Bridgin connects the outer kinetochore to centromeric chromatin

Shreyas Sridhar [1,4], Tetsuya Hori[2], Reiko Nakagawa [3], Tatsuo Fukagawa [2✉] & Kaustuv Sanyal [1,2✉]

The microtubule-binding outer kinetochore is coupled to centromeric chromatin through CENP-C[Mif2], CENP-T[Cnn1], and CENP-U[Ame1] linker pathways originating from the constitutive centromere associated network (CCAN) of the inner kinetochore. Here, we demonstrate the recurrent loss of most CCAN components, including certain kinetochore linkers during the evolution of the fungal phylum of Basidiomycota. By kinetochore interactome analyses in a model basidiomycete and human pathogen *Cryptococcus neoformans*, a forkhead-associated domain containing protein "bridgin" was identified as a kinetochore component along with other predicted kinetochore proteins. In vivo and in vitro functional analyses of bridgin reveal its ability to connect the outer kinetochore with centromeric chromatin to ensure accurate chromosome segregation. Unlike established CCAN-based linkers, bridgin is recruited at the outer kinetochore establishing its role as a distinct family of kinetochore proteins. Presence of bridgin homologs in non-fungal lineages suggests an ancient divergent strategy exists to bridge the outer kinetochore with centromeric chromatin.

[1] Molecular Mycology Laboratory, Molecular Biology and Genetics Unit, Jawaharlal Nehru Center for Advanced Scientific Research (JNCASR), Bangalore, India 560064. [2] Laboratory of Chromosome Biology, Graduate School of Frontier Biosciences, Osaka University, Suita, Osaka 565-0871, Japan. [3] Laboratory for Phyloinformatics, RIKEN Center for Biosystems Dynamics Research (BDR), Kobe, Japan. [4]Present address: Graduate School of Frontier Biosciences, Osaka University, Suita, Osaka 565-0871, Japan. ✉email: tfukagawa@fbs.osaka-u.ac.jp; sanyal@jncasr.ac.in

Accurate chromosome segregation ensures faithful transmission of the genetic material to progeny. The kinetochore is a multicomplex protein structure assembled on centromere DNA of each eukaryotic chromosome[1–4] and is attached to the spindle microtubules for accurate chromosome segregation[5]. Components involved in error correction mechanisms, including the spindle assembly checkpoint (SAC), are recruited at kinetochores to ensure the biorientation of sister chromatids in mitosis[6–8]. The inner kinetochore is composed of centromeric histone H3 variant CENP-A^{Cse4} [9–11] and the 16-member constitutive centromere-associated network (CCAN)[12–16]. The outer kinetochore members of the KMN (Knl1, Mis12, and Ndc80 complexes) network[1,17] are recruited to CCAN to form the kinetochore ensemble across eukaryotes, including budding yeast and vertebrates[18]. Additional components such as the 10-member Dam1 complex (Dam1C) in fungi and the three-member Ska complex in vertebrates localize to the outer kinetochore to ensure accurate kinetochore–microtubule interactions[19–24].

Genome sequencing data reveal that while the KMN network components are highly conserved[25–27], the inner kinetochore composition is variable across eukaryotes[25–29]. The Ndc80 complex (Ndc80C) of the KMN network directly binds to spindle microtubules[30–32]. CCAN proteins, CENP-C^{Mif2} and CENP-T^{Cnn1}, have been shown to independently link centromeric chromatin with the KMN network[33–38]. In addition, CENP-U^{Ame1} functions as a linker in budding yeast[39,40]. CENP-C^{Mif2} and CENP-T^{Cnn1}, through their amino(N)-termini, interact with the KMN network, while their carboxy(C)-termini interact with centromeric chromatin[41–43]. CENP-U^{Ame1} cooperatively with CENP-C^{Mif2} also binds to Mis12C, to ensure a linker function in budding yeast[39,40]. These three critical kinetochore linker proteins, CENP-C^{Mif2}, CENP-T^{Cnn1}, and CENP-U^{Ame1}, are often lost or significantly diverged during evolution. Despite the observed loss or functional divergence of linker proteins, the existence of other molecular innovations to bridge centromeric chromatin with the outer kinetochore to ensure accurate chromosome segregation remains largely unknown[21,28,44]. Although each of the kinetochore linker pathways has been shown to play important roles in chromosome segregation, the architectural dependence on a specific pathway varies across species[34,35,37].

Cryptococcus neoformans is a ubiquitous environmental fungus and an opportunistic pathogen causing fatal cryptococcal meningitis[45,46]. Our previous studies using this model basidiomycete suggested a stepwise kinetochore assembly on a long repetitive regional centromere[24,47].

In this work, by analyzing a number of fungal genomes, we strikingly find the recurrent absence of most CCAN components, including certain conventional linker proteins, CENP-T^{Cnn1} and CENP-U^{Ame1}, in the phylum Basidiomycota. We identify "bridgin", a previously undescribed protein, in the kinetochore interactome of *C. neoformans*. We predict the existence of bridgin homologs in eukaryotic lineages outside the fungal kingdom as well. Based on in vivo and in vitro functional analyses, we demonstrate that bridgin connects the outer kinetochore with centromeric chromatin revealing the existence of a distinct strategy to bridge the outer kinetochore and centromeric chromatin across eukaryotes.

Results

Recurrent independent loss events of CCAN proteins in Basidiomycota. To have a comprehensive understanding of the kinetochore composition in Basidiomycota, we analyzed putative kinetochore homologs using high-confidence protein homology searches combined with secondary and tertiary structure prediction. Species representing 31 fungal orders across the three subphyla, Pucciniomycotina, Ustilagomycotina, and Agaricomycotina, were considered. Additionally, representative species across 7 fungal phyla were included[48]. CENP-A^{Cse4}, the 16-member CCAN, and the 10-member KMN network were chosen for this study (Fig. 1a and Supplementary Data 1). Our analysis, in accordance with the previous reports[25,26,29], indicates the robust conservation of CENP-A^{Cse4} and the KMN network proteins across basidiomycetes. On the other hand, we observed recurrent loss of most CCAN proteins across 23 of the 31 basidiomycete orders (Fig. 1a). In the subphylum of Agaricomycotina, to which *C. neoformans* belongs, the loss event may have occurred early at the time of divergence of Wallemiales from other orders. Retention of CCAN proteins in a few discrete orders in the subphyla of Pucciniomycotina and Ustilogamycotina suggests the occurrence of multiple independent loss events of most CCAN subunits (Fig. 1a). The kinetochore linker CENP-C^{Mif2} was the only conserved CCAN component across basidiomycetes. Other known linker proteins, CENP-T^{Cnn1} and CENP-U^{Ame1}, were often lost together. Although the primary protein sequence conservation is low among CENP-T^{Cnn1} homologs, they share a typical protein architecture: an N-terminal α-helix composed of conserved hydrophobic residues, observed to occur within the PITH domain in Ustilaginales, and the CENP-T^{Cnn1} motif at the C terminus (Fig. 1b). As a result, the CENP-C^{Mif2} pathway remains the only known linker pathway among 23 of the 31 basidiomycete orders investigated. These observations suggest that additional factors may exist to compensate for the recurrent loss of most CCAN proteins in Basidiomycota.

Identification of the kinetochore interactome in *C. neoformans*. In search of such unknown kinetochore proteins, we attempted to analyze the kinetochore composition in the fungal phylum of Basidiomycota using a relatively well-studied model basidiomycetous yeast *C. neoformans* (Fig. 1a). To comprehensively determine the constitution of the kinetochore and its interactome in vivo, *C. neoformans* strains expressing 3×FLAG-tagged CENP-C^{Mif2}, Dsn1 (Mis12C), and Spc25 (Ndc80C) under control of the native promoter were generated. The functionality of the 3×FLAG-tagged proteins was validated by chromatin immunoprecipitation (ChIP) and co-immunoprecipitation (co-IP) assays (Supplementary Fig. 1a–c).

Mass-spectrometry (MS) analyses were performed after FLAG-IP for CENP-C^{Mif2}, Dsn1, and Spc25 from metaphase-enriched (mitotic index > 90%) cell population (Fig. 1c and Supplementary Data 2). Nearly all predicted inner kinetochore proteins (CENP-A^{Cse4} and CENP-C^{Mif2}) and outer kinetochore KMN network components were identified from each of the FLAG-IPs (Table 1 and Supplementary Data 2). Except for CENP-C^{Mif2}, no components of the CCAN were identified in the IP–MS (Table 1 and Supplementary Table. 2). Identified KMN network components included evolutionarily conserved but previously unannotated ORFs coding for subunits of Mis12C (CNAG_04300^{Nsl1}, CNAG_04479^{Nnf1}) and Knl1C (CNAG_03715^{Sos7}) (Table 1 and Supplementary Data 2). CNAG_03715^{Sos7} (Knl1C) was used as a candidate to authenticate the predicted ORFs as bona fide kinetochore proteins. CNAG_03715^{Sos7} was bioinformatically identified as the Sos7 homolog in *C. neoformans* using *Schizosaccharomyces pombe* Sos7[49] (Supplementary Fig. 1d, e). Microscopic observations of GFP-CnSos7 suggested a cell-cycle-dependent kinetochore-exclusive localization in *C. neoformans* (Fig. 1d and Supplementary Fig. 1f). Further analysis confirmed that while Sos7 was dispensable for cell viability, yet was essential for accurate chromosome segregation in *C. neoformans* (Supplementary Fig. 1g, h).
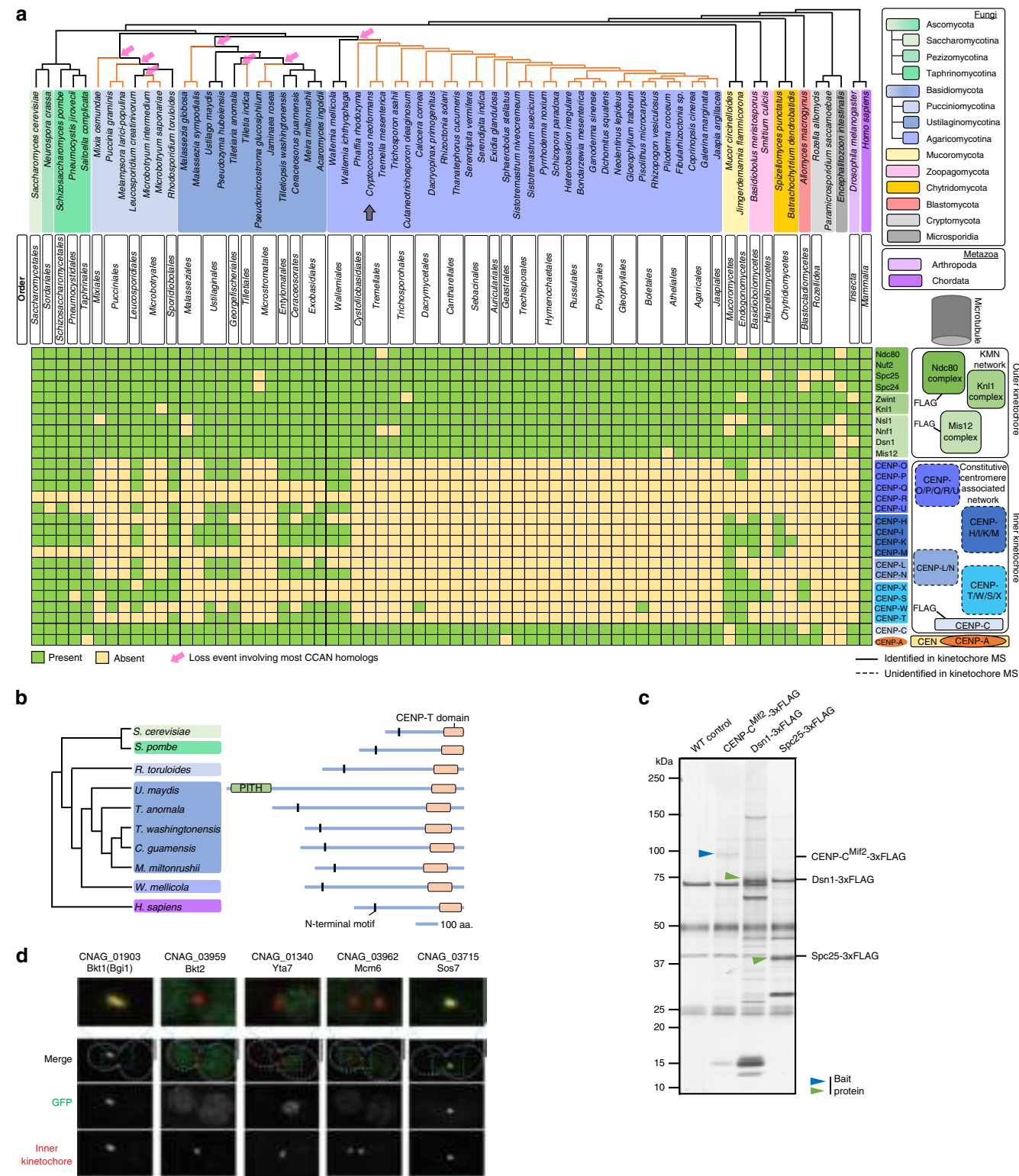
Taken together, the experimental identification of the kinetochore interactome confirms our bioinformatics prediction that *C. neoformans* retains a single known kinetochore linker pathway through CENP-C[Mif2].

Bridgin (Bgi1) is a basidiomycete kinetochore protein. Next, we attempted to find the existence of unknown kinetochore proteins compensating for the loss of CCAN subunits in Basidiomycota.

We gained confidence by identifying nearly all known structural kinetochore components from each FLAG-IP–MS experiment. Thus, we hypothesized that it was possible to identify previously undescribed kinetochore proteins, if any, from the identified common interactors of CENP-C[Mif2], Dsn1, and Spc25 (Supplementary Data 2). Among the identified proteins that fulfill this stringent criterion, we categorized two sets of proteins: (a) conserved among basidiomycetes with unknown function and

Fig. 1 Identification of the kinetochore interactome in *Cryptococcus neoformans*. a Conservation of kinetochore proteins across the mentioned species. The cladogram represents the relationship between the species; each subphylum is color-coded. The presence (green boxes) or absence (yellow boxes) of each kinetochore protein in every species is shown. Pink arrows indicate loss events of CCAN proteins, excluding CENP-C^{Mif2}. Orange lines in the cladogram refer to basidiomycete lineages that have lost most CCAN components. The gray arrow points to *C. neoformans*. *Right*, schematic of a typical kinetochore ensemble. FLAG-labeled proteins of kinetochore subcomplexes (CENP-C, Mis12C, and Ndc80C) in *C. neoformans* are indicted and used for the immunoprecipitation (IP)–mass-spectrometry (MS) identification of the kinetochore interactome. Protein complexes identified (solid borders) or those that remained unidentified (dotted borders) by the IP–MS experiment in *C. neoformans* are shown. **b** Domain architectures of identified CENP-T^{Cnn1} homologs among basidiomycetes. Color coding of subphyla is followed as in **a**. **c** Proteins eluted after FLAG-IP, of CENP-C^{Mif2} in the *C. neoformans* strain SHR896 (*CENP-C::CENP-C^{MIF2}-3×FLAG*), Dsn1 in SHR824 (*DSN1::DSN1-3×FLAG*), Spc25 in SHR823 (*SPC25::SPC25-3×FLAG*), and an untagged wild-type (WT) control (H99) after thiabendazole (TBZ) treatment, were separated on a gradient PAGE gel and silver-stained. Blue and green arrows mark the inner and outer kinetochore bait proteins, respectively. **d** Micrographs of *C. neoformans* cells at the M phase expressing GFP-tagged proteins identified by the IP–MS screen. Kinetochores are marked by inner kinetochore proteins, CENP-C^{Mif2}-mCherry for Bgi1^{Bkt1} in SHR876 (*BGI1::BGI1-V5-GFP*), Bkt2 in SHR897 (*BKT2::BKT2-V5-GFP*), and Yta7 in SHR842 (*YTA7::YTA7-V5-GFP*) or mCherry-CENP-A^{Cse4} for Mcm6 in SHR905 (*MCM6::MCM6-V5-GFP*) and Sos7 in SHR845 (*SOS7:: GAL7p-GFP-SOS7*). Scale bar, 3 μm. Source data are available as a Source Data file.

Complex	Protein	WT control		CENP-C^{Mif2}–3×FLAG		Dsn1–3xFLAG		Spc25-3×FLAG	
		C (%)	TSC	C (%)	TSC	C (%)	TSC	C (%)	TSC
Ndc80	Ndc80	–	–	16	8	46	113	61	1130
	Nuf2	2.7	1	10	4	44	67	69	683
	Spc25	19	7	21	6	64	90	84	581
	Spc24	–	–	23	6	73	51	82	313
Knl1	Sos7 (CNAG_03715)	–	–	11	4	55	74	57	32
	Knl1^{Spc105}	–	–	8	9	31	272	38	206
Mis12	Nsl1 (CNAG_04300)	–	–	15	10	67	84	47	24
	Nnf1 (CNAG_04479)	–	–	12	6	52	222	39	68
	Dsn1	–	–	26	20	61	607	48	127
	Mis12^{Mtw1}	–	–	30	11	72	287	49	56
CCAN	CENP-C^{Mif2}	–	–	48	272	35	66	–	–
CENP-A	CENP-A^{Cse4}	–	–	24	9	20	10	6	2

**Table 1 List of kinetochore proteins obtained in the kinetochore IP–MS.**

Percentage of amino acid sequence coverage (C,%) and the total spectrum counts (TSC) specific to the corresponding protein obtained by IP-MS are mentioned.

termed them basidiomycete kinetochore proteins (Bkts), which included CNAG_01903^{Bkt1}, CNAG_03959^{Bkt2}, and CNAG_02701^{Bkt3} (Supplementary Fig. 1i), and (b) known chromatin-interacting proteins with uncharacterized kinetochore function, that included CNAG_01340^{Yta7} and all components of the evolutionarily conserved DNA replication initiation factors Mcm2–7 of the MCM complex[50] (Supplementary Fig. 1j). We characterized CNAG_03962^{Mcm6} as the representative test candidate for the MCM complex. As a secondary screen, each of these five candidates was expressed as C-terminally V5-GFP epitope-tagged proteins from their endogenous locus. We could observe subcellular localization in strains expressing CNAG_01903^{Bkt1}, CNAG_03959^{Bkt2}, CNAG_01340^{Yta7}, or CNAG_03962^{Mcm6}. The protein encoded by CNAG_01903^{Bkt1} colocalized at the kinetochore with CENP-C^{Mif2} (Fig. 1d and Supplementary Fig. 1k). Other proteins did not show exclusive kinetochore localization, although some puncta of CNAG_01340^{Yta7} and CNAG_03962^{Mcm6} colocalized transiently with inner kinetochore proteins (CENP-C^{Mif2} or CENP-A^{Cse4}) at G1/S stage (Supplementary Fig. 1k). Recently, Yta7 was shown to act as a deposition factor for CENP-A^{Cse4} [51]. Based on the localization dynamics of Bkts, CNAG_01903^{Bkt1} was taken forward as a bona fide candidate kinetochore protein due to its exclusive kinetochore localization. Considering its identified function through this study, we refer to Bkt1 as "bridgin" (Bgi1) henceforth.

Multiple outer kinetochore receptors recruit bridgin. Cross-linked ChIP analysis suggested that bridgin localized to the centromere at M phase but absent at S phase (Fig. 2a and Supplementary Fig. 2a). We validated the localization by microscopic

observations (Supplementary Fig. 1k). Cellular pools of bridgin were relatively uniform through the cell cycle, suggesting that the G2/M kinetochore localization of bridgin is not due to a higher level of protein expression at the corresponding stage (Supplementary Fig. 2b). Further, bridgin localization at the kinetochore was not affected by the treatment of a microtubule-depolymerizing drug, thiabendazole (TBZ) (Supplementary Fig. 2c), suggesting that loss of spindle integrity did not influence the kinetochore localization of bridgin.

Our previous study suggested a stepwise assembly of the kinetochore in *C. neoformans*, yet no sequential assembly of kinetochore subcomplexes was established[24]. We investigated where bridgin localizes within the kinetochore localization hierarchy. We adopted a microscopy-based kinetochore interdependency assay strategy involving a test kinetochore protein in the background of a conditional kinetochore mutant[52] (Fig. 2b). Conditional mutants of CENP-A^{Cse4}, CENP-C^{Mif2}, Mis12^{Mtw1}, Nuf2, Knl1^{Spc105}, Dad1, and Dad2 in *C. neoformans* were generated by driving expression of each of these GFP/mCherry-tagged kinetochore proteins under the control of the *GAL7* promoter (*GAL7p*). Upon reduced protein levels due to nonpermissive conditions of the *GAL7p*-driven expression (Supplementary Fig. 2d, e), the kinetochore mutants displayed a lack of growth (Supplementary Fig. 2f) on account of reduced cell viability (Supplementary Fig. 2g). These results, in accordance with our previous report[53], validated the effective repression mediated by *GAL7p*. Using this experimental strategy, we determined that Mis12C and Ndc80C influence the stability of each other at the kinetochore (Supplementary Fig. 3a, b). Knl1C (Supplementary Fig. 3c–f) and Dam1C (Supplementary Fig. 3g–i)) independently require the Mis12C–Ndc80C platform to be localized
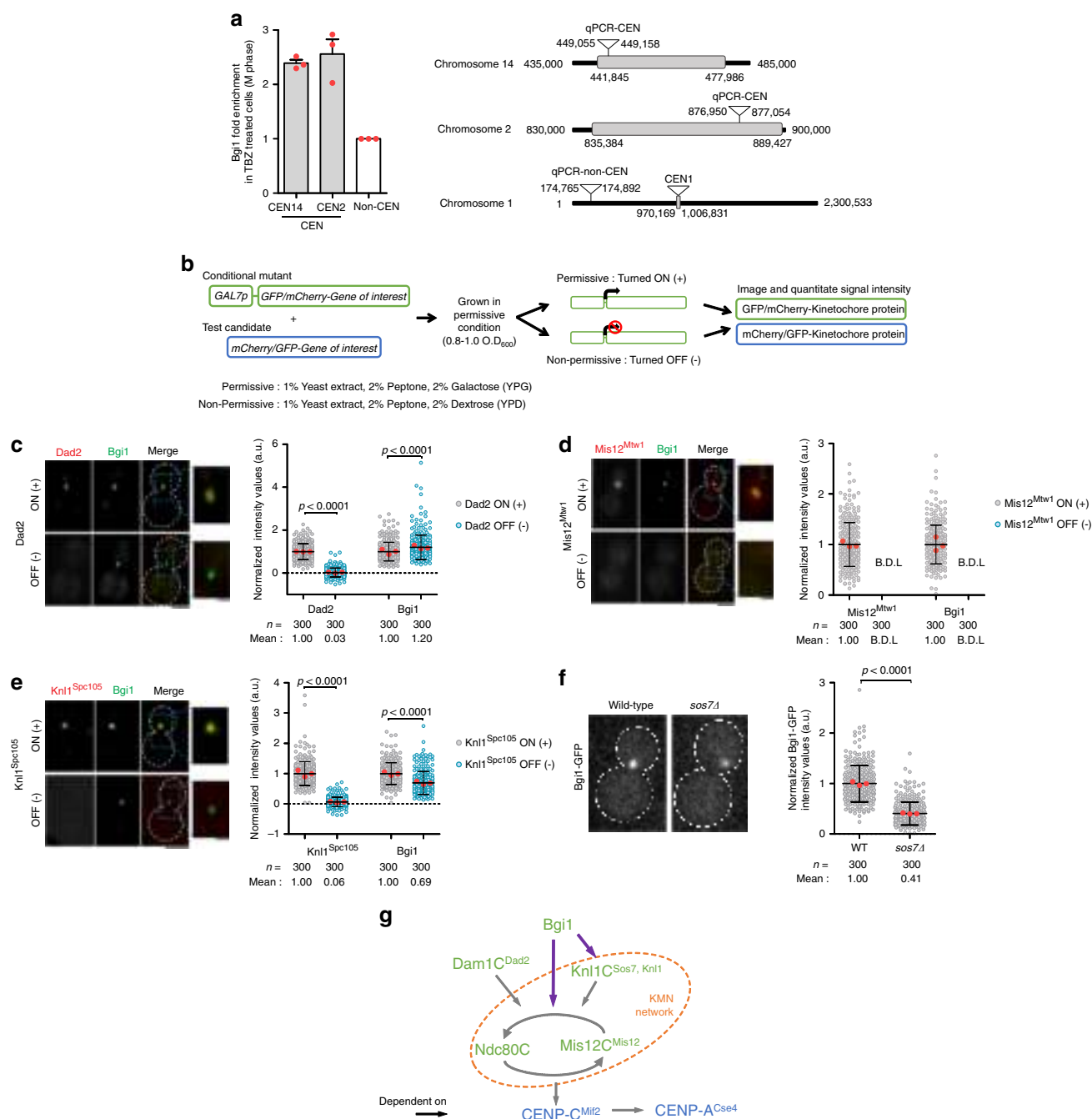
Fig. 2 Bridgin assembles onto the KMN network platform. a *Left*, enrichment of Bgi1 at M-phase kinetochores in SHR870 (*BGI1::BGI1-V5-GFP*) cells by cross-linked chromatin immunoprecipitation (ChIP) followed by quantitative polymerase chain reaction (qPCR). The data represent the mean ± standard deviation (S.D.) of three independent experiments. *Right*, schematic representing the location of qPCR primers used. Centromeric sequences are defined by grey boxes. **b** Steps followed to determine localization interdependency at the kinetochore among its subcomplexes. **c**–**e** *Left*, a representative image of Bgi1-V5-GFP in **c** SHR906 (*DAD2::GAL7p-mCherry-DAD2*), **d** SHR907 (*MIS12^MTW1^::GAL7p-mCherry-MIS12^MTW1^*) and **e** SHR945 (*KNL1^SPC105^::GAL7p-mCherry-KNL1^SPC105^*) in permissive (+) or nonpermissive (−) conditions of the *GAL7p*. Scale bar, 3 μm; 1 μm for magnified inset. *Right*, Bgi1-V5-GFP signals and **c** mCherry-Dad2, **d** mCherry-Mis12^Mtw1^, or **e** mCherry-Knl1^Spc105^ in cells with or without Dad2, Mis12^Mtw1^, or Knl1^Spc105^ expression, respectively, were quantified and normalized to their respective mean signals in the permissive condition of the *GAL7p*. Signals were measured after 18, 12, and 18 h in nonpermissive conditions for Dad2, Mis12^Mtw1^, and Knl1^Spc105^, respectively. For **d**, the strong influence of Mis12^Mtw1^ on kinetochore localization of bridgin resulted in signals that were below detectable levels (B.D.L). For **e**, bridgin signals in 24 out of 300 cells were below detectable levels. **f** *Left*, representative images of Bgi1-V5-GFP in wild-type H99 (*SOS7*) or in *sos7* null mutant SHR908 (*SOS7::sos7Δ*) cells. Scale bar, 2 μm. *Right*, Bgi1-V5-GFP signal intensities were quantified and normalized to the mean wild-type signal. **c**–**f** The results of three independent experiments with 100 cells each are represented. The red dot represents the mean of one experiment; mean ± S.D. is shown. For statistical comparison of differences between the samples, Mann–Whitney two-tailed analysis was applied, *p* values show significant differences. **g** Schematic describes the observed localization interdependency of inner (blue) and outer (green) kinetochore protein complexes at the *C. neoformans* kinetochore. Purple arrows highlight the dependency of bridgin on the KMN network. Source data are available as a Source Data file.
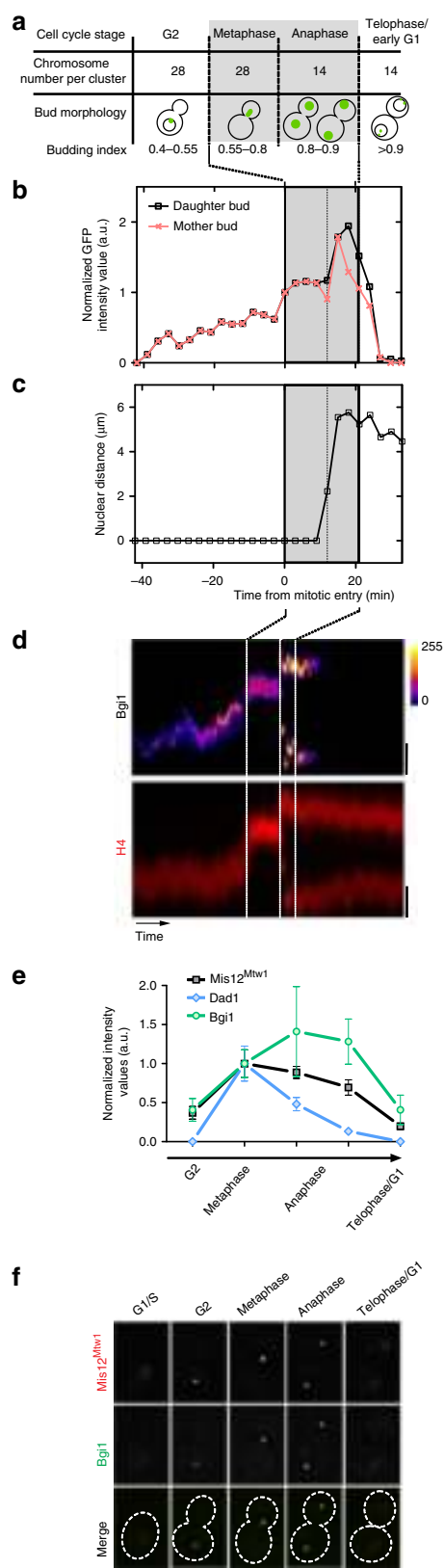
Fig. 3 Bridgin reaches its peak concentration at the kinetochore during anaphase. a Expected chromosome number per kinetochore cluster in the haploid-type strain H99/KN99 and spatial location of bridgin (green) with the corresponding budding index is tabulated. **b** Signal intensity measurements of Bgi1 in a single cell of the strain SHR843 (*BGI1::BGI1-V5-GFP, H4::H4-mCherry*), at an interval of 3 min were measured and normalized to the Bgi1 signal intensity at time point 0 (M-phase entry). **c** The nuclear marker histone H4 in SHR843 was utilized to measure internuclear distance. A cell is considered to have exited anaphase when nuclear distances have reached their maxima. **d** Kymograph of the cell used for signal measurements in **b**. Time interval represented 1 min for a total of 100 min. Scale bar, 2 µm. **e** Comparison of the protein levels of Bgi1 in SHR843, and representative outer kinetochore proteins, Mis12^{Mtw1} in SHR516 (*MIS12^{MTW1}::MIS12^{MTW1}-mCherry*), and Dad1 in CNVY120 (*KN99::H3p-GFP-DAD1*). Signal intensities of each kinetochore protein were quantified and normalized to the mean of their metaphase signals. The results represent the mean ± S.D. of five cells. **f** Co-localization of Bgi1 and Mis12^{Mtw1} in SHR869 (*BGI1::BGI1-V5-GFP, MIS12^{MTW1}::MIS12^{MTW1}-mCherry*) at various stages of the cell cycle in an asynchronous culture. Scale bar, 3 µm. Source data are available as a Source Data file.

Dam1C subunit Dad2 (Fig. 2c). On the other hand, bridgin's kinetochore localization was entirely compromised in the Mis12^{Mtw1} (Mis12C) conditional mutant (Fig. 2d). Considering that Mis12C and Ndc80C require each other for their kinetochore stability (Supplementary Fig. 3a, b), it would be difficult to distinguish their individual contributions to bridgin localization using the interdependency assay. Further, we suspect that bridgin's kinetochore localization would also be affected in Ndc80C conditional mutants. Downstream of the Mis12C–Ndc80C platform, the role of Knl1C components, Knl1^{Spc105} and Sos7, on bridgin's localization at the kinetochore, was tested. Bridgin levels at the kinetochore were reduced to 69% and 41% as compared to the wild type in the conditional mutant of Knl1^{Spc105} and the null mutant of *sos7*, respectively. Thus, both Knl1C components partially contributed to bridgin's kinetochore recruitment (Fig. 2e, f). Taken together, we demonstrate that the recruitment of bridgin at the kinetochore occurs downstream of Knl1C and Mis12C–Ndc80C, suggesting that bridgin has multiple binding sites within the KMN network (Fig. 2g).

Bridgin levels at the kinetochore peak at anaphase. To understand how bridgin dynamics is regulated during cell-cycle progression, we analyzed its signal intensities at the kinetochore. Bridgin localized to the kinetochore starting from G2 until telophase/G1 (Fig. 3a–d and Supplementary Fig. 1k). The signal intensities of bridgin at the kinetochore reached the peak immediately post-anaphase onset, attaining an average of ~150% of metaphase (Fig. 3b, d, e). The dynamic intensities of other transiently localized kinetochore proteins Mis12^{Mtw1} (KMN network) and Dad1 (Dam1C) were measured. Mis12^{Mtw1} localized to the kinetochore during G2 and persisted until telophase/G1, reaching the maximum signal intensity during metaphase (Fig. 3e). Dad1 (Dam1C) localized at the kinetochore exclusively postmitotic onset, reaching peak intensities at metaphase and reducing sharply to an almost undetectable level in late anaphase (Fig. 3e). Analysis of bridgin and Mis12^{Mtw1} localization in an asynchronous population further validated their co-localization to the kinetochore from G2 to telophase/G1 (Fig. 3f). Having observed bridgin's enrichment at the kinetochore, its dependence on the outer kinetochore KMN proteins, kinetochore localization dynamics, and spindle-independent kinetochore recruitment, we conclude that bridgin, as an outer kinetochore protein, is closely associated with the KMN platform.

at the kinetochore. Moreover, the Ndc80C–Mis12C platform requires the inner kinetochore protein CENP-C^{Mif2} for its kinetochore localization (Supplementary Fig. 3j). Observed interdependencies are summarized in Supplementary Fig. 3k.

Subsequent interdependency analyses with bridgin suggested that its localization to the kinetochore was independent of the
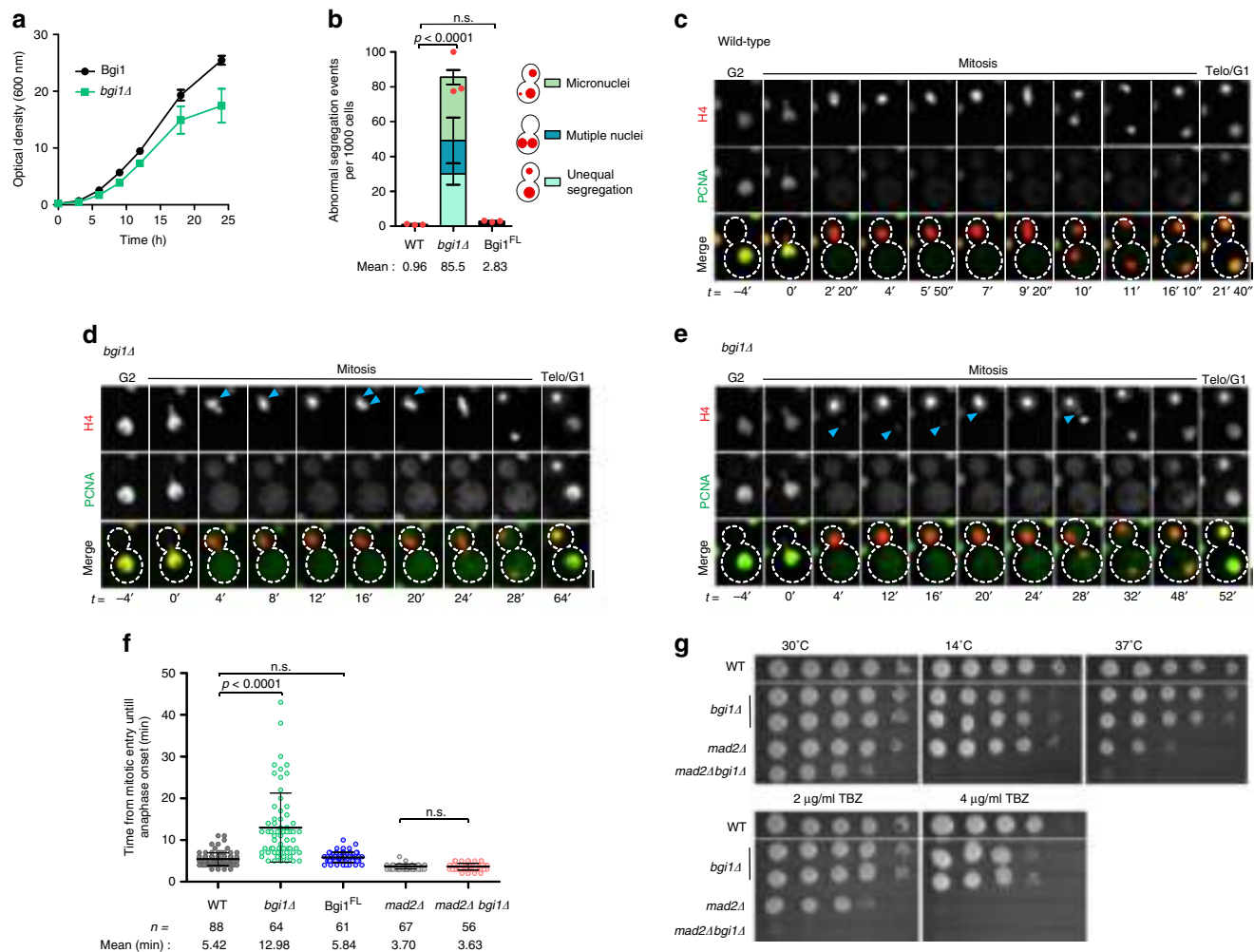
Fig. 4 Loss of bridgin function results in an increased rate of chromosome missegregation and impaired mitotic progression. a Growth curve of wild-type H99 and bridgin-null mutant strain SHR867 (*bgi1Δ::NEO*) at 30 °C. $N = 3$. **b** The rate of abnormal nuclear segregation events was measured at 30 °C using histone H4-mCherry in control CNVY121 (*BGI1*), bridgin-null mutant SHR832 (*bgi1Δ::HygB*), and the Bgi1[FL]-complemented SHR879 (*SHR832::3×FLAG-GFP-BGI1[FL]*) strains. The data represent the mean ± S.D. of three independent experiments. A red dot represents the net abnormal nuclear segregation events per 10^3 cells of one experiment. One-way ANOVA test followed by Dunn's multiple comparison test was used to calculate the statistical significance of differences between the net missegregation events across strains (the *p* values show the difference compared to wild type, ns nonsignificant). **c–e** Representative time-lapse images of GFP-PCNA and histone H4-mCherry dynamics in **c** control strain SHR854 (*BGI1*) or **d, e** bridgin-null mutant strain SHR873 cells at 30 °C. Time measurements were made from the mitotic onset. Cell-cycle stages were scored for either PCNA localization or chromatin condensation and nuclear migration into the daughter bud. Scale bar, 2 μm. **d** Blue arrows indicate a chromosome that is separated from the compact chromatin mass prior to anaphase onset. **e** Blue arrows point to a lagging chromosome at the onset of mitosis through anaphase. **f** Time from mitotic entry to anaphase onset at 30 °C was quantified using a nuclear segregation marker, histone H4, dynamics, and plotted for wild-type control SHR854 (*BGI1*), bridgin-null mutant SHR873 (*bgi1Δ::HygB*), the Bgi1[FL] reintegrated strain SHR879, *mad2* null mutant SHR741 (*mad2Δ: NEO*), and the double-mutant strain SHR866 (*mad2Δ::NEO, bgi1Δ::HygB*). An internuclear distance of >1 μm was considered as the entry into anaphase. Mean ± S.D. is indicated. *n* indicates the number of live cells measured. Each dot represents a single live-cell measurement. Kruskal–Wallis one-way analysis followed by Dunn's multiple comparison test was used to calculate the statistical significance of differences (the *p* values show the difference compared to their respective controls, wild type or *mad2Δ*). **g** Serial 10-fold dilutions starting from $2 × 10^5$ cells were spotted for strains wild type (H99), *bgi1Δ* (SHR867), *mad2Δ* (SHR741), and *mad2Δ bgi1Δ* (SHR866) are shown. Source data are available as a Source Data file.

Bridgin is essential for mitotic progression and accurate chromosome segregation. Bridgin-null (*bgi1Δ*) strains were generated to further characterize the function of bridgin as a kinetochore protein. *bgi1Δ* cells exhibited reduced growth rates (Fig. 4a), and ~20% loss in viability as compared to its wild-type parent strain (Supplementary Fig. 4a). The *bgi1Δ* mutant cells also displayed an ~90-fold increase in the gross missegregation rate, which may account for the reduced viability of *bgi1Δ* (Fig. 4b and Supplementary Fig. 4a). These defects of *bgi1Δ* cells were complemented by the reintegration of the full-length bridgin gene, Bgi1[FL], expressed under the control of its native promoter

(Fig. 4b and Supplementary Fig. 4a). We subsequently examined how *bgi1Δ* affected cell-cycle progression. For this, previously described microscopy-based markers such as histone H4 and PCNA were utilized to determine cell-cycle stages[24,53]. The results were summarized in Supplementary Fig. 4b. In the wild type, the nucleus migrated into the daughter cell as a coalesced mass at the onset of M phase, followed by its equal segregation in anaphase (Fig. 4c). On the other hand, *bgi1Δ* cells exhibited lagging chromosomes at the M phase, suggesting inaccurate kinetochore–microtubule attachments (Fig. 4d, e). The live-cell analysis revealed that while wild-type cells spent an average of 18
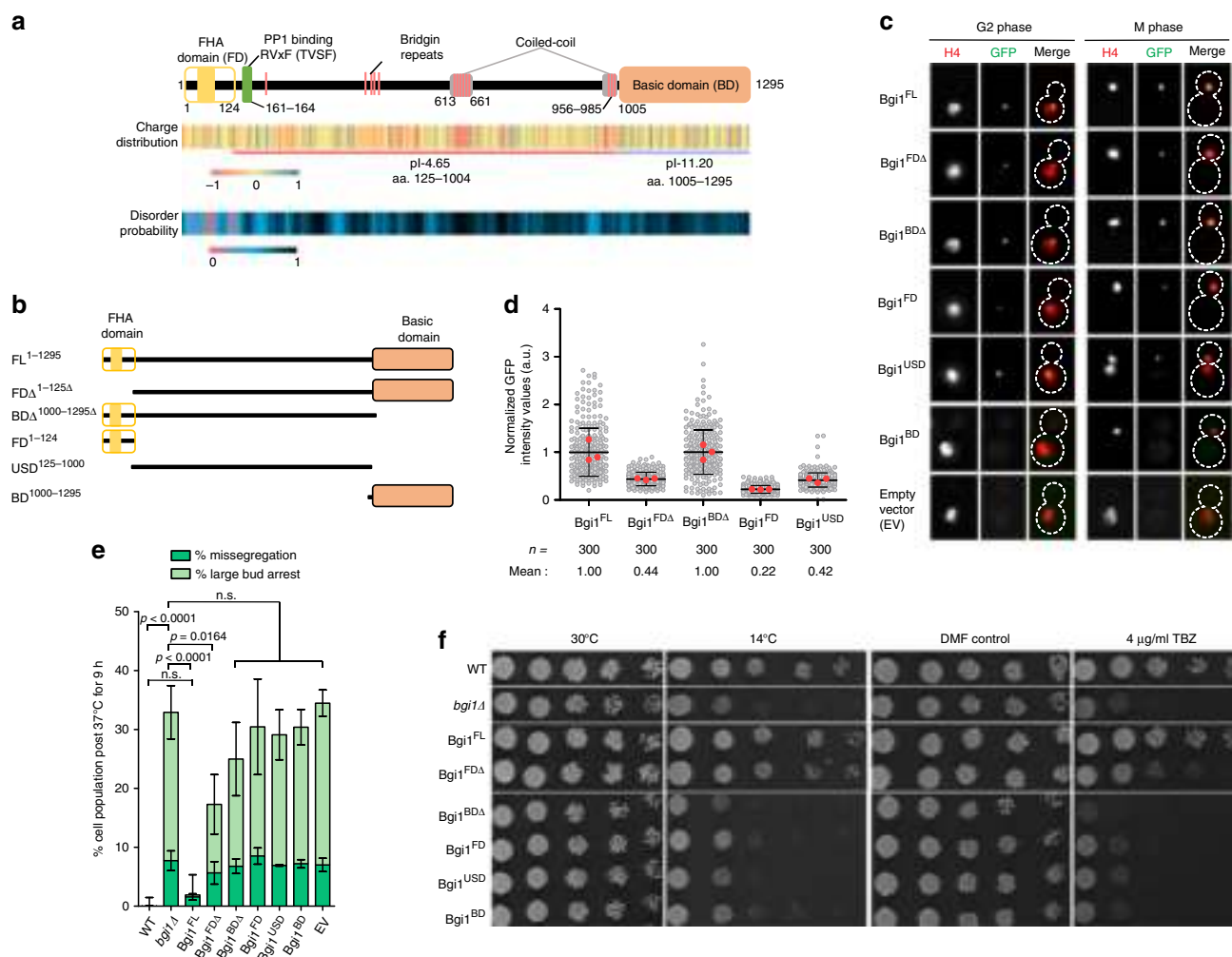
Fig. 5 The basic domain of bridgin is dispensable for its kinetochore recruitment but not for its function. a Schematic describing predicted features of bridgin. Charge distribution of amino acid residues was predicted with a window size of 2 using EMBOSS charge. The disorder probability of bridgin was calculated using IUPRED2A. **b** Schematic of domain deletion constructs of bridgin used in this study. Constructs were generated with a 3×FLAG-GFP tag at the amino terminus. **c** Representative cells of Bgi1FL (SHR879), Bgi1FDΔ (SHR913, *SHR832::3×FLAG-GFP-BGI1FDΔ*), Bgi1BDΔ (SHR880, *SHR832::3×FLAG-GFP-BGI1BDΔ*), Bgi1FD (SHR915, *SHR832::3×FLAG-GFP-BGI1FD*), Bgi1USD (SHR916, *SHR832::3×FLAG-GFP-BGI1USD*), Bgi1BD (SHR917, *SHR832::3×FLAG-GFP-BGI1BD*), and an empty vector (EV)/FLAG-GFP control (SHR918, *SHR832::3×FLAG-GFP*) depicting localization of the mentioned GFP constructs in G2 and M phases. Nuclear localization was scored for using the chromatin marker histone H4-mCherry. Scale bar, 3 μm. **d** Signal intensities of bridgin constructs mentioned in **c** were measured at metaphase and normalized to the mean intensity of Bgi1FL. The data represent the results of three independent experiments with 100 cells each. The red dot represents the mean of one experiment; mean ± S.D. is shown. **e** The extent of complementation of bridgin constructs mentioned in **c**, wild-type control CNVY121 (*BGI1*), and the bridgin-null mutant SHR832 was measured. Cells were grown at 30 °C until log phase and transferred to 37 °C. Indicated cell populations were measured 9 h post incubation at 37 °C. All values were normalized to wild-type control levels. Defects in nuclear segregation were measured as mentioned in Fig. 4b. The data represent the mean ± S.D. of three independent experiments. One-way ANOVA test with Tukey's multiple comparison test was used to calculate the statistical significance of differences between the net missegregation and large-bud arrest populations across strains. **f** Cells of varying numbers, 2×10^4, 2×10^3, 200, 100, and 50, of strains mentioned in **e** were spotted on plates as indicated. Dimethylformamide, DMF, solute control. Source data are available as a Source Data file.

min, *bgi1Δ* cells exhibited a delayed spending ~30 min in the M phase. This is an underrepresentation since ~10% of cells failed to exit the M phase even after >50 min (Supplementary Fig. 4c). Within the M phase, the delay occurred before anaphase onset (Fig. 4f and Supplementary Fig. 5f). Further analysis revealed that *bgi1Δ* cells were sensitive to spindle insults (Fig. 4g). Next, we tested whether *bgi1Δ*-associated defects are under the surveillance of SAC. *mad2Δ* in the background of *bgi1Δ* alleviated M-phase delay (Fig. 4f and Supplementary Fig. 4d). However, the double mutants (*mad2Δ bgi1Δ*) were conditionally synthetic lethal upon treatment of TBZ (2 μg/ml) or under conditions that compromise the spindle stability (14 and 37 °C) (Fig. 4g). Based on nuclear

segregation defects, sensitivities to spindle insults, and SAC-mediated delay in mitotic progression, we conclude that bridgin is important for accurate kinetochore–microtubule interactions required for high-fidelity chromosome segregation. These defects associated with *bgi1Δ* cells activate SAC response to correct for erroneous kinetochore–microtubule attachments.

The basic domain of bridgin is critical for its function. Next, we sought to understand how bridgin, being a part of the kinetochore, contributes to high-fidelity chromosome segregation. Bridgin is predicted to be 1295-amino acid (aa) long. Its N-terminal region (aa1–124) forms a forkhead-associated (FHA)

domain (a phosphopeptide recognition domain) followed by an unconventional putative PP1 docking site (aa161–164). In contrast, the remaining region of the protein is predicted to be largely unstructured (Fig. 5a). The C-terminal region (aa1005–1295) that has a pI of 11.20 was termed as the basic domain (BD). The unstructured domain (USD) was defined as the region spanning aa125–1004, which was acidic with a predicted pI of 4.65 (Fig. 5a), and contained 13 repeats of a bridgin consensus motif (Supplementary Fig. 5a and Supplementary Table 1).

Domain deletion constructs of bridgin (Supplementary Fig. 5b and Fig. 5b) wherein the various versions of bridgin are expressed under the control of its native promoter with an N-terminal 3×FLAG-GFP epitope tag. Each of these constructs was reintegrated into $bgi1\Delta$ cells expressing histone H4-mCherry to obtain strains expressing truncated bridgin proteins (Fig. 5b). Microscopic estimation of GFP signal intensities of the bridgin constructs suggested that the FHA domain (FD) and USD regions were able to localize independently of each other at the kinetochore, albeit to different extents of ~20% and ~40% of the Bgi1^{FL} level, respectively (Fig. 5c, d). Localization of Bgi1$^{BD\Delta}$ at the kinetochore was not significantly different from that of Bgi1^{FL} (Fig. 5c, d). Further, a lack of kinetochore localization of Bgi1^{BD} suggested that BD is dispensable for the kinetochore localization of bridgin. These results suggest that bridgin through its FD and USD makes multiple contacts at the kinetochore, consistent with the observation that bridgin is recruited downstream to multiple outer kinetochore KMN protein subunits (Fig. 2g).

To define the domains necessary for bridgin's function in chromosome segregation, we scored for complementation of the $bgi1\Delta$ phenotype at 37 °C. This assay condition was used for the ease of scoring due to the enhanced phenotype of M-phase delay (Fig. 5e). In addition, cell growth assays under conditions altering microtubule dynamics (Fig. 5f), and estimation of abnormal segregation defects at 30 °C was performed (Supplementary Fig. 5c). As expected, Bgi1^{FL} was able to suppress the $bgi1\Delta$ phenotype significantly (Fig. 5e, f). Partial complementation of phenotype was observed for the mutant expressing Bgi1$^{FD\Delta}$. No significant complementation was obtained for any other domain deletion constructs, including the Bgi1$^{BD\Delta}$ that localized to the kinetochore similar to Bgi1^{FL} levels (Fig. 5e, f). As observed at 37 °C, comparable results were observed for the measured rate of missegregation at 30 °C, albeit weak complementation was observed for Bgi1$^{BD\Delta}$ (Supplementary Fig. 5c). Taken together, we conclude that all domains, including BD, which is not involved in kinetochore localization of bridgin, are critical for restoring full bridgin function.

The C-terminal BD of bridgin interacts with chromatin. The impact of bridgin on SAC activity (Supplementary Fig. 5d, e) and spindle dynamics (measured at anaphase, Supplementary Fig. 5f) was tested and found to be unaltered in $bgi1\Delta$. Thus, these factors were ruled out as possible reasons for increased missegregation associated with $bgi1\Delta$ mutants. To address the role of BD toward bridgin function, FLAG-IP for Bgi1^{FL} (using 150 mM KCl and a more stringent condition of 300 mM KCl) and Bgi1$^{BD\Delta}$ (150 mM KCl) was performed and subjected to MS analysis (Fig. 6a and Supplementary Fig. 6a). A comparison of the specific interactors obtained across the Bgi1^{FL} and Bgi1$^{BD\Delta}$ suggested an enrichment of chromatin-interacting proteins, including inner kinetochore proteins, in Bgi1^{FL} IP as compared to Bgi1$^{BD\Delta}$ IP (Table 2 and Supplementary Data 3). In comparison, outer kinetochore KMN network proteins were identified across the Bgi1^{FL} and Bgi1$^{BD\Delta}$ IPs (Table 2). Proteins of the KMN network were among the top hits in Bgi1$^{BD\Delta}$ IP (Supplementary Data 3).

Based on the observation that chromatin-interacting proteins were less recovered in Bgi1$^{BD\Delta}$ IP, we hypothesized that BD might play a role in chromatin interactions. Through co-IP experiments, histone H4 was found to associate with Bgi1^{FL} (150 mM) as well as with the inner kinetochore component CENP-C^{Mif2} (Supplementary Fig. 6b). On the other hand, histone H4 was greatly reduced in Bgi1$^{BD\Delta}$ IP like the outer kinetochore protein IPs of Dsn1, Spc25, and Spc34 (Dam1C) (Supplementary Fig. 6b). These results raise the possibility that the bridgin–chromatin interaction in vivo occurs directly through BD and is not a consequence of mere bridgin assembly onto the centromere-localized KMN network platform. We further tested the direct interaction of purified bridgin BD with free DNA or reconstituted nucleosomes in vitro by electrophoretic mobility shift assay (EMSA). We found that BD was able to bind to DNA (Fig. 6b) and nucleosomes of varying compositions, containing histone H3 or CENP-A (Supplementary Fig. 6c). Taking together, MS results along with in vivo- and in vitro-binding assays strongly indicated that BD of bridgin interacts with DNA/chromatin.

Although BD seems to bind DNA nonspecifically in vitro, we cannot rule out the possibility that BD may exhibit specificity to *C. neoformans* centromere DNA/chromatin. Alternatively, other factors may restrict BD's centromeric localization in vivo. If the latter possibility is true, overexpression (OE) of bridgin may not alter its localization. The OE assay entailed that chromatin-unbound, but nuclear-localized proteins would diffuse into the cytoplasm owing to increased nuclear permeability in the M phase of *C. neoformans* (Fig. 6c). This is demonstrated by GFP-PCNA, a protein with well-established roles in S phase[54]. PCNA is nuclear localized in G2, but being unbound to chromatin in the M phase, it diffuses into the cytoplasm (see below). OE of GFP-Bgi1^{FL} (Supplementary 6d) resulted in its signals no longer being restricted to the kinetochores (Fig. 6d), rather signals were found to be overlapped with chromatin marked by histone H4 (Fig. 6e). Taking together, we conclude that bridgin interacts with bulk DNA/chromatin nonspecifically in vivo as well.

Further, using the OE strategy (Fig. 6e), we demonstrated that in the absence of BD, signals of Bgi1$^{BD\Delta}$-OE were restricted to a punctum, unlike Bgi1^{FL}-OE localization. Supporting the notion that bridgin localizes to the kinetochore through FD and USD, a single punctum for both overexpressed constructs, Bgi1^{FD} and Bgi1^{USD}, was observed. On the other hand, OE of constructs carrying BD such as Bgi1^{BD}-OE, Bgi1$^{FD\Delta}$-OE, or Bgi1^{FL}-OE, resulted in their localization to bulk chromatin. Thus, these observations suggested that BD is necessary and sufficient to bind chromatin in vivo, and the loss of BD in the OE constructs was sufficient to restrict bridgin localization to the kinetochore. Taken together, unlike other outer kinetochore proteins, it was surprising that bridgin was able to interact with centromeric chromatin.

Overall net positive charge of the BD is vital for bridgin function. Toward understanding the significance of bridgin binding to chromatin, we show that loss of bridgin function neither alters the previously described chromatin marks of CpG methylation at *C. neoformans* centromeres[47] (Supplementary Fig. 6e, f) nor H3K9me2 cellular pools previously shown to be predominantly enriched across centromeres and telomeres[55] (Supplementary Fig. 6g). To summarize, we observe that BD is dispensable for kinetochore localization of bridgin, but required for its function. However, it remains unclear how BD influences bridgin's function. We hypothesize two possibilities: either BD modulates bridgin function by interacting with specific proteins at chromatin, or BD does so by interacting directly with DNA/
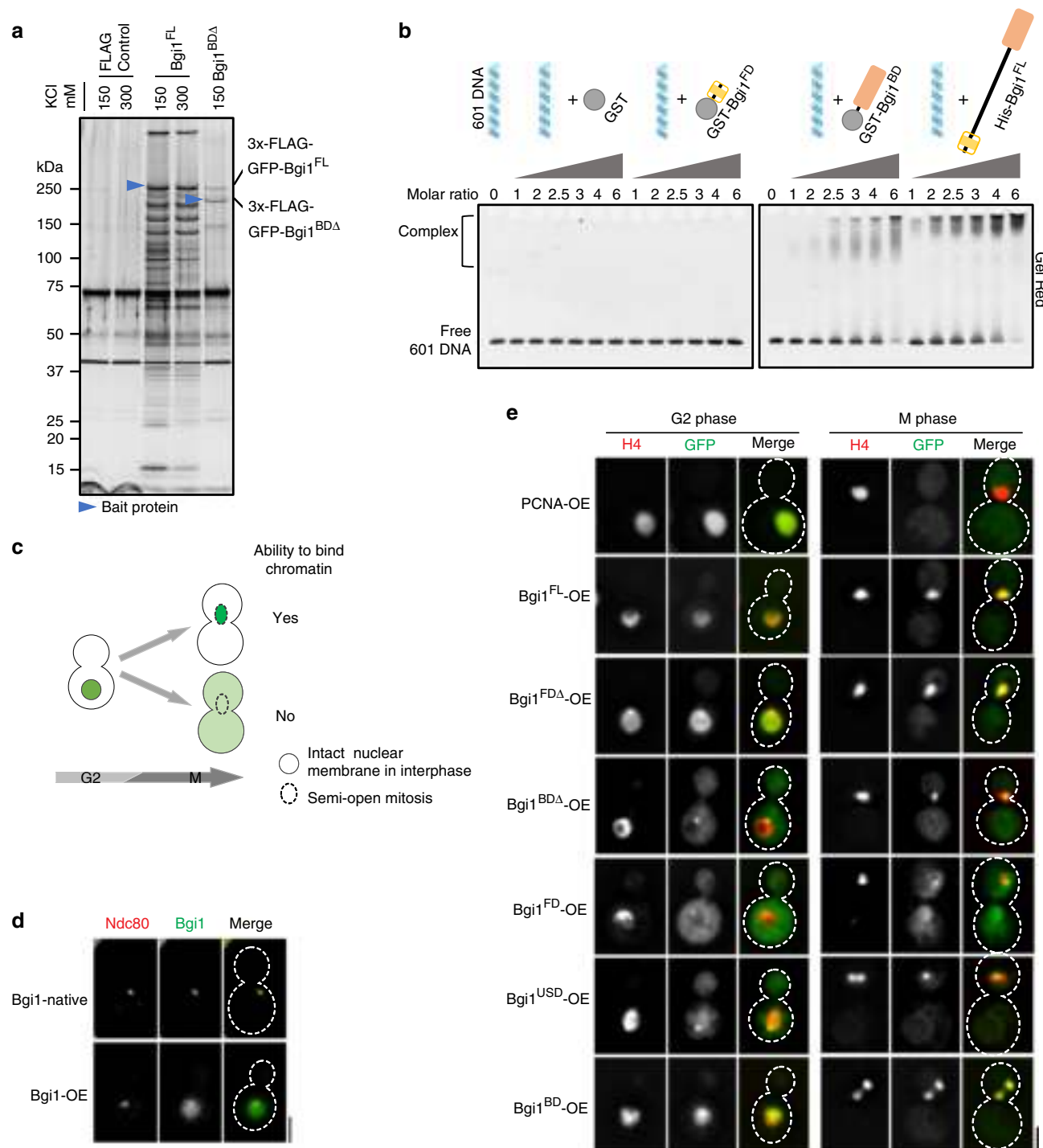
Fig. 6 The basic domain of bridgin has a property to interact with DNA nonspecifically in vitro and in vivo. a Proteins eluted from TBZ-treated cells after FLAG-IP of Bgi1^FL expressing SHR879, Bgi1^BDΔ expressing SHR880, and a FLAG control strain SHR942 expressing Bgi1 were separated on a gradient PAGE gel and silver-stained. Blue arrows mark the bait proteins. **b** Electrophoretic mobility shift assay (EMSA) samples were separated on a PAGE gel and stained with GelRed for visualization. **c** Chromatin-bound proteins (green) colocalize with the nuclear marker histone H4-mCherry in metaphase, while free nuclear proteins diffuse into the cytoplasm following the entry into mitosis. **d** Visualization of GFP-Bgi1 when expressed under the native (SHR847, *BGI1:: BGI1-V5-GFP*) or an overexpression (OE) *GAL7* promoter (SHR858, *BGI1::GAL7p-GFP-BGI1*). Outer kinetochore protein Ndc80-mCherry was used to mark the kinetochore. **e** Visualization of bridgin OE constructs. The expression of OE constructs of Bgi1^FL (SHR895, *SHR832::H3p-GFP-BGI1^FL*), Bgi1^FDΔ (SHR920, *SHR832::H3p-GFP-BGI1^FDΔ*), Bgi1^BDΔ (SHR921, *SHR832::H3p-GFP-BGI1^BDΔ*), Bgi1^FD (SHR922, *SHR832::H3p-GFP-BGI1^FD*), Bgi1^USD (SHR923, *SHR832::H3p-GFP-Bgi1^USD*) and Bgi1^BD (SHR924, *SHR832::3×FLAG-GFP-BGI1^BD*), and GFP-PCNA control (SHR854, *BGI1, CNVY121::H3p-GFP-PCNA*) was visualized in G2 and M phases. Histone H4-mCherry was used to mark chromatin. Scale bar, 3 μm. Source data are available as a Source Data file.

Table 2 List of chromatin-associated and kinetochore proteins obtained as interactors from bridgin IP's performed in Fig. 5a.

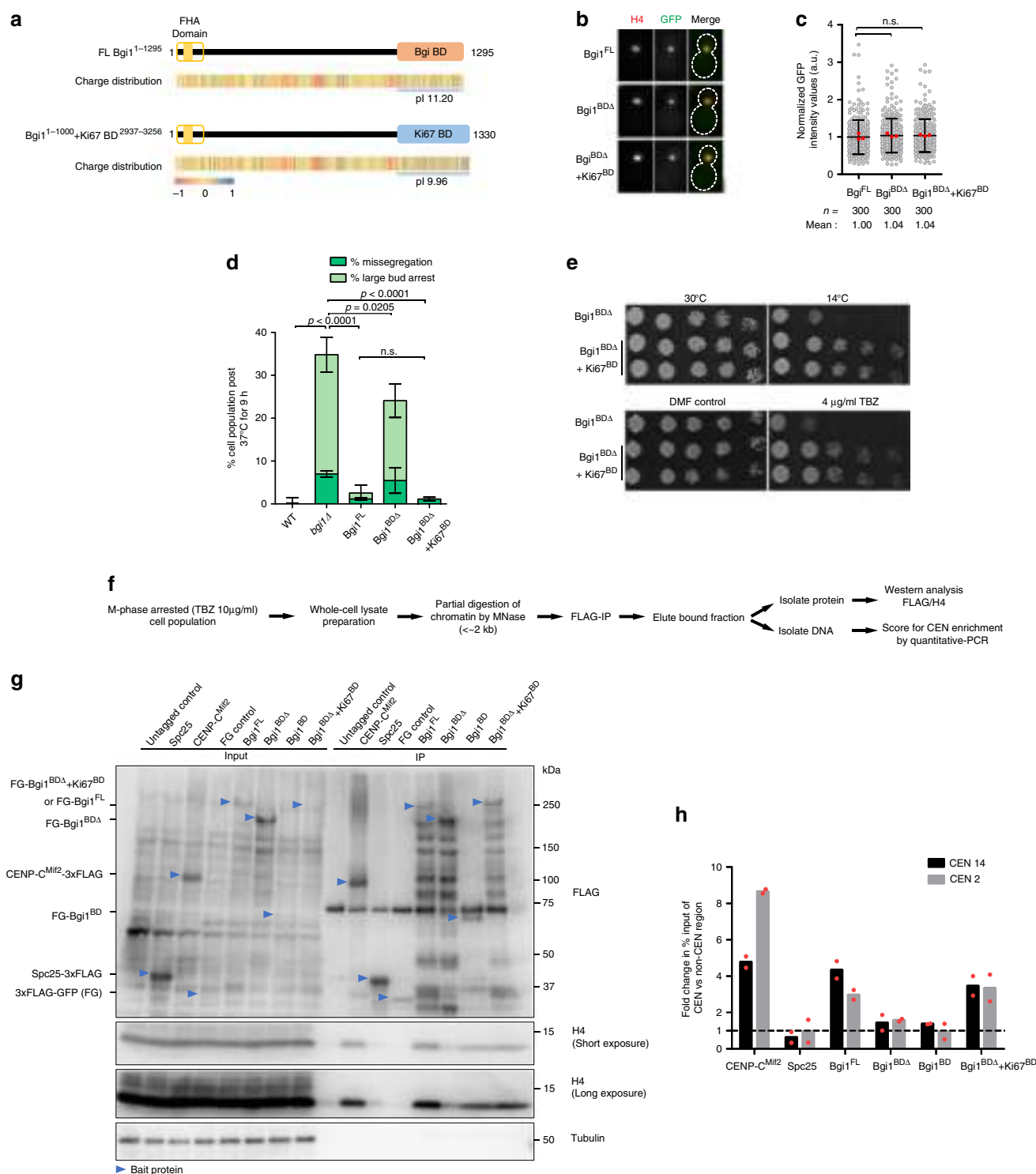| Complex/category	Protein	FLAG control 150 mM		FLAG control 300 mM		Bgi1^{FL} 150 mM KCl		Bgi1^{FL} 300 mM KCl		Bgi1$^{BD\Delta}$ 150 mM KCl	
		C (%)	TSC	C (%)	TSC	C (%)	TSC	C (%)	TSC	C (%)	TSC
Chromatin-associated proteins											
Histone proteins	H4	12	2	51	12	62	139	57	96	51	18
	H3	–	–	13	2	51	45	45	33	51	4
	H2A	6.9	1	12	3	46	30	45	27	12	3
	H2A.Z	6.5	1	6.5	2	47	13	49	13	6.5	2
	H2B	–	–	14	2	53	54	44	47	24	4
	H1	–	–	–	–	51	18	18	4	–	–
MCM complex	Mcm2	–	–	–	–	48	100	37	37	–	–
	Mcm3	–	–	–	–	56	136	45	56	–	–
	Mcm4	–	–	–	–	60	119	42	43	–	–
	Mcm5	–	–	–	–	58	85	35	28	–	–
	Mcm6	–	–	–	–	48	80	36	36	–	–
	Mcm7	–	–	–	–	62	119	39	43	–	–
DNA-binding proteins	Hmo1	–	–	–	–	36	27	–	–	–	–
	RFA1	–	–	–	–	66	70	63	53	–	–
	Ku70	–	–	–	–	40	24	44	35	–	–
	Ku80	–	–	–	–	41	34	20	15	–	–
Kinetochore proteins											
Ndc80	Ndc80	–	–	–	–	24	14	17	10	9.8	6
	Nuf2	–	–	–	–	30	14	31	12	27	14
	Spc25	–	–	–	–	39	15	46	20	18	6
	Spc24					39	12	57	20	21	4
Knl1	Sos7	–	–	–	–	37	12	35	16	2.5	1
	Knl1^{Spc105}	–	–	–	–	21	32	13	22	24	34
Mis12	Nsl1	–	–	–	–	15	3	30	6	3.9	1
	Nnf1	–	–	–	–	25	15	31	17	15	6
	Dsn1	–	–	–	–	34	22	27	19	16	10
	Mis12^{Mtw1}	–	–	–	–	37	11	53	21	14	5
CCAN	CENP–C^{Mif2}	–	–	–	–	4.7	4	–	–	–	–
CENP-A	CENP–A^{Cse4}	–	–	–	–	14	3	20	3	–	–
PP1	PP1	–	–	–	–	39	27	54	36	18	8
Bridgin	Bgi1	14	17	–	–	67	853	67	1113	30	87

Percentage amino acid sequence coverage (C,%) and the total spectrum counts (TSC) specific to the corresponding protein are tabulated.

chromatin. To prove further, we performed a domain-swap experiment by replacing bridgin BD$^{1005–1295}$ with an amino acid stretch of similar properties (length: ~300aa, unstructured, non-specific DNA binding and a pI of ~10) found in the basic region Ki67-BD$^{2937–3256}$ of human Ki67 (Fig. 7a). No evident sequence conservation was observed between the BD of Bgi1 and the basic region of Ki67 (Supplementary Fig 7a). Ki67 was previously shown to bind nonspecifically to DNA[56] and functions as a surfactant by coating chromosomes during mitosis[57]. However, Ki67 is not reported to interact with kinetochores[58]. We confirmed that OE of Ki67^{BD} localized to the entire nucleus, suggesting that it nonspecifically binds to DNA in *C. neoformans* (Supplementary Fig. 7b). Bgi1^{FL}, Bgi1$^{BD\Delta}$, and Bgi1$^{BD\Delta}$+Ki67^{BD} were expressed under the native bridgin promoter, and each of these versions of Bgi1 was found to be localized to the kinetochore with comparable intensities when integrated into a *bgi1Δ* background strain (Fig. 7b, c). Weak complementation of bridgin function by Bgi1$^{BD\Delta}$ over *bgi1Δ* was observed (Fig. 7d). Strikingly however, Bgi1$^{BD\Delta}$+Ki67^{BD} was able to complement defects observed in *bgi1Δ* to the extent that was comparable to Bgi1^{FL} (Fig. 7d and Supplementary Fig. 7c). These observations were additionally validated by complementation of phenotypes associated with the Bgi1$^{BD\Delta}$ mutant by the Bgi1$^{BD\Delta}$ + Ki67^{BD} chimeric protein in the spotting growth assay (Fig. 7e).

These results strongly indicated that the basic nature of BD contributes to bridgin function through its interaction with chromatin. To provide further evidence, we prepared partial MNase-digested chromatin (Supplementary Fig. 7d) to perform co-IP of various bridgin and kinetochore protein constructs with histone H4 (Fig. 7f-h). We also examined the enrichment of centromere DNA in these IP samples (Fig. 7h). These results show that Bgi1^{FL}, similar to the inner kinetochore linker protein CENP-C^{Mif2} but unlike the KMN network component Spc25, interacts with chromatin (Fig. 7g) and enriches at the centromere under these native conditions (Fig. 7h). Bgi1$^{BD\Delta}$, on the other hand, exhibited a reduced interaction with chromatin and did not strongly enrich at the centromeres (Fig. 7g, h). Further, Bgi1^{BD} was sufficient to interact with chromatin in vivo, but its binding was not restricted to centromere DNA (Fig. 7g, h). Since the Bgi1$^{BD\Delta}$+Ki67^{BD} chimera behaved like Bgi1^{FL}, we conclude that the intrinsic basic nature of BD, rather than the specific amino acid sequence, is necessary for bridgin function (Fig. 7g, h).

Bridgin maintains physiological levels of outer kinetochore proteins in the M phase. Having established that bridgin connects the outer kinetochore to centromeric chromatin after being recruited by multiple receptors at the KMN network, we were keen to examine how bridgin contributes to the kinetochore integrity. Fluorescent intensities of inner kinetochore proteins, CENP-A^{Cse4} and CENP-C^{Mif2}, and KMN network components of Mis12^{Mtw1}, Nuf2, and Knl1^{Spc105} were analyzed at metaphase and anaphase in wild-type and *bgi1Δ* cells. No significant difference in the levels of inner kinetochore proteins was observed (Supplementary Fig. 7e, f). However, the components of the KMN

a FL Bgi1$^{1-1295}$ — FHA Domain — Bgi BD — 1295 — Charge distribution — pI 11.20

Bgi1$^{1-1000}$+Ki67 BD$^{2937-3256}$ — FHA Domain — Ki67 BD — 1330 — Charge distribution — pI 9.96 — −1 0 1

b H4 / GFP / Merge — Bgi1^{FL}, Bgi1$^{BD\Delta}$, Bgi1$^{BD\Delta}$+Ki67^{BD}

c Normalized GFP intensity values (a.u.) — n.s. — Bgi^{FL}, Bgi1$^{BD\Delta}$, Bgi1$^{BD\Delta}$+Ki67^{BD}
n = 300 300 300
Mean: 1.00 1.04 1.04

d % cell population post 37°C for 9 h — % missegregation, % large bud arrest — WT, bgi1Δ, Bgi1 FL, Bgi1 BDΔ, Bgi1 BDΔ+Ki67-BD — $p < 0.0001$, $p = 0.0205$, $p < 0.0001$, n.s.

e Bgi1$^{BD\Delta}$ / Bgi1$^{BD\Delta}$ + Ki67^{BD} — 30°C / 14°C; DMF control / 4 µg/ml TBZ

f M-phase arrested (TBZ 10µg/ml) cell population → Whole-cell lysate preparation → Partial digestion of chromatin by MNase (<~2 kb) → FLAG-IP → Elute bound fraction → Isolate protein → Western analysis FLAG/H4; Isolate DNA → Score for CEN enrichment by quantitative-PCR

g FLAG; H4 (Short exposure); H4 (Long exposure); Tubulin — ► Bait protein

h Fold change in % input of CEN vs non-CEN region — CEN 14, CEN 2 — CENP-C^{Mif2}, Spc25, Bgi1^{FL}, Bgi1$^{BD\Delta}$, Bgi1^{BD}, Bgi1$^{BD\Delta}$+Ki67^{BD}

network, Mis12^{Mtw1} and Nuf2 in metaphase, and Mis12^{Mtw1} and Knl1^{Spc105} in anaphase were reduced by a small but significant extent (Supplementary Fig. 7g–i).

In summary, we identified a kinetochore protein bridgin, which derives its ability to coordinate accurate chromosome segregation in a unique way by interacting with the outer kinetochore and centromeric chromatin. We propose that bridgin connects the outer kinetochore with centromeric chromatin following its recruitment to the outer kinetochore KMN network that also restricts bridgin's binding to centromeric chromatin.

Discussion

Our in silico analysis suggested that most inner kinetochore CCAN proteins were lost multiple independent times in the fungal phylum of Basidiomycota. To understand the unusual kinetochore composition that might have compensated for the loss of CCAN proteins, we analyzed the kinetochore interactome in the model basidiomycete, an important human fungal pathogen *C. neoformans*. In this study, we identified an outer kinetochore protein bridgin, along with all other evolutionarily conserved kinetochore proteins that are predicted to be present in

Fig. 7 Interaction of bridgin with centromeric chromatin is established by the basic domain enriched with positively charged residues. a Schematic of the bridgin constructs wherein its basic domain was replaced with the basic DNA-binding domain from human Ki67. **b** Representative micrographs of Bgi1[FL] expressing SHR879, Bgi1[BDΔ] expressing SHR880, and the domain-swap chimera Bgi1[BDΔ] + Ki67[BD] expressing SHR926 (*SHR832::3×-FLAG-GFP-Bgi1[BDΔ]+Ki67[BD]*). Scale bar, 3 μm. **c** Signal intensities of bridgin constructs mentioned in **b** were measured at metaphase and normalized to the mean intensity of Bgi1[FL]. The data represent the results of three independent experiments with 100 cells each. The red dot represents the mean of one experiment; mean ± S.D. is shown. Kruskal–Wallis one-way analysis followed by Dunn's multiple comparison test was used to calculate the statistical significance of differences (the *p* values show the difference compared to Bgi1[FL]). **d** The extent of complementation of bridgin constructs mentioned in **c**, wild-type control CNVY121, and the bridgin-null mutant SHR832 was measured. Indicated cell populations were measured 9 h post incubation at 37°C. All values were normalized to wild-type control levels. Defects in nuclear segregation were measured as mentioned in Fig. 4b. The data represent the mean ± S.D. of three independent experiments. One-way ANOVA test with Tukey's multiple comparison test was used to calculate the statistical significance of differences between the net missegregation and large-bud arrest population across strains. **e** Cells of varying numbers, 2×10^4, 2×10^3, 200, 100, and 50, of Bgi1[BDΔ] expressing SHR880 and the domain-swap chimera Bgi1[BDΔ]+Ki67[BD] expressing SHR926 strains, were spotted on mentioned media plates. **f** Experimental design to determine Bgi1-BD interaction with chromatin in vivo by MNase IP. **g** FLAG immunoprecipitation of kinetochore proteins. 3×FLAG-tagged proteins of Spc25 (SHR861,*SPC25::SPC25-3×FLAG*), CENP-C[Mif2] (SHR896), FLAG-GFP (FG) control (SHR918), Bgi1[FL] (SHR879), Bgi1[BDΔ] (SHR880), Bgi1[BD] (SHR917), the domain-swap chimera Bgi1[BDΔ]+Ki67[BD] (SHR926), and untagged wild-type control (H99) were pulled down and levels of interacting histone H4 were determined. Bait proteins are marked by blue arrows. **h** Relative enrichment of FLAG-tagged proteins described in **g** at centromeres following the MNase IP. qPCR was performed using CEN2 and CEN14 and nonentromeric primer sets described in Fig. 2a followed by normalization to noncentromeric controls. The data represent the mean ± S.D. of two independent experiments. Source data are available as a Source Data file.

C. neoformans. Our bioinformatic predictions and experiments converge on the fact that the CENP-C[Mif2] pathway is the single known linker pathway connecting centromeric chromatin to the outer kinetochore in *C. neoformans*, reminiscent of *Drosophila melanogaster* and *Caenorhabditis elegans* like kinetochores[26,59–62]. We demonstrate that bridgin through its FD and unstructured domain (USD) is recruited by the outer kinetochore and interacts with centromeric chromatin via its basic domain (BD) (Fig. 8a). Strikingly, unlike previously defined linker proteins CENP-C[Mif2] [38,39,63], CENP-T[Cnn1] [14,28,64], or CENP-U[Ame1] [39,40], that require other inner kinetochore components for their recruitment[1], bridgin is recruited by the outer kinetochore KMN network. Thus, the identification and characterization of bridgin in this study reveal an alternate pathway originating at the KMN network to connect the outer kinetochore with centromeric chromatin (Fig. 8a).

We describe that bridgin BD alone cannot specifically localize to centromeric chromatin. In addition, we demonstrate that BD's binding to centromeric chromatin is reliant on its positively charged residues rather than the primary amino acid sequence. Based on these results, we hypothesize that the binding of BD to centromeric chromatin is a consequence of specific kinetochore recruitment of bridgin through its FD and USD (Fig. 8a). Further, optimum levels of bridgin help restrict its kinetochore localization. It is noteworthy that the kinetochore localization of bridgin is dependent on multiple KMN network components, Knl1[Spc105] and Sos7 of the Knl1C, and the Mis12C–Ndc80C platform. Coincidentally, Bgi1[FD] and Bgi1[USD] independently exhibit kinetochore localization. These observations prompt us to speculate that each of the individual domains, FD and USD, may regulate interactions with a specific KMN component and their synergistic activity is important for maintaining cellular levels of bridgin at kinetochores. We are unable to determine the independent contribution of Ndc80C and Mis12C, due to their interdependency, on bridgin recruitment. Reciprocally, the reduced levels of KMN components at the M phase in *bgi1Δ* cells may be a consequence of bridgin's influence on the maintenance of the KMN components. These questions and findings necessitate further examination in the future.

Dsn1 autoinhibition[34,35,37], which diminishes the Mis12C–CENP-C interaction, must be overcome by multiple kinetochore linker pathways. The essentiality of linker pathways to overcome Dsn1 autoinhibition is species-specific. While the Dsn1 autoinhibitory domain is absent in *D. melanogaster*[60,65], that possesses a single kinetochore linker pathway CENP-C, *C.*

neoformans retains the domain (Supplementary Fig. 8a). Although a recent study suggests Nnf1 to be the Dsn1 homolog in *D. melanogaster*, we were unable to identify the presence of the Dsn1 autoinhibitory domain in the suggested homolog[26]. Bridgin levels at the kinetochore reach the peak at anaphase, a stage of the cell cycle when Aurora B[Ipl1]-mediated Dsn1 phosphorylation is suggested to be countered by the phosphatase activity of PP1[34,35] (Fig. 8a). The sharp reduction of Aurora B[Ipl1] localization at anaphase kinetochores[53] and the conservation of the amino terminus of CENP-C[Mif2] (Supplementary Fig. 8b) are observed in *C. neoformans*. Taken together, we propose that the kinetochore architecture alters during the metaphase–anaphase transition and that the bridgin pathway functions to reinforce/stabilize the outer kinetochore. It will be intriguing to test whether the presence of Dsn1 autoinhibition can provide a constraint driving evolution/ maintenance of multiple outer kinetochore linker pathways required for outer kinetochore reinforcement in organisms with monocentric chromosomes.

Outer kinetochore proteins are found to be more conserved than their inner kinetochore counterparts, including linker proteins across eukaryotes[25,26]. Thus, additional connections mediated by factors recruited by KMN proteins, like the bridgin pathway, may provide cells with an effective alternative toward outer kinetochore reinforcement. Bridgin homologs are identified across all subphyla in Basidiomycota (Fig. 8b). Strikingly, an inability to identify bridgin homologs in specific orders correlate with the presence of multiple known linker pathways (Fig. 8b). It would be worth investigating whether the presence of the bridgin pathway allowed more flexibility in the retention/loss of specific linker pathways in basidiomycetes. Genome sequencing of many basidiomycetes would help address the correlation. Further, it would be intriguing to recognize the contribution of the multiple linker pathways in organisms like *Ustilago maydis*, which retained CENP-T, CENP-C in addition to bridgin.

Bridgin homologs could be identified in the basal ascomycetous fungi of the class Pneumocystidales, such as *Pneumocystis jirovecii* (causative organism of pneumonia), and Taphrinales. Moreover, the existence of bridgin-like proteins outside fungi may suggest an ancient origin of this family of kinetochore proteins containing an FD (Supplementary Fig. 8c and Supplementary Data 4). Ki67, a component of the mitotic chromosome periphery[57,66] but not reported to be associated with the kinetochore[58], is the identified bridgin homolog in metazoa (Supplementary Fig. 8c). This raises interesting questions regarding the functional divergence and neofunctionalization of bridgin
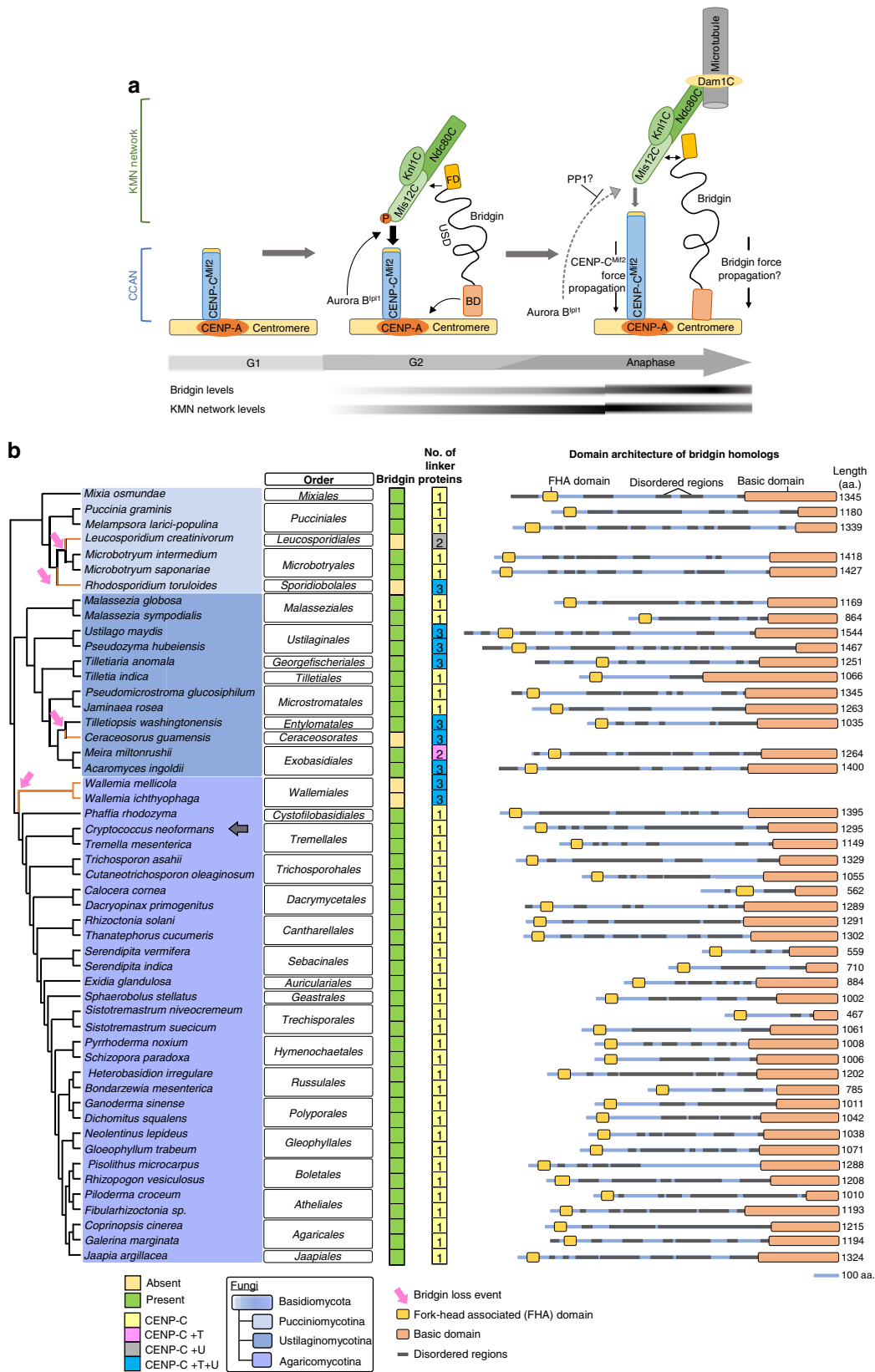
homologs invertebrates. Future experiments will reveal how bridgin is recruited onto kinetochore, how its chromatin binding is regulated, and to what extent bridgin can bear the load at the kinetochore.

There are limited experimental studies on kinetochores in nontraditional model eukaryotes[25,67] and it, therefore, remains unclear whether there are lineage-specific kinetochore components because proteins such as bridgin cannot be identified by means of bioinformatics using known kinetochore proteins as a query. Thus, the identification of such kinetochore components emphasizes the significance of studying nonconventional systems.

Fig. 8 Bridgin, a protein conserved across most basidiomycete species that lost one or more linker proteins, can bridge the outer kinetochore KMN network and centromeric chromatin. a A model describing bridgin as a kinetochore protein connecting the outer KMN network, through its FD and USD, and centromeric DNA directly via its BD. Restricted interaction of bridgin with centromeric chromatin in G2/M cells is a possible consequence of its outer kinetochore-specific recruitment. Increased bridgin localization is observed in anaphase. **b** Identification of bridgin homologs across Basidiomycota. *Left*, the presence (green box) or absence (yellow box) of a bridgin homolog is represented. The number of identified linker pathways has been mentioned and color-coded to represent the linker pathway(s) present. Bridgin loss events in basidiomycete lineages are represented by orange lines in the cladogram. *Right*, the domain architecture of identified bridgin homologs in Basidiomycota.

Methods

Homolog detection. All searches were carried out in the NCBI nonredundant protein database and the UniProtKB database as available in July 2019. Searches for kinetochore homologs were initially carried out using iterative HMMER[68] jackhammer searches (E value $\leq 10^{-3}$) with Pfam models for the mentioned kinetochore proteins. When available, models of both yeast and metazoan kinetochore homologs were considered. Obtained hits were validated by performing reciprocal HMMER searches. The secondary structure of obtained hits was validated using Jpred4 and tertiary structure prediction using HHpred[69] and/or Phyre2[70]. Protein sequences that were unable to produce hits upon reciprocal searches or failed to conform to expected secondary and tertiary structures were discarded. Further searches were performed with the same criteria using identified homologs phylogenetically closest to the species in question. Species considered in the study are mentioned in Supplementary Data 1, when homologs were not identified from a specific strain, an obtained homolog from another strain of the same species was considered. Obtained hits when possible were validated with the identified homologs from *C. neoformans*. Known kinetochore homologs from *S. cerevisiae*, *S. pombe*, *D. melanogaster*, and *H. sapiens* were used to draw the matrix of kinetochore homologs.

Toward identifying homologs of bridgin, the conserved FD was taken as the bait for subsequent iterative HMMER jackhammer searches. Obtained hits were further screened for overall protein architecture (amino-terminus FD, an unstructured central region, and a basic carboxy-terminus (Fig. 8b and Supplementary Fig. 8c). The probability of protein disorder was predicted using IUPred2A[71], and pI of the amino acid residues was predicted using ProtParam[72]. Using published multigene and genome-scale phylogenetic data from The Fungal Kingdom[48], JGI MycoCosm[73], Interactive Tree of Life (iTOL) v4[74], and Wang et al.[75], the cladograms were drawn showing the relationship among the considered species.

Sos7 and bridgin alignments were generated using T-COFFEE v11.0 and visualized using JalView v2.11.1.0.

Yeast strains and plasmids. A list of strains and plasmids used in the study can be found in Supplementary Table 2 and Supplementary Table 3, respectively. Primers used to generate the constructs are mentioned in Supplementary Table 4. Conditional kinetochore mutant strains were grown on 1% yeast extract, 2% peptone, and 2% galactose (YPG). All other strains were grown in 1% yeast extract, 2% peptone, and 2% dextrose (YPD) at 30 °C, 180 rpm, unless mentioned otherwise. Strains were maintained on YPD/YPG solidified with 2% agar and stored at 4 °C or −80 °C in 15% glycerol. Yeast strains are based on the haploid-type strain H99 or KN99 and generated by the standard procedure as previously described[24]. In brief, created native tagging and the *GAL7* promoter[76] replacement cassettes were excised from the plasmid construct, appropriate restriction enzymes linearized OE cassettes, and deletion cassettes were generated by overlap polymerase chain reaction (PCR) and transformed into *C. neoformans* strains of appropriate background by biolistic transformation[77]. Transformed cells were selected on drug selection in YPG for the *GAL7* promoter replacement strains to generate conditional mutants and YPD for all other strains.

Protein affinity purification and native ChIP. An overnight culture was inoculated at 0.1 OD$_{600}$ into fresh YPD, grown until ~0.7 OD$_{600}$, and treated with 10 μg/ml of TBZ for 3 h. Cells were harvested, washed once in water followed by one wash with binding buffer BB150 (25 mM HEPES, pH 8.0, 2 mM MgCl$_2$, 0.1 mM EDTA, 0.5 mM EGTA, 0.1% NP-40, 150 mM KCl, 1× complete EDTA-free protease inhibitor (Roche), 1× PhosStop (Roche), and 15% glycerol). Cells were resuspended in binding buffer (100 OD$_{600}$/ml). Bead beating was used to lyse the cell suspension until ~80% cell lysis was obtained. Lysates were centrifuged at 20,000g for 20 min, and the supernatant was collected. The extracted cell lysate was incubated with anti-FLAG M2 antibodies (Sigma) conjugated to Dynabeads™ M-280 sheep anti-mouse IgG (ThermoFisher Scientific) for 2 h at 4 °C, under constant rotation. Unbound proteins were collected as flow-through, and proteins bound to antibody-conjugated beads were washed five times with BB150 w/o glycerol. Invert mixing was followed during each wash. Bound proteins were eluted in BB150 w/o glycerol + 200 μg/ml of 3× FLAG peptide (Sigma). Two elutes of ½ volume each of the initial bead volumes was taken and pooled.

In all, 1 μg of Anti-FLAG M2 antibody was conjugated to 10 μl of Dynabeads™ M-280 sheep anti-mouse IgG (ThermoFisher Scientific) in 1× phosphate-buffered saline (PBS), pH 7.4, and incubated for 1 h at room temperature (RT). It was washed twice with 1× PBS and resuspended in PBS. These anti-FLAG-conjugated

beads were used for the lysate prepared from 100 OD$_{600}$ culture. Affinity purification samples that were processed subsequently for MS were started from a 2.25L culture, yielding ~4500 OD$_{600}$ cells. In all, 300 mM KCl, where mentioned, was used throughout the affinity purification experiment as part of the binding buffer yielding BB300.

For GFP affinity purification, GFP-Trap agarose beads (ChromoTek) were used. Bound proteins were eluted by boiling the beads for 10 min in 1× sample loading buffer (50 mM Tris-HCl, pH 6.8, 2% sodium dodecyl sulphate (SDS), 0.05% bromophenol blue, 10% glycerol, and 5% 2-mercaptoethanol) and the supernatant was collected. Other steps of the affinity purification protocol were kept the same as mentioned above.

For MNase co-IP and native ChIP, lysates prepared after bead beating were subject to centrifugation at 3300g for 10 min. Following this, the supernatant was treated with MNase at RT to obtain digested chromatin (~<2 Kb in length). Post digestion, the solution was centrifuged at 20,000g for 20 min. Subsequent steps followed were as described for the affinity purifications. For MNase native ChIP, DNA from the elute and input sample was extracted using GeneJET PCR purification kit (ThermoFisher Scientific). qPCR for input and IP was set up using centromere primers CN1 (CEN14)-5′-CCATCCAGTTCTTGCTTGAG-3′ 5′-GCA AGGAATGTGTTGTCTGG-3′ and CN3 (CEN2)-5′-CAGACCCTTCCTTCAG CCG-3′ 5′-TGGCAAGGAGTCGTCAGCG-3′ and noncentromeric primer set NC3 5′-GATCAAGTATAGGCGAAGG-3′ 5′-ATCTCTTATTCCCACTTCTACTC-3′ located ~825 kb away from the centromere on chromosome 1. The reactions were supplemented with SYBR green (BioRad). Data were obtained using StepOne v2.3 (Applied Biosystems). Values were normalized to percent input observed in centromeric versus noncentromeric regions and plotted using GraphPad Prism.

Immunoblot analysis. For whole-cell lysates, 3 OD$_{600}$ cells were harvested and resuspended in 15% TCA overnight. About 500-μl 0.5-mm glass beads were added, and samples were vortexed for a total time of 15min, with intermittent cooling on ice. They were centrifuged at 15,000g for 10 min, and the obtained pellet was washed twice with 100% acetone, air-dried, and resuspended in 1× sample loading buffer and boiled for 10 min. Samples were separated on a sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to Immobilon-P (Merck).

For Supplementary Figs. 1c, 2d, e, and 6g, primary antibody and secondary antibody dilutions were made in skim milk. Proteins bound by antibodies were detected with Clarity western ECL (BioRad) and visualized with Versadoc (BioRad). For Fig. 7g and Supplementary Fig. 1a, 2b, 6a, b, d primary and secondary antibody dilution were prepared in Signal Enhancer Hikari (Nacalai tesque). ChemiDoc Touch (BioRad) was used to visualize proteins reacting with antibody in the presence of the substrate ECL Prime (GE Healthcare). ImageJ[78,79] and Image lab v6.0.0 (BioRad) was used to visualize and process images. Antibodies used are tabulated in Supplementary Table 5. Uncropped blots are presented in the Source Data File.

Mass spectrometry. Affinity-purified samples were separated on an SDS-PAGE followed by silver staining. Silver-stained gels were visualized using an Epson GT-X980 scanner equipped with Epson Scan (Seiko Epson) v5.1.1.0. Isolated samples from the stained gel were Trypsin digested. Samples were subject to nano LC–MS–MS as described previously[80]. Briefly, after the gel slices were treated with DTT and iodoacetamide, in-gel digestion with 10 μg/ml modified trypsin (Sequencing grade, Promega) was performed at 37 °C for 16 h. The digested peptides were extracted with 1% trifluoroacetic acid and 50% acetonitrile, dried under a vacuum, and dissolved in 2% acetonitrile and 0.1% formic acid. The peptide mixtures were then fractionated by C18 reverse-phase chromatography (3 μm, ID 0.075 × 150 mm, CERI, ADVANCE UHPLC, AMR Inc.) at a flow rate of 300 nL/min with a linear gradient of 5–35% solvent B over 20 min. The compositions of solvents A and B were 0.1% TFA in water and 100% acetonitrile, respectively. The data-dependent acquisition was performed by a hybrid linear ion trap mass spectrometer (LTQ Orbitrap Velos Pro, ThermoFisher Scientific) with Advanced Captive Spray SOURCE (AMR Inc.). The mass spectrometer was programmed to carry out 11 successive scans. The first scan was performed as a full-MS scan over the range 350–1800 m/z using Oubitrap analyzer at a resolution of 60,000. The second-to- eleven scan events were detected by ion trap analyzer with automatic data-dependent MS/MS scans of the top 10 most abundant ion obtained in the first scan. MS/MS spectra were obtained by setting relative collision energy of 35% CID and exclusion time of 20 s for molecules of the same m/z value range.

Using the MASCOT ver2.6.2 search engine in Proteome Discoverer 2.1.1.21 and 2.2.0.388 (ThermoFisher Scientific), the obtained spectra peaks were assigned using the UniProt proteome database for *C. neoformans* H99 database (ID: UP000010091 20171201downloaded (7340 sequences)). Fragment tolerance 0.80 Da (Monoisotropic), parent tolerance 10 PPM (Monoisotropic), fixed modification of +57 on C (Carbamidomethyl), variable modification of +16 on M (oxidation) and +42 on peptide amino terminus (Acetyl), and allowing for a maximum of 2 missed cleavages for CENP-C^{Mif2}, Dsn1, and Spc25 and 3 missed cleavages for bridgin samples. Scaffold (version Scaffold_4.10.0, Proteome Software Inc., Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they exceeded specific database search engine thresholds. Protein identifications were accepted if they contained at least two identified peptides. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters. Identified protein hits from CENP-C^{Mif2}, Dsn1, Spc25, and their untagged controls can be found in Supplementary Data 2. To relatively quantitate protein abundance obtained within each of the experiments, Bgi1^{FL} 150 mM, Bgi1^{FL} 300 mM, and Bgi1^{BDΔ} 150 mM, emPAI[81] (exponentially modified protein abundance index) values were determined using Scaffold 4.10.0 (Proteome Software). The higher the emPAI score, the more abundant the protein is in the mixture. Supplementary Data 3 summarizes the identified interacting protein hits from Bgi1^{FL}, Bgi1^{BDΔ}, and their untagged control IPs.

Cross-linked ChIP and quantitative real-time PCR. ChIP assays were performed with some modification of previously described protocols[82,83]. In brief, 100 ml of Bgi1-GFP strain was grown until ~1 OD$_{600}$. Cross-linking was performed for 20 min using formaldehyde to a final concentration of 1% and incubated at RT with intermittent mixing. The reaction was quenched by the addition of 2.5 M glycine and further incubated for 5 min. Fixed cells were harvested by centrifugation and resuspended in 9.5 ml of deionized water, followed by the addition of 0.5 ml of 2-mercaptoethanol and incubated at 30 °C for 60 min at 180 rpm. Cells were pelleted and resuspended in 10 ml of spheroplasting buffer (1 M sorbitol, 0.1 M sodium citrate, and 0.01 M EDTA) containing 40 mg of lysing enzyme from *Trichoderma harzianum* (Sigma). Spheroplasts were washed once with 15 ml each of the following buffers: (1) 1× PBS, pH 7.4, (2) Buffer I (0.25% Triton X-100, 10 mM EDTA, 0.5 mM EGTA, and 10 mM Na-HEPES, pH 6.5), and (3) Buffer II (200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, and 10 mM Na-HEPES, pH 6.5). Following which the spheroplasts were resuspended in 1 ml of extraction buffer (50 mM HEPES, pH 7.4, 1% Triton X-100, 140 mM NaCl, 0.1% Na-Deoxycholate, and 1 mM EDTA) and sonicated to shear chromatin using a Bioruptor (Diagenode) for 30 cycles of 30-s on and 30-s off bursts at high-intensity setting. Sheared chromatin was isolated in the supernatant fraction after centrifugation for 15 min at 15,000g. Average chromatin fragment sizes ranged from 200 to 500 bp. In all, 100 μl, 1/10th the volume, of the chromatin fraction, was kept for input DNA preparation, the remaining chromatin volume was divided into two halves of 450 μl each for (+) antibody and (−) antibody. For (+) antibody, 20 μl of GFP-Trap agarose beads (ChromoTek) and 20 μl of blocked agarose beads (ChromoTek) was added to (−). The tubes were incubated for 8 h to overnight on a rotator at 4 °C. Following which the supernatant was isolated as flow-through, and the beads were washed twice with low-salt buffer, twice with the high-salt buffer, once with LiCl buffer, and twice with TE. Bound chromatin was eluted in two 250-μl elution using elution buffer. All three fractions (SM, (+Ab), and (−Ab)) were de-cross-linked (mixed with 20 μl of 5 M NaCl and incubated at 65 °C for 8 h to overnight), Proteinase K treated (10 μl of 0.5 M EDTA, 20 μl of 1 M Tris-HCl, pH 6.8, 40 mg of Proteinase K was added to the solution and incubated for up to 2 h at 45 °C). DNA was isolated using phenol:chloroform extraction followed by ethanol precipitation. Isolated DNA was air-dried and dissolved in 25 μl of deionized water containing 25 μg/ml of RNase (Sigma).

All three samples (SM, (+), and (−) antibody) were subject to real-time quantitative PCR. The reaction mixture was set up using the iTaq™ universal SYBR green Supermix (BioRad) with 1 μl of the undiluted (+Ab), (−Ab) DNA samples, and SM (diluted 1:50). qPCR primers described above for the native ChIP were used. Data were obtained using CFX manager v3.1 (BioRad). Enrichment at centromeric regions was normalized to a control noncentromeric region. Values were plotted using GraphPad Prism. For ChIP-PCR assays, the obtained PCR products from SM, (+), and (−) antibody were separated on 2% agarose gels and visualized on BioRad Gel Doc XR using Quantity One v4.6.9.

Microscopy and analysis. Overnight cultures grown in YPD were subcultured into fresh YPD at 0.1 OD$_{600}$ and grown until 0.4–0.6 OD$_{600}$. Cells were isolated, washed twice in 1× PBS, and mounted on slides. Images for Supplementary Fig. 1k (CNAG_01340) were acquired using the Airyscan mode in the Zeiss LSM 880 confocal system equipped with an Airyscan module, 63× Plan Apochromat 1.4 NA. Z stacks were obtained at an interval of 166 nm, 488/516- and 561/595-nm excitation/emission wavelengths were used as GFP and mCherry, respectively. Airyscan images were processed using Zen Black v2.3 (Zeiss) and visualized in ImageJ[78,79]. Images for Fig. 1d and Supplementary Fig. 1f, k were acquired in the Zeiss LSM 880 confocal system equipped with GaAsp photodetectors. Z stacks were obtained at an interval of 300 nm, 488 nm, and 561-nm excitation was used for GFP and

mCherry, respectively, and emission between 490–553 nm and 571–651 nm was captured. Images are represented as maximum-intensity projections.

Live-cell microscopy, images for kinetochore quantitation, and microscopy-based assays were acquired using the Zeiss Axio Observer 7, equipped with Definite Focus.2, Colibri 7 (LED light source), TempController 2000-2 (PECON), 100× Plan Apochromat 1.4 NA objective, pco.edge 4.2 sCMOS, and Andor iXon Ultra 897 electron-multiplying charge-coupled device (CCD). Zen 2.3 (blue edition) was used for image acquisition and controlling all hardware components. Filter set 92 HE with excitation 455–483 and 583–600 nm for GFP and mCherry, respectively, and corresponding emission was captured at 501–547 and 617–758 nm. To limit the time taken for an image, a complete Z stack was obtained for each channel before switching.

For live-cell microscopy, an overnight culture grown in YPD was subcultured into fresh YPD at ~0.1 OD$_{600}$ and grown for 2–3 generations until 0.4–0.8 OD$_{600}$. Cells were harvested, washed in 1× PBS, and resuspended in synthetic complete media with 3% dextrose. Cells were mounted onto an agarose pad (3% dextrose, 3% agarose in synthetic complete media) and sealed with petroleum jelly. Images were captured at time intervals of 0.5, 1, 2, or 4 min, as appropriate, with an EM gain of 300 and Z interval of 300 nm. Z-stack projection of images is represented. For live-cell quantitation of kinetochore signal, signal intensities were measured after the projection of Z stacks.

To study kinetochore interdependency, *GAL7p* conditional strains were grown overnight in YPG, subcultured at 0.2 OD$_{600}$, and grown until 0.8–1 OD$_{600}$. Cells were washed and resuspended in 1× PBS. Following which cells were inoculated into YPD (repressive, −) and YPG (permissive, +) at 0.1 OD$_{600}$. Timepoints chosen for the microscopic analysis of the repression phenotype were based on when the repressed kinetochore protein signals were no longer detectable at the kinetochore, which often coincided with an ~90% loss in cell viability (Supplementary Fig. 2g). Images were acquired after 6, 12, 15, 18, 9, and 18 h for CENP-C^{Mif2}, Mis12^{Mtw1}, Nuf2, Knl1^{Spc105}, Dad1, and Dad2, respectively. Z stack was obtained at an interval of 300 nm. Single Z slice representing the maximum intensity of the tagged kinetochore proteins was represented. Quantitation of kinetochore signal was performed from large budded cells (budding index ~0.55–0.90).

To estimate the population of large buds and cells with segregation defects, cells were grown until early-log phase 0.8–1 OD$_{600}$ after subculture from an overnight culture. They were imaged using the above-mentioned sCMOS camera with a Z interval of 300 nm. Cells with a budding index of >0.55 were considered as large budded cells in mitosis. Chromatin marked with a GFP/mCherry-tagged histone H4 construct was used to observe missegregation events. In total, >1000 cells were measured for each independent experiment to estimate segregation defects. To calculate the percent large budded population after 10 μg/ml TBZ treatment, >300 cells were measured for each independent experiment.

Images for the OE assay of bridgin strains are representative maximum-intensity projection images. KMN network and Dam1C components reach the peak intensity at metaphase, while bridgin exhibits maximum kinetochore intensity at anaphase. For quantitation of kinetochore signals, measurements were performed at the M phase to capture their signal at peak intensity, where possible. Components of the outer kinetochore are transiently localized to the kinetochore from G2 to M. Thus, clustered kinetochore signals in the daughter bud (budding index of ~0.55–0.8) provided us with a reliable stage to score for the kinetochore signal intensities. To maintain uniformity and ease of measurement, bridgin and inner kinetochore signal intensities were also measured at metaphase. Kinetochore signal measurement in interdependency assays and *bgi1Δ* background were measured from the in-focus Z plane exhibiting the most intense signal. Background signal measured from a region neighboring the kinetochore-measured signal in the same plane of the equal area was subtracted from the measured kinetochore intensity and normalized to the appropriate control and plot using GraphPad Prism v5.00 or v7.0a (GraphPad software). All acquired images were processed in ImageJ[78,79]. For images wherein brightness and contrast are modified, the settings were applied uniformly across the entire image.

Budding index calculation. Budding index of a cell is defined as the ratio obtained by

(1)   The diameter of the daughter cell/diameter of the mother cell

The diameter of the daughter and mother cell was measured along the mother–daughter axis using the line tool in ImageJ[79].

Generation of recombinant proteins. GST, GST-Bgi1^{FD} (residues 1–130), and GST-Bgi1^{BD} (residues 1000–1295) were expressed from pGEX-6P-1 (GE Healthcare) in Rosetta2 (DE3) (Merck). GST and GST-Bgi1^{BD} were induced for expression using 1 mM IPTG for 3 h at 37 °C. GST-FD was induced for expression overnight at 16 °C using 0.2 mM IPTG. Cells were harvested and lysed in lysis buffer (20 mM HEPES, pH 7.5, 300 mM NaCl, 1 mM EDTA, 0.5 mM TECP, and 1× complete EDTA-free protease inhibitor (Roche)) and 1× PBS with 1× complete EDTA-free protease inhibitor (Roche) for GST and GST-Bgi1^{FD}. GST fusion proteins were affinity purified using Glutathione sepharose 4b beads (GE Healthcare) and eluted using 20 mM glutathione. GST-Bgi1^{FD} and GST-Bgi1^{BD} were further purified using anion exchange chromatography. The column was

equilibrated using 20 mM Tris-HCl pH 7.5, 1 mM DTT. An elution gradient of 5–75% NaCl was achieved using 20 mM Tris-HCl, pH 7.5, 1 M NaCl, and 1 mM DTT. Relevant fractions were pooled, concentrated in Amicon-Ultra (Merck), frozen in liquid nitrogen, and stored at −80 °C.

His-Bgi1[FL] was expressed in SF9 cells. Cells were resuspended and lysed in binding buffer (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 5 mM imidazole). His-Bgi1[FL] was affinity purified with Ni-NTA agarose (GE Healthcare), eluted with 20 mM Tris-HCl, pH 8, 500 mM NaCl, and 500 mM imidazole. Purified protein was dialyzed against buffer containing Tris-HCl, pH 7.5, 1 mM DTT, and 100 mM NaCl. Samples were concentrated using Amicon-Ultra (Merck), frozen in liquid nitrogen, and stored at −80 °C. The absence of contaminating DNA was confirmed in all recombinant protein samples.

Viability assay. An overnight culture was inoculated into fresh YPD medium at 0.1 OD_{600} and grown to ~0.8 OD_{600}. The cell number was measured, followed by dilution of the cell suspension. Multiple dilutions containing 100–500 cells were subsequently plated on YPD solidified using 2% agar and grown for 2 days at 30 °C. The number of colonies formed was measured and plotted as normalized values to the wild-type control.

Serial dilution growth analysis. Cells were grown overnight, inoculated into fresh YPD at 0.2 OD_{600}, and grown until 0.8–1 OD_{600}. Following which cells were isolated and made up to 2 OD/ml in 1× PBS. Further dilutions were made as indicated in 1× PBS. In all, 2 µl of the cell suspension was transferred onto appropriate agar plates as mentioned and incubated for 2 days for 30 °C, DMF control and 2 µg/ml TBZ, 3 days for 30 °C + 4 µg/ml TBZ, and 37 °C and 7 days for 14 °C.

Electrophoretic mobility shift assays. Purified recombinant proteins of the mentioned molar ratio were incubated with 601 DNA (2.5 pM) or 1 pM of reconstituted nucleosomes in binding buffer (20 mM Tris, pH 7.5, 100 mM NaCl, 5% glycerol, and 1 mM DTT). They were incubated for 1 h at 4 °C and separated on a PAGE gel, stained with GelRed, and visualized using a UVP gel documentation system equipped with Doc-It LS Image acquisition software (UVP) v6.7.1. Further, the gels were stained with Coomassie to visualize the protein complexes and imaged using a scanner.

To generate chicken or human nucleosomes, histone H3.2 and histone H4 tetramer or CENP-A and histone H4 tetramer were mixed with histone H2A/H2B dimer with 601 DNA. KCl was added to the mixture to a final concentration of 2 M. The subsequent mixture was dialyzed to a salt gradient buffer (2 M to 200 mM KCl, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, and 10 mM DTT) overnight at 4 °C. The mixture was dialyzed to the final buffer (100 mM KCl, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, and 10 mM DTT) for 3 h at 4 °C. The pellet in the mixture was removed by centrifugation at 5000g at 4 °C. To stabilize the resultant nucleosome, the supernatant was kept in the incubator at 37 °C for 1 h. The reconstituted nucleosomes were checked by native-PAGE and detected by EtBr staining for DNA.

Estimation of DNA methylation. Genomic DNA was isolated from overnight cultures of wild type and bgi1Δ, using a modified glass bead protocol[76]. In brief, cells were suspended in a microfuge tube containing 500 µl of lysis buffer (50 mM Tris-HCl, pH 7.5, 20 mM EDTA, and 1% SDS) and 250 µl of glass beads. Cells were disrupted by vortexing for 5 min and centrifuged for 1 min at 15,000g. To the supernatant, 275 µl of 7 M ammonium acetate was added and incubated at 65 °C for 5 min and rapidly chilled on ice for 5 min. In all, 500 µl of chloroform was added, mixed, and centrifuged at 15,000g for 3 min. The supernatant containing DNA was precipitated with isopropanol, washed with 70% ethanol, dried, and resuspended in 50 µl of deionized water.

The isolated genomic DNA was digested separately with CpG methylation-sensitive (HhaI) or insensitive (HindIII) restriction enzymes (New England Biolabs) overnight with a no enzyme (uncut) control reaction. The digested DNA was diluted 1:50 and used for PCR amplification. Primer sets for PCR amplification of the centromeric region (5′-AGTCTCGTGTGGCTATGATT-3′ and 5′-GGATCT GCTTGACAGTGTCA-3′) and noncentromeric regions (5′-CCAACCGAAGCC CAAGACAA-3′ and 5′-TTGAAGGATGATCCGGCCGA-3′) were used. The obtained PCR products were subsequently separated by agarose gel electrophoresis using a 1% agarose gel and visualized by EtBr staining.

Statistics and reproducibility. Standard deviation and the mean of at least three independent experiments are mentioned in all plots unless mentioned otherwise. Statistical significance of differences was calculated as mentioned in the figure legends with one-way ANOVA with Dunnett's or Tukey's multiple comparison test, Mann–Whitney two-tailed, or Kruskal–Wallis one-way analysis followed by Dunn's multiple comparisons test. P values ≥ 0.05 were considered as non-significant (n.s.). Precise p values are mentioned within figures if significant and above $p = 0.0001$. All analyses were conducted using Microsoft Excel v2008 or GraphPad Prism version Windows v5.00 or v7.0a. Micrographs and immunoblotting images shown in figures are representative of at least two independent experiments with similar results.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

References

1. Musacchio, A. & Desai, A. A molecular view of kinetochore assembly and function. *Biology* **6**, 5 (2017).
2. McKinley, K. L. & Cheeseman, I. M. The molecular basis for centromere identity and function. *Nat. Rev. Mol. Cell Biol.* **17**, 16–29 (2015).
3. Fukagawa, T. & Earnshaw, W. C. The centromere: chromatin foundation for the kinetochore machinery. *Dev. Cell* **30**, 496–508 (2014).
4. Yadav, V., Sreekumar, L., Guin, K. & Sanyal, K. Five pillars of centromeric chromatin in fungal pathogens. *PLoS Pathog.* **14**, 1–7 (2018).
5. Suzuki, A. et al. How the kinetochore couples microtubule force and centromere stretch to move chromosomes. *Nat. Cell Biol.* **18**, 382–392 (2016).
6. Joglekar, A. P. & Kukreja, A. A. How kinetochore architecture shapes the mechanisms of its function. *Curr. Biol.* **27**, R816–R824 (2017).
7. Lara-Gonzalez, P., Westhorpe, F. G. & Taylor, S. S. The spindle assembly checkpoint. *Curr. Biol.* **22**, R966–R980 (2012).
8. Caldas, G. V. & DeLuca, J. G. Mad2 'opens' Cdc20 for BubR1 binding. *Mol. Cell* **51**, 3–4 (2013).
9. Earnshaw, W. C. & Rothfield, N. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma* **91**, 313–321 (1985).
10. Tachiwana, H. et al. Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature* **476**, 232–235 (2011).
11. Stoler, S., Keith, K. C., Curnick, K. E. & Fitzgerald-Hayes, M. A mutation in CSE4, an essential gene encoding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis. *Genes Dev.* **9**, 573–586 (1995).
12. Perpelescu, M. & Fukagawa, T. The ABCs of CENPs. *Chromosoma* **120**, 425–446 (2011).
13. Foltz, D. R. et al. The human CENP-A centromeric nucleosome-associated complex. *Nat. Cell Biol.* **8**, 458–469 (2006).
14. Hori, T. et al. CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. *Cell* **135**, 1039–1052 (2008).
15. Okada, M. et al. The CENP-H–I complex is required for the efficient incorporation of newly synthesized CENP-A into centromeres. *Nat. Cell Biol.* **8**, 446–457 (2006).
16. Yan, K. et al. Structure of the inner kinetochore CCAN complex assembled onto a centromeric nucleosome. *Nature* https://doi.org/10.1038/s41586-019-1609-1 (2019).
17. Varma, D. & Salmon, E. D. The KMN protein network—chief conductors of the kinetochore orchestra. *J. Cell Sci.* **125**, 5927–5936 (2012).
18. Joglekar, A. P., Bloom, K. & Salmon, E. D. In vivo protein architecture of the eukaryotic kinetochore with nanometer scale accuracy. *Curr. Biol.* **19**, 694–699 (2009).
19. Westermann, S. et al. Formation of a dynamic kinetochore-microtubule interface through assembly of the Dam1 ring complex. *Mol. Cell* **17**, 277–290 (2005).
20. Welburn, J. P. I. et al. The human kinetochore Ska1 complex facilitates microtubule depolymerization-coupled motility. *Dev. Cell* **16**, 374–385 (2009).

21. van Hooff, J. J. E., Snel, B. & Kops, G. J. P. L. Unique phylogenetic distributions of the Ska and Dam1 complexes support functional analogy and suggest multiple parallel displacements of Ska by Dam1. *Genome Biol. Evol.* **9**, 1295–1303 (2017).

22. Thakur, J. & Sanyal, K. The essentiality of the fungus-specific Dam1 complex is correlated with a one-kinetochore-one-microtubule interaction present throughout the cell cycle, independent of the nature of a centromere. *Eukaryot. Cell* **10**, 1295–1305 (2011).

23. Jenni, S. & Harrison, S. C. Structure of the DASH/Dam1 complex shows its role at the yeast kinetochore-microtubule interface. *Science* **360**, 552–558 (2018).

24. Kozubowski, L. et al. Ordered kinetochore assembly in the human-pathogenic basidiomycetous yeast *Cryptococcus neoformans*. *MBio* **4**, 1–8 (2013).

25. D'Archivio, S. & Wickstead, B. Trypanosome outer kinetochore proteins suggest conservation of chromosome segregation machinery across eukaryotes. *J. Cell Biol* **216**, 379–391 (2017).

26. Hooff, J. J., Tromer, E., Wijk, L. M., Snel, B. & Kops, G. J. Evolutionary dynamics of the kinetochore network in eukaryotes as revealed by comparative genomics. *EMBO Rep* **18**, 1559–1571 (2017).

27. Drinnenberg, I. A., deYoung, D., Henikoff, S. & Malik, H. S. Recurrent loss of CenH3 is associated with independent transitions to holocentricity in insects. *Elife* **3**, 1–19 (2014).

28. Schleiffer, A. et al. CENP-T proteins are conserved centromere receptors of the Ndc80 complex. *Nat. Cell Biol.* **14**, 604–613 (2012).

29. Navarro-Mendoza, M. I. et al. Early diverging fungus mucor circinelloides lacks centromeric histone CENP-A and displays a mosaic of point and regional centromeres. *Curr. Biol.* https://doi.org/10.1016/j.cub.2019.09.024 (2019).

30. Cheeseman, I. M., Chappie, J. S., Wilson-Kubalek, E. M. & Desai, A. The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. *Cell* **127**, 983–997 (2006).

31. DeLuca, J. G. et al. Kinetochore microtubule dynamics and attachment stability are regulated by Hec1. *Cell* **127**, 969–982 (2006).

32. Alushin, G. M. et al. The Ndc80 kinetochore complex forms oligomeric arrays along microtubules. *Nature* **467**, 805–810 (2010).

33. Hara, M. & Fukagawa, T. Where is the right path heading from the centromere to spindle microtubules? *Cell Cycle* **0**, 1–13 (2019).

34. Lang, J., Barber, A. & Biggins, S. An assay for de novo kinetochore assembly reveals a key role for the CENP-T pathway in budding yeast. *Elife*, **7**, 1–27 (2018).

35. Hara, M., Ariyoshi, M., Okumura, E., Hori, T. & Fukagawa, T. Multiple phosphorylations control recruitment of the KMN network onto kinetochores. *Nat. Cell Biol.* **20**, 1378–1388 (2018).

36. Nishino, T. et al. CENP-T provides a structural platform for outer kinetochore assembly. *EMBO J* **32**, 424–436 (2013).

37. Kim, S. & Yu, H. Multiple assembly mechanisms anchor the KMN spindle checkpoint platform at human mitotic kinetochores. *J. Cell Biol.* **208**, 181–196 (2015).

38. Milks, K. J., Moree, B. & Straight, A. F. Dissection of CENP-C-directed centromere and kinetochore assembly. *Mol. Biol. Cell* **20**, 4246–4255 (2009).

39. Dimitrova, Y. N., Jenni, S., Valverde, R., Khin, Y. & Harrison, S. C. Structure of the MIND complex defines a regulatory focus for yeast kinetochore assembly. *Cell* **167**, 1014–1027.e12 (2016).

40. Hornung, P. et al. A cooperative mechanism drives budding yeast kinetochore assembly downstream of CENP-A. *J. Cell Biol.* **206**, 509–524 (2014).

41. Kato, H. et al. A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science* **340**, 1110–1113 (2013).

42. Nishino, T. et al. CENP-T-W-S-X forms a unique centromeric chromatin structure with a histone-like fold. *Cell* **148**, 487–501 (2012).

43. Huisin'T Veld, P. J. et al. Molecular basis of outer kinetochore assembly on CENP-T. *Elife* **5**, 1–24 (2016).

44. Tromer, E. C., van Hooff, J. J. E., Kops, G. J. P. L. & Snel, B. Mosaic origin of the eukaryotic kinetochore. *Proc. Natl Acad. Sci. USA* **116**, 12873–12882 (2019).

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

46. Park, B. J. et al. Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS. *AIDS* **23**, 525–530 (2009).

47. Yadav, V. et al. RNAi is a critical determinant of centromere evolution in closely related fungi. *Proc. Natl. Acad. Sci. USA* **115**, 3108–3113 (2018).

48. Heitman, J. et al. *The Fungal Kingdom*. (American Society of Microbiology, 2017). https://doi.org/10.1128/9781555819583.

49. Jakopec, V., Topolski, B. & Fleig, U. Sos7, an essential component of the conserved *Schizosaccharomyces pombe* Ndc80-MIND-Spc7 complex, identifies a new family of fungal kinetochore proteins. *Mol. Cell. Biol.* **32**, 3308–3320 (2012).

50. Li, N. et al. Structure of the eukaryotic MCM complex at 3.8 Å. *Nature* **524**, 186–191 (2015).

51. Shahnejat-Bushehri, S. & Ehrenhofer-Murray, A. E. The ATAD2/ANCCA homolog Yta7 cooperates with Scm3 HJURP to deposit Cse4 CENP-A at the centromere in yeast. *Proc. Natl. Acad. Sci. USA* **117**, 5386–5393 (2020).

52. Thakur, J. & Sanyal, K. A coordinated interdependent protein circuitry stabilizes the kinetochore ensemble to protect CENP-A in the human pathogenic yeast *Candida albicans*. *PLoS Genet.* **8**, e1002661 (2012).

53. Varshney, N. et al. Spatio-temporal regulation of nuclear division by Aurora B kinase Ipl1 in *Cryptococcus neoformans*. *PLoS Genet.* **15**, e1007959 (2019).

54. Boehm, E. M., Gildenberg, M. S. & Washington, M. T. The many roles of PCNA in eukaryotic DNA replication. *Enzymes.* 231–254. https://doi.org/10.1016/bs.enz.2016.03.003 (2016).

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

56. MacCallum, D. E. & Hall, P. A. The biochemical characterization of the DNA binding activity of pKi67. *J. Pathol.* **191**, 286–298 (2000).

57. Cuylen, S. et al. Ki-67 acts as a biological surfactant to disperse mitotic chromosomes. *Nature* **535**, 308–312 (2016).

58. Sobecki, M. et al. The cell proliferation antigen Ki-67 organises heterochromatin. *Elife* **5**, 1–33 (2016).

59. Richter, M. M. et al. Network of protein interactions within the Drosophila inner kinetochore. *Open Biol.* **6**, 150238 (2016).

60. Liu, Y. et al. Insights from the reconstitution of the divergent outer kinetochore of *Drosophila melanogaster*. *Open Biol.* **6**, 150236 (2016).

61. Barth, T. K. et al. Identification of novel Drosophila centromere-associated proteins. *Proteomics* **14**, 2167–2178 (2014).

62. Drinnenberg, I. A., Henikoff, S. & Malik, H. S. Evolutionary turnover of kinetochore proteins: a ship of theseus? *Trends Cell Biol.* **26**, 498–510 (2016).

63. Petrovic, A. et al. Structure of the MIS12 complex and molecular basis of its interaction with CENP-C at human kinetochores. *Cell* **167**, 1028–1040.e15 (2016).

64. Bock, L. J. et al. Cnn1 inhibits the interactions between the KMN complexes of the yeast kinetochore. *Nat. Cell Biol.* **14**, 614–624 (2012).

65. Venkei, Z. et al. Spatiotemporal dynamics of Spc105 regulates the assembly of the Drosophila kinetochore. *Open Biol* **2**, 110032–110032 (2012).

66. Booth, D. G. & Earnshaw, W. C. Ki-67 and the chromosome periphery compartment in mitosis. *Trends Cell Biol.* **27**, 906–916 (2017).

67. Akiyoshi, B. & Gull, K. Discovery of unconventional kinetochores in kinetoplastids. *Cell* **156**, 1247–1258 (2014).

68. Potter, S. C. et al. HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–W204 (2018).

69. Zimmermann, L. et al. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* **430**, 2237–2243 (2018).

70. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).

71. Mészáros, B., Erdős, G. & Dosztányi, Z. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.* **46**, W329–W337 (2018).

72. Gasteiger, E. et al. *The Proteomics Protocols Handbook. The Proteomics Protocols Handbook Protein Identification and Analysis Tools on the ExPASy Server.* (Humana Press, 2005). https://doi.org/10.1385/1592598900.

73. Grigoriev, I. V. et al. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res.* **42**, 699–704 (2014).

74. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* **47**, W256–W259 (2019).

75. Wang, Q. M. et al. Multigene phylogeny and taxonomic revision of yeasts and related fungi in the Ustilaginomycotina. *Stud. Mycol.* **81**, 55–83 (2015).

76. Ruff, J. A., Lodge, J. K. & Baker, L. G. Three galactose inducible promoters for use in *C. neoformans* var. grubii. *Fungal Genet. Biol.* **46**, 9–16 (2009).

77. Davidson, R. C. et al. Gene disruption by biolistic transformation in serotype D strains of *Cryptococcus neoformans*. *Fungal Genet. Biol.* **29**, 38–48 (2000).

78. del Valle, J., de la Oliva, N., Muller, M., Stieglitz, T. & Navarro, X. Biocompatibility evaluation of parylene C and polyimide as substrates for peripheral nerve interfaces. in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)* 442–445 (IEEE, 2015). https://doi.org/10.1109/NER.2015.7146654.

79. Abràmoff, M. D., Magalhães, P. J. & Ram, S. J. Image processing with ImageJ Part II. *Biophotonics Int.* **11**, 36–43 (2005).

80. Oya, E. et al. H3K14 ubiquitylation promotes H3K9 methylation for heterochromatin assembly. *EMBO Rep.* **e48111**, 1–15 (2019).

81. Ishihama, Y. et al. Exponentially Modified Protein Abundance Index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell. Proteomics* **4**, 1265–1272 (2005).

82. Sanyal, K. & Baum, M. & Carbon, J. Centromeric DNA sequences in the pathogenic yeast *Candida albicans* are all different and unique. *Proc. Natl. Acad. Sci. USA* **101**, 11374–11379 (2004).

83. Dubin, M., Fuchs, J., Gräf, R., Schubert, I. & Nellen, W. Dynamics of a novel centromeric histone variant CenH3 reveals the evolutionary ancestral timing of centromere biogenesis. *Nucleic Acids Res.* **38**, 7526–7537 (2010).

Acknowledgements

Author contributions

S.S. and K.S. conceived the project. S.S., K.S., T.H., and T.F. designed the experiments. R. N. performed the mass spectrometry and S.S. performed the rest of the experiments. T.H. assisted with affinity purification for mass spectrometry. S.S. and K.S. wrote the paper. S.S., K.S., T.F., and T.H. revised the paper. All authors approved the final version. K.S. gathered funding and supervised the entire project.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-20161-9.

Correspondence and requests for materials should be addressed to T.F. or K.S.

Peer review information *Nature Communications* thanks Bungo Akiyoshi, Geert Kops, Kozo Tanaka and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

RNAi is a critical determinant of centromere evolution in closely related fungi

Vikas Yadav[a], Sheng Sun[b], R. Blake Billmyre[b], Bhagya C. Thimmappa[a], Terrance Shea[c], Robert Lintner[c], Guus Bakkeren[d], Christina A. Cuomo[c], Joseph Heitman[b], and Kaustuv Sanyal[a,1]

[a]Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560064, India; [b]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710; [c]Infectious Disease and Microbiome Program, Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, MA 02142; and [d]Agriculture and Agri-Food Canada, Summerland Research and Development Centre, Summerland, BC V0H1Z0, Canada

The centromere DNA locus on a eukaryotic chromosome facilitates faithful chromosome segregation. Despite performing such a conserved function, centromere DNA sequence as well as the organization of sequence elements is rapidly evolving in all forms of eukaryotes. The driving force that facilitates centromere evolution remains an enigma. Here, we studied the evolution of centromeres in closely related species in the fungal phylum of Basidiomycota. Using ChIP-seq analysis of conserved inner kinetochore proteins, we identified centromeres in three closely related *Cryptococcus* species: two of which are RNAi-proficient, while the other lost functional RNAi. We find that the centromeres in the RNAi-deficient species are significantly shorter than those of the two RNAi-proficient species. While centromeres are LTR retrotransposon-rich in all cases, the RNAi-deficient species lost all full-length retroelements from its centromeres. In addition, centromeres in RNAi-proficient species are associated with a significantly higher level of cytosine DNA modifications compared with those of RNAi-deficient species. Furthermore, when an RNAi-proficient *Cryptococcus* species and its RNAi-deficient mutants were passaged under similar conditions, the centromere length was found to be occasionally shortened in RNAi mutants. In silico analysis of predicted centromeres in a group of closely related *Ustilago* species, also belonging to the Basidiomycota, were found to have undergone a similar transition in the centromere length in an RNAi-dependent fashion. Based on the correlation found in two independent basidiomycetous species complexes, we present evidence suggesting that the loss of RNAi and cytosine DNA methylation triggered transposon attrition, which resulted in shortening of centromere length during evolution.

retrotransposons | *Cryptococcus* | *Ustilago* | experimental evolution | DNA methylation

The centromere is a specialized DNA locus that is required for assembly of a multiprotein complex, the kinetochore, which drives faithful chromosome segregation in eukaryotes. Centromeres can be classified as point centromeres of short DNA sequences of <400 bp (e.g., *Saccharomyces cerevisiae*), and regional centromeres which are long, and can range between a few kilobases (kb) (e.g., *Schizosaccharomyces pombe*) to several megabases (Mb) (e.g., humans and plants) (1, 2). The repetitive DNA present in regional centromeres (core and pericentromeres) consists of either arrays of satellite DNA or transposons or both. Transposons, despite being present in lower copy numbers than the satellites, are proposed to play a major role in regional centromere evolution (1, 3, 4). Transposon domestication at the centromere can also give rise to functional domains or repeats, including satellite DNA repeats in a centromere. The *dh/dg* repeats in the fission yeast or α-satellite repeats in humans are proposed to be the result of such domestication events (4). The presence of CENP-B, a centromere DNA binding protein that shows a high level of similarity with DNA transposons in humans and its homologs in other organisms, provides another line of evidence toward the role of transposons in the evolution of centromere

structure and function (5, 6). Notably, a large number of transposons present in centromeres are RNA transposons or retrotransposons, which propagate through RNA intermediates in the "copy–paste" mode. These elements are different from DNA transposons, which excise themselves from the original site and move to a new target site in the genome, and thus propagate in the "cut–paste" mode (7). Retrotransposons, owing to their mode of propagation, have been proposed as architects of the regional centromeres (4). These elements can produce multiple copies of themselves and integrate into the same centromere locus or centromeres of other chromosomes (3, 4, 8, 9).

A retroelement must be transcribed to generate its RNA intermediates required for transposition. A low level of transcription from the centromere in a range of organisms, including budding yeast, fission yeast, mouse, humans, and maize, has been reported (10, 11). Studies on centromere function in various organisms suggest that neither a complete absence of transcription nor a high level of transcription at the centromere supports assembly of the multisubunit kinetochore complex on the centromere DNA (10, 12). Long noncoding centromeric transcript RNAs are shown to be required for loading of critical kinetochore proteins such as CENP-C (13, 14). In fission yeast, transcripts from *dh/dg* repeat regions of

Significance

The "centromere paradox" refers to rapidly evolving and highly diverse centromere DNA sequences even in closely related eukaryotes. However, factors contributing to this rapid divergence are largely unknown. Here, we identified large regional, LTR retrotransposon-rich centromeres in a group of human fungal pathogens belonging to the *Cryptococcus* species complex. We provide evidence that loss-of-functional RNAi machinery and possibly cytosine DNA methylation trigger instability of the genome by activation of centromeric retrotransposons presumably suppressed by RNAi. We propose that RNAi, together with cytosine DNA methylation, serves as a critical determinant that maintains repetitive transposon-rich centromere structures. This study explores the direct link between RNAi and centromere structure evolution.

the centromere are targets of the RNA interference (RNAi) machinery and are required for heterochromatin formation (15, 16). Lack of proper heterochromatinization at *dh/dg* repeats causes abnormal centromere function leading to chromosome missegregation, although RNAi does not directly affect CENP-A loading in this organism (16). These results highlight the importance of transcription of centromeres. However, the level of transcription at the centromere must be regulated at an optimal level for proper kinetochore assembly and genome stability.

By regulating transcription from retrotransposons, present at the centromere or elsewhere in the genome, RNAi contributes to genome defense (17, 18). The functional RNAi pathway involves key proteins, including Dicer (Dcr), Argonaute (Ago), and RNA-dependent RNA polymerase (Rdp). The RNAi machinery targets double-stranded RNA being generated from retroelements/repeats leading to its degradation and in some cases promoting heterochromatinization through repressive histone modifications (such as H3K9 dimethylation) and/or DNA methylation (15, 19). Thus, by controlling expression of transposons at centromeres, RNAi could play an important regulatory role in the structural evolution of centromeres. However, except in *S. pombe*, no specific study has been conducted to explore a possible link between RNAi and centromere structure–function evolution (9).

The presence/absence of retrotransposons or repetitive DNA sequences can also contribute to rapid divergence of centromeres in a species complex. For example, centromeres in three related species of the *Schizosaccharomyces* clade, *S. pombe*, *Schizosaccharomyces japonicus*, and *Schizosaccharomyces octosporus*, exhibit divergence in the centromeric architecture (9). *S. japonicus* centromeres contain mostly transposons, whereas *S. pombe* and *S. octosporus* contain repeat-rich centromeres, and have lost most of the active transposons (9). Another group of fungi, the *Candida* species complex in the CTG clade, have been studied extensively with respect to centromere structure and function. *Candida albicans*, *Candida dubliniensis*, and *Candida lusitaniae* harbor repeatless centromeres where centromere DNA sequences are unique and different in each species (20–22). On the other hand, *Candida tropicalis* contains highly homogenized repeat-associated fission yeast-like centromeres (23). Both of these genera, *Schizosaccharomyces* and *Candida*, belong to the phylum Ascomycota of the fungal kingdom.

In this study, we identified centromeres in three closely related *Cryptococcus* species in the fungal phylum Basidiomycota. The three species, *Cryptococcus neoformans* (type strain H99), *Cryptococcus deneoformans* (reference isolate JEC21), and *Cryptococcus deuterogattii* (outbreak isolate R265), are haploid and diverged from a common ancestor as recently as 34 Mya (24–26). All three species are human pathogens and two of these primarily infect immunocompromised individuals, including HIV/AIDS patients, to cause cryptococcal meningitis, a leading cause of death in these patients. While *C. neoformans* and *C. deneoformans* are the most studied species in this complex, a recent rise in cases of drug resistance and infections in otherwise healthy individuals have attracted attention to the lesser explored species *C. deuterogattii* (27, 28). Unlike *C. neoformans* and *C. deneoformans*, *C. deuterogattii* lost key components of the functional RNAi machinery (29). Here, we show that centromeres in all three *Cryptococcus* species are regional centromeres featuring retroelements or their remnants. Notably, the RNAi-deficient *C. deuterogattii* species harbors significantly shorter centromeres compared with the RNAi-proficient species *C. neoformans* or *C. deneoformans*. We also predicted centromere locations in the *Ustilago* species complex, a group of plant pathogens in the Basidiomycota, and observed a similar correlation between the loss of RNAi and shortening of putative centromeres. Overall, this study provides mechanistic insights into the structural evolution of centromeres in an RNAi-dependent manner.

Results and Discussion

***Cryptococcus* Species Harbor Large Regional Centromeres.** We previously reported the probable structure of centromeres in

C. neoformans as long ORF-free regions that are rich in retrotransposable elements (25). In addition, centromeres in *C. deneoformans* were predicted as a single transposon-rich locus on each of the 14 chromosomes (26). In this study, we sought to identify the centromeres experimentally in three pathogenic species: *C. neoformans*, *C. deneoformans*, and *C. deuterogattii*, of the *Cryptococcus* species complex, including verifying centromere locations predicted earlier in *C. neoformans* and *C. deneoformans*.

To achieve this goal, two of the evolutionarily conserved inner kinetochore proteins, CENP-A and CENP-C, were identified in each of the three species (*SI Appendix*, Fig. S1 *A* and *B*). Both of these proteins were tagged with mCherry and showed localization patterns as reported previously (Fig. 1*A* and *SI Appendix*, Fig. S1*C*) (30). To identify functional centromeres in *C. neoformans*, we performed CENP-A and CENP-C chromatin immunoprecipitation (ChIP) followed by deep sequencing (ChIP-seq). ChIP-seq analysis revealed overlapping binding of both proteins at a single locus on each of the 14 chromosomes of *C. neoformans* (*SI Appendix*, Fig. S2 and Table S1). The binding regions of both proteins were largely overlapping and these regions spanned across the ORF-free, poorly transcribed regions on most chromosomes (Fig. 1*B* and *SI Appendix*, Fig. S3). The minor variations observed between the binding sites of these two proteins could be due to different sequencing approaches employed. While no specific patterns could be detected in CENP-A/CENP-C binding across these regions, occasional dips in the binding pattern of both proteins were observed (*SI Appendix, SI Materials and Methods*). The length of the CENP-A– and CENP-C–bound regions varied from 20 kb to 40 kb in *C. neoformans* unlike centromeres of many other fungi where the CENP-A–bound region remains nearly constant across chromosomes (2). A similar CENP-A binding pattern across the entire stretch of the centromere was found in *Neurospora crassa* (31). It is important to note here that centromeres in *C. neoformans* were not completely assembled, as they contain a few sequence gaps in the current assembly. We attempted to close these gaps by PacBio as well as Sanger sequencing (*SI Appendix, SI Materials and Methods*). This resulted in a significant improvement of the genome assembly as we could close the sequence gaps in 11 of 14 centromeres, leaving only *CEN3*, *CEN11*, and *CEN14* with some sequence gaps. While this is the best assembly obtained thus far for the *C. neoformans* genome, the presence of additional breaks/gaps cannot be excluded.

The *C. neoformans* genome shares a high level of gene synteny with that of *C. deneoformans* (*SI Appendix*, Fig. S4*A* and Table S2) (25, 32). By performing synteny analysis across centromere flanking regions of *C. neoformans*, we were able to predict putative centromeres in *C. deneoformans*. The predicted regions were large, ORF-free, poorly transcribed, and map to loci previously predicted as centromeres (*SI Appendix*, Figs. S3 and S4*B* and Table S1) (26). CENP-C ChIP-qPCR confirmed the authenticity of each of these regions as functional centromeres in all 14 chromosomes of *C. deneoformans* (Fig. 1*C*). Compared with the noncentromeric locus, we obtained significant enrichment of centromeric CENP-C binding using two pairs of primers located distantly on each of the 14 centromeres.

We performed PacBio and Nanopore sequencing for the *C. deuterogattii* genome and generated a complete, chromosome-level de novo genome assembly, improving on the previous assembly of 26 scaffolds (Fig. 2*A* and *SI Appendix, SI Materials and Methods*). A comparison of the *C. deuterogattii* genome with that of *C. neoformans* revealed a number of rearrangements between the two species (*SI Appendix*, Fig. S4*C* and Table S2). It was previously predicted that the *C. neoformans* and *C. deuterogattii* genomes have undergone an arm exchange involving chromosomes 1 and 2 (25). The new long-read assembly provides compelling evidence supporting this rearrangement. Next, we performed CENP-C ChIP-seq in *C. deuterogattii* and analyzed the data using the chromosomal assembly and obtained 14 binding peaks, one on each chromosome (Fig. 2*B* and *SI Appendix*, Fig. S4*D* and Table S1). Similar to *C. neoformans* and *C. deneoformans*, all

GENETICS
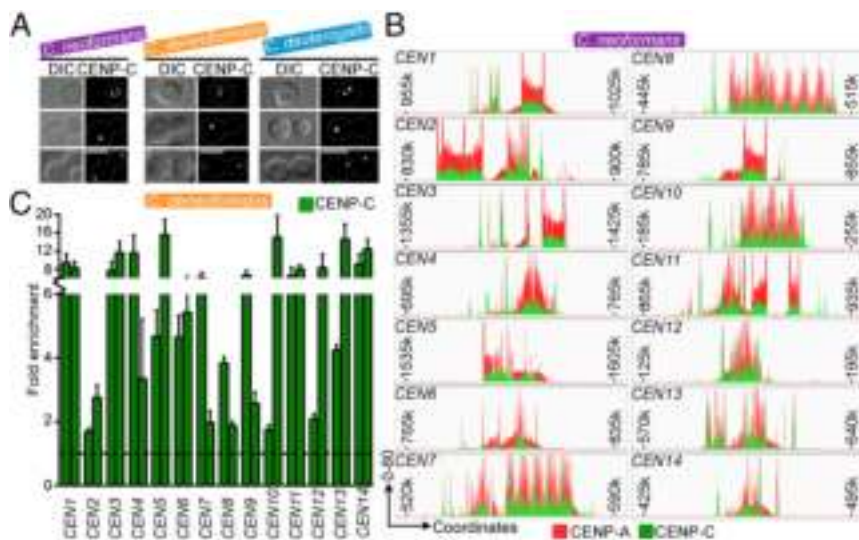
Fig. 1. The *Cryptococcus* species complex has large regional-type centromeres. (*A*) The subcellular localization patterns of a conserved kinetochore protein CENP-C at various cell cycle stages (interphase, prometaphase, and anaphase) in *C. neoformans*, *C. deneoformans*, and *C. deuterogattii*. (Scale bar, 5 μm.) (*B*) Overlapping binding of CENP-A (red) and CENP-C (green) on each of the 14 chromosomes of *C. neoformans* as revealed by ChIP-seq analysis of these proteins. An 80-kb DNA sequence harboring the centromere is shown for each of the 14 chromosomes. (*C*) CENP-C (mCherry)-ChIP-qPCR analysis confirmed enrichment of CENP-C on the predicted centromeres in *C. deneoformans*. Fold enrichment was normalized to a non-*CEN* region, the level of which is marked by the dotted line in the graph. Error bars represent SEM.

14 peaks mapped to ORF-free poorly transcribed regions in the *C. deuterogattii* genome (*SI Appendix*, Fig. S3).

Sequence analysis of the CENP-A/CENP-C–bound ORF-free centromere regions in all three species revealed the presence of retrotransposons Tcn1–Tcn6 (Fig. 3*A*) (33). While Tcn1–Tcn5 belong to the Ty3-*gypsy* family of retroelements, Tcn6 belongs to the Ty1-*copia* family of retroelements. The retroelements have a high degree of conservation (70% or more identity) across the three species (Dataset S1). However, they differ in their sequence from the centromeric retroelements of a closely related nonpathogenic *Cryptococcus* species, *Cryptococcus amylolentus* (*SI Appendix*, Fig. S5*A* and Dataset S1) (34). It is important to note that Tcn1–Tcn6 elements are not found in the *C. amylolentus* genome, indicating that these elements are specific to the pathogenic *Cryptococcus* species. Based on these results, we conclude that centromeres in pathogenic *Cryptococcus* species are large regional-type, and enriched in the same retroelements. These results suggest that neither the location nor the sequence elements of the centromeres have diverged significantly among these closely related fungal species. However, further analysis revealed that the centromeres in *C. deuterogattii* are significantly shorter, with an average length of around 14 kb compared with those of *C. neoformans* or *C. deneoformans*, which have an average length of 44 kb and 62 kb, respectively (Fig. 3*B*). Consequently, the CENP-C–bound regions were also found to be shortened in *C. deuterogattii* (5–15 kb) compared with *C. neoformans* (20–40 kb). Thus, we conclude that the *Cryptococcus* species have CENP-A/CENP-C–rich regional centromeres of varying lengths, but the centromere DNA sequence elements are highly conserved.

Loss of the RNAi Machinery and Shortening of Centromeres Are Correlated in *C. deuterogattii*. Ago, Dcr, and Rdp—the key proteins of the RNAi machinery—are all present in *C. neoformans* and *C. deneoformans* but all are absent in *C. deuterogattii* (29). We previously demonstrated that RNAi suppresses transcription and transposition of retrotransposons, including the Tcn elements in the RNAi-proficient species *C. neoformans* and *C. deneoformans* (35, 36). Real-time PCR analysis of two of the Tcn elements confirmed that the expression levels of these elements are higher in *C. deuterogattii* compared with *C. neoformans* (*SI Appendix*, Fig. S5*B*). To further investigate alterations in the centromere length observed, we performed a detailed analysis of the retrotransposons, Tcn1–Tcn6, in all three species. These retroelements are specifically enriched and mostly restricted (>95%) to centromeric regions in all three species (Fig. 3*C* and *SI Appendix*, Figs. S2 and S4). The most striking observation is that the centromeres in *C. neoformans* and *C. deneoformans* harbor a significant proportion of full-length retroelements (20–30%), whereas the *C. deuterogattii* centromeres are completely devoid of such full-length elements (Fig. 3*D*). Instead, retroelements present in *C. deuterogattii* centromeres are only remnants of transposons (Fig. 3*E*). They lack one or more of the essential domains (LTRs, reverse transcriptase, integrase) required for transposition activity, rendering them nonfunctional for further transposition. Thus, this analysis reveals that loss of RNAi in *C. deuterogattii* is correlated with the loss of full-length retrotransposons and an overall reduction in the length of the centromeres.

By comparison with *C. neoformans*, the genome of *C. deuterogattii* seems to have lost additional genes besides key enzymes in the RNAi pathway (24, 29, 37). These additional lost genes include genes involved in protein processing, protein degradation, and the mitochondrial oxidation pathway. In addition, we found
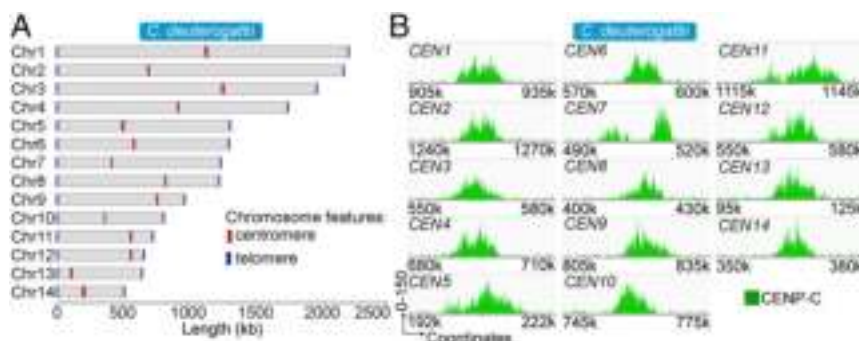
Fig. 2. ChIP sequencing and PacBio sequencing map centromeres on each of the 14 chromosomes in *C. deuterogattii*. (*A*) Map showing the 14 chromosomes in *C. deuterogattii* with the telomeres and centromeres marked. The centromeres are marked to scale, while the telomeres are marked as 10-kb regions for visualization purposes. (*B*) CENP-C (mCherry)-ChIP-seq identified 14 binding regions among 14 chromosomes in *C. deuterogattii*'s latest assembly. A 30-kb region spanning the CENP-C–bound region is shown for each chromosome.
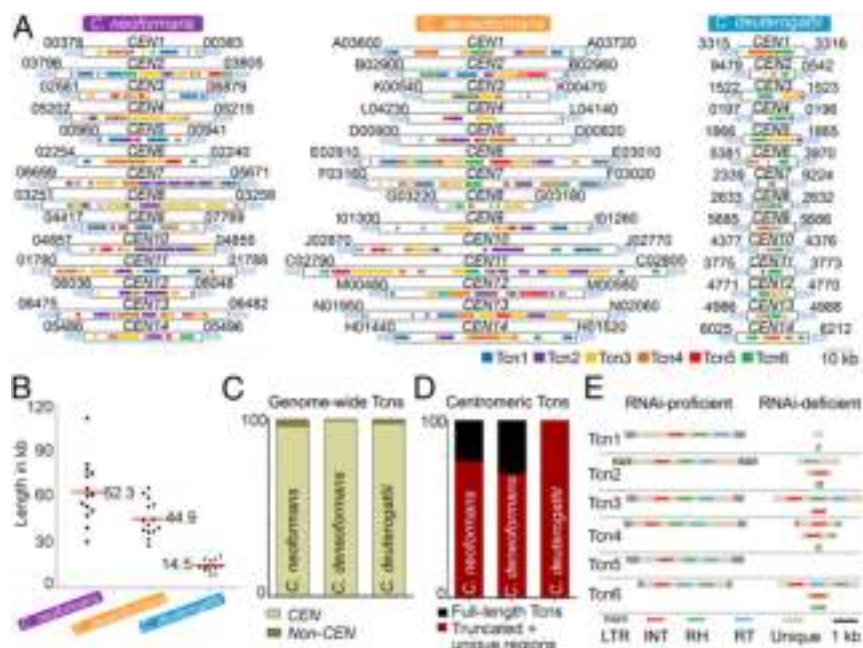
Fig. 3. Centromeres are enriched with retrotransposons in the *Cryptococcus* species complex. (*A*) The presence of various retrotransposons across the centromeres in *C. neoformans*, *C. deneoformans*, and *C. deuterogattii*. The numbers refer to centromere-flanking ORFs which are preceded by "CNAG_" for *C. neoformans*, "CN" for *C. deneoformans*, and "CNBG_" for *C. deuterogattii*. The diagrams are drawn to scale. (*B*) The length of each centromere of the respective species was plotted. Each dot represents one centromere, and the horizontal red line depicts the mean centromere length of the corresponding species. (*C*) Distribution of retrotransposons, Tcn1–Tcn6, in centromeres and across the genome in *C. neoformans*, *C. deneoformans*, and *C. deuterogattii*. (*D*) A bar diagram showing the distribution of full-length versus truncated Tcn elements at the centromeres in all three species. (*E*) Comparison of retrotransposon elements present in centromeres of RNAi-proficient (*C. neoformans* and *C. deneoformans*) and RNAi-deficient (*C. deuterogattii*) species. INT, integrase; LTR, long terminal repeat; RH, RNaseH; RT, reverse transcriptase; Unique, unique DNA sequence in each retroelement.

that the sole DNA methyltransferase encoding gene, *DNMT5*, is truncated in *C. deuterogattii* (*SI Appendix*, Fig. S6*A*). A recent study reported that putative centromeres of *C. neoformans* are methylated at the DNA level (38). Combining the available bisulfite sequencing results with our ChIP-seq results revealed that the centromere DNA sequence of *C. neoformans* is indeed extensively methylated (*SI Appendix*, Fig. S2). Incidentally, due to the presence of a truncated form of *DNMT5*, cytosine DNA methylation at centromeres is found to be absent in *C. deuterogattii* (*SI Appendix*, Fig. S6 *B* and *C*). The lack of centromere DNA methylation in *C. deuterogattii* was further supported by base modification analysis based on PacBio single-molecule real-time (SMRT) sequencing (Fig. 4*A* and *SI Appendix*, Fig. S6*D*). While *C. neoformans* centromeres harbor extensive base modifications, little or no methylation is associated with *C. deuterogattii* centromeres. Previous reports suggest that the DNA methylation suppresses recombination as well as regulates retrotransposition in various organisms (39, 40).

To strengthen support for our hypothesis on the roles of RNAi in centromere evolution, we also identified centromeres in another basidiomycete species complex, the *Ustilago* species complex. Similar to the *Cryptococcus* species complex, the *Ustilago* species complex harbors three species of which *Ustilago maydis* lost all three major proteins of the RNAi machinery along with the DNA methyltransferase while *Ustilago hordei* and *Ustilago bromivora* harbor the complete machinery (41, 42). In *Cryptococcus*, centromeres are present in large ORF-free and poorly transcribed regions of the genome. The centromeres in three *Cryptococcus* species are also syntenic with each other. Keeping in mind these centromeric features, we predicted centromeres in all three species of the *Ustilago* complex by in silico analysis and found that the RNAi-deficient species *U. maydis* possesses shorter centromeres (average length 12 kb) compared with both of the RNAi-proficient species *U. hordei* (average length 36 kb) and *U. bromivora* (average length 27 kb) (Fig. 4*B* and *SI*

Appendix, Fig. S7 and Table S3). The centromere locations identified in *U. maydis* matched well with the ones predicted previously (43). Earlier studies have reported that *U. maydis* harbors fewer transposons than *U. hordei* and *U. bromivora* (41, 42). These observations further confirmed the correlation between loss of RNAi and a reduction in centromere length. We further extended our analysis to fungal species harboring regional centromeres, and in which RNAi status is known, and found that species with all components of the RNAi machinery harbor longer centromeres than the species that lack one or more of the RNAi components (Fig. 4*C* and *SI Appendix*, Table S4).

Rather than a single optimized DNA sequence, centromeres are among the most rapidly changing loci in the genome despite having a conserved and essential function, a phenomenon termed the "centromere paradox" (44). A number of studies verified that centromeres are evolving rapidly, even among very closely related species of plants and animals (3, 8, 45). This process of rapid evolution of centromeres is well studied in fungal species, especially among the members of the Ascomycota (9, 23, 46). The nature of centromeres in another significant fungal phylum, the Basidiomycota, was unknown. In this report, we identified centromeres in three pathogenic species of the *Cryptococcus* species complex. Using multiple sequencing and analysis methods, we significantly improved the genome assembly for *C. neoformans* and obtained a full genome-wide, chromosome-level assembly for *C. deuterogattii*. The three *Cryptococcus* species also provided an opportunity to test the role of RNAi in centromere evolution. We discovered that centromeres in the RNAi-deficient species are shorter compared with the RNAi-proficient species. Analysis of the centromere DNA sequence revealed that the RNAi-deficient species possesses fewer and truncated retrotransposons compared with the RNAi-proficient species. Taken together, our study reveals that centromeres evolve rapidly among closely related species in the Basidiomycota phylum of fungi.

GENETICS
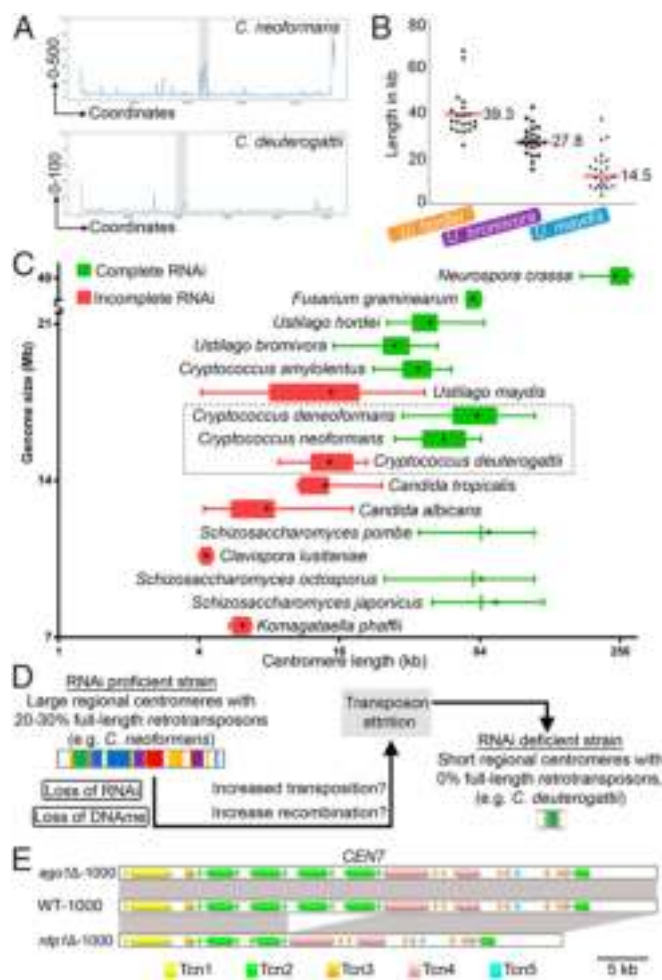
Fig. 4. RNAi as a key determinant of longer centromeres in closely related fungi. (*A*) Base modification analysis based on SMRT PacBio sequencing revealed a high level of methylation in *C. neoformans CEN14* (depicted as gray-shaded region) but not in *C. deuterogattii CEN14* (see *SI Appendix*, Fig. S6D for remaining centromeres). (*B*) Comparison of the predicted centromere length in *U. maydis*, *U. bromivora*, and *U. hordei*. Only 20 predicted centromeres are plotted for *U. hordei*, while all 23 are shown for both *U. maydis* and *U. bromivora.* Each dot represents one centromere, and the horizontal line depicts the mean value. (*C*) Graph showing the correlation between centromere (ORF-free region) lengths and status of RNAi among the fungal species. The star in each of the boxes represents the mean value, the boxes depict the range from the 25th percentile to 75th percentile values, and the terminal vertical lines mark the range of centromere lengths. Species lacking any one of the three key proteins (Ago, Dcr, or Rdp) of the RNAi machinery were considered to harbor incomplete RNAi machinery for this analysis (*SI Appendix*, Table S4). (*D*) Possible sequence of events that might have occurred due to loss of RNAi machinery and/or DNA methylation in an RNAi-proficient strain (*C. neoformans* or *C. deneoformans*) that led to a genome with truncated retrotransposons that are unable to transpose in an RNAi-deficient strain (*C. deuterogattii*). (*E*) PacBio sequencing results of strains passaged for 1,000 doublings revealed reduction in *CEN7* length in the *rdp1Δ* mutant but not in *C. neoformans* wild-type or the *ago1Δ* strains.

The majority of the siRNA in *C. neoformans* maps to centromeric retrotransposons, the level of which drops in RNAi-defective mutants of *C. neoformans* (36, 47). The key proteins required for a functional RNAi machinery were found to be absent in *C. deuterogattii* (29). We propose that loss of RNAi in this species might have led to amplification of retroelements, which in turn would have integrated into the centromere-proximal sites, leading to a transient elongation of the centromeres.

Consequently, the presence of retroelements in close vicinity to each other might have enhanced the rate of recombination between these elements. Recombination would cause shortening of these regions and might render retrotransposons inactive. Further, the absence of RNAi along with cytosine DNA methylation can contribute to enhancing the rate of recombination between these elements (39, 40, 48). It is notable that both transposition and recombination can be damaging to the genome if they are associated with loss of essential genes. Thus, cells that have no essential genes inactivated after transposition and recombination would survive. Because the partial loss of a centromere may not affect its function (49, 50), the probability of such events being tolerated at the centromere locus may be higher. Thus, RNAi-deficient strains with a stabilized genome will be likely to have shorter centromeres with truncated retroelements, similar to those of *C. deuterogattii* or *U. maydis*, while the intermediate population during evolution might have possessed centromeres of varying length (Fig. 4*D*).

We attempted to test our hypothesis by performing experimental evolution experiments. *C. neoformans* and its derived RNAi mutants were grown for 1,000 doublings under standard laboratory conditions (*SI Appendix*, Fig. S8 and *SI Materials and Methods*). The passaged strains did not show any obvious growth defects as measured by their generation time (*SI Appendix*, Table S5). PacBio sequencing of wild-type and mutant strains (both passaged and unpassaged) revealed two rearrangements in the centromeric regions of RNAi mutants compared with the wild-type grown under similar conditions (*SI Appendix*, Table S6). *CEN7* of *rdp1Δ* mutant exhibited a reduction in length in the 1,000 doubling passaged strains compared with the 1,000 doubling wild-type and unpassaged strains (Fig. 4*E*). In addition, both unpassaged and passaged RNAi mutants carried a shorter *CEN2* compared with the wild type (*SI Appendix*, Fig. S8 and Table S6). These results suggest that centromeres are prone to structural alterations in the absence of RNAi. Overall, our study provides evidence of RNAi in maintaining the structure of retrotransposon-rich centromeres in fungi. Experimental evolution in RNAi as well as cytosine DNA methylation mutants under conditions that favor retrotransposon expression, such as during the sexual cycle, will be tested in the future to gain further insight into this process.

Transposons play a major role in shaping the evolution of genomes, including the centromere, in multiple ways (18, 51). It has been proposed that the centromeric repeats present in fission yeast, maize, and the α-satellite repeats in human centromeres evolved from transposable elements (4). Transposons also have been shown to play an active role in centromere evolution among closely related species in plants (3). Based on a study in the *Schizosaccharomyces* group, it was proposed that loss of transposons observed in *S. pombe* centromeres occurred as a result of recombination between LTR elements that remained present in the *S. japonicus* genome (9). The loss of retrotransposons was also correlated with a shift in transposon regulation from RNAi to the CENP-B homolog in *S. pombe*, Cbp1 (5). In this study, we show structural changes in centromeres mediated by retrotransposons. Centromeres in the RNAi-proficient species *C. neoformans* and *C. deneoformans* harbor full-length retrotransposable elements, whereas RNAi-deficient *C. deuterogattii* centromeres contain only footprints of these elements. We propose that the truncation of retrotransposons could have occurred due to recombination among retroelements rendering them inactive. One such example is reported in angiosperms where retrotransposon truncation via illegitimate recombination resulted in genome size reduction (52, 53). Loss of full-length retroelements can prove to be advantageous for *C. deuterogattii* as a pathogen. Host conditions might induce retrotransposition and damage a pathogen's genome despite suppression of active transposition by RNAi in an RNAi-proficient species (54, 55). While loss of RNAi in such situations can be lethal due to unregulated transposition, truncation and inactivation of full-length retroelements can provide selective advantages. The *C. deuterogattii* genome is 1.3 Mb smaller compared with *C. neoformans* and centromere shortening accounts for approximately one-third of this reduction (total centromere lengths being >630 kb in

Yadav et al.

C. neoformans versus 203 kb in *C. deuterogattii*). A relatively faster generation time (*SI Appendix*, Table S5) of *C. deuterogattii* compared with *C. neoformans* or *C. deneoformans* may be due to the smaller genome of the former. Thus, we speculate that loss of full-length retroelements triggered by the loss of RNAi in *C. deuterogattii* could provide a pathogenic advantage over its related RNAi-proficient species in addition to other factors (27). There is as yet no established direct correlation between pathogenesis and the loss of RNAi along with full-length retrotransposon. The *Cryptococcus* species complex might prove to be a good model system with which to address such questions.

Materials and Methods

The strains and primers used in this study are listed in *SI Appendix*, Tables S7 and S8, respectively. *Cryptococcus* was grown in YPD (1% yeast extract, 2% peptone, and 2% dextrose) media at 30 °C. *Cryptococcus* cells were transformed using biolistics. Transformants were selected on YPD agar media containing 200 μg/mL of G418 (Sigma-Aldrich) or hygromycin

(Invitrogen). Details of all of the experimental procedures and sequence analysis are given in *SI Appendix, SI Materials and Methods*. All of the sequencing data (including ChIP-seq and PacBio sequencing) have been deposited under NCBI BioProject accession no. PRJNA395628 and the Nanopore data under PRJNA13691. The reference number for each study is provided in *SI Appendix*, Table S9.

1. Brown JD, O'Neill RJ (2014) *The Evolution of Centromeric DNA Sequences. eLS* (John Wiley & Sons, Hoboken, NJ).
2. Roy B, Sanyal K (2011) Diversity in requirement of genetic and epigenetic factors for centromere function in fungi. *Eukaryot Cell* 10:1384–1395.
3. Gao D, Jiang N, Wing RA, Jiang J, Jackson SA (2015) Transposons play an important role in the evolution and diversification of centromeres among closely related species. *Front Plant Sci* 6:216.
4. Wong LH, Choo KH (2004) Evolutionary dynamics of transposable elements at the centromere. *Trends Genet* 20:611–616.
5. Cam HP, Noma K, Ebina H, Levin HL, Grewal SI (2008) Host genome surveillance for retrotransposons by transposon-derived proteins. *Nature* 451:431–436.
6. Mateo L, González J (2014) *Pogo*-like transposases have been repeatedly domesticated into CENP-B-related proteins. *Genome Biol Evol* 6:2008–2016.
7. Goodier JL, Kazazian HH, Jr (2008) Retrotransposons revisited: The restraint and rehabilitation of parasites. *Cell* 135:23–35.
8. Comai L, Maheshwari S, Marimuthu MPA (2017) Plant centromeres. *Curr Opin Plant Biol* 36:158–167.
9. Rhind N, et al. (2011) Comparative functional genomics of the fission yeasts. *Science* 332:930–936.
10. Chan FL, Wong LH (2012) Transcription in the maintenance of centromere chromatin identity. *Nucleic Acids Res* 40:11178–11188.
11. Hall LE, Mitchell SE, O'Neill RJ (2012) Pericentric and centromeric transcription: A perfect balance required. *Chromosome Res* 20:535–546.
12. Pezer Z, Ugarković D (2008) Role of non-coding RNA and heterochromatin in aneuploidy and cancer. *Semin Cancer Biol* 18:123–130.
13. Du Y, Topp CN, Dawe RK (2010) DNA binding of centromere protein C (CENPC) is stabilized by single-stranded RNA. *PLoS Genet* 6:e1000835.
14. Scott KC (2013) Transcription and ncRNAs: At the cent(rome)re of kinetochore assembly and maintenance. *Chromosome Res* 21:643–651.
15. Volpe T, Martienssen RA (2011) RNA interference and heterochromatin assembly. *Cold Spring Harb Perspect Biol* 3:a003731.
16. Pidoux AL, Allshire RC (2005) The role of heterochromatin in centromere function. *Philos Trans R Soc Lond B Biol Sci* 360:569–579.
17. Dumesic PA, Madhani HD (2014) Recognizing the enemy within: Licensing RNA-guided genome defense. *Trends Biochem Sci* 39:25–34.
18. Fedoroff NV (2012) Presidential address. Transposable elements, epigenetics, and genome evolution. *Science* 338:758–767.
19. Matzke MA, Mosher RA (2014) RNA-directed DNA methylation: An epigenetic pathway of increasing complexity. *Nat Rev Genet* 15:394–408.
20. Padmanabhan S, Thakur J, Siddharthan R, Sanyal K (2008) Rapid evolution of Cse4p-rich centromeric DNA sequences in closely related pathogenic yeasts, *Candida albicans* and *Candida dubliniensis. Proc Natl Acad Sci USA* 105:19797–19802.
21. Kapoor S, Zhu L, Froyd C, Liu T, Rusche LN (2015) Regional centromeres in the yeast *Candida lusitaniae* lack pericentromeric heterochromatin. *Proc Natl Acad Sci USA* 112:12139–12144.
22. Sanyal K, Baum M, Carbon J (2004) Centromeric DNA sequences in the pathogenic yeast *Candida albicans* are all different and unique. *Proc Natl Acad Sci USA* 101:11374–11379.
23. Chatterjee G, et al. (2016) Repeat-associated fission yeast-like regional centromeres in the ascomycetous budding yeast *Candida tropicalis. PLoS Genet* 12:e1005839.
24. D'Souza CA, et al. (2011) Genome variation in *Cryptococcus gattii*, an emerging pathogen of immunocompetent hosts. *mBio* 2:e00342-10.
25. Janbon G, et al. (2014) Analysis of the genome and transcriptome of *Cryptococcus neoformans* var. *grubii* reveals complex RNA expression and microevolution leading to virulence attenuation. *PLoS Genet* 10:e1004261.
26. Loftus BJ, et al. (2005) The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans. Science* 307:1321–1324.
27. Bielska E, May RC (2016) What makes *Cryptococcus gattii* a pathogen? *FEMS Yeast Res* 16:fov106.
28. Dixit A, Carroll SF, Qureshi ST (2009) *Cryptococcus gattii*: An emerging cause of fungal disease in North America. *Interdiscip Perspect Infect Dis* 2009:840452.
29. Feretzaki M, Billmyre RB, Clancey SA, Wang X, Heitman J (2016) Gene network polymorphism illuminates loss and retention of novel RNAi silencing components in the *Cryptococcus* pathogenic species complex. *PLoS Genet* 12:e1005868.
30. Kozubowski L, et al. (2013) Ordered kinetochore assembly in the human-pathogenic basidiomycetous yeast *Cryptococcus neoformans. mBio* 4:e00614-13.
31. Smith KM, Phatale PA, Sullivan CM, Pomraning KR, Freitag M (2011) Heterochromatin is required for normal distribution of *Neurospora crassa* CenH3. *Mol Cell Biol* 31:2528–2542.
32. Sun S, Xu J (2009) Chromosomal rearrangements between serotype A and D strains in *Cryptococcus neoformans. PLoS One* 4:e5524.
33. Goodwin TJ, Poulter RT (2001) The diversity of retrotransposons in the yeast *Cryptococcus neoformans. Yeast* 18:865–880.
34. Sun S, et al. (2017) Fungal genome and mating system transitions facilitated by chromosomal translocations involving intercentromeric recombination. *PLoS Biol* 15:e2002527.
35. Janbon G, et al. (2010) Characterizing the role of RNA silencing components in *Cryptococcus neoformans. Fungal Genet Biol* 47:1070–1080.
36. Wang X, et al. (2010) Sex-induced silencing defends the genome of *Cryptococcus neoformans* via RNAi. *Genes Dev* 24:2566–2582.
37. Farrer RA, et al. (2015) Genome evolution and innovation across the four major lineages of *Cryptococcus gattii. mBio* 6:e00868-15.
38. Huff JT, Zilberman D (2014) Dnmt1-independent CG methylation contributes to nucleosome positioning in diverse eukaryotes. *Cell* 156:1286–1297.
39. Maloisel L, Rossignol JL (1998) Suppression of crossing-over by DNA methylation in *Ascobolus. Genes Dev* 12:1381–1389.
40. Mirouze M, et al. (2012) Loss of DNA methylation affects the recombination landscape in *Arabidopsis. Proc Natl Acad Sci USA* 109:5880–5885.
41. Laurie JD, et al. (2012) Genome comparison of barley and maize smut fungi reveals targeted loss of RNA silencing components and species-specific presence of transposable elements. *Plant Cell* 24:1733–1745.
42. Rabe F, et al. (2016) A complete toolset for the study of *Ustilago bromivora* and *Brachypodium* sp. as a fungal-temperate grass pathosystem. *eLife* 5:e20522.
43. Kämper J, et al. (2006) Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis. Nature* 444:97–101.
44. Henikoff S, Ahmad K, Malik HS (2001) The centromere paradox: Stable inheritance with rapidly evolving DNA. *Science* 293:1098–1102.
45. Melters DP, et al. (2013) Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biol* 14:R10.
46. Bensasson D, Zarowiecki M, Burt A, Koufopanou V (2008) Rapid evolution of yeast centromeres in the absence of drive. *Genetics* 178:2161–2167.
47. Dumesic PA, et al. (2013) Stalled spliceosomes are a signal for RNAi-mediated genome defense. *Cell* 152:957–968.
48. Ellermeier C, et al. (2010) RNAi and heterochromatin repress centromeric meiotic recombination. *Proc Natl Acad Sci USA* 107:8701–8705.
49. Mravinac B, et al. (2009) Histone modifications within the human X centromere region. *PLoS One* 4:e6602.
50. Steiner NC, Clarke L (1994) A novel epigenetic effect can alter centromere function in fission yeast. *Cell* 79:865–874.
51. Chuong EB, Elde NC, Feschotte C (2017) Regulatory activities of transposable elements: From conflicts to benefits. *Nat Rev Genet* 18:71–86.
52. Devos KM, Brown JK, Bennetzen JL (2002) Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis. Genome Res* 12:1075–1079.
53. Vitte C, Panaud O (2005) LTR retrotransposons and flowering plant genome size: Emergence of the increase/decrease model. *Cytogenet Genome Res* 110:91–107.
54. Horváth V, Merenciano M, González J (2017) Revisiting the relationship between transposable elements and the eukaryotic stress response. *Trends Genet* 33:832–841.
55. Nicolás FE, Torres-Martínez S, Ruiz-Vázquez RM (2013) Loss and retention of RNA interference in fungi and parasites. *PLoS Pathog* 9:e1003089.

GENETICS