論文 / 著書情報
Article / Book Information

Title	Homology length dictates the requirement for Rad51 and Rad52 in gene targeting in the Basidiomycota yeast Naganishia liquefaciens
Authors	Maierdan Palihati, Hideo Tsubouchi, Bilge Argunhan, Rei Kajitani, Omirgul Bakenova, Yong-Woon Han, Yasuto Murayama, Takehiko Itoh, Hiroshi Iwasaki
Citation	Current Genetics, Vol. 67, no. 6, pp. 919-936
Pub. date	2021, 12
Note	This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature ' s AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/s00294-021-01201-3.

Homology length dictates the requirement for Rad51 and Rad52 in gene targeting in the Basidiomycota yeast *Naganishia liquefaciens*

Maierdan Palihati[1,2], Hideo Tsubouchi[1,2,*], Bilge Argunhan[2,3], Rei Kajitani[1], Omirgul Bakenova[1], Yong-Woon Han[1,4], Yasuto Murayama[1,5], Takehiko Itoh[1], and Hiroshi Iwasaki[1,2,*]

[1]School and Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8503, Japan.

[2]Institute of Innovative Research, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8503, Japan.

[3]Present address: Section of Structural Biology, Faculty of Medicine, Imperial College London, South Kensington, London SW7 2AZ, UK.

[4]Present address: Laboratory for Integrative Genomics, RIKEN Center for Integrative Medical Sciences, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan.

[5]Present address: Center for Frontier Research, National Institute of Genetics, 1111, Yata, Mishima, Shizuoka 411-8540, Japan.

*Correspondence: htsubouchi@bio.titech.ac.jp and hiwasaki@bio.titech.ac.jp

ORCID

Maierdan Palihati: 0000-0002-9859-3873

Hideo Tsubouchi: 0000-0003-0814-8432

Rei Kajitani: 0000-0002-5013-0052

Bilge Argunhan: 0000-0002-6023-7654

Hiroshi Iwasaki: 0000-0002-0153-687

Abstract

Here, we report the development of methodologies that enable genetic modification of a Basidiomycota yeast, *Naganishia liquifaciens*. The gene targeting method employs electroporation with PCR products flanked by an 80 bp sequence homologous to the target. The method, combined with a newly devised CRISPR-Cas9 system, routinely achieves 80% gene targeting efficiency. We further explored the genetic requirement for this homologous recombination (HR)-mediated gene targeting. The absence of Ku70, a major component of the non-homologous end joining (NHEJ) pathway of DNA double-strand break repair, almost completely eliminated inaccurate integration of the marker. Gene targeting with short homology (80 bp) was almost exclusively dependent on Rad52, an essential component of HR in the Ascomycota yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. By contrast, the RecA homolog Rad51, which performs homology search and strand exchange in HR, plays a relatively minor role in gene targeting, regardless of the homology length (80 bp or 1 kb). The absence of both Rad51 and Rad52, however, completely eliminated gene targeting. Unlike Ascomycota yeasts, the absence of Rad52 in *N. liquefaciens* conferred only mild sensitivity to ionizing radiation. These traits associated with the absence of Rad52 are reminiscent of findings in mice.

Keywords

Basidiomycota, CRISPR-Cas9, DNA repair, gene targeting, homologous recombination, *Naganishia liquefaciens*

Declarations

Funding

This work was supported in part by Grants-in-Aids for Scientific Research (A) (18H03985 to H.I.), for Scientific Research (B) (18H02371 to H.T.), for Young Scientists (B) (17K15061 to B.A.), and for Early-Career Scientists (20K15713 to B.A.) from the Japan Society for the Promotion of Science (JSPS). H.T. is also supported by the Takeda Science Foundation.

Conflicts of interest/Competing interests

The authors have no conflict of interest to declare.

Ethics approval

Not applicable

Consent to participate

Not applicable

Consent for publication

All authors have agreed to publish this manuscript.

Availability of data and material

All relevant data are included in the manuscript. Requests for reagents or further information should be directed to H.T. (htsubouchi@bio.titech.ac.jp) or H.I. (hiwasaki@bio.titech.ac.jp).

Code availability

Not applicable

Authors' contributions

M.P., H.T. and O.B. conducted experiments. Y.-W.H., R.K. and T.I. are responsible for sequencing analyses. M.P., H.T. and H.I. are responsible for conceptualization and project design. H.T., B.A., Y.M. and H.I. supervised the study. M.P., H.T. and H.I. are responsible for data analysis. H.T., B.A., M.P. and H.I wrote the manuscript.

Introduction

Ascomycota and Basidiomycota represent two major phyla in the fungal kingdom, which split from a common ancestor around 1 billion years ago (Hedges et al. 2004; Blackwell 2011). In general, Ascomycota produce an ascus, which is a sac carrying spores, and include the *Penicillium*, *Candida*, and *Aspergillus genera*, as well as brewer's and baker's yeasts (Stajich et al. 2009). Several major model organisms, such as S*accharomyces cerevisiae*, *Schizosaccharomyces pombe* and *Neurospora crassa*, belong to the Ascomycota phylum.

Basidiomycota produce basidia, which are club-shaped structures that bear spores (Stajich et al. 2009). Basidiomycota include mushrooms, bracket fungi, polypores, and other fungi. Unlike Ascomycota, which are usually monokaryotic for most of their lives, Basidiomycota have long-lived dikaryotic states, and their hyphae grow and divide as dikaryon with two genetically distinct nuclei. Overall, Basidiomycota biology has been much less explored than Ascomycota.

Basidiomycota also include monocellular species that fall under the category of yeasts; the human pathogen *Cryptococcus neoformans* (Mochizuki et al. 1987) and the plant pathogen *Ustilago maydis* (O'Donnell and McLaughlin 1984) are such examples, and these organisms are amenable to similar molecular genetics approaches that have proven hugely successful in *S. cerevisiae* and *S. pombe* (Heitman et al. 2010). In particular, *C. neoformans* has recently emerged as a representative model organism of the Basidiomycota phylum. Despite similarities to *S. cerevisiae* in cell morphology and genome size, *C. neoformans* shares some biological features with Metazoan species. For example, the gene organization is much more complex than that of *S. cerevisiae*, with frequent introns (~ 5 introns per gene) and complex gene regulation including alternative splicing and antisense transcription (Loftus et al. 2005). Unlike *S. cerevisiae*, centromeres are not clustered in premitotic cells, which progressively go through ordered assembly of kinetochores towards mitosis. Furthermore, the

nuclear envelope partially breaks open as chromosomes separate. These traits are also reminiscent of Metazoan mitosis (Kozubowski et al. 2013).

A non-pathogenic *Cryptococcus* species, named strain N6, was originally isolated from the deep-sea sediment at a depth of ~ 6,500 m in the Japan Trench (Abe et al. 2001; Miura et al. 2001). This strain N6 was identified as *Cryptococcus liquefaciens* based on a phylogenetic analysis using the sequence of the divergent D1/D2 domain of 26S rDNA (Abe et al. 2006). However, recent studies have proposed that *C. liquefaciens* actually belongs to the genus *Naganishia* in the order Filobasidiales, while *C. neoformans* belongs to the order Tremellales (Liu et al. 2015). We recently drafted the whole genome sequence and gene structure of the strain N6 and used the obtained genomic data to construct a maximum-likelihood phylogenetic tree in the class Tremellomycetes (Han et al. 2020). While the genome annotation is ongoing, this result further supported the classification of strain N6 within the clade of the genus *Naganishia*, leading us to propose that this strain be named *Naganishia liquefaciens* N6 (Han et al. 2020). We do not currently have information about the ploidy or sexual cycle of this strain. Some features of *N. liquefaciens*, such as its non-pathogenicity, short doubling time, and simple culturing conditions, are similar to those for *S. cerevisiae*, which would potentially make this organism a suitable model to study Basidiomycota biology (Abe et al. 2001, 2006).

Here, we report the development of methodologies that enable genetic modification of *N. liquefaciens*. We have established an efficient gene targeting method whereby PCR amplification of a drug-resistant cassette with primers containing an 80 bp sequence homologous to the target, combined with CRISPR-Cas9, routinely achieves ~ 80% gene targeting efficiency. We further explored the genetic requirement for this homologous recombination (HR)-mediated gene targeting. The absence of Ku70, a major component of the non-homologous end joining (NHEJ) pathway of DNA double-strand break (DSB) repair, almost completely eliminated inaccurate integration of the marker. On the other hand, the

absence of Rad52, an essential component of HR in *S. cerevisiae* and *S. pombe*, eliminated gene targeting mediated by 80 bp homology. However, with a homology length of 1 kb, 10-40% of gene targeting was seen, suggesting that Rad52 is not absolutely required for gene targeting. The absence of Rad51, a RecA homolog that performs homology search and strand exchange in HR, led to a modest reduction (30-50% of the wild type strain) in gene targeting regardless of homology length (80 bp or 1 kb). The absence of both Rad51 and Rad52, however, completely eliminated gene targeting. Thus, in *N. liquifaciens*, gene targeting with short homology exclusively requires Rad52, but not Rad51, while Rad51 and Rad52 redundantly support gene targeting when ample homology exists.

Materials and methods

Strains and growth conditions

N. liquefaciens strain N6 (MP1), which was one of stocks originating from samples collected in the Japan Trench (Abe et al. 2001), was used as a wild-type strain. Deletion strains constructed in this study are listed in Table S1. *N. liquefaciens* stains were cultured essentially in the same manner as *S. cerevisiae* (Abe et al. 2001, 2006; Amberg et al. 2015). Briefly, they were grown in YPD (1% w/v yeast extract, 2% w/v Bacto peptone, and 2% w/v glucose) or the synthetic drop-out (SD) medium (0.17% yeast nitrogen base without amino acids and ammonium sulfate, 2% glucose, 0.5% ammonium sulfate and 0.2% drop-out mix) at 30°C, unless indicated otherwise. Nourseothricin sulfate (GoldBio), G418 (Nacalai Tesque, Japan) and hygromycin B (InvivoGen) were used for selection at final concentrations of 100 μg/ml, 100 μg/ml and 25 μg/ml, respectively. Deletion strains were constructed by replacing the gene of interest with a marker cassette (see below) encoding nourseothricin N-acetyl transferase (*NAT*), hygromycin B phosphotransferase (*HYG*) or aminoglycoside 3'-phosphotransferase (*NEO*).

Identification of *N. liquefaciens* genes encoding homologs of DSB repair proteins

To identify *N. liquefaciens* genes encoding DSB repair proteins, the amino acid sequence for Rad51, Rad52, or Ku70 of *S. cerevisiae* was used as a query to search for its *N. liquefaciens* homolog using BLASTP against the database of *N. liquefaciens* predicted proteins (Genbank PRJDB10172). Multiple sequence alignment was then performed using Clustal Omega with default settings (Sievers et al. 2011) (Fig. S1).

DNA primers used for PCR

DNA primers used in this study are listed in Table S2 (synthesized by Eurofins or Fasmac).

Genomic DNA purification

N. liquefaciens N6 cells cultured in 5 ml of YPD were collected (3,500 × g, 5 min) and washed in protoplast buffer which is McIlvain buffer [pH 6.0] (0.1 M citric acid solution and 0.2 M disodium hydrogenphosphate at the ratio of 36.8 : 63.2 [v/v]) containing 0.3 M sodium tartrate. The cell pellet was resuspended in 0.5 ml westase solution (0.5% [w/v] of westase [Takara Bio, Japan] in protoplast buffer) and incubated for 1 h at 30°C. Cells were washed once in 1 ml of 50 mM Tris-HCl (pH 7.5), resuspended in 500 μl of TE 50:20 (50 mM Tris-HCl (pH 7.5) 20 mM EDTA), then mixed well with 50 μl of 10% SDS. After 30 min incubation at 65 °C, 200 μl of 5 M potassium acetate was added to the sample and stored on ice for 15 min. The sample was centrifugated at 15,000 rpm for 10 min to obtain the cell lysate, which was used for genomic DNA preparation by essentially the same procedure as for *S. cerevisiae* (Amberg et al. 2015). Alternatively, genomic DNA was purified using Dr. GenTLE™ High Recovery kit (Takara Bio, Japan) from approximately $1x10^8$ cells cultured in YPD. s

Plasmid construction

The *NAT* gene in pFA6a-natMX6 (Hentges et al. 2005) was amplified by PCR (primers Pr-22 and Pr-25). The *TEF1* promoter (*Ptef1*) of *N. liquefaciens* was amplified by PCR using the genomic DNA as a template (primers Pr-36 and Pr-23). The resultant two fragments were fused via overlap-extension PCR (Higuchi et al. 1988) using primers Pr-36 and Pr-25. The resultant *Ptef1-NAT* fragment was replaced with the XbaI-BamHI fragment of pBluescript II SK(+), giving pBS-Ptef1-NAT.

The *NEO* gene in pcDNA3 (Invitrogen) was amplified by PCR (primers Pr-200 and Pr-201). The *ACT1* promoter of *N. liquefaciens* was amplified by PCR using the genomic DNA as a template (primers Pr-198 and 199). The resultant two fragments were fused by

overlap-extension PCR using primers Pr-198 and Pr-201. The resultant *Pact1-NEO* fragment was replaced with the XbaI-BamHI fragment of pBluescript II SK(+), giving pBS-Pact1-NEO.

The *HYG* gene in pFA6a-hygMX6 (Hentges et al. 2005) was amplified by PCR (primers Pr-203 and Pr-25). The *ACT1* promoter of *N. liquefaciens* was amplified by PCR using the genomic DNA as a template (primers Pr-198 and Pr-202). The resultant two fragments were fused by overlap-extension PCR using primers Pr-198 and Pr-25. The resultant *Pact1-HYG* fragment was replaced with the XbaI-BamHI fragment of pBluescript II SK(+), giving pBS-Pact1-HYG.

A Cas9 expression construct was created by combining the endogenous *ACT1* promoter (primers Pr-198 and Pr-51) and a human codon optimized SV40 NLS- containing Cas9 fragment (primers Pr-50 and Pr-39) from pX330 (Cong et al. 2013) by overlap PCR (primers Pr-198 and Pr-39). The fused fragment was cloned at the pBluescript II SK (+) XbaI site to create pBS Pact1:Cas9. To construct the gRNA expression cassette for the production of single RNA, an endogenous *N. liquefaciens* U6 promoter (primers Pr-43 and Pr-44) was fused to the gRNA scaffold fragment (primers Pr-41 and Pr-42) from pX330 by PCR using the primers Pr-43 and Pr-42. The resulting gRNA expression cassette was cloned at the pBS Pact1:Cas9 HindIII site to create pBS Pact1:Cas9 Pu6:gDNA, giving pM101.

For designing and cloning DNA for gRNA expression, we followed the protocols described previously (Wang et al. 2014) to eliminate less favorable gRNA targets. Since the U6 promoter was employed to drive gRNA expression, a 20 nt sequence that has G at the 5' end of the target sequence (Mali et al. 2013) followed by a PAM sequence was manually chosen. Oligos were designed and synthesized with a 20 nt gRNA sequence GN_{19} (N for any base) containing 5' overhang TTTC and 3' overhang CAAA to facilitate cloning into the BbsI site of pM101. To anneal oligos, 1 µl of each oligo (100 µM) was mixed with 8 µl of

annealing buffer (10 mM Tris [pH 7.5], 50 mM NaCl, and 1 mM EDTA), boiled for 5 min at 95°C, and cooled to room temperature. The annealed oligos were ligated at the BbsI site of pM101. Targeting was performed with the following plasmid templates and primer pairs, with the primer pairs shown in brackets. *HIS3*, *ADE2*, *LEU2*, *TRP2, RAD51, RAD52* and *KU70* loci, *NAT*, *NEO* and *HYG* are: pM101-gHIS3 (Pr-290, Pr-291), pM101-gADE2 (Pr-284, Pr-285), pM101-gLEU2 (Pr-310, Pr-311), pM101-gTRP2 (Pr-312, Pr-313), pM101-gRAD51 (Pr-346, Pr-347), pM101-gRAD52 (Pr-361, Pr-362), pM101-gKU70 (Pr-500, Pr-501), pM101-gNAT (Pr-415, Pr-416), pM101-gNEO (Pr-620, Pr-621) and pM101-gHYG (Pr-622, Pr-623).

Transformation of *N. liquefaciens*

A fresh colony of *N. liquefaciens* was cultured in 5 ml of YPD medium at 30°C with shaking for ~ 15 h. The culture was diluted in 50 ml of YPD to yield an OD_{600} of 0.2–0.3. Cells were then cultured for 4 ~ 6 h until they reached an OD_{600} of 0.6–1.0. Cells were harvested by centrifugation (3,500 × g, 5 min, 4°C). The pelleted cells were washed with ice-cold water and then with electroporation buffer (EB: 10 mM Tris-HCl [pH 7.5], 1 mM $MgCl_2$, and 270 mM sucrose). The cells were resuspended in 10 ml of EB containing 4 mM DTT. After incubation on ice for 15–30 min, cells were collected, washed with EB, and resuspended in 200 μl of EB. The cell suspension (45 μl) was mixed with 5 μl of DNA (3 μg) in a 0.2-cm electroporation cuvette (BioRad) and used for transformation by electroporation (BioRad Gene Pulser, 0.75 KV, 25 μF, and ∞ Ω). The electroporated cells were then suspended in 1 ml of YPD and incubated at 30°C for 2 h before being plated onto the appropriate selection medium. The plates were typically incubated for 3 days at 30°C. A chemical transformation method using lithium acetate (Ito et al. 1983), a typical method for transforming *S. cerevisiae*, yielded no transformants.

Preparation of gene targeting fragments

Gene targeting fragments with 80 bp homology arms were prepared as follows. PCR was carried out using an appropriate primer set (see below) and pBS-Ptef1-NAT or pBS-Pact1-NEO as a template. The primer set consisted of two primers, forward and reverse, whose size is 100 nt (20 nt for annealing to the template and 80 nt for gene targeting). The PCR products were purified using the MonoFas DNA purification kit (GL Sciences, Japan).

Gene targeting fragments with 0.5 kb or 1 kb homologous arms were prepared as follows. The 5' and 3' homologous arms were amplified by PCR with a primer set and the *N. liquefaciens* genomic DNA as a template. The sequence of each primer is shown in Table. S2. For 0.5 kb homologous arms: *HIS3* (Pr-160-Pr-157, Pr-161-Pr-158); *ADE2* (Pr-52-Pr-54, Pr-53-Pr-55); *LEU2* (Pr-377-Pr-375, Pr-378-Pr-376); *TRP2* (Pr-299-Pr-399, Pr-302-Pr-400). For 1 kb homologous arms: *HIS3* (Pr-156-Pr-157, Pr-159-Pr-158); *ADE2* (Pr-63-Pr-54, Pr-394-Pr-55); *LEU2* (Pr-373-Pr-375, Pr-374-Pr-376); *TRP2* (Pr-397-Pr-399, Pr-398-Pr-400). The 5' and 3' homologous arms were then fused to a drug marker by overlap-extension PCR (Higuchi et al. 1988). The PCR products were purified using MonoFas DNA purification kit (GL Sciences, Japan).

Gene targeting fragments with 0.5 kb or 1 kb homology arms for split marker replacement were prepared as follows. The 5' and 3' homologous arms were amplified by PCR with a primer set and the *N. liquefaciens* genomic DNA as a template. The sequence of each primer is shown in Table. S2. For 0.5 kb homologous arms: *HIS3* (Pr-160-Pr-157, Pr-161-Pr-158); *ADE2* (Pr-52-Pr-54, Pr-53-Pr-55); *LEU2* (Pr-377-Pr-375, Pr-378-Pr-376); *TRP2* (Pr-299-Pr-399, Pr-302-Pr-400). For 1 kb homologous arms: *HIS3* (Pr-156-Pr-157, Pr-159-Pr-158); *ADE2* (Pr-63-Pr-54, Pr-394-Pr-55); *LEU2* (Pr-373-Pr-375, Pr-374-Pr-376); *TRP2* (Pr-397-Pr-399, Pr-398-Pr-400). The *NAT* gene on pBS Ptef1-NAT was amplified by PCR with primers Pr-90 and Pr-91. In the third round of PCR, the 5' and 3' homology arms were then fused to the 5' and 3' regions of the *NAT* gene amplified above, respectively, by overlap

PCR, yielding the 5' split marker and 3' split marker fragments. All the PCR products were purified using MonoFas DNA purification kit (GL Sciences, Japan).

Strain construction

A gene of interest was replaced with a drug resistant marker by the one-step replacement method or combined with the CRISPR/Cas9 method using the plasmids described above. MP108 (*rad51*::*NAT ku70*::*NAT*) was constructed by transforming MP87 (*rad51*::*HYG ku70*::*NEO*) using a mixture of pM101-gHYG targeting the *HYG* cassette, the *rad51*::*NAT* donor fragment, pM101-gNEO targeting the *NEO* cassette and the *ku70*::*NAT* donor fragment. nourseothricin-resistant clones sensitive to both HYG and G418 were selected and gene replacement was confirmed by PCR (Pr-144 and Pr-145 for checking *rad51*::*NAT*, Pr-504 and Pr-505 for *ku70*::*NAT*). MP113 (*rad51*::*NAT rad52*::*HYG ku70*::*NAT*) was constructed by transforming MP108 (*rad51*::*NAT ku70*::*NAT*) with the mixture of pM101-gRAD52 targeting *RAD52* with the *rad52*::*HYG* donor fragment. HYG-resistant clones were selected and gene replacement was confirmed by PCR (Pr-174-Pr175 for checking *rad52::HYG*).

Ectopic integration of foreign DNA into the *TRP2* locus for complementation analysis

pM175, the plasmid for ectopic integration, consists of the following three components. (i) the *TPR2* gene (2.2 kb); (ii) a gene to be integrated (Gene X); (iii) the 200 bp fragment downstream of *TRP2*. The genomic fragment carrying the *TRP2* gene was amplified by PCR (primers Pr-394 and Pr-395) using *N.liquefaciens* genomic DNA as a template, which was cloned at the SalI site of pBluescript II SK (+) to construct pBS *TRP2*. The 200 bp fragment downstream of *TRP2* was amplified (primers Pr-475 and Pr-476) and replaced with the BamHI-XbaI fragment of pBS *TRP2* to obtain pM175. *GENE X* can be cloned at multiple cloning sites. The fragment carrying *TRP2* and *GENE X* can be released by digesting the

construct with NotI, which is to be used with the CRSPR/Cas9 construct (pM101-gNAT) targeting the *NAT* marker integrated at the *TRP2* locus.

Cloning and sequence determination of Rad51, Rad52 and Ku70 cDNA

Total RNA was extracted from ~1 x 10^8 cells at late-log phase using the Nucleospin RNA kit (Macherey-Nagel). cDNA was synthesized using ReverTra Ace-α- (Toyobo) according to the manufacturer's instructions with the following materials and conditions; 0.25 μg of the total RNA was used with primers listed in Table S2 (*RAD51*, Pr-459; *RAD52*, Pr-551; *KU70*, Pr-553) or 10 pmol of oligo (dT) 20, followed by 20 min of 1st strand synthesis at 42ºC. cDNAs for Rad51, Rad52 and Ku70 were PCR-amplified using primers listed in Table S2: *RAD51*, Pr-458-Pr-459; *RAD52*, Pr-550-Pr-551; *KU70*, Pr-552-Pr-553. The amplified cDNA was then cloned into the XbaI site of pBluescript II SK(+), giving pBS-RAD51, pBS-RAD52 and pBS-KU70 respectively. Fidelity of the cloned cDNA was confirmed by sequencing.

Evaluation of gene targeting efficiency

We considered that the drug-resistant transformants had gone through correct gene targeting if they also showed an auxotrophic phenotype associated with gene disruption (e.g., a *his3::NAT* transformant should show both nourseothricin resistance and histidine auxotrophy). The transformants where the *ADE2* gene was replaced by *NAT* formed pink colonies like the *ade2* mutant of *S. cerevisiae*. Gene targeting efficiency was expressed as the ratio of the number of transformants showing auxotrophic phenotypes divided by the number of all antibiotic resistant transformants. Difference in gene targeting efficiencies was evaluated using unpaired two-tailed t-test.

Fourteen random transformants that showed the genetic trait for accurate gene targeting in each strain background were further examined by PCR to amplify the area containing the integration target site using the following primers. For 1 kb homologous arms:

HIS3 (Pr-195-Pr-196); *ADE2* (Pr-122-Pr-74). For 80 bp homologous arms: *HIS3* (Pr-410-Pr-411); *ADE2* (Pr-257-Pr-258). Among 14 transformants obtained with 80 bp homologous arms, four transformants were further examined for their 5' and 3' integration junctions by sequencing PCR amplicons with forward (Pr-410) and reverse (Pr-411) primers. DNA sequencing was performed using BigDye Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems).

Genomic DNA sequences of the genes used in this study

Genomic DNA sequence of the genes used in this study are available in the DDBJ/EMBL/GenBank databases under the following accession numbers: MT185598 for *HIS3*, MT185599 for *ADE2*, MT185600 for *LEU2*, MT185601 for *TRP2*, MT185602 for *RAD51*, MT185603 for *RAD52,* MT185604 for *KU70*, MT210101 for *ACT1* and MT210102 for *TEF1*. The draft genome sequence of *N. liquefaciens* is available in the DDBJ/EMBL/GenBank databases (accession number BLZA00000000.1) (Han et al. 2020).

Results

Establishing a transformation system for *Naganishia* yeast

Securing a genetic marker is the first step for genetic manipulation of a given organism. We employed three pairs of a cytotoxic drug and a cognate gene that confers resistance to the host (Fig. 1a): nourseothricin and *NAT* (nourseothricin N-acetyl transferase); hygromycin B and *HYG* (hygromycin B phosphotransferase); and G418 and *NEO* (aminoglycoside 3'-phosphotransferase). To express these genes in *N. liquefaciens*, 990 bp directly upstream of the start codon of the translation elongation factor 1-alpha (*Ptef1*) gene was fused to the coding sequence of *NAT*. Similarly, 1 kb directly upstream of the start codon of the actin gene, *ACT1*, was fused to the coding sequences of *NEO* and *HYG*. We first sought to replace the *HIS3* gene with *NAT* using a construct where the *NAT* marker is flanked by 1 kb sequences corresponding to the upstream or downstream regions of the *HIS3* coding sequence. Electroporation was used to deliver the gene targeting DNA fragments into the cell. We examined a range of voltages, from 0.25 to 1.25 kV, and found that 0.75 kV yielded the most transformants (Fig. 1b). Under this condition, approximately 10-to-50 stable *NAT*-resistant colonies were routinely obtained using 3 μg of targeting DNA fragments and 7.5×10^7 cells. About 25% of the transformants obtained as nourseothricin-resistant also exhibited histidine auxotrophy (Fig. 1c, Fig. S2). This suggests that gene targeting in *N. liquefaciens* is not as efficient as in *S. cerevisiae* (Amberg et al. 2015).

Gene targeting efficiency depends on the length of homologous arms

We further investigated the relationship between the length of homologous arms and gene targeting. *his3::NAT* deletion constructs with three different arm lengths (80 bp, 500 bp, and 1 kb) were tested for their gene targeting efficiency. As described above, gene targeting

efficiency is around 25% with 1 kb of flanking homology (Fig. 1c). The efficiency with 500 bp homologous arms was equivalent to that with 1 kb homology, while 80 bp homology reduced the targeting efficiency to approximately half of that observed with 1 kb homology.

To further establish a relationship between gene targeting efficiency and homology length, several other loci were examined, namely *ADE2*, *LEU2,* and *TRP2* (Fig. 1d-f). Overall, a similar trend to the *HIS3* locus was seen at these loci. The highest targeting efficiency (10-20%) was obtained with a homology length of 1 kb, while the efficiency was slightly less (10-15%) with 0.5 kb homology. The efficiency was the lowest (5-10%) with 80 bp homology. Together, these data reveal a modest correlation between the length of homology sequence attached to the targeting construct and the actual targeting efficiency.

Gene targeting with 80 bp homology to the *HIS3* locus was employed to compare gene targeting efficiency with the *NAT*, *NEO,* and *HYG* markers. The *NEO* marker supported the formation of a similar number of transformants as *NAT,* while it was lower with the *HYG* marker (~20% of *NAT;* Fig. S3A). The gene targeting efficiency of *HYG* was also lower than the other two markers (~50% reduction; Fig. S3B).

Split marker transformation increases gene targeting efficiency

It has been reported that gene targeting is improved by the split marker method (Fairhead et al. 1996; Fu et al. 2006; Lin et al. 2015), where a marker is split into two fragments, each carrying only part of the drug selection marker. These two fragments share some overlap, which allows them to recombine to form a fully functional, drug-resistance gene if a cell takes up both fragments (Fig. 2a). Although the split marker transformation was originally developed in *S. cerevisiae* (Fairhead et al. 1996), it was also found to be effective in *C. neoformans* (Fu et al. 2006; Lin et al. 2015). Two gene targeting constructs, each carrying an 80 bp arm homologous to the upstream or downstream region of the *HIS3* coding sequence and also sharing part of the *NAT* marker (440 bp), were used together for transformation. Gene targeting efficiency was ~40%, which is approximately twofold higher

than that with a normal intact marker (Fig. 2b). However, this method led to a substantial reduction in transformation efficiency, with only 3-to-10 transformants obtained with 3 μg of targeting DNA as opposed to 10-to-50 transformants with the intact marker.

CRISPR-Cas9 expression promotes efficient gene targeting

The overall low efficiency of gene targeting and transformation prompted us to apply the CRISPR/Cas9 system to *N. liquefaciens* (Mashiko et al. 2013; Jacobs et al. 2014; Arras et al. 2016). The *CAS9* gene was placed under the control of the promoter sequence of the *N. liquefaciens ACT1* gene, which supports constitutive gene expression. A gRNA sequence was placed under the control of a U6-like promoter of *N. liquefaciens*. The U6 promoter, a strong promoter for RNA polymerase III, is typically used for gRNA transcription. A 20 nt target sequence can be cloned at the BbsI site so that gRNA targeting a locus of interest is expressed. This CRISPR/Cas9 vector (pM101) was combined with the previously used gene targeting fragments, now serving as a donor, and gene targeting was directed to the loci tested above.

We first chose the *HIS3* locus for gene targeting with the construct with 1 kb homologous arms combined with the expression of Cas9-gHIS3 (Cas9 with a gRNA targeting *HIS3*). Among the nourseothricin-resistant transformants, ~80% also exhibited histidine auxotrophy (Fig. 3a). Similarly, by using the CRISPR/Cas9 system, highly efficient gene targeting (70-80%) was achieved at the *ADE2*, *LEU2*, *TRP2* loci (Fig. 3b-d). Furthermore, even when the length of the homologous arms was reduced to 500 bp, similarly efficient gene targeting (70-80%) was seen at the *HIS3*, *ADE2*, *LEU2* and *TRP2* loci (Fig. 3a-d). Next, we reduced the homology length even further to 80 bp, which is small enough to be synthesized as part of a primer used to PCR-amplify a drug-resistance marker. Remarkably, the gene targeting efficiency still remained as high as that with 1 kb homologous arms (70-80%) at all the loci tested (Fig. 3a-d).

We then examined the accuracy of gene targeting more closely, since the histidine auxotrophy phenotype could also arise from an event not associated with HR-mediated gene targeting (e.g., marker insertion into a DSB caused at the *HIS3* locus, or coincidental DSB-induced mutagenesis at *HIS3* and an ectopic integration of the marker). From the experiment where *HIS3* was targeted with the 1 kb homology construct, we randomly picked 14 nourseothricin-resistant, histidine auxotrophic transformants and examined their *HIS3* locus by diagnostic PCR (Fig. 3e (i)). All 14 transformants showed a 3.5 kb band, consistent with the expected amplicon size for HR-mediated marker replacement. We also examined by PCR if the plasmid expressing Cas9-gHIS3 (pM101-gHIS3) is maintained in transformant cells, possibly via integration into the genome. All 14 transformants tested above failed to show any amplification for the Cas9 ORF (Fig. 3e (ii)), suggesting that the Cas9 expression plasmid taken up by cells is eventually lost during colony formation. We also examined transformants from the experiment where *HIS3* was targeted with the construct with 80 bp homologous arms. Again, from the 14 randomly picked, nourseothricin-resistant, histidine auxotrophic colonies, all showed a 1.9 kb band, consistent with the expected amplicon size for HR-mediated marker replacement (Fig. 3f). Taken together, CRISPR-Cas9 drastically increased the efficiency of gene targeting in *N. liquefaciens* even when the donor DNA carried as little as 80 bp homology.

A system that facilitates ectopic integration of foreign DNA at the *TRP2* locus enables complementation analysis

Complementation testing is of critical importance to establish a causal relationship between a phenotype and a genotype. In *N. liquefaciens*, however, a plasmid system is not available. Thus, we decided to develop a system that facilitates the integration of foreign DNA into the *TRP2* locus by taking advantage of our efficient Cas9 system. This system utilizes a base-strain where the native *TRP2* is replaced with *NAT*. A plasmid was constructed to facilitate

integration of foreign DNA at *trp2::NAT* (Fig 4a). This plasmid (pM175) contains the upstream region of the *TRP2* locus along with the *TRP2* gene itself, a multiple cloning site where a gene of interest (designated as *GENE X*) can be inserted, and the sequence downstream of *TRP2*. The whole fragment can be released from the plasmid by NotI digestion and used for transformation along with the Cas9 plasmid that induces a DSB inside the *NAT* marker at the *TRP2* locus. The Cas9-induced DSB facilitates HR between the *TRP2* locus and the donor fragment containing *GENE X*.

To examine if this system can be used for complementation testing, a gene encoding the Rad51 homolog in *N. liquefaciens* was identified (see below, and Materials and methods) and deleted with the *NEO* marker. Rad51 is the central player in HR, and cells become sensitive to DNA damage in its absence due to defects in recombinational DNA repair (see below). A *trp2::NAT rad51::NEO* strain was transformed with the plasmid expressing Cas9 and the donor fragment carrying the *RAD51* gene. Transformants showing both tryptophan prototrophy and nourseothricin sensitivity were further analyzed by PCR for the correct integration of the *TRP2-RAD51* fragment at the *TRP2* locus (Fig. 4b,c). The correct integrant restored ultraviolet light (UV) resistance to a level indistinguishable from the wild type strain, indicating that the *rad51* null mutant was indeed complemented by the ectopically integrated wild-type *RAD51* gene (Fig. 4d).

Identification of *RAD51*, *RAD52* and *KU70* homologs in *N. liquefaciens*

Next, we sought to investigate the mechanistic relationship between DSB repair and gene targeting, and therefore focused on three proteins: Ku70, Rad51 and Rad52. Ku70 is a central component of the NHEJ pathway (Critchlow and Jackson 1998) whose absence has been shown to lead to a substantial increase in gene targeting efficiency in other organisms (Ninomiya et al. 2004; Pöggeler and Kück 2006; Fennessy et al. 2014). Rad51 is a RecA homolog that performs homology search and strand exchange in HR, while Rad52 is an

auxiliary factor of Rad51 (San Filippo et al. 2008). Rad52 also plays a Rad51-independent role (San Filippo et al. 2008). We identified *N. liquefaciens* genes encoding homologs of Rad51, Rad52 and Ku70 (Materials and methods, Fig. S1).

Rad51 from *N. liquefaciens* displays high conservation throughout its whole sequence with other Rad51 orthologues, showing 70.8%, 65.3% and 73% amino acid identity with its *S. pombe*, *S. cerevisiae*, and human counterparts, respectively (Fig. S1a). In the case of Rad52, however, conservation is limited to the N-terminal half, with 30.9%, 28.8% and 31% amino acid sequence identity when compared with the *S. pombe*, *S. cerevisiae*, and human counterparts, respectively (Fig. S1b). Ku70 is the least conserved, showing 25.2%, 24% and 27.4% identity in amino acid sequence with the *S. pombe*, *S. cerevisiae* and human counterparts, respectively (Fig. S1c).

Gene targeting is drastically improved by the absence of Ku70 in *Naganishia* yeast

Our results so far suggested that gene targeting is not very efficient in *N. liquefaciens*, which is typically seen when NHEJ, as opposed to HR, plays a predominant role in repairing DSBs (Critchlow and Jackson 1998). Thus, we first turned our attention to the NHEJ pathway. To examine if suppressing NHEJ improves gene targeting efficiency in *N. liquefaciens*, the gene encoding Ku70 was deleted.

The absence of Ku70 dramatically improved gene targeting efficiency, regardless of the length of homology arms (1 kb or 80 bp) and the locus examined (*HIS3* or *ADE2*) (Fig. 5a, Table S3). Without Cas9 induction (denoted as "conventional"), the efficiencies of gene targeting in the *ku70* mutant reached > 95%, while those for the wild-type strain were around 10-20% (Fig. 5a). A similar trend was observed when gene targeting was assisted by the Cas9 system (denoted as "Cas9 mediated") except that gene targeting efficiencies were already high (~ 80%) in the wild-type background, and they became even higher in the absence of Ku70 (~ 100%; Fig. 5b, Table S3).

Thus, it is likely that the low efficiency of gene targeting in *N. liquefaciens* is largely attributable to the predominant utilization of the NHEJ pathway.

The length of homology specifies the requirement for Rad51 and Rad52 in gene targeting in *Naganishia* yeast

The mechanism of HR is highly relevant to gene targeting (Mehta and Haber 2014). Rad51 and Rad52 play central roles in HR. Thus, their possible roles in gene targeting were examined by deleting them in *N. liquefaciens*.

In conventional gene targeting experiments where *HIS3* is targeted with 1 kb homologous sequences, the absence of Rad51 or Rad52 caused a mild reduction in gene targeting efficiency, to approximately 50% of the wild-type strain (Fig. 5c, Table S3). This reduction was almost completely suppressed by introducing the *ku70* mutation; gene targeting efficiency was ≥ 90% in wild type, *rad51*, and *rad52* strains in the absence of Ku70. These results suggest that, although both Rad51 and Rad52 are important for gene targeting, neither Rad51 nor Rad52 is essential when the homology length is 1 kb. However, no gene targeting was observed in the *rad51 rad52* double mutant with or without Ku70, arguing that Rad51 and Rad52 redundantly support gene targeting under this condition.

Next, the length of homology was reduced to 80 bp (Fig. 5c, Table S3). The absence of Rad51 lead to a reduction in gene targeting that was comparable to what we observed with 1 kb homology (~ 30% of the wild type strain), and again, this reduction was robustly suppressed by introduction of the *ku70* mutation (~ 80% in the *rad51* strain and ~ 100% in the wild-type strain). The absence of Rad52, however, completely eliminated gene targeting, and this reduction was not suppressed by the absence of Ku70. Thus, with 80 bp homology, Rad52 is indispensable for gene targeting, while Rad51 is not. A similar trend was seen when gene targeting was induced by Cas9 except that the overall targeting efficiencies were much higher (Fig. 5d, Table S3).

In the strains employed (wild type, *rad51*, *ku70*, and *rad51 ku70* strains with or without Cas9), homology-mediated accurate integration of the *NAT* marker at the *HIS3* locus was confirmed by PCR (14 clones examined per genotype, Fig. S4) and subsequent sequencing of the integration junctions (four clones examined per genotype, see Materials and methods for details). Unlike the experiments without Cas9, four transformants showing nourseothricin resistance and histidine auxotrophy arose in the *rad52* mutant when Cas9 was expressed. Three of them are likely the outcome of illegitimate recombination events since their PCR amplicons, which span the *HIS3* locus, were larger than would be expected if it was via HR (Fig. S4). One showed the amplicon size consistent with an HR-mediated event, which was further validated by sequencing its integration junctions. The results obtained at the *ADE2* locus essentially mirrored the results at the *HIS3* locus, arguing for locus independency of this trend (Fig. S5a,b, Table S3).

The *rad52* mutant is as resistant to DNA damage as the wild type strain in *Naganishia* yeast

Historically, genes involved in HR or NHEJ were found to be mutated in mutant strains/cell lines showing hyper-sensitivity to various DNA damaging sources including IR and UV (Game and Mortimer 1974; Jeggo 1998). Thus, we next probed the functional relationship between gene targeting and DNA damage repair. The mutants employed in the gene targeting experiments above were examined for their sensitivity to UV and ionizing radiation (IR).

The *rad51* mutant exhibited moderate sensitivity to UV, while *ku70* or *rad52* mutants were essentially indistinguishable from the wild-type strain (Fig. 6a). However, the *ku70* or *rad52* mutation slightly exacerbated the sensitivity of *rad51.* Furthermore, the *rad51 rad52 ku70* triple mutant exhibited substantially higher sensitivity to UV than either the *rad51 rad52* or *rad51 ku70* double mutants.

These strains also showed a similar pattern in response to IR treatment (Fig. 6b). The *rad51* mutant displayed much higher sensitivity than the wild-type, *rad52,* and *ku70* strains, but unlike what was observed with UV, the *rad52* and *ku70* mutants also showed subtle but discernable sensitivity to IR. The *ku70* mutation, and to a lesser extent the *rad52* mutation, both further sensitized the *rad51* mutant to IR. As was observed with UV treatment, the *rad51 rad52 ku70* triple mutant exhibited the highest sensitivity to IR. Notably, the wild-type strain barely exhibited IR sensitivity at 600 Gy, suggesting that *N. liquefaciens* has high tolerance to IR (Fig. 6b). This is reminiscent of other Basidiomycota species known to be highly tolerant to IR (Holloman et al. 2007; Jung et al. 2016).

Discussion:

Here, with the aim of establishing *N. liquefaciens* as a model organism, we have developed advanced genetic tools that enable efficient gene targeting and complementation analysis in this organism. The non-pathogenic nature of *N. liquefaciens*, as indicated by its inability to grow at 37°C (Fig. S6a), along with the short doubling time (~120 min, Fig. S6b), and the ability to thrive under conditions that are practically the same as those for *S. cerevisiae*, are highly advantageous traits as a model organism. Furthermore, we recently determined the draft genome sequence of this organism (Han et al. 2020). Together, this has enabled the application of molecular genetics approaches to *N. liquefaciens*, which provide further means to promote in-depth study of Basidiomycota biology.

Naganishia liquefaciens, a Basidiomycota budding yeast

N. liquefaciens N6 was isolated from deep-sea sediments at a depth of ~6,500 m (Abe et al. 2001). Whole genome sequencing placed this organism under the phylum of Basidiomycota despite its apparent morphological similarity to the Ascomycota yeast *S. cerevisiae* (Han et al. 2020). *N. liquefaciens* and *S. cerevisiae* are both budding yeasts. Media commonly used for culturing *S. cerevisiae* can be used for culturing *N. liquefaciens* without further modification (Abe et al. 2001, 2006), which will make this organism accessible to researchers already working with *S. cerevisiae*. The doubling time of *N. liquefaciens* in rich media is ~120 min at 30°C, which is comparable to that for *S. cerevisiae*.

There are at least two Basidiomycota yeasts being used for studying basic biology: the human pathogen *Cryptococcus neoformans* (Mochizuki et al. 1987) and the plant pathogen *Ustilago maydis* (O'Donnell and McLaughlin 1984). In particular, *C. neoformans* has recently emerged as a representative model organism of the Basidiomycota phylum. Unlike *C. neoformans*, *N. liquefaciens* is incapable of growth at 37°C, implying that if it is ingested into the human body, it would be unable to proliferate. This is a favorable trait for a model organism.

Although both *Naganishia* and *Cryptococcus* are Basidiomycota yeasts, they are not closely related, likely having split ~ 250 million years ago (Zhao et al. 2017). *Ustilago* (Ustilaginales) is separated from *Naganishia* (Filobasidiales) and *Cryptococcus* (Tremellales) even further, ~ 450 million years ago (Han et al. 2020). Given that *S. cerevisiae* and *S. pombe*, the two most widely investigated Ascomycota yeasts, diverged from a common ancestor around 300 to 400 million years ago (Sipiczki 2000), studying multiple Basidiomycota yeasts distantly-related from each other also has the potential to provide unique insights into basic biological mechanisms.

Gene targeting is inefficient in *Naganishia* yeast

In this work, we showed that gene targeting is relatively inefficient in *N. liquefaciens*. The overall efficiency was loosely correlated with the size of homology arms, with the best efficiency of around 25% obtained with 1 kb homology arms.

The most commonly used approach for gene targeting is to take advantage of the so called “ends-out” recombination (Pâques and Haber 1999), where the selective marker is flanked by a different length of the targeted sequence. The length of flanking homologous DNA necessary for gene targeting varies from species to species. Homology as short as 30-45 bp is sufficient to achieve successful gene targeting in *S. cerevisiae* (Manivasakam et al. 1995), whereas much longer homology, varying from 80 bp to several kb, is often necessary in most Ascomycota yeast species including *S. pombe* (Klinner and Schäfer 2004). HR-mediated gene targeting is extremely low in Metazoans including mice and humans (Capecchi 2005).

One factor that contributes to a reduction in gene targeting is usage of the NHEJ pathway. A DSB can be repaired accurately by using HR, or inaccurately by NHEJ. *S. cerevisiae* is exceptional in that DSBs are almost exclusively repaired through HR, and gene targeting is extremely efficient. If NHEJ is predominantly used for DSB repair,

gene targeting becomes accordingly inefficient. This is likely to be the case in *Naganishia* yeast because the absence of Ku70, a main component of the NHEJ pathway (Critchlow and Jackson 1998), dramatically improved gene targeting efficiency. A similar improvement by suppressing NHEJ functions has been seen in many other model organisms, from Ascomycota to vertebrates (Ninomiya et al. 2004; Pöggeler and Kück 2006; Iiizumi et al. 2008; Fennessy et al. 2014).

Introducing a DSB at the target locus by a newly devised CRISPR/Cas9 system in *Naganishia* yeast achieved highly efficient gene targeting. CRISPR/Cas9 is now a well-established means to improve gene targeting efficiency and has proven to be hugely successful in organisms where NHEJ is the predominant DSB repair pathway, including humans. Importantly, with this CRISPR/Cas9 system, gene targeting efficiency reached ~80% with the length of homology as short as 80 bp. In the fungus species *Aspergillus fumigatus*, a donor DNA flanked by 35~50 bp homology arms supports efficient HR-mediated gene targeting when coupled with the CRISPR-Cas9 system (Al Abdallah et al. 2017). On the other hand, in *C. neoformans*, a donor DNA with 50 bp homology is not sufficient (Fan and Lin 2018). A further investigation will address if *N. liquefaciens* and *C. neoformans* share a similar mechanism for gene targeting. A homology length as short as 80 bp was enough to achieve efficient gene targeting. This means that gene targeting fragments (i.e., donor sequences) can be prepared rapidly and economically by PCR using 100 bp primers containing 80 bp of homology flanking the target site.

Genetic requirement for gene targeting in *Naganishia* yeast

To explore the genetic requirement for gene targeting in *N. liquefaciens*, two different experimental systems were employed. First, conventional gene targeting employing just a targeting DNA fragment with either short (80 bp) or long (500 bp and 1 kb) homologous sequences at its ends was tested. Second, the CRISPR/Cas9 system was employed along

with the above targeting fragment, which served as a donor. In either system, the condition where the NHEJ pathway is suppressed via deletion of *KU70* was also examined. Moreover, because gene targeting is a form of HR, we also examined the requirement for two critical components of the HR machinery: Rad51, the enzyme responsible for conducting homology search and strand exchange between homologous DNA molecules; and Rad52, which is essential for the recruitment of Rad51 to DSBs in Ascomycota yeasts (San Filippo et al. 2008). Rad52 also plays a Rad51-independent role in DNA repair (San Filippo et al. 2008). For this reason, HR defects caused by mutating *RAD52* are severer than those observed in the *rad51* mutant. The absence of each or both proteins was investigated in our gene targeting system.

There are a few general trends seen at the test loci (*HIS3* and *ADE2*). When the length of the homology arm is 1 kb, in either the conventional or Cas9-mediated system, the absence of Rad51 or Rad52 reduced gene targeting efficiency to ~50% of the wild-type level. This reduction was almost completely suppressed by deleting *KU70*. In the absence of both Rad51 and Rad52, however, gene targeting was completely abolished irrespective of homology length, employment of Cas9, or the absence of Ku70. These results argue that gene targeting is redundantly supported by Rad51 and Rad52 when there is significant homology (1 kb). The absence of Rad51 also caused a mild reduction in gene targeting when the length of the homology arm is 80 bp, both in the conventional and Cas9-mediated systems, although the defect was largely rescued by introducing the *ku70* mutation. In the absence of Rad52, however, gene targeting was almost completely eliminated under all tested conditions when 80 bp of homology was employed. These observations suggest that gene targeting relies almost exclusively on Rad52 when homology is limited (80 bp), with Rad51 relegated to a minor role.

The overall trend described above is essentially in line with previous work implicating Rad52, but not Rad51, in playing a predominant role in gene targeting in *S. cerevisiae* (Schiestl et al. 1994). Rad52 not only promotes Rad51 activity, but often functions

1 independently of Rad51, for example, in single-strand annealing (SSA) and break-induced
2 replication (BIR) (Symington 2002). Rad52 has an activity to facilitate annealing of
3 complementary strands (Mortensen et al. 1996; Sugiyama et al. 2006; Bugreev et al. 2007),
4 which is likely to play an essential role in gene targeting when homology length is short.
5 It is intriguing that gene targeting does not exclusively require either Rad51 or Rad52
6 when homology length is long enough (1 kb). Given that no gene targeting happens in the
7 absence of both Rad51 and Rad52, this argues that Rad51 can function independently of
8 Rad52. The *rad52* mutant is almost completely epistatic to *rad51* in *S. cerevisiae*, but there
9 are certain genetic conditions where Rad51 can function in the absence of Rad52 in *S. pombe*.
10 If the *rad52* mutation is combined with a mutation in the *FBH1* gene, which encodes a helicase
11 that negatively regulates Rad51 assembly, Rad51 can function semi-independently of Rad52
12 (Morishita et al. 2005; Osman et al. 2005). Substantial gene targeting still occurred in the
13 absence of Rad52 in *N. liquefaciens*. Consistently, little-to-no sensitivity to UV or IR was
14 observed in the *rad52* mutant. These observations suggest that homologous recombination
15 takes place without Rad52 in *N. liquefaciens* (discussed below).
16 We did observe a very small number of illegitimate integration events in the *rad51*
17 *rad52 ku70* triple mutant. This indicates that other DSB repair pathway(s), such as the
18 microhomology-mediated end joining (MMEJ) pathway, may also be operating in *N.*
19 *liquafaciens*. Consistent with this possibility, previous studies have shown that MMEJ is
20 dependent on neither Rad52 nor Ku proteins (Ma et al. 2003; Decottignies 2007).
21
22 **Basidiomycota yeasts and Metazoan biology**
23 Rad51, Rad52 and Ku70 are highly conserved proteins in eukaryotic species including
24 Basidiomycota and Ascomycota yeasts (Fig. S1). Rad51 in *C. neoformans* and *U. maydis* is
25 important for HR and repairing damaged DNA similarly to that in *N. liquefaciens* and other
26 Ascomycota yeasts such as *S. cerevisiae* and *S. pombe* (Ferguson et al. 1997; Jung et al.
27 2016). Interestingly, the *ku70* null mutant is not viable in *U. maydis* (de Sena-Tomás et al.

2015), while in *C. neoformans*, cells without Ku70/80 are viable but exhibit slightly higher sensitivity to phleomycin, a radio-mimetic agent (Goins et al. 2006). This observation is consistent with the involvement of *C. neoformans* Ku70 in DSB repair and also in line with our result showing that the *ku70* mutant displays mild sensitivity to IR. The *ku70* null mutant in *S. cerevisiae* and *S. pombe* shows little, if any, sensitivity to IR or radio-mimetic agents (Boulton and Jackson 1996; Manolis et al. 2001).

Interestingly, the absence of Rad52 in *N. liquefaciens* does not cause severe sensitivity to IR, which is in stark contrast to the phenotypes of the *rad52* mutant in the Ascomycota yeasts *S. cerevisiae* and *S. pombe*. This result is in agreement with observations obtained in another Basidiomycota yeast, *U. maydis*, where the *rad52* mutation causes no major defects in DNA repair (Kojic et al. 2008). Despite a level of UV and IR resistance that is comparable to the wild type strain, gene targeting in the *N. liquefaciens* *rad52* mutant was severely defective when the homology length is short (80 bp). Given that gene targeting involves some mechanisms related to those employed in BIR, it is possible that Rad52 plays an essential role in BIR in this organism, just like in Ascomycota yeasts (Anand et al. 2013). This would also suggest that the role of Rad52 in gene targeting is largely dispensable for repairing UV or IR-damaged DNA. If Rad52 is indeed exclusively required for BIR and SSA in this organism, BIR and SSA might play only a minor role in DNA damage repair. Overall, these traits associated with the absence of Rad52 are reminiscent of those found in mice. It has been shown that the absence of Rad52 does not cause major HR defects (Rijkers et al. 1998; Yamaguchi-Iwai et al. 1998), while BIR-associated phenomena, especially those related to telomere maintenance (Verma et al. 2019; Zhang et al. 2019) and mitotic DNA synthesis (Murfuni et al. 2013; Bhowmick et al. 2016) are specifically impaired. It would be interesting to examine if Rad52 is also important for supporting efficient gene targeting in vertebrates, especially when homology length attached to a gene targeting construct is rather short (~ 80 bp).

Basidiomycota and Ascomycota represent two major phyla of the fungal kingdom. Ascomycota yeasts, especially *S. cerevisiae* and *S. pombe*, are widely accepted as the simplest eukaryotic models (Forsburg 2005). Their common use is mainly attributable to technical advantages such as a short doubling time, simple cultivation conditions, and the common availability of various molecular genetic tools. However, some biological processes characteristic of higher eukaryotes, such as nuclear envelope dynamics during open mitosis, cannot be studied using Ascomycota models. By contrast, Basidiomycota yeasts, such as the human pathogen *C. neoformans* (Kozubowski et al. 2013) and the plant pathogen *U. maydis,* undergo semi-open mitosis, which is reminiscent of Metazoan mitosis (Straube et al. 2005). In addition, *C. neoformans* is the most intron-rich fungal species (Csuros et al. 2011). Moreover, a genome-wide comparison of the predicted proteome of *U. maydis*, *S. cerevisiae* and humans revealed that human proteins share more similarity to those of *U. maydis* than those of *S. cerevisiae* (Münsterkötter and Steinberg 2007; Steinberg and Perez-Martin 2008). Thus, in addition to the Ascomycota model yeasts, Basidiomycota yeasts could provide unique insights into Metazoan biology.

Acknowledgments

We are grateful to the Biomaterials Analysis Division, Open Facility Center, Tokyo Institute of Technology for sequence analysis. We also thank Yumiko Kurokawa and all members of the Iwasaki laboratory for stimulating discussion.

Figure Legends

Figure 1. The establishment of gene targeting systems in *N. liquefaciens*. **a** Constructs for gene targeting developed in this study. *GOI*, gene of interest. **b** Optimization of electroporation conditions (see Materials and methods for details). **c-f** Gene targeting efficiency was examined by transformation at the indicated loci. Error bars, standard deviation. n = 3 for all measurements. Statistical significance was determined by unpaired two-tailed t-test (n.s., not significant; * $p < 0.05$; ** $p < 0.01$).

Figure 2. Split marker transformation improves gene targeting efficiency. **a** Schematics of gene targeting using an intact marker (i) or split marker fragments (ii). **b** Gene targeting efficiencies using the intact marker and split marker approaches. Error bars, standard deviation. n = 9 for all measurements. Statistical significance was determined by unpaired two-tailed t-test (*, $p < 0.05$).

Figure 3. Cas9 expression promotes gene targeting. **a-d** Gene targeting fragments were employed with a newly developed Cas9-expressing plasmid and gene targeting efficiency was measured at the indicated loci. Error bars, standard deviation. n = 3 for all measurements. **e** 14 transformants showing nourseothricin resistance and histidine auxotrophy obtained with targeting DNA carrying 1 kb homologous arms and the Cas9 system, were randomly selected and both correct gene targeting (i) and possible random integration of the *CAS9* gene (ii) were examined. **f** Same as (**e** (i)) but targeting DNA with 80 bp homologous arms was employed. Primers used for PCR are much closer to the *HIS3* coding sequence, thus PCR amplicons are smaller than those in (**e** (i)) (see Materials and methods).

Figure 4. Development of a system that facilitates ectopic integration of foreign DNA. **a** Schematic of the strategy for integrating foreign DNA at the *TRP2* locus. A strain whose *TRP2* gene is replaced by the *NAT* marker is used as the base strain. NotI digestion of pM175 releases the DNA fragment carrying the *TRP2* gene, a multicloning site (MCS) where a foreign DNA (*GENE X*) can be cloned, and flanking sequences homologous to the regions upstream and downstream of the *TRP2* coding sequence. The released DNA and the plasmid expressing Cas9-gNAT (pM101-gNAT) facilitates targeted integration of the DNA fragment at the *TRP2* locus using HR. Arrowheads indicate the sites of DNA digestion. **b** An example of using the foreign DNA integration system to introduce the wild-type *RAD51* gene into the *TRP2* locus of the *rad51* null mutant. The relationship between genotypes and the annealing sites of PCR primers used for diagnostic PCR are shown. Primers used are: #1, Pr-144; #2, Pr-145; #3, Pr-292; #4, Pr-293; #5, Pr-313; #6, Pr-422. **c** Correct integration of the wild-type *RAD51* gene at the *TRP2* locus was verified by PCR using primers shown in (**b)**. **d** The three strains used in (**b**) were examined for their sensitivity to UV. Serial 10-fold dilutions of the indicated three strains were spotted onto two YPD plates, one irradiated with UV while the other left unirradiated (control). Strain A, wild type (MP17); strain B, the *rad51* null mutant before *RAD51* integration (MP21); strain C, a transformant showing tryptophan prototrophy and nourseothricin sensitivity (MP56). Δ denotes deletion of a gene.

Figure 5. Impact of DSB repair mutations on gene targeting efficiency. **a** Gene targeting efficiency was examined using strains with or without Ku70. Different lengths of homologous arms and targeted loci were employed as indicated. **b** As in (**a**) except gene targeting was promoted by the Cas9 system. **c** As in (**a**) except that various combinations of DSB repair mutations were examined as indicated. **d** Same as (**c**) except gene targeting was promoted by the Cas9 system. The data used for wild type and the *ku70* single mutant strains in (**c**) and (**d**) are the same as those in (**a**) and (**b**). Strains used are (Δ denotes deletion of a

gene): wild type, MP1; *ku70*Δ, MP72; *rad51*Δ, MP35; *rad52*Δ, MP33; *rad51*Δ *rad52*Δ, MP112; *ku70*Δ, MP72; *rad51*Δ *ku70*Δ, MP87; *rad52*Δ *ku70*Δ, MP75; *rad51*Δ *rad52*Δ *ku70*Δ, MP113. Error bars, standard deviation. n = 9 in "conventional" experiments. n = 3 in "Cas9 mediated" experiments except that n = 9 for strains carrying the *rad52* mutation. Statistical significance was determined by unpaired two-tailed t-test (n.s., not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

Figure 6. The absence of Rad52 confers little-to-no sensitivity to DNA damage. **a** The indicated mutants were examined for their sensitivity to UV. **b** Same as (**a**) except that sensitivity to IR was examined. Error bars, standard deviation. n = 3 for all measurements. Strains used are (Δ denotes deletion of a gene): wild type, MP1; *rad51*Δ, MP35; *rad52*Δ, MP33; *rad51*Δ *rad52*Δ, MP39; *ku70*Δ, MP72; *rad51*Δ *ku70*Δ, MP87; *rad52*Δ *ku70*Δ, MP75; *rad51*Δ *rad52*Δ *ku70*Δ, MP89.

References

Abe F, Minegishi H, Miura T, et al (2006) Characterization of cold- and high-pressure-active polygalacturonases from a deep-sea yeast, *Cryptococcus liquefaciens* strain N6. Biosci Biotechnol Biochem 70:296–299. https://doi.org/10.1271/bbb.70.296

Abe F, Miura T, Nagahama T, et al (2001) Isolation of a highly copper-tolerant yeast, *Cryptococcus sp.*, from the Japan Trench and the induction of superoxide dismutase activity by Cu2+. Biotechnol Lett 23:2027–2034. https://doi.org/org/10.1023/A:1013739232093

Al Abdallah Q, Ge W, Fortwendel JR (2017) A Simple and Universal System for Gene Manipulation in *Aspergillus fumigatus*: In Vitro-Assembled Cas9-Guide RNA Ribonucleoproteins Coupled with Microhomology Repair Templates. mSphere 2:. https://doi.org/10.1128/mSphere.00446-17

Amberg DC, Burke D, Strathern J (2015) Methods in yeast genetics : a Cold Spring Harbor Laboratory course manual. Cold Spring Harbor Laboratory Press

Anand RP, Lovett ST, Haber JE (2013) Break-Induced DNA Replication. Cold Spring Harb Perspect Biol 5:a010397–a010397. https://doi.org/10.1101/cshperspect.a010397

Arras SDM, Chua SMH, Wizrah MSI, et al (2016) Targeted Genome Editing via CRISPR in the Pathogen *Cryptococcus neoformans*. PLoS One 11:e0164322. https://doi.org/10.1371/journal.pone.0164322

Bhowmick R, Minocherhomji S, Hickson ID (2016) RAD52 Facilitates Mitotic DNA Synthesis Following Replication Stress. Mol Cell 64:1117–1126. https://doi.org/10.1016/j.molcel.2016.10.037

Blackwell M (2011) The fungi: 1, 2, 3 ... 5.1 million species? Am J Bot 98:426–38. https://doi.org/10.3732/ajb.1000298

Boulton SJ, Jackson SP (1996) *Saccharomyces cerevisiae* Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways. EMBO J 15:5093–5103. https://doi.org/10.1002/j.1460-2075.1996.tb00890.x

Bugreev D V, Hanaoka F, Mazin A V (2007) Rad54 dissociates homologous recombination intermediates by branch migration. Nat Struct Mol Biol 14:746–53. https://doi.org/10.1038/nsmb1268

Capecchi MR (2005) Gene targeting in mice: Functional analysis of the mammalian genome for the twenty-first century. Nat Rev Genet 6:507–512. https://doi.org/10.1038/nrg1619

Cong L, Ran FA, Cox D, et al (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339:819–23. https://doi.org/10.1126/science.1231143

Critchlow SE, Jackson SP (1998) DNA end-joining: from yeast to man. Trends Biochem Sci 23:394–398. https://doi.org/10.1016/S0968-0004(98)01284-5

Csuros M, Rogozin IB, Koonin E V. (2011) A Detailed History of Intron-rich Eukaryotic Ancestors Inferred from a Global Survey of 100 Complete Genomes. PLoS Comput Biol 7:e1002150. https://doi.org/10.1371/journal.pcbi.1002150

de Sena-Tomás C, Yu EY, Calzada A, et al (2015) Fungal Ku prevents permanent cell cycle arrest by suppressing DNA damage signaling at telomeres. Nucleic Acids Res 43:2138–2151. https://doi.org/10.1093/nar/gkv082

Decottignies A (2007) Microhomology-Mediated End Joining in Fission Yeast Is Repressed by Pku70 and Relies on Genes Involved in Homologous Recombination. Genetics 176:1403–1415. https://doi.org/10.1534/genetics.107.071621

Fairhead C, Llorente B, Denis F, et al (1996) New vectors for combinatorial deletions in yeast chromosomes and for gap-repair cloning using “split-marker” recombination. Yeast 12:1439–57. https://doi.org/10.1002/(SICI)1097-0061(199611)12:14%3C1439::AID-YEA37%3E3.0.CO;2-O

Fan Y, Lin X (2018) Multiple applications of a transient CRISPR-Cas9 coupled with electroporation (TRACE) system in the *cryptococcus neoformans* species complex. Genetics 208:1357–1372. https://doi.org/10.1534/genetics.117.300656

Fennessy D, Grallert A, Krapp A, et al (2014) Extending the Schizosaccharomyces pombe molecular genetic toolbox. PLoS One 9:. https://doi.org/10.1371/journal.pone.0097683

Ferguson DO, Rice MC, Rendi MH, et al (1997) Interaction between *Ustilago maydis REC2* and *RAD51* genes in DNA repair and mitotic recombination. Genetics 145:243–51

Forsburg SL (2005) The yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*: models for cell biology research. Gravit Space Biol Bull 18:3–9

Fu J, Hettler E, Wickes BL (2006) Split marker transformation increases homologous integration frequency in *Cryptococcus neoformans*. Fungal Genet Biol 43:200–12. https://doi.org/10.1016/j.fgb.2005.09.007

Game JC, Mortimer RK (1974) A genetic study of x-ray sensitive mutants in yeast. Mutat Res 24:281–92. https://doi.org/10.1016/0027-5107(74)90176-6

Goins CL, Gerik KJ, Lodge JK (2006) Improvements to gene deletion in the fungal pathogen *Cryptococcus neoformans*: Absence of Ku proteins increases homologous recombination, and co-transformation of independent DNA molecules allows rapid complementation of deletion phenotypes. Fungal Genet Biol 43:531–544. https://doi.org/10.1016/j.fgb.2006.02.007

Han Y-W, Kajitani R, Morimoto H, et al (2020) Draft Genome Sequence of *Naganishia liquefaciens* Strain N6, Isolated from the Japan Trench. Microbiol Resour Announc 9:19–21. https://doi.org/10.1128/MRA.00827-20

Hedges SB, Blair JE, Venturi ML, Shoe JL (2004) A molecular timescale of eukaryote evolution and the rise of complex multicellular life. BMC Evol Biol 4:2. https://doi.org/10.1186/1471-2148-4-2

Heitman J, Kozel TR, Kwon-Chung KJ, et al (eds) (2010) Cryptococcus. ASM Press, Washington, DC, USA

Hentges P, Van Driessche B, Tafforeau L, et al (2005) Three novel antibiotic marker cassettes for gene disruption and marker switching in *Schizosaccharomyces pombe*. Yeast 22:1013–1019. https://doi.org/10.1002/yea.1291

Higuchi R, Krummel B, Saiki RK (1988) A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic Acids Res 16:7351–67. https://doi.org/10.1093/nar/16.15.7351

Holloman WK, Schirawski J, Holliday R (2007) Towards understanding the extreme radiation resistance of Ustilago maydis. Trends Microbiol 15:525–529. https://doi.org/10.1016/j.tim.2007.10.007

Iiizumi S, Kurosawa A, So S, et al (2008) Impact of non-homologous end-joining deficiency on random and targeted DNA integration: Implications for gene targeting. Nucleic Acids Res 36:6333–6342. https://doi.org/10.1093/nar/gkn649

Ito H, Fukuda Y, Murata K, Kimura A (1983) Transformation of intact yeast cells treated with alkali cations. J Bacteriol 153:163–168. https://doi.org/10.1128/jb.153.1.163-168.1983

Jacobs JZ, Ciccaglione KM, Tournier V, Zaratiegui M (2014) Implementation of the CRISPR-Cas9 system in fission yeast. Nat Commun 5:5344. https://doi.org/10.1038/ncomms6344

Jeggo PA (1998) Identification of genes involved in repair of DNA double-strand breaks in mammalian cells. Radiat Res 150:S80-91

Jung KW, Yang DH, Kim MK, et al (2016) Unraveling fungal radiation resistance regulatory networks through the genome-wide transcriptome and genetic analyses of *Cryptococcus neoformans*. MBio 7:1–15. https://doi.org/10.1128/mBio.01483-16

Klinner U, Schäfer B (2004) Genetic aspects of targeted insertion mutagenesis in yeasts. FEMS Microbiol Rev 28:201–223. https://doi.org/10.1016/j.femsre.2003.10.002

Kojic M, Mao N, Zhou Q, et al (2008) Compensatory role for Rad52 during recombinational repair in *Ustilago maydis*. Mol Microbiol 67:1156–1168. https://doi.org/10.1111/j.1365-2958.2008.06116.x

Kozubowski L, Yadav V, Chatterjee G, et al (2013) Ordered Kinetochore Assembly in the Human-Pathogenic Basidiomycetous Yeast *Cryptococcus neoformans*. MBio 4:1–8. https://doi.org/10.1128/mBio.00614-13

Lin X, Chacko N, Wang L, Pavuluri Y (2015) Generation of stable mutants and targeted gene deletion strains in *Cryptococcus neoformans* through electroporation. Med Mycol 53:225–34. https://doi.org/10.1093/mmy/myu083

Liu XZ, Wang QM, Göker M, et al (2015) Towards an integrated phylogenetic classification of the Tremellomycetes. Stud Mycol 81:85–147. https://doi.org/10.1016/j.simyco.2015.12.001

Loftus BJ, Fung E, Roncaglia P, et al (2005) The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. Science (80-) 307:1321–1324. https://doi.org/10.1126/science.1103773

Ma J-L, Kim EM, Haber JE, Lee SE (2003) Yeast Mre11 and Rad1 Proteins Define a Ku-Independent Mechanism To Repair Double-Strand Breaks Lacking Overlapping End Sequences. Mol Cell Biol 23:8820–8828. https://doi.org/10.1128/MCB.23.23.8820-8828.2003

Mali P, Yang L, Esvelt KM, et al (2013) RNA-Guided Human Genome Engineering via Cas9. Science (80-) 339:823–826. https://doi.org/10.1126/science.1232033

Manivasakam P, Weber SC, McElver J, Schiestl RH (1995) Micro-homology mediated PCR targeting in *Saccharomyces cerevisiae*. Nucleic Acids Res 23:2799–2800. https://doi.org/10.1093/nar/23.14.2799

Manolis KG, Nimmo ER, Hartsuiker E, et al (2001) Novel functional requirements for non-homologous DNA end joining in *Schizosaccharomyces pombe*. EMBO J 20:210–21. https://doi.org/10.1093/emboj/20.1.210

Mashiko D, Fujihara Y, Satouh Y, et al (2013) Generation of mutant mice by pronuclear injection of circular plasmid expressing Cas9 and single guided RNA. Sci Rep 3:3355. https://doi.org/10.1038/srep03355

Mehta A, Haber JE (2014) Sources of DNA Double-Strand Breaks and Models of Rec. Cold Spring Harb Perspect Biol 6:1–19. https://doi.org/10.1101/cshperspect.a016428

Miura T, Abe F, Inoue A, et al (2001) Purification and characterization of novel extracellular endopolygalacturonases from a deep-sea yeast, *Cryptococcus sp. N6*, isolated from the Japan Trench. Biotechnol Lett 23:1735–1739. https://doi.org/org/10.1023/A:1012488115482.

Mochizuki T, Tanaka S, Watanabe S (1987) Ultrastructure of the mitotic apparatus in *Cryptococcus neoformans*. J Med Vet Mycol 25:223–33

Morishita T, Furukawa F, Sakaguchi C, et al (2005) Role of the *Schizosaccharomyces pombe* F-Box DNA Helicase in Processing Recombination Intermediates. Mol Cell Biol 25:8074–8083. https://doi.org/10.1128/mcb.25.18.8074-8083.2005

Mortensen UH, Bendixen C, Sunjevaric I, Rothstein R (1996) DNA strand annealing is promoted by the yeast Rad52 protein. Proc Natl Acad Sci U S A 93:10729–34. https://doi.org/10.1073/pnas.93.20.10729

Münsterkötter M, Steinberg G (2007) The fungus *Ustilago maydis* and humans share disease-related proteins that are not found in *Saccharomyces cerevisiae*. BMC Genomics 8:473. https://doi.org/10.1186/1471-2164-8-473

Murfuni I, Basile G, Subramanyam S, et al (2013) Survival of the Replication Checkpoint Deficient Cells Requires MUS81-RAD52 Function. PLoS Genet 9:. https://doi.org/10.1371/journal.pgen.1003910

Ninomiya Y, Suzuki K, Ishii C, Inoue H (2004) Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. Proc Natl Acad Sci 101:12248–12253. https://doi.org/10.1073/pnas.0402780101

O'Donnell KL, McLaughlin DJ (1984) Postmeiotic Mitosis, Basidiospore Development, and Septation in Ustilago Maydis. Mycologia 76:486–502. https://doi.org/10.1080/00275514.1984.12023869

Osman F, Dixon J, Barr AR, Whitby MC (2005) The F-Box DNA Helicase Fbh1 Prevents Rhp51-Dependent Recombination without Mediator Proteins. Mol Cell Biol 25:8084–8096. https://doi.org/10.1128/mcb.25.18.8084-8096.2005

Pâques F, Haber JE (1999) Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. Microbiol Mol Biol Rev 63:349–404

Pöggeler S, Kück U (2006) Highly efficient generation of signal transduction knockout mutants using a fungal strain deficient in the mammalian ku70 ortholog. Gene 378:1–10. https://doi.org/10.1016/j.gene.2006.03.020

Rijkers T, Van Den Ouweland J, Morolli B, et al (1998) Targeted Inactivation of Mouse *RAD52* Reduces Homologous Recombination but Not Resistance to Ionizing Radiation. Mol Cell Biol 18:6423–6429. https://doi.org/10.1128/MCB.18.11.6423

San Filippo J, Sung P, Klein H (2008) Mechanism of Eukaryotic Homologous Recombination. Annu Rev Biochem 77:229–257. https://doi.org/10.1146/annurev.biochem.77.061306.125255

Schiestl RH, Zhu J, Petes TD (1994) Effect of mutations in genes affecting homologous recombination on restriction enzyme-mediated and illegitimate recombination in *Saccharomyces cerevisiae*. Mol Cell Biol 14:4493–4500. https://doi.org/10.1128/MCB.14.7.4493

Sievers F, Wilm A, Dineen D, et al (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539. https://doi.org/10.1038/msb.2011.75

Sipiczki M (2000) Where does fission yeast sit on the tree of life? Genome Biol 1:1–4

Stajich JE, Berbee ML, Blackwell M, et al (2009) Primer – The Fungi. Curr Biol 19:R840–R845. https://doi.org/10.1016/j.cub.2009.07.004.Primer

Steinberg G, Perez-Martin J (2008) *Ustilago maydis*, a new fungal model system for cell biology. Trends Cell Biol 18:61–67. https://doi.org/10.1016/j.tcb.2007.11.008

Straube A, Weber I, Steinberg G (2005) A novel mechanism of nuclear envelope break-down in a fungus: nuclear migration strips off the envelope. EMBO J 24:1674–1685. https://doi.org/10.1038/sj.emboj.7600644

Sugiyama T, Kantake N, Wu Y, Kowalczykowski SC (2006) Rad52-mediated DNA annealing after Rad51-mediated DNA strand exchange promotes second ssDNA capture. EMBO J 25:5539–5548. https://doi.org/10.1038/sj.emboj.7601412

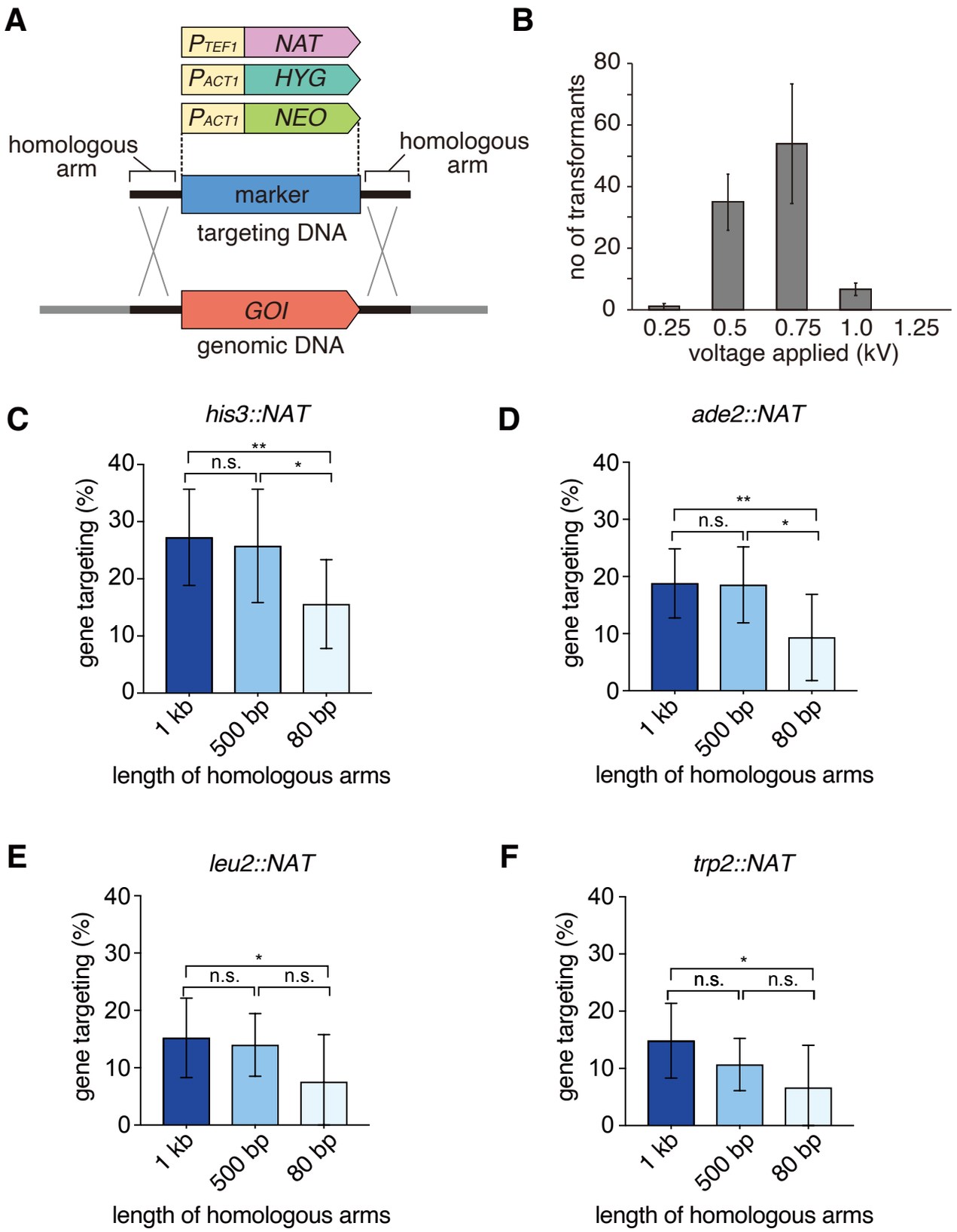
Verma P, Dilley RL, Zhang T, et al (2019) RAD52 and SLX4 act nonepistatically to ensure telomere stability during alternative telomere lengthening. Genes Dev 33:221–235. https://doi.org/10.1101/gad.319723.118

Wang T, Wei JJ, Sabatini DM, Lander ES (2014) Genetic Screens in Human Cells Using the CRISPR-Cas9 System. Science (80-) 343:80–84. https://doi.org/10.1126/science.1246981

Yamaguchi-Iwai Y, Sonoda E, Buerstedde J-M, et al (1998) Homologous Recombination, but Not DNA Repair, Is Reduced in Vertebrate Cells Deficient in RAD52. Mol Cell Biol 18:6430–6435. https://doi.org/10.1128/MCB.18.11.6430

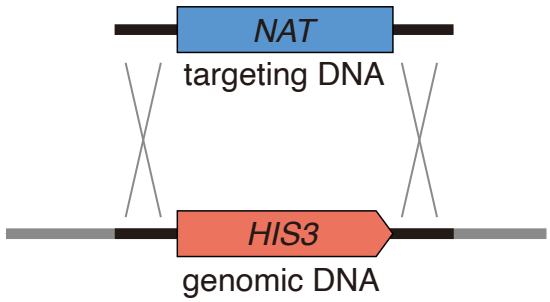
Zhang J-M, Yadav T, Ouyang J, et al (2019) Alternative Lengthening of Telomeres through Two Distinct Break-Induced Replication Pathways. Cell Rep 26:955-968.e3. https://doi.org/10.1016/j.celrep.2018.12.102

1 Zhao RL, Li GJ, Sánchez-Ramírez S, et al (2017) A six-gene phylogenetic overview of
2 Basidiomycota and allied phyla with estimated divergence times of higher taxa and a
3 phyloproteomics perspective
4

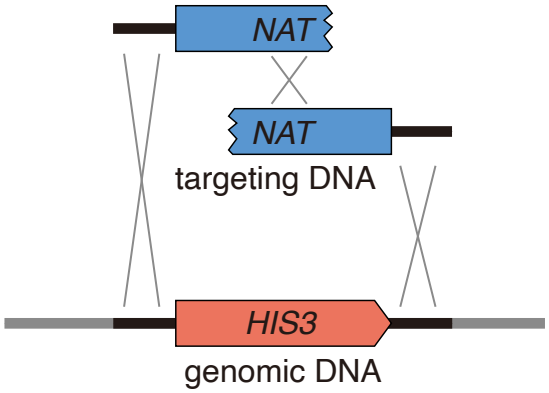

A
PTEF1
NAT
PACT1
HYG
PACT1
NEO
homologous arm
homologous arm
marker
targeting DNA
GOI
genomic DNA
B
no of transformants
80
60
40
20
0
0.25
0.5
0.75
1.0
1.25
voltage applied (kV)
C
his3::NAT
gene targeting (%)
40
30
20
10
0
**
n.s.
*
1 kb
500 bp
80 bp
length of homologous arms
D
ade2::NAT
gene targeting (%)
40
30
20
10
0
**
n.s.
*
1 kb
500 bp
80 bp
length of homologous arms
E
leu2::NAT
gene targeting (%)
40
30
20
10
0
*
n.s.
n.s.
1 kb
500 bp
80 bp
length of homologous arms
F
trp2::NAT
gene targeting (%)
40
30
20
10
0
*
n.s.
n.s.
1 kb
500 bp
80 bp
length of homologous arms

A

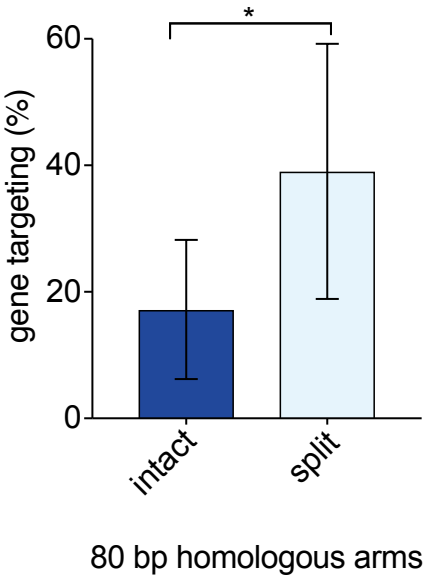
(i) intact marker

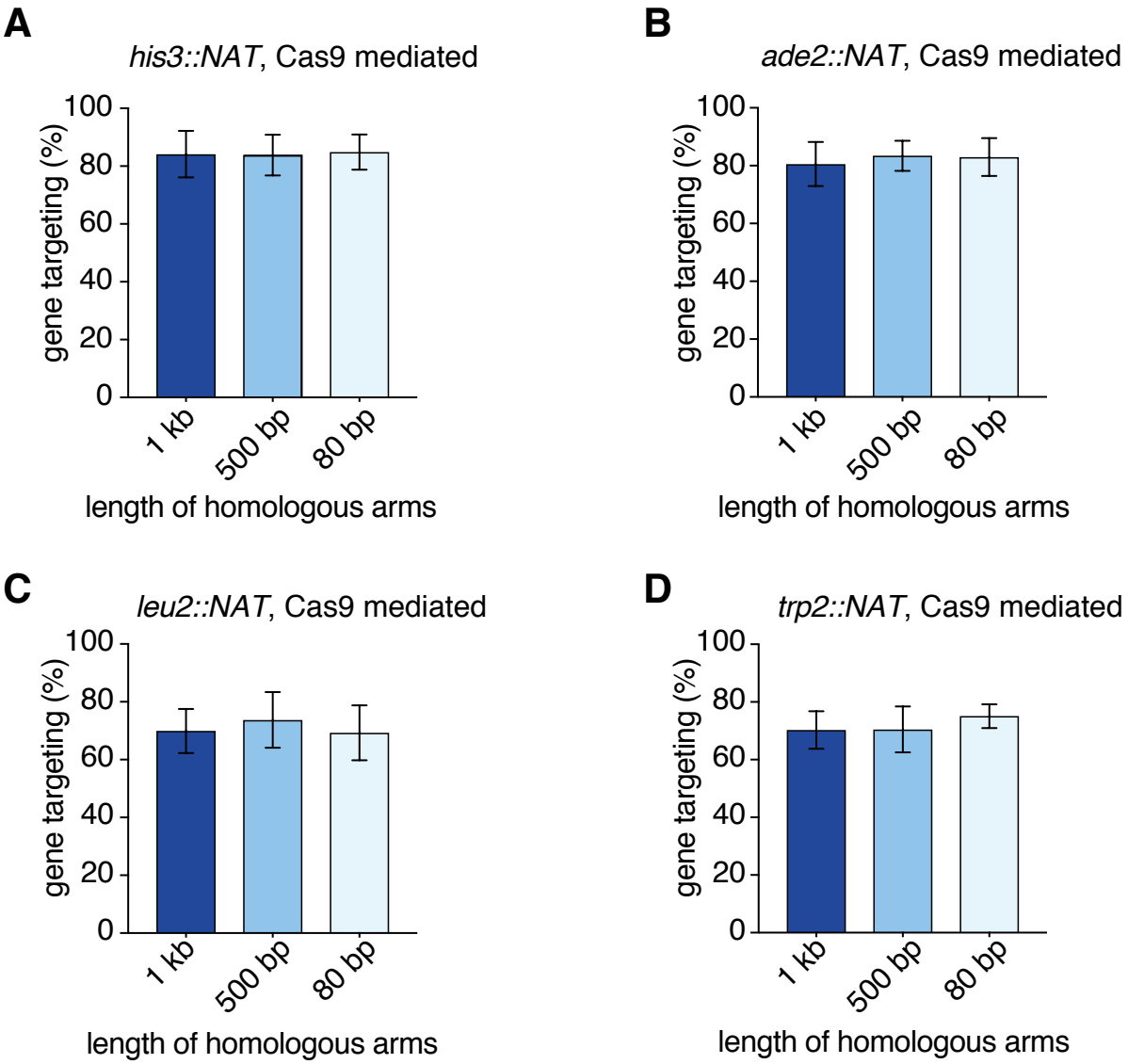


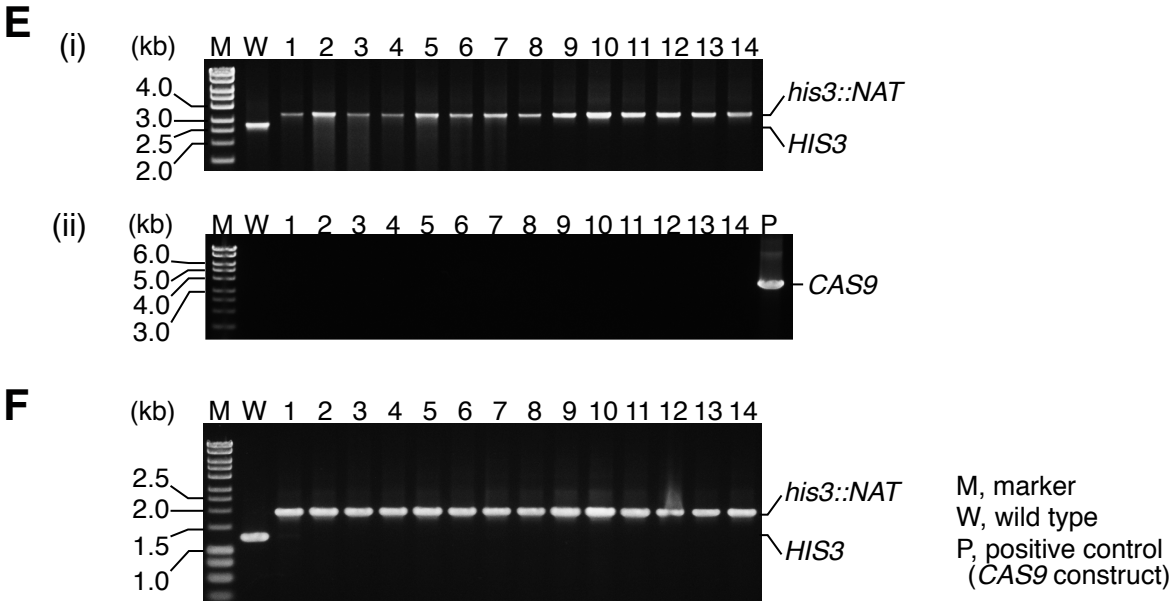
(ii) split marker

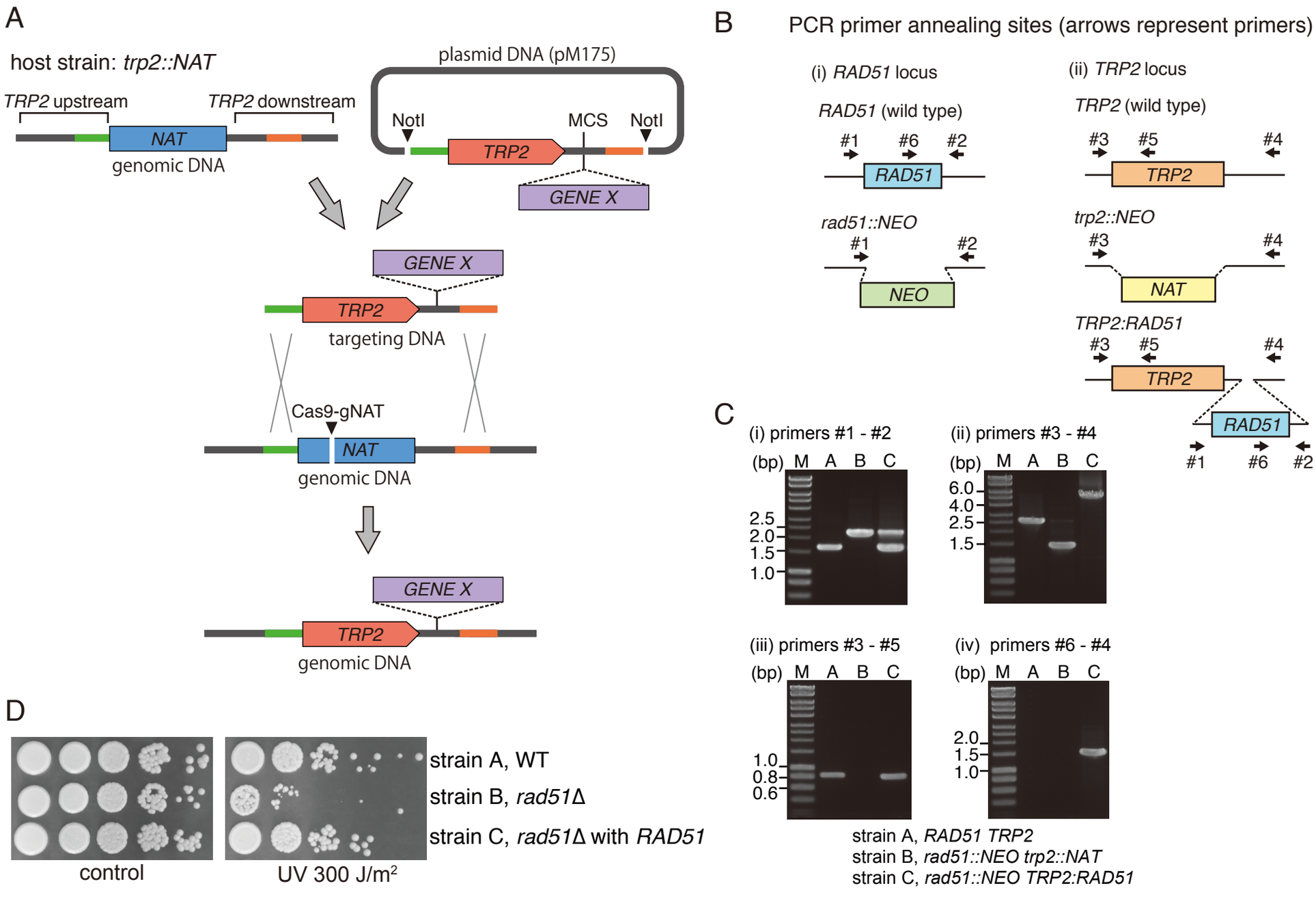


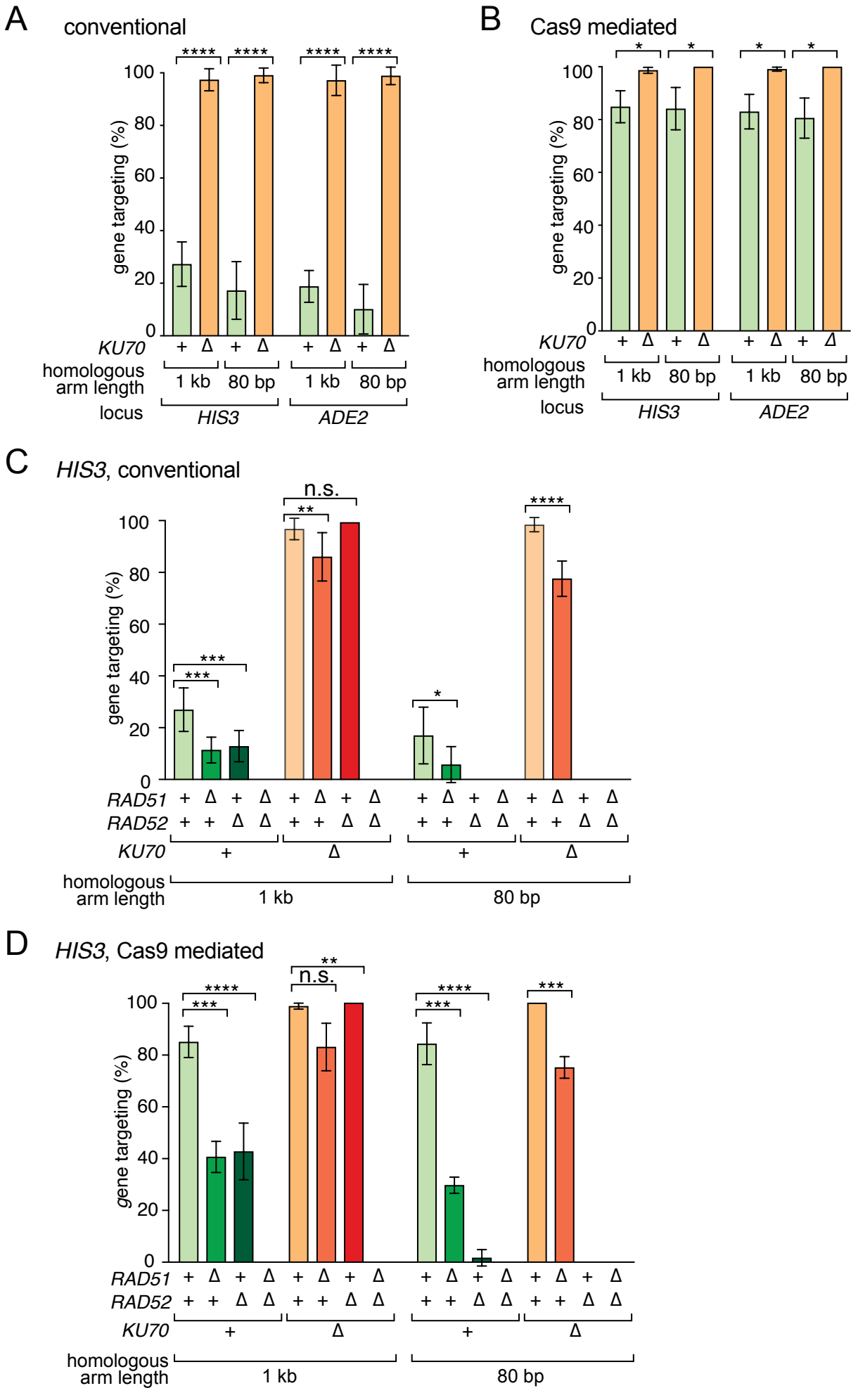
B



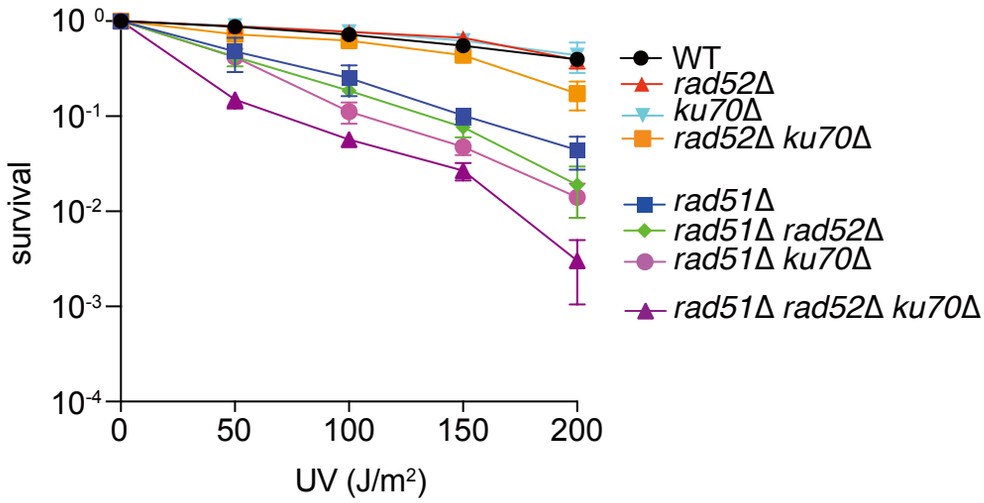

A
his3::NAT, Cas9 mediated
gene targeting (%)
100
80
60
40
20
0
1 kb
500 bp
80 bp
length of homologous arms
B
ade2::NAT, Cas9 mediated
gene targeting (%)
100
80
60
40
20
0
1 kb
500 bp
80 bp
length of homologous arms
C
leu2::NAT, Cas9 mediated
gene targeting (%)
100
80
60
40
20
0
1 kb
500 bp
80 bp
length of homologous arms
D
trp2::NAT, Cas9 mediated
gene targeting (%)
100
80
60
40
20
0
1 kb
500 bp
80 bp
length of homologous arms


E
(i)
(kb)
M W 1 2 3 4 5 6 7 8 9 10 11 12 13 14
4.0
3.0
2.5
2.0
his3::NAT
HIS3
(ii)
(kb)
M W 1 2 3 4 5 6 7 8 9 10 11 12 13 14 P
6.0
5.0
4.0
3.0
CAS9
F
(kb)
M W 1 2 3 4 5 6 7 8 9 10 11 12 13 14
2.5
2.0
1.5
1.0
his3::NAT
HIS3
M, marker
W, wild type
P, positive control
(CAS9 construct)

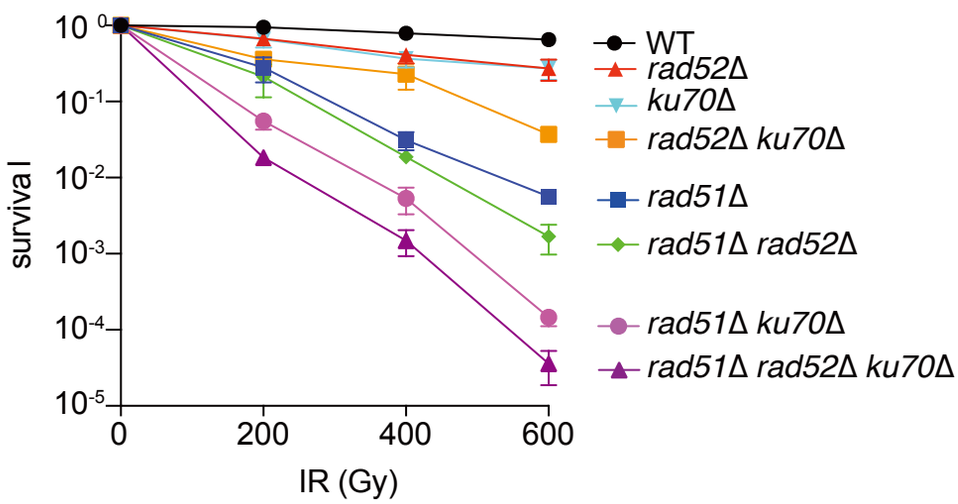

A
host strain: trp2::NAT
plasmid DNA (pM175)
TRP2 upstream
TRP2 downstream
NAT
genomic DNA
NotI
MCS
NotI
TRP2
GENE X
GENE X
TRP2
targeting DNA
Cas9-gNAT
NAT
genomic DNA
GENE X
TRP2
genomic DNA
B
PCR primer annealing sites (arrows represent primers)
(i) RAD51 locus
RAD51 (wild type)
#1
#6
#2
RAD51
rad51::NEO
#1
#2
NEO
(ii) TRP2 locus
TRP2 (wild type)
#3
#5
#4
TRP2
trp2::NEO
#3
#4
NAT
TRP2:RAD51
#3
#5
#4
TRP2
RAD51
#1
#6
#2
C
(i) primers #1 - #2
(bp) M A B C
2.5
2.0
1.5
1.0
(ii) primers #3 - #4
(bp) M A B C
6.0
4.0
2.5
1.5
(iii) primers #3 - #5
(bp) M A B C
1.0
0.8
0.6
(iv) primers #6 - #4
(bp) M A B C
2.0
1.5
1.0
strain A, RAD51 TRP2
strain B, rad51::NEO trp2::NAT
strain C, rad51::NEO TRP2:RAD51
D
control
UV 300 J/m2
strain A, WT
strain B, rad51Δ
strain C, rad51Δ with RAD51


A
conventional
B
Cas9 mediated
gene targeting (%)
KU70
homologous arm length
1 kb
80 bp
locus
HIS3
ADE2
C
HIS3, conventional
RAD51
RAD52
n.s.
D
HIS3, Cas9 mediated

A


10 0
10-1
10-2
10-3
10-4
survival
0
50
100
150
200
UV (J/m2)
WT
rad52Δ
ku70Δ
rad52Δ ku70Δ
rad51Δ
rad51Δ rad52Δ
rad51Δ ku70Δ
rad51Δ rad52Δ ku70Δ

B


10 0
10-1
10-2
10-3
10-4
10-5
survival
0
200
400
600
IR (Gy)
WT
rad52Δ
ku70Δ
rad52Δ ku70Δ
rad51Δ
rad51Δ rad52Δ
rad51Δ ku70Δ
rad51Δ rad52Δ ku70Δ

Supplementary Information

Homology length dictates the requirement for Rad51 and Rad52 in gene targeting in the Basidiomycota yeast *Naganishia liquefaciens*

Maierdan Palihati[1,2], Hideo Tsubouchi[1,2,*], Bilge Argunhan[2,3], Rei Kajitani[1], Omirgul Bakenova[1], Yong-Woon Han[1,4], Yasuto Murayama[1,5], Takehiko Itoh[1], and Hiroshi Iwasaki[1,2,*]

[1]School and Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8503, Japan.

[2]Institute of Innovative Research, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8503, Japan.

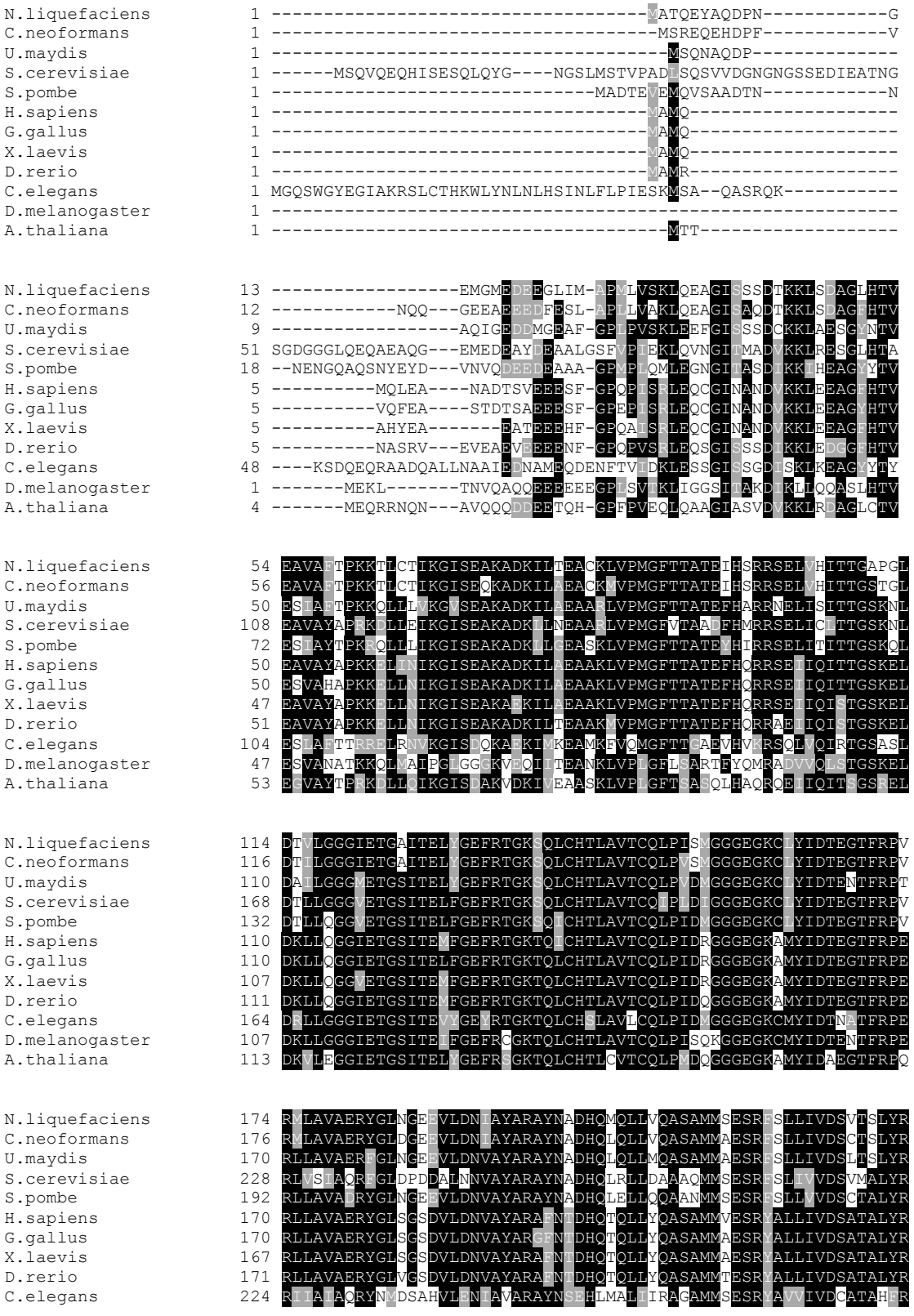
[3]Present address: Section of Structural Biology, Faculty of Medicine, Imperial College London, South Kensington, London SW7 2AZ, UK.

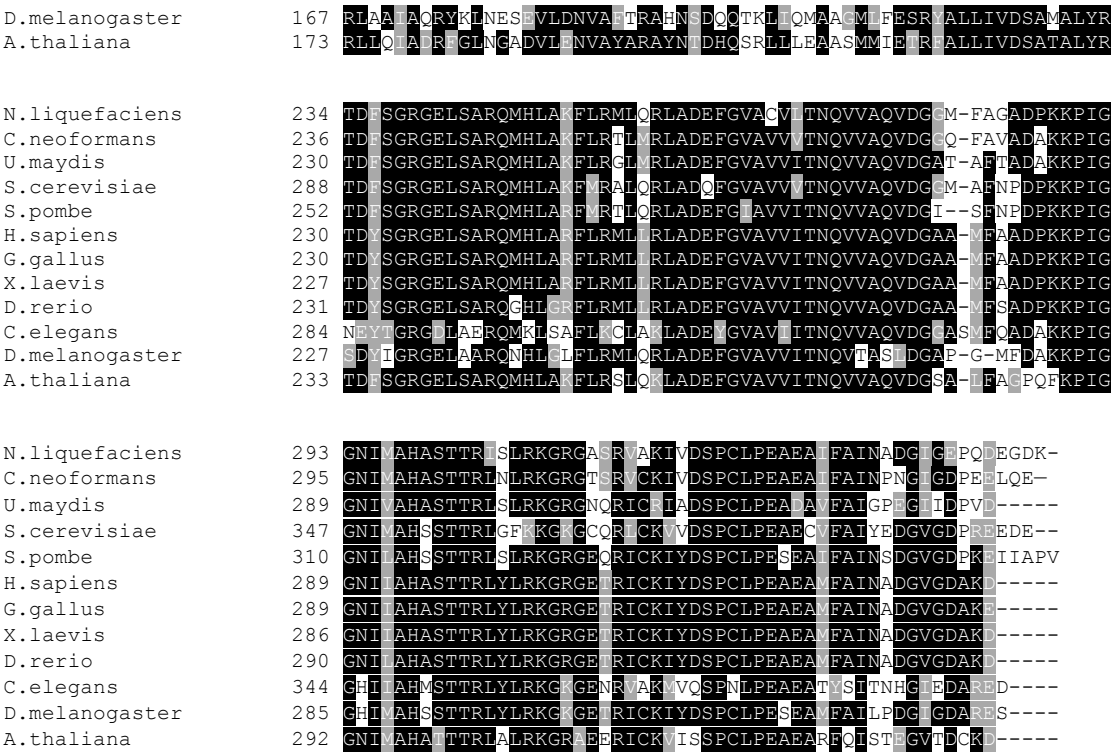
[4]Present address: Laboratory for Integrative Genomics, RIKEN Center for Integrative Medical Sciences, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan.

[5]Present address: Center for Frontier Research, National Institute of Genetics, 1111, Yata, Mishima, Shizuoka 411-8540, Japan.

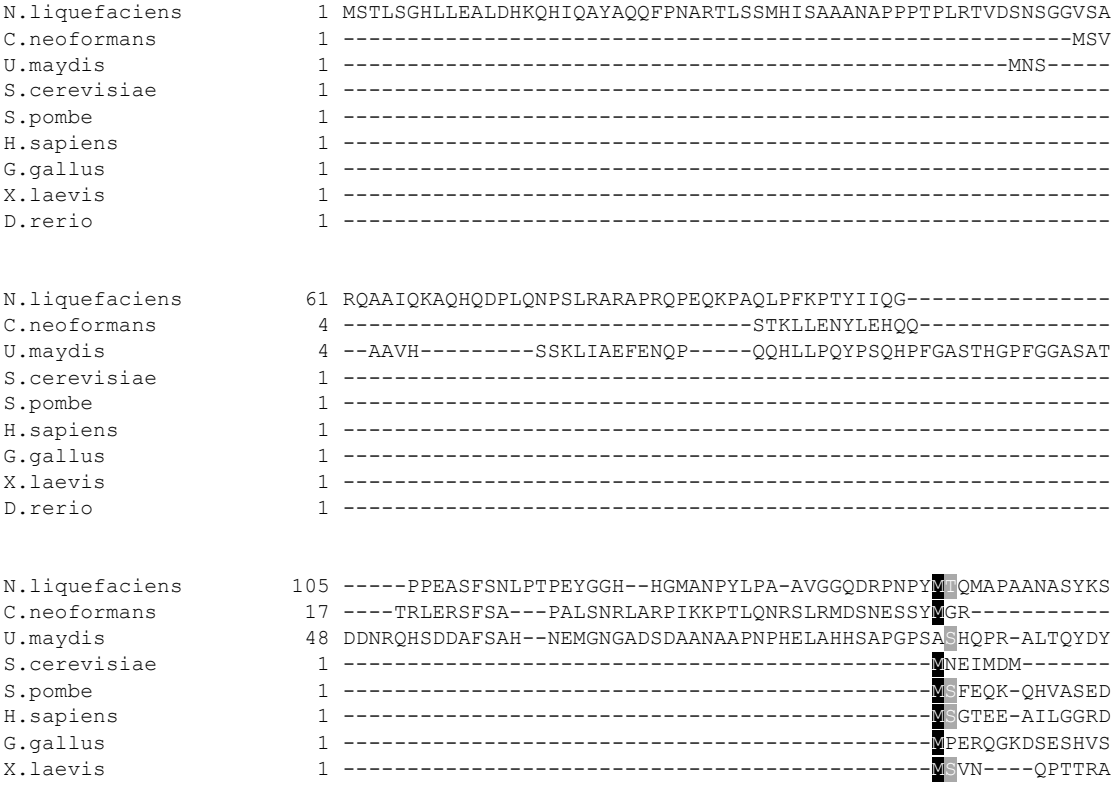
*Correspondence: htsubouchi@bio.titech.ac.jp and hiwasaki@bio.titech.ac.jp

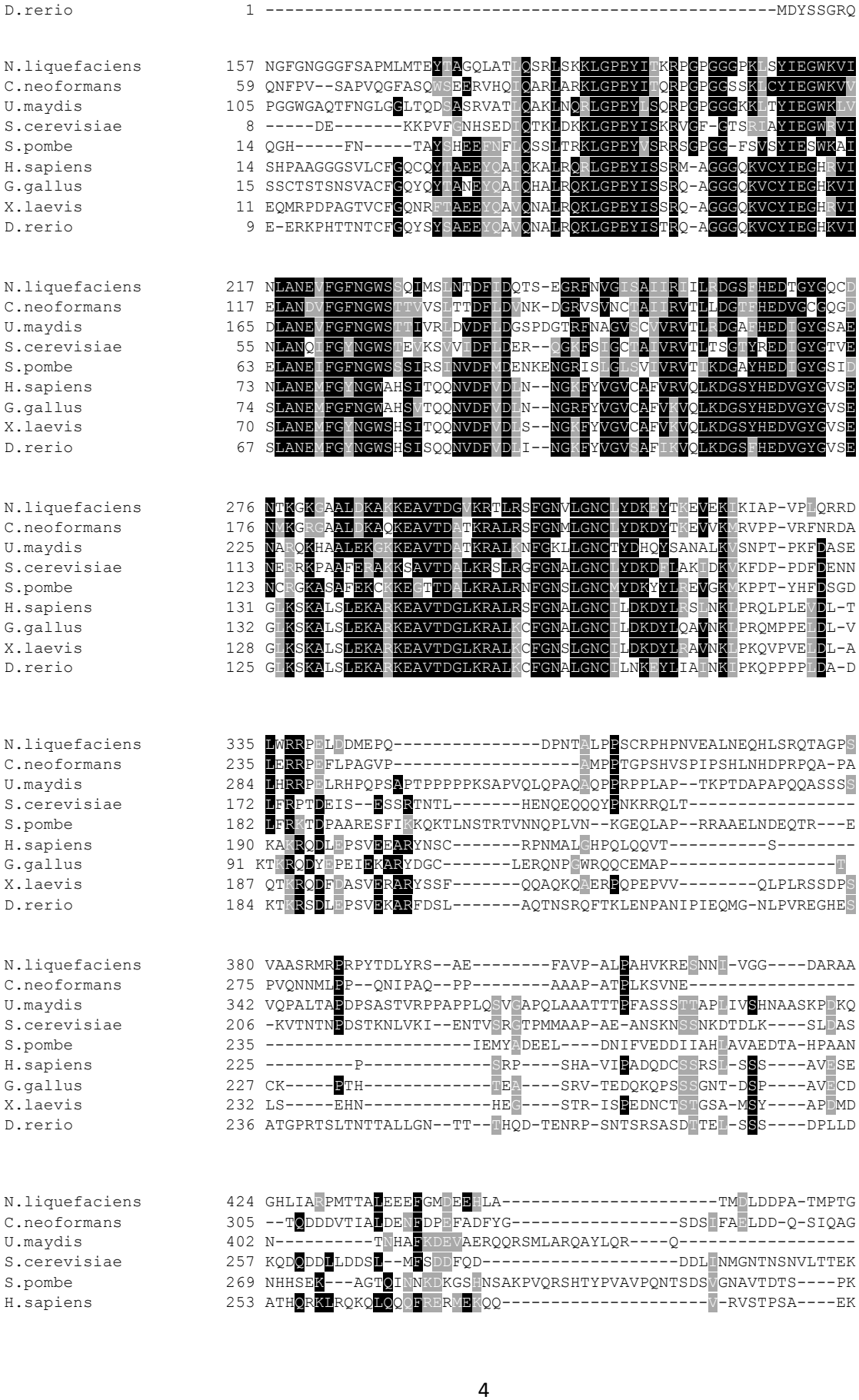
A.

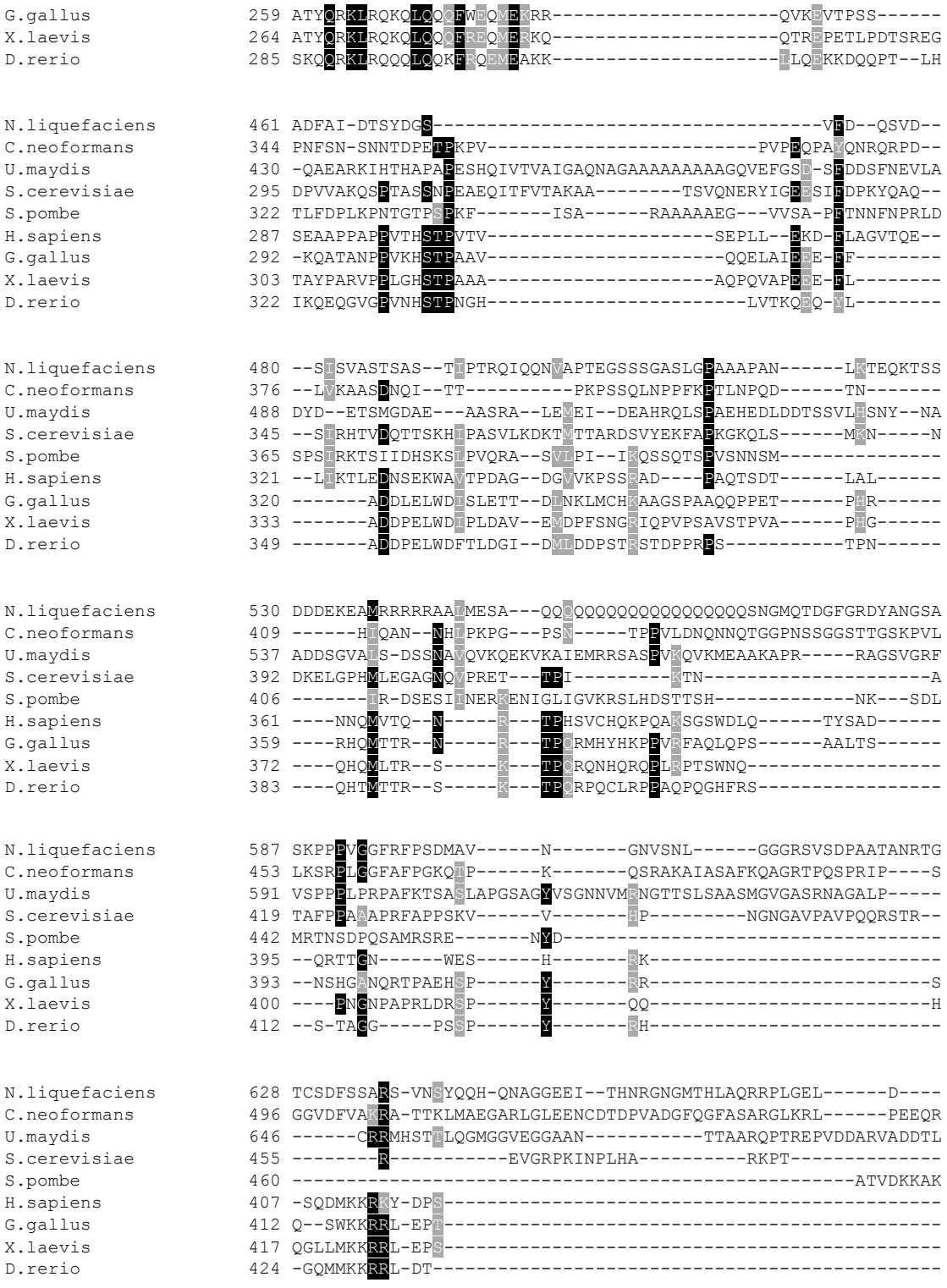

N.liquefaciens
C.neoformans
U.maydis
S.cerevisiae
S.pombe
H.sapiens
G.gallus
X.laevis
D.rerio
C.elegans
D.melanogaster
A.thaliana


D.melanogaster 167
A.thaliana 173
N.liquefaciens 234
C.neoformans 236
U.maydis 230
S.cerevisiae 288
S.pombe 252
H.sapiens 230
G.gallus 230
X.laevis 227
D.rerio 231
C.elegans 284
D.melanogaster 227
A.thaliana 233
N.liquefaciens 293
C.neoformans 295
U.maydis 289
S.cerevisiae 347
S.pombe 310
H.sapiens 289
G.gallus 289
X.laevis 286
D.rerio 290
C.elegans 344
D.melanogaster 285
A.thaliana 292

B.


N.liquefaciens 1
C.neoformans 1
U.maydis 1
S.cerevisiae 1
S.pombe 1
H.sapiens 1
G.gallus 1
X.laevis 1
D.rerio 1
N.liquefaciens 61
C.neoformans 4
U.maydis 4
S.cerevisiae 1
S.pombe 1
H.sapiens 1
G.gallus 1
X.laevis 1
D.rerio 1
N.liquefaciens 105
C.neoformans 17
U.maydis 48
S.cerevisiae 1
S.pombe 1
H.sapiens 1
G.gallus 1
X.laevis 1

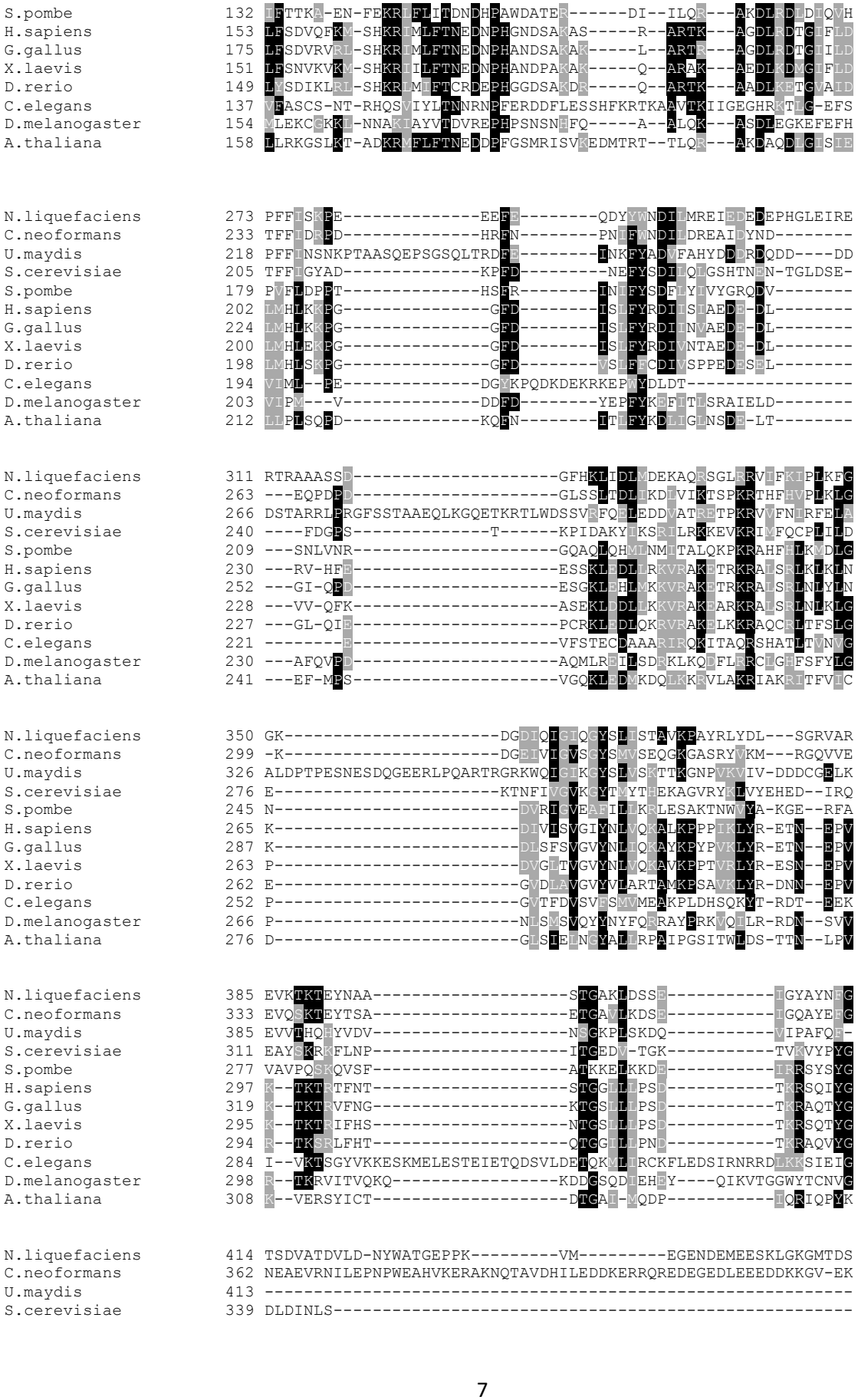

D.rerio 1 MDYSSGRQ
N.liquefaciens 157 NGFGNGGGFSAPMLMTEYTAGQLATLQSRLSKKLGPEYITKRPGPGGGPKLSYIEGWKVI
C.neoformans 59 QNFPV--SAPVQGFASQWSEERVHQIQARLARKLGPEYITQRPGPGGSSKLCYIEGWKVV
U.maydis 105 PGGWGAQTFNGLGGLTQDSASRVATLQAKLNQRLGPEYLSQRPGPGGGKKLTYIEGWKLV
S.cerevisiae 8 -----DE-------KKPVFGNHSEDIQTKLDKKLGPEYISKRVGF-GTSRIAYIEGWRVI
S.pombe 14 QGH-----FN-----TAYSHEEFNFLQSSLTRKLGPEYVSRRSGPGG-FSVSYIESWKAI
H.sapiens 14 SHPAAGGGSVLCFGQCQYTAEEYQAIQKALRQRLGPEYISSRM-AGGGQKVCYIEGHRVI
G.gallus 15 SSCTSTSNSVACFGQYQYTANEYQAIQHALRQKLGPEYISSRQ-AGGGQKVCYIEGHKVI
X.laevis 11 EQMRPDPAGTVCFGQNRFTAEEYQAVQNALRQKLGPEYISSRQ-AGGGQKVCYIEGHRVI
D.rerio 9 E-ERKPHTTNTCFGQYSYSAEEYQAVQNALRQKLGPEYISTRQ-AGGGQKVCYIEGHKVI
N.liquefaciens 217 NLANEYFGFNGWSSQIMSLNTDFIDQTS-EGRFNVGISAIIRTILRDGSFHEDTGYGQCD
C.neoformans 117 ELANDVFGFNGWSTTVVSLTTDFLDVNK-DGRVSVNCTAIIRVTLLDGTFHEDVGCGQGD
U.maydis 165 DLANEYFGFNGWSTTIVRLDVDFLDGSPDGTRFNAGVSCVVRVTLRDGAFHEDIGYGSAE
S.cerevisiae 55 NLANQIFGYNGWSTEVKSVVIDFLDER--QGKFSIGCTAIVRVTLTSGTYREDIGYGTVE
S.pombe 63 ELANEIFGFNGWSSSIRSLNVDFMDENKENGRISLGLSVTVRVTLKDGAYHEDIGYGSID
H.sapiens 73 NLANEYFGYNGWAHSITQQNVDFVDLN--NGKFYVGVCAFVRVQLKDGSYHEDVGYGVSE
G.gallus 74 SLANEYFGFNGWAHSVTQQNVDFVDLN--NGRFYVGVCAFVTVQLKDGSYHEDVGYGVSE
X.laevis 70 SLANEYFGYNGWSHSITQQNVDFVDLS--NGKFYVGVCAFVAVQLKDGSYHEDVGYGVSE
D.rerio 67 SLANEYFGYNGWSHSISQQNVDFVDLI--NGKFYVGVSAFIRVQLKDGSFHEDVGYGVSE
N.liquefaciens 276 NTKGKCAALDKAKKEAVTDGVKRTLRSFGNVLGNCIYDKEYTKEVEKIKIAP-VPLQRRD
C.neoformans 176 NVKGRCAALDKAQKEAVTDATKRALRSFGNMLGNCIYDKDYTKEVVKMRVPP-VRFNRDA
U.maydis 225 NAFQKHAALEKGKKEAVTDATKRALKNFGKLLGNCTYDHQYSANALKVSNPT-PKFDASE
S.cerevisiae 113 NERRKPAAFERAKKSAVTDALKRSLRGFGNALGNCIYDKDTIAKIDKVKFDP-PDFDENN
S.pombe 123 NCRGKASAFEKCKKEGTTDALKRALRNFGNSLGNCMYDKEYLREVGKMKPPT-YHFDSGD
H.sapiens 131 GLKSKALSLEKARKEAVTDGLKRALRSFGNALGNCILDKDYLRSLNKLPRQLPLEVDL-T
G.gallus 132 GLKSKALSLEKARKEAVTDGLKRALKCFGNALGNCILDKDYLQAVNKLPRQMPPEIDL-V
X.laevis 128 GLKSKALSLEKARKEAVTDGLKRALKCFGNSLGNCILDKDYLRAVNKLPKQVPVEIDL-A
D.rerio 125 GLKSKALSLEKARKEAVTDGLKRALKCFGNALGNCILNKEYLIAINKIPKQPPPPIDA-D
N.liquefaciens 335 LWRRPELLDMEPQ---------------DPNTALPPSCRPHPNVEALNEQHLSRQTAGPS
C.neoformans 235 LERRPEFLPAGVP--------------------AMPPTGPSHVSPIPSHLNHDPRPQA-PA
U.maydis 284 LHRRPELRHPQPSAPTPPPPPKSAPVQLQPAQAQPPRPPLAP--TKPTDAPAPQQASSSS
S.cerevisiae 172 LFRPTDEIS--ESSRTNTL-------HENQEQQQYPNKRRQLT----------------
S.pombe 182 LFRKTDPAARESFIKKQKTLNSTRTVNNQPLVN--KGEQLAP--RRAAELNDEQTR---E
H.sapiens 190 KAKRQDLLPSVEEARYNSC-------RPNMALGHPQLQQVT----------S--------
G.gallus 91 KTKRQDYLPEIEKARYDGC-------LERQNPGWRQQCEMAP----------------T
X.laevis 187 QTKRQDFLASVERARYSSF-------QQAQKQAERPQPEPVV--------QLPLRSSDPS
D.rerio 184 KTKRSDLLPSVEKARFDSL-------AQTNSRQFTKLENPANIPIEQMG-NLPVREGHES
N.liquefaciens 380 VAASRMRERPYTDLYRS--AE--------FAVP-ALPAHVKRESNNI-VGG----DARAA
C.neoformans 275 PVQNNMLEP--QNIPAQ--PP--------AAAP-ATPLKSVNE----------------
U.maydis 342 VQPALTAEDPSASTVRPPAPPLQSVCAPQLAAATTTPFASSSTTAPLIVSHNAASKPDKQ
S.cerevisiae 206 -KVTNTNFDSTKNLVKI--ENTVSRGTPMMAAP-AE-ANSKNSSNKDTDLK----SLDAS
S.pombe 235 ----------------------IEMYADEEL----DNIFVEDDIIAHLAVAEDTA-HPAAN
H.sapiens 225 ----------P--------------SRP----SHA-VIPADQDCSSRSL-SSS----AVESE
G.gallus 227 CK-----ETH--------------TEA----SRV-TEDQKQPSSSGNT-DSP----AVECD
X.laevis 232 LS-----EHN--------------HEG----STR-ISPEDNCTSIGSA-MSY----APDMD
D.rerio 236 ATGPRTSLTNTTALLGN--TT--THQD-TENRP-SNTSRSASDITEL-SSS----DPLLD
N.liquefaciens 424 GHLIARPMTTALEEEFGMEEHLA----------------------TMDLDDPA-TMPTG
C.neoformans 305 --TQDDDVTIALDENFDPEFADFYG-----------------SDSIFAELDD-Q-SIQAG
U.maydis 402 N----------TNHAFKDEVAERQQRSMLARQAYLQR----Q------------------
S.cerevisiae 257 KQDQDDLLDDSL--MFSDDFQD--------------------DDLINMGNTNSNVLTTEK
S.pombe 269 NHHSEK---AGTQINNKDKGSHNSAKPVQRSHTYPVAVPQNTSDSVGNAVTDTS----PK
H.sapiens 253 ATHQRKLRQKQLQQDFREREKQQ---------------------V-RVSTPSA----EK

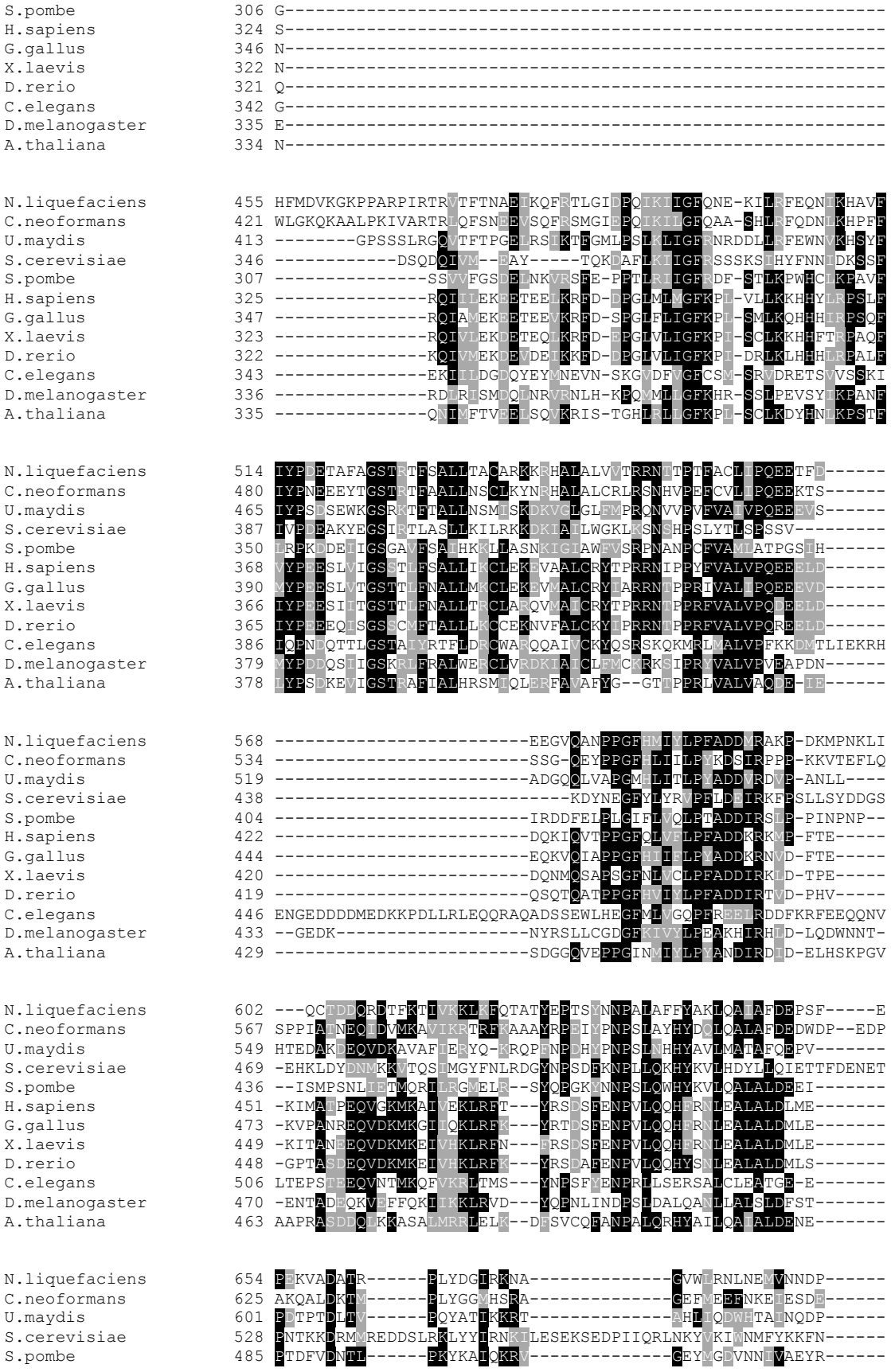


```
G.gallus        ------------------------------------------------
X.laevis        ------------------------------------------------
D.rerio         ------------------------------------------------
```

C.

```
N.liquefaciens          1 --MSQAKTQGVSFHPSQLTGKATSLNPGNEWKDETE----------DDDDLD---EN-QF
C.neoformans            1 ------------------MSSYYNRGDAPSWEALDQD--------GLDDVID---TS-EY
U.maydis                1 ------------------MPKAYFVNKREAARAGNSD--------TDDDDDDELEPLDLD
S.cerevisiae            1 --------------------------MRSVTNAFG-N---------SGELNDQVDETGYR
S.pombe                 1 ------------------------------------------MENDEQIDETEN
H.sapiens               1 ------------------M---------SGWESYYKTEGDEEAE---EEQEDNLEASGDY
G.gallus                1 MEMWVLGEVGMAVLSAAAM---------ADWVSYYRGDGPDEEEDGEQQEEEGPEAVADY
X.laevis                1 ------------------M---------TEWGDHFVNQEVED----DEEQKTGESDASAF
D.rerio                 1 ------------------M---------ANWGNYFPNEDDVDEQ-------DEEESGGDY
C.elegans               1 ------------------------------------------MDETRFEDDDFEG
D.melanogaster          1 ------------------M---------STWNP----ENDVDLLSGS---EDEEDVSMKR
A.thaliana              1 ------------------ME---------LDPDDVFRDEDE-----------DPENDFFQE

N.liquefaciens         45 KFVVKDHVLFCIDASSSMQEPLPDYMFKGEEFTKSLDTDIASTVIGKGKTPLHIALEVVH
C.neoformans           31 AYASRDHILFCIDAAQSMHKPYPDTTDE-------------SGRVVRGRSALHQALDVAQ
U.maydis               35 ATFQKDMVLFCIDAGPSMHRIDPATN----------------------TSPLYSALKAAA
S.cerevisiae           25 KFDIHEGLLFCIPLSETMEKESSDLEY---------------------KSPLLEILESLD
S.pombe                13 FAIGKYALFVIEVSPSMLDPVD-----------------------EFTPSSLQVALICAY
H.sapiens              31 KYSGRDSLIFLVDASKAMETSQS----------------------EDELTPFDVSIQCIQ
G.gallus               52 RFSGRDSLIFLVDASKAMETPYE----N-----------------EEAATPFDVTMQCIR
X.laevis               30 RFSGRDSLIFLVDASKPMETSID----------------------GELTPFDITLQCIR
D.rerio                27 KSSGRDSLVFLVDASKEMEIKGE----------------------NGEPSNFDVTMQCVR
C.elegans              14 STANKKYTFFIIDGNPAMETTSK----------------------AGEPPEFKLAIKLIL
D.melanogaster         27 DYHGRFALLFVVDANLQTAG---------------------------VERLLEALNIIR
A.thaliana             24 KEASKEFVVYLIDASPKMECSTCPSEEE-----------------DKQESHFHLAVSCIA

N.liquefaciens        105 AYEKAKALVAPQDSVGVLFYNVDHDTQPPTLANGERSETFRRGT-VLYQPLR-QVNVEE
C.neoformans           78 QIQRAKVLISGPDDSVGLILYNVDPSAV------AEDPGNYQPGN-YVFQTLR-TINAEE
U.maydis               73 SLMQKKIISSPHDHVGVLIFNCADT-----LFHSVKPGEYYKGS-YELQSVR-QVNVPD
S.cerevisiae           64 ELMSQLVITRPGTAIGCYFYYCNREDA--------KEGI------YELFPLR-DINATF
S.pombe                51 QLAAQRVITNPSDIMGVLLYGTESSTGR-------FANQM-----MLL-----DIDPPD
H.sapiens              69 SVYISKIISSDRDLLAVFYGTEKDKNS-VN----FKNIY-----VLQ-----ELDNPG
G.gallus               91 NVYTSKIISSDKDLLSVVFYGMENNKNS-AD----FKHIY-----VLQ-----ELDNPG
X.laevis               67 SVYASKIISSDHDLLSVVFFGTRESTNC-DP----FKHIC-----VLH-----DLDTPG
D.rerio                65 SVYTSKIISSDRDLVALVFYGTEQSKNPRNS----FKHIY-----VYH-----DLESPG
C.elegans              52 DEIVRVCCSRSLNNHIGVIVTSTKNSETEG------LENS------TLLVPVG-VLGQEE
D.melanogaster         59 TAFISGLLVNDKDLIGLIFANTKHSPPPLEA--SALDNIVMPDNCAVFLPLR-QLTKPI
A.thaliana             67 QSLKAHIINRSNDEIAICGTNTREKKNL-QD----LNGIY-----VFNVPERDSIDRPT

N.liquefaciens        162 IKRVRALLQEANIQLGEQSEDE-SGRSQPNILGETF--KPVNPDEILDIADVFE--CCNH
C.neoformans          129 MKRLAKLMQTAKEQYEAQGDDEAVETTEPEILRETF--PPIEMSHEMNIANVIQ--TCNF
U.maydis              125 TYNLKALLHEAELD--------------PKHLYTVL--PPAEKQMRIDWAL------ANA
S.cerevisiae          108 MKKLNDLLEDLS--------------SGRISLYDYFMFQQTCSEKQVRLSVLFTFMLDTF
S.pombe                93 AERIKSLQSFEK--------------DFQFSK---E--KFKPCSCQVSLSSVLV--HCSV
H.sapiens             113 AKRILELDQFKG--------------QQGQKR--FQ--DMVGHGSDYSLSEVLW--VCAN
G.gallus              135 AKRILLIDQYRG--------------DEGRVL--FR--ETFGHNADYSLCEALW--ACSN
X.laevis              111 AKRILLLDKYKE--------------EKGRAL--FC--DTICCGGDFSLCEALW--LCSN
D.rerio               110 AKRVQDIDKLKG--------------DKGGQF--AE--KTVGSG-ETSLCEALW--CCSN
C.elegans              99 VNKIKEIAE------------------EENLLSAVN--NYGGDHHKSDLSNVLN--YCKR
D.melanogaster        115 VEHYLE---FMG--------------GVETQFADVY--GLAEPDGRGRFDLMIR--LCIE
A.thaliana            116 ARLIKEFDLIEE--------------SFDKEIGSQT--GIVSDSRENSLYSALV--VAQA

N.liquefaciens        217 VFRDAGTKLSGTKRVFLVTNNDRPEIQVENEDL-RIDPRRPARTR---FIDLNNLGVSVD
C.neoformans          185 LFRDGGTQLKGNKRVFWITDNDMPP---------GMNNRQPARTS---YGDLTTYGVAAE
U.maydis              163 GVAVVVAANAGSKRVFWITDNDDPHPMSIDP-K-ATKARRACLDK---MNDFRKTGVRIE
S.cerevisiae          154 LEEIPSQKQLSNKRVFLFTDIDKPQEAQDIDERARLRRL---------TIDIFDNKVNFA
```


S.pombe 132
H.sapiens 153
G.gallus 175
X.laevis 151
D.rerio 149
C.elegans 137
D.melanogaster 154
A.thaliana 158
N.liquefaciens 273
C.neoformans 233
U.maydis 218
S.cerevisiae 205
S.pombe 179
H.sapiens 202
G.gallus 224
X.laevis 200
D.rerio 198
C.elegans 194
D.melanogaster 203
A.thaliana 212
N.liquefaciens 311
C.neoformans 263
U.maydis 266
S.cerevisiae 240
S.pombe 209
H.sapiens 230
G.gallus 252
X.laevis 228
D.rerio 227
C.elegans 221
D.melanogaster 230
A.thaliana 241
N.liquefaciens 350
C.neoformans 299
U.maydis 326
S.cerevisiae 276
S.pombe 245
H.sapiens 265
G.gallus 287
X.laevis 263
D.rerio 262
C.elegans 252
D.melanogaster 266
A.thaliana 276
N.liquefaciens 385
C.neoformans 333
U.maydis 385
S.cerevisiae 311
S.pombe 277
H.sapiens 297
G.gallus 319
X.laevis 295
D.rerio 294
C.elegans 284
D.melanogaster 298
A.thaliana 308
N.liquefaciens 414
C.neoformans 362
U.maydis 413
S.cerevisiae 339


S.pombe 306
H.sapiens 324
G.gallus 346
X.laevis 322
D.rerio 321
C.elegans 342
D.melanogaster 335
A.thaliana 334
N.liquefaciens 455
C.neoformans 421
U.maydis 413
S.cerevisiae 346
S.pombe 307
H.sapiens 325
G.gallus 347
X.laevis 323
D.rerio 322
C.elegans 343
D.melanogaster 336
A.thaliana 335
N.liquefaciens 514
C.neoformans 480
U.maydis 465
S.cerevisiae 387
S.pombe 350
H.sapiens 368
G.gallus 390
X.laevis 366
D.rerio 365
C.elegans 386
D.melanogaster 379
A.thaliana 378
N.liquefaciens 568
C.neoformans 534
U.maydis 519
S.cerevisiae 438
S.pombe 404
H.sapiens 422
G.gallus 444
X.laevis 420
D.rerio 419
C.elegans 446
D.melanogaster 433
A.thaliana 429
N.liquefaciens 602
C.neoformans 567
U.maydis 549
S.cerevisiae 469
S.pombe 436
H.sapiens 451
G.gallus 473
X.laevis 449
D.rerio 448
C.elegans 506
D.melanogaster 470
A.thaliana 463
N.liquefaciens 654
C.neoformans 625
U.maydis 601
S.cerevisiae 528
S.pombe 485

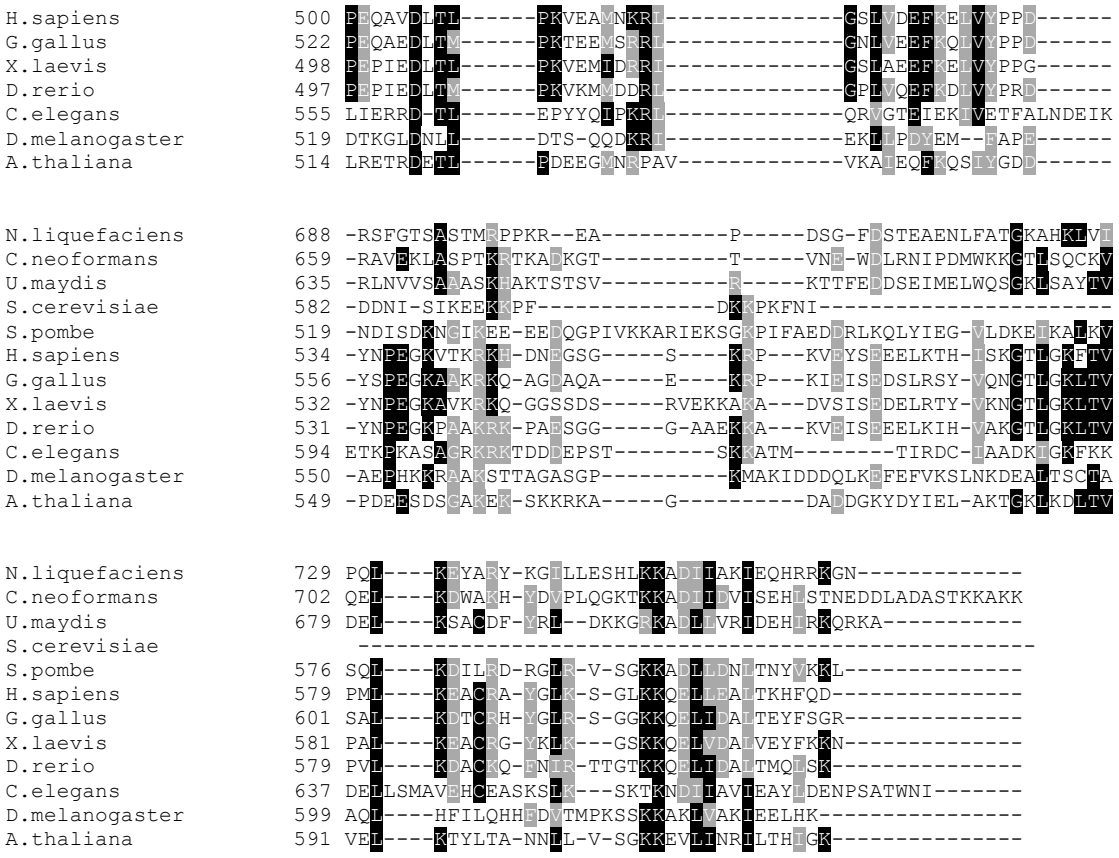


Figure S1. Protein sequence alignment of Rad51, Rad52 and Ku70. (A) Rad51 protein sequence alignment. Genbank accession numbers are as follows: *Naganishia liquefaciens* MT185602; *Cryptococcus neoformans* AFR92850; *Ustilago maydis* AAC61878; *Saccharomyces cerevisiae* AAA34948; *Schizosaccharomyces pombe* CAA80399; *Homo sapiens* AAF69145; *Gallus gallus* AAB26354; *Xenopus laevis* AAH88930; *Danio rerio* NP_998371; *Caenorhabditis elegans* CAB61038; *Drosophila melanogaster* AAF57005; *Arabidopsis thaliana* AAB37762. (B) Rad52 protein sequence alignment. Accession numbers are as follows: *Naganishia liquefaciens* MT185603; *Cryptococcus neoformans* AFR93075; *Ustilago maydis* KIS67122; *Saccharomyces cerevisiae* DAA09866; *Schizosaccharomyces pombe* CAA91896; *Homo sapiens* AAS00097; *Gallus gallus* NP_00116123; *Xenopus laevis* AAH99014; *Danio rerio* AAY43162. (C) Ku70 protein sequence alignment. Accession numbers are as follows: *Naganishia liquefaciens* MT185604; *Cryptococcus neoformans* AFR96952; *Ustilago maydis* KIS70074; *Saccharomyces cerevisiae* DAA10185; *Schizosaccharomyces pombe* CAA22471; *Homo sapiens* AAW34364; *Gallus gallus* BAA32018; *Xenopus laevis* BAA76953; *Danio rerio* NP_956198; *Caenorhabditis elegans* CAB55094; *Drosophila melanogaster* AAF54631; *Arabidopsis thaliana* OAP15286. Black shading denotes identity, grey shading denotes conservative change. Multiple sequence alignments were performed using Clustal Omega.

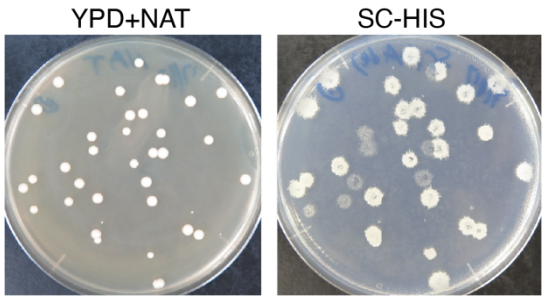


Figure S2. An example of gene targeting assay, wherein the *HIS3* gene was targeted with the *NAT* marker with 1 kb homologous arms. nourseothricin-resistant colonies were initially selected on YPD medium containing nourseothricin. The transformants were then replica-plated onto a synthetic complete medium lacking histidine (SC-HIS) to examine if the *NAT* marker replaced the *HIS3* gene.

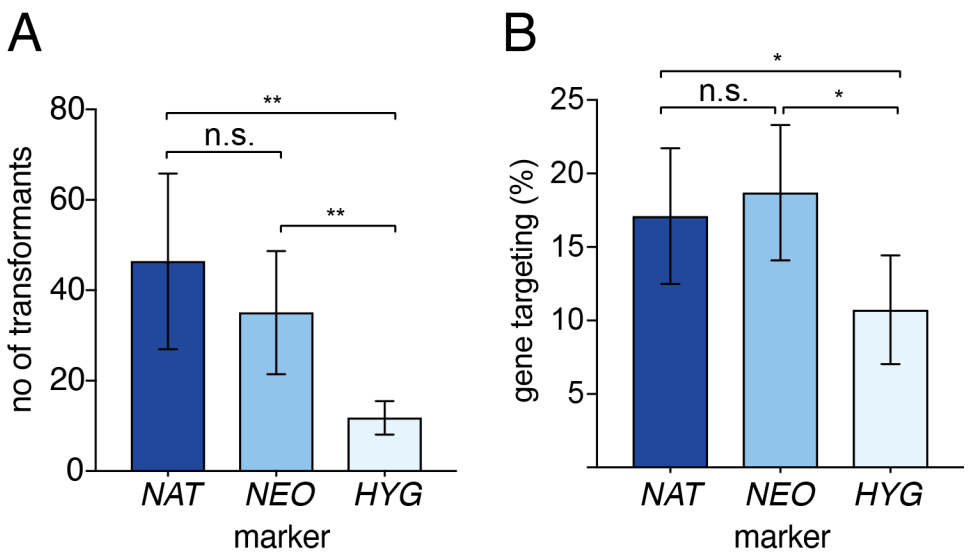


Figure S3. Comparison of transformation efficiency between the *NAT*, *NEO* and *HYG* markers. In each experiment, 3 μg of PCR-generated targeting DNA with 80 bp homology to the *HIS3* locus was used. (A) The total number of drug-resistant colonies. (B) Gene targeting efficiency obtained from each transformation. Error bars, standard deviation. n = 3 for all measurements. Statistical significance was determined by unpaired two-tailed t-test (n.s., not significant; * $p < 0.05$; ** $p < 0.01$).

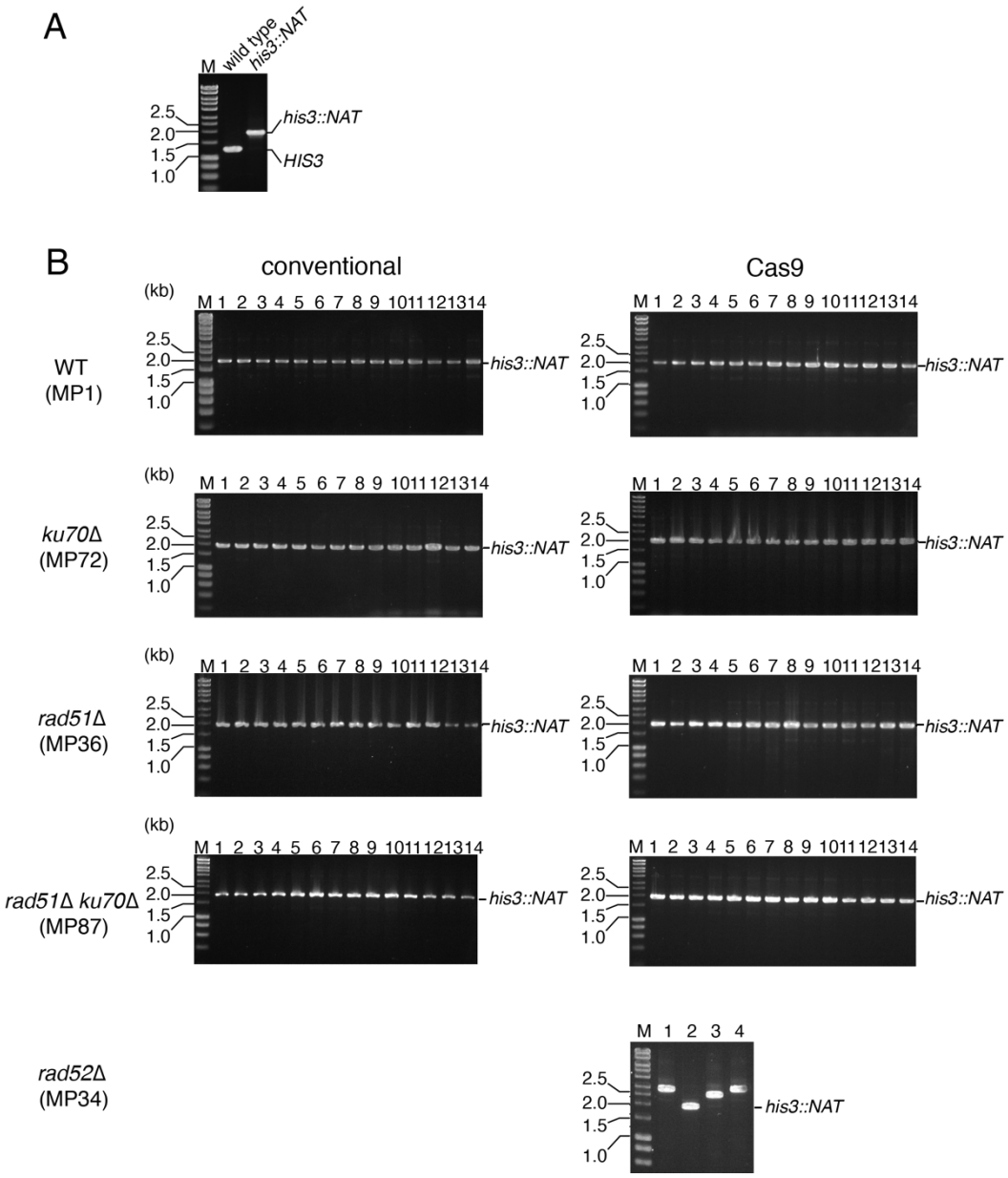


Figure S4. Gene targeting accuracy was examined by diagnostic PCR. **a** PCR amplicons carrying the wild-type *HIS3* gene and *his3::NAT* were separated by agarose gel electrophoresis. Primers used are: Pr-410 and Pr-411. **b** Gene targeting accuracy using the targeting DNA with 80 bp homology to the *HIS3* locus was examined by diagnostic PCR as in (**a**). 14 random transformants that showed nourseothricin resistance and histidine auxotrophy in the indicated strains, without Cas9 expression (conventional) or with Cas 9 expression (Cas9), were analyzed, except for *rad52*Δ, where only four transformants were obtained with Cas9 expression and none without Cas9. M, marker.

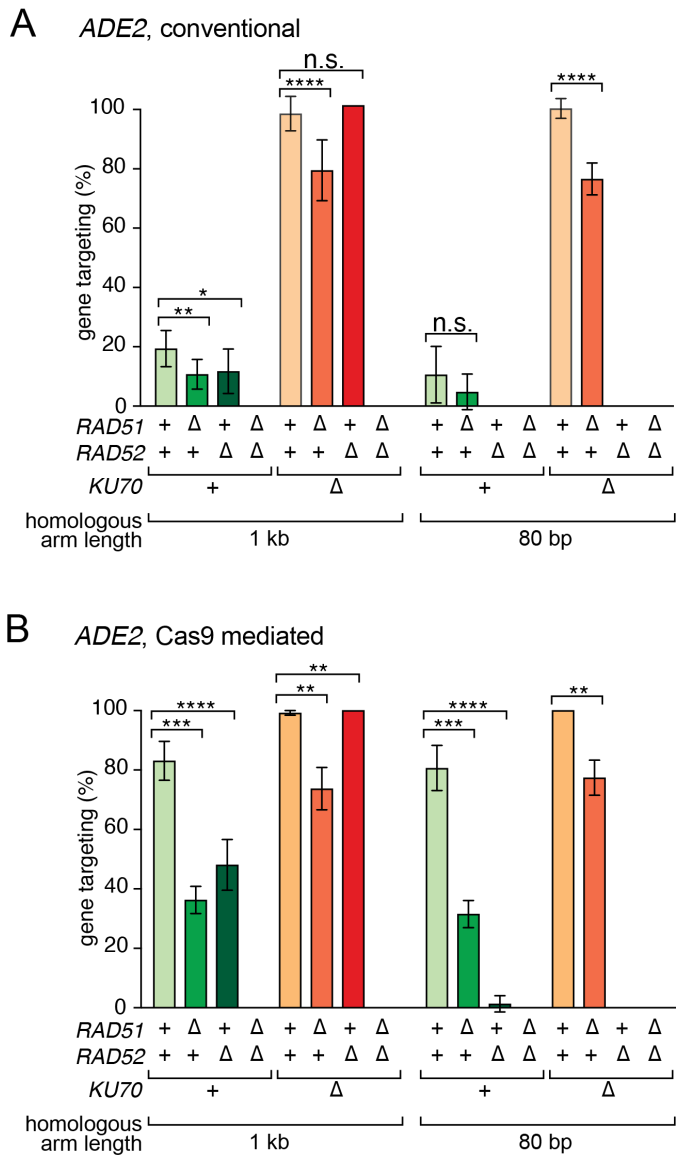


Figure S5. Impact of the absence of DSB repair genes on gene targeting efficiency. (A) Either 1 kb or 80 bp homology arms were employed in targeting DNA to the *ADE2* locus. (B) Same as (A) except gene targeting was promoted by the Cas9 system. Error bars, standard deviation. The data for wild type and the *ku70* single mutant strains presented in this figure are the same as those used in Figure 5A,B. Strains used are (Δ denotes deletion of a gene): wild type, MP1; *rad51Δ*, MP35; *rad52Δ*, MP33; *rad51*Δ *rad52*Δ, MP112; *ku70*Δ, MP72; *rad51*Δ *ku70*Δ, MP87; *rad52*Δ *ku70*Δ, MP75; *rad51*Δ *rad52*Δ *ku70*Δ, MP113. n = 9 in "conventional" experiments. n = 3 in "Cas9 mediated" experiments except that n = 9 for strains carrying the *rad52* mutation. Statistical significance was determined by unpaired two-tailed t-test (n.s., not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

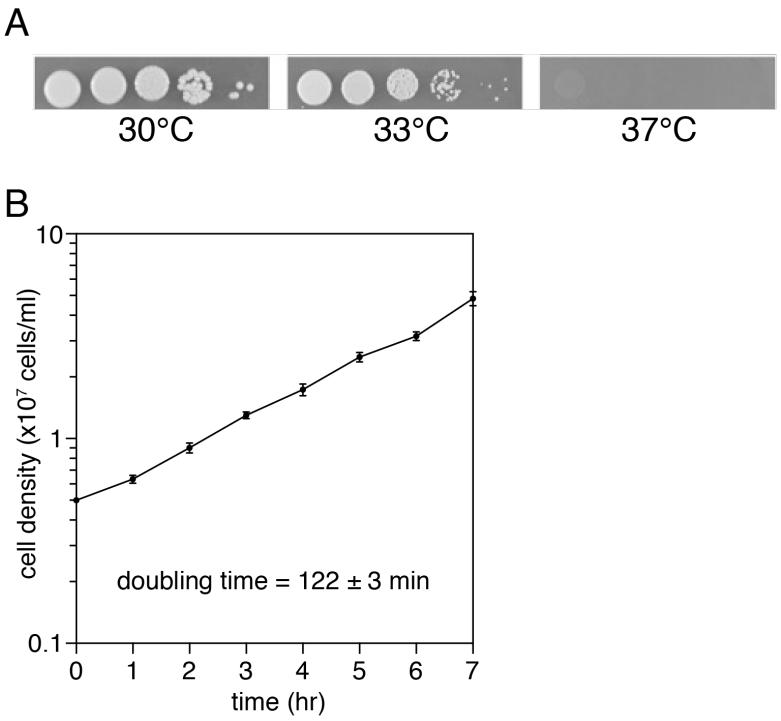


Figure S6. Cell growth of *N. liquefaciens*. (A) Analysis of temperature sensitivity of *N. liquefaciens*. 10-fold serial dilutions were made from a mid-log phase culture, and 5 *µ*l was spotted onto YPD. Plates were then incubated for three days at 30°C, 33°C, and 37°C, as indicated. (B) Growth curve of *N. liquefaciens*. An overnight liquid culture of the wild type strain (MP1) was diluted to 0.5 x 10^7 cells/ml in YPD and cell density was measured at indicated time points. n=3. Error bars, standard deviation.

Table S1. ***N. liquefaciens*** **N6 strains used in this study**

Strain	Genotype	Source
MP1	wild type	(Abe et al. 2006)
MP17	*trp2::NAT*	This study
MP21	*trp2::NAT, rad51::NEO*	This study
MP56	*rad51::NEO TRP2:RAD51*	This study
MP35	*rad51::NAT*	This study
MP36	*rad51::NEO*	This study
MP33	*rad52::NAT*	This study
MP34	*rad52::NEO*	This study
MP72	*ku70::NEO*	This study
MP39	*rad51::NAT rad52::NEO*	This study
MP87	*rad51::HYG ku70::NEO*	This study
MP75	*rad52::HYG ku70::NEO*	This study
MP89	*rad51::HYG rad52::NAT ku70::NEO*	This study
MP108	*rad51::NAT ku70::NAT*	This study
MP112	*rad51::NAT rad52::HYG*	This study
MP113	*rad51::NAT rad52::HYG ku70::NAT*	This study

Table S2. Primers used in this study

Primer ID	Sequence (5'->3')
Pr-36	ACCTCTAGATTCAAGAATCTCGTGAAATGC
Pr-23	CGTCAAGAGTGGTACCCATTTTGTTAGGTTTTTGT
Pr-22	AAAAACCTAACAAAATGGGTACCACTCTTGACGAC
Pr-25	AACGGATCCCAGTATAGCGACCAGCATTC
Pr-198	AACTCTAGAAGACCGTGACGAGCATAACG
Pr-199	GTTCAATCATTGTGATTGATTTAGATGTCTATGGC
Pr-200	ATCAATCACAATGATTGAACAAGATGGATTGCACG
Pr-201	ACCGGATCCCAGTATAGCGACCAGCATTCTGGGCGAAGAACTCCAGCAT
Pr-202	TTTTACCCATTGTGATTGATTTAGATGTCTATGGC
Pr-203	ATCAATCACAATGGGTAAAAAGCCTGAACTCACC
Pr-84	ACCACTCTTGACGACACGG
Pr-85	TGACGTTGGTGACCTCCAGC
Pr-90	AACAACCTCGGCGAATTC
Pr-91	GTATAGCGACCAGCATTC
Pr-359	CGCATTTGTTAATTCCCATTCGCCTCGTCCTCCCTCCTCATACATCGATTCCG TCCGCCAAACGAAATCAGCCATTGAAATCAACAACCTCGGCGAATTC
Pr-205	CAGGAGTAGATGACGCCTATGGTTTGCGCGCGGCCTGAATATTTACCGCCC TGGAAAGCCTGTTGTAATGACAACGGAACCAGTATAGCGACCAGCATTC
Pr-356	TGCATCCGCCCCGAGAGCGAGCAGCGGGGGGGCATGCAAGAAGATGATTT CTCGACGGAATCAGGGAATCGCATGCCCCCTCAACAACCTCGGCGAATTC
Pr-277	ACACGCGTGCGACGCAGCGTACCGACCGGTCTGAAATCGTTTTATAATCTT GCTATAACCGAAATACATTGGACGCGGCGCAGTATAGCGACCAGCATTC
Pr-357	AGAGTCCTCATCGGCTGGTCTGGGGTGTCTGGGGTGTCTGGGGTGTCTGCG CTGCATATTCCAAACATCGTATCAACATCTCAACAACCTCGGCGAATTC
Pr-360	TTTTTCCTCTCTGGCGCAGCTCGAAAGCGTCATTGTACAAGAACCACTTGTA TTTTTCTTACTTCACACAATGAAGACCTCAGTATAGCGACCAGCATTC
Pr-315	CCTCAGCCGGCCTCTCCCTGCTTCGGCTCGATCGCTCGCAAAAGGTTGAGG CACTTTGACTCACCTGAAGTTTCAGGAATTCAACAACCTCGGCGAATTC
Pr-272	GGGCCTGAACGAGCTTCCGCTGGCTTTCACTGGATGTAGTATCGTACCTATG CAGATGATACAACGGCGTAGTCTTAGACCAGTATAGCGACCAGCATTC
Pr-410	ACAAGCAGTCTGATGGCGAA
Pr-411	ATGGTAGGCGCGCAATATCT
Pr-257	TGACTGCTGCGATGAGAGAC
Pr-258	ACCGACTGGTCTGAAATGGA
Pr-156	ATGGCGTATGCAGTGCAGTG
Pr-157	GAATTCGCCGAGGTTGTTTTTCAATGGCTGATTTCGTT
Pr-158	GAATGCTGGTCGCTATACGTTCCGTTGTCATTACAACA
Pr-159	TAACGGAGAGATCTCCTCCG
Pr-195	ATTGCTTGGGCCATCTGATC
Pr-196	AGGAAGATGCAGCTGCTGAT
Pr-160	ACTCGCTTCGCAAGATAGAG
Pr-161	AGCAAGCCAAAGCAGAATTC
Pr-397	GGTGTGGACACAAAGATTCCTGATTGAGAA
Pr-399	GAATTCGCCGAGGTTGTTATTCCTGAAACTTCAGGTGA
Pr-398	GTTGTCTCCTGACGACCCTATCTCTTTTCT
Pr-400	GAATGCTGGTCGCTATACGTCTAAGACTACGCCGTTGT
Pr-299	ATCCTTGATGTCTTTGCTGG
Pr-302	GAACTAGCGGCTGAAATCGA
Pr-63	ACGATTGAGCAGTCCGAGGA
Pr-54	GAATTCGCCGAGGTTGTTGGGGGCATGCGATTCCCTGA
Pr-55	GAATGCTGGTCGCTATACCGCCGCGTCCAATGTATTTC
Pr-394	CTCCCTGACCCTGAGCTTGTTTTCGAGAAT
Pr-52	AAGCTGATGCAATCTGGTATGC
Pr-53	ATCAACCGGCAAGTCGAATG
Pr-122	AGAAACACGTCCAACACGAG
Pr-74	AGAAGTCTTTGGCGGGAGCT
Pr-373	ATCTTGGATGTCACATCGCT
Pr-375	GAATTCGCCGAGGTTGTTGATGTTGATACGATGTTTGG
Pr-376	GAATGCTGGTCGCTATACAGGTCTTCATTGTGTGAAGT
Pr-374	TCAAACAACGGACATCTCGC
Pr-377	ACCTCTACATCGATGTCCTG
Pr-378	CGTCTCCGACTCTCCTCACT
Pr-51	TATAGTCCATTGTGATTGATTTAGATGTCTATGGC
Pr-50	TAAATCAATCACAATGGACTATAAGGACCACGACG
Pr-39	ACCTCTAGATCCCCAGCATGCCTGCTATTC

Pr-43	AACAAGCTTGAACAGCTGGATGACACTGA
Pr-44	AGGTCTTCTCGAAGACCCGAAAACTACTTGAGAAACAC
Pr-41	GGGTCTTCGAGAAGACCTGTTTTAGAGCTAGAAATAGC
Pr-42	AACAAGCTTAAAAAAGCACCGACTC
Pr-406	AAGAAATTCAAGGTGCTGGGCAACA
Pr-407	TGGTCAGGGTAAACAGGTGGATGAT
Pr-290	TTTCGAAGACTGTGCTATCGCTCT
Pr-291	AAACAGAGCGATAGCACAGTCTTC
Pr-284	TTTCGCATGCTCGATGTCAATCAG
Pr-285	AAACCTGATTGACATCGAGCATGC
Pr-310	TTTCGAAGAAGACGGACCTCAACC
Pr-311	AAACGGTTGAGGTCCGTCTTCTTC
Pr-312	TTTCGTCTGTTCCGTTCAATATCC
Pr-313	AAACGGATATTGAACGGAACAGAC
Pr-394	ACCGTCGACACTCATGCTGCGCTTACCTC
Pr-395	ACCGTCGACCCATGGGTTAACTACAGGGAAGCCAAAGAAAGG
Pr-465	ACCGTTAACAGTGTCTTGCGTTGATGCGG
Pr-466	AACGTTAACACTGCTCCGTAATGCACGAG
Pr-346	TTTCGGCGCAATCACCGAACTCTA
Pr-347	AAACTAGAGTTCGGTGATTGCGCC
Pr-332	CGAGAGGTGGAAGCCACACTTCCTTTTCCACGCATAGTTCTGCCACACGCA AGAATCTCCGACAACAACTCATCGCGGCGAGACCGTGACGAGCATAACG
Pr-333	TTGTTTGAGGTGCTGAGACGAGCGATTCGAAATCAAATGGATACAACAACG ACGTTTGACTGTCCGTTCAGCTTAGGGTTCAGTATAGCGACCAGCATTC
Pr-399	CGAGAGGTGGAAGCCACACTTCCTTTTCCACGCATAGTTCTGCCACACGCA AGAATCTCCGACAACAACTCATCGCGGCGTCAACAACCTCGGCGAATTC
Pr-400	TTGTTTGAGGTGCTGAGACGAGCGATTCGAAATCAAATGGATACAACAACG ACGTTTGACTGTCCGTTCAGCTTAGGGTTCAGTATAGCGACCAGCATTC
Pr-144	ACCGGGACCCATTACACTAC
Pr-145	TGCATGCTGATAGTCCTCCG
Pr-475	ACCGGATCCTCGACTACATCGGCCGGAAG
Pr-476	AACTCTAGAGGTACCAGTTTGAATTCGGCCTGCTC
Pr-292	ACCGTGCTCCTAATGTAGCG
Pr-293	TCAGTCTTCCGGCCGATGTA
Pr-422	TCAGACCTGTCCGTATGCTG
Pr-458	ACCTCTAGAATGGCGACCCAAGAATACGC
Pr-459	ACCTCTAGATTACTTGTCACCTTCGTCTT
Pr-550	ACCTCTAGACATATGATGTCGACGCTGTCGGGACA
Pr-551	ACCTCTAGACATATGTCATGCTTTGCTGCGTTTGG
Pr-552	ACCTCTAGAATGTCCCAAGCTAAAACCCA
Pr-553	ACCTCTAGATCAATTCCCTTTGCGCCTGT
Pr-361	TTTCGTACGGCGGTCATCACGGCA
Pr-362	AAACTGCCGTGATGACCGCCGTAC
Pr-322	GCATCTCGAAGCACATCACCACACCACCAGCCAGACAAACCTCCAAGAGA AAGCTCGGATCCACGTCTTCCAGTGTTTCCAGACCGTGACGAGCATAACG
Pr-323	GTATCATCCTACTCTACATTCACATTCATACGCGCGTCCCGTATCGTGTCTG TATGCACAAAGACCCGTGGCAGGACAGTCAGTATAGCGACCAGCATTC
Pr-403	GCATCTCGAAGCACATCACCACACCACCAGCCAGACAAACCTCCAAGAGA AAGCTCGGATCCACGTCTTCCAGTGTTTCCTCAACAACCTCGGCGAATTC
Pr-404	GTATCATCCTACTCTACATTCACATTCATACGCGCGTCCCGTATCGTGTCTG TATGCACAAAGACCCGTGGCAGGACAGTCAGTATAGCGACCAGCATTC
Pr-174	ATTCCACTCTCGGTCTCGGT
Pr-175	TCCGAAGATACGAGTCGCTG
Pr-500	TTTCGATATCGCATCGACCGTCAT
Pr-501	AAACATGACGGTCGATGCGATATC
Pr-506	CCATCTATCAAGCACCCACCCTTCCTCGCACAATCGACAGGAAGCGATCAA GAATACTTGCTAATAAATGCGAGATCATCAGACCGTGACGAGCATAACG
Pr-507	CCCGTTCACACGGCGTTGGCGGATAAGAAAAAGCAATGCGGTGGAAGGGC ATCAGCACTTTTTGAGAGCCATCAACCCCTCAGTATAGCGACCAGCATTC
Pr-583	CCATCTATCAAGCACCCACCCTTCCTCGCACAATCGACAGGAAGCGATCAA GAATACTTGCTAATAAATGCGAGATCATCTCAACAACCTCGGCGAATTC
Pr-584	CCCGTTCACACGGCGTTGGCGGATAAGAAAAAGCAATGCGGTGGAAGGGC ATCAGCACTTTTTGAGAGCCATCAACCCCTCAGTATAGCGACCAGCATTC
Pr-504	ACCTGAGAGCTGATCGATGG
Pr-505	AGACACGCTTATCGGAATCG
Pr-415	TTTCGAGGCCATCGAGGCACTGGA
Pr-416	AAACTCCAGTGCCTCGATGGCCTC
Pr-620	TTTCGACTGGGCACAACAGACAAT
Pr-621	AAACATTGTCTGTTGTGCCCAGTC
Pr-622	TTTCGAGGGCGTGGATATGTCCTG

Pr-623	AAACCAGGACATATCCACGCCCTC

Table S3. Number of transformants analyzed in Figures 5 and S5

Δ denotes deletion of a gene.

(1) *HIS3*

	wild type	*rad51*Δ	*rad52*Δ	*rad51*Δ *rad52*Δ	*ku70*Δ	*ku70*Δ *rad51*Δ	*ku70*Δ *rad52*Δ	*ku70*Δ *rad51*Δ *rad52*Δ
1 kb (conventional)	124	111	69	28	96	107	23	14
80 bp (conventional)	104	116	27	45	88	86	15	15
1 kb (Cas9)	387	453	107	39	216	361	65	11
80 bp (Cas9)	524	388	164	41	517	309	15	13

(2) *ADE2*

	wild type	*rad51*Δ	*rad52*Δ	*rad51*Δ *rad52*Δ	*ku70*Δ	*ku70*Δ *rad51*Δ	*ku70*Δ *rad52*Δ	*ku70*Δ *rad51*Δ *rad52*Δ
1 kb (conventional)	132	108	74	46	93	95	24	14
80 bp (conventional)	115	124	39	59	92	88	14	15
1 kb (Cas9)	380	401	148	35	341	339	74	15
80 bp (Cas9))	391	367	101	59	435	443	17	15