R* locus resembles that of other *P/R a2* strains, whereas the right side aligns more closely with the *P/R a1* allele (**Fig 3A**). Gene genealogies further revealed trans-specific polymorphism in several genes. While deep trans-specific polymorphism has been well-documented for *STE3*, *STE12*, and *MF* genes, which were ancestrally integrated into the *P/R* locus across *Cryptococcus* and *Kwoniella* [23, 46, 66, 77], a shallower trans-specific

polymorphism pattern was observed for *NDC80*, *BSP3*, *ATG7*, and *UBP6*. Specifically, *NDC80* exhibited trans-specific polymorphism within clades E/F and within clade G (**Fig 3B**), while *ATG7*, *BSP3*, and *UBP6* displayed this pattern exclusively within clade G (**S5C Fig**). Notably, the *NDC80* allele at the right edge of the *P/R* locus in *K. ovata* and *K. endophytica* clustered more closely with *a1* alleles from other species (**Fig 3B**), and a similar pattern was found in *K. dendrophila*, together pointing to a shared evolutionary history where the extant *P/R* locus resulted from past intra-*P/R* recombination. While this event may have initially produced strains with intercompatible pheromone and receptor pairs, potentially enabling self-filamentous growth, subsequent evolutionary pressures appear to have led to the differential loss of *STE3*α and *MF*α mating-type alleles, ultimately eliminating self-filamentation. Consistent with this, solo cultivation of these species in V8 pH 5 media did not induce hyphal growth, indicating that extant species are no longer self-fertile and instead retain relics of the ancestral recombination event. Collectively, these findings unveil a complex evolutionary history of the *P/R* locus in *Kwoniella*, marked by dynamic structural rearrangements, recombination, and lineage-specific expansions.

Evolution and expansion of the *P/R* locus in *Cryptococcus*

Compared to *Kwoniella*, the *P/R* locus in tetrapolar *Cryptococcus* species has undergone significant expansion, averaging 95.4 kb in clade B and 86.9 kb in clade D, a difference that is not statistically significant (P = 0.063, Mann–Whitney U Test; **Figs 2** and **S4**; **S2 Appendix**). It also exhibits substantially more complex rearrangements between mating types, precluding the precise reconstruction of the evolutionary events leading to the extant configurations (**Figs 2**, **S4B**, and **S4E**). Another striking feature is the consistently lower GC content of the *P/R* locus relative to genome-wide averages (**Figs 2**, **S4A** and **S4D**; **S4 Appendix**). For instance, in *C. amylolentus* CBS6039 (clade B), the GC content of the *P/R* locus is 51.24%, compared to a genome-wide GC content of 53.36% (*t*-test: t = 65.03, P < 0.0001; permutation test: P = 0.0288), with Z-scores

ranging from –1.38 to –4.00 across clade B strains. This trend is even more pronounced in clade D species, where the GC content difference can approach 10%. For example, in *Cryptococcus* sp. DSM108351, the *P/R* locus GC content is 49.17%, compared to a genome-wide average of 59.24% (*t*-test: $t = 199.19$, $P < 0.0001$; permutation test: $P = 0.002$), with Z-scores ranging from –6.25 to –7.22 across clade D strains. Further analysis of codon composition revealed a significant enrichment of AT-rich codons in the *P/R* locus compared to genome-wide sequences (**S4 Appendix**). For instance, in *Cryptococcus* sp. DSM108351, the *P/R* locus contains 13,429 AT-rich codons versus 13,390 GC-rich codons, whereas genome-wide sequences contain 2,877,796 AT-rich codons versus 3,475,722 GC-rich codons ($\chi^2 = 245.86$, $P < 0.0001$). Similar patterns were observed in clade C species, such as *C. amylolentus* CBS6039 ($\chi^2 = 21.22$, $P < 0.0001$; **S4 Appendix**). These findings suggest that the lower GC content of the *P/R* locus is likely driven by a bias towards AT-rich codons, potentially resulting from mutation accumulation associated with reduced or suppressed recombination in this region.

Despite these shared features, the *P/R* locus gene content shows striking variation between *Cryptococcus* species of clades B and D (**Fig 1C**). While both clades share a core set of 7 genes (*MF* with two to four copies, *STE3*, *STE12*, *BSP3*, *MYO2*, *STE20*, and *RPL39*; **S3 Appendix**), several genes within the *P/R* locus in clade B species are instead located in the right flanking region of the locus in clade D species (**Figs 2A-B**, **S4B**, and **S4E**) and similarly positioned outside, but proximal to, the *P/R* locus in *Kwoniella* species (**Figs 2C-D**, **S4H**, **S4K**, and **S4N**). Additionally, two genes (*ATG7* and *C2prot*) located within the *Kwoniella P/R* locus are also found within the *P/R* locus in two of the four clade D species, but absent from the locus of other tetrapolar *Cryptococcus* clades. Interestingly, although only 7 genes are shared between clades B and D, the *P/R* loci are similar in overall size. This similarity arises because clade D species harbor up to 15 genes (**Figs 1C**, **2B**, and **S4E**) that are absent from the *P/R* locus in other *Cryptococcus* or *Kwoniella* clades. Extending these comparisons across all tetrapolar *Cryptococcus* and *Kwoniella*

clades revealed that only four genes (*MF*, *STE3*, *STE20*, and *STE12*) are conserved within the *P/R* loci, likely reflecting the ancestral configuration (**Figs 1C-D**, **S3 Appendix**). These findings suggest that the ancestral *P/R* locus consisted of a very small set of genes and independently expanded in the different lineages by incorporating distinct gene sets through lineage-specific rearrangements.

A final notable distinction of the *P/R* locus in *Cryptococcus* is the presence of multiple pheromone genes per mating type [34, 46]. Unlike *Kwoniella*, which retains a single pheromone gene per mating type, all heterothallic *Cryptococcus* species, including the newly identified clade D species, exhibit multiple copies, with up to four pheromone genes per mating type. These pheromone genes are often arranged in divergent orientations (**Figs 2A-B**, **S4B**, and **S4E**), a configuration that may act as inverted repeats. Such an arrangement facilitates the formation of inversion loops during recombination, driving both historical and potentially ongoing structural rearrangements, and the expansion of the *P/R* locus in *Cryptococcus*.

MAT locus evolution and structural dynamics in *Cryptococcus* pathogens

The fusion of the *P/R* and *HD* loci into a single nonrecombining *MAT* locus represents a pivotal transition in the evolution of breeding systems in basidiomycetes, marking a shift from tetrapolar to bipolar systems. This transition has been extensively documented in the *Cryptococcus* lineage that includes all of the human pathogens (clade A). Building on previous work, we expanded the analysis of the *MAT* locus to all recognized species in this clade (**Figs 1**, **4**, **S6**, and **S7**; **S1 Appendix**). For *C. deneoformans*, the analysis included the well-studied congenic strains JEC21α and JEC20**a**, their progenitor strains NIH12α and NIH433**a**, and another *MAT***a** strain (NIH430) (**S6 Fig** and **S1 Appendix**). For *C. neoformans*, we examined multiple strains of both mating types across its four main lineages (VNI, VNII, VNBI, and VNBII), including eight newly sequenced genomes assembled telomere-to-telomere (T2T) (**Fig 4A** and **S1**

Appendix). This comprehensive dataset offered unprecedented resolution for investigating intra-specific structural dynamics of the *MAT* locus.

Synteny analysis across species revealed that the *MAT***a** allele is generally longer than the *MAT*α allele in both *C. neoformans* ($P < 0.0001$, Mann–Whitney U Test) and the *C. gattii* species complex ($P = 0.04142$, Mann–Whitney U Test; **Fig 4C and S2 Appendix**). A similar trend is observed in *C. deneoformans*, although limited sampling precluded statistical support. The *MAT***a** allele also exhibits greater structural diversity than the *MAT*α allele among closely related species, such as in the *C. gattii* complex (**S7 Fig**). This pattern extends to *C. neoformans*, where pronounced rearrangements, predominantly inversions, have occurred even among *MAT***a** strains of the same VN lineage (e.g., VNI strains Bt130 vs. IUM96-2828, or strain Bt206 compared to other VNBII strains; **Fig 4D**). Notably, the breakpoints for these rearrangements are often associated with the pheromone genes, supporting the hypothesis that multiple pheromone genes may act as mediators of structural changes. In contrast, the *MAT*α allele structure in both *C. neoformans* and the *C. gattii* species complex is more conserved, with only limited inversions involving the two pairs of pheromone genes flanking the *PRT1* and *ZNF1* genes (**Figs 4D** and **S7**). Interestingly, a comparison of the entire *MAT* chromosome across *C. neoformans* strains showed that most of the intraspecific structural variation is confined to the *MAT* locus, with no major structural changes detected outside this region, apart from differences within the centromeric regions, which are inherently more dynamic (**Fig 4B**).

Beyond these structural changes, species-specific gene losses (e.g., *NCP1* and *NCP2*, which are absent in most species) and TEs further drove structural differences in *MAT* allele configurations (**S7 Fig**). TE activity, in particular, may play a significant role as the *C. deneoformans MAT* locus has been estimated to contain over five times more TEs than the rest of the genome, excluding centromeric regions [16, 34, 92]. In addition to previously identified TE remnants, our analysis revealed novel transposon insertions. Specifically, two distinct KDZ

transposons, recently characterized in *C. neoformans* as large mobile elements (~11 kb) with terminal inverted repeats (TIRs) and target site duplications (TSDs) [93], were detected within the *MAT* locus in 5 of the 23 analyzed strains. *KDZx* was identified in strains 125.91, C45, PMHc1023, and T4, while *KDZ1* was found in strain Ze90-1 (**Fig 4D)**. Together, these findings underscore the dynamic evolution of the *MAT* locus in *Cryptococcus* pathogens, shaped by an interplay of structural rearrangements, gene loss, and TE activity.

In *C. neoformans*, the VNI lineage is hypothesized to have initially diverged with only the *MAT*α allele, potentially due to a bottleneck associated with a small founding population, and later acquired the *MAT***a** allele via introgression through mating between a VNBI *MAT***a** strain and VNI *MAT*α strain [94]. To date, only five VNI *MAT***a** strains have been identified [94-99], two of which were fully sequenced (T2T) and annotated in this study (125.91 and Bt130), while a third strain (IUM96-2828) was sequenced in a previous study [100] and annotated as part of this work. However, synteny analysis revealed significant structural divergence between the *MAT***a** alleles of VNBI strains (Ftc555-1 and Bt63) and these VNI strains (**Fig 4D**), raising questions about the proposed introgression hypothesis.

To address this, we reanalyzed SNP data leveraging newly generated reference genome sequences. Genome-wide SNP analysis, using the VNBI *MAT***a** strain Ftc555-1 as the reference, revealed markedly lower SNP densities within the *MAT***a** locus in VNI strains compared to VNBII strains (**Figs 5A** and **5B**). In contrast, across the rest of the genome, VNI *MAT***a** strains exhibited higher divergence from the VNBI reference, while VNBII strains showed lower SNP densities (**Fig 5A** and **5B**). Interestingly, the typically high SNP densities between VNI and VNBI strains are restored immediately downstream of the *MAT* locus and approximately five genes upstream, consistent with previous reports of recombination hotspots flanking the *MAT* locus [101]. Phylogenetic analyses further supported these observations; VNBI strains clustered more closely with VNBII strains in trees inferred from genome-wide and chr. 5-specific SNPs (**Figs 5C** and **5D**),

whereas those constructed with *MAT*-specific SNPs showed closer clustering of VNI and VNBI *MAT***a** strains, supporting a more recent shared ancestry for the *MAT***a** allele (**Fig 5E**). Together, these findings support and refine the introgression model illustrated in **Fig 5F**, suggesting that lineage-specific rearrangements occurred following the introgression event to account for the observed structural differences.

Evidence for independent fusion of *P/R* and *HD* loci in *Cryptococcus* sp. 3

Our initial screening of *MAT* gene content revealed that the *P/R* and *HD* loci-associated genes are adjacent on the same chromosome in *Cryptococcus* sp. 3, indicating a fusion of these two regions (**Fig 1E**). While only a single strain (CMW60451) is currently available for this species, limiting precise determination of the *MAT* locus length, the minimum estimated size is ~90 kb based on the distance between the two most distal pheromone genes (**Fig 6C**). If the boundaries are extended to include additional genes known to be part of the *MAT* locus in pathogenic *Cryptococcus* species, the *MAT* locus could span ~120 kb, from *RPL22* on the left to *STE12* on the right (**Fig 6C**). Under this broader definition, a total of 15 genes are shared between *Cryptococcus* sp. 3 and clade A species (**Figs 1C** and **6C,** and **S3 Appendix**), including for example four *MF*α pheromone gene copies and three genes typically associated with the *HD* locus (*RPL22*, *CAP1*, and *SPO14*). Additionally, two distinct regions within the *MAT* locus contain TEs: one harboring a complete KDZ-like transposon and another with only remnants of this element (**Fig 6C**).

Further analysis uncovered an unusual combination of *MAT* alleles. Unlike the bipolar *Cryptococcus* pathogens, where *MAT*α strains carry *HD1* (*SXI1*) linked to *STE3*α and *MAT***a** strains have *HD2* (*SXI2*) linked to *STE3***a**, *Cryptococcus* sp. 3 retains only *SXI2*, which is instead linked to *STE3*α. Although this distinct configuration suggested an independent fusion event at the origin of the *MAT* locus in *Cryptococcus* sp. 3, the uncertain phylogenetic placement of clade C

relative to clade A left open the possibility that the two clades share a more recent common ancestor. In this scenario, the fusion of the *P/R* and *HD* loci could have occurred in their shared ancestor rather than independently in both lineages.

To test this hypothesis, we analyzed the phylogenetic clustering patterns of genes within the predicted *MAT* region shared with pathogenic *Cryptococcus* species. Genes ancestrally recruited to the *P/R* locus, such as *STE3, STE12,* and *MYO2*, were excluded from the analysis because they exhibit deeper trans-specific polymorphism predating the divergence of *Cryptococcus* species [23, 46, 66, 77], making them less informative. Instead, we focused on the genealogies of four genes (*BSP3, PRT1, SPO14,* and *CAP1*), which are part of the *MAT* locus in pathogenic species and display mating-type-specific clustering within them [34]. The analysis revealed that *Cryptococcus* sp. 3 sequences do not group with either the **a** or α allele-specific clusters from pathogenic species; rather, their sequences are distinct and positioned outside these clusters, which is more consistent with an independent fusion event than one inherited from a common ancestor (**S8 Fig**).

Next, to investigate the chromosomal rearrangements underlying the linkage of *P/R* and *HD* loci in *Cryptococcus* sp. 3, synteny blocks were reconstructed with SynChro [102], employing *K. shandongensis* as reference. The choice of *K. shandongensis* was guided by several factors: (i) it is tetrapolar, with the *P/R* and *HD* loci located on separate chromosomes, representing the ancestral configuration (**Figs 1**, **S3N**, and **S4N**); (ii) *Kwoniella* species exhibit significantly fewer interchromosomal rearrangements compared to *Cryptococcus* species, likely due to their simpler and smaller centromeres, as well as a lower TE density, given that a higher TE density is a known factor promoting genomic rearrangements in *Cryptococcus* [48, 103-105]; and (iii) it retains a 14-chromosome karyotype, which is the predicted ancestral state for both genera [48] (**Figs 6** and **S9A**).

The nuclear genome of *Cryptococcus* sp. 3, fully assembled into 13 complete chromosomes (**S1 Appendix**), represents a reduction from the ancestral 14-chromosome karyotype. This reduction raised the possibility that the observed linkage of the *P/R* and *HD* loci could be a result of this karyotype transition. Specifically, the fusion of chromosomes containing the *P/R* and *HD* loci during chromosomal rearrangement could account for their linked configuration in this species. To test this hypothesis, we examined which centromere was lost during the karyotype transition. Synteny analysis and centromere mapping traced the missing centromere to a region on chr. 3, corresponding to the *CEN4* region in *C. neoformans* H99 (**S9B Fig**). This region no longer contains the LTR-rich sequences typically associated with *Cryptococcus* centromeres (**S9B Fig**) [35, 48, 106, 107]. Additional comparisons with *K. shandongensis* suggest that this centromere loss possibly resulted from an intercentromeric recombination event leading to the inactivation and eventual loss of one of the resulting centromeres (**S9B Fig**). Because the *MAT* locus in *Cryptococcus* sp. 3 resides on chr. 11 (**Fig 6A**), which is unrelated to the chromosome affected by the centromere loss, this indicates that the physical linkage of the *P/R* and *HD* loci (i.e., colocation on the same chromosome) in *Cryptococcus* sp. 3 arose through a separate mechanism, independent of the reduction from 14 to 13 chromosomes.

To further investigate the mechanism underlying the tetrapolar-to-bipolar transition in *Cryptococcus* sp. 3, we compared its *MAT*-containing chromosome to chromosomes harboring the *P/R* and *HD* loci in *K. shandongensis*. Synteny analysis revealed significant alignment between chr. 11 of *Cryptococcus* sp. 3 and large regions of chrs. 12 and 5 in *K. shandongensis* carrying the *P/R* and *HD* loci, respectively (**Fig 7A**). Additionally, a small region derived from the *P/R* chromosome is embedded towards the left end of chr. 11 within a region otherwise predicted to have originated from the *HD* chromosome. This suggests that the physical linkage of the *P/R* and *HD* loci in *Cryptococcus* sp. 3 likely began with a chromosomal translocation that brought the two

loci onto the same chromosome, initially separated by a considerable distance. A subsequent inversion then relocated the *HD* locus closer to the *P/R* locus, resulting in their fusion (events 1-3 in **Fig 7A**). Notably, the centromere flanking regions of chr. 11 in *Cryptococcus* sp. 3 show considerable synteny to the *CEN12*-flanking region in the *K. shandongensis P/R* chromosome. This suggests that the two centromeres share a similar evolutionary origin and that the translocation event did not directly involve centromeric recombination but rather occurred elsewhere in the chromosomes (**Fig 7A**).

Revisiting the model of *P/R* and *HD* loci linkage in *Cryptococcus* pathogens

Based on these findings, which reflect a broader genomic analysis and species sampling than any previous study, we revisited a previously proposed model for the physical linkage and subsequent fusion of *P/R* and *HD* loci in *Cryptococcus* pathogens. This model, originally formulated from comparisons between the tetrapolar species *C. amylolentus* (clade B) and the bipolar pathogen *C. neoformans* (clade A), posited that the physical linkage was initiated by a chromosomal translocation mediated by ectopic recombination between repetitive elements within centromeric regions of *P/R*- and *HD*-containing chromosomes. This translocation would place the two *MAT* loci on the same chromosome, initially separated by the centromere. Subsequent rearrangements, such as inversions or transpositions, were hypothesized to bring the two loci into closer proximity, ultimately leading to their fusion and establishing the contiguous *MAT* locus characteristic of *Cryptococcus* pathogens [35, 36].

To reassess this hypothesis, we expanded our analysis to include multiple *Cryptococcus* species spanning the four clades, focusing on synteny and chromosomal architecture. In *C. neoformans*, the *MAT* locus resides on chr. 5. Our analysis revealed that its centromere (*CEN5*) shares conserved synteny on both flanking regions with *CEN12* of *K. shandongensis*, as well as with *CEN11* of *Cryptococcus* sp. 3 (clade C) and *CEN9* of *Cryptococcus* sp. DSM108351 (clade

D) (**Fig 6B**). These conserved synteny patterns indicate a shared evolutionary origin for these centromeres, strongly suggesting that they were not directly involved in the translocation that linked the *P/R* and *HD* loci in *C. neoformans*. In contrast, analysis of the *P/R* chromosome (chr. 10) in *C. amylolentus* (clade B) reveals that its centromere-flanking regions align with two distinct chromosomes, indicative of intercentromeric recombination and pointing to a derived rather than ancestral centromere structure. Comparison between *C. neoformans* and *K. shandongensis* further suggests a model in which the physical linkage and subsequent fusion of the *P/R* and *HD* loci in *C. neoformans* arose through a mechanism similar to that proposed for *Cryptococcus* sp. 3, involving a chromosomal translocation outside the centromeric region, followed by an inversion event that brought the two loci into closer proximity (**Fig 7B**). In *C. neoformans*, however, the inversion seems to have relocated the *P/R* locus closer to the *HD* locus near the left chromosomal end (events 3 in **Fig 7B**). These broader analyses refine the earlier model and suggest that the physical linkage of *P/R* and *HD* loci in *Cryptococcus* species likely arose through alternative mechanisms, independent of centromeric recombination.

Tetrapolar to bipolar and pseudobipolar transitions in *Kwoniella*

We recently reported that karyotype reduction within *Kwoniella* occurred both progressively and independently over time, often accompanied by the formation of “giant” chromosomes (up to 18 Mb) through repeated chromosome fusion events [48]. While such events could potentially lead to physical linkage of *P/R* and *HD* loci, only two *Kwoniella* species harbor both loci on the same chromosome (**Fig 1B**).

The first instance is in *K. europaea* (clade E), where the *HD* and *P/R* loci are located on the same chromosome but remain far apart (~8.95 Mb) (**Figs 1E** and **S4G**), representing a pseudobipolar configuration. This transition appears to have resulted from a previously documented chromosomal translocation specific to *K. europaea* [48]. In this event, the progenitor

552 “giant” chr. 1 of clade E species underwent a translocation with chr. 2, giving rise to the extant
553 pseudobipolar arrangement. Given their physical distance, the two loci are still expected to
554 recombine.

555 The second instance is in *K. fici*, where the *HD* locus is located less than 13 kb away from
556 the *P/R* locus gene cluster, both localized near the terminal region of chr. 1 (**Fig 8D**). This
557 chromosome in *K. fici* is also a “giant” chromosome, likely formed independently through similar
558 chromosomal fusion events as those documented for *Kwoniella* clades E and G species [48].
559 Interestingly, following the divergence from *Kwoniella* clade I (*K. shandongensis/K.*
560 *newhampshirensis*), an ancestral fusion of *K. shandongensis* chrs. 11 and 12 resulted in the
561 emergence of the *P/R* chromosome, observed as a single chromosome in most *Kwoniella* species
562 (exemplified by chr. 2 in *Kwoniella* sp. DSM27419; **Fig 8A**) [48]. In *K. fici*, however, this ancestral
563 fusion product has fragmented further; a portion now resides at the terminal region of *K. fici* chr.
564 1, while another segment is located on chr. 4 (**Fig 8C**). Similarly, the majority of the *K.*
565 *shandongensis HD* (chr. 5), aligns with the opposite end of *K. fici* chr. 1, although a small portion
566 is present on chr. 4, adjacent to the *P/R*-derived region. This pattern suggests that translocations
567 likely contributed to the genesis of the *P/R* and *HD* loci fusion in *K. fici*, with blocks derived from
568 the ancestral *P/R* and *HD* chromosomes brought together during these events. Notably, only a
569 few key *HD*-associated genes (*SXI1*, *SXI2*, and *RPL22*) are positioned near the *P/R* locus gene
570 cluster at the right end of chr. 1 (**Fig 8D)**. This apparently bipolar arrangement retains both *HD*
571 genes (*SXI1* and *SXI2*), contrasting with the presence of only one *HD* gene in opposite mating
572 types of bipolar *Cryptococcus* species. Together, these observations suggest intricate
573 rearrangements leading to this unique derived configuration in *K. fici*.

574

575

576

Analysis of sexual reproduction in *Kwoniella mangrovensis*

Kwoniella mangrovensis (clade G) is known for its ability to reproduce sexually in the lab [51, 65] and characterized by having only three chromosomes. To better understand its reproductive biology, we analyzed the outcomes of sexual reproduction, focusing on recombination, ploidy variation, and potential self-fertility in the progeny.

The two originally *K. mangrovensis* sequenced strains, CBS8507 (*P/R a1*, *HD b1*) and CBS8886 (*P/R a2*, *HD b1*), were found to share identical *HD b1* alleles, making them incompatible for mating (**Fig 9A**). However, Guerreiro et al. [65] previously characterized a third strain, CBS10435, as haploid and compatible with CBS8507. We Illumina-sequenced CBS10435 and confirmed that this strain carries the *P/R a2* allele and a distinct *HD b2* allele, rendering it compatible with CBS8507 (**Fig 9A**). Next, we performed crosses between CBS8507 and CBS10435, which were able to mate on both CMA and V8 pH 5 media. These crosses produced basidia with characteristic globose and lageniform morphologies, consistent with prior descriptions of the species (**Fig 9B**) [51]. Progeny were recovered by dissecting eight clusters of cells germinating from basidia along the hyphae (see **S5 Appendix** for details). A total of 51 F1 progeny were obtained from mating patches from both media types. To ensure that we avoided analyzing duplicated genotypes potentially resulting from asexual clonal propagation, we screened the progeny for evidence of recombination applying a PCR-RFLP approach with two markers per chromosome, positioned on either side of the centromere (**S5 Appendix**). Putative recombination events (allele exchange) as well as heterozygosity at several loci were detected, and nine F1 progeny with distinct genotypes were selected for whole-genome sequencing. SNPs identified between the parental strains served as markers to infer inheritance patterns associated with recombination and loss of heterozygosity (LOH). Additionally, we combined read coverage analysis with FACS to assess ploidy in these isolates.

Our results showed that the majority of F1 progeny were non-haploid (**Figs 9C-D**, **S10, and S11**). Genotyping and FACS results were largely concordant; however, some discrepancies (e.g., MP11 and MP18) suggested variability within cell populations, possibly stemming from intrinsic genomic instability. Indeed, mapping of reads to a combined reference genome (CBS8507 and CBS10435) revealed uneven coverage for certain chromosomes. For example, chr. 2 of CBS10435 exhibited low coverage in progeny MP59 (**Fig 9D**). Although CBS8507 and CBS10435 were previously reported as haploid [65], our FACS analysis revealed that CBS10435 predominantly consists of diploid cells, with a subpopulation of haploids (**Figs 9C** and **S11**). This difference in ploidy between the parental strains likely explains the observed aneuploidy in the progeny.

Analysis of the *MAT* gene content across the F1 progeny showed that some carried compatible *P/R* and *HD* alleles (*a1a2b1b2*), potentially enabling self-fertility (**Fig 9C**). To investigate this, we tested the F1 progeny under mating conditions and compared their phenotypes to a cross of the parental strains (**Figs 9E-F**, and **S12**). On CMA and V8 pH 5 media, progeny with fully compatible *MAT* alleles (MP16, MP18, MP54) produced hyphae and basidia similar to the parental cross. Some hyphal growth was also observed in solo culture of strains MP49, MP51 and MP59, exclusively on V8 pH 5, but the filaments were shorter and had no structures resembling basidia (**S12 Fig**). These isolates showed a phenotype similar to previously reported self-filamentous *K. mangrovensis* strains, which are characterized by highly branched, irregular hyphae that lack visible clamp cells [65]. Taken together, our analyses reveal that sexual reproduction in *K. mangrovensis* involves recombination and can generate non-haploid progeny with the potential for self-fertility.

Discussion

Our broad comparative genomic analysis of *Cryptococcus* and *Kwoniella MAT* loci provided new insights into their evolution and structural dynamics. These include distinct expansions of the *P/R* locus and additional independent events of *HD-P/R* fusion, which are summarized in **Fig 10** as a visual roadmap of key transitions.

The archetypal basidiomycete *MAT* loci are thought to consist of a compact *P/R* region, with a single pheromone receptor and one pheromone gene, and an unlinked *HD* region containing two divergently transcribed homeodomain transcription factors (*HD1* and *HD2*). While this configuration is retained in some lineages, such as in *Malassezia* species [20], various studies have shown that in most other basidiomycetes, the *HD* locus remains largely conserved (except in some mushroom-forming species that have undergone segmental duplications of the *HD* genes; reviewed in [9, 28]), whereas the *P/R* locus has experienced lineage-specific expansions and structural rearrangements. For example, in *Ustilago maydis*, the size difference between the two allelic forms of the *P/R* locus (4.5 kb in the *a1* allele vs. 8 kb in the *a2* allele) results from the presence of two additional genes (*lga2* and *rga2*) in the *a2* allele that direct mitochondrial uniparental DNA inheritance (mito-UPI) [108, 109]. Similarly, several *Microbotryomycete*s (*Pucciniomycotina*) yeasts harbor substantially expanded *P/R* loci incorporating additional genes [26, 110, 111]. Likewise, tetrapolar *Cryptococcus* species exhibit an expanded *P/R* locus [35, 46, 77], and in *C. amylolentus* this region has also been implicated in mito-UPI [112].

Our analysis of the *P/R* locus across tetrapolar *Cryptococcus* and *Kwoniella* revealed distinct evolutionary trajectories, presumably originating from a shared ancestral locus containing four genes (*MF, STE3, STE12,* and *STE20*) (**Fig 10**, event 1). In *Kwoniella*, the *P/R* locus remains relatively compact across species, averaging 29.3 kb with a core set of nine conserved genes (**Fig 10**, event 2). Its evolution has involved (i) mating-type-specific changes, including the insertion of an additional gene of unknown function and *STE12* truncation in *a2* alleles; (ii) lineage-specific

small-scale expansions through the incorporation of two additional genes (*BSP3* and *UBP6*) via inversions; and (iii) intra-*P/R* locus recombination in *K. ovata*, *K. endophytica*, and *K. dendrophila*, generating a mosaic structure. The latter suggests that recombination suppression may not be fully established, though whether such recombination events are rare or recurrent remains unclear.

In contrast, the *P/R* locus in tetrapolar *Cryptococcus* species has undergone significant expansion, averaging 95.4 kb in clade B and 86.9 kb in clade D. Despite their similar sizes, substantial differences in gene content suggest that these expansions occurred independently after the divergence of the two lineages from their most recent common ancestor (**Fig 10**, event 5). A defining feature of *Cryptococcus P/R* loci is the duplication of pheromone (*MF*) genes, with each mating type harboring multiple copies, typically arranged in divergent orientations. This pattern is also observed in expanded *P/R* loci of *Microbotryomycetes* yeasts [26, 110, 111]. These duplications may have evolved to strengthen mating partner recognition, suggesting that high pheromone production is an important feature of the mating process in these species. However, they also introduce structural instability by facilitating the formation of inverted repeats, which can promote inversion loops and allele-specific structural rearrangements, potentially extending regions of recombination suppression between *P/R* alleles.

One way to mitigate structural instability would be to cluster pheromone gene pairs closer together, thereby restricting the extent of rearrangements. Consistent with this, we observed reduced structural variation in *MAT*α strains of *C. neoformans* and *C. gattii* species complex, where the two pheromone gene pairs are positioned in close proximity, separated by only two genes (*ZNF1* and *PRT1*). In contrast, *MAT***a** strains exhibit a more dispersed arrangement of pheromone genes, which appears to contribute to greater structural diversity across strains, including the VNI-specific rearrangements that arose following introgression of the *MAT***a** allele from the VNBI lineage. Once established, the presence of multiple pheromone gene copies may be maintained as they could serve as templates for intrallelic repair via gene conversion, a non-

673 reciprocal exchange that can occur during meiosis, primarily between homologous sequences but
674 also at lower frequencies between sister chromatids, or within duplicated regions on a single
675 chromosome during mitosis [113, 114]. If confirmed, this mechanism would resemble the role of
676 palindromic sequences in preserving gene integrity on the human Y chromosome [115-117].

677 A striking feature of the *P/R* locus in tetrapolar *Cryptococcus* species is its consistently
678 lower GC content compared to genome-wide averages, particularly in clades B and D, where it is
679 significantly enriched for AT-rich codons. One likely explanation is that recombination suppression
680 in this region has allowed AT-biased mutational pressure to accumulate over time, a pattern
681 observed in other fungi, including another *Tremellomycetes* lineage, *Trichosporonales*, where *MAT*
682 loci also exhibit reduced GC content [18]. Additionally, this AT enrichment may reflect relaxed
683 selective pressure on codon usage, particularly if *P/R* locus genes are expressed primarily under
684 specific mating conditions rather than constitutively. Indeed, genes with stage-specific expression
685 may experience weaker selection for translational efficiency, leading to reduced codon
686 optimization, as observed in *Arabidopsis thaliana*, where genes with restricted tissue-specific
687 expression tend to show lower codon adaptation [118]. The *frq* gene in *Neurospora crassa*
688 provides another example of how codon usage bias can influence regulatory processes:
689 optimization of *frq* codons disrupts circadian clock function, highlighting how non-optimal codon
690 usage can be functionally relevant [119]. These findings raise the possibility that codon usage
691 within the *P/R* locus may be shaped not only by recombination suppression and mutational biases
692 but also by selective constraints related to mating-type regulation and expression timing.

693 Another contributing factor could be the reduction of GC-biased gene conversion (gBGC),
694 which in many species counteracts AT-biased mutations in recombining regions [120-122]. If
695 suppressed recombination also limits gBGC in the *P/R* locus, this could further contribute to
696 localized GC depletion. Interestingly, the absence of a similar GC drop in the *MAT* locus of
697 *Cryptococcus* pathogens, despite the presence of recombination suppression, suggests that the

genomic GC background plays a role. In high-GC genomes (e.g., clades B and D), reduced recombination may lead to a more pronounced AT accumulation, whereas in *Cryptococcus* pathogens, which have a lower genomic GC content overall [48], the absence of such a shift suggests their genomes may already be at equilibrium under mutation pressure. Further investigation of codon adaptation indices, genome-wide relative synonymous codon usage (gw-RSCU), recombination rates, and expression profiles across diverse *Cryptococcus* species will be essential for disentangling the relative contributions of mutation bias, selection, and recombination dynamics in shaping the evolution of *P/R* loci regions.

Beyond the previously documented tetrapolar-to-bipolar transition in pathogenic *Cryptococcus* lineages, we identified three additional transitions: two resulting in a bipolar configuration, in *Cryptococcus* sp. 3 and *K. fici* (events 6 and 4 in **Fig 10**, respectively), and one leading to a pseudobipolar arrangement in *K. europaea* (**Fig 10**, event 3). Our findings suggest that the transition to bipolarity in *Cryptococcus* sp. 3 likely occurred independently of that observed in *Cryptococcus* pathogens (Clade A). First, we confirmed that the fusion of *P/R* and *HD* loci is unrelated to reduction from 14 to 13 chromosomes in this species and instead was initiated by a chromosomal translocation followed by an inversion event. Second, the *MAT* locus in *Cryptococcus* sp. 3 retains only *SXI2* (*HD2*), which is linked to the *STE3*α allele, differing from *Cryptococcus* pathogens, where *SXI2* is associated with *STE3***a**. Third, phylogenetic analyses revealed that many genes within the *MAT* locus in *Cryptococcus* sp. 3 do not group with **a** or α allele-specific clusters from pathogenic species. If the *MAT* locus fusion predated the divergence of these clades, one would expect to see trans-specific polymorphism, with alleles of the same mating type clustering across species. However, this pattern is not observed, which is more consistent with an independent fusion event in *Cryptococcus* sp. 3 rather than inheritance from a common ancestor with Clade A. Nevertheless, the absence of additional strains for this species,

limits our ability to fully delineate the *MAT* locus boundaries and confirm whether the opposite *MAT* allele carries the expected *SXI1/HD1–STE3***a** combination.

Furthermore, it remains uncertain whether the initial translocation event that brought the *P/R* and *HD* loci onto the same chromosome occurred independently in *Cryptococcus* sp. 3 or in a common ancestor of clades A and C. This uncertainty stems from two main factors. First, the phylogenetic relationships among clades A (bipolar), B (tetrapolar), and C (which includes *Cryptococcus* sp. 3) remain unresolved. Second, although our synteny analyses indicate that the translocation breakpoints at the origin of the physical linkage of the *MAT* loci appear to differ between *Cryptococcus* sp. 3 and pathogenic species (which would favor the possibility of independent translocation events), these differences might also result from distinct chromosomal rearrangements that accumulated along the two lineages. Further sampling and additional comparisons will be required to conclusively resolve these relationships and gain deeper insight into the broader implications of these transitions.

The composition of *MAT* loci can provide insight into both past and more recent transition events. For example, *K. fici* also underwent a transition to fused *MAT* loci, yet both *SXI1* (*HD1*) and *SXI2* (*HD2*) have been retained. This could suggest that the fusion event is relatively recent, and that insufficient time has passed for one *HD* gene to decay in each mating type, as observed in *Cryptococcus* pathogens, where remnants of *SXI2* persist next to *SXI1* in the *MAT*α alleles of some species [123]. However, retaining both *HD* genes in each allele is the most common organization in other bipolar basidiomycetes with fused *MAT* loci, including several *Microbotryum* [27], *Sporisorium* [30], and *Ustilago* species [31, 124, 125]. As in these fungi, this configuration may be functionally advantageous in *K. fici*, potentially reflecting differences in mating strategies or genomic constraints. For instance, if both *HD* genes have additional roles beyond mating that do not require heterodimerization, their retention could be selectively maintained, although this possibility remains largely unexplored in basidiomycetes.

In contrast, beyond *Cryptococcus* pathogens and *Cryptococcus* sp. 3 identified here, the only other documented case of a bipolar system with fused *MAT* loci containing a single *HD* gene per mating type occurs in the *Trichosporonales* [18]. A few other known cases of *HD* gene loss in heterothallic bipolar species, albeit distinct, include: (i) *Microbotryum violaceum* on *Silene caroliniana*, where *HD1* is lost in *a2* strains, while *a1* retains both *HD* genes [27]; (ii) *Microbotryum superbum*, where loss of *HD* function in mating compatibility may have relaxed constraints on *HD2* maintenance in *a1* strains, leading to its disruption, while *a2* strains retains both *HD* genes [126]; and (iii) xerotolerant *Wallemia* species, where *HD1* is retained in *MAT a1*, while the opposite mating type appears to lack any *HD* gene, potentially indicating another case of *HD* loss-of-function in mating-type determination [127]. Because *HD* compatibility is no longer an independent determinant of mating success in bipolar species, maintaining both *HD* genes in each *MAT* allele may no longer be necessary. Thus, their loss could result from genomic streamlining, eliminating redundancy once *MAT* loci became fused and recombination was suppressed.

The tetrapolar-to-bipolar transition in *Cryptococcus* pathogens has been hypothesized to result from a chromosomal translocation mediated by ectopic recombination between repetitive elements in the centromeric regions of the *P/R*- and *HD*-containing chromosomes [35]. This model had support from several observations: (i) *Cryptococcus* centromeres are enriched with species-specific long-terminal-repeat retrotransposons (LTRs) from the Ty3/*Gypsy* and Ty1/*copia* families shared among different centromeres, potentially facilitating recombination [23, 46, 48, 106, 107]; (ii) chromosomal arm exchanges have been identified during the divergence of *Cryptococcus* lineages from their common ancestor [35, 105]; and (iii) the repair of experimentally induced double-strand breaks at centromere-specific transposons in *C. neoformans* frequently results in chromosomal translocations [105].

Our expanded comparative analysis, however, indicates that the centromere of the *C. neoformans MAT* chromosome, along with extended flanking regions, is conserved across most

Cryptococcus species and in *K. shandongensis*, suggesting an ancestral organization. In contrast, in *C. amylolentus* this region aligns to two distinct chromosomes, implying that intercentromeric recombination occurred as a derived rather than ancestral feature. Therefore, these findings provide an alternative perspective to previous interpretations that proposed intercentromeric recombination as the primary event initiating the tetrapolar-to-bipolar transition in *Cryptococcus* pathogens.

Our analyses suggest that translocations outside centromeric regions may have played a more prominent role in repositioning the *P/R* and *HD* loci onto the same chromosome (**Fig 7**). Such non-centromeric translocations appear to underlie the transition to bipolarity in several *Ustilaginales* species, including *Sporisorium scitamineum* [30] and in the common ancestor of *Ustilago hordei*, *U. nuda*, and *U. bromivora* [31]. Non-centromeric translocations events are also readily apparent across other genomic regions when comparing genomes of different *Cryptococcus* species [48] and even among strains within the same species, as observed in *C. neoformans* [100, 128]. Whether such translocations are influenced by TEs has not yet been systematically analyzed in *Cryptococcus*; however, studies in other model organisms, including *Saccharomyces cerevisiae*, have shown that translocations can occur via homologous recombination between retrotransposons on different chromosomes [129]. Given the abundance of TEs in *Cryptococcus* genomes [48], a similar mechanism could have contributed to the chromosomal relocation of *MAT* loci and, ultimately, to the tetrapolar-to-bipolar transition.

Mating-type chromosomes with genetically linked *P/R-HD* loci can also arise from the fusion of entire ancestral *P/R* and *HD* chromosomes, as reported in *Microbotryum violaceum* on *Silene paradoxa* [27]. In *Kwoniella*, repeated and independent chromosome fusions have dramatically reduced chromosome numbers, with some extant species retaining only 3 (clade E) or 5 (clades G and H) chromosomes, descending from a 14-chromosome ancestral karyotype [48]. Yet, aside from the bipolar arrangement in *K. fici* (clade H) and the pseudobipolar

arrangement in *K. europaea* (clade E), no other instances of *P/R-HD* linkage were detected in *Kwoniella*. This observation reinforces the view that while chromosomal fusions and translocations provide the mechanistic basis for tetrapolar-to-bipolar transitions, they do not inherently drive shifts in breeding systems. Instead, whether a species remains tetrapolar or transitions to bipolarity likely depends on other related factors such as mating ecology (e.g., the degree of outcrossing vs. selfing/inbreeding), selection for mating efficiency, and long-term genomic stability.

The colonization of new habitats, including host-jumps, can lead to significantly altered dispersal patterns and changes in mating success, which may drive genetic rearrangements and facilitate adaptation to new niches. While sexuality is very frequent in some fungal groups (e.g., mushroom-forming basidiomycetes), asexual reproduction is widespread across the fungal kingdom, and approximately 20% of known fungal species lack documented evidence of sexual cycles [130, 131]. This is particularly evident in many fungal pathogens of animals, where sexual reproduction is rarely observed or remains entirely unknown [132, 133].

Among the species analyzed in this study, sexual cycles are documented in only 8 of 17 *Cryptococcus* species and 2 of 18 *Kwoniella* species. Research on these fungi is significantly hindered by the difficulty of culturing them, resulting in many species being known from only a single isolate [47]. Unlike many mushroom-forming basidiomycetes, most *Tremellomycetes* are parasites of lichens and other fungi, producing inconspicuous fruiting bodies or reproductive structures embedded within host tissues. This makes direct observation of their sexual reproduction in the field and the isolation of compatible partners for experimental studies particularly challenging. While genomic analysis of *MAT* loci content and organization allows for inferences about potential breeding systems (heterothallism vs. homothallism), experimental mating assays remain the most definitive approach for confirming sexual reproduction and viable progeny production. Expanding sampling efforts and discovering new species in their natural environments are crucial for uncovering the full diversity of *Tremellomycetes*, gaining deeper

822 insights into their reproductive strategies, and understanding their ecological roles and
823 evolutionary trajectories [47, 49].

824 Our experimental analysis confirmed a sexual cycle in *Cryptococcus decagattii*,
825 demonstrating its ability to undergo mating and sporulation. However, we were unable to induce
826 sexual reproduction in several other lineages, including *C. tetragattii* and *Cryptococcus* sp. 4. The
827 failure to observe mating under laboratory conditions may stem from suboptimal culture conditions
828 rather than an inherent inability to reproduce sexually, underscoring the need to identify
829 appropriate mating conditions and compatible partners. Notably, *Cryptococcus* sp. 3 is closely
830 related to *C. depauperatus*, a homothallic species, yet it does not undergo sexual reproduction
831 alone, suggesting that it is not homothallic. This is consistent with its genomic composition, which
832 indicates a heterothallic, bipolar *MAT* configuration.

833 One possible approach to overcome these challenges and induce the sexual state is the
834 targeted deletion of *CRG1*, a regulator of G protein signaling that negatively regulates Gα proteins
835 activated through pheromone receptors, thereby downregulating signaling from a pheromone-
836 bound *STE3* receptor [134, 135]. *C. neoformans*, *crg1*Δ mutants are hypersensitive to mating
837 signals, exhibit hyperfilamentation, and show enhanced mating efficiency, particularly under
838 conditions that would otherwise be non-permissive for sexual reproduction [135, 136]. Given that
839 *CRG1* is conserved across *Cryptococcus* and *Kwoniella*, its deletion in cryptic lineages with
840 compatible mating-type strains could facilitate mating and reveal latent sexual cycles, contingent
841 on establishing these species as genetically tractable.

842 In *K. mangrovensis*, progeny analysis from a mating cross revealed significant ploidy
843 variation, with a high incidence of aneuploid strains. The primary factor contributing to this
844 outcome appears to be the difference in ploidy between the parental strains: while CBS8507 is
845 haploid, FACS analysis indicated that CBS10435 predominantly consists of diploid cells. As
846 observed in other fungi, haploid-diploid mating can result in chromosome mis-segregation and

aneuploid progeny due to improper chromosome pairing and segregation during meiosis. This effect may be further exacerbated in *K. mangrovensis* by its highly reduced karyotype (3 chrs.), where errors in the segregation of its disproportionately large chromosome may have an outsized impact on ploidy balance. Despite this, genomic comparisons of F1 isolates with parental strains revealed patterns consistent with recombination, with chromosome segments mapping to different parents, supporting the occurrence of meiotic exchange. Additionally, loss of heterozygosity (LOH) was observed in some regions. While ploidy estimates from genome-wide read depth analysis and FACS were generally concordant, some discrepancies suggest either intrinsic genomic instability or methodological differences in DNA content estimation, potentially stemming from differences in cell wall composition, nuclear DNA organization, or cell cycle stage, which could affect staining efficiency and fluorescence intensity in FACS-based ploidy assessment.

Interestingly, some of the *K. mangrovensis* F1 progeny, by virtue of their diploid/aneuploid status, carried compatible *P/R* and *HD* alleles (*a1a2b1b2*), suggesting the potential for self-fertility. Under mating conditions these strains formed hyphae and basidia similar to those produced by the parental cross, with some even exhibiting more extensive filamentation. Further investigation of the *K. mangrovensis* life cycle, including controlled crosses with haploid strains, synchronized cultures, and genomic analyses of progeny, will be necessary to clarify the extent to which aneuploidy results from chromosomal instability and is tolerated versus differences in parental ploidy during mating.

Overall, our findings highlight both the challenges and opportunities in understanding the reproductive biology and evolutionary trajectories of *Cryptococcus* and *Kwoniella*. Comparative genomic analyses revealed distinct evolutionary paths shaping *MAT* loci architecture, including independent *P/R* locus expansions, diverse modes of *HD-P/R* fusion, and convergent transitions in mating systems. Additional analysis of sexual cycles will require expanded strain sampling, optimized mating conditions, and genetic approaches such as *CRG1* deletion to uncover latent

reproductive potential. Beyond laboratory investigations, assessing sexual recombination in natural populations and exploring the hidden diversity of *Cryptococcus* species — particularly in their likely African origin [49] — will be crucial for understanding their speciation, ecology, and adaptation, as well as the evolutionary forces shaping *MAT* locus evolution and reproductive transitions, with possible implications for pathogenic emergence.

Materials and methods

Strains and media

Strains studied in this work were routinely grown on YPD medium (10 g/L yeast extract, 20 g/L Bacto Peptone, 20 g/L dextrose, and 20 g/L agar) unless indicated otherwise. *C. neoformans* and *C. deneoformans* strains were incubated at 30°C, while other *Cryptococcus* and *Kwoniella* species strains were grown at room temperature (20-23°C). A complete list of strains used in this study is provided in **S1 Appendix**. Strain NCYC1536 was obtained from the National Collection of Yeast Cultures (Norwich, UK), and strain CMW60451 was isolated in October 2022, from a bark beetle (*Lanurgus* sp.) infesting twigs of *Widdringtonia cedarbergensis* in the Cederberg Mountains, South Africa.

Genomic DNA extraction

High-molecular weight (HMW) DNA was extracted with a cetyltrimethylammonium bromide (CTAB) extraction as previously described [105], minimizing sharing stress during sample preparation. DNA quality was evaluated by determining the A260/A280 and A260/A230 ratios on NanoDrop spectrophotometer (Thermo). Integrity and fragment size were analyzed using clamped homogeneous electric field (CHEF) electrophoresis, and gDNA for short-read whole-genome sequencing (Illumina) was extracted with a phenol:chloroform-based protocol, both as previously

described [48]. DNA concentration for all samples was measured with Qubit dsDNA Assay Kits (Invitrogen) on the Qubit fluorometer.

Genome sequencing

Whole-genome sequencing was performed with PacBio, Nanopore, and Illumina technologies. For *K. heveanensis* BCC8398 and *K. mangrovensis* CBS10435, Illumina sequencing at the Broad Institute Genomics Platform utilized “fragment” and “jumping” libraries, constructed and sequenced as previously specified [48]. All other Illumina sequencing was conducted at the Duke University Sequencing and Genomic Technologies (DUSGT) Core, with libraries prepared with Kapa HyperPlus library kit and sequenced as paired-end 2 x 150 bp reads on various Illumina platforms. For PacBio sequencing, 15- to 20-kb insertion-size libraries were prepared and run on a PacBio RS II or Sequel (2.0 chemistry) system at the DUSGT. Nanopore sequencing was carried out in-house. Single-strain libraries were prepared with either the SQK-LSK108 or SQK-LSK110 kit, whereas up to three DNA samples were barcoded with the SQK-LSK109 and EXP-NBD103/EXP-NBD104, or SQK-NBD114.24 kits. Libraries were prepared according to the manufacturer’s protocols, either individually or pooled, and sequenced on R9 flow cells (FLO-MN106) or R10 flow cells (FLO-MIN114) for 48-72 hours at default voltage on a MinION Mk1B or MinION Mk1C system. The MinION software version available at the time of each run was applied. Further details on specific genome sequencing platforms, basecalling, and demultiplexing are provided in **S1 Appendix**.

Genome assembly

Complete genomes were assembled with Canu [137] using Nanopore or PacBio sequencing data and default parameters. The accuracy of the assemblies was improved through

initial error correction with Medaka (https://github.com/nanoporetech/medaka) for Nanopore-based assemblies, followed by up to five rounds of iterative polishing with Pilon v1.22 [138] (`--fix all`) for all assemblies, using Illumina reads aligned to the first pass-polished assembly with BWA-MEM v0.7.17-r1188 [139]. Contigs containing only rDNA sequences, detected by Barrnap (https://github.com/tseemann/barrnap) (`--kingdom euk`), or those classified as mitochondrial DNA, were excluded from the final nuclear genome assembly. To confirm assembly completeness and evaluate telomeric regions, Nanopore/PacBio and Illumina reads were realigned to the Canu-corrected assembly using minimap2 v2.9-r720 [140] and BWA-MEM, respectively, and read coverage profiles were examined in the Integrative Genomics Viewer (IGV) [141]. Draft genome assemblies for *K. europaea* PYCC6162 and *K. botswanensis* CBS12717 were generated with SPAdes v3.15.3 using default settings, while the assembly for *K. heveanensis* BCC8398 and *K. mangrovensis* CBS10435 was constructed with Allpaths [142]. Genome assemblies and raw sequencing data have been deposited in DDBJ/EMBL/GenBank under the BioProject numbers listed in **S1 Appendix**, which also provides specific details on sequencing platforms, basecalling, demultiplexing, and assembly parameters for each genome.

Gene prediction and annotation

Gene models were predicted on repeat-masked assemblies using either BRAKER2 v2.1.5 [143] or Funannotate v1.8.9 (https://github.com/nextgenusfs/funannotate), following previously described methodologies [20, 23, 48]. Manual inspection and correction were performed exclusively for the mating-type gene models, including the prediction and addition of short mating pheromone precursor genes across all species. Functional annotation was integrated into the final gene models using the Funannotate *"annotate"* module, which incorporated data for PFAM and InterPro domains, Gene Ontology (GO) terms, fungal transcription factors, Cluster of Orthologous Genes (COGs), secondary metabolites, Carbohydrate-Active Enzymes (CAZYmes), secreted

proteins, proteases (MEROPS), and Benchmarking Universal Single-Copy Orthologs (BUSCO) groups. InterPro domain data, COG annotations, and secondary metabolite predictions were obtained by InterProScan v5.55-88.0, eggNOGmapper v.2.1.7 (eggNOG DB version: 5.0.2) [144], and AntiSMASH v6.1.0 [145], before being passed to Funannotate *annotate* with the options *--iprscan, --eggnog*, and *--antismash*. Specific parameters are provided in **S1 Appendix**.

Ortholog identification and sequence alignment

To construct the phylogenomic data matrix, single-copy orthologs (SC-OGs) were identified across *Cryptococcus* and *Kwoniella* species, as well as three outgroup species (*Tremella mesenterica* ATCC28783, GCA_004117975.1; *Saitozyma podzolica* DSM27192, GCA_003942215.1; and *Bullera alba* JCM2954, GCA_001600095.1) using OrthoFinder v3.0.1b1 with options: *-M msa -S diamond_ultra_sens -I 1.5 -M msa -A mafft -T fasttree -t 48 -a 6*. A total of 3,086 SC-OGs shared among all species were identified. The amino acid sequences of these SC-OGs were individually aligned with MAFFT v7.310 [146] using *--localpair --maxiterate 1000* and trimmed with TrimAl v1.4.rev22 [147] using the options *-gappyout -keepheader*. The same approach was applied to construct the phylogenomic data matrix for *C. neoformans* strains representing different VN groups, along with three outgroup strains of *C. deneoformans*. This analysis identified 5,439 SC-OGs shared across all strains, which were subsequently aligned and trimmed as described above.

Species phylogeny and estimation of topological support

The updated phylogeny of *Cryptococcus* and *Kwoniella* species (**Figs 1A** and **S1**) was inferred using two complementary strategies: (i) concatenation-based partitioned maximum likelihood (ML) phylogeny reconstruction in IQ-TREE v2.1.6 [148], and (ii) gene-based

coalescence analysis in ASTRAL v5.7.8 [149]. For the concatenation approach, individual amino acid alignments for 3,086 single-copy orthologs (SC-OGs) were combined into a partitioned supermatrix (52 taxa, with 3,086 partitions and 1,686,691 sites) using the *"-p"* option in IQ-TREE. The edge-linked proportional partition model [150] was applied to account for differences in evolutionary rates across partitions, and the best-fit substitution model for each partition was identified with ModelFinder [151] based on the Bayesian information criterion. The ML tree was inferred with the parameters *--seed 54321 -m MFP -msub nuclear -B 1000 -alrt 1000 -T 14*, incorporating 1,000 ultrafast bootstrap (UFboot) replicates [152] and Shimodaira–Hasegawa approximate likelihood ratio tests (SH-aLRT) for branch support assessment. For the coalescence-based approach, ML gene trees were independently constructed for each SC-OG alignment using IQ-TREE's *"-S"* option, which performs both model selection and tree inference for individual alignments. The resulting trees were then input to ASTRAL to reconstruct a species phylogeny under default settings. Quartet support values were computed with the *-t 2* option, providing quartet support metrics for the primary topology (q1) and alternative topologies (q2, q3), and local posterior probability (LPP) support. Genealogical concordance and topological support for branches in the concatenated ML tree were evaluated using the gene concordance factor (gCF) and site concordance factor (sCF) metrics as implemented in IQ-TREE. This analysis utilized both the best-scoring concatenated ML tree (concat.treefile) and the set of gene trees (loci.treefile), with options *-t –gcf -p --scf 100*. The concatenation-based approach was also employed to resolve relationships within *C. neoformans* strains representing different VN groups (**Fig 4A**). Phylogenetic trees were visualized and annotated with iTOL v7.

Gene genealogies

Amino acid sequences of genes of interest were extracted from the corresponding OGs identified by OrthoFinder. These sequences were manually inspected and reannotated as needed.

Curated protein sequences were aligned and trimmed as above, and ML phylogenies were reconstructed and visualized as above. Detailed model parameters are provided in the tree files accessible at https://doi.org/10.5281/zenodo.14851287.

Synteny analyses, *MAT* loci delineation, and centromere identification

Conserved synteny blocks in pairwise genome comparisons were determined using SynChro [102], with the delta parameter set to 3 to ensure high stringency. Comparisons shown in **Figs 6A**, **8A** and **S9A** employed the *K. shandongensis* genome assembly as the reference. Detailed linear synteny plots comparing chromosomes and specific genomic regions, including the *MAT* loci, were generated with EasyFig v2.2.2 [153] using BLASTN. Centromeres positions were determined by in silico analysis as previously described [48], combining the detection of centromere-associated LTR elements and synteny analysis. Centromere lengths were estimated as the intergenic regions between flanking centromeric genes (**S1 Appendix**). *MAT* loci were initially identified using BLAST searches, with *C. neoformans*-derived *MAT* genes and their flanking proteins as queries. The *HD* loci were defined as the regions spanning the *HD1* and *HD2* genes. The *P/R* loci were delineated based on structural comparisons between mating types. For species with available strains of opposite mating types, the *P/R* locus was defined as the region where synteny between the mating types is disrupted, with the boundaries corresponding to the restoration of synteny. For species with only a single strain analyzed, the *P/R* locus length was inferred based on the distance between *P/R*-flanking genes, as determined from comparisons involving opposite mating types. Statistical analyses of *MAT* loci lengths across species, clades, or mating-types were conducted with Python3 with Pandas, Seaborn, Matplotlib, and SciPy libraries. Differences between the two groups were assessed using the two-sided Mann–Whitney U test. Scripts and raw data are available at https://doi.org/10.5281/zenodo.14851287. To

enhance figure clarity and readability, plots were refined by adjusting color schemes, modifying labels, and manually adding additional features using Adobe Illustrator.

Analysis of gene content in *MAT* loci, their frequency across species, and gene essentiality classification

A curated gene presence/absence matrix was constructed following the delineation of the *MAT* loci across species. Genes were categorized as “present” (coded as “1”), "absent" (coded as “0”), “unclear” (coded as “?”), or “pseudogene” (coded as “pseudo”). Unclear cases indicated instances where gene presence could not be conclusively determined due to incomplete or uncertain *MAT* locus boundaries. For frequency calculations, pseudogenes were treated as present, while unclear cases were treated as absent. The presence of either *SXI1* or *SXI2* was scored as present and combined for frequency quantification. The matrix also included information on gene essentiality, exclusively inferred from studies on *C. neoformans* H99 (i.e., gene essentiality in other species was extrapolated solely from H99-based predictions). Essentiality classifications were based on experimentally validated studies [34, 154, 155] and predictions from a recent high-throughput transposon mutagenesis and sequencing (TN-seq) study (https://bbillmyre.shinyapps.io/Crypto_TN_seq_viewer/) [91]. Genes without orthologs in H99 were categorized as “unknown” due to the absence of data for classification. A custom R script (`0_calculate_gene_presence_frequency.R`) was developed to analyze and plot gene presence frequencies, sorted by frequency. The presence/absence matrix was also visualized as a heatmap using a custom R script (`1_plot_gene_matrix_based_on_frequency.R`), with genes ordered by their frequency across strains. While *SXI1* and *SXI2* were shown as separate columns in the heatmap, their ranking reflected their combined presence frequency. These analyses and visualizations were conducted in R using the dplyr, ggplot2, ComplexHeatmap, and circlize packages. Additionally, gene presence/absence across clades was analyzed using a custom

Python script (*2_compare_MAT_gene_content_across_clades.py*) that calculates clade-specific and pairwise gene presence, identify genes shared across all specified clades, and highlight genes unique to individual clades. Scripts are available at https://doi.org/10.5281/zenodo.14851287, and the gene presence/absence input matrices are provided in **S2 and S3 Appendices**.

GC content and codon composition analysis

Genome-wide GC content and deviations from the mean GC% were analyzed using a custom Python script (*0_gc_content_analysis_and_plots.py*), employing a non-overlapping sliding window approach (1 kb for whole-genome analyses, 0.5 kb for zoomed-in views). Regions of interest (e.g., *P/R* loci) were highlighted using BED file coordinates. The resulting plots were further refined in Adobe Illustrator for producing **Fig 2**. To assess whether the *P/R* locus exhibited lower GC content than genome-wide averages, we used another custom Python script (*1b_gc_content_analysis.py*) that calculated GC content for the *P/R* locus and compared it to 100 randomly sampled regions of the same size, either genome-wide or on the same chromosome. A one-sample *t*-test assessed whether the *P/R* locus GC content differed significantly from the sampled mean, and a permutation test (10,000 iterations) generated a null distribution to calculate *P*-values and Z-scores. The analysis was automated using the shell script (*1a_run_script_vs_wg.sh*), consolidating results into a summary table (**S4 Appendix**). Codon usage analysis was conducted with a third custom Python script (*3_codon_usage_analysis.py*), which compared codon frequencies in the *P/R* locus and genome-wide coding sequences. Codons were classified as AT- or GC-rich, and their counts were compared between the two regions. Chi-square tests ($P < 0.05$) evaluated: (i) differences in codon usage (64 codons) and (ii) AT/GC composition differences. For codon usage, the null hypothesis assumed the codon distributions in the *P/R* locus matched genome-wide distributions. Scripts are raw data are

accessible at https://doi.org/10.5281/zenodo.14851287 and raw data is also summarized in **S4 Appendix**.

Variant analysis and SNP distribution in *C. neoformans MAT*a strains

Multi-genome variant analysis was performed with the Snippy pipeline v4.6.0 (https://github.com/tseemann/snippy) using the newly assembled genome of *C. neoformans* Ftc555-1 (VNBI) as the reference, and the following parameters: *--cpus 10, --unmapped, --mincov 10, and --minfrac 0.9*. Paired-end reads were obtained from the NCBI SRA database and are provided in **S1 Appendix**. Three datasets were independently used: (i) the entire genome, (ii) only the *MAT* locus region (chr_5:153,773-277,230), and (iii) only chr. 5. The *MAT* locus and chr. 5 were extracted from the reference genome to generate region-specific GenBank files, which were used as inputs for Snippy to identify SNPs and construct region-specific phylogenies. Visualization of SNV distributions was based on the whole-genome dataset and included both chromosome-wide and zoomed-in views of specific regions. A custom Python script was used to process merged SNV datasets, compute total SNP counts for each strain, and generate visualizations with gradient color-coding for SNP density and annotated regions of interest. Final composite figures were refined using Adobe Illustrator. Core SNP alignments (core.aln) from each dataset were used for phylogenetic reconstruction with IQ-TREE2, incorporating 10,000 UFboot replicates and SH-aLRT tests for branch support. Scripts for SNV analysis and visualization are available at https://doi.org/10.5281/zenodo.14851287.

Chromosome composition analysis of *K. mangrovensis* progeny

To characterize the chromosome composition of *K. mangrovensis* progeny, a combined nuclear reference genome was constructed using the genome assemblies of the two parental

strains, CBS8507 and CBS10435. For CBS8507, the original assembly consisted of four contigs, two of which represented the same chromosome but were fragmented at the rDNA array. To streamline downstream analyses and facilitate data interpretation, these two contigs were merged into a single contig. A small gap (5 Ns) was introduced to connect the contigs in the correct orientation, aligning the rDNA genes consistently in the same direction and resulting in a finalized 3-chromosome assembly. For CBS10435, the draft assembly, which comprised 37 contigs, was reordered and reoriented to match the 3-chromosome assembly of CBS8507. This was achieved using the D-GENIES tool [156], which employs minimap2 for genome alignment. Raw Illumina paired-end reads from selected *K. mangrovensis* progeny, along with the combined reference genome, were subsequently processed through the sppIDer pipeline [157]. The sppIDer workflow sequentially mapped the Illumina short reads to the combined reference genome, applied stringent quality filtering (MQ > 3), and generated depth-of-coverage plots. Chromosome number and ploidy for each progeny were estimated by integrating the depth-of-coverage data from sppIDer with flow cytometry results.

Mating assays and phenotyping on mating inducing media

Mating assays using *K. mangrovensis* strains CBS8507 and CBS10435 were performed as previously described [65, 66]. Equal amounts of cells from each strain were mixed on V8 agar (10 g/L yeast extract, 20 g/L Bacto Peptone, 20 g/L dextrose, 20 g/L agar; pH 5) or corn meal agar (CMA; 15 g/L corn meal agar, 5 g/L agar) and incubated in the dark at room temperature (approximately 20-23°C) for up to one month. Plates were regularly monitored for the development of mating structures. Random progeny (F1) was recovered by microdissecting individual cells from cell clusters [78] embedded in or on the surface of the mating media (V8 and CMA) at the edges of mating patches. A total of 51 progeny were collected from dissection plates, grown axenically on YPD plates, and stored as glycerol stocks at -80 °C. Selected F1 progeny (MP11, MP15, MP16,

MP18, MP49, MP50, MP51, MP54 and MP59) were assessed for hyphal formation and mating structures on V8 pH5 and CMA media, compared to the parental cross (CBS8507 x CBS10432) and the solo culture of each parental strain. Assays were repeated 3 separate times to ensure consistent phenotypic observations.

Potentially compatible strain pairs from *Cryptococcus* species lacking documented sexual reproduction (IND107 x CBS11718; CBS11687 x 7685027; and DSM108351 x NCYC1536) were tested for mating ability under conditions known to induce mating in *C. neoformans* and *C. deneoformans* (V8 media, in the dark, and at room temperature) for up to a month, with regular monitoring [78]. Solo cultures of each strain were also evaluated for their capacity to undergo sexual reproduction without a compatible partner. Similar experiments were performed for strains of newly described species lacking known compatible partners. All assays were repeated three times to confirm phenotype consistency.

Analysis of *K. mangrovensis* progeny by PCR-restriction fragment length polymorphism (RFLP) analysis

Kwoniella mangrovensis has only three chromosomes. To analyze the F1 progeny from the cross between strains CBS8507 and CBS10432, six primer pairs (one per chromosomal arm) were selected for the initial screening. PCR primers were manually designed to produce a distinct pattern between the two parental alleles after digestion of the PCR products with specific restriction enzymes. As a positive control for heterozygosity, PCR products from a 1:1 DNA mixture of both parental strains were amplified for each primer pair, followed by digestion with the corresponding restriction enzyme. The results of this analysis and primer sequences are provided in **S5 Appendix**.

Microscopy

The edge of yeast colonies and mating patches were examined and photographed to assess hyphal growth, basidia, and spores. Imaging was performed with a Zeiss Axio Scope.A1 microscope equipped with an Axiocam Color camera, using the ZEN Lite V3.4 software.

Fluorescence-activated cell-sorting (FACS)

Fluorescence-activated cell sorting (FACS) was performed to determine the ploidy of *K. mangrovensis* strains CBS8507, CBS10432, and nine progeny strains from their cross, as previously described [158]. Briefly, strains were grown overnight at room temperature (21-23°C) on YPD medium, harvested, and washed with PBS. Cells were fixed using 2 ml of 70% ethanol at 4°C overnight. Fixed cell pellets were washed with 1 ml of NS buffer (10 mM Tris-HCl pH 7.2, 250 mM sucrose, 1 mM EDTA pH 8.0, 1 mM $MgCl_2$, 0.1mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.4 mM phenylmethylsulfonyl fluoride, and 7 mM β-mercaptoethanol) and then stained with 5 µl of propidium iodide (0.5 mg/ml) in 180 µl NS buffer with 20 µl of RNase (10 mg/ml) at 4°C overnight. Lastly, 50 µl of stained cells were diluted in 2 ml of 50 mM Tris-HCl pH 8.0 and sonicated for 1 min before analysis at the Duke Cancer Institute Flow Cytometry Shared Resource. Data were collected from 10,000 cells using the FL1 channel on a Becton-Dickinson FACScan and analyzed with FlowJo software. Strains JEC21 and XL143 were respectively used as haploid and diploid controls [159].

Funding Statement

This study was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under awards R01 AI050113-20 (J.H.), R01 AI039115-27 (J.H.), and R01 AI33654-08 (J.H.). Funding for the collection of *Cryptococcus* isolates from bark beetles was provided by the Harry Oppenheimer Fellowship Award of The Oppenheimer Memorial Trust (to M.J.W) and the SARChI Chair in Fungal Genomics, supported by the Department of Science and Innovation (DSI)/National Research Foundation (NRF), South Africa (to B.D.W.). JH is Co-Director and Fellow of the CIFAR program Fungal Kingdom: Threats & Opportunities. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

We thank Vikas Yadav for assistance with Nanopore data acquisition, Terrance Shea and Christina Cuomo for providing the genomes assemblies of *K. mangrovensis* CBS8886 and CBS10435, as well as *K. heveanensis* BCC8398, and the Broad Institute Genomics Platform for generating Illumina sequencing data for these three strains. We also thank Fred Dietrich for computational resources and Prof Francois Roets for providing permits to collect specimens in South Africa.

Author contributions

Conceptualization: Marco A. Coelho, Márcia David-Palma, Joseph Heitman

Data Curation: Marco A. Coelho, Márcia David-Palma

Formal analysis: Marco A. Coelho

Funding acquisition: Brenda D. Wingfield, Michael J. Wingfield, Joseph Heitman

Investigation: Marco A. Coelho, Márcia David-Palma, Nam Pham, Seonju Marincowitz, Janneke Aylward, Sheng Sun

Project Administration: Brenda D. Wingfield, Michael J. Wingfield, Joseph Heitman

Resources: Andrey M. Yurkov, Brenda D. Wingfield, Michael J. Wingfield, Joseph Heitman

Software: Marco A. Coelho

Supervision: Joseph Heitman

Visualization: Marco A. Coelho, Márcia David-Palma

Writing – original draft: Marco A. Coelho, Márcia David-Palma

Writing – review & editing: Marco A. Coelho, Márcia David-Palma, Nam Pham, Seonju Marincowitz, Janneke Aylward, Andrey M. Yurkov, Brenda D. Wingfield, Michael J. Wingfield, Sheng Sun, Joseph Heitman

References

1. Otto SP, Lenormand T. Resolving the paradox of sex and recombination. Nat Rev Genet. 2002;3(4):252-61. Epub 2002/04/23. doi: 10.1038/nrg761. PubMed PMID: 11967550.
2. de Visser JA, Elena SF. The evolution of sex: empirical insights into the roles of epistasis and drift. Nat Rev Genet. 2007;8(2):139-49. Epub 2007/01/19. doi: 10.1038/nrg1985. PubMed PMID: 17230200.
3. Goodenough U, Heitman J. Origins of eukaryotic sexual reproduction. Cold Spring Harb Perspect Biol. 2014;6(3). Epub 2014/03/05. doi: 10.1101/cshperspect.a016154. PubMed PMID: 24591519; PubMed Central PMCID: PMCPMC3949356.
4. Hawksworth DL, Lucking R. Fungal diversity revisited: 2.2 to 3.8 million species. Microbiol Spectr. 2017;5(4). doi: 10.1128/microbiolspec.FUNK-0052-2016. PubMed PMID: 28752818; PubMed Central PMCID: PMCPMC11687528.
5. Heitman J, Howlett BJ, Crous PW, Stukenbrock EH, James TY, Gow NAR. The Fungal Kingdom: American Society of Microbiology; 2017.
6. Begerow D, Kemler M, Feige A, Yurkov A. Parasitism in yeasts. In: Buzzini P, Lachance M-A, Yurkov A, editors. Yeasts in Natural Ecosystems: Ecology. Cham: Springer International Publishing; 2017. p. 179-210.
7. Lee SC, Ni M, Li W, Shertz C, Heitman J. The evolution of sex: a perspective from the fungal kingdom. Microbiol Mol Biol Rev. 2010;74(2):298-340. Epub 2010/05/29. doi: 10.1128/MMBR.00005-10. PubMed PMID: 20508251; PubMed Central PMCID: PMCPMC2884414.
8. Lin X, Heitman J. Mechanisms of homothallism in Fungi and transitions between heterothallism and homothallism. In: Heitman J, Kronstad JW, Taylor JW, Casselton LA, editors. Sex in fungi: Molecular determination and evolutionary implications. Washington D.C.: ASM Press; 2007. p. 35-57.
9. Coelho MA, Bakkeren G, Sun S, Hood ME, Giraud T. Fungal Sex: the Basidiomycota. Microbiol Spectr. 2017;5(3):10.1128/microbiolspec.FUNK-0046-2016. Epub 2017/06/10. doi: 10.1128/microbiolspec.FUNK-0046-2016. PubMed PMID: 28597825; PubMed Central PMCID: PMCPMC5467461.
10. Kües U, James TY, Heitman J. Mating type in basidiomycetes: unipolar, bipolar, and tetrapolar patterns of sexuality. In: Pöggeler S, Wöstemeyer J, editors. Evolution of fungi and Fungal-like organisms. The Mycota. 6. Heidelberg, Germany: Springer Berlin Heidelberg; 2011. p. 97-160.
11. James TY, Srivilai P, Kues U, Vilgalys R. Evolution of the bipolar mating system of the mushroom *Coprinellus disseminatus* from its tetrapolar ancestors involves loss of mating-type-specific pheromone receptor function. Genetics. 2006;172(3):1877-91. Epub 2006/02/08. doi: 10.1534/genetics.105.051128. PubMed PMID: 16461425; PubMed Central PMCID: PMCPMC1456265.
12. Yi R, Tachikawa T, Ishikawa M, Mukaiyama H, Bao D, Aimi T. Genomic structure of the *A* mating-type locus in a bipolar basidiomycete, *Pholiota nameko*. Mycol Res. 2009;113(Pt 2):240-8. Epub 20081119. doi: 10.1016/j.mycres.2008.11.002. PubMed PMID: 19049868.
13. James TY, Lee M, van Diepen LT. A single mating-type locus composed of homeodomain genes promotes nuclear migration and heterokaryosis in the white-rot fungus *Phanerochaete chrysosporium*. Eukaryot Cell. 2011;10(2):249-61. Epub 20101203. doi: 10.1128/EC.00212-10. PubMed PMID: 21131435; PubMed Central PMCID: PMCPMC3067404.
14. Olson A, Aerts A, Asiegbu F, Belbahri L, Bouzid O, Broberg A, et al. Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen. New Phytol. 2012;194(4):1001-13. Epub 20120328. doi: 10.1111/j.1469-8137.2012.04128.x. PubMed PMID: 22463738.
15. Bakkeren G, Kronstad JW. Linkage of mating-type loci distinguishes bipolar from tetrapolar mating in basidiomycetous smut fungi. Proc Natl Acad Sci U S A. 1994;91(15):7085-9. Epub 1994/07/19. doi: 10.1073/pnas.91.15.7085. PubMed PMID: 7913746; PubMed Central PMCID: PMCPMC44343.

16. Lengeler KB, Fox DS, Fraser JA, Allen A, Forrester K, Dietrich FS, et al. Mating-type locus of *Cryptococcus neoformans*: a step in the evolution of sex chromosomes. Eukaryot Cell. 2002;1(5):704-18. doi: 10.1128/ec.1.5.704-718.2002. PubMed PMID: 12455690; PubMed Central PMCID: PMCPMC126754.

17. Badouin H, Hood ME, Gouzy J, Aguileta G, Siguenza S, Perlin MH, et al. Chaos of rearrangements in the mating-type chromosomes of the anther-smut fungus microbotryum lychnidis-dioicae. Genetics. 2015;200(4):1275-84. Epub 2015/06/06. doi: 10.1534/genetics.115.177709. PubMed PMID: 26044594; PubMed Central PMCID: PMCPMC4574255.

18. Sun S, Coelho MA, Heitman J, Nowrousian M. Convergent evolution of linked mating-type loci in basidiomycete fungi. PLOS Genet. 2019;15(9):e1008365. Epub 2019/09/07. doi: 10.1371/journal.pgen.1008365. PubMed PMID: 31490920; PubMed Central PMCID: PMCPMC6730849.

19. Gioti A, Nystedt B, Li W, Xu J, Andersson A, Averette AF, et al. Genomic insights into the atopic eczema-associated skin commensal yeast *Malassezia sympodialis*. mBio. 2013;4(1):e00572-12. Epub 2013/01/24. doi: 10.1128/mBio.00572-12. PubMed PMID: 23341551; PubMed Central PMCID: PMCPMC3560662.

20. Coelho MA, Ianiri G, David-Palma M, Theelen B, Goyal R, Narayanan A, et al. Frequent transitions in mating-type locus chromosomal organization in *Malassezia* and early steps in sexual reproduction. Proc Natl Acad Sci U S A. 2023;120(32):e2305094120. Epub 2023/07/31. doi: 10.1073/pnas.2305094120. PubMed PMID: 37523560; PubMed Central PMCID: PMCPMC10410736.

21. David-Palma M, Sampaio JP, Gonçalves P. Genetic dissection of sexual reproduction in a primary homothallic basidiomycete. PLOS Genet. 2016;12(6):e1006110. Epub 2016/06/22. doi: 10.1371/journal.pgen.1006110. PubMed PMID: 27327578; PubMed Central PMCID: PMCPMC4915694.

22. Cabrita A, David-Palma M, Brito PH, Heitman J, Coelho MA, Goncalves P. Multiple pathways to homothallism in closely related yeast lineages in the Basidiomycota. mBio. 2021;12(1). Epub 2021/02/18. doi: 10.1128/mBio.03130-20. PubMed PMID: 33593979; PubMed Central PMCID: PMCPMC8545103.

23. Passer AR, Clancey SA, Shea T, David-Palma M, Averette AF, Boekhout T, et al. Obligate sexual reproduction of a homothallic fungus closely related to the *Cryptococcus* pathogenic species complex. eLife. 2022;11. Epub 2022/06/18. doi: 10.7554/eLife.79114. PubMed PMID: 35713948; PubMed Central PMCID: PMCPMC9296135.

24. Wang YW, McKeon MC, Elmore H, Hess J, Golan J, Gage H, et al. Invasive Californian death caps develop mushrooms unisexually and bisexually. Nat Commun. 2023;14(1):6560. Epub 2023/10/25. doi: 10.1038/s41467-023-42317-z. PubMed PMID: 37875491; PubMed Central PMCID: PMCPMC10598064.

25. Lin X, Hull CM, Heitman J. Sexual reproduction between partners of the same mating type in *Cryptococcus neoformans*. Nature. 2005;434(7036):1017-21. doi: 10.1038/nature03448. PubMed Central PMCID: PMCPMID: 15846346.

26. Maia TM, Lopes ST, Almeida JM, Rosa LH, Sampaio JP, Goncalves P, et al. Evolution of mating systems in basidiomycetes and the genetic architecture underlying mating-type determination in the yeast *Leucosporidium scottii*. Genetics. 2015;201(1):75-89. Epub 2015/07/17. doi: 10.1534/genetics.115.177717. PubMed PMID: 26178967; PubMed Central PMCID: PMCPMC4566278.

27. Branco S, Carpentier F, Rodriguez de la Vega RC, Badouin H, Snirc A, Le Prieur S, et al. Multiple convergent supergene evolution events in mating-type chromosomes. Nat Commun. 2018;9(1):2000. Epub 2018/05/23. doi: 10.1038/s41467-018-04380-9. PubMed PMID: 29784936; PubMed Central PMCID: PMCPMC5962589.

28. Nieuwenhuis BP, Billiard S, Vuilleumier S, Petit E, Hood ME, Giraud T. Evolution of uni- and bifactorial sexual compatibility systems in fungi. Heredity. 2013;111(6):445-55. Epub 2013/07/11. doi: 10.1038/hdy.2013.67. PubMed PMID: 23838688; PubMed Central PMCID: PMCPMC3833681.

29. James TY. Why mushrooms have evolved to be so promiscuous: Insights from evolutionary and ecological patterns. Fungal Biol Rev. 2015;29(3):167-78. doi: https://doi.org/10.1016/j.fbr.2015.10.002.

30. Taniguti LM, Schaker PD, Benevenuto J, Peters LP, Carvalho G, Palhares A, et al. Complete genome sequence of *Sporisorium scitamineum* and biotrophic interaction transcriptome with sugarcane. PLOS ONE. 2015;10(6):e0129318. Epub 20150612. doi: 10.1371/journal.pone.0129318. PubMed PMID: 26065709; PubMed Central PMCID: PMCPMC4466345.

31. Depotter JRL, Okmen B, Ebert MK, Beckers J, Kruse J, Thines M, et al. High nucleotide substitution rates associated with retrotransposon proliferation drive dynamic secretome evolution in smut pathogens. Microbiol Spectr. 2022;10(5):e0034922. Epub 20220816. doi: 10.1128/spectrum.00349-22. PubMed PMID: 35972267; PubMed Central PMCID: PMCPMC9603552.

32. Carpentier F, Rodriguez de la Vega RC, Branco S, Snirc A, Coelho MA, Hood ME, et al. Convergent recombination cessation between mating-type genes and centromeres in selfing anther-smut fungi. Genome Res. 2019;29(6):944-53. Epub 20190501. doi: 10.1101/gr.242578.118. PubMed PMID: 31043437; PubMed Central PMCID: PMCPMC6581054.

33. Duhamel M, Carpentier F, Begerow D, Hood ME, Rodriguez de la Vega RC, Giraud T. Onset and stepwise extensions of recombination suppression are common in mating-type chromosomes of *Microbotryum* anther-smut fungi. J Evol Biol. 2022;35(12):1619-34. Epub 20220310. doi: 10.1111/jeb.13991. PubMed PMID: 35271741; PubMed Central PMCID: PMCPMC10078771.

34. Fraser JA, Diezmann S, Subaran RL, Allen A, Lengeler KB, Dietrich FS, et al. Convergent evolution of chromosomal sex-determining regions in the animal and fungal kingdoms. PLOS Biol. 2004;2(12):e384. doi: 10.1371/journal.pbio.0020384. PubMed PMID: 15538538; PubMed Central PMCID: PMCPMC526376.

35. Sun S, Yadav V, Billmyre RB, Cuomo CA, Nowrousian M, Wang L, et al. Fungal genome and mating system transitions facilitated by chromosomal translocations involving intercentromeric recombination. PLOS Biol. 2017;15(8):e2002527. Epub 2017/08/12. doi: 10.1371/journal.pbio.2002527. PubMed PMID: 28800596; PubMed Central PMCID: PMCPMC5568439.

36. Sun S, Coelho MA, David-Palma M, Priest SJ, Heitman J. The evolution of sexual reproduction and the mating-type locus: links to pathogenesis of *Cryptococcus* human pathogenic fungi. Annu Rev Genet. 2019;53(1):417-44. doi: 10.1146/annurev-genet-120116-024755. PubMed PMID: 31537103.

37. Branco S, Badouin H, Rodriguez de la Vega RC, Gouzy J, Carpentier F, Aguileta G, et al. Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism. Proc Natl Acad Sci U S A. 2017;114(27):7067-72. Epub 2017/06/21. doi: 10.1073/pnas.1701658114. PubMed PMID: 28630332; PubMed Central PMCID: PMCPMC5502610.

38. Hartmann FE, Duhamel M, Carpentier F, Hood ME, Foulongne-Oriol M, Silar P, et al. Recombination suppression and evolutionary strata around mating-type loci in fungi: documenting patterns and understanding evolutionary and mechanistic causes. New Phytol. 2021;229(5):2470-91. Epub 2020/10/29. doi: 10.1111/nph.17039. PubMed PMID: 33113229; PubMed Central PMCID: PMCPMC7898863.

39. Hagen F, Khayhan K, Theelen B, Kolecka A, Polacheck I, Sionov E, et al. Recognition of seven species in the *Cryptococcus gattii*/*Cryptococcus neoformans* species complex. Fungal Genetics and Biology. 2015;78:16-48. doi: https://doi.org/10.1016/j.fgb.2015.02.009.

40. Farrer RA, Chang M, Davis MJ, van Dorp L, Yang D-H, Shea T, et al. A new lineage of *Cryptococcus gattii* (VGV) discovered in the Central Zambezian Miombo Woodlands. mBio. 2019;10(6):e02306-19. doi: 10.1128/mBio.02306-19.

41. Fisher MC, Denning DW. The WHO fungal priority pathogens list as a game-changer. Nat Rev Microbiol. 2023;21(4):211-2. doi: 10.1038/s41579-023-00861-x. PubMed PMID: 36747091; PubMed Central PMCID: PMCPMC9901396.

42. Casalini G, Giacomelli A, Antinori S. The WHO fungal priority pathogens list: a crucial reappraisal to review the prioritisation. Lancet Microbe. 2024;5(7):717-24. Epub 20240409. doi: 10.1016/S2666-5247(24)00042-9. PubMed PMID: 38608682.

43. Ginns J, Bernicchia A. *Filobasidiella lutea*: parasitism of *Hypochnicium vellereum*. Karstenia. 2000;40:49-51.

44. Findley K, Rodriguez-Carres M, Metin B, Kroiss J, Fonseca A, Vilgalys R, et al. Phylogeny and phenotypic characterization of pathogenic *Cryptococcus* species and closely related saprobic taxa in the Tremellales. Eukaryot Cell. 2009;8(3):353-61. Epub 2009/01/20. doi: 10.1128/EC.00373-08. PubMed PMID: 19151324; PubMed Central PMCID: PMCPMC2653247.

45. Liu XZ, Wang QM, Goker M, Groenewald M, Kachalkin AV, Lumbsch HT, et al. Towards an integrated phylogenetic classification of the Tremellomycetes. Stud Mycol. 2015;81:85-147. doi: 10.1016/j.simyco.2015.12.001. PubMed PMID: 26955199; PubMed Central PMCID: PMCPMC4777781.

46. Passer AR, Coelho MA, Billmyre RB, Nowrousian M, Mittelbach M, Yurkov AM, et al. Genetic and genomic analyses reveal boundaries between species closely related to *Cryptococcus* pathogens. mBio. 2019;10(3):e00764-19. doi: 10.1128/mBio.00764-19. PubMed PMID: 31186317; PubMed Central PMCID: PMCPMC6561019.

47. Kachalkin AV, Turchetti B, Inácio J, Carvalho C, Mašínová T, Pontes A, et al. Rare and undersampled dimorphic basidiomycetes. Mycol Prog. 2019. doi: 10.1007/s11557-019-01491-5.

48. Coelho MA, David-Palma M, Shea T, Bowers K, McGinley-Smith S, Mohammad AW, et al. Comparative genomics of the closely related fungal genera *Cryptococcus* and *Kwoniella* reveals karyotype dynamics and suggests evolutionary mechanisms of pathogenesis. PLOS Biol. 2024;22(6):e3002682. Epub 2024/06/06. doi: 10.1371/journal.pbio.3002682. PubMed PMID: 38843310; PubMed Central PMCID: PMCPMC11185503 declared that no competing interests exist.

49. Coelho MA, David-Palma M, Aylward J, Pham NQ, Visagie CM, Fuchs T, et al. Decoding *Cryptococcus*: from African biodiversity to worldwide prevalence. PLOS Pathogens. 2025;21(2):e1012876. doi: 10.1371/journal.ppat.1012876.

50. Chen R, Jiang YM, Wei SC, Wang QM. *Kwoniella shandongensis* sp. nov., a basidiomycetous yeast isolated from soil and bark from an apple orchard. Int J Syst Evol Microbiol. 2012;62(Pt 11):2774-7. Epub 2012/03/27. doi: 10.1099/ijs.0.039172-0. PubMed PMID: 22447698.

51. Statzell-Tallman A, Belloch C, Fell JW. *Kwoniella mangroviensis* gen. nov., sp.nov. (*Tremellales*, *Basidiomycota*), a teleomorphic yeast from mangrove habitats in the Florida Everglades and Bahamas. FEMS Yeast Res. 2008;8(1):103-13. Epub 2007/10/27. doi: 10.1111/j.1567-1364.2007.00314.x. PubMed PMID: 17961172.

52. Wang L, Zhai B, Lin X. The link between morphotype transition and virulence in *Cryptococcus neoformans*. PLOS Pathog. 2012;8(6):e1002765. Epub 20120621. doi: 10.1371/journal.ppat.1002765. PubMed PMID: 22737071; PubMed Central PMCID: PMCPMC3380952.

53. Heitman J, Carter DA, Dyer PS, Soll DR. Sexual reproduction of human fungal pathogens. Cold Spring Harb Perspect Med. 2014;4(8):a019281. Epub 2014/08/03. doi: 10.1101/cshperspect.a019281. PubMed PMID: 25085958; PubMed Central PMCID: PMCPMC4109574.

54. Zhao Y, Lin X. *Cryptococcus neoformans*: sex, morphogenesis, and virulence. Infect Genet Evol. 2021;89:104731. Epub 20210123. doi: 10.1016/j.meegid.2021.104731. PubMed PMID: 33497839; PubMed Central PMCID: PMCPMC8092418.

55. Phadke SS, Feretzaki M, Heitman J. Unisexual reproduction enhances fungal competitiveness by promoting habitat exploration via hyphal growth and sporulation. Eukaryot Cell. 2013;12(8):1155-9. Epub 20130621. doi: 10.1128/EC.00147-13. PubMed PMID: 23794511; PubMed Central PMCID: PMCPMC3754539.

56. Fu C, Thielhelm TP, Heitman J. Unisexual reproduction promotes competition for mating partners in the global human fungal pathogen *Cryptococcus deneoformans*. PLOS Genet. 2019;15(9):e1008394. Epub 2019/09/20. doi: 10.1371/journal.pgen.1008394. PubMed PMID: 31536509; PubMed Central PMCID: PMCPMC6772093.

57. Wang L, Lin X. The morphotype heterogeneity in *Cryptococcus neoformans*. Curr Opin Microbiol. 2015;26:60-4. Epub 20150619. doi: 10.1016/j.mib.2015.06.003. PubMed PMID: 26094087.

58. Fu MS, Liporagi-Lopes LC, Dos Santos SRJ, Tenor JL, Perfect JR, Cuomo CA, et al. Amoeba predation of *Cryptococcus* neoformans results in pleiotropic changes to traits associated with virulence. mBio. 2021;12(2). Epub 20210427. doi: 10.1128/mBio.00567-21. PubMed PMID: 33906924; PubMed Central PMCID: PMCPMC8092252.

59. Sauters TJC, Roth C, Murray D, Sun S, Floyd Averette A, Onyishi CU, et al. Amoeba predation of *Cryptococcus*: A quantitative and population genomic evaluation of the accidental pathogen hypothesis. PLOS Pathog. 2023;19(11):e1011763. Epub 20231113. doi: 10.1371/journal.ppat.1011763. PubMed PMID: 37956179; PubMed Central PMCID: PMCPMC10681322.

60. Idnurm A. Isolation of a fungal calcineurin A mutant suggests that amoebae can counter-select virulence attributes of microbes. Med Mycol. 2023;61(2). doi: 10.1093/mmy/myad013. PubMed PMID: 36708172; PubMed Central PMCID: PMCPMC9927882.

61. Price CTD, Hanford HE, Al-Quadan T, Santic M, Shin CJ, Da'as MSJ, et al. Amoebae as training grounds for microbial pathogens. mBio. 2024;15(8):e0082724. Epub 20240708. doi: 10.1128/mbio.00827-24. PubMed PMID: 38975782; PubMed Central PMCID: PMCPMC11323580.

62. Velagapudi R, Hsueh YP, Geunes-Boyer S, Wright JR, Heitman J. Spores as infectious propagules of *Cryptococcus neoformans*. Infect Immun. 2009;77(10):4345-55. Epub 2009/07/22. doi: 10.1128/IAI.00542-09. PubMed PMID: 19620339; PubMed Central PMCID: PMCPMC2747963.

63. Walsh NM, Botts MR, McDermott AJ, Ortiz SC, Wuthrich M, Klein B, et al. Infectious particle identity determines dissemination and disease outcome for the inhaled human fungal pathogen *Cryptococcus*. PLOS Pathog. 2019;15(6):e1007777. Epub 2019/06/28. doi: 10.1371/journal.ppat.1007777. PubMed PMID: 31247052; PubMed Central PMCID: PMCPMC6597114.

64. Ortiz SC, Hull CM. Biogenesis, germination, and pathogenesis of *Cryptococcus* spores. Microbiol Mol Biol Rev. 2024;88(1):e0019623. Epub 20240305. doi: 10.1128/mmbr.00196-23. PubMed PMID: 38440970; PubMed Central PMCID: PMCPMC10966950.

65. Guerreiro MA, Springer DJ, Rodrigues JA, Rusche LN, Findley K, Heitman J, et al. Molecular and genetic evidence for a tetrapolar mating system in the basidiomycetous yeast *Kwoniella mangrovensis* and two novel sibling species. Eukaryot Cell. 2013;12(5):746-60. Epub 2013/03/26. doi: 10.1128/ec.00065-13. PubMed PMID: 23524993; PubMed Central PMCID: PMCPmc3647770.

66. Metin B, Findley K, Heitman J. The mating type locus (*MAT*) and sexual reproduction of *Cryptococcus heveanensis*: insights into the evolution of sex and sex-determining chromosomal regions in fungi. PLOS Genet. 2010;6(5):e1000961. Epub 2010/05/27. doi: 10.1371/journal.pgen.1000961. PubMed PMID: 20502678; PubMed Central PMCID: PMCPMC2873909.

67. Cheng T, Veselská T, Křížková B, Švec K, Havlíček V, Stadler M, et al. Insight into the genomes of dominant yeast symbionts of European spruce bark beetle, *Ips typographus*. Front Microbiol. 2023;14:1108975. Epub 20230403. doi: 10.3389/fmicb.2023.1108975. PubMed PMID: 37077248; PubMed Central PMCID: PMCPMC10106607.

68. Basson JR, Roets F, Wingfield MJ, Aylward J. Bark beetles and their associated fungi infesting native *Widdringtonia* species in the Western Cape province of South Africa. Afr Entomol. 2024;32:1-8.

69. Crous PW, Luangsa-Ard JJ, Wingfield MJ, Carnegie AJ, Hernandez-Restrepo M, Lombard L, et al. Fungal Planet description sheets: 785-867. Persoonia. 2018;41:238-417. Epub 2019/02/08. doi: 10.3767/persoonia.2018.41.12. PubMed PMID: 30728607; PubMed Central PMCID: PMCPMC6344811.

70. Li AH, Yuan FX, Groenewald M, Bensch K, Yurkov AM, Li K, et al. Diversity and phylogeny of basidiomycetous yeasts from plant leaves and soil: Proposal of two new orders, three new families, eight new genera and one hundred and seven new species. Stud Mycol. 2020;96:17-140. Epub 20200128. doi: 10.1016/j.simyco.2020.01.002. PubMed PMID: 32206137; PubMed Central PMCID: PMCPMC7082220.

71. Leufvén A, Bergström G, Falsen E. Interconversion of verbenols and verbenone by identified yeasts isolated from the spruce bark beetle *Ips typographus*. J Chem Ecol. 1984;10(9):1349-61. doi: 10.1007/BF00988116. PubMed PMID: 24317586.

72. Yurkov AM, Rohl O, Pontes A, Carvalho C, Maldonado C, Sampaio JP. Local climatic conditions constrain soil yeast diversity patterns in Mediterranean forests, woodlands and scrub biome. FEMS Yeast Res. 2016;16(1):fov103. doi: 10.1093/femsyr/fov103. PubMed PMID: 26568202.

73. Kwon-Chung KJ. A new genus, *Filobasidiella*, the perfect state of *Cryptococcus neoformans*. Mycologia. 1975;67:1197-200.

74. Kwon-Chung KJ. A new species of *Filobasidiella*, the sexual state of *Cryptococcus neoformans* B and C serotypes. Mycologia 1976;68:943-6.

75. Kwon-Chung KJ. Morphogenesis of *Filobasidiella neoformans*, the sexual state of *Cryptococcus neoformans*. Mycologia. 1976;68:821-33.

76. Rodriguez-Carres M, Findley K, Sun S, Dietrich FS, Heitman J. Morphological and genomic characterization of *Filobasidiella depauperata*: a homothallic sibling species of the pathogenic *Cryptococcus* species complex. PLOS ONE. 2010;5(3):e9620. doi: 10.1371/journal.pone.0009620.

77. Findley K, Sun S, Fraser JA, Hsueh Y-P, Averette AF, Li W, et al. Discovery of a modified tetrapolar sexual cycle in *Cryptococcus amylolentus* and the evolution of *MAT* in the *Cryptococcus* species complex. PLOS Genet. 2012;8(2):e1002528. doi: 10.1371/journal.pgen.1002528.

78. Sun S, Priest SJ, Heitman J. *Cryptococcus neoformans* mating and genetic crosses. Curr Protoc Microbiol. 2019:e75. Epub 2019/01/21. doi: 10.1002/cpmc.75. PubMed PMID: 30661293.

79. Raper J, Flexer A. Mating systems and evolution of the basidiomycetes. Evolution in the higher basidiomycetes. Knoxville, TN: University of Tennessee Press; 1971. p. 149-67.

80. Fraser JA, Hsueh Y-P, Findley KM, Heitman J. Evolution of the mating-type locus: the basidiomycetes. Sex in Fungi2007. p. 19-34.

81. Phadke SS, Feretzaki M, Clancey SA, Mueller O, Heitman J. Unisexual reproduction of *Cryptococcus gattii*. PLOS ONE. 2014;9(10):e111089. Epub 2014/10/23. doi: 10.1371/journal.pone.0111089. PubMed PMID: 25337713; PubMed Central PMCID: PMCPMC4206507.

82. Chang YC, Wickes BL, Miller GF, Penoyer LA, Kwon-Chung KJ. *Cryptococcus neoformans STE12*α regulates virulence but is not essential for mating. J Exp Med. 2000;191(5):871-82. doi: 10.1084/jem.191.5.871. PubMed PMID: 10704467; PubMed Central PMCID: PMCPMC2195848.

83. Yue C, Cavallo LM, Alspaugh JA, Wang P, Cox GM, Perfect JR, et al. The *STE12*α homolog is required for haploid filamentation but largely dispensable for mating and virulence in *Cryptococcus neoformans*. Genetics. 1999;153(4):1601-15. doi: 10.1093/genetics/153.4.1601. PubMed PMID: 10581270; PubMed Central PMCID: PMCPMC1460866.

84. Burger F, Daugeron MC, Linder P. Dbp10p, a putative RNA helicase from *Saccharomyces cerevisiae*, is required for ribosome biogenesis. Nucleic Acids Res. 2000;28(12):2315-23. doi: 10.1093/nar/28.12.2315. PubMed PMID: 10871363; PubMed Central PMCID: PMCPMC102738.

85. Parnell KM, Bass BL. Functional redundancy of yeast proteins Reh1 and Rei1 in cytoplasmic 60S subunit maturation. Mol Cell Biol. 2009;29(14):4014-23. Epub 20090511. doi: 10.1128/MCB.01582-08. PubMed PMID: 19433447; PubMed Central PMCID: PMCPMC2704747.

86. Mizushima N, Noda T, Yoshimori T, Tanaka Y, Ishii T, George MD, et al. A protein conjugation system essential for autophagy. Nature. 1998;395(6700):395-8. doi: 10.1038/26506. PubMed PMID: 9759731.

87. Zhu YH, Hyun J, Pan YZ, Hopper JE, Rizo J, Wu JQ. Roles of the fission yeast UNC-13/Munc13 protein Ync13 in late stages of cytokinesis. Mol Biol Cell. 2018;29(19):2259-79. Epub 20180725. doi: 10.1091/mbc.E18-04-0225. PubMed PMID: 30044717; PubMed Central PMCID: PMCPMC6249806.

88. Wigge PA, Kilmartin JV. The Ndc80p complex from *Saccharomyces cerevisiae* contains conserved centromere components and has a function in chromosome segregation. J Cell Biol. 2001;152(2):349-60. doi: 10.1083/jcb.152.2.349. PubMed PMID: 11266451; PubMed Central PMCID: PMCPMC2199619.

89. Guterman A, Glickman MH. Complementary roles for Rpn11 and Ubp6 in deubiquitination and proteolysis by the proteasome. J Biol Chem. 2004;279(3):1729-38. Epub 20031027. doi: 10.1074/jbc.M307050200. PubMed PMID: 14581483.

90. Sridhar S, Hori T, Nakagawa R, Fukagawa T, Sanyal K. Bridgin connects the outer kinetochore to centromeric chromatin. Nat Commun. 2021;12(1):146. Epub 20210108. doi: 10.1038/s41467-020-20161-9. PubMed PMID: 33420015; PubMed Central PMCID: PMCPMC7794384.

91. Billmyre RB, Craig CJ, Lyon J, Reichardt C, Eickbush MT, Zanders SE. Saturation transposon mutagenesis enables genome-wide identification of genes required for growth and fluconazole resistance in the human fungal pathogen *Cryptococcus neoformans*. bioRxiv. 2024:2024.07.28.605507. doi: 10.1101/2024.07.28.605507.

92. Loftus BJ, Fung E, Roncaglia P, Rowley D, Amedeo P, Bruno D, et al. The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. Science. 2005;307(5713):1321-4. Epub 2005/01/18. doi: 10.1126/science.1103773. PubMed PMID: 15653466; PubMed Central PMCID: PMCPMC3520129.

93. Huang J, Larmore CJ, Priest SJ, Xu Z, Dietrich FS, Yadav V, et al. Distinct evolutionary trajectories following loss of RNA interference in *Cryptococcus neoformans*. Proc Natl Acad Sci U S A. 2024;121(47):e2416656121. Epub 20241113. doi: 10.1073/pnas.2416656121. PubMed PMID: 39536081; PubMed Central PMCID: PMCPMC11588098.

94. Desjardins CA, Giamberardino C, Sykes SM, Yu CH, Tenor JL, Chen Y, et al. Population genomics and the evolution of virulence in the fungal pathogen *Cryptococcus neoformans*. Genome Res. 2017;27(7):1207-19. Epub 2017/06/15. doi: 10.1101/gr.218727.116. PubMed PMID: 28611159; PubMed Central PMCID: PMCPMC5495072.

95. Lengeler KB, Wang P, Cox GM, Perfect JR, Heitman J. Identification of the *MAT***a** mating-type locus of *Cryptococcus neoformans* reveals a serotype A *MAT***a** strain thought to have been extinct. Proc Natl Acad Sci U S A. 2000;97(26):14455-60. doi: 10.1073/pnas.97.26.14455. PubMed PMID: 11121047; PubMed Central PMCID: PMCPMC18940.

96. Viviani MA, Esposto MC, Cogliati M, Montagna MT, Wickes BL. Isolation of a *Cryptococcus neoformans* serotype A *MAT***a** strain from the Italian environment. Med Mycol. 2001;39(5):383-6. doi: 10.1080/mmy.39.5.383.386. PubMed PMID: 12054047.

97. Viviani MA, Nikolova R, Esposto MC, Prinz G, Cogliati M. First European case of serotype A *MAT***a** *Cryptococcus neoformans* infection. Emerg Infect Dis. 2003;9(9):1179-80. doi: 10.3201/eid0909.020770. PubMed PMID: 14528882; PubMed Central PMCID: PMCPMC3016796.

98. Litvintseva AP, Marra RE, Nielsen K, Heitman J, Vilgalys R, Mitchell TG. Evidence of sexual recombination among *Cryptococcus neoformans* serotype A isolates in sub-Saharan Africa. Eukaryot Cell. 2003;2(6):1162-8. doi: 10.1128/EC.2.6.1162-1168.2003. PubMed PMID: 14665451; PubMed Central PMCID: PMCPMC326657.

99. Ashton PM, Thanh LT, Trieu PH, Van Anh D, Trinh NM, Beardsley J, et al. Three phylogenetic groups have driven the recent population expansion of *Cryptococcus neoformans*. Nat Commun. 2019;10(1):2035. Epub 20190502. doi: 10.1038/s41467-019-10092-5. PubMed PMID: 31048698; PubMed Central PMCID: PMCPMC6497710.

100. Yadav V, Sun S, Heitman J. Uniparental nuclear inheritance following bisexual mating in fungi. eLife. 2021;10. Epub 2021/08/03. doi: 10.7554/eLife.66234. PubMed PMID: 34338631; PubMed Central PMCID: PMCPMC8412948.

101. Hsueh YP, Idnurm A, Heitman J. Recombination hotspots flank the *Cryptococcus* mating-type locus: implications for the evolution of a fungal sex chromosome. PLOS Genet. 2006;2(11):e184. Epub 2006/11/07. doi: 10.1371/journal.pgen.0020184. PubMed PMID: 17083277; PubMed Central PMCID: PMCPMC1630710.

102. Drillon G, Carbone A, Fischer G. SynChro: a fast and easy tool to reconstruct and visualize synteny blocks along eukaryotic chromosomes. PLOS ONE. 2014;9(3):e92621. Epub 2014/03/22. doi:

10.1371/journal.pone.0092621. PubMed PMID: 24651407; PubMed Central PMCID: PMCPMC3961402.

103. D'Souza CA, Kronstad JW, Taylor G, Warren R, Yuen M, Hu G, et al. Genome variation in *Cryptococcus gattii*, an emerging pathogen of immunocompetent hosts. MBio. 2011;2(1):e00342-10. Epub 2011/02/10. doi: 10.1128/mBio.00342-10. PubMed PMID: 21304167; PubMed Central PMCID: PMCPMC3037005.

104. Morrow CA, Lee IR, Chow EW, Ormerod KL, Goldinger A, Byrnes EJ, 3rd, et al. A unique chromosomal rearrangement in the *Cryptococcus neoformans* var. *grubii* type strain enhances key phenotypes associated with virulence. MBio. 2012;3(2). doi: 10.1128/mBio.00310-11. PubMed PMID: 22375073; PubMed Central PMCID: PMCPMC3302566.

105. Yadav V, Sun S, Coelho MA, Heitman J. Centromere scission drives chromosome shuffling and reproductive isolation. Proc Natl Acad Sci U S A. 2020;117(14):7917-28. Epub 2020/03/21. doi: 10.1073/pnas.1918659117. PubMed PMID: 32193338; PubMed Central PMCID: PMCPMC7149388.

106. Janbon G, Ormerod KL, Paulet D, Byrnes EJ, 3rd, Yadav V, Chatterjee 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. 2014;10(4):e1004261. Epub 2014/04/20. doi: 10.1371/journal.pgen.1004261. PubMed PMID: 24743168; PubMed Central PMCID: PMCPMC3990503.

107. Yadav V, Sun S, Billmyre RB, Thimmappa BC, Shea T, Lintner R, et al. RNAi is a critical determinant of centromere evolution in closely related fungi. Proc Natl Acad Sci U S A. 2018. Epub 2018/03/07. doi: 10.1073/pnas.1713725115. PubMed PMID: 29507212.

108. Urban M, Kahmann R, Bolker M. The biallelic *a* mating type locus of *Ustilago maydis*: remnants of an additional pheromone gene indicate evolution from a multiallelic ancestor. Mol Gen Genet. 1996;250(4):414-20. Epub 1996/03/07. doi: 10.1007/BF02174029. PubMed PMID: 8602158.

109. Fedler M, Luh KS, Stelter K, Nieto-Jacobo F, Basse CW. The *a2* mating-type locus genes *lga2* and *rga2* direct uniparental mitochondrial DNA (mtDNA) inheritance and constrain mtDNA recombination during sexual development of *Ustilago maydis*. Genetics. 2009;181(3):847-60. Epub 2008/12/24. doi: 10.1534/genetics.108.096859. PubMed PMID: 19104076; PubMed Central PMCID: PMCPMC2651059.

110. Coelho MA, Gonçalves P, Sampaio JP. Evidence for maintenance of sex determinants but not of sexual stages in red yeasts, a group of early diverged basidiomycetes. BMC Evol Biol. 2011;11(1):249. doi: 10.1186/1471-2148-11-249.

111. Huang J, Hu P, Ye L, Shen Z, Chen X, Liu F, et al. Pan-drug resistance and hypervirulence in a human fungal pathogen are enabled by mutagenesis induced by mammalian body temperature. Nat Microbiol. 2024;9(7):1686-99. Epub 2024/06/20. doi: 10.1038/s41564-024-01720-y. PubMed PMID: 38898217.

112. Sun S, Fu C, Ianiri G, Heitman J. The pheromone and pheromone receptor mating-type locus is involved in controlling uniparental mitochondrial inheritance in *Cryptococcus*. Genetics. 2020;214(3):703-17. Epub 2020/01/01. doi: 10.1534/genetics.119.302824. PubMed PMID: 31888949; PubMed Central PMCID: PMCPMC7054021.

113. Jackson JA, Fink GR. Gene conversion between duplicated genetic elements in yeast. Nature. 1981;292(5821):306-11. Epub 1981/07/23. doi: 10.1038/292306a0. PubMed PMID: 6265790.

114. Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW. The double-strand-break repair model for recombination. Cell. 1983;33(1):25-35. Epub 1983/05/01. doi: 10.1016/0092-8674(83)90331-8. PubMed PMID: 6380756.

115. Hawley RS. The human Y chromosome: rumors of its death have been greatly exaggerated. Cell. 2003;113(7):825-8. doi: 10.1016/s0092-8674(03)00470-7. PubMed PMID: 12837241.

116. Rozen S, Skaletsky H, Marszalek JD, Minx PJ, Cordum HS, Waterston RH, et al. Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. Nature. 2003;423(6942):873-6. doi: 10.1038/nature01723. PubMed PMID: 12815433.

117. Connallon T, Clark AG. Gene duplication, gene conversion and the evolution of the Y chromosome. Genetics. 2010;186(1):277-86. Epub 2010/06/17. doi: 10.1534/genetics.110.116756. PubMed PMID: 20551442; PubMed Central PMCID: PMCPMC2940292.

118. Camiolo S, Farina L, Porceddu A. The relation of codon bias to tissue-specific gene expression in *Arabidopsis thaliana*. Genetics. 2012;192(2):641-9. Epub 2012/08/07. doi: 10.1534/genetics.112.143677. PubMed PMID: 22865738; PubMed Central PMCID: PMCPMC3454886.

119. Zhou M, Guo J, Cha J, Chae M, Chen S, Barral JM, et al. Non-optimal codon usage affects expression, structure and function of clock protein FRQ. Nature. 2013;495(7439):111-5. Epub 2013/02/19. doi: 10.1038/nature11833. PubMed PMID: 23417067; PubMed Central PMCID: PMCPMC3629845.

120. Pessia E, Popa A, Mousset S, Rezvoy C, Duret L, Marais GA. Evidence for widespread GC-biased gene conversion in eukaryotes. Genome Biol Evol. 2012;4(7):675-82. Epub 2012/05/26. doi: 10.1093/gbe/evs052. PubMed PMID: 22628461; PubMed Central PMCID: PMCPMC5635611.

121. Marais G. Biased gene conversion: implications for genome and sex evolution. Trends Genet. 2003;19(6):330-8. Epub 2003/06/13. doi: 10.1016/S0168-9525(03)00116-1. PubMed PMID: 12801726.

122. Mugal CF, Weber CC, Ellegren H. GC-biased gene conversion links the recombination landscape and demography to genomic base composition: GC-biased gene conversion drives genomic base composition across a wide range of species. Bioessays. 2015;37(12):1317-26. Epub 2015/10/08. doi: 10.1002/bies.201500058. PubMed PMID: 26445215.

123. Byrnes EJ, 3rd, Li W, Ren P, Lewit Y, Voelz K, Fraser JA, et al. A diverse population of *Cryptococcus gattii* molecular type VGIII in southern Californian HIV/AIDS patients. PLOS Pathog. 2011;7(9):e1002205. Epub 2011/09/13. doi: 10.1371/journal.ppat.1002205. PubMed PMID: 21909264; PubMed Central PMCID: PMCPMC3164645.

124. Lee N, Bakkeren G, Wong K, Sherwood JE, Kronstad JW. The mating-type and pathogenicity locus of the fungus *Ustilago hordei* spans a 500-kb region. Proc Natl Acad Sci U S A. 1999;96(26):15026-31. Epub 1999/12/28. doi: 10.1073/pnas.96.26.15026. PubMed PMID: 10611332; PubMed Central PMCID: PMCPMC24767.

125. Liang SW, Huang YH, Chiu JY, Tseng HW, Huang JH, Shen WC. The smut fungus *Ustilago esculenta* has a bipolar mating system with three idiomorphs larger than 500 kb. Fungal Genet Biol. 2019;126:61-74. Epub 2019/02/23. doi: 10.1016/j.fgb.2019.02.007. PubMed PMID: 30794950.

126. Lucotte EA, Jay P, Rougemont Q, Boyer L, Cornille AA, Snirc A, et al. Repeated loss of function at HD mating-type genes and of recombination suppression without mating-type locus linkage in anther-smut fungi. bioRxiv. 2024. Epub 2024/03/18. doi: 10.1101/2024.03.03.583181. PubMed PMID: 38496565; PubMed Central PMCID: PMCPMC10942285.

127. Gostinčar C, Sun X, Zajc J, Fang C, Hou Y, Luo Y, et al. Population genomics of an obligately halophilic basidiomycete *Wallemia ichthyophaga*. Front Microbiol. 2019;10:2019. Epub 2019/09/26. doi: 10.3389/fmicb.2019.02019. PubMed PMID: 31551960; PubMed Central PMCID: PMCPMC6738226.

128. Priest SJ, Yadav V, Roth C, Dahlmann TA, Kück U, Magwene PM, et al. Uncontrolled transposition following RNAi loss causes hypermutation and antifungal drug resistance in clinical isolates of *Cryptococcus neoformans*. Nat Microbiol. 2022;7(8):1239-51. Epub 2022/08/03. doi: 10.1038/s41564-022-01183-z. PubMed PMID: 35918426.

129. Lemoine FJ, Degtyareva NP, Lobachev K, Petes TD. Chromosomal translocations in yeast induced by low levels of DNA polymerase a model for chromosome fragile sites. Cell. 2005;120(5):587-98. Epub 2005/03/16. doi: 10.1016/j.cell.2004.12.039. PubMed PMID: 15766523.

130. Hawksworth DL, Kirk PM, Sutton BC, Pelger DN. Ainsworth and Bisby's Dictionary of the fungi. 8th ed. Oxon, UK: CABI; 1995.

131. Nieuwenhuis BP, James TY. The frequency of sex in fungi. Philos Trans R Soc Lond B Biol Sci. 2016;371(1706). doi: 10.1098/rstb.2015.0540. PubMed PMID: 27619703; PubMed Central PMCID: PMCPMC5031624.

132. Sexton AC, Howlett BJ. Parallels in fungal pathogenesis on plant and animal hosts. Eukaryot Cell. 2006;5(12):1941-9. Epub 20061013. doi: 10.1128/EC.00277-06. PubMed PMID: 17041185; PubMed Central PMCID: PMCPMC1694825.

133. Morrow CA, Fraser JA. Sexual reproduction and dimorphism in the pathogenic basidiomycetes. FEMS Yeast Res. 2009;9(2):161-77. Epub 2009/02/18. doi: 10.1111/j.1567-1364.2008.00475.x. PubMed PMID: 19220864.

134. Nielsen K, Cox GM, Wang P, Toffaletti DL, Perfect JR, Heitman J. Sexual cycle of *Cryptococcus neoformans* var. *grubii* and virulence of congenic **a** and α isolates. Infect Immun. 2003;71(9):4831-41. Epub 2003/08/23. PubMed PMID: 12933823; PubMed Central PMCID: PMCPMC187335.

135. Wang P, Cutler J, King J, Palmer D. Mutation of the regulator of G protein signaling Crg1 increases virulence in *Cryptococcus neoformans*. Eukaryot Cell. 2004;3(4):1028-35. doi: 10.1128/EC.3.4.1028-1035.2004. PubMed PMID: 15302835; PubMed Central PMCID: PMCPMC500888.

136. Fraser JA, Subaran RL, Nichols CB, Heitman J. Recapitulation of the sexual cycle of the primary fungal pathogen *Cryptococcus neoformans* var. *gattii*: implications for an outbreak on Vancouver Island, Canada. Eukaryot Cell. 2003;2(5):1036-45. doi: 10.1128/ec.2.5.1036-1045.2003. PubMed PMID: 14555486.

137. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 2017;27(5):722-36. Epub 2017/03/17. doi: 10.1101/gr.215087.116. PubMed PMID: 28298431; PubMed Central PMCID: PMCPMC5411767.

138. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLOS ONE. 2014;9(11):e112963. Epub 2014/11/20. doi: 10.1371/journal.pone.0112963. PubMed PMID: 25409509; PubMed Central PMCID: PMCPMC4237348.

139. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60. Epub 2009/05/20. doi: 10.1093/bioinformatics/btp324. PubMed PMID: 19451168; PubMed Central PMCID: PMCPMC2705234.

140. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34(18):3094-100. Epub 2018/05/12. doi: 10.1093/bioinformatics/bty191. PubMed PMID: 29750242; PubMed Central PMCID: PMCPMC6137996.

141. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24-6. Epub 2011/01/12. doi: 10.1038/nbt.1754. PubMed PMID: 21221095; PubMed Central PMCID: PMCPMC3346182.

142. Gnerre S, Maccallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, et al. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc Natl Acad Sci U S A. 2011;108(4):1513-8. Epub 2010/12/29. doi: 10.1073/pnas.1017351108. PubMed PMID: 21187386; PubMed Central PMCID: PMCPMC3029755.

143. Bruna T, Hoff KJ, Lomsadze A, Stanke M, Borodovsky M. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. NAR Genom Bioinform. 2021;3(1):lqaa108. Epub 2021/02/13. doi: 10.1093/nargab/lqaa108. PubMed PMID: 33575650; PubMed Central PMCID: PMCPMC7787252.

144. Cantalapiedra CP, Hernandez-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol. 2021;38(12):5825-9. Epub 2021/10/02. doi: 10.1093/molbev/msab293. PubMed PMID: 34597405; PubMed Central PMCID: PMCPMC8662613.

145. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, et al. antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res. 2021;49(W1):W29-W35. Epub 2021/05/13. doi: 10.1093/nar/gkab335. PubMed PMID: 33978755; PubMed Central PMCID: PMCPMC8262755.

146. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772-80. Epub 2013/01/19. doi: 10.1093/molbev/mst010. PubMed PMID: 23329690; PubMed Central PMCID: PMCPMC3603318.

147. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25(15):1972-3. Epub 2009/06/10. doi: 10.1093/bioinformatics/btp348. PubMed PMID: 19505945; PubMed Central PMCID: PMCPmc2712344.

148. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020;37(5):1530-4. Epub 2020/02/06. doi: 10.1093/molbev/msaa015. PubMed PMID: 32011700; PubMed Central PMCID: PMCPMC7182206.

149. Zhang C, Rabiee M, Sayyari E, Mirarab S. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics. 2018;19(Suppl 6):153. Epub 2018/05/11. doi: 10.1186/s12859-018-2129-y. PubMed PMID: 29745866; PubMed Central PMCID: PMCPMC5998893.

150. Chernomor O, von Haeseler A, Minh BQ. Terrace aware data structure for phylogenomic inference from supermatrices. Syst Biol. 2016;65(6):997-1008. Epub 2016/04/29. doi: 10.1093/sysbio/syw037. PubMed PMID: 27121966; PubMed Central PMCID: PMCPMC5066062.

151. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14(6):587-9. Epub 2017/05/10. doi: 10.1038/nmeth.4285. PubMed PMID: 28481363; PubMed Central PMCID: PMCPMC5453245.

152. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018;35(2):518-22. Epub 2017/10/28. doi: 10.1093/molbev/msx281. PubMed PMID: 29077904; PubMed Central PMCID: PMCPMC5850222.

153. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. Bioinformatics. 2011;27(7):1009-10. Epub 2011/02/01. doi: 10.1093/bioinformatics/btr039. PubMed PMID: 21278367; PubMed Central PMCID: PMCPMC3065679.

154. Ianiri G, Fang YF, Dahlmann TA, Clancey SA, Janbon G, Kuck U, et al. Mating-type-specific ribosomal proteins control aspects of sexual reproduction in *Cryptococcus neoformans*. Genetics. 2020;214(3):635-49. Epub 2019/12/29. doi: 10.1534/genetics.119.302740. PubMed PMID: 31882399; PubMed Central PMCID: PMCPMC7054023.

155. Bian Z, Xu Z, Peer A, Choi Y, Priest SJ, Akritidou K, et al. Essential genes encoded by the mating-type locus of the human fungal pathogen *Cryptococcus neoformans*. bioRxiv. 2024. Epub 2024/12/16. doi: 10.1101/2024.12.02.626420. PubMed PMID: 39677606; PubMed Central PMCID: PMCPMC11642766.

156. Cabanettes F, Klopp C. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. PeerJ. 2018;6:e4958. Epub 2018/06/12. doi: 10.7717/peerj.4958. PubMed PMID: 29888139; PubMed Central PMCID: PMCPMC5991294.

157. Langdon QK, Peris D, Kyle B, Hittinger CT. sppIDer: A Species Identification Tool to Investigate Hybrid Genomes with High-Throughput Sequencing. Mol Biol Evol. 2018;35(11):2835-49. Epub 2018/09/06. doi: 10.1093/molbev/msy166. PubMed PMID: 30184140; PubMed Central PMCID: PMCPMC6231485.

158. Fu C, Heitman J. *PRM1* and *KAR5* function in cell-cell fusion and karyogamy to drive distinct bisexual and unisexual cycles in the *Cryptococcus pathogenic* species complex. PLOS Genet. 2017;13(11):e1007113. Epub 2017/11/28. doi: 10.1371/journal.pgen.1007113. PubMed PMID: 29176784; PubMed Central PMCID: PMCPMC5720818.

159. Priest SJ, Coelho MA, Mixao V, Clancey SA, Xu Y, Sun S, et al. Factors enforcing the species boundary between the human pathogens *Cryptococcus neoformans* and *Cryptococcus deneoformans*. PLOS Genet. 2021;17(1):e1008871. Epub 2021/01/20. doi: 10.1371/journal.pgen.1008871. PubMed PMID: 33465111; PubMed Central PMCID: PMCPMC7846113.

Figures

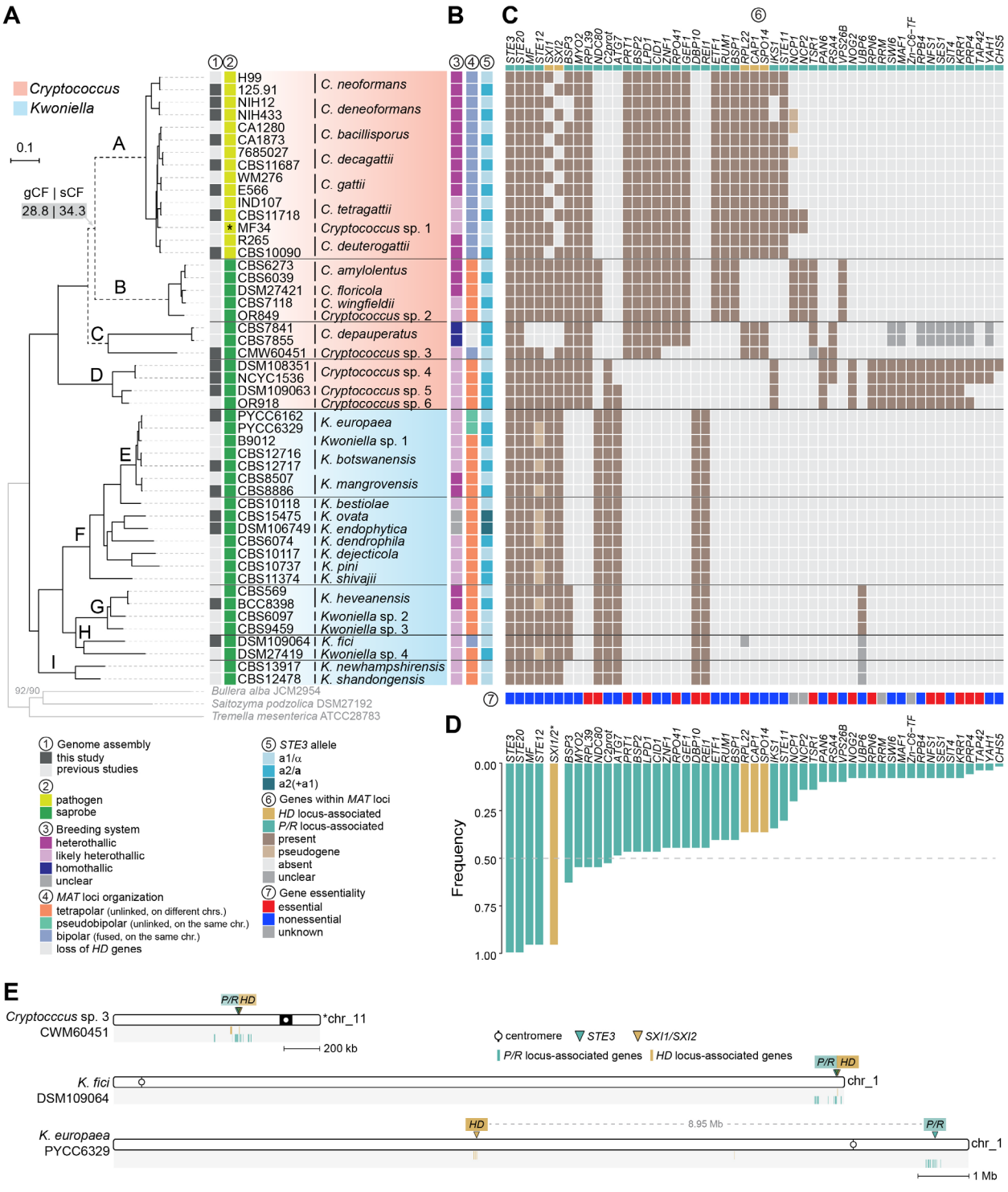


Fig 1. Phylogeny, breeding systems, and *MAT* loci organization in *Cryptococcus* and *Kwoniella*. (A) Phylogenetic tree of strains analyzed in this study, inferred by maximum likelihood analysis of a concatenated data matrix comprising protein alignments of 3,086 single-copy genes. Branch lengths are shown as the number of substitutions per site (scale bar). All branches are supported >95% based on SH-aLTR and UFboot tests, unless otherwise indicated. Gene concordance factor (gCF) and site concordance factor (sCF) were assessed to evaluate genealogical concordance across branches, revealing an ambiguous placement of clades A, B and C (dashed branches; see S1 Fig for details). An asterisk (*) denotes that VGV has not yet been associated with human infections. **(B)** Predicted breeding systems, *MAT* locus organization, and mating-type identity of the *P/R* locus. Breeding systems are classified as heterothallic or homothallic for species with defined sexual cycles, and as likely heterothallic for species inferred to exhibit heterothallism based on *MAT* gene content. **(C)** Heatmap of gene presence/absence and pseudogene annotations within the inferred *MAT* loci across species. Genes are ordered by their frequency of presence, with the most conserved genes on the left. Genes marked as “unclear” indicate instances where there is uncertainty regarding their presence within the *MAT* loci because these regions could not be precisely determined (see text for details). *BSP3* and *IKS1*, which were evicted from the *MAT* locus in *C. deneoformans* [34], are listed as absent in this species. Gene essentiality, as predicted or experimentally validated in strain H99 in prior studies rather than direct assessments in each species, is displayed at the bottom. **(D)** Gene presence frequency plot summarizing panel B. Pseudogenes were treated as present for frequency calculations, while unclear cases were treated as absent. The identification of either *SXI1* or *SXI2* was scored as “present” and combined for frequency quantification. **(E)** *MAT* loci locations in *Cryptococcus* sp. CWM60451, *K. fici*, and *K. europaea*, illustrating bipolar and pseudobipolar arrangements, respectively. Genes typically found either within or near the *HD* and *P/R* loci in tetrapolar species are designated as *HD*- and *P/R*-associated genes and their respective chromosomal locations are indicated by gold or teal vertical bars. The data underlying this Figure can be found in **S1 Appendix** and at https://doi.org/10.5281/zenodo.14851287.

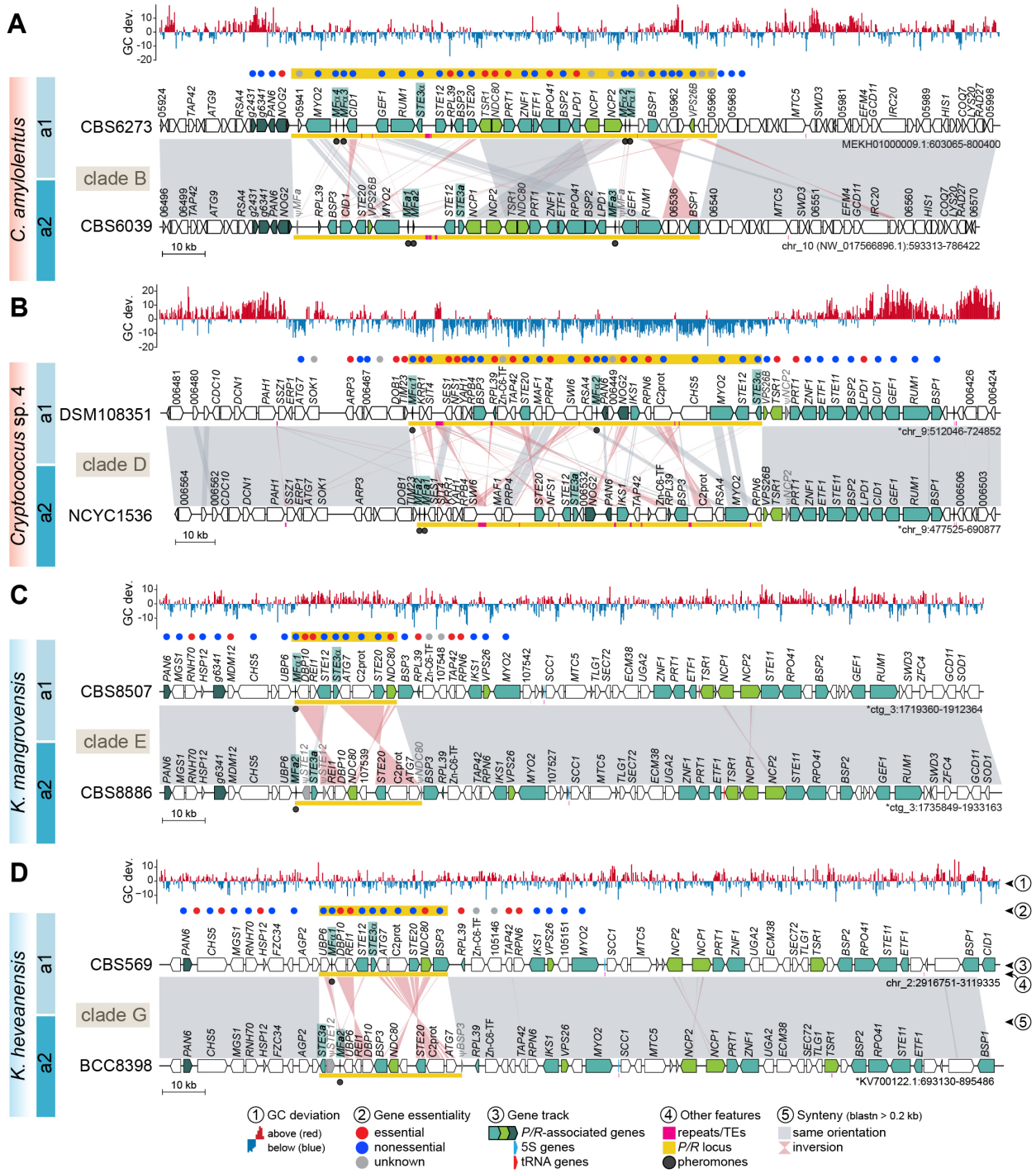


Fig 2. Synteny analysis of the *P/R* locus in tetrapolar *Cryptococcus* and *Kwoniella* species. (A-B) Detailed views of the *P/R* locus of representative *Cryptococcus* species from clades B and D. In both clades, the *P/R* locus exhibits lower GC content compared to other genomic regions. In clade D species, the predicted *P/R* locus contains only a subset of the genes found within the *P/R* locus of clade B species or the *MAT* locus of pathogenic *Cryptococcus* species. The additional genes in clade B or pathogenic species are instead located in the immediate downstream region, suggesting independent expansion of this genomic segment in clade D species. **(C-D)** *P/R* locus region in representative *Kwoniella* species of clades E and G.

In *a2* strains, the *STE12* gene appears truncated. Remnants of the *NDC80* gene (in clade E/F species) or the *BSP3* gene (in clade G species) are observed at the right edge of the *P/R a2* allele, likely resulting from inversion events. In all panels, *P/R*-associated genes are colored teal if their orthologs in *Cryptococcus* pathogens are located within the *MAT* locus, with darker teal indicating genes positioned in the flanking regions, and bright green for genes present within the *P/R* locus of *C. amylolentus*, but absent in most *Cryptococcus* pathogens. The *P/R* allele of each strain (*a1* or *a2*) is indicated on the left. Chromosomes inverted relative to their original assembly orientations are marked with asterisks. GC content is depicted as the deviation from the genome average, calculated in 0.5 kb non-overlapping windows. Gene essentiality within and in the immediate vicinity of the *P/R* locus is inferred from predictions or experimental validations in *C. neoformans* H99 rather than direct assessments in these species. Other features are annotated as shown in the key. See S4Fig for additional comparisons. The data underlying this Figure can be found at https://doi.org/10.5281/zenodo.14851287.

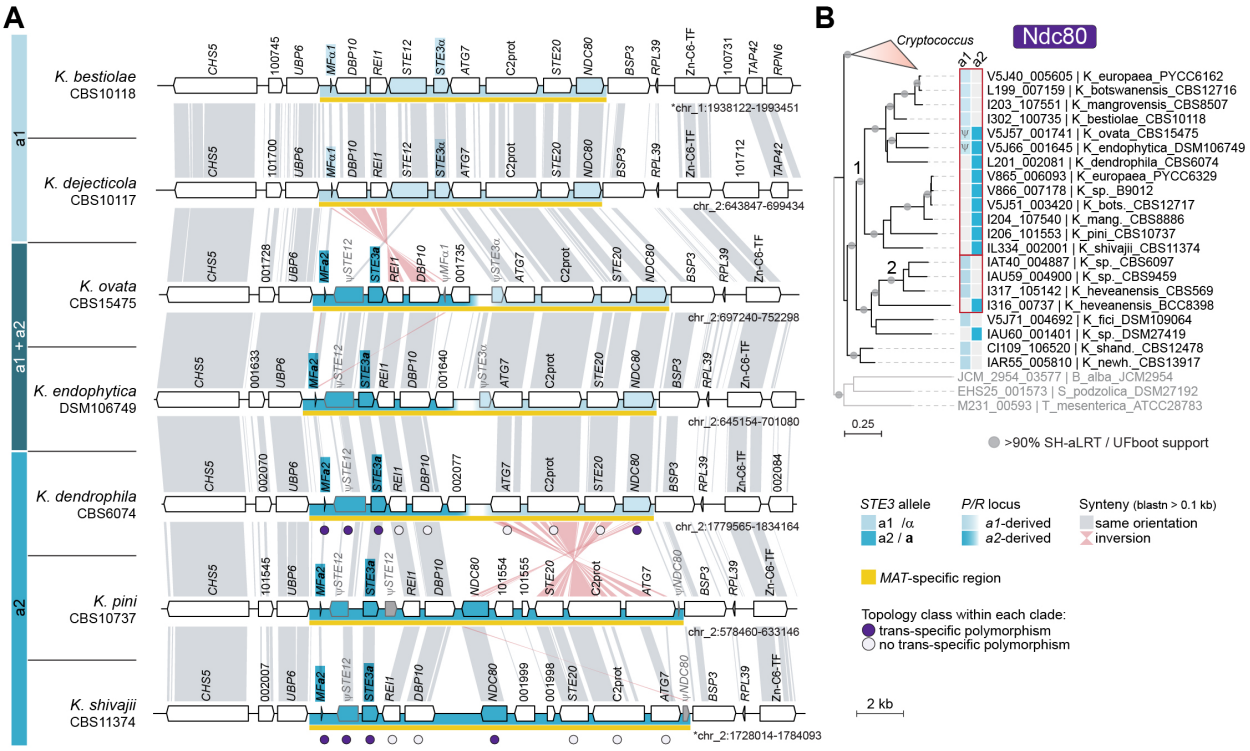


Fig 3. The *P/R* locus in *K. ovata*, *K. endophytica*, and *K. dendrophila* likely arose through intra-*P/R* recombination. **(A)** Synteny analysis comparing *P/R a1* and *P/R a2* alleles in select *Kwoniella* species from clade F. The *P/R* loci of *K. ovata* and *K. endophytica* contain *STE3***a** and *MF***a** alleles, along with truncated (ψ) versions of *STE3*α. Additionally, *K. ovata* harbors a truncated *MF*α gene. Synteny analysis indicates that the left side of the *P/R* locus in *K. ovata*, *K. endophytica*, and *K. dendrophila* is structurally more similar to *P/R a2* alleles of *K. pini* and *K. shivajii*, while the right side aligns more closely with the *P/R a1* alleles of *K. bestiolae* and *K. dejecticola*. **(B)** Gene genealogy of the *NDC80* gene inferred with IQ-TREE2 (model JTT+G4) showing trans-specific polymorphism across clade E/F species (labeled as 1) and within clade G species (labeled as 2). The *NDC80* alleles of *K. ovata*, *K. endophytica*, and *K. dendrophila* cluster more closely with *a1* alleles than with other *a2* alleles. The data underlying this Figure can be found at https://doi.org/10.5281/zenodo.14851287.

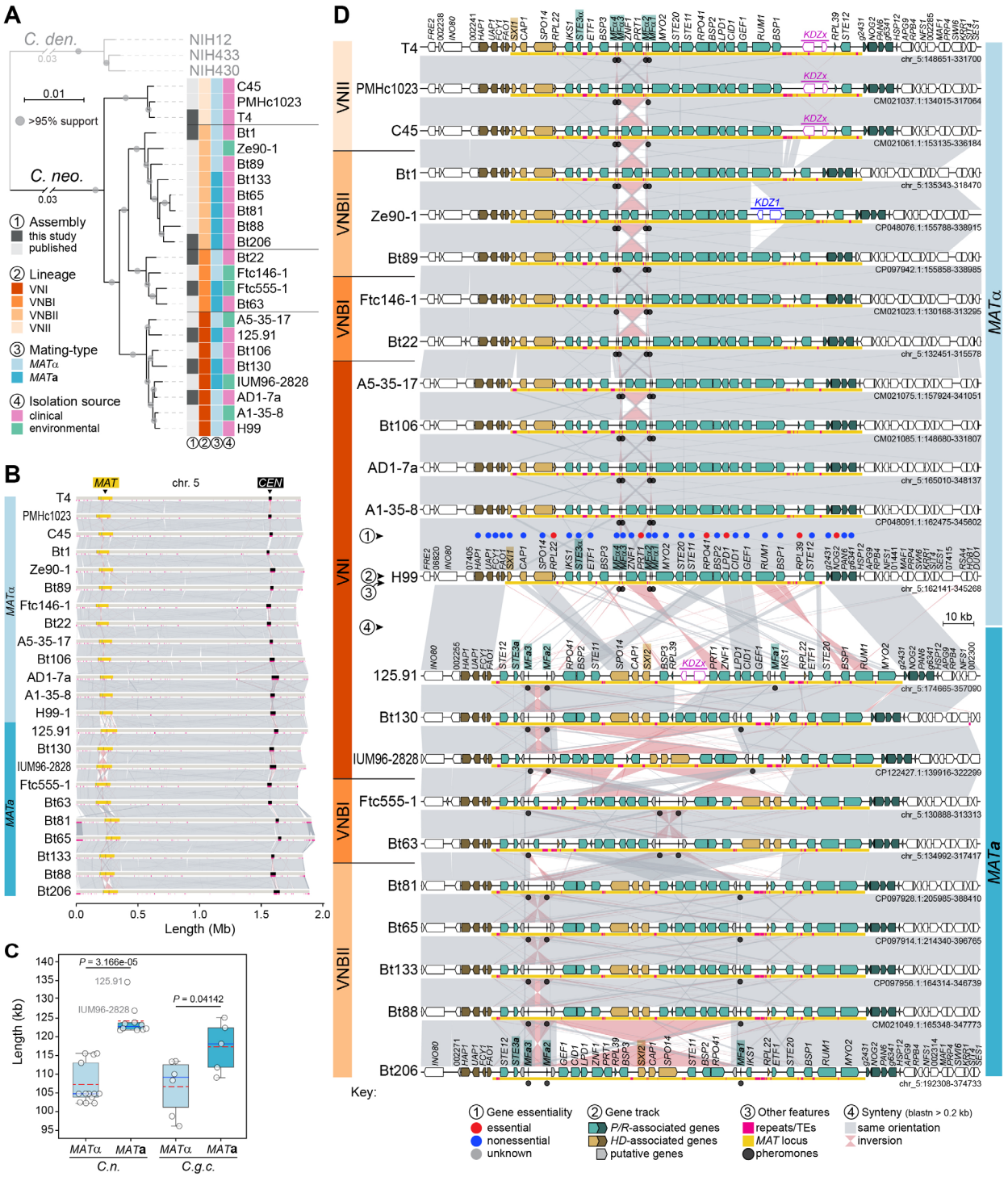


Fig 4. Synteny analysis of the mating-type locus (*MAT*α and *MAT*a) across different *Cryptococcus neoformans* lineages. (A) Phylogenetic tree of representative *C. neoformans* strains from different VN lineages, inferred using maximum likelihood analysis of a concatenated matrix of protein alignments from 5,439 single-copy genes. Branch support is based on SH-aLTR and UFboot tests. Mating types were determined based on the allelic version of the pheromone receptor gene (*STE3*). **(B)** Synteny analysis of chr. 5, where the *MAT* locus resides, showing significant conservation across strains, except within the *MAT*

and centromere regions, which exhibit rearrangements. Strains are organized by mating type and lineage, as in panel D. **(C)** Box plot comparing *MAT* locus size between α and **a** strains of *C. neoformans* (*C.n.*) and *C. gattii* species complex (*C.g.c.*). The red dashed line, blue line, and boxes denote the mean value, median value, and interquartile range, respectively. Outliers are labeled with the corresponding strain name. Statistical significance (*P*-values) was determined by the Mann–Whitney U test. **(D)** Detailed synteny of the *MAT* locus, highlighting strain- and mating-type-specific rearrangements. The *MAT***a** configuration is more variable both across and within VN lineages, existing in three distinct configurations, whereas the *MAT*α structure is more conserved, with limited rearrangements between ancestrally linked homeodomain (*HD*) and pheromone receptor (*P/R*) genes. Observed rearrangements appear to result from inversions mediated by identical copies of pheromone genes. KDZ transposons are detected within the *MAT* locus of certain strains. Genes typically associated with the *P/R* and *HD* loci in tetrapolar species are colored green and gold, respectively, with darker shades denoting genes flanking the *MAT* locus in *Cryptococcus* pathogens. Gene essentiality within the *MAT* locus, as predicted or experimentally validated in strain H99, is indicated by color-coded circles as given in the key. The data underlying this Figure can be found at https://doi.org/10.5281/zenodo.14851287.

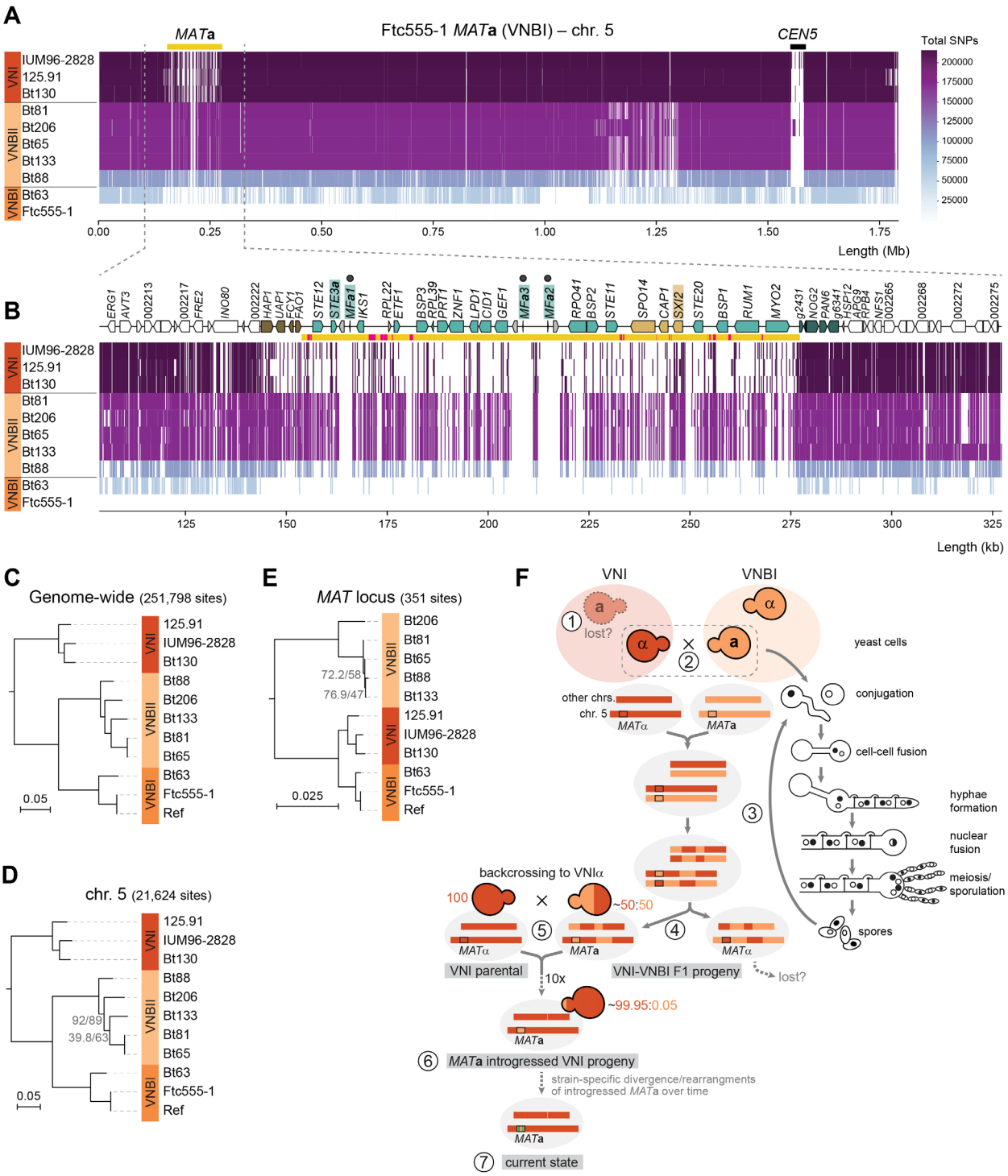


Fig 5. SNP distribution, phylogenetic analysis, and revised introgression model of the *MAT*a allele from VNBI into VNI. (A) Genome-wide distribution of SNPs across chromosome 5 in *C. neoformans MAT*a strains from different VN lineages relative to the *C. neoformans* Ftc555-1 VNBI genome reference. Empty spaces in the plots represent either sites identical to the reference genome (monomorphic sites) or regions excluded due to low coverage or repetitive/duplicated sequences. Strains on the y-axis are ordered by the total number of SNPs across the genome, with a color gradient (right) indicating SNP counts: darker colors

represent higher SNP densities relative to the reference (strains positioned at the top), while lighter colors indicate fewer SNPs (strains at the bottom). Reads from the reference strain were also mapped as a control. The x-axis shows genomic coordinates. **(B)** Zoomed-in view of the *MAT* locus region plus 50 kb flanking regions, with genes and other features represented as in Fig 4. Notably, VNI strains exhibit a markedly lower number of SNPs at the *MAT***a** locus relative to the VNBI reference, compared to VNBII strains. **(C)** Midpoint-rooted tree based on genome-wide SNPs across all chromosomes, illustrating the evolutionary relationships of the strains. **(D)** Midpoint-rooted tree constructed using SNPs from chromosome 5 only, showing similar relationships to the genome-wide analysis. **(E)** Midpoint-rooted tree based on SNPs restricted to the *MAT* locus, highlighting a closer relationship between VNI and VNBI strains. In all trees, branches are shown as the number of substitutions per site (scale bars) and have >95% support based on SH-aLTR and UFboot tests, unless otherwise indicated. “Ref” represents the reference genome assembly employed for read mapping and is thus identical to Ftc555-1 reads, serving as a control to validate the SNP calling pipeline **(F)** Proposed model for the introgression of the VNBI *MAT***a** allele into VNI, shown alongside the key stages of the *C. neoformans* sexual cycle (right). Step 1: possible loss of the *MAT***a** allele in the VNI lineage, likely due to a bottleneck; step 2: mating and cell-cell fusion between a VNBI (*MAT***a**) and a VNI (*MAT*α) strain; step 3: formation of dikaryotic hyphae, nuclear fusion, meiosis, recombination, and sporulation; step 4: some of the resulting F1 progeny emerged with a ~50:50 VNBI:VNI genome composition, retaining the VNBI *MAT***a** locus; step 5: repeated backcrossing of the *MAT***a** F1 progeny to VNI *MAT*α strains progressively purging VNBI genomic regions while retaining the introgressed *MAT***a** allele; step 6: after ~10 backcrosses, the genome became ~99.95% VNI, with only the VNBI *MAT***a** locus and possibly a small fraction (~0.05%) of the VNBI genome remaining; step 7: over time, the introgressed VNBI *MAT***a** locus diverged and underwent rearrangements within the VNI background, reflecting strain-specific changes. The data underlying this Figure can be found at https://doi.org/10.5281/zenodo.14851287.

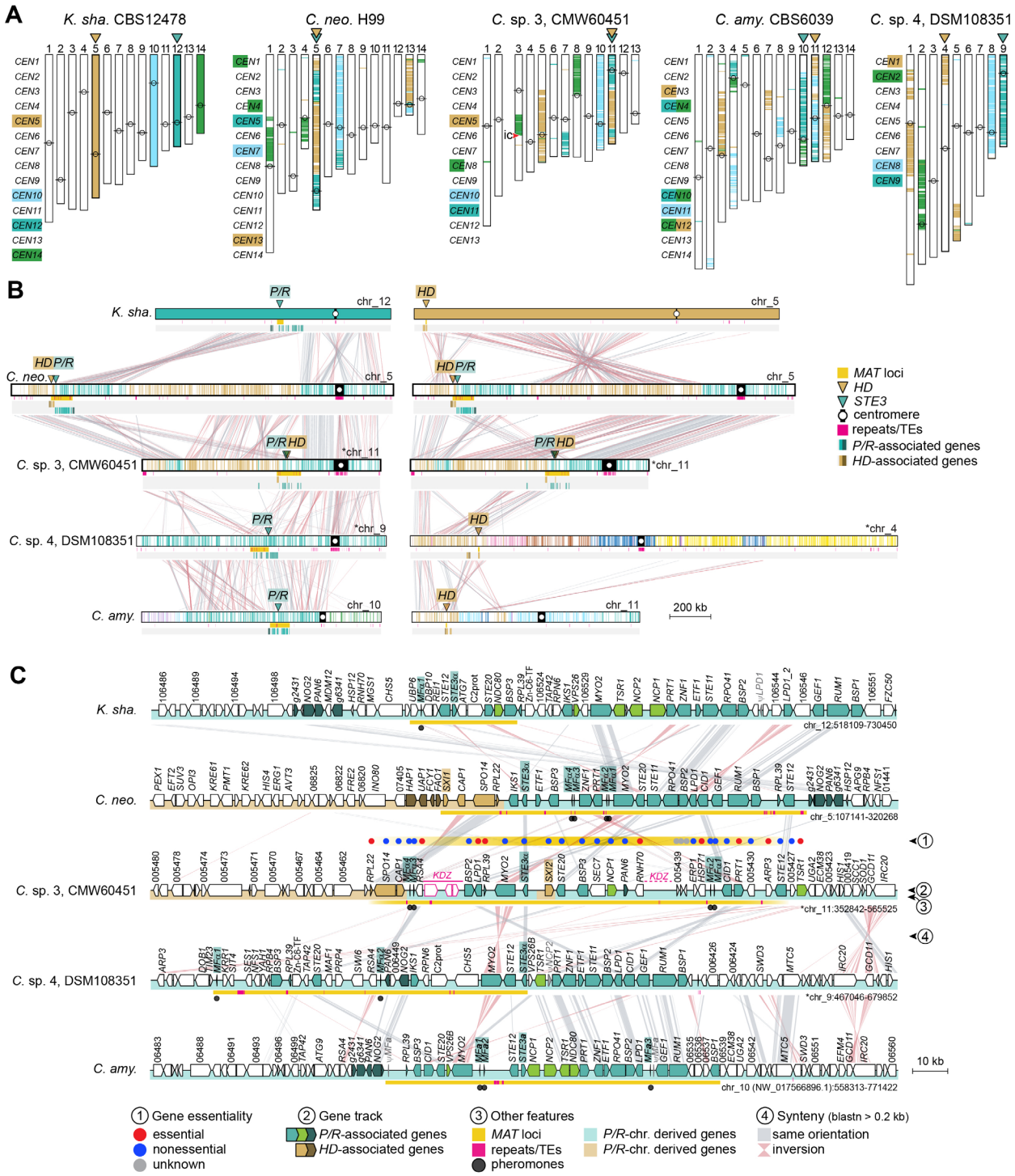


Fig 6. Transition to linked *P/R-HD* loci in *Cryptococcus* sp. CMW60451 and *Cryptococcus* pathogens. **(A)** The karyotype of *K. shandongensis* (with 14 chrs.) served as the reference for reconstructing synteny blocks in pairwise comparisons. For simplicity of visualization, only synteny blocks corresponding to *P/R*- and *HD*-containing chromosomes (chrs. 12 and 5, colored teal and gold, respectively) were plotted in representative *Cryptococcus* species of clades A (*C. neo.* H99), B (*C. amy.* CBS6039), C (*C.* sp. CMW60451), and D (*C. sp.* DSM108351). Besides the *MAT*-containing chromosomes, synteny blocks of

two additional chromosomes in *K. shandongensis* (chrs. 10 and 14), which correspond to a large portion of *P/R* and *HD* chromosomes in *C. amylolentus* were also plotted (color-coded light blue and green, respectively). Other synteny blocks can be visualized in **S9A Fig**). A red arrowhead pinpoints the predicted location of an inactivated centromere (ic) in *Cryptococcus* sp. CMW60451. **(B)** Linear chromosome plots depicting gene synteny conservation across species with zoomed-in views depicted in **(C)**.

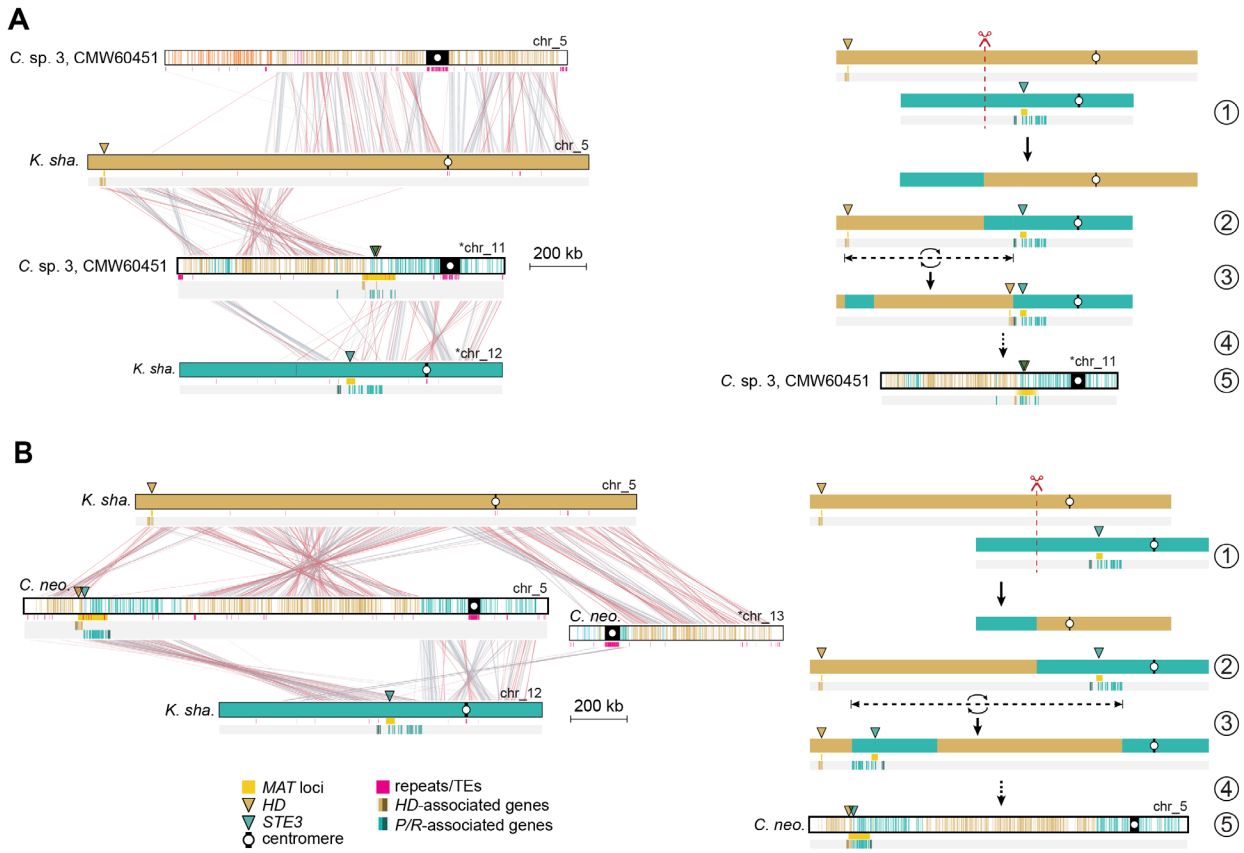


Fig 7. Hypothesized mechanisms underlying the tetrapolar-to-bipolar transitions in *Cryptococcus* sp. 3 and in *Cryptococcus* pathogens. Comparative synteny analysis between *K. shandongensis* and **(A)** *Cryptococcus* sp. 3 or **(B)** *C. neoformans*, illustrating the inferred sequence of chromosomal rearrangements leading to this transition. In both panels, the left side displays pairwise alignments, with syntenic regions connected by red (collinear) and gray (inverted) links, while the right side depicts a stepwise model of the proposed structural changes. The transition likely began with (1) an initial chromosomal break and reciprocal translocation (red dashed line), leading to (2) the repositioning of the *P/R* locus onto the same chromosome as the *HD* locus. (3) In *Cryptococcus* sp. 3, a subsequent inversion relocated the *HD* locus closer to the presumed ancestral position of the *P/R* locus, whereas in *Cryptococcus* pathogens, an inversion instead shifted the *P/R* locus toward the original location of the *HD* locus. (4) Additional structural modifications further refined the newly linked configuration, ultimately resulting in (5) the extant *MAT* locus organization observed in each species. These models suggest that despite independent evolutionary trajectories, both lineages underwent convergent genomic rearrangements that facilitated the transition to bipolar mating systems.

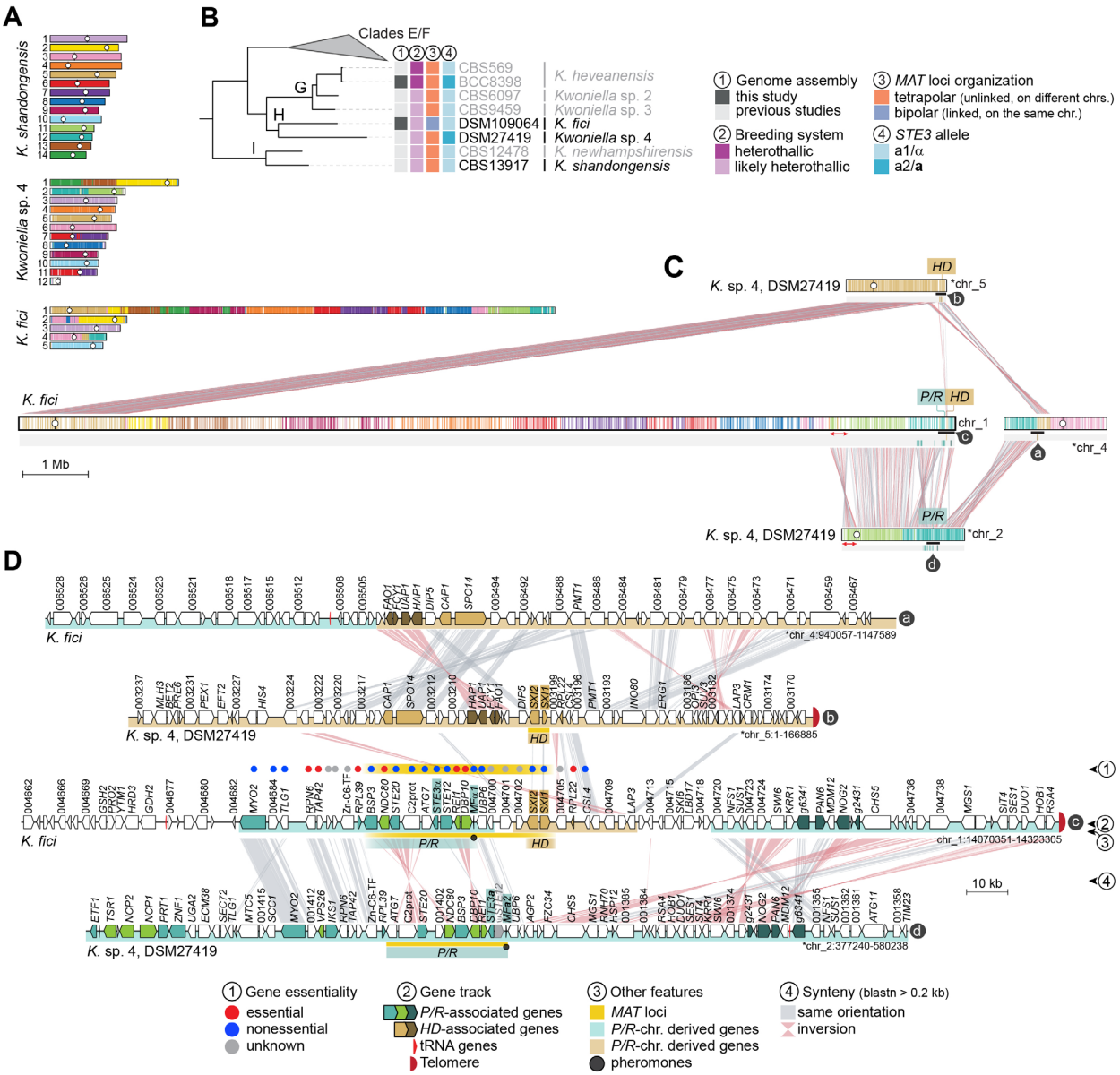
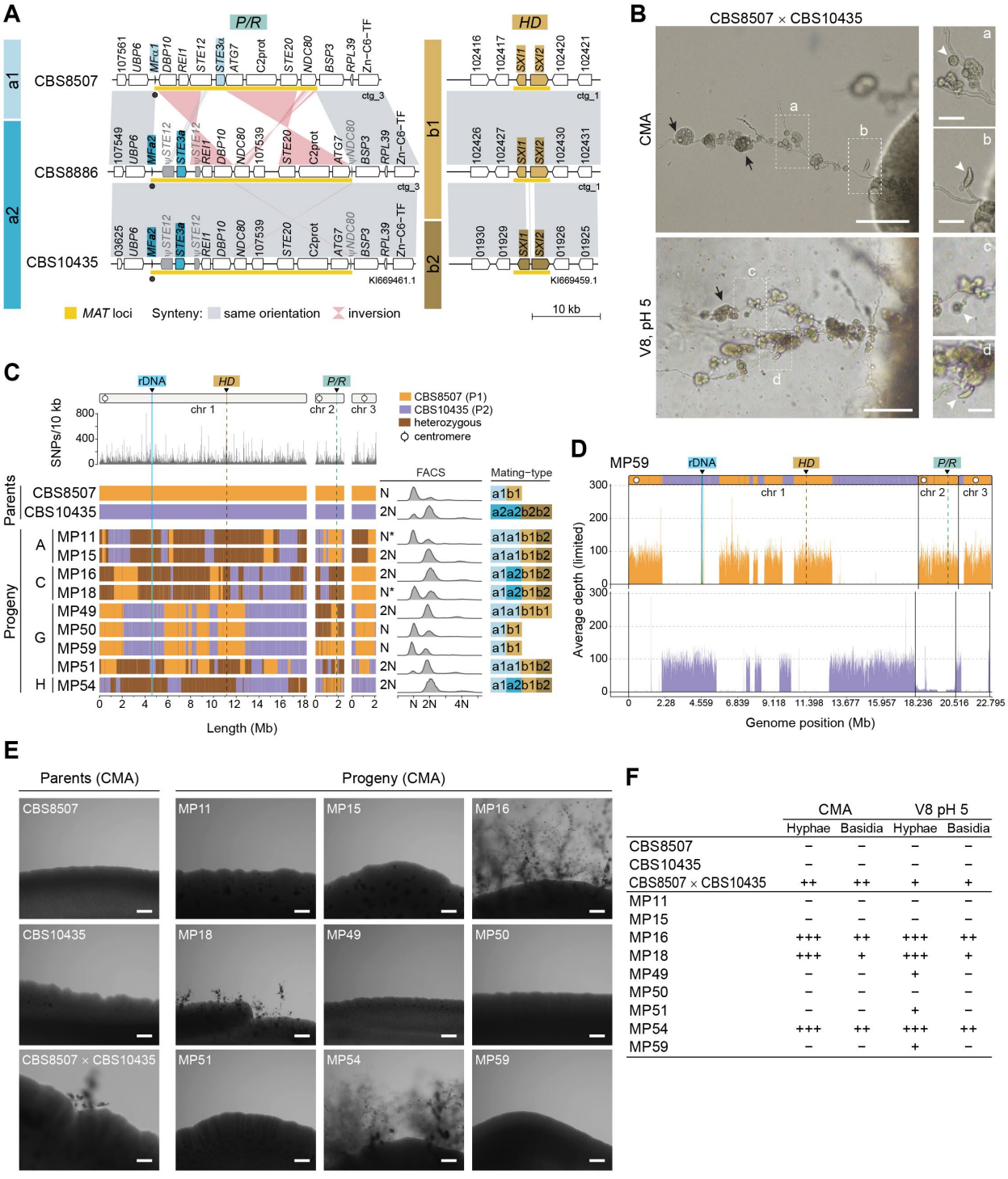


Fig 8. Linkage of *P/R* and *HD* loci in *Kwoniella fici* and *MAT* locus structure (A) The karyotype of *K. shandongensis* (with 14 chrs.) served as a reference for reconstructing synteny blocks in pairwise comparisons with *Kwoniella* sp. 4 and *K. fici*. **(B)** Phylogenetic relationships among selected *Kwoniella* species, highlighting their *MAT* locus organization and inferred breeding systems. **(C)** Synteny comparison between *K. fici* and its closest relative *Kwoniella* sp. 4, illustrating chromosomal rearrangements underlying the linkage of *P/R* and *HD* loci in *K. fici*. A red double-headed arrow marks that the centromere-proximal regions of chr. 2 in *Kwoniella* sp. 4 correspond to a region near the fusion point on the *K. fici* "giant" chromosome (chr. 1), whereas the telomere-proximal regions of *Kwoniella* sp. 4 chr. 2 align with more internalized regions. This suggests that a large pericentric inversion targeting the centromere-adjacent region is associated with this fusion event, as previously reported for other *Kwoniella* species [48]. **(D)** Detailed gene-level organization of the *MAT* locus in *K. fici* compared to *Kwoniella* sp. 4. Each track (labeled a–d in dark circles) corresponds to a specific region marked by a black bar in panel C. Chromosomes inverted relative to their original assembly orientations are marked with asterisks.



	CMA		V8 pH 5	
	Hyphae	Basidia	Hyphae	Basidia
CBS8507	–	–	–	–
CBS10435	–	–	–	–
CBS8507 × CBS10435	++	++	+	+
MP11	–	–	–	–
MP15	–	–	–	–
MP16	+++	++	+++	++
MP18	+++	+	+++	+
MP49	–	–	+	–
MP50	–	–	–	–
MP51	–	–	+	–
MP54	+++	++	+++	++
MP59	–	–	+	–

Fig 9. Sexual reproduction and recombination in *Kwoniella mangrovensis*. **(A)** Synteny analysis of the *P/R* and *HD* loci among three *K. mangrovensis* strains of different mating types (CBS8507, *a1b1*; CBS8886, *a2b1*; and CBS10435, *a2b2*). Note the pronounced divergence in the N-terminal domains of both *HD1* (*SXI1*) and *HD2* (*SXI2*) gene products, which are known to mediate mating-type specificity in other basidiomycetes. **(B)** Micrographs showing hyphal filaments extending from colony peripheries in a sexual cross between CBS8507 and CBS10435 after 2 weeks of incubation on CMA and V8 pH 5 media, in the

dark, at room temperature. Insets depict two types of basidia: globose (a–c) and lageniform (b–d). Black arrows indicate clusters of cells emerging near basidia (either at the surface or embedded), scored as potential meiotic progeny (basidiospores). Single cells from these clusters were isolated using a micromanipulator, cultured into colonies, and genotyped by PCR-RFLP (see S5 Appendix). Putative recombinants were further analyzed using Illumina sequencing. Scale bars = 100 μm (25 μm in insets). **(C)** Genotypic analysis of selected meiotic progeny. SNP density between the parental strains was calculated as the number of SNPs per 10 kb (top). The genotypes of 9 segregants, derived from four distinct cell clusters, were inferred from SNP data and are depicted as follows: orange for regions inherited from CBS8507, purple for regions inherited CBS10435, and brown for heterozygous regions (i.e. inherited from both parents). Instances of recombination or loss of heterozygosity (LOH) are detected by changes in genotype along the chromosomes. Discrepancies in ploidy, as inferred from FACS and sequencing read coverage, suggest potential genomic instability (marked by asterisks). Mating-type identity was inferred through sequencing, coverage analysis, and FACS. **(D)** Sequencing coverage plot for progeny MP59 with color-coded contributions from each parent. Haplotypes blocks inferred from SNP data are overlaid for each chromosome for comparison (additional progeny data is presented in S8 Fig). **(E)** Self-filamentation phenotype of *K. mangrovensis* progeny. Parental strains CBS8507 and CBS10435 (grown individually and in co-culture) and their recovered progeny were cultivated on CMA at room temperature. Self-filamentation was assessed after 2 weeks of incubation. While neither parental strain exhibited self-filamentation in solo culture, their co-culture produced hyphal filaments and basidia. Progeny MP16, MP18, and MP54 also exhibited robust self-filamentation on both CMA and V8 pH 5 (see S9Fig), whereas progeny MP49, MP51, and MP59 displayed weaker filamentous growth, detectable only on V8 pH 5 (S9 Fig). Scale bars = 200 μm. **(F)** Summary of the growth phenotype of the recovered progeny. The production of mycelium and basidia was classified as: extensive (+++), when observed across the entire periphery of the mating patch; moderate (++), when restricted to specific areas of the mating patch; poor (+), when limited to a single spot of the matching patch and slower to develop; and negative (–), when no filamentous growth was observed. Results represent observations from two independent tests.

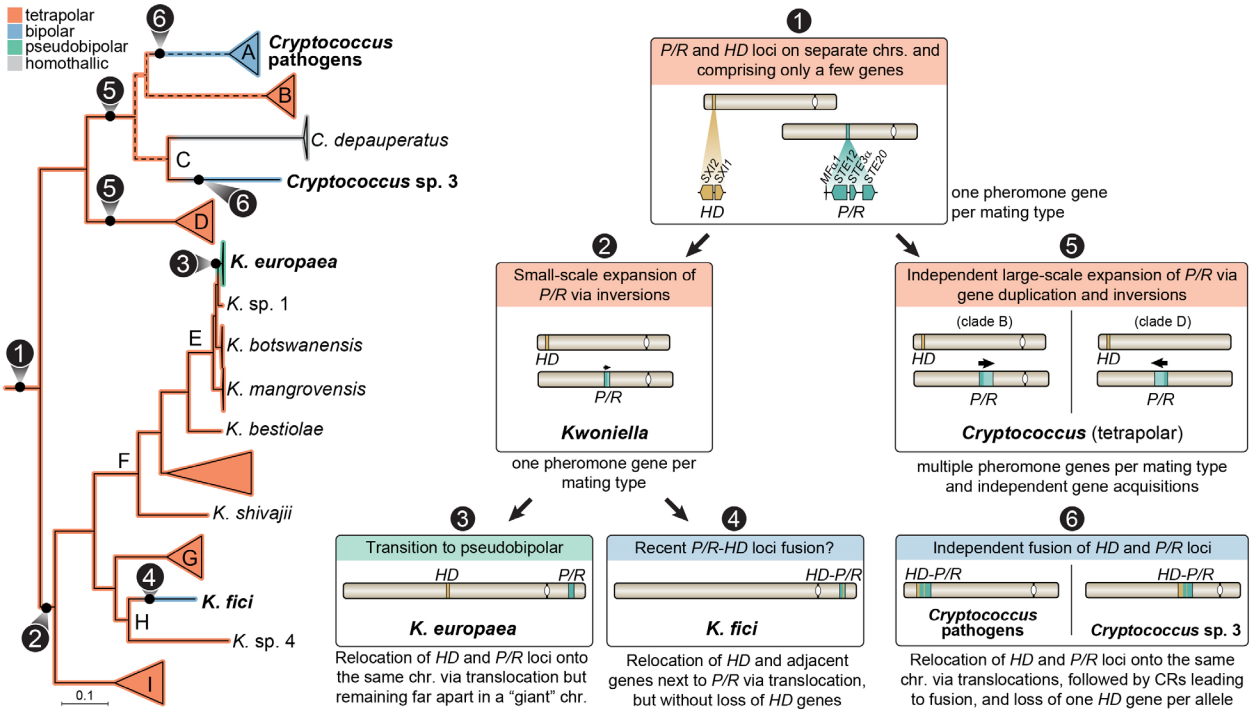


Fig 10. Summary of the evolutionary transitions in *MAT* locus organization across *Kwoniella* and *Cryptococcus* species. Schematic representation of key chromosomal rearrangements and evolutionary events underlying transitions from tetrapolar to bipolar and pseudobipolar mating configurations. The phylogenetic tree highlights inferred changes in *MAT* locus structure, with color-coded branches representing both extant and reconstructed *MAT* configurations. Dashed lines indicate unresolved phylogenetic relationships among clades A, B, and C. Insets summarize distinct evolutionary stages, including the ancestral organization of *P/R* and *HD* loci (1), small- and large-scale expansions of the *P/R* locus in *Kwoniella* and *Cryptococcus* (2, 5), relocation of *P/R* and *HD* loci onto the same chromosome leading to pseudobipolar or fused *MAT* configurations in *Kwoniella* (3, 4), and independent *HD-P/R* fusion events in *Cryptococcus* pathogens and *Cryptococcus* sp. 3 (6). These transitions highlight the diverse pathways by which chromosomal rearrangements have shaped *MAT* locus evolution in these fungal lineages.