A Candida albicans CRISPR system permits genetic engineering of essential genes and gene families.

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

Citation

As Published
http://dx.doi.org/10.1126/sciadv.1500248

Publisher
American Association for the Advancement of Science (AAAS)

Version
Final published version

Accessed
Sun Jun 18 10:27:54 EDT 2017

Citable Link
http://hdl.handle.net/1721.1/96395

Terms of Use
Creative Commons Attribution-Noncommercial-Share Alike

Detailed Terms
http://creativecommons.org/licenses/by-nc/4.0/
A *Candida albicans* CRISPR system permits genetic engineering of essential genes and gene families

Valmik K. Vyas,¹ M. Inmaculada Barrasa,¹ Gerald R. Fink¹,²*–

*Corresponding author. E-mail: gfink@wi.mit.edu*

**MOLECULAR GENETICS**

*Candida albicans* is a pathogenic yeast that causes mucosal and systematic infections with high mortality. The absence of facile molecular genetics has been a major impediment to analysis of pathogenesis. The lack of meiosis coupled with the absence of plasmids makes genetic engineering cumbersome, especially for essential functions and gene families. We describe a *C. albicans* CRISPR system that overcomes many of the obstacles to genetic engineering in this organism. The high frequency with which CRISPR-induced mutations can be directed to target genes enables easy isolation of homozygous gene knockouts, even without selection. Moreover, the system permits the creation of strains with mutations in multiple genes, gene families, and genes that encode essential functions. This CRISPR system is also effective in a fresh clinical isolate of undetermined ploidy. Our method transforms the ability to manipulate the genome of *Candida* and provides a new window into the biology of this pathogen.

**INTRODUCTION**

*Candida albicans*, the major fungal pathogen of humans, causes infections that can be fatal in immunocompromised individuals (1–3). The absence of facile molecular genetics has been a major impediment to analysis of pathogenesis. *Candida* is diploid, lacks any known meiotic phase, and has no plasmid system. Diploidy is a problem because obtaining a phenotype for a recessive mutation requires engineering the mutation in both alleles of a gene. In the absence of meiosis and plasmids, diploidy makes it difficult to investigate genes encoding vital functions. In more tractable organisms, loss-of-function mutations in essential genes can be maintained in a diploid and identified after meiosis. Alternatively, they can be maintained in mutant strains with a plasmid carrying a functional copy. The difficulty in making multiple knockouts becomes insurmountable when analyzing gene families, where simultaneous mutation at multiple loci is required for a phenotype. The *Candida* genome is populated by many gene families, including more than 120 drug efflux pumps (4–6). This redundancy impedes analysis of the resistance to antifungal agents because the construction of multiple mutations in the members of these families is beyond current technology. These pumps also give *Candida* a high inherent drug resistance, rendering all but one drug resistance marker useless.

An added complexity to genetics in *Candida* is that the chromosome number is not rigidly controlled, so that many strains contain one or more additional copies of a chromosome (2n + 1) (7–10). Such variation in chromosome number can amplify the difficulty in achieving a homozygous recessive mutation. Moreover, it is a serious hurdle to the analysis of fresh clinical isolates, which can have multiple aneuploidies (8). Although isolates of different ploidy have been described, they are unstable and mostly revert to the diploid state (11).

We have developed a CRISPR system in *C. albicans* that overcomes many of these roadblocks by permitting efficient genome editing. Our system, shown in Fig. 1D, consists of a *Candida*-compatible Cas9 nuclelease and a synthetic guide RNA (sgRNA) that directs Cas9 to cleave regions in the genome that hybridize to the 20–base pair (bp) guide (or protospacer) from the sgRNA when it is followed by the sequence NGG (PAM or the protospacer adjacent motif). This system has been successfully imported to diverse kingdoms ranging from fungi to plants and animals [reviewed in (12, 13)]. However, most of these systems do not pose the unique set of constraints found in *Candida*.

The CRISPR system we have constructed in *Candida* provides an efficient method for generating homozygous mutations of a target gene in one transformation in both laboratory and clinical isolates. Moreover, in a single transformation, it is possible to obtain homozygous mutations of multiple family members or in genes of vital function. The ability to analyze essential genes provides an opportunity to explore potential antifungal targets. Computational analysis shows that most genes in the *Candida* genome can be uniquely targeted by our system.

**RESULTS**

**A CRISPR system for Candida**

To create a CRISPR system for *Candida*, several challenges had to be overcome: the Cas9 gene had to be recoded because the leucine CUG codon is predominantly translated as serine, there are no autonomously replicating plasmids, and there are no expression systems for small RNAs. To express a *Candida*-compatible Cas9, we synthesized a *Candida/Saccharomyces* codon–optimized version of Cas9 (*CaCas9*) that avoids the use of the CUG codon, ensuring compatibility with all *Candida* clade species. The *Cas9* gene is fused to sequences encoding the SV40 nuclear localization signal and FLAG tag for in-frame fusion to the N′ end of the gene. The *Cas9* from this construct is expressed from the constitutive *ENO1* promoter at the plasmid integration site. Because there are no autonomously replicating plasmids in *Candida*, we integrated this construct by transformation into SC5314 at the *ENO1* locus. The RNA polymerase III (Pol III) promoter *SNR52* was used to express sgRNAs necessary for Cas9 targeting.

For most genes, *Candida* diploids require knockout of both alleles of a gene to obtain a phenotype. As a proof of principle for our *Candida* CRISPR system, we chose to target *ADE2* because the *ade2* mutation confers an easily visible red phenotype. The red *ade2* phenotype is manifest among white *ADE2/ADE2* diploids only if both alleles of the *ADE2* gene are simultaneously nonfunctional (*ade2/ade2*).

We created two systems based on the design principles listed above. The "Duet system," shown in Fig. 1A, requires the sequential integration of two plasmids. Integration of the *Cas9* expression plasmid at the
ENO1 locus is first selected with nourseothricin (Nat). By induction of the flippase gene and subsequent excision of the Nat<sup>R</sup> gene, it is possible to use this marker again for selection. The second plasmid for expression of the sgRNA against ADE2 (targeted to the RP10 locus) is cotransformed with a mutagenic double-stranded oligonucleotide. This oligonucleotide is complementary to ADE2 and contains a mutation to the PAM sequence and a premature UAA stop codon (sequences shown in Fig. 2B). No defect in the growth rate of Cas9-expressing yeast cells was detected on YPD medium (see Materials and Methods).

The "Solo system" (Fig. 1B) consolidates the CRISPR system with the Nat resistance marker. Targeting additional genes would require the introduction of additional guides. To this end, we created a version of the Solo plasmid with a recyclable Nat cassette (fig. S1) that permits the introduction of additional guide sequences to target other loci. Both the Duet and Solo systems feature simplified ligation of annealed oligos into the site created with BsmBI, leaving no extraneous sequences (Fig. 1C).

**Highly efficient CRISPR mutagenesis in Candida.** Both the Duet and Solo systems produce red ade2/ade2 transformants at high frequency (Fig. 2A and figs. S2A and S3B); each required a functional Cas9, an sgRNA against ADE2, and the complementary repair template spanning the cut site. In the absence of any one of these components, only white ADE2<sup>+</sup> colonies were obtained (figs. S2 and S3). The Duet system produced 20 to 40% red colonies among the transformants, and these were authentic CRISPR-induced mutations, as sequencing of the ade2/ade2 mutants revealed the UAA and the PAM mutation in the ade2 gene (Fig. 2B). The Solo system was more efficient than the Duet system; 60 to 80% of the transformants were red ade2/ade2 mutants (Fig. 2A and fig. S3B). The frequency of targeting was so high that transformation with Solo plasmid and the repair template for ade2 without any selection for integration of either the Solo Cas9 cassette or the repair template yielded red ade2/ade2 mutants at a rate of 2 to 3% (fig. S3D).

The Solo system appears to be generally applicable for mutagenesis, because we were able to easily create mutations or truncations in URA3, RAS1, MltA1, MltA2, and TPK2 (Fig. 2 and Fig. S4). Transformation plates for RAS1V13 mutants provided an easy visual phenotype for identification based on colony morphology or glycogen staining with iodine (Fig. 2D). Our isolation of RAS1 truncation mutants is particularly significant, because this mutation significantly reduces the growth rate (Fig. 2E) (14). From our transformation plates, we obtained slow-growing isolates at a similar frequency to that of wrinkly colonies for RAS1V13.

The high efficiency of the Candida CRISPR system in making homozygous knockouts suggested that it might be possible to knock out multiple members of a gene family with a single guide. To test this possibility, we attempted to knock out both CDR1 and CDR2, members of the multigene drug efflux pump encoding family. Loss of cdr1 or cdr2 increases sensitivity to the clinically useful azole antifungal agents (15). To this end, we designed an sgRNA that targeted both genes and a repair template that had homology to both CDR1 and CDR2. Our repair template contained a stop codon as well as a unique restriction site, which enabled us to quickly genotype transformants (Fig. 3A). Among the transformants, we identified drug-sensitive mutants of both genes and a repair template that had homology to both CDR1 and CDR2. Our repair template contained a stop codon as well as a unique restriction site, which enabled us to quickly genotype transformants (Fig. 3A).
This experiment showed that four loci can be targeted with high efficiency with a single guide. Moreover, it demonstrates that one does not need a visible phenotype to identify the intended transformants. The *Candida* CRISPR system is so efficient that ~20% of the transformants were drug-sensitive. Thus, even mutants with modest phenotypic differences from wild type can now be easily identified.

Drug resistance to azoles is a problem in the clinical treatment of *Candida* infections. Although several mechanisms contribute to this resistance [reviewed in (16)], up-regulation of drug pumps is a common cause. To determine whether our CDR1/CDR2 CRISPR guides could be used to characterize a recent fluconazole–hyper-resistant clinical isolate Can90, we transformed this strain with the appropriate guides and repair templates, as we did for SC5314. We easily identified Can90, we transformed this strain with the appropriate guides and repair templates, as we did for SC5314. We easily identified *cdr1/cdr1 cd2/cd2* homozygous double mutants (three of seven transformants tested) and showed that these transformants no longer displayed the hyper-resistance to fluconazole or cycloheximide displayed by the parental clinical isolate Can90 (Fig. 3B and fig. S5).

The ease of *Saccharomyces* genetics largely rests on the ability to easily produce multiple mutations in a given strain. Without the ability to make recombinant haploids through meiosis, this is a difficult feat to achieve in *Candida*. To see if we could circumvent this limitation with our CRISPR system, we cotransformed our Solo CDR system alongside the sgRNA-expressing Duet ADE2 vector. We were able to identify strains that were simultaneously mutated at *ADE2*, *CDR1*, and *CDR2* (six loci) from a single transformation (Fig. 3C).

**CRISPR to target essential functions**

We obtained homozygous loss-of-function mutations in essential genes of *C. albicans* using the CRISPR system by creating conditional alleles. Null alleles of *DCR1*, which is required for ribosomal RNA processing, are lethal at low temperature but viable at high temperature (17). Transformation of SC5314 was carried out using the Solo CRISPR plasmid containing a guide directed against *DCR1* and a repair template that introduced a stop codon. The transformation plates were incubated at 37°C, and transformants were screened for growth at either 37° or 16°C to identify candidate *dcr1/dcr1* mutants. We isolated a number of *dcr1/dcr1* mutants that failed to grow at 16°C and contained the signature nonsense mutation (Fig. 4A and fig. S4).

Another approach to obtaining null mutations in lethal functions is to replace the resident functional genes with the gene under the control of the inducible MAL2 promoter. To test if we could quickly introduce a regulatable promoter for *SNF1*, which is essential (18, 19), we created a guide that cut in the *SNF1* promoter region and inserted a *MAL2* promoter fragment with flanking homology to resident sequences, permitting *SNF1* to be transcribed on maltose but not glucose. We plated transformation mixtures onto selective maltose plates, and replica-plated these onto maltose (permissive) or glucose (restrictive) medium. We were able to identify several transformants that only grew in maltose and confirmed that they were maltose promoter integrants (Fig. 4B and fig. S6B), verifying the essential nature of *SNF1*.

Both previous attempts to knock out *SNF1* function relied on the failure to obtain a homozygous gene replacement (18, 19) without the presence of *SNF1* elsewhere in the genome. This indirect evidence suggests that the Snf1 function is essential and implies that the kinase activity of Snf1 is required. It does not rule out the possibility that only the protein itself, but not the kinase activity, is required. To discriminate between these possibilities, we generated Solo system guides for *SNF1*, and repair templates that mutate Lys81 to arginine in the ATP (adenosine 5’-triphosphate)–binding pocket. Mutation at this conserved position either eliminates or vastly diminishes kinase activity in *Saccharomyces* and human Snf1/AMPK (20, 21). We noticed that our *K81R* CRISPR transformation plates contained ~40% wrinkled colonies (fig. S6A), which, upon further analysis, turned out to be homozygous for *snf1*-K81R (fig. S6, B

---

**Fig. 2. *C. albicans* CRISPR is an efficient mutagenesis system.** (A) *Candida* CRISPR efficiently mutagenizes both *ADE2* loci in SC5314. The transformation mix contained pV1081 and a mutagenic repair template (see Materials and Methods and the Supplementary Materials). Omission of Cas9, sgRNA, or a repair template generates both *ade2* mutants. (B) Sequence of the *ADE2* locus in wild-type (WT) and mutant isolates. (C) Assay for *ura3/ura3* transformant on 5-fluoroorotic acid (FOA) plates. FOA permits growth of *ura3/ura3* but not *URA3* strains. (D) Wrinkled colony morphology of *RAS1V13* on selective plates (top) and glycogen accumulation defect/wrinkled colony morphology of *RAS1V13* (bottom). Glycogen accumulation is visualized by exposing yeast to iodine vapors, which stains glycogen red. WT (left) has a smooth morphology and stains red due to accumulated glycogen (left), whereas *RAS1V13* (right) has a wrinkled morphology and fails to stain. (E) Truncation of *RAS1* at position 13 [*ras1(TAA)13*] reduces growth rate.
and C). The snf1-K81R/snf1-K81R strains are unable to grow on maltose (Fig. 4B), consistent with the *Saccharomyces snf1* mutant’s failure to grow on non-glucose carbon sources (20, 22). The additional phenotypes of cold sensitivity (Fig. 4C) and defective filamentous growth (Fig. 4D) are also seen in *snf1* mutants in *Saccharomyces* (23–25). In addition, we found that *snf1*-K81R was hypersensitive to fluconazole, suggesting that Snf1’s stress response function is required for activation of fluconazole resistance (fig. S6).

**CRISPR-accessible sites in the genome**

We searched the most recent diploid assembly of the *C. albicans* genome database (26) for Cas9 recognition motifs—N20 followed by a PAM sequence—and selected only those sequences that overlap with annotated features. Of the 6466 genes in the *Candida* genome, 6341 can be targeted uniquely by 601,770 guides. Of those guides, 551,175 can direct cleavage at both alleles, whereas 59,595 target only one of the two. A small subset of these guides target more than one location in the same gene (genes with internal repeats). The sequences of each of these guides can be found in the Supplementary Data Files. In addition, we identified 49,195 guides that target more than one putative gene sequence, without targeting non-genic sequences. Such sequences can be found for 6023 genes. These can be used to target certain motifs or gene families for simultaneous mutagenesis, as we did with *CDR1* and *CDR2*, and are listed in the Supplementary Data Files.

---

**Fig. 3.** CRISPR permits simultaneous targeting of *CDR1* and *CDR2*, which mediate resistance to fluconazole and cycloheximide. (A) Sequence of *CDR1* and *CDR2* loci and verification by digestion. (B) Mutation of *CDR1* and *CDR2* sensitizes SC5314 (left) and fluconazole-resistant clinical isolate Can90 (right) to fluconazole (0.41 μg/ml for SC5314, 200 μg/ml for Can90). Different fluconazole concentrations were used for each strain background, because the Can90 isolate had much greater resistance. Solid lines indicate medium without fluconazole, and dotted lines indicate medium with fluconazole. (C) Simultaneous mutation of three genes (six sites) in a single transformation, and the resulting phenotypes. Left panel is YPD, and right panel is YPD plus cycloheximide at 400 μg/ml. The poorer growth on petri plates of the ade2 cdr1 cdr2 triple mutants is reflected in liquid growth on fluconazole. The ade2 *CDR1 CDR2* has a doubling time of 6 hours, whereas the ade2 cdr1 cdr2 mutant has a doubling time of 12 hours when grown in fluconazole (1.2 μg/ml).
**DISCUSSION**

The CRISPR system we have developed circumvents many of the challenges unique to the genetic manipulation of *C. albicans*. A major impediment has been the paucity of antibiotic resistance markers, which, coupled with diploidy and variable transformation frequency, makes knockouts of a single function a considerable task. Our system enables a single transformation experiment to mutate both copies of a gene or to delete several copies of a multigene family, resulting in a discernable phenotype. Furthermore, CRISPR-induced mutations are observed at such a high frequency that they can be made without selection. Using a combination of guides, we showed that one can knock out both copies of a gene or gene families. Gene families, which have been historically difficult to study, can now be modified in a single experiment.

Our finding that CRISPR works effectively in a recent antifungal-resistant clinical isolate suggests a route to characterize clinical isolates of drug-resistant strains of *Candida*. The contribution of each of the many mechanisms that render *Candida* resistant to antifungals—changes in ergosterol biosynthesis, up-regulation of multidrug efflux and uptake pumps, changes in cell wall composition, and overexpression or mutation of drug target genes—can now be directly measured in clinical isolates using appropriate guides.

**MATERIALS AND METHODS**

**Strains and media**

*C. albicans* strain SC5314 was used for all experiments unless otherwise noted. The fluconazole-resistant *C. albicans* strain Can90 was provided by the Massachusetts General Hospital. Yeast strains were grown in YPD medium supplemented with 0.27 mM uridine and selected using Nat at a concentration of 200 μg/ml. Transformations were performed using the lithium acetate method (27). Flipout of Nat<sup>R</sup> gene from Cas9-expressing Duet vector pV1025 was done by induction of flippase by growth in Difco yeast carbon base with bovine serum albumin, and screening for isolates that had lost the Nat<sup>R</sup> gene. Filamentation experiments were performed with yeast grown overnight in liquid YPD, washed twice in RPMI 1640 medium (cat. #22400-105, Life Technologies) supplemented with 10% FBS, and incubated in RPMI + 10% FBS for the indicated time at a starting optical density (OD) of 0.1. Growth curves were performed in a clear-bottom 96-well plate and incubated with shaking at 30°C in a Tecan Saphire<sup>2</sup> plate reader, reading OD at 600 nm every 5 min for the indicated time. YPD-grown overnight yeast cultures were used to inoculate these wells to an initial OD of 0.05. CRISPR-mutagenized loci were verified by sequence analysis of PCR products amplified from the target locus and by restriction digest where applicable.

**Plasmids/DNA**

Plasmids for CaCas9 Duet and Solo systems are listed in the Supplementary Materials. The CaCas9 DNA was synthesized by BioBasic, with codons optimized for expression in both *C. albicans* and *Saccharomyces cerevisiae*. All key components were verified by sequencing and restriction analysis, and vector sequences will be provided upon request. Solo and/or Duet vectors (5 to 10 μg) were linearized by digesting...
with Kpn1 and Sac1 before transformation for efficient targeting to the ENO1 and/or the RP10 locus. Purified repair templates (3 μg) were transformed along with the guide expression plasmids for the Solo or Duet systems. Repair templates were generated with 60-bp oligonucleotide primers containing 20-bp overlap at their 3' ends centered on the desired mutation point. Primers were extended by thermocycling with ExTaq. Most guides were either immediately adjacent to or within 15 bp of the desired mutagenesis point. Phosphorylated and annealed guide sequence–containing primers were ligated into CIP (calf intestinal phosphatase)–treated BsmBI-digested parent vectors as depicted in Fig. 1C. Correct clones were identified by sequencing.

Computational analysis
We searched the diploid C. albicans genome sequence for matches to the patterns N_{20}(NGG) or (CCN)_{N_{20}} and selected only sequences that overlapped with features found in the most recent gff file available from the Candida Genome Database (C. albicans_SSC3514_version-A22-s05-m01-r03_features.gff), excluding the chromosomes themselves. We removed any targets that have 6Ts in the 20 bp before the NGG, because this would result in premature termination from Pol III promoters. Because matches 13 nucleotides proximal to a PAM sequence (NGG or CCN) would also result in a cut to the genome, we searched all sites that would be targeted by each 13 bp proximal to any PAM motif in the genome (Supplementary Data Files). We also did the same search with 12 bp for a stricter cutoff. We annotated and classified the target sequences based on the number of genes and intergenic regions they targeted.

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/1/3/e1500248/DC1

Table S1. Plasmids used in this study. Correct clones were identified by sequencing.

REFERENCES AND NOTES