

1 A FACS-optimized screen identifies regulators of genome stability in *Candida albicans*

2

3

4 Raphaël Loll-Krippelber^{a, b, c*}, Adeline Feri^{a, b, c}, Marie Nguyen^d, Corinne Maufrais^e, Jennifer
5 Yansouni^f, Christophe d'Enfert^{a, b} and Mélanie Legrand^{a, b#}

6

7 Unité Biologie et Pathogénicité Fongiques, Département Mycologie, Institut Pasteur, Paris,
8 France^a; INRA USC2019, Paris, France^b; Univ. Paris Diderot, Sorbonne Paris Cité, Paris,
9 France^c; Plate-Forme de Cytométrie, Imagopole, Institut Pasteur, Paris, France^d; Centre
10 d'Informatique en Biologie, Institut Pasteur, Paris, France^e; BIOASTER, Genomics and
11 transcriptomics, Lyon, France^f

12

13 Running head: Loss Of Heterozygosity screen in *Candida albicans*

14

15 #Address correspondence to Mélanie Legrand, mlegrand@pasteur.fr.

16 *Present address: Donnelly Centre for Cellular and Biomolecular Research, University of
17 Toronto, Toronto, Canada

18

19

20

21

22

23

24

25 **ABSTRACT**

26 Loss-of-heterozygosity (LOH) plays important roles in genome dynamics, notably during
27 tumorigenesis. In the fungal pathogen *Candida albicans*, LOH contributes to the acquisition
28 of antifungal resistance. In order to investigate the mechanisms that regulate LOH in *C.*
29 *albicans*, we have established a novel method combining an artificial heterozygous locus
30 harboring the BFP and GFP fluorescent markers and flow cytometry to detect LOH events at
31 the single cell level. Using this fluorescence-based method, we have confirmed that elevated
32 temperature, treatment with methyl methanesulfonate and inactivation of the Mec1 DNA
33 damage checkpoint kinase triggered an increase in the frequency of LOH. Taking advantage
34 of this system, we have searched for *C. albicans* genes whose overexpression triggered an
35 increase in LOH and identified four candidates, some of which are known regulators of
36 genome dynamics with human homologues contributing to cancer progression. Hence, the
37 approach presented here will allow implementing new screens to identify genes that are
38 important for genome stability in *C. albicans* and more generally in eukaryotic cells.

39

40 **INTRODUCTION**

41 Normally found as a harmless commensal organism, *Candida albicans* is also the major
42 fungal pathogen of humans and is capable of causing serious and even life-threatening
43 diseases when the immune system of the host is compromised (1). Although *C. albicans* is
44 mostly found as a diploid organism, haploid and tetraploid forms have been observed in the
45 laboratory and upon passage of *C. albicans* in animal models of infection (2-4). The formation
46 of tetraploids results from mating between diploids of opposite mating type that have
47 undergone the so-called white-opaque phenotypic switch (5). Meiosis is thought to have been
48 lost in *C. albicans* and the formation of haploids from diploids or diploids from tetraploids
49 results from concerted chromosome loss (2, 6). The *C. albicans* genome is highly plastic
50 undergoing a number of important genome rearrangements, such as Loss Of Heterozygosity

51 (LOH) events, aneuploidies and the formation of isochromosomes (7). In particular, LOH has
52 been shown to occur during commensal carriage (8) and during systemic infection in a mouse
53 model (9) and to contribute to the acquisition of antifungal resistance (10, 11). Stressful
54 conditions such as high temperature, oxidative stress or azole antifungal treatment trigger an
55 increase in LOH frequencies as well as changes in the mechanisms leading to LOH: while
56 azole treatment and temperature cause an increase in LOH due to chromosome loss and
57 reduplication, oxidative stress results in an increase in gene conversion events (12). Yet, little
58 is known about the molecular mechanisms that control LOH in *C. albicans* despite the
59 apparent importance of LOH in the biology of this species. Overall, it has been shown that
60 knock-out mutations in genes involved in base-excision repair, nucleotide-excision repair and
61 mismatch repair had little impact on LOH frequencies in *C. albicans* (13, 14). In contrast, null
62 mutations in the *MRE11* and *RAD50* genes involved in homologous recombination and in the
63 *MEC1*, *RAD53* and *DUN1* genes involved in the DNA damage checkpoint pathway result in
64 increased frequencies of LOH (13, 15, 16).

65 Evaluation of LOH frequencies in *C. albicans* has up to now relied on the use of several
66 counter-selectable marker genes such as *URA3* and *GAL1* that confer distinctive phenotypes
67 when in a heterozygous state or in a homozygous null state (17-19). *URA3* encodes the
68 orotidine-5-phosphate decarboxylase; while *URA3/URA3* and *URA3/ura3* strains are
69 prototrophic for uridine and sensitive to 5-fluoroorotic acid (5-FOA), *ura3/ura3* strains are
70 auxotrophic for uridine and resistant to 5-FOA. *GAL1* encodes a galactokinase; while
71 *GAL1/GAL1* and *GAL1/gal1* strains are sensitive to 2-deoxygalactose (2-DG), *gal1/gal1*
72 strains are resistant to 2-DG. Hence, measuring the number of spontaneous 5-FOA-resistant or
73 2-DG-resistant clones that arise from a *URA3/ura3* or a *GAL1/gal1* strain, respectively,
74 provides a measure of LOH frequencies in these strains. Although robust, these systems are
75 labor-intensive, costly and lack the flexibility that would be required to implement a genetic

76 screen for *C. albicans* genes regulating LOH. With the development of collections of knock-
77 out constructs or overexpression plasmids, such screens are becoming amenable to *C.*
78 *albicans* (20-22) and could provide new insights in the mechanisms that control genome
79 stability in this species and other eukaryotes.

80 Here, we report the development of a novel LOH-reporter system that combines
81 fluorescent markers and flow cytometry to detect LOH in *C. albicans* and is amenable to
82 high-throughput screening. We show that this system can report changes in LOH frequencies
83 triggered by different physical and chemical stresses, and knock-out mutations. Moreover,
84 using a newly developed collection of overexpression plasmids for 124 genes whose
85 orthologs in *S. cerevisiae* are involved in DNA replication, recombination and repair, we
86 show that the LOH-reporter system can be used to identify genes whose overexpression
87 triggers genome instability, specifically through LOH events. Overall, the characterization of
88 new players in genome maintenance will allow a better understanding of how genomic
89 instability may contribute to *C. albicans* success as a commensal and pathogen.

90

91 **MATERIALS AND METHODS**

92 **Strains and growth conditions**

93 *C. albicans* strains were routinely cultured at 30°C in YPD medium (1% yeast extract, 2%
94 peptone, 2% dextrose), synthetic complete (SC) medium with the omission of the appropriate
95 amino-acids (23) or synthetic dextrose (SD) medium (0.67% yeast nitrogen base, 2%
96 dextrose). Solid media were obtained by adding 2% agar. All *C. albicans* strains used in this
97 work were derived from SN148 (24) and are listed in Table 1. For selection of transformants,
98 nourseothricin was used at a final concentration of 300 µg/mL in YPD.
99 For the LOH assay by FACS, aliquots of prototrophic cells were taken from the -80°C stock
100 and grown in 96-well plates overnight in YPD at 30°C with rotatory shaking for recovery.

101 The next day, these cultures were inoculated in SC-His-Arg to eliminate the cells that had
102 undergone LOH and have lost the *BFP-HIS1* or *GFP-ARG4* cassettes as a result of the
103 freezing/thawing step. The SC-His-Arg cultures were grown overnight at 30°C with rotatory
104 shaking. At this point, the cells were ready for possible treatments before LOH analysis by
105 flow cytometry.

106 Overexpression from P_{TET} was induced by the addition of anhydrotetracycline (ATc, Fisher
107 Bioblock Scientific, 3 µg/ml) in YPD or SD at 30°C. Overexpression experiments were
108 carried out in the dark as ATc is light sensitive.

109 For heat shock, *C. albicans* strain CEC3989 (Table 1) was grown in YPD overnight at 30°C,
110 37°C or 39°C and allowed to recover overnight in YPD before flow cytometry analysis.

111 For genotoxic stress, CEC3989 was cultured for 30 minutes in YPD + 0.01%, 0.02% or
112 0.03% Methyl Methanesulfonate (MMS) at 30°C. The cells were collected, washed with fresh
113 YPD and allowed to recover in rich medium at 30°C overnight before flow cytometry
114 analysis.

115

116 **Plasmid constructions**

117 Oligonucleotides used in this work are listed in Tables S1 and S2 in the supplemental
118 material. Plasmids used or generated in this work are listed in Table S3 in the supplemental
119 material. Plasmid constructions are described in the supplemental material (Text S1) and
120 yielded six plasmids harboring 1) either the *GFP* or *BFP* genes placed under the control of the
121 constitutively-active promoter of the *C. albicans TDH3* gene (25) and 2) either the *ARG4*
122 (26), *HIS1* (27), *URA3* (26) or *SAT1* (27) transformation markers (Figure S1).

123

124 ***C. albicans* strains construction**

125 *Multiple heterozygous markers strain construction*

126 One allele of the *GALI* gene was replaced with the *CmLEU2* marker gene by PCR-mediated
127 transformation of SN148 (24). The resulting strain was named CEC1027. The P_{TDH3} -*BFP-*
128 *HIS1* and P_{TDH3} -*GFP-ARG4* cassettes were PCR-amplified from plasmids pECC727-BH
129 (ECC727) and pECC596-GA (ECC596) with oligonucleotides K7 yFP-Fwd and K7 yFP-Rev
130 carrying homologous sequences to the genomic DNA located between *PGA59* and *PGA62* on
131 chromosome 4 (Ch4). The PCR products were sequentially transformed in CEC1027
132 according to Gola *et al.* (26) yielding CEC2814. For activation of the P_{TET} promoter, strain
133 CEC2814 was eventually transformed with *SacII*,*KpnI*-digested pNIMX that encodes a
134 tetracycline-controlled transactivator (21), yielding CEC2824. Strain CEC2683 was obtained
135 through successive transformation of SN148 (24) with the P_{TDH3} -*BFP-SAT1* and P_{TDH3} -*GFP-*
136 *ARG4* cassettes PCR-amplified from plasmids pECC729-BS (ECC729) and pECC596-GA
137 (ECC596) using oligonucleotides K7 yFP-Fwd and K7 yFP-Rev. CEC2683 was subsequently
138 transformed with *StuI*-linearized *Cip30* (28) to generate CEC3194, the prototrophic parental
139 strain.

140 A *mec1* null mutant was constructed by replacing both alleles with the *LEU2* and *HIS1*
141 markers using a PCR-based method (29) in CEC2683, yielding first the heterozygous
142 *MEC1/mec1* mutant CEC3172 and then the null *mec1/mec1* mutant CEC3183. CEC3172 and
143 CEC3183 were then transformed with *StuI*-linearized *Cip30* (28) or *Cip10* (30) to generate
144 CEC3203 and CEC3216, the prototrophic *mec1* heterozygous and homozygous mutants.

145 *C. albicans* overexpression strains

146 *StuI*-digested or *I-SceI*-digested *Cip10*- P_{TET} -GTW derivatives were transformed into strain
147 CEC2824 as described in Chauvel *et al.* (21), resulting in 124 overexpression strains (see
148 Table S2 in the supplemental material) which together with the CEC3989 control strain
149 derived from CEC2824 transformed with *StuI*-linearized *Cip10*- P_{TET} -GTW will be referred to
150 as the 3R overexpression collection.

151 All *C. albicans* transformants were verified by PCR using primers inside the transformed
152 plasmid/cassette and primers in the genomic DNA regions of insertion in order to assess
153 proper integration of the plasmid/cassette in the *C. albicans* genome.

154

155 **Flow cytometry analysis**

156 Cells from overnight cultures were diluted at 1:1000 in 1X PBS in BD Falcon tubes (Ref
157 352054) and 10^6 cells were analyzed by flow cytometry in 96-well plate format using a
158 MACSQuant (Miltenyi) flow cytometer. We used the 405 nm and 488 nm lasers to excite the
159 BFP and GFP proteins and the 425/475 and 500/550 filters to detect the BFP and GFP
160 emission signals.

161

162 **LOH screen with the BFP/GFP system**

163 *Primary screen*

164 Strains in the 3R overexpression collection were screened for LOH frequencies by flow
165 cytometry following overnight growth at 30°C in YPD containing ATc (3 µg/ml). To evaluate
166 the frequency of LOH at the GFP or BFP loci, we first analyzed the flow cytometry output for
167 14 independent cultures of the control strain (CEC3989) with FlowJo 7.6. All profiles were
168 very similar. We created gates to define the Bfp⁺ Gfp⁻ (mono-BFP) and Bfp⁻ Gfp⁺ (mono-
169 GFP) populations in one of the control sample and applied these gates to the rest of the data
170 set for the control cultures. The percentage of Bfp⁺ Gfp⁻ cells and Bfp⁻ Gfp⁺ cells was
171 exported in excel format. The mean was calculated, as well as the standard deviation (SD).
172 The same gates were applied to all the mutants, resulting in a percentage of Bfp⁺ Gfp⁻ cells
173 and Bfp⁻ Gfp⁺ cells for each mutant strain. For each mutant, a Z-score was calculated and we
174 selected mutants that showed a Z-score superior to 5 for the Bfp⁻ Gfp⁺ population
175 (corresponding to five standard deviations from the mean value of the control strain).

176

177 *Secondary screen*

178 Because the overexpression-induced morphology changes observed in some of the candidate
179 overexpression mutants (data not shown) could distort the flow cytometry results, a secondary
180 screen was carried out to confirm the candidate genes. All mutants identified in the primary
181 screen were grown in triplicates in YPD and YPD+ATc (3 µg/ml) for 8 hours and re-
182 inoculated in YPD alone to recover overnight at 30°C. The next day, 10⁶ cells in the yeast
183 form, as confirmed by microscopy, were analyzed by flow cytometry and a t-test was
184 performed to validate the candidates.

185

186 **Sorting of cells having undergone a LOH event**

187 The control strain and the candidate mutants were grown in the presence of ATc (3 µg/ml) for
188 8 hours and allowed to recover overnight in YPD alone. Fluorescence-activated cell sorting of
189 these cultures, using a FACS Aria III cell sorter (BD Biosciences), was performed into 1.5 ml
190 Eppendorf tubes at a rate of 10,000 evt/s. 1000 cells from the Bfp- Gfp+ and Bfp+ Gfp-
191 populations were sorted, recovered in 400 µl YPD, and immediately plated on YPD plates.
192 The plates were incubated at 30°C for 2 days. For the Bfp- Gfp+ population, 48 colonies were
193 inoculated in 800 µl YPD in 96-well plates and grown overnight at 30°C. For the Bfp+ Gfp-
194 population, we observed a colony size phenotype and therefore 24 small colonies and 24 large
195 colonies were inoculated in 800 µl YPD in 96-well plates and grown overnight at 30°C.
196 Cultures were then analyzed by flow cytometry to validate mono-fluorescence and spotted on
197 YPD, SC-HIS-ARG, SC-HIS and SC-ARG to confirm the LOH events by checking the loss
198 of the auxotrophic marker associated with the BFP or GFP marker. We next investigated the
199 molecular mechanisms giving rise to these LOH events by SNP-typing of 16 isolates, when
200 possible. Genomic DNA was extracted from the sorted clones and PCR-RFLP was performed

201 to assess the extent of the LOH. We used SNPs 95 and 156, which are part of *AluI* and *TaqI*
202 restriction sites, respectively (31). The SNP-typing analysis was carried out for both the Bfp+
203 Gfp- (small and large colonies) and Bfp- Gfp+ populations and was performed as described
204 previously (31).

205

206 **Whole genome sequencing and sequence analysis**

207 Genomic DNAs isolated from the *C. albicans* isolates were processed to prepare libraries for
208 Illumina sequencing. The TruSeq Nano DNA Sample Prep kit (Illumina) was used according
209 to the manufacturer's recommendations. DNA was randomly fragmented by sonication to an
210 average fragment length of 550 base pairs and Illumina adapters were blunt-end ligated to the
211 fragments. The final libraries were amplified by PCR followed by sequencing on a MiSeq
212 platform, using the v3 chemistry. 300 nt paired-end reads were aligned to the *C. albicans*
213 strain SC5314 reference genome assembly 21, available from CGD (32, 33) using BWA (34).
214 Single nucleotide variants (SNVs) between the sequenced genomes and the reference genome
215 were identified using GATK (35) at positions with a sequencing depth equal or above 18X.
216 Heterozygous SNVs were defined as positions where 15% or more of the calls showed one
217 allele and 85% or less of the calls showed a second allele. Homozygous SNVs were defined
218 as positions where more than 90% of the calls differed from the reference genome.
219 Sequencing depth and heterozygous SNP density maps were constructed as described in Loll-
220 Krippleber et al. (16). Homozygous SNP density maps were constructed by counting the
221 number of homozygous positions per 10 kb region and plotting these values.

222

223 **RESULTS**

224 **A new reporter system for LOH**

225 As an alternative to the *GALI* and *URA3* systems routinely used for LOH studies in *C.*
226 *albicans* (12, 18), we developed a new system combining the use of fluorescent markers and
227 flow cytometry, more suitable for high-throughput studies. This system relies on the
228 construction of an artificial locus on Ch4 in the intergenic region between the *PGA59* and
229 *PGA62* ORFs (36). This 9 kb region has been used previously as a platform for integration of
230 *C. albicans* two-hybrid plasmids (37) and its modification is thought to be neutral to *C.*
231 *albicans* biology as its deletion did not result in any phenotype (36). As shown in Fig. 1A, one
232 of the Ch4 homologs was engineered to carry the gene coding for the Green Fluorescent
233 Protein (GFP; (38)) under the control of the promoter of the constitutively, highly-expressed
234 *TDH3* gene (25) and linked to an auxotrophic marker (*ARG4*, *HIS1*, *URA3*) or the
235 nourseothricin resistance marker (*SATI*). The other homolog was engineered to carry the gene
236 coding for the Blue Fluorescent Protein (BFP; Evrogen) placed under the control of the *TDH3*
237 promoter and linked to another auxotrophic marker or the nourseothricin resistance marker. A
238 series of cassettes have been developed that associate the *BFP* or *GFP* genes to different
239 selection markers (*SATI*, *HIS1*, *URA3* and *ARG4*), allowing their use in various genetic
240 backgrounds (see Fig. S1 in the supplemental material). In this set-up, LOH events can be
241 detected by flow cytometry since cells that undergo LOH at the *PGA59-62* locus will produce
242 only one of the two fluorescent proteins. As shown in Figure 1B, FACS analysis of strain
243 CEC2683, which carries the *BFP-SATI/GFP-ARG4* system showed that the majority of the
244 cells (>99%) co-expressed the two fluorescent proteins BFP and GFP. In addition, mono-GFP
245 and mono-BFP populations were observed at frequencies of 0.020% and 0.033%,
246 respectively. In order to verify that the cells from these populations had undergone genuine
247 LOH events, we recovered cells from each population by cell-sorting and characterized 76
248 and 52 cells from the mono-GFP and mono-BFP populations, respectively. 95% (72/76) and
249 94% (49/52) of the cells recovered from both populations were mono-fluorescent (data not

250 shown). Moreover, spotting assays on appropriate media indicated that all mono-fluorescent
251 cells displayed the phenotypes expected from a concomitant loss of the *BFP* and *SAT1* genes
252 (mono-GFP cells) or the *GFP* and *ARG4* genes (mono-BFP cells), *ie* sensitivity to
253 nourseothricin or arginine auxotrophy, respectively (data not shown). This initial analysis of a
254 BFP/GFP strain revealed a spontaneous LOH frequency of $2-3 \times 10^{-4}$ at this locus.
255 Interestingly, the cells recovered from the Bfp- Gfp- population did not grow, indicating that
256 these correspond to non viable cells. Fluorescence plots of mono-BFP (CEC4007), mono-GFP
257 (CEC3998), unlabeled (SN100) and BFP/GFP-labeled strains are presented in Figure S2A in
258 the supplemental material.

259

260 **LOH increase at the *PGA59-62* locus in *mec1* mutants**

261 The Mec1 kinase is a central regulator of the DNA damage checkpoint in all eukaryotes (39,
262 40). A previous study has shown that inactivation of the *C. albicans MEC1* gene results in an
263 increase in LOH at the *GALI* locus (15). In order to test whether an increase in LOH could be
264 detected with the BFP/GFP system, we inactivated the *MEC1* gene in a strain carrying this
265 system. Results presented in Figure 2A and Table 2 showed a statistically significant increase
266 in the proportion of mono-GFP or mono-BFP cells in *mec1* null mutants as compared to the
267 wild-type strain. Interestingly, we also observed an increase in the proportion of Bfp- Gfp-
268 cells, suggesting that inactivation of *MEC1* increases cell mortality (Fig. 2A). Cell death in
269 the Bfp- Gfp- population was confirmed by propidium iodide staining and microscopic
270 analysis of these cells after cell sorting (data not shown). Hence, the BFP/GFP LOH reporter
271 system could detect an increase in LOH frequencies by flow cytometry upon inactivation of
272 the *MEC1* gene.

273

274 **LOH increase at the *PGA59-62* locus in response to different stresses**

275 Forche *et al.* (12) have shown an increase in LOH in response to different stresses that are
276 reminiscent of the conditions encountered by fungi during human infection, e.g. elevated
277 temperatures, oxidative stress and the presence of azole antifungals. We evaluated the
278 frequency of LOH in a control strain carrying the BFP/GFP system when grown overnight at
279 30°C, 37°C or 39°C. Although cells were filamenting after 7 hours growth at 37°C or 39°C,
280 microscopic analysis of cells after an overnight growth at 37°C or 39°C revealed that the
281 majority of the cells reverted to the yeast form and were therefore suitable for flow cytometry
282 analysis (data not shown). Still, we allowed the cells to recover overnight in YPD before flow
283 cytometry analysis. Similarly to the data obtained with the *GALI* system (12), the BFP/GFP
284 system revealed that elevated temperatures promote LOH in *C. albicans* (Fig. 2B and Table
285 2). The fluorescence plots presented in Figure S2B confirmed that the distribution of mono-
286 labeled cells remains the same in absence or presence of heat stress.

287 We also treated the control strain carrying the BFP/GFP system with the genotoxic agent
288 MMS (Methyl Methanesulfonate). After a 30-minute treatment at different concentrations of
289 MMS, the cells were allowed to recover in rich medium overnight. Flow cytometry analysis
290 revealed a LOH frequency that increased together with MMS concentrations (Fig. 2C and
291 Table 2).

292 Taken together, these data indicated that the BFP/GFP LOH reporter system was suitable to
293 detect increases in LOH frequencies resulting from physical or chemical stresses.

294

295 **Construction of the 3R overexpression collection**

296 We reasoned that the BFP/GFP LOH reporter system could provide an efficient means for the
297 identification of genes that control LOH in *C. albicans*. Therefore, we implemented an
298 overexpression screen aimed at identifying *C. albicans* genes whose overexpression would
299 alter wild-type LOH frequencies. To this aim, we generated a collection of *C. albicans*

300 overexpression strains focused on genes which, based on their GO annotations or the function
301 of their orthologs in *S. cerevisiae*, were likely to be involved in different aspects of genome
302 maintenance such as DNA repair, DNA replication, recombination, chromosome segregation,
303 cell cycle, and telomere maintenance. PCR products extending from the start codon to the
304 penultimate codon of 204 selected genes (see Table S2 in the supplemental material) were
305 cloned into the pDONR207 donor vector. Following Sanger and Illumina/Solexa sequence
306 validation, a total of 151 (74%) derivatives of pDONR207 were obtained. ORFs cloned into
307 pDONR207 were subsequently transferred into uniquely barcoded CIp10- P_{TET} -GTW plasmids
308 (20, 21). A total of 147 (72%) CIp10- P_{TET} -GTW derivatives were obtained and subsequently
309 introduced at the *RPS1* locus in *C. albicans* strain CEC2824 that harbors a *GFP-HIS1/BFP-*
310 *ARG4* system at the *PGA59-62* locus and the pNIMX plasmid encoding a tetracycline-
311 controlled transactivator (21). Eventually, 124 *C. albicans* P_{TET} -driven OE strains were
312 obtained indicating an overall 60.8% success rate, slightly below that reported previously by
313 Chauvel et al. (21) for a non overlapping set of overexpression plasmids, possibly because of
314 the larger size of the cloned ORFs in our study (27% above 2500 bp vs none in the study of
315 Chauvel et al.). Together with the CEC3989 control strain harboring the empty CIp10- P_{TET} -
316 GTW plasmid, these 124 overexpression strains will be referred to below as the 3R
317 overexpression collection.

318

319 **FACS-optimized LOH screens upon gene overexpression**

320 Prior to analyzing the overexpression mutants for LOH frequencies under inducing
321 conditions, we evaluated the LOH frequency in the control strain CEC3989. Fourteen
322 independent YPD+ATc (3 μ g/mL) cultures of the control strain were analyzed by flow
323 cytometry. For each analysis, 1 million cells were analyzed and the mean frequency of LOH
324 and standard deviation were calculated. We found that LOH occurred at an average frequency

325 of $2.2 \times 10^{-4} \pm 0.4 \times 10^{-4}$. Subsequently, we submitted the entire 3R overexpression collection to
326 a primary screen whereby 1 million cells of each overexpression strain were analyzed by flow
327 cytometry after an overnight induction. Thirty-three overexpression strains were selected that
328 had a Z-score (for the frequency of Bfp- Gfp+ cells) above 5 (see Fig. S3 in the supplemental
329 material). In the secondary screen, we allowed the cells to recover in rich media after an 8
330 hours induction in YPD + ATc3 and re-analyzed the 33 overexpression strains identified in
331 the primary screen in triplicates. While we saw no difference in Bfp+ Gfp- and Bfp- Gfp+
332 frequencies between the control strain and the overexpression mutants in absence of ATc
333 treatment, we confirmed the increase in LOH at the *PGA59-62* locus after ATc treatment for 4
334 of these genes, namely *BIMI* (10-fold for the Bfp+ Gfp- cells and 13-fold for the Bfp- Gfp+
335 cells), *CDC20* (32- and 38-fold, respectively), *RAD51* (7- and 8-fold, respectively) and
336 *RAD53* (13-fold) (Table 3). Notably, we could not identify any overexpression strain that had
337 a significantly reduced LOH frequency possibly owing to the relatively low LOH frequency
338 in the control strain.

339

340 **Characterization of the molecular mechanisms leading to LOH**

341 LOH at the BFP/GFP locus could result from gene conversion, mitotic crossing over (MCO)
342 or break-induced replication (BIR), chromosome truncation or chromosome loss. Further to
343 identifying genes whose overexpression increased LOH frequencies, we tested whether
344 overexpression of these genes could also alter the balance between these different types of
345 events as observed in the control strain.

346 Fluorescence-activated cell sorting was used in order to recover individual mono-fluorescent
347 cells and subsequently analyze these cells by SNP-RFLP typing that can reveal the molecular
348 mechanisms – gene conversion, MCO-BIR or chromosome truncation, chromosome loss –
349 underlying the LOH event responsible for the loss of the BFP or GFP marker on Ch4 (Fig.

350 3A; (31)). Cells sorted from the Bfp- Gfp+ control and mutant populations and plated on YPD
351 gave rise to colonies homogeneous in size. For the control strain, SNP-RFLP typing showed
352 that LOH in the cells that had yielded these colonies had arisen at somewhat equal frequency
353 through BIR/MCO or chromosome truncation, these two types of events being
354 undistinguishable by SNP-RFLP typing, and gene conversion (Fig. 3B). A complete loss of
355 the homolog carrying the BFP marker was never observed. A similar pattern was observed
356 upon overexpression of *CDC20* and *RAD53*. In contrast, we detected an increase in the
357 proportion of BIR/MCO or chromosome truncation events versus gene conversions resulting
358 in the loss of the BFP marker in the *RAD51* and *BIMI* overexpression mutants (Fig. 3B).
359 Notably, colonies of different size were observed when cells sorted from the Bfp+ Gfp-
360 populations were plated on YPD (data not shown). This included colonies of a size similar to
361 those obtained from the Bfp- Gfp+ populations and smaller colonies. This small colony
362 phenotype was observed for the control strain and all overexpression mutants tested. In the
363 control strain, 80% and 20% of the mono-BFP cells were derived from small and large
364 colonies, respectively (Fig. 3C). We observed an increase in the proportion of the large
365 colony-derived mono-BFP cells in the *CDC20*, *RAD51* and *RAD53* overexpression mutants
366 (Fig. 3C). SNP-RFLP typing of 16 mono-BFP cells from small and large colonies revealed
367 that, unlike the Bfp- Gfp+ population, chromosome loss was the major mechanism
368 responsible for the loss of the GFP marker in the small colony-derived cells in both the
369 control strain and the overexpression mutants, other LOH events resulting from MCO-BIR or
370 chromosome truncation (Fig. 3D). On the other hand, SNP-RFLP typing revealed that LOH
371 was often the result of recombination-mediated events or chromosome truncations in the
372 larger colonies of the control strain and the *RAD51*, *BIMI* and *RAD53* overexpression mutants
373 (Fig. 3E). In contrast, chromosome loss was the only mechanism leading to LOH in the Bfp+
374 Gfp- cells of the *CDC20* overexpression mutants (Fig. 3D and 3E). SNP-RFLP typing data for

375 both types of colonies were normalized to the proportion of small vs large colonies and
376 combined to represent the percentage of each molecular mechanism (gene conversion,
377 BIR/MCO or chromosome truncation, chromosome loss) in the overall Bfp+ Gfp- population
378 (Fig. 3F).

379 Although the homozygous status of both SNPs 156 and 95 is routinely recognized as a mark
380 of LOH chromosome loss on Ch4 (31), multiple gene conversion events could also be
381 responsible for the concomitant homozygous status of these SNPs. In order to determine the
382 extent of the LOH, we performed whole genome sequencing of 3 small Bfp+ Gfp- colonies
383 obtained from an overexpression mutant that did not show any change in LOH frequency
384 relative to the control and in which LOH were the result of chromosome loss, as demonstrated
385 by SNP-RFLP typing. Heterozygous and homozygous SNP density maps against the *C.*
386 *albicans* SC5314 HapA and HapB reference sequences (41) revealed that cells from the 3
387 small colonies had lost the entire Ch4B carrying the GFP marker (Fig. 3G). The observation
388 that Ch4 had remained disomic (data not shown) indicates that the other homolog 4A has been
389 reduplicated. Homozygosis of Ch4A also accounts for the histidine auxotrophy displayed by
390 the Bfp+ Gfp- cells, despite the maintenance of the *BFP-HIS1* cassette at the *PGA59-62* locus
391 (data not shown). Indeed, the inactive *HIS4* allele reported by Gómez-Raja *et al.* (42) is
392 carried by Ch4A.

393 Overall, our SNP-RFLP analysis revealed several interesting features: (i) the molecular
394 mechanisms causing the LOH differed according to the Ch4 homolog undergoing the LOH
395 event, with chromosome loss being only observed for Ch4B; (ii) this was reflected by the
396 occurrence of small colony size clones when all or part of Ch4B was lost, (iii) when
397 chromosome loss occurred, overexpression of *CDC20* favored this mechanism, (iv) *RAD51*
398 overexpression favored BIR/MCO events, (v) *BIMI* overexpression favored BIR/MCO events

399 unless chromosome loss was possible; and (vi) *RAD53* overexpression favored BIR/MCO
400 events when chromosome loss was possible.

401 **DISCUSSION**

402 Here, we have presented a FACS-optimized genetic system to detect LOH in *C. albicans* and
403 its application to the identification of genes whose overexpression results in an increase in
404 LOH frequency and changes in the frequencies of the molecular events that are at the origin of
405 LOH.

406 The new BFP/GFP system displays appealing features over other genetic systems currently
407 used to study LOH in *C. albicans*. Indeed, we provide a powerful and robust tool that allows
408 rare event analysis and high-throughput LOH detection since up to 96 samples can be
409 processed in a single run. This system is robust and reproducible because LOH is studied on a
410 cellular scale and a large number of cells are analyzed. Unlike the systems based on the *GALI*
411 or *URA3* markers, which require large amounts of expensive drugs, the BFP/GFP system does
412 not require any costly consumables. The use of drugs such as 2-deoxygalactose or 5-
413 fluoroorotic acid also raises the issue of exposing the cells to a selective pressure, which
414 might distort the evaluation of LOH frequencies. In contrast, the BFP/GFP system allows the
415 measurement of spontaneous LOH frequencies since cells are not exposed to any stress. As
416 the other LOH reporter systems, the BFP/GFP system is in theory flexible since the markers
417 can be inserted anywhere in the genome thus enabling determination of LOH frequencies at
418 different genomic locations and therefore the study of site specific LOH events.

419 The BFP/GFP system was used to screen a collection of 124 *C. albicans* overexpression
420 mutants to identify genes whose overexpression would result in an increase in LOH
421 frequency. Our screen identified 4 candidates as regulators of genome stability in *C. albicans*.
422 Null or conditional deletion mutants have been obtained and studied for 3 of these genes:
423 *CDC20* (43), *RAD51* (44) and *RAD53* (45). Yet, the role of these genes in genome integrity

424 was never addressed, except for *RAD53* whose deletion increases LOH frequency (16).
425 Reports of *C. albicans* overexpression mutants of the 4 candidates genes could not be found
426 in the literature. In contrast, null or conditional deletion mutants, as well as overexpression
427 mutants, have been described in *S. cerevisiae* for the 4 candidate genes. Although genome
428 stability was investigated in the deletion mutants, the question of genome integrity was not
429 addressed in the overexpression mutants.

430 Our screen identified *CDC20* as the gene whose overexpression triggers the highest increase
431 in LOH frequency. In *S. cerevisiae*, the essential *CDC20* gene encodes an activator of
432 ubiquitin-protein ligase activity and is required for metaphase/anaphase transition. At this
433 stage of the mitotic cycle, Cdc20p enables the APC (Anaphase-Promoting Complex) to
434 properly degrade Securin allowing the degradation of the cohesin rings that link the two sister
435 chromatids. Several studies have shown that degradation of Cdc20p is an essential and
436 conserved mechanism to maintain the spindle assembly checkpoint (46, 47). Therefore, the
437 sustained presence of the Cdc20 protein could perturb genome integrity and our observation
438 that overexpression of *C. albicans CDC20* favors LOH, especially through chromosome loss
439 when these types of events can occur, is consistent with this hypothesis. Conditional *CDC20*
440 deletion mutants have been characterized in *C. albicans* by Chou et al. (43). Their work
441 suggested that, similarly to *ScCDC20*, *CaCDC20* was important for the metaphase-to-
442 anaphase transition and mitotic exit. Interestingly, *CDC20* has a distinct role in
443 morphogenesis in *C. albicans*, which is not the case in *S. cerevisiae*.

444 Rad53 and Rad51, two proteins involved in DNA damage checkpoint and DNA repair, were
445 identified as causing an increase in LOH upon overexpression. Rad53 is a kinase involved in
446 the DNA-damage checkpoint whose deletion in *C. albicans* causes a growth defect, an
447 increased sensitivity to several DNA-damaging agents and a defect in genotoxic-stress
448 induced filamentation (45). An increase in LOH could also be observed in *CaRAD53* deletion

449 mutants (16). The recombinase Rad51 plays a major role in homologous recombination
450 during DNA double strand break repair by searching for sequence homology and promoting
451 strand pairing. Deletion of *RAD51* in *C. albicans* results in a decreased growth rate and an
452 increased sensitivity to various DNA-damaging agents (44). On the other hand, genome
453 integrity has not been investigated in the *CaRAD51* deletion mutants. In *S. cerevisiae*, *RAD51*
454 overexpression has been shown to promote genome instability, probably by inhibiting the
455 accurate repair of double-strand breaks by homologous recombination (48). Similarly, *RAD51*
456 overexpression in the mouse leads to an increase in genome instability, notably LOH events
457 (49). Our observation that, in addition to increasing LOH frequency, overexpression of
458 *RAD51* and *RAD53* favored BIR/MCO events is consistent with their role in DNA double
459 strand break repair, their overexpression being likely to unbalance the DNA damage response
460 and favor recombination-mediated events.

461 In this work, a structural component of the microtubule skeleton of yeast encoded by *BIMI*
462 has been identified as increasing LOH frequency upon overexpression. While *ScBIMI*
463 deletion has been shown to result in increased sensitivity to various DNA-damaging agents
464 (50) and increased chromosome instability (51), the effect of *BIMI* overexpression on genome
465 integrity has not been investigated, probably in part because of the strong growth defect
466 associated with *BIMI* overexpression in *S. cerevisiae* (52-54). Similarly, we also observed a
467 growth defect associated to *BIMI* overexpression in *C. albicans* (data not shown). Our
468 observation that *BIMI* overexpression, in addition to promoting LOH, favors BIR/MCO
469 events was unexpected as Bim1 is involved in chromosome segregation, and therefore its
470 overexpression would be more likely to cause chromosome loss. We hypothesized that the
471 overexpression-induced reduced growth rate associated with *BIMI* overexpression could
472 favor the occurrence of DNA breaks and thus explain the increase in BIR/MCO events or
473 chromosome truncations in the *BIMI* overexpression mutant.

474 Our SNP-typing analysis of sorted cells revealed an additional important aspect. Indeed, we
475 have observed that the molecular mechanisms giving rise to LOH can be homolog-specific
476 with chromosome loss never affecting Ch4A. Notably, loss of all or part of Ch4B was
477 accompanied by variations in colony size. As cells in small colonies have almost all arisen
478 through chromosome loss (Fig. 3D) and yield large colonies (data not shown), we hypothesize
479 that small colonies reflect a slower growth rate of cells that are monosomic for Ch4A and
480 eventually duplicate this chromosome. Chromosome homozygosis was long thought to be
481 hampered in *C. albicans* by the presence of recessive lethal mutations dispersed throughout
482 the genome. The existence of haploids (2), completely homozygous *rad52*-derived isolates
483 (55) and partly homozygous diploid parasexual progenies (56) demonstrated that
484 homozygosis of each one of the 8 chromosomes can occur. Nevertheless, a bias in the
485 homolog being retained was reported for certain chromosomes (2, 55). This was the case for
486 Ch4, for which one homolog could never be lost (2, 55). Our observation that Ch4A could
487 never be lost suggests the presence of at least one recessive lethal allele on Ch4B. Because
488 BIR/MCO or chromosome truncation-mediated LOH occurred in mono-GFP cells, we can
489 narrow the recessive lethal allele(s)-carrying region to the right arm of Ch4B or the left arm of
490 Ch4B between the *PGA59/62* locus and the centromere. Genome sequencing of these mono-
491 GFP cells will reveal the extent of the LOH and allow to refine the position of this or these
492 hypothetical recessive lethal alleles. While homozygosis along the entire smaller
493 chromosomes, Ch5, Ch6 and Ch7, were often observed, MCO-BIR events or chromosome
494 truncations have been reported to be responsible for most LOH on both Ch1 homologs (12,
495 13). Altogether, these observations suggest that the mechanisms underlying LOH in *C.*
496 *albicans* could be chromosome-specific, as well as homolog-specific. The presence of
497 essential alleles that cannot be lost might indeed select for cells that have repaired DSB
498 through BIR, especially in the case of Ch1, one of the largest chromosomes in *C. albicans*

499 (3.2 Mb). On the other hand, the loss of one of Ch4 homologs, which is half the size of Ch1
500 (1.6 Mb), might be less detrimental to the cell. Alternatively, Ch4 might be more prone to
501 mitotic nondisjunction events or to unrepaired DNA lesions.

502 Some expected genes such as *TUB2* were not recovered in our screen (57). Yet, this may
503 reflect the inherent limitations of a genetic screen. First, a number of the usual suspects were
504 shown to play a role in genome stability based on deletion mutant study, which does not
505 imply that their overexpression triggers an increase in LOH frequency. Second, the level of
506 overexpression achieved in this study might be insufficient to cause a significant phenotype.
507 Because the proteins are produced in an untagged form, we have no means of verifying their
508 production. Third, the lack of phenotype could result from the presence of peptides encoded
509 by the *att* sequences at the amino- and carboxy-terminal ends of the proteins, which could
510 alter their proper folding and activity.

511 Living cells have evolved different mechanisms in order to ensure genome integrity (58).
512 Interestingly, these mechanisms are conserved across species and cancer genetic studies have
513 linked malfunctions in these processes to the genome instability observed in human cancer
514 cells (59, 60). Therefore, the characterization of all the players involved in genome integrity
515 maintenance and a better understanding of the associated signaling pathways are very
516 important from the human perspective. In this respect, functional genomic approaches in
517 model eukaryotic species such as *S. cerevisiae* have uncovered many components required for
518 genome maintenance and integrity (61). Genome-wide genetic screens in this species have
519 examined deletion mutants for growth defect upon exposure to different DNA-damaging
520 agents (62-65) or for synthetic genetic interaction with other mutant genes important for the
521 cellular response to DNA damage (66-69). Other screens identified mutants with increased
522 genomic instability by investigating the mutation rates within a specific gene (70) or the rates
523 of gross chromosomal rearrangements (51, 71-75). While several studies have used *S.*

524 *cerevisiae* to study genome stability, drastic genome changes are not well tolerated in this
525 organism thus correlating with a high fitness cost (76). Interestingly, *C. albicans* is more
526 tolerant to genome changes, which is reminiscent of the situation in cancer cells that continue
527 to divide rapidly despite undergoing massive genome rearrangements. The tools presented
528 here, in association with the *C. albicans* ORFeome we are currently constructing with the
529 group of Dr Carol Munro in Aberdeen (20), will allow *C. albicans* to become a model of
530 choice to study mechanisms involved in eukaryotic genome integrity.

531

532 **ACKNOWLEDGMENTS**

533 We thank Christiane Bouchier and Laurence Ma from the Pasteur Plateforme Genomique
534 (PF1) for Illumina sequencing of the overexpression plasmids. We are thankful to members of
535 the BIOASTER genomics and transcriptomics core facility, namely Lilia Boucinha and
536 Frédéric Reynier, for help with genome sequencing. We are also grateful to other members of
537 Fungal Biology and Pathogenicity Unit for helpful discussions.

538 This work was supported by the French Government's Investissement d'Avenir program,
539 Laboratoire d'Excellence "Integrative Biology of Emerging Infectious Diseases" (grant #
540 ANR-10-LABX- 62-IBEID). RL was the recipient of a PhD grant from INRA and Institut
541 Pasteur. ML was the recipient of a postdoctoral fellowship from Institut Pasteur (Bourse
542 Roux).

543

544 **REFERENCES**

- 545 1. **Edmond MB, Wallace SE, McClish DK, Pfaller MA, Jones RN, Wenzel RP.** 1999.
546 Nosocomial bloodstream infections in United States hospitals: a three-year analysis.
547 Clin. Infect. Dis. **29**:239-244.

- 548 2. **Hickman MA, Zeng G, Forche A, Hiraakawa MP, Abbey D, Harrison BD, Wang**
549 **YM, Su CH, Bennett RJ, Wang Y, Berman J.** 2013. The 'obligate diploid' *Candida*
550 *albicans* forms mating-competent haploids. *Nature* **494**:55-59.
- 551 3. **Hull CM, Raisner RM, Johnson AD.** 2000. Evidence for mating of the "asexual"
552 yeast *Candida albicans* in a mammalian host. *Science* **289**:307-310.
- 553 4. **Magee BB, Magee PT.** 2000. Induction of mating in *Candida albicans* by construction
554 of MTL α and MTL α strains. *Science* **289**:310-313.
- 555 5. **Miller MG, Johnson AD.** 2002. White-opaque switching in *Candida albicans* is
556 controlled by mating-type locus homeodomain proteins and allows efficient mating.
557 *Cell* **110**:293-302.
- 558 6. **Bennett RJ, Johnson AD.** 2003. Completion of a parasexual cycle in *Candida*
559 *albicans* by induced chromosome loss in tetraploid strains. *EMBO J.* **22**:2505-2515.
- 560 7. **Selmecki A, Forche A, Berman J.** Genomic plasticity of the human fungal pathogen
561 *Candida albicans*. *Eukaryotic cell* **9**:991-1008.
- 562 8. **Diogo D, Bouchier C, d'Enfert C, Bournoux ME.** 2009. Loss of heterozygosity in
563 commensal isolates of the asexual diploid yeast *Candida albicans*. *Fungal Genet. Biol.*
564 **46**:159-168.
- 565 9. **Forche A, Magee PT, Selmecki A, Berman J, May G.** 2009. Evolution in *Candida*
566 *albicans* populations during a single passage through a mouse host. *Genetics* **182**:799-
567 811.
- 568 10. **Coste A, Turner V, Ischer F, Morschhauser J, Forche A, Selmecki A, Berman J,**
569 **Bille J, Sanglard D.** 2006. A mutation in *Tac1p*, a transcription factor regulating
570 *CDR1* and *CDR2*, is coupled with loss of heterozygosity at chromosome 5 to mediate
571 antifungal resistance in *Candida albicans*. *Genetics* **172**:2139-2156.

- 572 11. **Dunkel N, Blass J, Rogers PD, Morschhauser J.** 2008. Mutations in the multi-drug
573 resistance regulator MRR1, followed by loss of heterozygosity, are the main cause of
574 MDR1 overexpression in fluconazole-resistant *Candida albicans* strains. *Mol.*
575 *Microbiol.* **69**:827-840.
- 576 12. **Forche A, Abbey D, Pisithkul T, Weinzierl MA, Ringstrom T, Bruck D, Petersen**
577 **K, Berman J.** 2011. Stress alters rates and types of loss of heterozygosity in *Candida*
578 *albicans*. *mBio* **2**.
- 579 13. **Legrand M, Chan CL, Jauert PA, Kirkpatrick DT.** 2007. Role of DNA mismatch
580 repair and double-strand break repair in genome stability and antifungal drug
581 resistance in *Candida albicans*. *Eukaryotic cell* **6**:2194-2205.
- 582 14. **Legrand M, Chan CL, Jauert PA, Kirkpatrick DT.** 2008. Analysis of base excision
583 and nucleotide excision repair in *Candida albicans*. *Microbiology* **154**:2446-2456.
- 584 15. **Legrand M, Chan CL, Jauert PA, Kirkpatrick DT.** 2011. The contribution of the
585 S-phase checkpoint genes MEC1 and SGS1 to genome stability maintenance in
586 *Candida albicans*. *Fungal Genet. Biol.*
- 587 16. **Loll-Krippelber R, d'Enfert C, Feri A, Diogo D, Perin A, Marcet-Houben M,**
588 **Bougnoux ME, Legrand M.** 2013. A study of the DNA damage checkpoint in
589 *Candida albicans*: uncoupling of the functions of Rad53 in DNA repair, cell cycle
590 regulation and genotoxic stress-induced polarized growth. *Mol. Microbiol.* **91**:452-
591 471.
- 592 17. **Boeke JD, LaCroute F, Fink GR.** 1984. A positive selection for mutants lacking
593 orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance.
594 *Mol. Gen. Genet.* **197**:345-346.

- 595 18. **Forche A, May G, Beckerman J, Kauffman S, Becker J, Magee PT.** 2003. A
596 system for studying genetic changes in *Candida albicans* during infection. *Fungal*
597 *Genet. Biol.* **39**:38-50.
- 598 19. **Gorman JA, Gorman JW, Koltin Y.** 1992. Direct selection of galactokinase-
599 negative mutants of *Candida albicans* using 2-deoxy-galactose. *Curr. Genet.* **21**:203-
600 206.
- 601 20. **Cabral V, Chauvel M, Firon A, Legrand M, Nesseir A, Bachellier-Bassi S,**
602 **Chaudhari Y, Munro CA, d'Enfert C.** 2012. Modular gene over-expression
603 strategies for *Candida albicans*. *Methods Mol. Biol.* **845**:227-244.
- 604 21. **Chauvel M, Nesseir A, Cabral V, Znaidi S, Goyard S, Bachellier-Bassi S, Firon**
605 **A, Legrand M, Diogo D, Naulleau C, Rossignol T, d'Enfert C.** 2012. A Versatile
606 Overexpression Strategy in the Pathogenic Yeast *Candida albicans*: Identification of
607 Regulators of Morphogenesis and Fitness. *PloS one* **7**:e45912.
- 608 22. **Noble SM, French S, Kohn LA, Chen V, Johnson AD.** Systematic screens of a
609 *Candida albicans* homozygous deletion library decouple morphogenetic switching and
610 pathogenicity. *Nat. Genet.* **42**:590-598.
- 611 23. **Rose M, Winston F, Hieter P.** 1990. *Methods in Yeast Genetics. A Cold Spring*
612 *Harbour Laboratory Course Manual.* Cold Spring Harbour Laboratory Press.
- 613 24. **Noble SM, Johnson AD.** 2005. Strains and strategies for large-scale gene deletion
614 studies of the diploid human fungal pathogen *Candida albicans*. *Eukaryotic cell* **4**:298-
615 309.
- 616 25. **Delgado ML, Gil ML, Gozalbo D.** 2003. *Candida albicans* TDH3 gene promotes
617 secretion of internal invertase when expressed in *Saccharomyces cerevisiae* as a
618 glyceraldehyde-3-phosphate dehydrogenase-invertase fusion protein. *Yeast* **20**:713-
619 722.

- 620 26. **Gola S, Martin R, Walther A, Dunkler A, Wendland J.** 2003. New modules for
621 PCR-based gene targeting in *Candida albicans*: rapid and efficient gene targeting using
622 100 bp of flanking homology region. *Yeast* **20**:1339-1347.
- 623 27. **Schaub Y, Dunkler A, Walther A, Wendland J.** 2006. New pFA-cassettes for PCR-
624 based gene manipulation in *Candida albicans*. *J. Basic Microbiol.* **46**:416-429.
- 625 28. **Dennison PM, Ramsdale M, Manson CL, Brown AJ.** 2005. Gene disruption in
626 *Candida albicans* using a synthetic, codon-optimised Cre-loxP system. *Fungal Genet.*
627 *Biol.* **42**:737-748.
- 628 29. **Wilson RB, Davis D, Mitchell AP.** 1999. Rapid hypothesis testing with *Candida*
629 *albicans* through gene disruption with short homology regions. *J. Bacteriol.* **181**:1868-
630 1874.
- 631 30. **Murad AM, Lee PR, Broadbent ID, Barelle CJ, Brown AJ.** 2000. CIp10, an
632 efficient and convenient integrating vector for *Candida albicans*. *Yeast* **16**:325-327.
- 633 31. **Forche A, Steinbach M, Berman J.** 2009. Efficient and rapid identification of
634 *Candida albicans* allelic status using SNP-RFLP. *FEMS yeast research* **9**:1061-1069.
- 635 32. **Binkley J, Arnaud MB, Inglis DO, Skrzypek MS, Shah P, Wymore F, Binkley G,**
636 **Miyasato SR, Simison M, Sherlock G.** 2014. The *Candida* Genome Database: the
637 new homology information page highlights protein similarity and phylogeny. *Nucleic*
638 *Acids Res* **42**:D711-716.
- 639 33. **van het Hoog M, Rast TJ, Martchenko M, Grindle S, Dignard D, Hogues H,**
640 **Cuomo C, Berriman M, Scherer S, Magee BB, Whiteway M, Chibana H, Nantel**
641 **A, Magee PT.** 2007. Assembly of the *Candida albicans* genome into sixteen
642 supercontigs aligned on the eight chromosomes. *Genome biology* **8**:R52.
- 643 34. **Li H, Durbin R.** 2009. Fast and accurate short read alignment with Burrows-Wheeler
644 transform. *Bioinformatics* **25**:1754-1760.

- 645 35. **McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A,**
646 **Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA.** 2010. The Genome
647 Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA
648 sequencing data. *Genome Res.* **20**:1297-1303.
- 649 36. **Moreno-Ruiz E, Ortu G, de Groot PW, Cottier F, Loussert C, Prevost MC, de**
650 **Koster C, Klis FM, Goyard S, d'Enfert C.** 2009. The GPI-modified proteins Pga59
651 and Pga62 of *Candida albicans* are required for cell wall integrity. *Microbiology*
652 **155**:2004-2020.
- 653 37. **Stynen B, Van Dijck P, Tournu H.** 2010. A CUG codon adapted two-hybrid system
654 for the pathogenic fungus *Candida albicans*. *Nucleic Acids Res* **38**:e184.
- 655 38. **Cormack BP, Bertram G, Egerton M, Gow NA, Falkow S, Brown AJ.** 1997.
656 Yeast-enhanced green fluorescent protein (yEGFP): a reporter of gene expression in
657 *Candida albicans*. *Microbiology* **143 (Pt 2)**:303-311.
- 658 39. **Kato R, Ogawa H.** 1994. An essential gene, *ESR1*, is required for mitotic cell growth,
659 DNA repair and meiotic recombination in *Saccharomyces cerevisiae*. *Nucleic Acids*
660 *Res* **22**:3104-3112.
- 661 40. **Zhou BB, Elledge SJ.** 2000. The DNA damage response: putting checkpoints in
662 perspective. *Nature* **408**:433-439.
- 663 41. **Muzzey D, Schwartz K, Weissman JS, Sherlock G.** 2013. Assembly of a phased
664 diploid *Candida albicans* genome facilitates allele-specific measurements and provides
665 a simple model for repeat and indel structure. *Genome biology* **14**:R97.
- 666 42. **Gomez-Raja J, Andaluz E, Magee B, Calderone R, Larriba G.** 2008. A single
667 SNP, G929T (Gly310Val), determines the presence of a functional and a non-
668 functional allele of *HIS4* in *Candida albicans* SC5314: detection of the non-functional
669 allele in laboratory strains. *Fungal Genet. Biol.* **45**:527-541.

- 670 43. **Chou H, Glory A, Bachewich C.** 2011. Orthologues of the anaphase-promoting
671 complex/cyclosome coactivators Cdc20p and Cdh1p are important for mitotic
672 progression and morphogenesis in *Candida albicans*. *Eukaryotic cell* **10**:696-709.
- 673 44. **Garcia-Prieto F, Gomez-Raja J, Andaluz E, Calderone R, Larriba G.** 2010. Role
674 of the homologous recombination genes RAD51 and RAD59 in the resistance of
675 *Candida albicans* to UV light, radiomimetic and anti-tumor compounds and oxidizing
676 agents. *Fungal Genet. Biol.* **47**:433-445.
- 677 45. **Shi QM, Wang YM, Zheng XD, Lee RT, Wang Y.** 2007. Critical role of DNA
678 checkpoints in mediating genotoxic-stress-induced filamentous growth in *Candida*
679 *albicans*. *Mol. Biol. Cell* **18**:815-826.
- 680 46. **Stevenson LF, Kennedy BK, Harlow E.** 2001. A large-scale overexpression screen
681 in *Saccharomyces cerevisiae* identifies previously uncharacterized cell cycle genes.
682 *Proc. Natl. Acad. Sci. U. S. A.* **98**:3946-3951.
- 683 47. **Yu L, Pena Castillo L, Mnaimneh S, Hughes TR, Brown GW.** 2006. A survey of
684 essential gene function in the yeast cell division cycle. *Mol. Biol. Cell* **17**:4736-4747.
- 685 48. **Paffett KS, Clikeman JA, Palmer S, Nickoloff JA.** 2005. Overexpression of Rad51
686 inhibits double-strand break-induced homologous recombination but does not affect
687 gene conversion tract lengths. *DNA repair* **4**:687-698.
- 688 49. **Richardson C, Stark JM, Ommundsen M, Jasin M.** 2004. Rad51 overexpression
689 promotes alternative double-strand break repair pathways and genome instability.
690 *Oncogene* **23**:546-553.
- 691 50. **Kapitzky L, Beltrao P, Berens TJ, Gassner N, Zhou C, Wuster A, Wu J, Babu**
692 **MM, Elledge SJ, Toczyski D, Lokey RS, Krogan NJ.** 2010. Cross-species
693 chemogenomic profiling reveals evolutionarily conserved drug mode of action. *Mol*
694 *Syst Biol* **6**:451.

- 695 51. **Yuen KW, Warren CD, Chen O, Kwok T, Hieter P, Spencer FA.** 2007. Systematic
696 genome instability screens in yeast and their potential relevance to cancer. Proc. Natl.
697 Acad. Sci. U. S. A. **104**:3925-3930.
- 698 52. **Schwartz K, Richards K, Botstein D.** 1997. BIM1 encodes a microtubule-binding
699 protein in yeast. Mol. Biol. Cell **8**:2677-2691.
- 700 53. **Sopko R, Huang D, Preston N, Chua G, Papp B, Kafadar K, Snyder M, Oliver**
701 **SG, Cyert M, Hughes TR, Boone C, Andrews B.** 2006. Mapping pathways and
702 phenotypes by systematic gene overexpression. Mol. Cell **21**:319-330.
- 703 54. **Yoshikawa K, Tanaka T, Ida Y, Furusawa C, Hirasawa T, Shimizu H.** 2011.
704 Comprehensive phenotypic analysis of single-gene deletion and overexpression strains
705 of *Saccharomyces cerevisiae*. Yeast **28**:349-361.
- 706 55. **Andaluz E, Bellido A, Gomez-Raja J, Selmecki A, Bouchonville K, Calderone R,**
707 **Berman J, Larriba G.** 2011. Rad52 function prevents chromosome loss and
708 truncation in *Candida albicans*. Mol. Microbiol. **79**:1462-1482.
- 709 56. **Forche A, Alby K, Schaefer D, Johnson AD, Berman J, Bennett RJ.** 2008. The
710 Parasexual Cycle in *Candida albicans* Provides an Alternative Pathway to Meiosis for
711 the Formation of Recombinant Strains. PLoS Biol **6**:e110.
- 712 57. **Burke D, Gasdaska P, Hartwell L.** 1989. Dominant effects of tubulin overexpression
713 in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **9**:1049-1059.
- 714 58. **Harper JW, Elledge SJ.** 2007. The DNA damage response: ten years after. Mol. Cell
715 **28**:739-745.
- 716 59. **Pfau SJ, Amon A.** 2012. Chromosomal instability and aneuploidy in cancer: from
717 yeast to man. EMBO reports **13**:515-527.
- 718 60. **Negrini S, Gorgoulis VG, Halazonetis TD.** 2010. Genomic instability--an evolving
719 hallmark of cancer. Nat Rev Mol Cell Biol **11**:220-228.

- 720 61. **Kolodner RD, Putnam CD, Myung K.** 2002. Maintenance of genome stability in
721 *Saccharomyces cerevisiae*. *Science* **297**:552-557.
- 722 62. **Bennett CB, Lewis LK, Karthikeyan G, Lobachev KS, Jin YH, Sterling JF, Snipe**
723 **JR, Resnick MA.** 2001. Genes required for ionizing radiation resistance in yeast. *Nat.*
724 *Genet.* **29**:426-434.
- 725 63. **Birrell GW, Giaever G, Chu AM, Davis RW, Brown JM.** 2001. A genome-wide
726 screen in *Saccharomyces cerevisiae* for genes affecting UV radiation sensitivity. *Proc.*
727 *Natl. Acad. Sci. U. S. A.* **98**:12608-12613.
- 728 64. **Chang M, Bellaoui M, Boone C, Brown GW.** 2002. A genome-wide screen for
729 methyl methanesulfonate-sensitive mutants reveals genes required for S phase
730 progression in the presence of DNA damage. *Proc. Natl. Acad. Sci. U. S. A.*
731 **99**:16934-16939.
- 732 65. **Hanway D, Chin JK, Xia G, Oshiro G, Winzeler EA, Romesberg FE.** 2002.
733 Previously uncharacterized genes in the UV- and MMS-induced DNA damage
734 response in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **99**:10605-10610.
- 735 66. **Bellaoui M, Chang M, Ou J, Xu H, Boone C, Brown GW.** 2003. Elg1 forms an
736 alternative RFC complex important for DNA replication and genome integrity. *EMBO*
737 *J.* **22**:4304-4313.
- 738 67. **Mayer ML, Pot I, Chang M, Xu H, Aneliunas V, Kwok T, Newitt R, Aebersold R,**
739 **Boone C, Brown GW, Hieter P.** 2004. Identification of protein complexes required
740 for efficient sister chromatid cohesion. *Mol. Biol. Cell* **15**:1736-1745.
- 741 68. **Tong AH, Evangelista M, Parsons AB, Xu H, Bader GD, Page N, Robinson M,**
742 **Raghibizadeh S, Hogue CW, Bussey H, Andrews B, Tyers M, Boone C.** 2001.
743 Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science*
744 **294**:2364-2368.

- 745 69. **Warren CD, Eckley DM, Lee MS, Hanna JS, Hughes A, Peyser B, Jie C, Irizarry**
746 **R, Spencer FA.** 2004. S-phase checkpoint genes safeguard high-fidelity sister
747 chromatid cohesion. *Mol. Biol. Cell* **15**:1724-1735.
- 748 70. **Huang ME, Rio AG, Nicolas A, Kolodner RD.** 2003. A genomewide screen in
749 *Saccharomyces cerevisiae* for genes that suppress the accumulation of mutations. *Proc.*
750 *Natl. Acad. Sci. U. S. A.* **100**:11529-11534.
- 751 71. **Alabrudzinska M, Skoneczny M, Skoneczna A.** 2011. Diploid-specific [corrected]
752 genome stability genes of *S. cerevisiae*: genomic screen reveals haploidization as an
753 escape from persisting DNA rearrangement stress. *PloS one* **6**:e21124.
- 754 72. **Andersen MP, Nelson ZW, Hetrick ED, Gottschling DE.** 2008. A Genetic Screen
755 for Increased Loss of Heterozygosity in *Saccharomyces cerevisiae*. *Genetics*
756 **179**:1179-1195.
- 757 73. **Kanellis P, Gagliardi M, Banath JP, Szilard RK, Nakada S, Galicia S, Sweeney**
758 **FD, Cabelof DC, Olive PL, Durocher D.** 2007. A screen for suppressors of gross
759 chromosomal rearrangements identifies a conserved role for PLP in preventing DNA
760 lesions. *PLoS genetics* **3**:e134.
- 761 74. **Smith S, Hwang JY, Banerjee S, Majeed A, Gupta A, Myung K.** 2004. Mutator
762 genes for suppression of gross chromosomal rearrangements identified by a genome-
763 wide screening in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.*
764 **101**:9039-9044.
- 765 75. **Strome ED, Wu X, Kimmel M, Plon SE.** 2008. Heterozygous screen in
766 *Saccharomyces cerevisiae* identifies dosage-sensitive genes that affect chromosome
767 stability. *Genetics* **178**:1193-1207.

768 76. **Torres EM, Sokolsky T, Tucker CM, Chan LY, Boselli M, Dunham MJ, Amon A.**
769 2007. Effects of aneuploidy on cellular physiology and cell division in haploid yeast.
770 Science **317**:916-924.

771 77. **Barelle, C.J., Manson, C.L., MacCallum, D.M., Odds, F.C., Gow, N.A. and**
772 **Brown, A.J.** (2004) GFP as a quantitative reporter of gene regulation in *Candida*
773 *albicans*. *Yeast*, 21, 333-340.

774

775

776 **FIGURE LEGENDS**

777 **Figure 1. The BFP/GFP system.**

778 **(A) Artificial heterozygous locus on Chromosome 4.** BFP and GFP expression cassettes
779 were integrated in the gene-free *PGA59/PGA62* intergenic region on chromosome 4. One
780 homolog carries the gene coding for the BFP while the other carries the gene coding for the
781 GFP. Both *BFP* and *GFP* genes are under the control of the promoter of the constitutively,
782 highly-expressed *TDH3* gene and linked to a selection marker. When LOH takes place at this
783 locus, the cells express only one of the fluorescent proteins and LOH events can be detected
784 by flow cytometry. **(B) Flow cytometry analysis of a WT BFP/GFP strain (CEC2683).**
785 Cells were grown overnight in rich media, diluted in 1X PBS and analyzed on a MACSQuant
786 cytometer (Miltenyi Biotech). 10^6 events are displayed.

787

788 **Figure 2. Flow cytometry analysis reveals an increase in LOH in *mec1* null mutants and** 789 **in response to physical and chemical stresses.**

790 **(A) Flow cytometry analysis of homozygous *mec1* deletion strains.** The WT BFP/GFP
791 strain CEC3194 and the *mec1ΔΔ* mutant were grown overnight in YPD at 30°C and analyzed
792 by flow cytometry. **(B) Flow cytometry analysis upon heat treatments.** The WT BFP/GFP

793 strain CEC3989 was grown overnight in YPD at 30°C, 37°C or 39°C and allowed to recover
794 overnight in YPD before flow cytometry analysis. **(C) Flow cytometry analysis in presence**
795 **of the DNA-damaging agent MMS.** The WT BFP/GFP strain CEC3989 was treated with
796 increasing concentrations of MMS for 30 minutes. Cells were then allowed to recover
797 overnight in fresh YPD and analyzed by flow cytometry.

798

799 **Figure 3. SNP typing reveals a shift in the molecular mechanisms leading to LOH upon**
800 **overexpression of some of the candidate genes.**

801 **(A) Map of Chromosome 4 and localization of the BFP/GFP system.** SNPs used for RFLP
802 characterization are indicated with black triangles. The telomere proximal SNP156 is part of a
803 TaqI restriction site. One allele contains the TaqI site while the other does not. The
804 centromere proximal SNP95 is located in the middle of a AluI restriction site. One allele
805 contains the AluI site while the other does not. The heterozygosity status of these SNPs
806 provides information on the molecular mechanisms that give rise to LOH. **(B) SNP-RFLP**
807 **typing in the Bfp- Gfp+ cells.** Histogram presenting the proportion of BIR/MCO or
808 chromosome truncation vs chromosome loss in the Bfp- Gfp+ population. BIR/MCO or
809 chromosomal truncation events correspond to isolates that have maintained a heterozygous
810 SNP95 and displayed a homozygous SNP156. Chromosome loss events correspond to isolates
811 in which both SNP95 and SNP156 became homozygous. Gene conversion events correspond
812 to isolates in which both SNP95 and SNP156 remained heterozygous. **(C) Proportion of**
813 **small vs large colonies giving rise to true mono-BFP cells.** **(D) SNP-RFLP typing in the**
814 **small colony-derived Bfp+ Gfp- cells.** **(E) SNP-RFLP typing in the large colony-derived**
815 **Bfp+ Gfp- cells.** **(F) SNP-RFLP typing in the Bfp+ Gfp- cells.** The normalized percentage
816 of each molecular mechanism, for instance gene conversion, was calculated as follow:
817 normalized percentage GC = (% small colonies × % GC in small colonies) + (% large

818 colonies \times % GC in large colonies). **(G) Heterozygous and homozygous SNV density maps**
819 **of 1 large and 3 small colony-derived cells that have undergone LOH at the BFP/GFP**
820 **locus on Ch4.** The number of heterozygous SNVs in each 10kb region along Ch4 was
821 computed and is represented, revealing a LOH by gene conversion in the large colony-derived
822 cells and by chromosome loss and reduplication on Ch4 in the small colony-derived cells.
823 Sequencing reads were mapped on the HapA or HapB reference sequences and yielded
824 similar heterozygous SNP density maps. Only the map resulting from mapping on HapB is
825 shown (Left panel). The number of homozygous SNVs in each 10kb region along HapA
826 (Middle panel) or HapB (Right panel) was computed and is represented. The high density of
827 homozygous SNVs upon mapping on HapB indicates that chromosome loss, in the small
828 colony-derived cells, affected Ch4B.

829

830 **TABLE FOOTNOTES**

831 **Table 1. Yeast strains used in this study.**

832

833 **Table 2. LOH increase in the *mec1* null mutant and in response to physical and chemical**
834 **stresses.**

835 ^a Three independent colonies were tested for the *mec1* mutants. Six independent cultures were
836 tested for the physical and chemical stresses. Standard error mean (SEM) were determined
837 with GraphPad Prism Software.

838 ^b Ratio between the LOH frequency in a given mutant and that in the parental strain.

839 ^c Ratio between the LOH frequency in presence of the stress and that in absence of the stress.

840 ^d A Mann–Whitney test was performed to test if the BFP loss frequency in the different
841 mutants were statistically different from the rate observed in the control strains/conditions.

842

843 **Table 3. LOH quantification in the 4 candidate overexpression mutants validated in the**
844 **secondary LOH screen.**

845 ^a Ratio between the LOH frequency in a given mutant and that in the control strain CEC3989.