Lawrence Berkeley National Laboratory
Recent Work

Title

Early Diverging Fungus Mucor circinelloides Lacks Centromeric Histone CENP-A and Displays a Mosaic of Point and Regional Centromeres.

Permalink
https://escholarship.org/uc/item/20z1h4bj

Journal
Current biology : CB, 29(22)

ISSN
0960-9822

Authors
Navarro-Mendoza, María Isabel
Pérez-Arques, Carlos
Panchal, Shweta
et al.

Publication Date
2019-11-01

DOI
10.1016/j.cub.2019.09.024

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

1 **Early diverging fungus *Mucor circinelloides* lacks**
2 **centromeric histone CENP-A and displays a mosaic of point and**
3 **regional centromeres**

4 María Isabel Navarro-Mendoza[1]#, Carlos Pérez-Arques[1]#, Shweta
5 Panchal[2], Francisco E. Nicolás[1], Stephen J. Mondo[3,4], Promit Ganguly[2],
6 Jasmyn Pangilinan[3], Igor V. Grigoriev[3,5], Joseph Heitman[6]*, Kaustuv
7 Sanyal[2]*, and Victoriano Garre[1]*

8 [1]Department of Genetics and Microbiology, Faculty of Biology, University
9 of Murcia, Murcia, 30100, Spain; [2]Molecular Biology and Genetics Unit,
10 Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur,
11 Bangalore 560064, India; [3]US Department of Energy Joint Genome
12 Institute, Walnut Creek, California, USA; [4]Bioagricultural Science and Pest
13 Management Department, Colorado State University, Fort Collins, CO
14 80521, USA; [5]Department of Plant and Microbial Biology, University of
15 California Berkeley, Berkeley, CA 94598, USA; [6]Department of Molecular
16 Genetics and Microbiology, Duke University Medical Center, NC, USA

17 #Equally contributed; *Corresponding authors

18 Author Contact: Victoriano Garre, vgarre@um.es; Kaustuv Sanyal,
19 sanyal@jncasr.ac.in; Joseph Heitman, heitm001@duke.edu.

20 Lead Contact: Victoriano Garre, vgarre@um.es

21 **Summary**

22 Centromeres are rapidly evolving across eukaryotes, despite
23 performing a conserved function to ensure high fidelity chromosome

segregation. CENP-A chromatin is a hallmark of a functional centromere in most organisms. Due to its critical role in kinetochore architecture, the loss of CENP-A is tolerated in only a few organisms, many of which possess holocentric chromosomes. Here, we characterize the consequence of the loss of CENP-A in the fungal kingdom. *Mucor circinelloides*, an opportunistic human pathogen, lacks CENP-A along with the evolutionarily conserved CENP-C, but assembles a monocentric chromosome with a localized kinetochore complex throughout the cell cycle. Mis12 and Dsn1, two conserved kinetochore proteins were found to colocalize to a short region, one in each of nine large scaffolds, comprised of an ~200-bp AT-rich sequence followed by a centromere-specific conserved motif that echoes the structure of budding yeast point centromeres. Resembling fungal regional centromeres, these core centromere regions are embedded in large genomic expanses devoid of genes yet marked by Grem-LINE1s, a novel retrotransposable element silenced by the Dicer-dependent RNAi pathway. Our results suggest that these hybrid features of point and regional centromeres arose from the absence of CENP-A, thus defining novel mosaic centromeres in this early-diverging fungus.

Introduction

Accurate chromosome segregation is crucial to maintain genome integrity during cell division. The timely attachment of microtubules to centromere DNA is essential to achieve proper chromosome segregation. This is accomplished by a specialized multilayered protein complex, the

48 kinetochore which links microtubules to centromere DNA. This protein
49 bridge is divided into two layers – the inner and outer kinetochore. The
50 fundamental inner kinetochore protein is the histone H3 variant CENP-A. It
51 binds directly to centromere DNA and lays the foundation to recruit other
52 essential proteins of the kinetochore complex, playing a fundamental role
53 in centromere structure and function, and hence, precise chromosome
54 segregation [1,2]. CENP-A is also found at all identified neocentromeres
55 [3] and at the active centromeres of dicentric chromosomes [4], acting as
56 an epigenetic determinant of centromeric identity.

57 Despite its conserved function, the centromere is one of the most
58 rapidly evolving regions of the genome [5]. This so-called “centromere
59 paradox” has led to centromeres of diverse sizes and content. The first
60 centromeres identified in *Saccharomyces cerevisiae* were found to be
61 point centromeres - small regions of ~120 bp defined by specific DNA
62 sequences [6,7]. In contrast to point centromeres described in only a few
63 budding yeasts of the phylum Ascomycota, most other fungi and
64 metazoans have regional centromeres that are larger, ranging from a few
65 kilobases to several megabases [8]. Regional centromeres are often
66 interspersed with repetitive sequences and are mostly defined by
67 epigenetic factors rather than DNA sequence per se [9]. While the
68 centromeres described thus far are confined to one region of the DNA,
69 there are organisms in which their entire chromosomes are loaded with
70 kinetochores and display a parallel separation of the two chromatids. Such
71 chromosomes possess holocentromeres and have been found in several
72 clades of insects [10] and nematodes like *Caenorhabditis elegans [11].*

73 Despite this remarkable variation in the types of centromeres, most
74 organisms have their centromere loci defined by the binding of CENP-A.
75 However, there are a few exceptions, including some insect lineages and
76 kinetoplastids, which have independently lost CENP-A. The kinetoplastid
77 *Trypanosoma brucei* has evolved a unique set of proteins to perform the
78 role of the kinetochore complex [12], whereas all insects that have
79 recurrently lost CENP-A have transitioned from monocentricity to
80 holocentricity [10].

81 Neither holocentricity nor lack of CENP-A has ever been reported in
82 the fungal kingdom until a recent study showed that this protein could be
83 absent in the Mucoromycotina subphylum [13], which belongs to the early
84 diverging fungi. Kinetochore structure and centromere function are well-
85 studied in species of the Dikarya, whereas centromere identity as well as
86 the kinetochore architecture remain largely unexplored in basal fungi
87 because these are challenging to manipulate genetically. *Mucor*
88 *circinelloides*, a fungus of the subphylum Mucoromycotina, is an exception
89 and molecular tools are available to modify its genome [14]. *M.*
90 *circinelloides*, as with other Mucorales, causes an infectious disease
91 known as mucormycosis, which have been associated with high mortality
92 rates due to rhino-orbital-cerebral, pulmonary, or cutaneous infections
93 [15]. Despite its medical significance and research efforts to identify new
94 antifungal targets, the genome biology and cell cycle of this fungus
95 remain poorly understood. Using evolutionarily conserved kinetochore
96 proteins as tools, we investigated the dynamics of the kinetochore
97 complex during nuclear division and unveiled the unique characteristics of

98 the centromeres of *M. circinelloides*. Studying distinct and divergent
99 factors involved in the cell cycle of this organism will open avenues for the
100 development of species-specific antifungal drugs.

101 **Results**

102 **Last Mucoralean common ancestor lost CENP-A and CENP-C, and**
103 **the descendant Mucoromycotina lack centromere-specific histone**
104 **variants**

105 A recent study of kinetochore evolution in eukaryotes reported the
106 striking absence of two well conserved inner kinetochore proteins, CENP-A
107 and CENP-C, in two species of the Mucoromycotina [13]. To ascertain the
108 presence or loss of well-studied kinetochore proteins across the entire
109 subphylum, we analyzed the curated proteomes of 75 fungal species,
110 including 55 species of Mucoromycotina and 20 species from other fungal
111 clades (Table S1). The kinetochore protein sequences from *Mus musculus*,
112 *Ustilago maydis*, *S. cerevisiae*, and *Schizosaccharomyces pombe* served
113 as queries to conduct iterative Blast and Hmmer searches against these
114 proteomes. These sequences were confirmed as putative orthologs by the
115 presence of conserved Pfam domains or by obtaining a matching first hit
116 in a reciprocal Blastp search against the initial four species: mouse, *U.*
117 *maydis*, budding, and fission yeasts (Table S1, protein sequences at
118 Mendeley data DOI: 10.17632/wpyyb58h8v.1).

119 To confirm the presence or absence of centromere-specific histone
120 H3 CENP-A in each fungal species, 289 histone H3 homologs were
121 analyzed, identifying 27 putative CENP-A proteins (Figure S1, Table S1,

122 protein sequences at Mendeley data DOI: 10.17632/wpyyb58h8v.1). These
123 proteins share common features with well-known CENP-A proteins [16],
124 such as an N-terminal tail that varies significantly in length and sequence
125 compared to the canonical H3 counterparts (Figure S1B), the absence of a
126 conserved glutamine residue in the first α-helix (69Q) and a phenylalanine
127 residue (84F) (Figure S1C), a longer and divergent first loop (Figure S1C),
128 and a non-conserved C-terminal tail (Figure S1D). In addition, the histone
129 fold domain (HFD) sequences of the putative CENP-A proteins have
130 diverged rapidly from the conserved canonical H3 sequences [16] (Figure
131 S1E). Nine additional H3 homologs also share several CENP-A common
132 features but were regarded as rare histones because they lack these
133 conserved substitutions. The analysis revealed that CENP-A was absent in
134 most of the Mucoromycotina clade, specifically in the orders
135 Umbelopsidales and Mucorales; nevertheless, we identified putative CENP-
136 A proteins in all Endogonalean species (Figure 1). Similarly, CENP-C is
137 present in the Endogonales and Umbelopsidales but not in the Mucorales
138 (Figure 1). These findings suggest that the last Mucoromycotina common
139 ancestor possessed CENP-A, which was lost after the Endogonales
140 diverged from the rest of the clade. In the absence of CENP-A, CENP-C was
141 presumably lost soon after. In addition, CENP-Q is absent in all Mucoralean
142 species, and CENP-U is lost in the whole subphylum. These proteins have
143 been recurrently lost across the fungal kingdom, as evidenced by their
144 absence in species of all phyla of early-diverging fungi and most species
145 in Basidiomycota (Figure 1). The absence of orthologs that were present in
146 closely-related species should be taken with caution because this could be

147 a result of the incomplete state of their genome assemblies and
148 annotation.

149 Despite these relevant losses, the remaining kinetochore proteins
150 were well conserved among most species of Mucoromycotina, especially
151 in *M. circinelloides*. Therefore, we hypothesized that another histone
152 protein could have taken the role of CENP-A as a part of centromeric
153 nucleosome for kinetochore assembly. We searched the *M. circinelloides*
154 genome sequence for histone protein-coding genes and examined their
155 sequences seeking distinctive features of histone variants. Genes
156 encoding mammalian histone variants have intronic sequences and
157 polyadenylation [poly(A)] signals, in contrast to canonical histones H3.1
158 and H3.2 [17]. However, fungal canonical histones may have intronic
159 sequences and poly(A) signals [18]. Thus, the analysis of possible variants
160 focused on the presence of substitutions in relevant amino acid residues
161 targeted by post-translational modifications [19] rather than intronic and
162 poly(A) sequences. The *M. circinelloides* genome contains four histone H3-
163 and three histone H4-coding genes, named *histone H three* (*hht1-4*) and
164 *histone H four* (*hhf1-3*), respectively (Figure S2A), and all of them feature
165 a putative poly(A) signal (5'-AATAAA-3'). Three of the four histone H3-
166 coding genes are intronless and their protein sequence identical, except
167 the intron-containing gene *hht4* (Figure S2A). Also, the amino acid
168 residues at positions 32, 43, and 97 in the Hht4 protein sequence differ
169 compared to the conserved H3 sequence (Figure S2B). Similarly, histone
170 H4-coding genes *hhf2* and *hhf3* lack introns and encode identical proteins,

171 whereas *hhf1* has intronic sequences (Figure S2A) and its product has a
172 longer N-terminal tail (Figure S2C).

173 To test if the minor differences observed in Hht4 and Hhf1 could
174 have resulted in a centromere-specific binding protein, the cellular
175 localization was analyzed and compared to their canonical histone
176 counterpart Hhf3. We constructed alleles that contained each histone-
177 coding gene fused in-frame at the C-terminus with the enhanced green
178 fluorescent protein (eGFP) flanked by each of the histone gene native
179 promoter and terminator sequences, and a selectable marker (Figure S3).
180 Homokaryotic strains expressing fluorescent histones Hht4, Hhf1, or Hhf3
181 were obtained by transforming a wild-type strain with their respective
182 eGFP-tagged alleles. Integration by homologous recombination at the
183 corresponding native loci was confirmed by PCR (Figure S3 and Table S2).

184 *M. circinelloides* is a saprophytic fungus with defined sexual and
185 asexual lifecycles [20]. The asexual reproductive cycle is the preferred
186 culture method to maintain this organism under laboratory conditions, due
187 to fast developing mycelia and sporangiospores that contain asexual
188 spores [14]. Asexual spores from these strains were observed with a
189 confocal microscope under two stages of germination: ungerminated and
190 4-h pre-germinated spores, that had started swelling and forming germ
191 tubes. In both conditions, all fluorescent histone fusions were localized
192 encompassing the entire nucleus instead of forming kinetochore-like
193 nuclear clusters (Figure S2D). Most species of the Mucoromycotina are
194 coenocytic fungi and have multinucleated spores, as does the *M.*

195 *circinelloides* wild-type strain used in this study which has large spores
196 with a high number of nuclei [21]. We established that nuclear division in
197 *M. circinelloides* is asynchronous by observing the nuclear distribution of
198 Hht4-eGFP protein in pre-germinated spores during a time-lapse imaging
199 session (Video S1). Altogether, these findings indicate that either none of
200 these histone variants functions like CENP-A in kinetochore assembly or
201 *Mucor* centromeres are holocentric in nature. We could test these two
202 possibilities by studying other kinetochore proteins that are evolutionarily
203 conserved and easily identifiable in *M. circinelloides.*

204 **Mis12 and CENP-T complexes show discrete nuclear localization**

205 Analyzing the nuclear localization and assembly dynamics of the
206 kinetochore complex of *M. circinelloides* could offer insights into its
207 architecture. In the absence of CENP-A or another similar histone H3
208 variant, and CENP-C, that could anchor the kinetochore formation, we
209 focused our attention on the evolutionarily conserved and obvious
210 homologs of inner kinetochore proteins that are expected to be closer to
211 the centromere DNA in this organism (Figure 2A). We designed
212 fluorescent kinetochore fusion proteins of Mis12, Dsn1, and CENP-T
213 because these are well-conserved in almost all species of the
214 Mucoromycotina. First, we constructed alleles to tag the Mis12 and Dsn1
215 proteins with the red fluorescent protein mCherry, both at their C-termini
216 and N-termini, aiming to integrate the alleles in their endogenous loci
217 following a similar strategy as for the epitope tagging of the histone genes
218 described above. Homologous recombination in the desired transformants

219 was confirmed by PCR, and six colonies from the four tagged alleles were
220 checked for fluorescent localization. Unfortunately, fluorescent signals
221 were not detected in any of the four different tagged strains possibly due
222 to low levels of expression.

223 As an alternative approach, the Mis12 and Dsn1 proteins were
224 tagged with mCherry at each end, N-terminal and C-terminal, separately,
225 and overexpressed from a strong promoter (P*zrt1*) in *M. circinelloides*
226 (Figures 2B and S3). The alleles were designed to target integration into
227 the *carRP* locus, a gene encoding an enzyme involved in carotenogenesis.
228 Thus, the mutant strains obtained should lack colored-carotenes and
229 display an albino phenotype [14]. After confirming the integration of the
230 alleles in the *carRP* locus by PCR (Figures S3E and S3F, Table S2), one
231 strain harboring each construct was used as a parental strain to integrate
232 the allele for H3-eGFP expression at its own locus as previously described,
233 allowing monitoring colocalization of the kinetochore proteins within the
234 nucleus. Once the double integration was confirmed, two independent
235 strains of each type were selected for fluorescent localization.

236 Mis12 and Dsn1 displayed a clear localization signal forming small
237 fluorescent puncta within the nucleus marked by histone H3 (Figures 2C
238 and 2D, respectively). The microscopy screening confirmed that the
239 expression of both N-terminal and C-terminal tagged proteins was similar.
240 The Mis12 and Dsn1 localization patterns revealed a single cluster of
241 kinetochores as observed in many fungal species. This result strongly

242 suggests that *M. circinelloides* has monocentric chromosomes with a
243 localized kinetochore even in the absence of CENP-A.

244 Next, another evolutionarily conserved inner kinetochore protein
245 CENP-T, which is known to interact with DNA, was tagged with eGFP at the
246 C-terminus following the same procedure used for the histone tagging and
247 integrated by homologous recombination at the endogenous locus in
248 strains expressing the Mis12-mCherry and Dsn1-mCherry fusion proteins.
249 Integration of the alleles at the desired loci was confirmed by PCR in both
250 parental backgrounds (Figure S3D), and the localization of the kinetochore
251 proteins was examined by analyzing their fluorescent signal in pre-
252 germinated spores. CENP-T colocalized with Mis12 and Dsn1 in each
253 nucleus (Figures 2E and 2F, respectively), indicating that all three proteins
254 assemble together in the kinetochore ensemble. Time-lapse imaging
255 during spore germination of the double-tagged strains allowed the
256 analysis of nuclear division and kinetochore structure during the cell cycle
257 in live cells. The Mis12 and Dsn1 proteins clustered at the nuclear
258 periphery during division (Figure 2G), a typical feature of fungal
259 kinetochore proteins, suggesting all these three proteins are constitutively
260 present at the kinetochore during the cell cycle.

261 ***M. circinelloides* exhibits small mosaic centromeres associated**
262 **with RNAi-suppressed L1-like retrotransposons**

263 The centromere regions of early-diverging fungi have never been
264 described, possibly because of their genetic intractability. The generation
265 of kinetochore-tagged strains of *M. circinelloides* offered the opportunity

266 to identify the nature of centromeres in an organism that lost CENP-A and
267 CENP-C. To map *M. circinelloides* centromeres, chromatin
268 immunoprecipitation followed by next generation sequencing (ChIP-seq)
269 was performed with Mis12- and Dsn1-mCherry proteins. Illumina
270 sequencing was conducted in two independently immunoprecipitated (*IP*
271 DNA) samples for each kinetochore protein, one for Mis12-mCherry and
272 the other for Dsn1-mCherry, using RFP-Trap MA beads (Chromotek) for
273 mCherry immunoprecipitation. Input control samples consisting of total
274 sheared DNA (*Input* DNA) were also sequenced for each tagged strain, as
275 well as mock binding control (*beads only* DNA) samples with MA beads
276 without antibody. Raw sequencing data were processed, and reads were
277 aligned to a newly generated PacBio genome assembly of *Mucor*
278 *circinelloides* f. *lusitanicus* MU402 strain (available at JGI
279 http://genome.jgi.doe.gov/Muccir1_3/Muccir1_3.home.html, hereafter the
280 *M. circinelloides* genome), the parental strain of the kinetochore protein-
281 tagged strains. Out of the whole genome assembly of 24 scaffolds, which
282 has an average GC content of 42.19%, 11 major scaffolds encompass
283 more than 90% of the genome.

284 Both Mis12 and Dsn1 *IP* DNA sequence reads were compared to the
285 corresponding control *Input* DNA reads to identify statistically significant
286 kinetochore protein-enriched genomic loci in both *IP* DNA samples (FDR ≤
287 $5x10^{-5}$, fold enrichment ≥ 1.6; Table S3). The result of this analysis was
288 visualized in a genome browser to ensure that the enrichment was robust
289 and absent in either the *Input* or *beads only* DNA controls (Figure S4). By
290 this approach, we detected nine significantly enriched peaks that

291 overlapped in both Mis12 and Dsn1 *IP* DNA samples, and thus we defined
292 kinetochore protein-bound regions as core centromeres (CCs) (Figure 3A)
293 that were designated by the corresponding scaffold number. Each of these
294 core *CEN* sequences was located in a different major scaffold of the
295 genome assembly (Figure 3B), indicating that each chromosome was
296 assembled into an average of one to three scaffolds. The identification of
297 five repeated regions matching the telomeric sequence (5'-TTAGGG-3') at
298 the 3'-end of scaffolds 2, 3, 8, 9, and 10, suggested that *CC2* and *CC9*
299 display a subtelomeric localization (Figure 3B). Because the length of the
300 overlapping peaks was minimally different in the two kinetochore *IP* DNA
301 samples, DNA sequence enriched with at least one of the kinetochore
302 proteins was assumed to be the length of kinetochore binding region on
303 each chromosome. In addition, contiguous enriched regions were added to
304 each *CC*, defining the core centromere regions (Table S3).

305 The average length of core centromere regions of *M. circinelloides* is
306 941 bp (Figure 4A). These unusually short regions are consistently AT-rich
307 (71% average) (Figure 4B), resembling point centromeres of *S. cerevisiae*
308 and its related budding yeast species. Further analysis revealed a 41-bp
309 DNA sequence motif conserved in all nine kinetochore protein-bound core
310 centromere regions (Figure 4C). An extensive search of the presence of
311 this motif across the whole genome of *M. circinelloides* revealed that the
312 41-bp sequence is centromere-specific, occurs as a single-copy element at
313 each core centromere, and is absent in the rest of the genome. Moreover,
314 a closer inspection of the core centromere regions revealed a conserved
315 pattern. Each core centromere is composed of a minor peak enriched of

316 kinetochore proteins spaced by a highly AT-rich stretch spanning the
317 major binding peak of kinetochore proteins that starts with the 41-bp
318 conserved motif (Figure 4D).

319 Surrounding the core centromere regions, both upstream and
320 downstream, are stretches of genomic sequences completely devoid of
321 genes. These pericentric regions are considerably larger than the core and
322 vary in length, ranging from 15 to 75 kb (Figure 5A). These pericentric
323 regions contain a variable number of large repeats oriented away from the
324 core centromere regions (Figure S5). These repeats cluster in nine groups
325 according to their similarity (≥ 95% identity and ≥ 70% coverage across
326 their entire sequence). Clusters 1, 2 and 3 contain all of the full-length
327 elements, which were termed autonomous elements, while the remaining
328 clusters are composed of incomplete repeats or remnants (Figure S6A). A
329 pairwise comparison of all of the repeats revealed a high similarity in their
330 aligned sequences (≥ 90% identity), though the truncated sequences of
331 remnants showed higher differences accounting for the gaps in the
332 alignment.

333 Each autonomous repeat comprises two sequential open reading
334 frames (ORFs) with an overlapping codon. The first ORF encodes a product
335 containing several zinc finger domains, and the second codes for a protein
336 with AP endonuclease and reverse transcriptase domains (Figure 5B,
337 Table S4). Based on their high similarity, their ORF domain architecture,
338 the presence of a poly(A) signal at their 3’-end, and the absence of long
339 terminal repeats (LTR) at either end, they were classified as a single non-

340 LTR retrotransposon. Phylogenetic profiling revealed that these
341 retrotransposons belong to the LINE1 clade (Figure S6B) and are closely
342 related to other LINE1-like retrotransposons of Mucoromycotina species,
343 so they were termed Genomic Retrotransposable Element of
344 Mucoromycotina (Grem) LINE1-like. A search of the *M. circinelloides*
345 genome revealed that the Grem-LINE1s are found specifically in the
346 pericentromeric regions; or at the non-telomeric ends of three small
347 scaffolds, oriented away from the end (Table S4). Four major scaffolds end
348 abruptly in a putative centromere region, suggesting that they may be
349 linked to these minor scaffolds that contain the Grem-LINE1 elements at
350 one end. Broadening the search across 55 genomes of species of
351 Mucoromycotina identified the Grem-LINE1 elements in most Mucorales
352 and Umbelopsidales (≥ 30% identity and ≥ 50% coverage), but they could
353 not be found in any of the Endogonalean genomes, indicating an inverse
354 correlation between the presence of Grem-LINE1s and CENP-A (Figure 1
355 and Table S5).

356 RNAi silences expression and prevents transposition of
357 retrotransposable elements in several fungal species featuring functional
358 RNAi [22,23]. *M. circinelloides* possess an elaborate RNAi mechanism that
359 protects the genome against invading elements, that could prevent the
360 movement of retrotransposable elements. The two Dicer enzymes,
361 particularly Dcl2, and one of the three Argonaute proteins (Ago1), play a
362 pivotal role in this protective RNAi-related pathway [24]. To analyze the
363 transcriptional landscape of the pericentric regions, previously generated
364 and publicly available transcriptomic raw data for mRNA and small RNAs

365 were reprocessed and aligned to the *M. circinelloides* genome assembly.
366 The pericentric regions were almost devoid of transcription in the wild-
367 type strain, although the retrotransposons were modestly expressed.
368 Indeed, a high number of sRNAs corresponding to the retrotransposons
369 were detected (Figures 5C and S6), indicating an inverse correlation
370 between mRNA and sRNA levels. This inverse correlation was confirmed
371 by analyzing the transcriptomic profile of mutant strains lacking essential
372 components of the RNAi machinery, Ago1 or Dcl1/Dcl2. In these RNAi-
373 deficient mutants, the pattern was reversed; high mRNA levels from the
374 retrotransposons were observed, while the production of sRNAs
375 originating from those regions was considerably lower than in the wild-
376 type strain. These findings suggest that the retrotransposons contained in
377 the pericentric regions of *M. circinelloides* are silenced by the Dicer-
378 dependent RNAi pathway.

379 In most organisms, deposition of CENP-A at the centromere-specific
380 nucleosomes is correlated to depletion of histone H3 [25–27]. Because
381 CENP-A is absent in *M. circinelloides*, we quantified the presence of
382 histone H3 at the centromere-specific nucleosomes by performing ChIP
383 with anti-histone H3 antibodies and a qPCR analysis. The occupancy of
384 histone H3 was examined by qPCR in four regions across the scaffold_2:
385 the spanning kinetochore protein-binding region (core centromere, *CC2*),
386 the ORF-free pericentromeric region, the centromere-flanking ORFs and
387 the control region far-CEN ORF ~2 Mb away from the centromere (Figure
388 5D). There are no significant differences in histone H3 occupancy across
389 the analyzed regions, indicating that H3 is not depleted at either the

390 kinetochore protein binding region or the pericentric region compared to
391 the flanking ORFs and the control far-CEN ORF (Figure 5D).

392 **Discussion**

393 Mucoralean species are the causative agents of mucormycosis, an
394 emerging fungal infection with extremely poor clinical outcomes, probably
395 as a consequence of their innate resistance to all current antifungal drugs
396 [28,29] and their ability to evade host innate immunity [30,31]. In spite of
397 their clinical relevance, limited information is available on essential
398 biological processes in these fungal pathogens, especially nuclear division
399 and cell cycle. In this study, we identified major kinetochore components
400 in *M. circinelloides*, and utilized them to identify and characterize nine
401 chromosomal loci as centromeres.

402 In most eukaryotes, the centrochromatin is composed of
403 nucleosomes containing the histone H3 variant CENP-A as a hallmark of
404 centromere identity. The most striking observation of our study was the
405 absence of this centromere determining protein in all Mucoralean species.
406 Absence of CENP-A was accompanied by loss of CENP-C, while most other
407 kinetochore protein orthologs were present in these species. Loss of a
408 previously-thought indispensable protein like CENP-A raised questions on
409 the evolutionary bases of centromere and kinetochore structure. Absence
410 of CENP-A had been previously observed in certain insect orders including
411 butterflies and moths, bugs and lice, earwigs, and dragonflies which in all
412 known cases led to a transition to holocentricity, suggesting that the

413 presence of a centromere-specific protein is dispensable for segregation
414 of a chromosome with localized kinetochore[10].

415 Our study is the first to demonstrate loss of core kinetochore
416 proteins in the fungal kingdom, reinforcing that the absence of CENP-A
417 can be tolerated in eukaryotic organisms. Surprisingly, punctate
418 localization patterns of other conserved kinetochore proteins including
419 Mis12, Dsn1, and CENP-T throughout the cell cycle indicate that these
420 organisms still retain a monocentric arrangement of a canonical
421 kinetochore structure, as opposed to the holocentric insects that lack
422 CENP-A. Kinetochore proteins in *M. circinelloides* were observed to be
423 localized at the periphery of each nucleus in the asexual spores as well as
424 mycelia. Three types of nuclear divisions have been studied in
425 multinucleated fungi – synchronous, asynchronous, and parasynchronous
426 [32]. In all ascomycetes except *S. pombe* and *Zymoseptoria tritici*,
427 kinetochore proteins are centromere-localized and clustered throughout
428 the cell cycle [33,34]. In the basidiomycete *C. neoformans,* kinetochore
429 proteins transition between clustering and declustering states in different
430 stages of the cell cycle [35].

431 We have observed asynchronous nuclear divisions during
432 germination of asexual spores of *M. circinelloides*, a multinucleated and
433 coenocytic fungus. The kinetochore was found to be consistently present
434 as a cluster in all stages of mitosis, implying that the basic kinetochore
435 structure in this organism is constitutively assembled throughout the cell
436 cycle. Overall, these findings hint that the role of CENP-A is being carried

437 out by either a distinct protein that has a newly evolved function or by
438 one of the known kinetochore proteins like CENP-T, which also has a
439 histone-fold DNA binding domain. We also tried to find whether one of the
440 histone H3 variants plays a role of centromeric histone, but none of them
441 displayed kinetochore-like localization, indicating that they are probably
442 not involved in replacing CENP-A function in chromosome segregation
443 mediated by the kinetochore machinery. Usually, histone H3 is depleted
444 at the centromeres as compared to other non-centromeric loci in the
445 genome. We observed the presence of H3 at the relatively short
446 kinetochore protein-binding region at a level that is similar to the
447 pericentric region or gene bodies. Thus, we propose that centromeres in
448 this organism contain histone H3 at the centromere core. Further studies
449 will provide the exact composition of centromere-specific nucleosomes in
450 this basal fungus.

451 A ChIP-based analysis using Dsn1 and Mis12 revealed the first
452 glimpse of the centromere structure of this Mucoromycotina species. The
453 centromeres in this organism possess features of both point and regional
454 centromeres, resulting in a unique mosaic; thus, we named them mosaic
455 centromeres (Figure 6). One of their distinctive features is a small
456 kinetochore protein binding domain (~750-1150 bp) defined by the
457 presence of an AT-rich sequence stretch (~70-80% AT, ~100 bp), which is
458 followed by a conserved centromere-specific 41-bp sequence motif. This
459 motif is only found in *M. circinelloides* centromeres and marks the start of
460 the major kinetochore-binding region. Overall, this pattern resembles the
461 distribution of conserved elements in *S. cerevisiae* point centromeres, in

462 which two conserved Centromere Determining Elements (CDEI and CDEIII)
463 are separated by an AT-rich non-conserved DNA sequence stretch (CDEII)
464 [7]. The kinetochore of *S. cerevisiae* binds CDEI and CDEIII in a sequence-
465 dependent manner through Cbf1 [36] and the CBF3 complex [37,38],
466 respectively. The presence of a conserved DNA motif in the centromeres
467 of *M. circinelloides* suggests that it could be a binding site for an unknown
468 centromere-specific protein in this species.

469 Other features of *M. circinelloides* mosaic centromeres are
470 reminiscent of fungal regional centromeres like the ones described in
471 several *Candida* spp. *[39,40]*. The pericentric regions of *M. circinelloides*
472 are large (~15-75 kb), gene-free, and transcriptionally silenced sequences
473 that are interspersed by Grem-LINE1s, which are repeats of a LINE1-like
474 non-LTR retrotransposable element. Several retroelements have been
475 identified in the centromeres and neocentromeres of highly diverged
476 eukaryotes: non-LTR LINE1-like elements in mammals [41–46] and *Jockey*
477 elements in the fruit fly [47], and LTR retrotransposons in plants [48] and
478 fungi [23,49]. Based on these findings, a model has been proposed that
479 CENP-A is recruited by genomic sequences rich in repeats and
480 transposable elements and thus, gives rise to the centromeres [50]. The
481 Grem-LINE1s are arranged in a systematic pattern, oriented away from
482 the core centromeric regions and thus forming a large inverted repeat in
483 most of the centromeres. Similar organization and specific orientation of
484 inverted repeats also appears flanking the core centromeres of *S. pombe*
485 *[51,52]* and *Candida tropicalis [40]*, performing an important role in
486 centromere function [40,53]. *M. circinelloides* is the first organism to

39
40

487 display an enrichment in retrotransposable elements in their centromeric
488 regions in the absence of CENP-A, suggesting that these non-LTR
489 elements determine the centromere identity. Thus, we hypothesize that
490 identifying Grem-LINE1 clusters in other species of Mucoromycotina could
491 help mapping the location of the putative pericentromeric regions of the
492 corresponding species.

493 RNAi machinery is essential to control the mobility of transposable
494 elements, but also to maintain centromere structure and stability [23].
495 Fungi that have lost essential components of the RNAi machinery have
496 shorter pericentric regions and fewer transposable elements than closely
497 related RNAi-proficient species. *M. circinelloides* has a functional and
498 complex RNAi system with Dicer-dependent and -independent pathways
499 [24]. Our results demonstrate that the expression of the centromeric
500 Grem-LINE1s is suppressed by the Dicer-dependent RNAi pathway, which
501 may have resulted in accumulation of these elements in the centromeres
502 [23].

503 Overall this study has uncovered for the first time the centromeres
504 of an early-diverging fungus, which are characterized by a mosaic of
505 features from both point and regional centromeres found in the fungal
506 kingdom (Figure 6). Several studies on rapidly-evolving centromeres
507 suggest that point centromeres may be a recently evolved state as
508 compared to regional centromeres [9]. Therefore, we hypothesize that the
509 last mucoralean common ancestor had regional centromeres with
510 pericentric regions that harbored retrotransposable elements contained

511 by an active RNAi machinery. The loss of CENP-A triggered an
512 evolutionary rearrangement from typical regional centromeres to the
513 smaller kinetochore-binding regions that defines the mosaic centromeres
514 of *M. circinelloides*.

515 **Acknowledgments**

516 This work was supported by the Ministerio de Economía y
517 Competitividad, Spain (grant number BFU2015-65501-P, cofinanced by
518 FEDER, and RYC-2014-15844) and Ministerio de Ciencia, Innovación y
519 Universidades, Spain (grant number PGC2018-097452-B-I00, cofinanced
520 by FEDER). CPA and MINM were supported by predoctoral fellowships from
521 the Ministerio de Educación, Cultura y Deporte, Spain (grants number
522 FPU14/01983 and FPU14/01832 respectively). We acknowledge financial
523 support from Tata Innovation Fellowship, Department of Biotechnology,
524 Government of India (BT/HRD/35/01/03/2017), DBT-BUILDER programme
525 support at JNCASR (BT/INF22/SP27679/2018) and intramural support from
526 JNCASR to KS. SP is supported by DST-WOS-A postdoctoral fellowship (SR/
527 WOS-A/LS-207/2018). Supported in part by NIH/NIAID R37 grant AI39115-
528 21 and R01 grant AI50113-15 to JH, and by the CIFAR program Fungal
529 Kingdom: Threats & Opportunities, for which JH serves as Co-Director and
530 Fellow. We thank B. Suma at the confocal imaging facility of the Molecular
531 Biology and Genetics Unit, JNCASR for assistance in imaging. Work
532 conducted by the US Department of Energy Joint Genome Institute, a DOE
533 Office of Science User Facility, was supported by the Office of Science of
534 the US Department of Energy under Contract No. DE-AC02-05CH11231.

535 We thank Alexander Idnurm (University of Melbourne, Australia), Luis M.
536 Corrochano (University of Seville, Spain), Timothy Y. James (University of
537 Michigan, USA), Mathew E. Smith (University of Florida, USA), Joseph W.
538 Spatafora (Oregon State University, USA), Jason E. Stajich (University of
539 California-Riverside, USA), and Gregory Bonito (Michigan State University,
540 USA) for sharing fungal genome sequences to analyze the distribution of
541 kinetochore proteins. We also thank Thomas D. Petes (Duke University
542 School of Medicine, USA) and Beth A. Sullivan (Duke University School of
543 Medicine, USA) for critical reading of the manuscript and the members of
544 KS lab, VG lab, and JH lab for valuable discussions during bi-weekly Skype
545 meetings.

546 **Author Contributions**

547 KS, JH, VG, and FEN conceived and supervised the study. MINM and
548 CPA analyzed the distribution of kinetochore proteins in fungi, generated
549 and validated the plasmids and strains, performed the ChIP-seq
550 experiments, identified and defined the retrotransposable elements, and
551 analyzed the mRNA and sRNA-seq data. MINM, CPA and PG analyzed the
552 ChIP-seq data. MINM, CPA, PG and SP characterized the centromeric loci.
553 SP, MINM and CPA did the fluorescence microscopy imaging. SP performed
554 the ChIP-qPCR experiments. IVG, JP and SJM provided the genome
555 assembly and gene annotation. MINM, CPA and SP designed the figures
556 and wrote the manuscript with support from KS, JH, and VG. VG, JH, FEN
557 and KS provided funding. All authors revised the manuscript and approved
558 the final version.

559 **Declaration of Interest**

560 The authors declare no competing interests.

561 **Main-text figure legends**

562 **Figure 1. Distribution of the kinetochore complex across the**
563 **Mucoromycotina subphylum.** Concise kinetochore schematic showing
564 the most conserved protein complexes in eukaryotes is shown on the left
565 corresponding to the kinetochore proteins analyzed in the matrix on the
566 right. The matrix displays the presence or absence of 26 kinetochore
567 proteins across a cladogram of 51 fungal lineages (top) to show the
568 relationships among them. Every fungal phylum is represented and color-
569 coded, differentiating the Mucoromycota into the Glomeromycotina,
570 Mortierellomycotina, and Mucoromycotina subphyla to provide a special
571 emphasis on the latter. Arrows at divergence events in the cladogram
572 mark the hypothetical loss of proteins CENP-A (red) and CENP-C (blue) in
573 these clades. See also Figure S1 and Table S1.

574 **Figure 2. Inner and outer kinetochore proteins colocalize in *M.***
575 ***circinelloides* nuclei in the absence of CENP-A and CENP-C. (A)**
576 Graphical representation of Mis12, Dsn1, and CENP-T sequence features
577 showing individual Pfam domains PF05859, PF08202, PF15511
578 respectively identified in *M. circinelloides*. **(B)** Schematic to represent C-
579 terminal tagging of histone H3 and kinetochore proteins with mCherry
580 (red circles), eGFP (green circles). **(C-F)** Confocal microscope imaging
581 showing the cellular localization of well-conserved mCherry-tagged outer
582 kinetochore proteins Mis12 **(C, E)** and Dsn1 **(D, F)**, eGFP-tagged histone

583 H3 **(C, D)**, and eGFP-tagged inner kinetochore CENP-T **(E, F)** in pre-
584 germinated spores of *M. circinelloides* strains expressing fluorescent
585 fusion proteins. **(G)** Time-lapsed confocal image displaying the cellular
586 localization of mCherry-tagged Mis12 and eGFP-tagged histone H3 in a
587 germinative tube sprouting from a spore. A discontinuous white perimeter
588 outlines the germinative tube in the fluorescence images, and a white
589 square indicates a nuclear division event. Cartoon for the zoomed image
590 represents the localization and signal intensity of Mis12-mCherry in the
591 nucleus. A calibrated scale (white bar) is provided for size comparison (5
592 µm) and mCherry fluorescent signal is colored as magenta in the merged
593 images in **C-G**. See also Figures S2 and S3, and Table S2.

594 **Figure 3. *M. circinelloides* displays nine centromeres. (A)** Putative
595 kinetochore-binding regions showing color-coded enrichment of
596 immunoprecipitated DNA (*IP* DNA) from Mis12 and Dsn1 mCherry-tagged
597 strains compared to their corresponding input (*Input* DNA) and binding
598 controls (*Beads only*). 1 kb flanking sequences from the center of the
599 enrichment peak are shown. Major (M) and minor (m) kinetochore-binding
600 regions are indicated. Each kinetochore-protein data was obtained from a
601 pool of duplicated IP DNA samples. **(B)** Scaled ideogram of the nine
602 scaffolds (white) of *M. circinelloides* genome assembly that contain a
603 kinetochore-binding region (blue), showing the telomeric repeats (red).
604 See also Figure S4 and Table S3.

605 **Figure 4. The centromeres of *M. circinelloides* are short, AT rich**
606 **and harbor a conserved DNA motif. (A)** Size distribution of the nine

607 core centromere (CC) regions. **(B)** GC content across the nine core
608 centromere (CC) sequences. The median (red line) and standard deviation
609 (black lines) are shown in **A** and **B**. **(C)** 41 bp centromere-specific DNA
610 motif conserved in all nine CC regions and absent in the remainder of the
611 genome. A score of 2 bits in the logo indicates that the nucleotide is
612 present at that position in all nine CC regions. **(D)** A genomic view of
613 *CEN11* illustrating all nine core centromere regions. Kinetochore-binding
614 region enrichment (IP coverage) as the average of both
615 immunoprecipitation signals minus *Input* and *beads only* controls (left
616 axis, blue), GC content across the region (right axis, black), and the
617 centromere-specific DNA motif described in **e** (red) are shown. See also
618 Figure S4 and Table S3.

619 **Figure 5. *M. circinelloides* core centromeres are present in large**
620 **ORF-free pericentric regions having retrotransposable elements**
621 **regulated by the Dicer-dependent RNAi pathway and bind H3**
622 **nucleosomes. (A)** Genomic sequences lacking annotated genes (light
623 blue) that flank the *CC* regions (dark blue) served as the reference points
624 to calculate the length both upstream and downstream. Black circles at
625 either end of the regions indicate an abrupt, non-telomeric end of scaffold.
626 **(B)** Schematic of the Grem-LINE1 interspersed in the pericentromeric
627 regions. Open Reading Frames (ORF) and protein domains predicted from
628 their coding sequences are shown as colored boxes [Zf, zinc finger
629 (PF00098 and PF16588); AP, AP endonuclease (PTHR22748); RVT, reverse
630 transcriptase (PF00078); and zf-RVT, zinc-binding in reverse transcriptase
631 (PF13966)]. **(C)** *CEN4* is exemplified by low GC content, lack of genes and

632 transcripts, which are shown in different tracks displaying the kinetochore-
633 binding region enrichment (IP, an average of both *IP* DNA signals minus
634 *Input* and *beads only* controls), annotation of genes (red blocks) and
635 transposable elements (light blue blocks), *CEN*-specific DNA motif position
636 (vertical line) and direction (arrow), GC content, and transcriptomic data
637 of mRNA (green) and sRNAs (red) in *M. circinelloides* wild-type, *ago1,* and
638 double *dcl1 dcl2* deletion mutant strains after 48 h of growth in rich
639 media. **(D)** Genomic location of *CEN2* showing the regions studied for
640 histone H3 occupancy as labelled and colored rectangles. ChIP assays
641 were performed using polyclonal antibodies against histone H3. Primers
642 were designed for *CC2* region, pericentric regions (1L, 2L, 1R, 2R),
643 flanking ORFs (ORF-L, ORF-R) and a far-*CEN* control ORF ~2 Mb away from
644 *CEN2* (Far-*CEN*). *IP* samples were analyzed by real-time PCR using these
645 primers. The *y*-axis denotes the qPCR value as a percentage of the total
646 chromatin input with standard error mean (SEM), from each region tested.
647 The experiment was repeated three times with similar results. See also
648 Figures S5 and S6, and Tables S2, S3 and S4.

649 **Figure 6**. ***M. circinelloides* possesses a mosaic of point and**
650 **regional centromeres.** Length plot of the core centromere (kinetochore-
651 bound region, blue circles) and the full centromere (defined as the region
652 required for centromere function, red dots) of representative fungi
653 belonging to the Mucoromycota (green panel), Ascomycota (blue panel)
654 and Basidiomycota (purple panel). The size axis is divided in short (≤1 kb,
655 gray area) and long regions (>1 kb, brown area). The centromeres of *M.*
656 *circinelloides* (CEN5), *S. cerevisiae* (CEN3 [7]), *Candida albicans* (CEN7

657 [39]), *C. tropicalis* (Scnt 3 [40]), *Magnaporthe oryzae* Guy11 (CEN2 [49]),
658 *S. pombe* (CEN2 [52]), and *C. neoformans* (CEN10 [23]) are used for
659 comparison and classified as point, regional and mosaic centromeres
660 according to their structural features. In organisms where functionality of
661 the centromere regions could not be experimentally determined, the ORF-
662 free region spanning the kinetochore binding region is considered as the
663 full centromere. Line diagrams represent the structural features of each
664 centromere (drawn to scale). The ORF-free region at the centromere is
665 defined by black vertical boundaries and flanking ORFs are shown as red
666 arrows. Core centromeres (rectangles); kinetochore-bound regions (blue
667 area); centromere-specific DNA motifs next to an AT-rich core (green
668 lines); Grem-LINE1, Grasshopper, GYMAG, Maggy, Tcn retrotransposons
669 (black arrows); inverted repeats (purple arrows); and other DNA repeats
670 (light and dark orange arrows) are shown. Repeats/retrotransposons sizes
671 are not drawn to scale. CC, core centromere; CDE, centromere DNA
672 elements; LR, left repeat; RR, right repeat; otr, outer repeats containing
673 dg/dh elements; imr, innermost repeats; cnt, central core.

674 **STAR Methods**

675 Lead contact and materials availability

676 Further information and requests for resources and reagents should
677 be directed to and will be fulfilled by the Lead Contact, Victoriano Garre
678 (vgarre@um.es).

679 Experimental model and subject details

680 **Fungal strains and culture conditions**

681 All the strains used and generated in this work derive from *M.*
682 *circinelloides* f. *lusitanicus* CBS277.49. The double auxotroph MU402 [14]
683 (Ura^-, Leu^-) or the MU636 (Ura^+, Leu^-) strain was used for transformation
684 with cassettes containing the selectable marker for uracil *pyrG* or for
685 leucine *leuA*, respectively. Single tagged strains in kinetochore proteins
686 MU840, MU842, MU844, and MU846 were used as parental strains for
687 transformation with the Hht4-eGFP cassette to obtain double-tagged
688 strains. All the generated strains in this work are listed in the Key
689 Resources Table.

690 After transformation, to select uracil prototrophy *M. circinelloides*
691 spores were plated in minimal medium with Casamino Acids (MMC) [14],
692 and for leucine prototrophy spores were plated in medium yeast nitrogen
693 base (YNB) [14], both adjusted at pH 3.2 for transformation and colony
694 isolation. *M. circinelloides* spores were cultured in rich medium yeast-
695 peptone-glucose (YPG)[14] for optimal growth and sporulation at pH 4.5.
696 The microscopy and ChIP assays were performed with pre-germinated
697 spores growing in YPG pH 4.5 for 3 hours at 250 rpm and 26 ºC.

698 Method details

699 **Construction of the fluorescent strains**

700 Alleles containing fluorescent fusion protein coding-sequences were
701 designed to either integrate them in their corresponding endogenous
702 locus, or in the *carRP* locus. The Hht4-eGFP, Hhf1-eGFP, Hhf4-eGFP and

703 CENP-T-eGFP alleles contained the ORF of each gene fused in-frame at
704 their C-termini with the eGFP sequence, followed by the *leuA* gene as a
705 selectable marker. The alleles were obtained by overlapping PCR using 5'-
706 modified primers harboring restriction sites for easy cloning. Briefly, each
707 allele was obtained by fusing the 1-kb fragment upstream the stop codon
708 of each gene, the 0.7-kb eGFP fragment (ending in a stop codon), the 3.4-
709 kb *leuA* fragment, and the 1-kb sequence downstream the stop codon of
710 each tagged gene. All the linear alleles were digested with appropriate
711 restriction enzymes, ligated into the plasmid Bluescript SK (+), and cloned
712 into *Escherichia coli*.

713 The Mis12-mCherry and Dsn1-mCherry constructions were designed
714 to allow the integration and ectopic overexpression of the fused proteins
715 in the *carRP* locus. First, a universal mCherry-expression vector was
716 constructed, containing the strong promoter *Pzrt1*, the ORF of the
717 mCherry coding-gene including the stop codon, and the *pyrG* gene as a
718 selectable marker (named pMAT1915). To do that, a fragment containing
719 the 1-kb downstream and upstream sequences from the *carRP* gene
720 flanking the *Pzrt1* promoter was amplified with an inverse PCR using the
721 plasmid pMAT1477 [14] as a template. Then, the *mCherry* and *pyrG*
722 fragments were added by In-Fusion cloning (Takara) using primers with
723 appropriate overlapping 5'-tails (Table S2). Once the universal pMAT1915
724 vector was constructed, it was used for the C-terminal tagging of *mis12* or
725 *dsn1*, fusing their ORFs without the stop codon upstream and in-frame
726 with the *mCherry* ORF by In-Fusion cloning. The resulting plasmids were
727 pMAT1917 and pMAT1919, respectively for *mis12* and *dsn1.* For the N-

728 terminal tagging, the plasmid pMAT1915 was inverse-amplified using
729 primers that exclude the stop codon of the *mCherry* fragment, and the
730 *mis12* and *dsn1* ORFs, including their stop codons, were cloned in-frame.
731 The resulting plasmids were pMAT1916 and pMAT1918 for mCherry-Mis12
732 and mCherry-Dsn1 expression, respectively.

733 Stellar™ competent cells (Takara) were used for transformation by
734 heat-shock following the supplier procedures. The correct construction of
735 all the generated plasmids was confirmed by restriction analysis and
736 Sanger sequencing.

737 For transformation, the alleles were released from the vector
738 backbone by restriction digestion and used to transform *M. circinelloides*
739 protoplasts by electroporation [14]. Spores from individual colonies were
740 collected and plated in selective media to ensure homokaryosis. After at
741 least five vegetative cycles, the integration in the target locus was
742 confirmed by PCR using primers that amplified the whole locus (Table S2),
743 producing PCR-fragments that allowed discrimination between the mutant
744 and the wild-type locus (Figures S3). Genomic DNA purification was
745 performed following the procedure previously published [54].

746 **Microscopy imaging and acquisition**

747 Freshly isolated *M. circinelloides* spores were washed twice with 1X
748 phosphate buffered saline (PBS). The spores were resuspended in PBS and
749 10 µl suspension was placed on a slide, covered with a coverslip and
750 sealed with clear nail polish. The images were acquired at room
751 temperature using laser scanning inverted confocal microscope LSM 880-

752 Airyscan (Zeiss, Plan Apochromat 63x, NA oil 1.4) equipped with highly
753 sensitive photo-detectors. The filters used were GFP/FITC 488, mCherry
754 561 for excitation and GFP/FITC 500/550 band pass, mCherry 565/650
755 long pass for emission. Z- stack images were taken at every 0.5 µm and
756 processed using Zeiss Zen software. All the images were displayed after
757 the maximum intensity projection of images using Zeiss Zen software.

758 **Live cell imaging**

759 Freshly isolated *M. circinelloides* spores were incubated in YPG
760 growth medium at 26ºC for 2 h. Then, the spores to be imaged were
761 pelleted at 3000 xg and washed twice with 1X phosphate buffered saline
762 (PBS). The spores were resuspended in PBS and 10µl suspension was
763 placed on a slide containing thin 2% agarose patch containing 2%
764 dextrose and covered with a coverslip. Live cell imaging was performed at
765 30ºC in a temperature-controlled chamber of an inverted confocal
766 microscope (Zeiss, LSM-880) using a Plan Apochromat 100X NA oil 1.4
767 objective and GaAsp photodetectors. Images were collected at 60-s
768 intervals with 0.5 µm Z-steps using GFP/FITC 488, mCherry 561 for
769 excitation. GFP/FITC 500/550 band pass or GFP 495/500 and mCherry
770 565/650 or mCherry 580-750 band pass for emission. All the images were
771 displayed after the maximum intensity projection of images at each time
772 using Zeiss Zen software.

773 **Ortholog search**

774 Seventy five fungal proteomes were retrieved from the Joint
775 Genome Institute (JGI) Mycocosm genome portal [55] (Table S1). BLAST+

776 [56] v2.7.1 individual protein-protein BLASTp, iterative BLASTp (PSI-
777 BLAST) and iterative HMMER v3.2.1 (http://www.hmmer.org) jackhammer
778 searches (E-value ≤ $1x10^{-5}$) were conducted against these proteomes
779 using *M. musculus*, *U. maydis*, *S. cerevisiae*, and *S. pombe* protein
780 sequences (Table S1) as queries. In addition, an hmmsearch was launched
781 using the Pfam-A [57] HMM models for each kinetochore protein [Pfam
782 gathering threshold (GA) as cut-off value]. All the putative orthologs found
783 were used to perform a reciprocal BLASTp against *M. musculus*, *U.*
784 *maydis*, *S. cerevisiae*, and *S. pombe* protein sequences. Also, Pfam-A
785 protein domains were searched in all the possible orthologs using HMMER
786 hmmscan. Protein sequences that either lacked appropriate Pfam domains
787 or failed to produce a hit in a reciprocal blastp search were discarded,
788 resulting in a matrix of putative orthologs for each given species (Table
789 S1). The cladogram showing the relationship among species was drawn
790 using the tree data from JGI Mycocosm [55] and the interactive Tree of
791 Life[58] (iTOL v4) tool. To determine CENP-A presence or absence in each
792 given species, all 289 protein sequences of histone H3 putative orthologs
793 (Table S1) were aligned using MUSCLE[59] v3.8.1551, and the
794 phylogenetic relationship of their HFD was inferred by the neighbor-joining
795 method employing the Jones, Taylor, and Thornton (JTT) substitution
796 model [60], with a bootstrap procedure of 1,000 iterations (MEGA[61]
797 v10.0.5).

798 **Chromatin immunoprecipitation and sequencing**

799 The ChIP protocol for *M. circinelloides* was adapted from previous
800 studies in other fungal models [23]. Briefly, spores from the strains MU842
801 (expressing Mi12-mCherry) and MU846 (expressing Dsn1-mCherry) were
802 used for independent IP experiments. 10^7 spores/mL of these strains were
803 germinated in 100 ml of liquid YPG pH 4.5 medium, shaking at 250 rpm at
804 26ºC for 3 hours. Then, cells were fixed by adding formaldehyde at a final
805 concentration of 1% for 20 minutes. The fixation was quenched by adding
806 glycine to a final concentration of 135 mM. Cells were pelleted and ground
807 with liquid nitrogen, and 150 mg of ground pellet was resuspended in ChIP
808 lysis Buffer (50 mM HEPES pH 7.4, 150 mM NaCl, 1% Triton X-100, 0.1%
809 DOC, and 1 mM EDTA), with a protease inhibitor cocktail. The sample was
810 sonicated using a Bioruptor (Diogenode) with pulses of 30s ON/OFF for 50
811 cycles to obtain a chromatin fragmentation between 100 and 300 bp.

812 The lysates were clarified by centrifugation at 12,000 xg for 10 min
813 at 4ºC. At this step, 100 µL were frozen at -20ºC for Input DNA control.
814 The samples were then divided into 450 µL for the antibody
815 immunoprecipitation and 450 µL for the mock control. For *IP* 25 µL of Red
816 Fluorescent Protein (RFP)-Trap Magnetic Agarose (MA) beads (Chromotek)
817 were added and 25 µL of MA beads (Chromotek) to the binding control.
818 Samples were incubated at 4ºC overnight. After incubation, the beads
819 were washed twice with 1 mL of the low salt wash buffer (20 mM Tris pH
820 8.0, 150 mM NaCl, 0.1% SDS, 1 % Triton X-100, and 2 mM EDTA), twice
821 with 1 mL of the high salt wash buffer (20 mM Tris pH 8.0, 500 mM NaCl,
822 0.1% SDS, 1% Triton X-100 and 2 mM EDTA), once with the LiCl wash
823 buffer (10 mM Tris pH 8.0, 1 mM EDTA, 0.25 M LiCl, 1% NP40, and 1%

824 DOC), and finally once with TE (10mM Tris pH 8.0, and 1 mM EDTA). To
825 elute the DNA, 250 µL of TES (50 mM Tris pH 8.0, 10 mM EDTA, 1% SDS)
826 was added to each sample and incubated for 10 min at 65ºC, twice. The
827 immunoprecipitated samples and the Input controls were de-crosslinked
828 by incubating at 65ºC overnight. The samples were treated with 199 µg of
829 RNase A (Sigma Aldrich) and 190 µg of Proteinase K (Sigma Aldrich) for 2
830 hours at 50ºC. For DNA purification, 1 volume of
831 phenol:chloroform:isoamyl alcohol (25:24:1) was added, the aqueous
832 phase was recovered, and then 1 volume of chloroform:isoamyl alcohol
833 (24:1) was added. After centrifugation, the aqueous phase was recovered,
834 and DNA was precipitated with 20 µg of glycogen (Thermo Fisher), 1/10
835 volume of Na-Acetate 3 M (pH 5.2), and 1 volume of ethanol 100%. Pulled
836 DNA from duplicated samples were precipitated together for sequencing.
837 Samples incubated at -20ºC overnight were centrifuged for 10 min at
838 12,000 xg and the pellets were washed with ethanol 70%. Each sample
839 was resuspended in 30 µL of MilliQ water. The quality of the DNA was
840 analyzed by QuBit (Thermo Fisher) and sent to Novogene, which prepared
841 the libraries using TruSeq ChIP Library Prep Kit and sequenced them with
842 Illumina Hiseq 2500 High-Output v4 single-end reads of 50 bp.

843 Histone H3-ChIP assays were performed in the same way with the
844 following changes. The protoplasts were obtained as described before [14]
845 and then sonicated as above. Native ChIP was performed [23] using anti-
846 H3 antibody (Abcam ab1791). The DNA pellet was dissolved in 25 µl of
847 MilliQ water. All samples (*Input*, *IP* with or without antibodies) were used
848 for PCR. The *Input* and *IP* DNA was subsequently used for qPCR using the

849 primers listed in Table S2 and SensiFAST™ SYBR® No ROX Kit. Histone H3
850 enrichment was determined by the percentage input method using the
851 formula: $100\,x\,2^{adjusted\,Ct\,Input-adjusted\,Ct\,IP}$. The adjusted Ct value is the dilution
852 factor (log_2 of dilution factor) subtracted from the Ct value of the *Input* or
853 *IP* DNA samples. Three technical replicates were taken for qPCR analysis
854 and standard error of mean was calculated.

855 **Sequencing data analysis**

856 Raw reads from Chromatin Immunoprecipitation and RNA
857 sequencing (ChIP- and RNA-seq) were quality-checked with FASTQC
858 v0.11.8 and adapters removed with Trim Galore! v0.6.2
859 (http://www.bioinformatics.babraham.ac.uk/projects/). Reads were aligned
860 to the *Mucor circinelloides* f. *lusitanicus* MU402 reference genome
861 Muccir1_3 (referred to as the *M. circinelloides* genome throughout the
862 text, available at JGI
863 http://genome.jgi.doe.gov/Muccir1_3/Muccir1_3.home.html), using the
864 Burrows-Wheeler Aligner [62] (BWA v0.7.8) algorithm for ChIP- and sRNA-
865 seq reads and STAR [63] v2.7.1a for mRNA-seq data. The number of
866 overlapping aligned reads per 25-bp bin was used as a measure of
867 coverage, obtaining normalized bigWig coverage files with deepTools [64]
868 v3.2.1 bamCoverage function. For RNA-seq data, coverage was
869 normalized to bins per million mapped reads (BPM). For ChIP-seq data,
870 reads were normalized per genomic content (1x normalization). ChIP
871 enrichment peaks were identified by Model-based Analysis of ChIP-seq
872 [65] (MACS2 v2.1.1) callpeak function and sorted by fold-enrichment and

873 FDR values (Table S3). Enriched peaks (fold-enrichment ≥ 1.6, FDR ≤
874 $5x10^{-5}$) were manually curated by analyzing the *IP* coverage with the
875 Integrative Genomics Viewer [66] (IGV v2.4.1) genome browser, ensuring
876 that peaks were present in both Mis12 and Dsn1 *IP* samples, and absent in
877 either the input and binding controls. The resulting centromeric sequences
878 were analyzed with Multiple Em for Motif Elicitation [67] (MEME v5.0.5) to
879 discover conserved DNA motifs on both strands. GC content every 25-bp
880 bin was calculated with bedTools [68] v2.28. Normalized coverage, GC
881 content and genetic elements annotation files were rendered with either
882 Gviz [69] v1.29.0 for whole-chromosome visualization or deepTools
883 pyGenomeTracks module for close-in genomic views.

884 **Transposable element analysis**

885 The *M. circinelloides* genome assembly of strain MU402 was
886 searched for repeats using RepeatScout [70] v1.0.5 and RepeatMasker
887 (http://www.repeatmasker.org) v4.0.9. The repeats located in the
888 pericentromeric regions were aligned using MAFFT [71] v7.407 and
889 clustered with CD-HIT [72] v4.8.1, generating groups of sequences that
890 shared ≥ 95% identity and ≥ 70% coverage and calculating their *p*-
891 distance (nucleotide changes per site). A BLASTn search against *M.*
892 *circinelloides* genome assembly was conducted using a representative
893 repeated sequence from each cluster to locate similar sequences across
894 the entire genome, obtaining hits with ≥ 80% identity and ≥ 20%
895 coverage (Table S4). The pericentromeric repeats were examined with
896 getorf from Emboss [73] v6.6.0, obtaining all ORFs ≥ 600 nt. Protein

897 domains in these ORFs translated sequences were predicted by
898 hmmsearch using the Pfam-A v32.0 database. The reverse transcriptase
899 (RVT) domains found in the repeated sequences were aligned with
900 MUSCLE to create a consensus sequence, and this consensus sequence
901 was aligned with 281 RVT domains of well-known non-LTR
902 retrotransposons from Repbase [74] 24.04. A neighbor-joining
903 phylogenetic tree was inferred from this alignment using the JTT
904 substitution model and a bootstrap procedure of 1000 iterations (MEGA
905 v10.0.5). To assess the presence of the retrotransposable elements in
906 other species of Mucoromycotina, a tBLASTn search was launched against
907 the 55 available genomes at JGI (Table S5) using the translated sequence
908 of both ORFs combined, recovering hits with ≥ 50% coverage and E-value
909 $\leq 10^{-5}$.

910 Quantification and statistical analysis

911 Statistical details of experiments are detailed in the Results, Figures
912 and Figure legends, including the number of biological and technical
913 replicates, the definition of center, dispersion and precision measures
914 (mean, SEM and SD). Statistical analyses were performed with GraphPad
915 Prism 6.

916 Data and code availability

917 ChIP-seq raw data and processed files are deposited at the Gene
918 Expression Omnibus (GEO) repository under GSE132687 accession
919 number and are publicly available. PacBio raw data and assembly of the

920 genome of *M. circinelloides* f. *lusitanicus* MU402 are publicly available at
921 the JGI website
922 (http://genome.jgi.doe.gov/Muccir1_3/Muccir1_3.download.html) and
923 subjected to the JGI Data Usage Policy. Both mRNA and sRNA-seq data
924 were already available at the Sequence Read Archive (SRA) and can be
925 accessed with the following run accession numbers: SRR1611144 (R7B
926 wild-type strain mRNA) [75], SRR1611151 (*ago1* deletion mutant strain
927 mRNA) [75], SRR1611171 (double *dcl1 dcl2* deletion mutant strain mRNA)
928 [75], SRR039123 (R7B wild-type strain sRNA) [76], SRR836082 (*ago1*
929 deletion mutant strain sRNA) [77], SRR039128 (double *dcl1 dcl2* deletion
930 mutant strain sRNA) [76]. The protein sequences of the kinetochore
931 orthologs identified in this study are accessible at Mendeley data (DOI:
932 10.17632/wpyyb58h8v.1), as well as all the main and supplementary
933 material.

934 **Legends for supplemental Excel tables and videos**

935 **Video S1. *M. circinelloides* displays asynchronous nuclear**
936 **division. Related to Figure 2.** A pre-germinated spore of *M.*
937 *circinelloides* showing the distribution of histone H3-eGFP fusion protein
938 encompassing the nuclei. Red arrows mark single events of nuclear
939 division.

940 **Table S1. Kinetochore proteins distribution across fungal**
941 **lineages. Related to Figure 1.** Orthologs identified in each species
942 analyzed, showing their correspondent ID in the JGI platform, Pfam

943 domains (name, ID, coordinates within the translated sequence, e-value
944 and score).

945 **Table S3. Centromere location and Mis12 and Dsn1 IP/Input**
946 **peaks. Related to Figures 3, 4, 5 and S4. (A)** Genomic coordinates of
947 the core centromere (CC) defined by the both Mis12 and Dsn1 kinetochore
948 proteins binding regions; and of the entire centromere, including the
949 upstream and downstream pericentric regions. (**B, C**) MACS2 called peaks
950 for Mis12 IP/Input (**B**) and Dsn1 IP/Input (**C**). Rows colored red are above
951 established threshold (FDR $\leq 5\times10^{-7}$, fold enrichment $\geq$ 1.6).

952 **Table S4. Centromeric Grem-LINE1s domain architecture. Related**
953 **to Figures 5, S5 and S6.** Protein domains [name and ID (Pfam or
954 Panther), start and end in the translated product, as well as predicted e-
955 value, score, and bias] for the ORFs [numbered (#) and showing start and
956 end in the DNA sequence] found in each Grem-LINE1 (showing genomic
957 location and length).

958 **Table S5. Grem-LINE1s distribution across Mucoromycotina.**
959 **Related to Figures 1 and S6.** Genomic location of Grem-LINE1s across
960 species of Mucoromycotina showing the percentage of identity compared
961 to a representative of *M. circinelloides* Grem-LINE1s.

962 **References**

963 1. McKinley, K.L., and Cheeseman, I.M. (2016). The molecular basis for
964 centromere identity and function. Nat. Rev. Mol. Cell Biol. *17*, 16–29.
965 2. Cheeseman, I.M. (2014). The kinetochore. Cold Spring Harb.

966 Perspect. Biol. *6*, a015826.

967 3. Marshall, O.J., Chueh, A.C., Wong, L.H., and Choo, K.H.A. (2008).
968 Neocentromeres: new insights into centromere structure, disease
969 development, and karyotype evolution. Am. J. Hum. Genet. *82*, 261–
970 82.

971 4. Earnshaw, W.C., and Migeon, B.R. (1985). Three related centromere
972 proteins are absent from the inactive centromere of a stable
973 isodicentric chromosome. Chromosoma *92*, 290–6.

974 5. Henikoff, S. (2001). The centromere paradox: stable inheritance with
975 rapidly evolving DNA. Science *293*, 1098–1102.

976 6. Clarke, L., and Carbon, J. (1980). Isolation of a yeast centromere and
977 construction of functional small circular chromosomes. Nature *287*,
978 504–9.

979 7. Clarke, L., and Carbon, J. (1985). The structure and function of yeast
980 centromeres. Annu. Rev. Genet. *19*, 29–55.

981 8. Friedman, S., and Freitag, M. (2017). Centrochromatin of fungi. In
982 Centromeres and kinetochores. Progress in Molecular and Subcellular
983 Biology, B. Black, ed. (Springer, Cham), pp. 85–109.

984 9. Malik, H.S., and Henikoff, S. (2009). Major evolutionary transitions in
985 centromere complexity. Cell *138*, 1067–1082.

986 10. Drinnenberg, I.A., DeYoung, D., Henikoff, S., and Malik, H.S. (2014).
987 Recurrent loss of CenH3 is associated with independent transitions to
988 holocentricity in insects. eLife *3*, 717–750.

989 11. Steiner, F.A., and Henikoff, S. (2014). Holocentromeres are dispersed
990 point centromeres localized at transcription factor hotspots. eLife *3*,

991 e02025.

992 12. Akiyoshi, B., and Gull, K. (2014). Discovery of unconventional
993 kinetochores in kinetoplastids. Cell *156*, 1247–1258.

994 13. van Hooff, J.J., Tromer, E., van Wijk, L.M., Snel, B., and Kops, G.J.
995 (2017). Evolutionary dynamics of the kinetochore network in
996 eukaryotes as revealed by comparative genomics. EMBO Rep. *18*,
997 e201744102.

998 14. Nicolás, F.E., Navarro-Mendoza, M.I., Pérez-Arques, C., López-García,
999 S., Navarro, E., Torres-Martínez, S., and Garre, V. (2018). Molecular
1000 tools for carotenogenesis analysis in the mucoral *Mucor*
1001 *circinelloides*. In Microbial Carotenoids. Methods in Molecular Biology,
1002 C. Barreiro and J. L. Barredo, eds. (Humana Press, New York, NY).

1003 15. Prakash, H., and Chakrabarti, A. (2019). Global epidemiology of
1004 mucormycosis. J. Fungi *5*, E26.

1005 16. Malik, H.S., and Henikoff, S. (2003). Phylogenomics of the
1006 nucleosome. Nat. Struct. Biol. *10*, 882–891.

1007 17. Osley, M. (1991). The regulation of histone synthesis in the cell cycle.
1008 Annu. Rev. Biochem. *60*, 827–861.

1009 18. Yun, C.S., and Nishida, H. (2011). Distribution of introns in fungal
1010 histone genes. PLoS One *6*, 4–7.

1011 19. Luense, L.J., Wang, X., Schon, S.B., Weller, A.H., Lin Shiao, E., Bryant,
1012 J.M., Bartolomei, M.S., Coutifaris, C., Garcia, B.A., and Berger, S.L.
1013 (2016). Comprehensive analysis of histone post-translational
1014 modifications in mouse and human male germ cells. Epigenetics
1015 Chromatin *9*, 24.

1016 20. Schipper, M.A.A. (1973). A study on variability in *Mucor hyemalis* and
1017 related species. Stud. Mycol. *4*, 1–40.
1018 21. Li, C.H., Cervantes, M., Springer, D.J., Boekhout, T., Ruiz-Vazquez,
1019 R.M., Torres-Martinez, S.R., Heitman, J., and Lee, S.C. (2011).
1020 Sporangiospore size dimorphism is linked to virulence of *Mucor*
1021 *circinelloides*. PLoS Pathog. *7*, e1002086.
1022 22. Smith, K.M., Phatale, P.A., Sullivan, C.M., Pomraning, K.R., and
1023 Freitag, M. (2011). Heterochromatin is required for normal
1024 distribution of *Neurospora crassa* CenH3. Mol. Cell. Biol. *31*, 2528–42.
1025 23. Yadav, V., Sun, S., Billmyre, R.B., Thimmappa, B.C., Shea, T., Lintner,
1026 R., Bakkeren, G., Cuomo, C.A., Heitman, J., and Sanyal, K. (2018).
1027 RNAi is a critical determinant of centromere evolution in closely
1028 related fungi. Proc. Natl. Acad. Sci. *115*, 3108–3113.
1029 24. Torres-Martínez, S., and Ruiz-Vázquez, R.M. (2016). RNAi pathways in
1030 *Mucor*: A tale of proteins, small RNAs and functional diversity. Fungal
1031 Genet. Biol. *90*, 44–52.
1032 25. Mizuguchi, G., Xiao, H., Wisniewski, J., Smith, M.M., and Wu, C.
1033 (2007). Nonhistone Scm3 and histones CenH3-H4 assemble the core
1034 of centromere-specific nucleosomes. Cell *129*, 1153–1164.
1035 26. Thakur, J., Talbert, P.B., and Henikoff, S. (2015). Inner kinetochore
1036 protein interactions with regional centromeres of fission yeast.
1037 Genetics *201*, 543–61.
1038 27. Kapoor, S., Zhu, L., Froyd, C., Liu, T., and Rusche, L.N. (2015).
1039 Regional centromeres in the yeast *Candida lusitaniae* lack
1040 pericentromeric heterochromatin. Proc. Natl. Acad. Sci. *112*, 12139–

1041 12144.

1042 28. Calo, S., Shertz-Wall, C., Lee, S.C., Bastidas, R.J., Nicolás, F.E.,
1043 Granek, J.A., Mieczkowski, P., Torres-Martínez, S., Ruiz-Vázquez, R.M.,
1044 Cardenas, M.E., *et al.* (2014). Antifungal drug resistance evoked via
1045 RNAi-dependent epimutations. Nature *513*, 555–558.

1046 29. Caramalho, R., Tyndall, J.D.A., Monk, B.C., Larentis, T., Lass-Flörl, C.,
1047 and Lackner, M. (2017). Intrinsic short-tailed azole resistance in
1048 mucormycetes is due to an evolutionary conserved aminoacid
1049 substitution of the lanosterol 14α-demethylase. Sci. Rep. *7*, 3–12.

1050 30. Andrianaki, A.M., Kyrmizi, I., Thanopoulou, K., Baldin, C., Drakos, E.,
1051 Soliman, S.S.M., Shetty, A.C., McCracken, C., Akoumianaki, T.,
1052 Stylianou, K., *et al.* (2018). Iron restriction inside macrophages
1053 regulates pulmonary host defense against *Rhizopus* species. Nat.
1054 Commun. *9*, 3333.

1055 31. Pérez-Arques, C., Navarro-Mendoza, M.I., Murcia, L., Lax, C.,
1056 Martínez-García, P., Heitman, J., Nicolás, F.E., and Garre, V. (2019).
1057 *Mucor circinelloides* thrives inside the phagosome through an Atf-
1058 mediated germination pathway. mBio *10*, 1–15.

1059 32. Gladfelter, A.S., Hungerbuehler, A.K., and Philippsen, P. (2006).
1060 Asynchronous nuclear division cycles in multinucleated cells. J. Cell
1061 Biol. *172*, 347–62.

1062 33. Liu, X., McLeod, I., Anderson, S., Yates, J.R., and He, X. (2005).
1063 Molecular analysis of kinetochore architecture in fission yeast. EMBO
1064 J. *24*, 2919–30.

1065 34. Schotanus, K., Soyer, J.L., Connolly, L.R., Grandaubert, J., Happel, P.,

1066 Smith, K.M., Freitag, M., and Stukenbrock, E.H. (2015). Histone
1067 modifications rather than the novel regional centromeres of
1068 *Zymoseptoria tritici* distinguish core and accessory chromosomes.
1069 Epigenetics Chromatin *8*, 41.
1070 35. Kozubowski, L., Yadav, V., Chatterjee, G., Sridhar, S., Yamaguchi, M.,
1071 Kawamoto, S., Bose, I., Heitman, J., and Sanyal, K. (2013). Ordered
1072 kinetochore assembly in the human-pathogenic basidiomycetous
1073 yeast Cryptococcus neoformans. mBio *4*, e00614-13.
1074 36. Hemmerich, P., Stoyan, T., Wieland, G., Koch, M., Lechner, J., and
1075 Diekmann, S. (2000). Interaction of yeast kinetochore proteins with
1076 centromere-protein/transcription factor Cbf1. Proc. Natl. Acad. Sci.
1077 *97*, 12583–12588.
1078 37. Sorger, P.K., Doheny, K.F., Hieter, P., Kopski, K.M., Huffaker, T.C., and
1079 Hyman, A.A. (1995). Two genes required for the binding of an
1080 essential *Saccharomyces cerevisiae* kinetochore complex to DNA.
1081 Proc. Natl. Acad. Sci. *92*, 12026–12030.
1082 38. Pietrasanta, L.I., Thrower, D., Hsieh, W., Rao, S., Stemmann, O.,
1083 Lechner, J., Carbon, J., and Hansma, H. (1999). Probing the
1084 *Saccharomyces cerevisiae* centromeric DNA (CEN DNA)-binding
1085 factor 3 (CBF3) kinetochore complex by using atomic force
1086 microscopy. Proc. Natl. Acad. Sci. *96*, 3757–3762.
1087 39. Sanyal, K., Baum, M., and Carbon, J. (2004). Centromeric DNA
1088 sequences in the pathogenic yeast *Candida albicans* are all different
1089 and unique. Proc. Natl. Acad. Sci. *101*, 11374–11379.
1090 40. Chatterjee, G., Sankaranarayanan, S.R., Guin, K., Thattikota, Y.,

1091 Padmanabhan, S., Siddharthan, R., and Sanyal, K. (2016). Repeat-
1092 associated fission yeast-like regional centromeres in the
1093 ascomycetous budding yeast *Candida tropicalis*. PLoS Genet. *12*,
1094 e1005839.
1095 41. Miga, K.H., Newton, Y., Jain, M., Altemose, N., Willard, H.F., and Kent,
1096 W.J. (2014). Centromere reference models for human chromosomes
1097 X and Y satellite arrays. Genome Res. *24*, 697–707.
1098 42. Chueh, A.C., Northrop, E.L., Brettingham-Moore, K.H., Choo, K.H.A.,
1099 and Wong, L.H. (2009). LINE retrotransposon RNA is an essential
1100 structural and functional epigenetic component of a core
1101 neocentromeric chromatin. PLoS Genet. *5*, e1000354.
1102 43. Carbone, L., Harris, R.A., Mootnick, A.R., Milosavljevic, A., Martin,
1103 D.I.K., Rocchi, M., Capozzi, O., Archidiacono, N., Konkel, M.K., Walker,
1104 J.A., *et al.* (2012). Centromere remodeling in *Hoolock leuconedys*
1105 (Hylobatidae) by a new transposable element unique to the gibbons.
1106 Genome Biol. Evol. *4*, 648–658.
1107 44. de Sotero-Caio, C.G., Cabral-de-Mello, D.C., Calixto, M. da S., Valente,
1108 G.T., Martins, C., Loreto, V., de Souza, M.J., and Santos, N. (2017).
1109 Centromeric enrichment of LINE-1 retrotransposons and its
1110 significance for the chromosome evolution of Phyllostomid bats.
1111 Chromosom. Res. *25*, 313–325.
1112 45. Longo, M.S., Carone, D.M., Green, E.D., O’Neill, M.J., and O’Neill, R.J.
1113 (2009). Distinct retroelement classes define evolutionary breakpoints
1114 demarcating sites of evolutionary novelty. BMC Genomics *10*, 334.
1115 46. O’Neill, R.J.W., O’Neill, M.J., and Graves, J.A.M. (1998).

1116 Undermethylation associated with retroelement activation and
1117 chromosome remodelling in an interspecific mammalian hybrid.
1118 Nature *393*, 68–72.
1119 47. Chang, C.-H., Chavan, A., Palladino, J., Wei, X., Martins, N.M.C.,
1120 Santinello, B., Chen, C.-C., Erceg, J., Beliveau, B.J., Wu, C.-T., *et al.*
1121 (2019). Islands of retroelements are major components of *Drosophila*
1122 centromeres. PLoS Biol. *17*, e3000241.
1123 48. Presting, G.G. (2018). Centromeric retrotransposons and centromere
1124 function. Curr. Opin. Genet. Dev. *49*, 79–84.
1125 49. Yadav, V., Yang, F., Reza, M.H., Liu, S., Valent, B., Sanyal, K., and
1126 Naqvi, N.I. (2019). Cellular dynamics and genomic identity of
1127 centromeres in cereal blast fungus. mBio *10*, e01581-19.
1128 50. Klein, S.J., and O'Neill, R.J. (2018). Transposable elements: genome
1129 innovation, chromosome diversity, and centromere conflict.
1130 Chromosom. Res. *26*, 5–23.
1131 51. Steiner, N.C., Hahnenberger, K.M., and Clarke, L. (1993).
1132 Centromeres of the fission yeast Schizosaccharomyces pombe are
1133 highly variable genetic loci. Mol. Cell. Biol. *13*, 4578–87.
1134 52. Pidoux, A.L., and Allshire, R.C. (2004). Kinetochore and
1135 heterochromatin domains of the fission yeast centromere.
1136 Chromosom. Res. *12*, 521–534.
1137 53. Baum, M., Ngan, V.K., and Clarke, L. (1994). The centromeric K-type
1138 repeat and the central core are together sufficient to establish a
1139 functional *Schizosaccharomyces pombe* centromere. Mol. Biol. Cell *5*,
1140 747–61.

1141 54. Navarro-Mendoza, M.I., Pérez-Arques, C., Murcia, L., Martínez-García,
1142 P., Lax, C., Sanchis, M., Capilla, J., Nicolás, F.E., and Garre, V. (2018).
1143 Components of a new gene family of ferroxidases involved in
1144 virulence are functionally specialized in fungal dimorphism. Sci. Rep.
1145 *8*, 7660.
1146 55. Grigoriev, I. V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R.,
1147 Riley, R., Salamov, A., Zhao, X., Korzeniewski, F., *et al.* (2014).
1148 MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic
1149 Acids Res. *42*, D699–D704.
1150 56. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J.,
1151 Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and
1152 applications. BMC Bioinformatics *10*, 421.
1153 57. El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter,
1154 S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A., *et al.*
1155 (2019). The Pfam protein families database in 2019. Nucleic Acids
1156 Res. *47*, D427–D432.
1157 58. Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4:
1158 recent updates and new developments. Nucleic Acids Res. *47*,
1159 W256–W259.
1160 59. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high
1161 accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.
1162 60. Jones, D.T., Taylor, W.R., and Thornton, J.M. (1992). The rapid
1163 generation of mutation data matrices from protein sequences.
1164 Comput. Appl. Biosci. *8*, 275–82.
1165 61. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and

1166 Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis
1167 using maximum likelihood, evolutionary distance, and maximum
1168 parsimony methods. Mol. Biol. Evol. *28*, 2731–2739.
1169 62. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment
1170 with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.
1171 63. Dobin, A., and Gingeras, T.R. (2015). Mapping RNA-seq Reads with
1172 STAR. In Current Protocols in Bioinformatics (Hoboken, NJ, USA: John
1173 Wiley & Sons, Inc.), pp. 11.14.1-11.14.19.
1174 64. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter,
1175 A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next
1176 generation web server for deep-sequencing data analysis. Nucleic
1177 Acids Res. *44*, W160–W165.
1178 65. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S.,
1179 Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., *et al.*
1180 (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biol. *9*,
1181 R137.
1182 66. Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M.,
1183 Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics
1184 viewer. Nat. Biotechnol. *29*, 24–26.
1185 67. Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi,
1186 L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for
1187 motif discovery and searching. Nucleic Acids Res. *37*, W202–W208.
1188 68. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of
1189 utilities for comparing genomic features. Bioinformatics *26*, 841–842.
1190 69. Hahne, F., and Ivanek, R. (2016). Visualizing genomic data using Gviz

1191 and Bioconductor. In Statistical genomics, E. Mathé and S. Davis,
1192 eds. (Humana Press, New York, NY), pp. 335–351.
1193 70. Price, A.L., Jones, N.C., and Pevzner, P.A. (2005). De novo
1194 identification of repeat families in large genomes. Bioinformatics *21*,
1195 i351–i358.
1196 71. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence
1197 alignment software version 7: improvements in performance and
1198 usability. Mol. Biol. Evol. *30*, 772–780.
1199 72. Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated
1200 for clustering the next-generation sequencing data. Bioinformatics
1201 *28*, 3150–3152.
1202 73. Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European
1203 Molecular Biology Open Software Suite. Trends Genet. *16*, 276–7.
1204 74. Bao, W., Kojima, K.K., and Kohany, O. (2015). Repbase update, a
1205 database of repetitive elements in eukaryotic genomes. Mob. DNA *6*,
1206 11.
1207 75. Nicolás, F.E., Vila, A., Moxon, S., Cascales, M.D., Torres-Martínez, S.,
1208 Ruiz-Vázquez, R.M., and Garre, V. (2015). The RNAi machinery
1209 controls distinct responses to environmental signals in the basal
1210 fungus *Mucor circinelloides*. BMC Genomics *16*, 237.
1211 76. Nicolas, F.E., Moxon, S., de Haro, J.P., Calo, S., Grigoriev, I. V., Torres-
1212 Martínez, S., Moulton, V., Ruiz-Vázquez, R.M., and Dalmay, T. (2010).
1213 Endogenous short RNAs generated by Dicer 2 and RNA-dependent
1214 RNA polymerase 1 regulate mRNAs in the basal fungus Mucor
1215 circinelloides. Nucleic Acids Res. *38*, 5535–5541.

1216 77. Cervantes, M., Vila, A., Nicolás, F.E., Moxon, S., de Haro, J.P., Dalmay,
1217 T., Torres-Martínez, S., and Ruiz-Vázquez, R.M. (2013). A single
1218 argonaute gene participates in exogenous and endogenous RNAi and
1219 controls cellular functions in the basal fungus *Mucor circinelloides*.
1220 PLoS One *8*, e69283.

1221