nature microbiology

ARTICLES

DOI: 10.1038/s41564-017-0043-0

A CRISPR–Cas9-based gene drive platform for genetic interaction analysis in *Candida albicans*

Rebecca S. Shapiro[1,2,3], Alejandro Chavez[3,4,5,11], Caroline B. M. Porter[1,2], Meagan Hamblin[2], Christian S. Kaas[3,5,6], James E. DiCarlo[5,7], Guisheng Zeng[8], Xiaoli Xu[8], Alexey V. Revtovich[9], Natalia V. Kirienko[9], Yue Wang[8,10], George M. Church[3,5*] and James J. Collins[1,2,3*]

Candida albicans **is the leading cause of fungal infections; yet, complex genetic interaction analysis remains cumbersome in this diploid pathogen. Here, we developed a CRISPR–Cas9-based 'gene drive array' platform to facilitate efficient genetic analysis in *C. albicans*. In our system, a modified DNA donor molecule acts as a selfish genetic element, replaces the targeted site and propagates to replace additional wild-type loci. Using mating-competent *C. albicans* haploids, each carrying a different gene drive disabling a gene of interest, we are able to create diploid strains that are homozygous double-deletion mutants. We generate double-gene deletion libraries to demonstrate this technology, targeting antifungal efflux and biofilm adhesion factors. We screen these libraries to identify virulence regulators and determine how genetic networks shift under diverse conditions. This platform transforms our ability to perform genetic interaction analysis in *C. albicans* and is readily extended to other fungal pathogens.**

Fungal pathogens are a leading cause of human mortality among the ever-growing immunocompromised population. The most pervasive cause of fungal infections is *Candida albicans*, an opportunistic pathogen present in the human microbiome[1]. As with other fungal infections, treating *C. albicans* is challenging due to its close evolutionary relationship with its human host. Indeed, identifying antifungal agents without toxicity to humans has proved incredibly challenging—there are few classes of antifungals in clinical use and the emergence of antifungal resistance is common, resulting in limited treatment options for these deadly infections[1]. In addition, *C. albicans* readily forms robust biofilms on medical devices—including urinary and venous catheters—that are highly resistant to antifungal treatment[2] and result in the highest crude mortality rate for medical-device-associated infections[3].

A thorough understanding of the biology and pathogenesis of *C. albicans* demands a comprehensive genetic toolkit to dissect the complex cellular signalling associated with virulence. One of the most powerful ways to rapidly determine underlying genetic interactions and global network topologies is by generating double-deletion mutants and comparing their resultant phenotype with that of the parental single-gene mutants[4]. This deceptively simple approach allows one to predict whether two genes operate in the same, parallel or different biochemical pathways; in doing so, one can uncover roles for uncharacterized genes and assign new activities to previously studied factors. To date, this type of research has been limited by *C. albicans*' diploid nature, incomplete sexual cycle and atypical codon usage. Developing new technologies to overcome these technical hurdles and facilitate large-scale genetic analysis is of imminent importance.

The RNA-guided endonuclease Cas9 has revolutionized our ability to perform precise genetic modifications within a diversity of organisms, including fungal pathogens[5–9]. By changing the standard architecture of the donor DNA molecule delivered to replace the native genomic locus targeted by Cas9, one can create a gene drive, which is a selfish genetic element used as a mechanism to bias inheritance within engineered populations. A gene drive will readily replace its genetic target site and, on mating, will propagate and replace the incoming wild-type locus at high efficiency[10] (Fig. 1). A further enabling technology is the discovery of rare, mating-competent *C. albicans* haploids, which can be genetically manipulated and mated to form stable diploid cells[11].

Here, we develop a CRISPR–Cas9-based genome-editing platform, which we use to create single- and double-gene deletion mutants in *C. albicans* with unprecedented efficiency, and perform the first large-scale genetic epistasis analysis in a fungal pathogen. By exploiting Cas9-based gene drives and mating-competent *C. albicans* haploid lineages, we develop the 'gene drive array' (GDA)—a strategy for tractable genome manipulation and rapid generation of homozygous deletion mutants in the diploid pathogen. We showcase the utility of this technology for the facile generation of double-deletion libraries, and uncover complex genetic interactions underlying drug resistance and biofilm formation within this clinically relevant fungal pathogen.

Results

A *C. albicans* genetic deletion technology using a Cas9-based gene drive platform. To overcome the technical limitations of performing genetic interaction analysis in *C. albicans*, we developed

[1]Department of Biological Engineering, Institute for Medical Engineering and Science, Synthetic Biology Center, Massachusetts Institute of Technology, Cambridge, MA, 02139 USA. [2]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [3]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA. [4]Department of Pathology, Massachusetts General Hospital, Boston, MA 02114, USA. [5]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. [6]Department of Expression Technologies 2, Novo Nordisk A/S, Maaloev 2760, Denmark. [7]Department of Ophthalmology, Columbia University, New York, NY 10032, USA. [8]Institute of Molecular and Cell Biology, Agency for Science, Technology & Research, 61 Biopolis Drive (Proteos), Singapore 138673, Singapore. [9]Department of BioSciences, Rice University, Houston, TX 77005, USA. [10]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117549, Singapore. Present Address: [11]Department of Pathology and Cell Biology, Columbia University College of Physicians and Surgeons, New York, 10032 NY, USA.
Rebecca S. Shapiro and Alejandro Chavez contributed equally to this work. *e-mail: jimjc@mit.edu; gchurch@genetics.med.harvard.edu

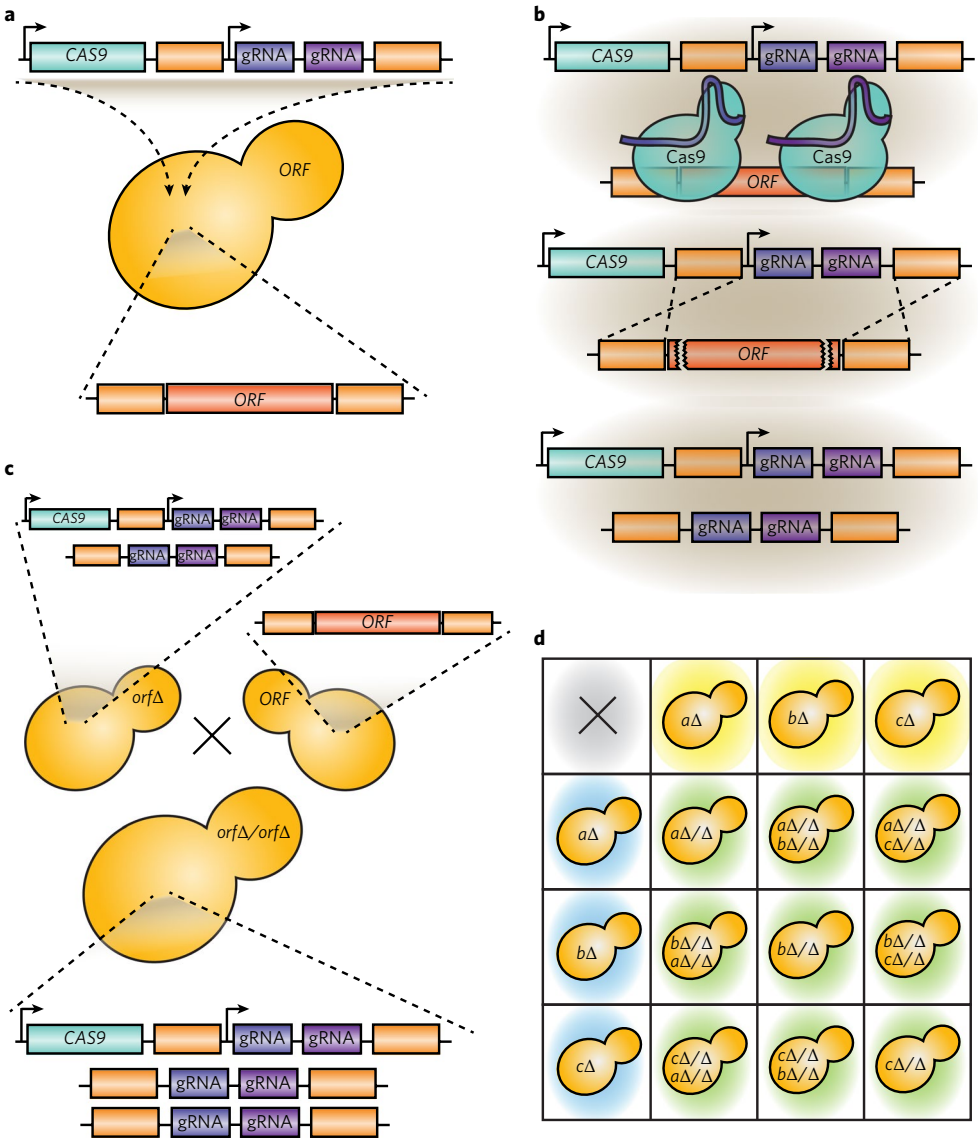

Fig. 1 | A synthetic Cas9 gene drive system for targeted homozygous deletion in *C. albicans*. **a**, The gene drive system can be used to target any gene in the *C. albicans* genome ('*ORF*') for deletion. Cas9 has been optimized for expression in *C. albicans*, and two small guide RNAs (gRNAs) are flanked by homologous sequences upstream and downstream of the *ORF* targeted for deletion. This construct is transformed into *C. albicans* on a plasmid that stably incorporates into the genome. **b**, When the gene drive is transformed into *C. albicans*, Cas9 is expressed and targeted by the gRNAs to regions at the 5′ and 3′ of the *ORF* inducing double-strand breaks. These double-strand breaks are repaired through homologous recombination, using the regions upstream and downstream of the gRNA on the plasmid as a repair template, leaving a complete deletion of the *ORF*. **c**, In our system, when a *C. albicans* haploid strain containing a gene drive (*orf*Δ) is mated to a strain containing a wild-type copy of the same *ORF*, the resultant diploid cell will be a homozygous deletion of the *ORF* (*orf*Δ/*orf*Δ) because the gene drive further propagates into the incoming wild-type locus during mating. **d**, Combinatorial mating of *C. albicans* haploids containing gene drives targeting different genes will result in a matrix of *C. albicans* diploid strains that are homozygous double-gene deletion mutants.

a CRISPR—Cas9-based gene drive array (GDA) platform that can efficiently generate single- and double-homozygous gene deletions. In our system, a plasmid contains *CAS9* and a pair of guide RNAs (gRNAs) that direct Cas9 to create double-strand breaks in the open reading frame (*ORF*) of the gene targeted for deletion. The gRNA module is flanked by regions of homology to the sequences upstream and downstream of the Cas9-targeted locus (Fig. 1a). When Cas9 cuts the *ORF*, the cell uses the homologous sequences present within the gRNA module to repair the break; by doing so, the entire *ORF* is deleted and replaced with the targeting gRNAs, generating a 'drive allele' (Fig. 1b). When drive-containing haploid cells are mated to wild-type cells, the gRNA-modified locus will initiate another round of cutting, which converts the incoming wild-type allele into a drive-containing variant (Fig. 1c). Once cells contain a functional drive, they can readily convert a heterozygous deletion into a homozygous deletion, facilitating the rapid generation of homozygous deletion mutants in this diploid pathogen.

This technology can be combined with mating-competent *C. albicans* haploid lineages[11] to create higher-order genetic mutants. Transforming our gene drive into haploids and mating them with wild-type haploids yields diploids that are deleted for the target gene on both chromosomes (Fig. 1c). Thus, our system functions counter to classical Mendelian genetics, yielding homozygous deletion strains following mating. Similarly, mating *C. albicans* haploids that

each contain a drive targeting different genes for deletion yields a diploid that is a homozygous double deletion for these two genes (Fig. 1d). This GDA platform enables us to analyse higher-order genetic mutants, and positions us to conduct large-scale genetic interaction analysis in *C. albicans*.

Optimized gene-drive deletion platform in *C. albicans* allows for efficient and precise creation of homozygous single- and double-gene deletions. Designing our GDA platform in *C. albicans* required several optimization steps. We designed gene drive plasmids targeting the adenine biosynthesis gene *ADE2*, as deletion of this gene yields an easily quantifiable phenotype—red colony pigmentation. We confirmed that integration of *C. albicans*-optimized Cas9 into the *C. albicans* genome at the *NEUT5L* locus, in contrast to the commonly used *ACT1* locus, has no quantifiable effect on fungal fitness (Supplementary Fig. 1). We then created 20 unique *ADE2* gene drive variants, in which we modified factors including Cas9 promoters, gRNA promoters, structural variants of Cas9 and gRNA target cut site(s) (Supplementary Text File 1). These *ADE2* gene drive variants were transformed into *C. albicans* haploids, and *ade2*Δ cells were chosen on the basis of red colony pigmentation. These red colonies were mated to wild-type haploids, and integration of the gene drive into the wild-type homologous chromosome was quantified on the basis of the number of *ade2*Δ/*ade2*Δ red diploid colonies.

We found that, unlike in our previous work in *Saccharomyces cerevisiae* where a single gRNA was sufficient to initiate efficient allele replacement[10], it was difficult to find a single gRNA that was able to induce efficient driving following mating in *C. albicans* (Fig. 2a). The choice of where the guide cuts within the target gene dictates the amount of exonucleolytic processing required before homology to the donor template is found. Therefore, we hypothesized that targeting Cas9 close to the regions of homology between the gRNA donor arms and the wild-type locus may limit the need for extensive DNA end-processing, and aid the cell in using the gRNA as a donor template. When we used two gRNAs targeting the 5′ and 3′ ends of the *ADE2* gene near the regions of donor homology, we found a robust increase in drive efficiency (~86–98%; Fig. 2a,b). This effect was not due to having more gRNA in the system, because vectors containing a tandem repeat of the same gRNA or two different guides targeting areas distal to the homology arms failed to increase in drive efficiency (Fig. 2a). Whole-genome sequencing of *ade2*Δ/*ade2*Δ strains showed deletion of the *ADE2* gene, without additional genomic deletions, confirming the specificity of our GDA system (Supplementary Fig. 2 and Supplementary Table 1).

Before generating double-mutant libraries, we used fluorescence-activated cell sorting to quantify the ploidy and stability of haploid lineages over passaging and through the transformation process (Supplementary Fig. 3). We selected the most stable haploid strains to use for genome editing. We further optimized media and environmental conditions for both mating and white-opaque switching (a precursor to *C. albicans* mating[10,12]), and confirmed that our mating-derived diploid strain retained virulence in a mouse model of infection (Supplementary Fig. 3). We next assessed the ability of our system to efficiently generate diploid homozygous double-gene deletions. We created a haploid strain containing a *LEU2*-disabling drive allele and mated it to a haploid cell of the opposite mating type containing our previous *ADE2* drive. Of the resulting diploid cells, ~88% were validated by PCR to be homozygous double deletions (*ade2*Δ/*ade2*Δ *leu2*Δ/*leu2*Δ), and showed the expected auxotrophic phenotype (Fig. 2c).

Efficient creation of two large pairwise gene deletion matrices. Given our ability to efficiently generate homozygous double-gene deletion mutants by mating, we scaled up our approach to create large pairwise gene deletion matrices, targeting genes involved in fungal virulence. We selected two categories of genes to target for deletion: efflux pumps that play roles in antifungal resistance[13], and adhesin genes involved in cellular adherence and biofilm formation[14]. We created drives targeting 10 efflux pump genes and 12 adhesin genes (Supplementary Table 2). The resulting constructs were transformed into *MAT**a*** and *MATα C. albicans* haploids, and each *MAT**a*** haploid was mated against each *MATα* haploid within the same drive set to generate diploid double-deletion matrices, comprising 100 and 144 diploid mutants, respectively (Supplementary Table 2). Given the homologous nature of the genes within these gene families, we performed whole-genome sequencing on several deletion strains to ensure the specificity of our gene

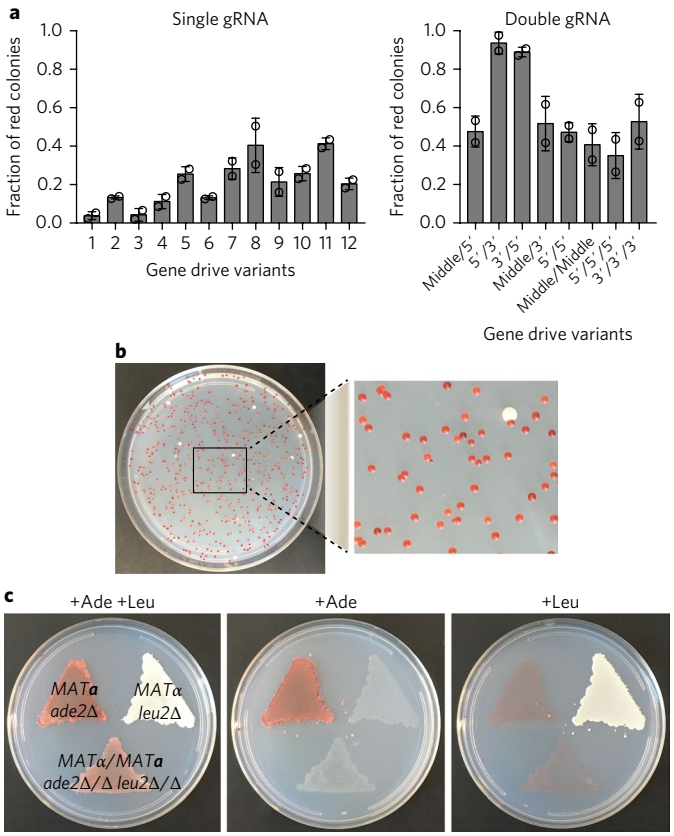

Fig. 2 | Optimized *C. albicans* gene drive efficiently creates homozygous double-gene deletion strains. a, Twenty unique gene drive constructs, targeting *ADE2* for deletion, were tested for their driving efficiency by counting the number of red (ade−) and white (ade+) diploid colonies obtained on mating a haploid *ADE2* drive strain with a wild-type haploid strain. Each of the different gene drive variants is described in Supplementary Text File 1, and contained either a single or double gRNA. The highest rate of gene driving to generate ade− diploids was obtained by constructs in which two gRNAs targeted the 5′ and 3′ end of the *ADE2* gene, respectively. The data represent mean fraction of red colonies from duplicate transformations; error bars are standard deviation. **b**, An example of a mating plate using construct 5′/3′ from **a**. A haploid strain containing construct 14 was mated with a wild-type haploid lineage, and the resultant diploid selection plate yielded ~97% red ade− colonies. The image is representative of four repetitions of the same mating. **c**, Similar to our *ADE2* design, a *LEU2*-targeting gene drive was developed. Haploid lineages containing either the *LEU2* gene drive or the *ADE2* gene drive were created to generate leu− and ade− auxotrophic mutants, respectively. These haploid mutants were mated, and resultant diploids were leu− ade− auxotrophs. Strains were grown on minimal medium, supplemented with adenine (+Ade) or leucine (+Leu) as indicated. The images are representative of three experimental replicates.

drive constructs amongst closely related genes (for example, *CDR1* and *CDR2*) (Supplementary Fig. 2 and Supplementary Table 3). Our sequencing showed that only the intended genes, and no additional genes, were deleted (Supplementary Fig. 2 and Supplementary Table 3), confirming the specificity of our GDA system.

We assessed the fitness of the mutant library strains by measuring growth using the optical density of each strain and normalizing it to that of the wild-type strain (Fig. 3a). To identify epistatic relationships between the genes, we assessed genetic interactions using a multiplicative model[15,16], which predicts double mutant fitness to be the product of the corresponding parental single-gene mutant fitness measurements[15–17] (Supplementary Table 4). Using this model, double mutants with greater than predicted fitness (positive epsilon score) indicate a positive genetic interaction, which may indicate genetic suppression that occurs when one mutation rescues the fitness defect of another, or a coequal interaction, which occurs when the fitness cost of the single-gene mutants is indistinguishable from that of the corresponding double mutant[15]. Those with fitness levels less than predicted (negative epsilon score) indicate a negative genetic interaction, indicative of genes that operate in parallel or redundant pathways with a common downstream function. We calculated significant genetic interaction scores amongst transporters, and amongst adhesins (Fig. 3b and Supplementary Table 4), and found 24 positive interactions (~53%) and 1 negative interaction (~2%) amongst transporters, and 49 positive interactions (~74%) and no negative interactions amongst adhesin genes.

We evaluated concordance between reciprocal pairs in our deletion libraries (that is, *cdr1*Δ *MAT**a*** x *cdr2*Δ *MATα* versus *cdr2*Δ *MAT**a*** x *cdr1*Δ *MATα*) by comparing genetic interaction scores (epsilon, ε). Among the efflux genes, 3 out of 44 pairs had epsilon values with different signs between pairs (for example, *cdr3*Δ*tpo3*Δ); however, no significant difference was found between epsilon values for any pair. Among the adhesin genes, all epsilon values between pairs had the same sign and no significant difference was found between the interaction scores for each pair. This analysis indicates that resultant diploid strains are highly reproducible and indicative of the underlying biology rather than off-target mutations.

In constructing GDA libraries, we identified an unexpected synthetic lethal interaction between two efflux pump genes: *YOR1* and *TPO3*. Mating together haploids containing gene drives targeting these genes individually yielded no diploid colonies, despite repeated attempts. We further validated the unviable nature of this double deletion using classic genetic techniques, where a wild-type or *yor1*Δ single-gene mutant strain was transformed with a knockout construct targeting *TPO3* for deletion (or the reciprocal transformation of *YOR1* knockout into a *tpo3*Δ strain). The wild-type transformation yielded significantly more colonies than the *yor1*Δ transformation (Supplementary Fig. 4); 100% of *yor1*Δ transformants were found to still contain a wild-type allele of *TPO3*. Therefore, we were not able to generate this double mutant using our GDA system, or classic genetic techniques. This result highlights a powerful application for the GDA platform—the ability to identify synthetic lethal interactions based on mutants that cannot form viable diploids; identification of synthetic lethal interactions can, in turn, assist in the functional characterization of genes.

High-throughput analysis of efflux pump genes and susceptibility to antifungal perturbations in *C. albicans*. Next, we investigated the role of efflux pump mutants in mediating susceptibility to antifungal stress. We grew the efflux pump deletion library in the presence of one of the most commonly administered antifungal drugs, fluconazole. Growth under fluconazole stress was assessed by optical density and normalized to wild-type growth (Fig. 4a). From this fitness measure, we identified fluconazole-impaired growth in almost all of our double-gene mutants. In particular, we found that *CDR11*, a poorly characterized ABC transporter, has the most

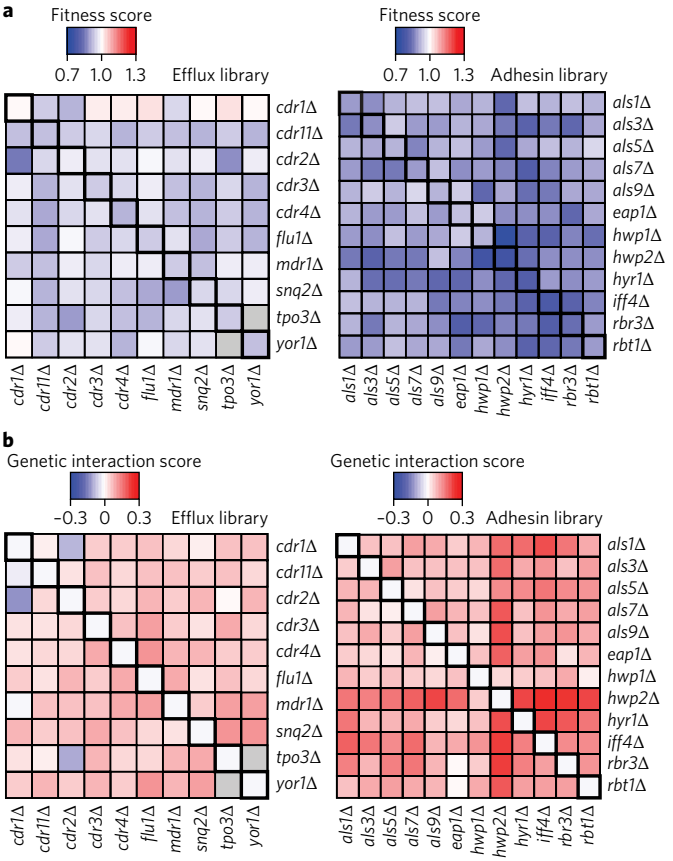

Fig. 3 | Fitness and genetic interaction analysis of two double-gene deletion virulence libraries. a, Heatmaps depicting average fitness of single- and double-gene deletions for the efflux and adhesin deletion libraries. Fitness is defined as optical density (OD) at 600 nm of the mutant normalized to the OD of the wild-type parental strain. Single-gene knockout fitness values are represented along the diagonal. Blue represents fitness values less than one, meaning the mutant grew worse than the wild type; red presents fitness values greater than one, meaning the knockout grew better than the wild type. The visible discrepancy among the *cdr1* mutants is a result of these mutants having fitness values close to one, the point at which the colour bar changes from blue to red. **b**, Heatmaps depicting average genetic interaction scores, ε, of double-gene deletions for the efflux and adhesin deletion libraries. ε is defined as the difference between the measured fitness and the expected fitness (based on a multiplicative model (see Methods)). Blue represents a negative interaction score and red represents a positive interaction score; values along the diagonal are zero, indicating no interaction for single-gene knockouts.

widespread impact on fluconazole sensitivity. We used the multiplicative model to calculate a genetic interaction score amongst transporter genes in response to fluconazole stress (Fig. 4b and Supplementary Table 4), and mapped statistically significant positive and negative interactions on a genetic interaction map (Fig. 4c). The genetic interaction map highlighted two negative genetic interactions, between *CDR1* and *CDR2* and between *TPO3* and *CDR2*. *CDR1* and *CDR2* are clinically important for resistance to antifungal agents, as they are frequently found to be overexpressed in resistant strains isolated from patients[18]. Little has been described about *TPO3*, which we find to play an important role in susceptibility to fluconazole, given its profound growth defect in the absence of *CDR2* (Fig. 4a,b).

To better understand how efflux pumps mediate susceptibility to a diversity of antifungal perturbations, we grew the *C. albicans* efflux

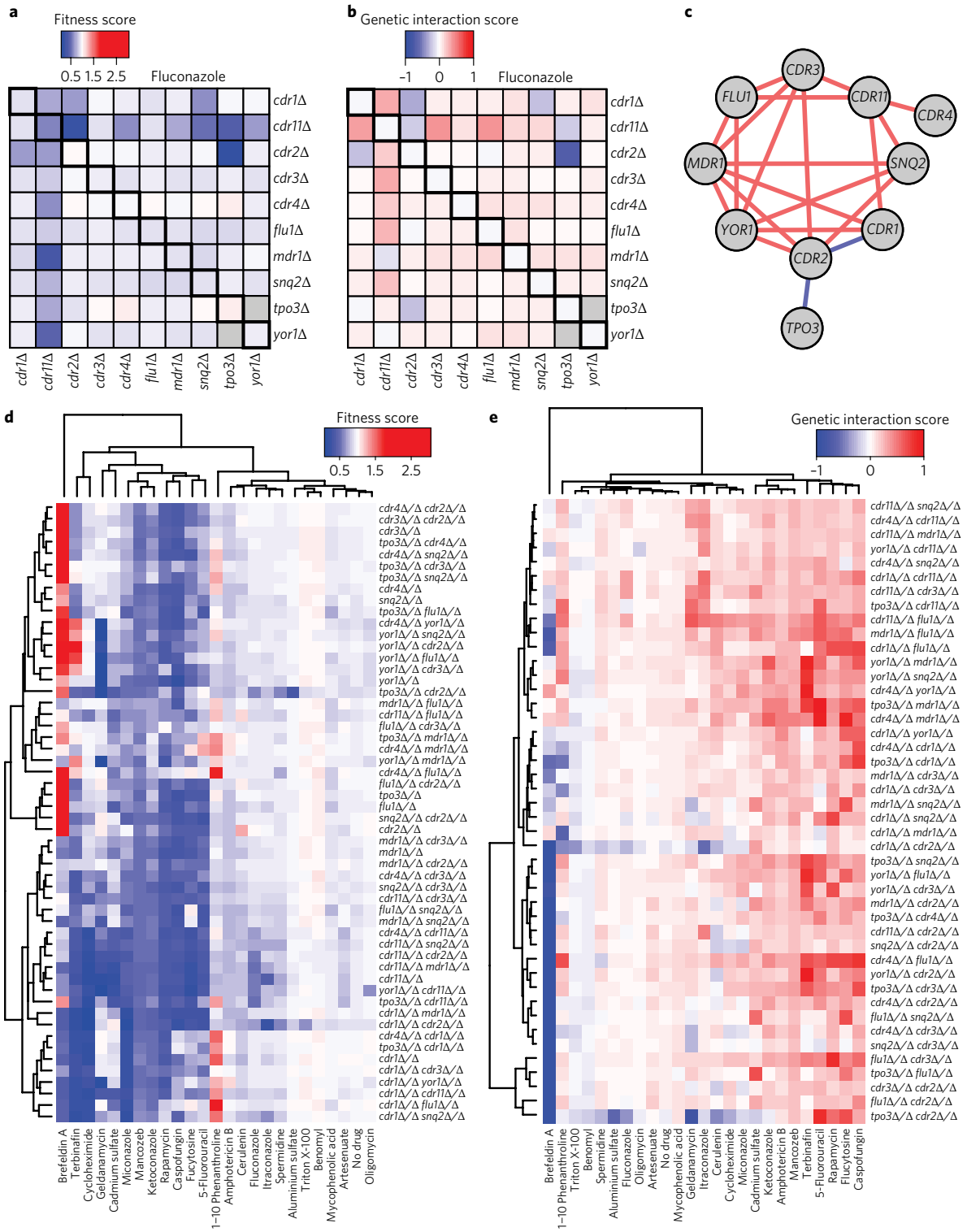

Fig. 4 | Double-deletion matrix of *C. albicans* transporter genes reveals condition-specific sensitivity and genetic interactions. a, Heatmap depicting average fitness of transport mutants (single- and double-gene deletions) under fluconazole perturbation. Fitness is defined as OD at 600 nm of the mutant normalized to the OD of the wild-type parental strain. Single knockout fitness values are represented along the diagonal. Blue represents fitness values less than one, meaning the mutant grew worse than the wild type; red presents fitness values greater than one, meaning the knockout grew better than the wild type. **b**, Heatmap depicting average genetic interaction scores, ε, of double transport gene deletions under fluconazole perturbation. ε is defined as the difference between the measured fitness and the fitness based on a multiplicative model (see Methods). Blue represents a negative interaction score and red represents a positive interaction score; values along the diagonal are zero, indicating no interaction for single-gene knockouts. **c**, Genetic interaction map in which red indicates a significantly positive interaction and blue represents a significantly negative interaction. Significant positive interactions are defined as an adjusted P value < 0.05 and $\varepsilon > 0$; significant negative interactions are defined as an adjusted P value < 0.05 and $\varepsilon < 0$. The maps were generated using Cytoscape. **d**,**e**, Hierarchical clustering of average fitness scores (**d**) and genetic interaction scores (**e**) of single- and double-gene knockouts under a diversity of antifungal perturbation conditions.

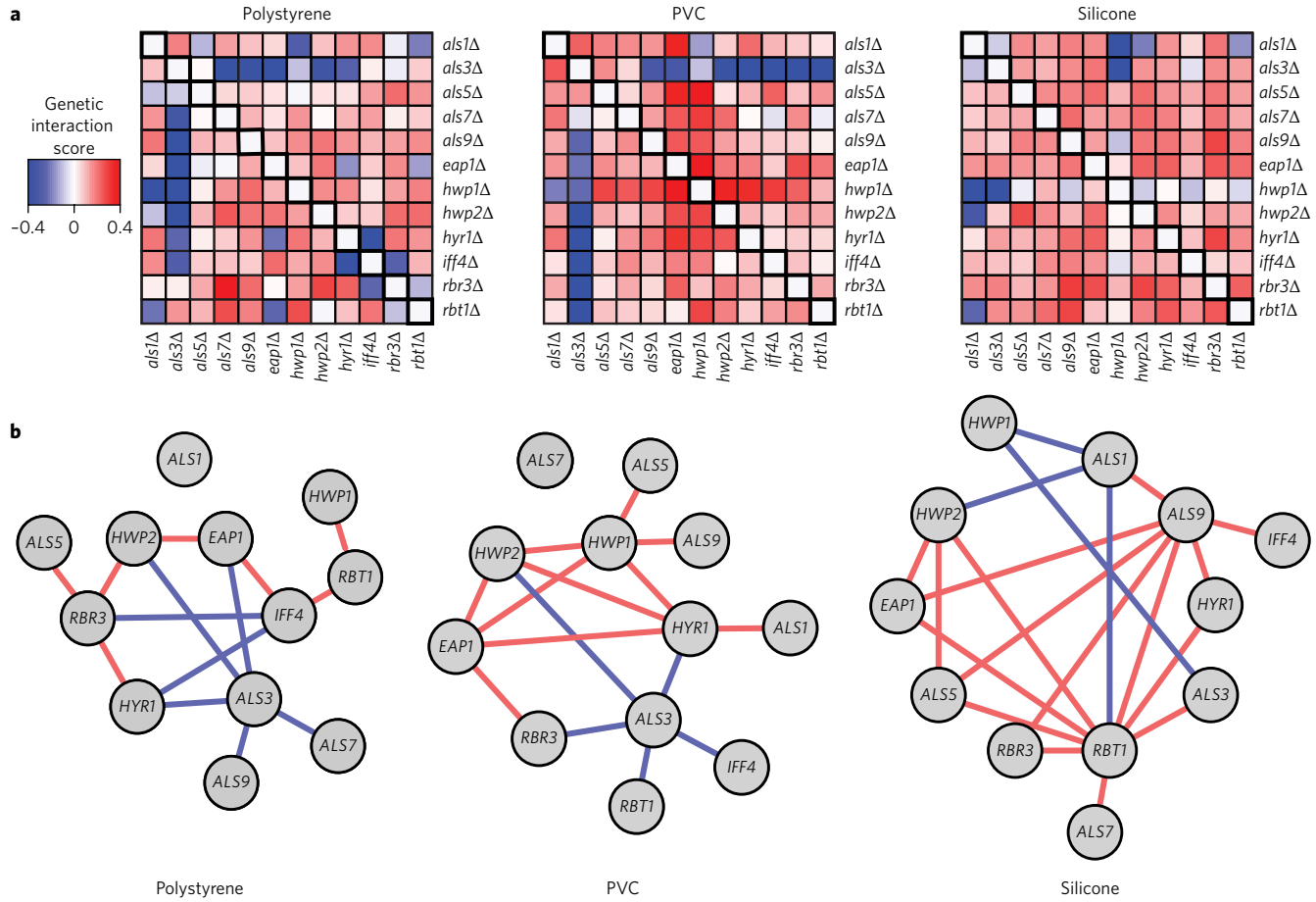

Fig. 5 | *C. albicans* double adhesin deletion matrix highlights critical factors and genetic interactions for biofilm growth. a, Heatmaps depicting average genetic interaction scores (ε) of single and double adhesin gene deletions grown on a polystyrene, PVC or silicone biofilm. ε is defined as the difference between the measured fitness and the expected fitness based on a multiplicative model, where fitness is defined as OD at 490 nm of the mutant and normalized to the OD of the wild-type parental strain (see Methods). Blue represents a negative interaction score and red represents a positive interaction score; values along the diagonal are zero, indicating no interaction for single-gene knockouts. **b**, Genetic interaction map in which red indicates a significantly positive interaction and blue represents a significantly negative interaction. Significant positive interactions are defined as an adjusted P value < 0.05 and $\varepsilon > 0$; significant negative interactions are defined as an adjusted P value < 0.05 and $\varepsilon < 0$. The maps were generated using Cytoscape.

pump deletion matrix under 24 environmental conditions, including stressors targeting the cell membrane, cell wall, DNA synthesis, protein synthesis and other perturbations (Supplementary Table 5). We calculated cellular fitness and performed hierarchical clustering to evaluate growth trends across perturbations and knockout strains (Fig. 4d). We observed that the perturbations split into two groups: those with fitness values less than wild type, suggesting that *C. albicans* is more susceptible to the perturbation without the efflux pumps, and those that are similar to wild type, suggesting either that the fungus is less reliant on the pumps at the tested concentrations or that there is greater redundancy under certain conditions. The latter is evidenced with growth in aluminum sulfate, where almost all mutants exhibit wild-type fitness, with the exception of the *tpo3*Δ/Δ *cdr2*Δ/Δ double-deletion mutant, which is severely impaired in growth (Fig. 4d). Brefeldin A is a notable exception—in this condition, many mutants tolerate stress better than wild type (Fig. 4d), suggesting an upregulation of compensatory mechanisms to mediate stress tolerance. Looking at clusters across conditions can reveal similarities in mutant sensitivities. For example, the imidazoles (ketoconazole and miconazole) cluster closely together, and the triazoles (fluconazole and itraconazole) cluster together, but the two classes cluster away from each other. This highlights unique efflux pump dependencies for these two classes of azole antifungals and suggests that the triazoles may be less reliant on pumps or that there is more genetic redundancy regulating their efflux (Fig. 4d).

We calculated interaction scores and find that genetic interactions vary across stress conditions, although they are mostly positive (Fig. 4e and Supplementary Table 4). This is not unexpected, given that efflux pump genes probably share common cellular functions (Fig. 4e). The negative genetic interactions identified between *CDR1* and *CDR2* and between *TPO3* and *CDR2* following fluconazole treatment are conserved amongst many other conditions tested (Fig. 4e), suggesting that these factors play redundant roles in pumping out toxic xenobiotics. Of the genetic interaction score, 90.9% had had the same sign as their reciprocal pair and for 95.5% of pairs no significant difference was found between the interaction scores, confirming the reproducibility of our mutants. These results allow us to assess genetic interactions between efflux pumps under stress conditions, and identify genetic combinations that sensitize *C. albicans* to antifungal perturbations.

High-throughput analysis of adhesin genes and biofilm growth in *C. albicans*. Finally, we addressed the function of adhesin gene deletion mutants in biofilm formation. The *C. albicans* adhesin deletion matrix was grown under biofilm-forming conditions; three distinct materials were used to mimic medical device surfaces on

which biofilms form[19,20]: polystyrene, polyvinyl chloride (PVC) and silicone. We quantified biofilm growth using an XTT metabolic reduction assay, assessed the ability of single- and double-adhesin mutants to form robust biofilms, and assessed epistatic relationships between adhesin genes under each growth condition.

We found both conserved and divergent genetic interactions between adhesin genes across conditions, some of which are robustly conserved. We observed a strong signature for positive genetic interactions between *HWP2* and *EAP1* across the three biofilm growth conditions, and conserved negative interactions between *HWP1* and both *ALS1* and *ALS3* (Fig. 5a). Other interactions (such as *ALS3* and *RBT1*) were found to fluctuate, highlighting how genetic interactions may be unique under distinct growth conditions (Fig. 5b). Again, we found good reproducibility between reciprocal knockout pairs, with 93.9% of significant interaction scores having the same sign between pairs and 100% of pairs having no significant difference in interaction scores. Through this network analysis, we readily identified factors (for example, *ALS3*) that serve as a highly interconnected hub of negative genetic interactions under all biofilm growth conditions (Fig. 5a,b). This suggests that targeting *ALS3* may be a powerful strategy to severely weaken the ability of *C. albicans* to form biofilms.

Discussion

Here, we presented the development of GDA: a platform for the generation of combinatorial deletion mutants in *C. albicans*. Our CRISPR–Cas9-based gene drive platform has been adapted from a gene drive system in *S. cerevisiae*[10], but has been significantly optimized for *C. albicans*, enabling systematic genetic interaction analysis in this fungal pathogen, which has unique virulence traits and many divergent cellular signalling pathways compared with *S. cerevisiae*[21–23]. We used our technology to determine how the multitude of drug pumps and cellular adhesins in *C. albicans* interact to govern tolerance to xenobiotics and biofilm formation, respectively.

Despite advances in genome editing technology, there remain challenges to manipulating *C. albicans* genetically, and creating higher-order genetic mutants. The GDA system marks an advance in our ability to perform efficient and higher-order genetic analysis in this pathogen; and given the nature of our system, it guarantees that all wild-type loci of a targeted gene are fully deleted, even in spite of potential genome rearrangements. At the genome scale, a systematic double-deletion library using current state-of-the-art Cas9-based methods or older marker recycling approaches would require over 18 million transformations and an intractable amount of effort and resources (see Supplementary Note 1). In contrast, the GDA platform requires only a single panel of *MAT**a*** and *MATα* haploid cells—each with a drive targeting a single gene (~12,000 transformations). These mutant gene panels could then be mated en masse to create all double-mutant combinations.

Our GDA screening platform allowed us to readily identify a previously uncharacterized, synthetic lethal interaction between the transporter genes *TPO3* and *YOR1*. Neither transporter is well annotated in *C. albicans*, but *TPO3* is predicted to have a role in polyamine transport, and *YOR1* a role in organic anion transport, based on homology to *S. cerevisiae* genes[24,25]. As polyamines are critical organic cations, we hypothesize that deletion of the two factors simultaneously may disrupt cellular ion homeostasis, in a manner that is lethal to the cell. This is reinforced by the finding that the *S. cerevisiae* orthologues *YOR1* and *TPO2* show similar genetic interaction profiles[26]. The fact that such an interaction has not been detected in *S. cerevisiae* may be due to an increased redundancy of polyamine paralogues[27].

Dissecting the combined roles of efflux pumps is of particular interest for the newly emerging fungal pathogen *Candida auris*. *C. auris* is multidrug resistant, with some strains exhibiting resistance to all classes of antifungal drugs[28], and genome sequencing has revealed that a significant portion of the *C. auris* genome encodes for efflux pumps[29]. Since a putative mating type locus has been identified in the genome[29], it is possible that a mating-based GDA platform—similar to the one presented here—could be developed in *C. auris*. High-throughput functional genetic analysis could prove invaluable for studying the role of these numerous transporter genes in this highly resistant and deadly fungal pathogen.

By focusing on adhesin genes, we identified genetic combinations that render *C. albicans* unable to form robust biofilms, thereby impairing fungal pathogenesis. For example, we identified the adhesin *ALS3* as a critical regulator of biofilm formation. While the role of *ALS3* in biofilm formation has been established[30], we uncover redundancies between *ALS3* and other adhesins, and find that deletion of *ALS3*, in combination with several other adhesin genes, severely impaired the ability of *C. albicans* to form biofilms on all of the substrates tested. Als3 has been proposed as a candidate target for invasive candidiasis infections[31], and our work lends support for its candidacy and further highlights how simultaneously targeting additional adhesin proteins, such as Hwp1 or Hwp2, might significantly enhance the efficacy of such a therapeutic.

Beyond *C. albicans*, the principles uncovered within this work should enable the adaptation of our gene drive platform to other clinically relevant, but often neglected, fungal pathogens such as *C. auris*. Employing gene drives in higher eukaryotes has been considerably successful, suggesting that the principles outlined here could further enable gene drives to be used as a plastic genetic tool to understand complex genetic interactions in higher-order organisms.

Methods

Microbial strains. *E. coli* NEB 10-beta was used for cloning and was cultured in LB medium at 37 °C. *C. albicans* strains were cultured in YPD medium at 30 °C or 37 °C. The parental haploid strains used in this work are the URA− *MATα* strain fRS29 (also known as GZY803)[11] (*MATα his4Δ ura3::HIS4*) and the HIS− *MAT**a*** strain fRS32 (also known as GZY892) (*MAT**a** ura3::imm434 gal1::URA3 his4Δ*; created from parental strain YJB12881 (Haploid XI) from ref. [11]).

Plasmid construction. To construct the base *NEUT5L* locus integrating Cas9 drive-containing plasmids, a combination of Gibson assembly, genomic PCR and IDT gene blocks was employed. The full sequence of the integrating plasmids is shown in Supplementary Text File 1, and we have made the plasmids available via Addgene (reference numbers 89576, 89577 and 89578).

Sequences of the generated gene drives are shown in Supplementary Text File 1. To design the most effective dual-gRNA-containing drives, gene blocks were ordered from IDT for Gibson assembly into our Cas9 integrating vector, which had first been digested with NgoMIV. Drives were designed by taking 225 base pairs upstream and 200 base pairs downstream of the target region to be deleted and using those sequences as homology arms to aid the drive being inserted into the proper locus. The gene drives were then designed to contain two gRNAs that are transcribed from a single *SNR52* promoter as a tandem fusion, without any linker sequences between the gRNAs. Of note, to aid in the in vitro synthesis and amplification of the drive cassette, two different gRNA tail sequences were used in all dual-guide designs. The spacers for targeting the genome were selected by using the Broad gRNA scorer and picking the best guide that was located within 125 base pairs of the homology arms that were used. If all guides were low scoring (score < 0.5), new homology arms were chosen to enable a different region to be selected for guide targeting. Once designs were completed, sequences containing the 5' homology arm–SNR52 promoter–gRNA1–gRNA2-3' homology arm were ordered as gene blocks. An example of a gene block that was ordered is given in Supplementary Text File 1.

Haploid strain construction. Gene drive plasmids were linearized using PacI. The gene drive region was additionally amplified from the plasmid, using primers at either end of the regions of homology. Linearized plasmid and purified PCR products were transformed into haploid *C. albicans* cells, and selected on YPD plates containing 250 μg ml^{-1} of nourseothricin (NAT). NAT-resistant colonies were genotyped by PCR to confirm proper integration of the drive plasmid and absence of any wild-type alleles of the target gene.

Ten efflux pump haploid deletion mutants were constructed in both the *MATα* and *MAT**a*** background (*cdr1Δ, cdr2Δ, cdr3Δ, cdr4Δ, cdr11Δ, flu1Δ, mdr1Δ, snq2Δ, tpo3Δ* and *yor1Δ*). Twelve adhesin haploid deletion mutants were constructed in both the *MATα* and *MAT**a*** background (*als1Δ, als3Δ, als5Δ, als7Δ, als9Δ, eap1Δ,*

*hwp1*Δ, *hwp2*Δ, *hyr1*Δ, *iff4*Δ, *rbr3*Δ and *rbt1*Δ). In addition, wild-type haploid control strains were generated in both the *MATα* and *MAT**a*** backgrounds by integrating a control NAT-resistance plasmid containing the Cas9 gene but lacking a gene drive.

For synthetic lethal analysis, *tpo3*Δ and *yor1*Δ strains were constructed using a NAT-flipper-based knockout strategy[32,33], where the recyclable NAT cassette (pJK863[33]) was amplified using primers containing homology upstream and downstream of the gene targeted for deletion, as previously described[34,35]. NAT-resistant transformants were selected and verified for deletion, and the NAT cassette was removed after growth in YNB–BSA media. For determination of synthetic lethal status, these *tpo3*Δ and *yor1*Δ strains, as well as a wild-type haploid strain, were again transformed with the opposite deletion construct, amplified from the NAT cassette.

Mating and diploid strain construction. Haploid *C. albicans* strains were struck onto YPD pH 6.0 plates, containing 5 μg ml^{-1} phloxine B, and incubated at 27 °C with 5% CO_2 for seven days to induce white-opaque switching. Pink-coloured opaque colonies were then selected for mating. ***a*** and *α* mating-type colonies (his− and ura−) were mixed together on YPD plates and incubated at room temperature for 5–7 days. The patch of mating cells was resuspended in 250 μl of PBS, and diploid colonies were selected by plating onto SD plates to select for his+ ura+ diploid colonies. Diploid colonies were re-patched onto SD plates, and genotyped by PCR to confirm proper deletion of both target genes and the absence of any wild-type genes.

Deletion library construction. Two libraries were constructed, containing single- or double-deletion mutants in drug efflux pumps and adhesin genes, respectively. Each library was constructed by pairwise mating of the 10 efflux, or 12 adhesin haploid *MATα* x *MAT**a*** mutants, yielding a library of 100 diploid efflux deletion strains and 144 diploid adhesin deletion strains, respectively. Each library contains one diploid strain representing each of the single-gene deletions, and two independent diploid strains representing each of the possible double-gene deletions. For the efflux pump library, this includes 10 diploid homozygous single-gene deletion mutants, and two sets of 45 diploid homozygous double-gene deletion mutants. For the adhesin library, this includes 12 diploid homozygous single-gene deletion mutants, and two sets of 66 diploid homozygous double-gene deletion mutants. Each library also contains wild-type diploid strains, generated by mating *MATα* and *MAT**a***, each containing a control Cas9–NAT plasmid, as well as wild-type diploid strains lacking the Cas9–NAT plasmid to ensure no effects of Cas9–NAT on any growth conditions. Both the drug efflux and adhesin libraries were created in duplicate versions.

Ploidy analysis. Ploidy of *C. albicans* lineages was assessed by FACS analysis, as previously described[11]. Briefly, strains were grown overnight in deep-well, 96-well plates in 250 μl of SDC media at 30 °C, and then subcultured 1:100 and grown for 4–6 hours to obtain mid-log-phase cells. Cells were pelleted and resuspended in TE buffer (50 mM Tris pH 8.0, 50 mM EDTA) in a round-bottom, 96-well plate, and fixed with 95% ethanol, overnight at −20 °C. Cells were then washed with TE buffer and treated with RNAse A (1 mg ml^{-1} in TE buffer), overnight at 37 °C. Cells were then treated with proteinase K (5 mg ml^{-1} in TE buffer) for 30 minutes at 37 °C. Cells were then washed with TE and stained with SYBR green solution (1:100 in TE buffer) and incubated overnight, at room temperature and protected from light. The SYBR-stained cells were then pelleted and resuspended in fresh TE buffer, transferred to a 96-well PCR tube plate, and sonicated in a water bath sonicator for two minutes. Twenty microlitres of sonicated cells were diluted in 180 μl of TE buffer for FACS analysis. FACS analysis was performed on the BD LSRFortessa (BD Biosciences) and analysed using FlowJo software. Ten thousand cells were counted for each strain tested. Ploidy values were calculated by comparing the peak locations in experimental samples with those of known haploid and diploid controls.

Archiving. Haploid and diploid deletion strains were maintained in 20% glycerol stocks at −80 °C in 96-well plates, and spotted onto YPD agar plates for downstream use.

Antifungal drug testing. The diploid drug efflux pump library was screened for resistance or susceptibility to diverse antifungal agents and other cellular perturbations. Antifungal drugs and cellular perturbation conditions were selected on the basis of susceptibility profiles of drug efflux pump mutants from the *Candida* Genome Database (CGD) (Supplementary Table 5), and susceptibility profiles of drug efflux pump mutant orthologues from the *Saccharomyces* Genome Database (SGD). Overnight cultures of the library were grown in round-bottom, 96-well plates in YPD media at 30 °C. Cells were then subcultured at a 1:100 dilution into fresh YPD media ± drugs in flat-bottom, 96-well plates, and incubated for 24–48 hours at 37 °C, shaking at 900 r.p.m. Drug concentrations were chosen on the basis of the annotations and previous literature from CGD and SGD, as listed in Supplementary Table 5. Growth was measured on a spectrophotometer, measuring optical density $OD_{600\,nm}$. For each assay, two versions of the diploid efflux pump library, including wild-type control strains, were screened in the presence and absence of drug. Each drug condition was tested in duplicate assays, yielding a total of four replicates for each knockout.

Biofilm growth. Biofilms were grown as previously described[36,37], with some modifications. The adhesin diploid deletion library was grown overnight in 96-well plates, in YPD media at 30 °C. Subsequently, each strain was subcultured into RPMI media (supplemented to 2% glucose), in 96-well plate format. Plates used for biofilm growth were made of polystyrene (Corning), polyvinyl chloride (PVC; Corning) or silicone (E&K Scientific). Biofilms were allowed to form for 24 hours, at 37 °C. The biofilms were then washed twice with PBS to remove non-adherent cells, and resuspended in fresh RPMI media and incubated at 37 °C for another 24 hours. Biofilms were again washed twice with PBS to remove non-adherent cells, and metabolic activity was measured using XTT. Ninety microlitres of 1 mg ml^{-1} XTT and 10 μl of 320 μg ml^{-1} phenazine methosulfate were added to each well and allowed to incubate for 2 hours at 37 °C. The supernatant was transferred to a clean 96-well plate, and absorbance was measured at 490 nm. For each assay, two versions of the diploid efflux pump library, including wild-type control strains, were screened in each of the different 96-well plates (polystyrene, PVC or silicone). Each assay was performed in duplicate, yielding a total of four replicates per strain.

Analysis of double-deletion mutant data. The fitness of each double-deletion mutant replicate, W, was calculated by dividing the $OD_{600\,nm}$ growth measurements of each mutant under a given condition by the averaged $OD_{600\,nm}$ measurements of the corresponding wild-type control strains under that same condition. The genetic interaction score was determined on the basis of the deviation of the measured fitness from the expected fitness of the double-deletion strains[38]. This deviation, ε, was defined as: $\varepsilon_{xy} = W_{xy} - W_x \times W_y$, where W_x and W_y are the fitness values for strains deleted for gene x and y, respectively, and W_{xy} is the measured fitness for the double-deletion strain. Genetic interaction maps were generated in Cytoscape (Cytoscape 3.4.0[39]), with a positive interaction defined as having an adjusted P value < 0.05 and $\varepsilon > 0$, and a negative interaction defined as having an adjusted P value < 0.05 and $\varepsilon < 0$. P values were calculated between the eight measured fitness values (forward (W_{xy}) and reverse (W_{yx}) knockout strains from duplicate assays) and four expected fitness values ($W_x \times W_y$ from duplicate assays) using Student's t-test (two-sided) in R and adjusted for multiple hypotheses using the Benjamini and Hochberg method. Clustered heatmaps were generated in R using heatmap.2[40] with the dist and hclust functions, using an average of either the eight replicate fitness values (W_{xy}) or genetic interaction scores (ε_{xy}). Un-clustered, condition-specific heatmaps were also generated using heatmap.2; in these heatmaps, averages from reciprocal pairs are split between the lower and upper triangles, respectively, and the diagonal values represent the single knockouts. All data analyses were performed using RStudio[41] version 1.0.136 and R version 3.2.3[42].

Sequencing library preparation. Sequencing library preparation was performed as described previously[43,44]. Bacterial strains were grown to saturation in LB media, and total genomic DNA was extracted from 1 ml of cultures using the PureLink Pro-96 Genomic Purification Kit (Life Technologies, catalogue number K1821-04A) and quantified using the Quant-iT High Sensitivity DNA Assay Kit (Life Technologies, catalogue number Q-33120), according to the manufacturer's instructions. Dual-barcoded genomic DNA libraries were prepared as previously described[43]. Briefly, 0.5 ng of genomic DNA was fragmented and tagged with sequencing adaptors, using the Nextera DNA Library Preparation Kit (Illumina, catalogue number FC-121-1030). The library was then amplified using KAPA Library Amplification Kits (KAPA Biosystems, catalogue number KK2611) and Illumina TruSeq primers. PCR clean-up and DNA size selection were performed using a magnetic bead solution made as previously described[43,45]. For a subset of samples, DNA fragment size distributions were analysed using High Sensitivity DNA Analysis Kits (Agilent Technologies, catalogue number 5067-4626) on the BioAnalyzer 2100 (Agilent Technologies). The average DNA fragment size was ~350 base pairs. Libraries were quantified using the Quant-iT High Sensitivity DNA Assay Kit and pooled so as to achieve approximately equal molar concentrations of each sample in the pool. Samples were sequenced at the Broad Institute Genomics Platform (Cambridge, MA) using the Illumina NextSeq platform, yielding 169 paired-reads total.

Sequencing analysis. The FASTQC tool (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to evaluate the quality of the fastq files before and after treatment. The FASTX Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) was used to remove the adaptamers (fastx_trimmer) and trim the ends for base pairs with a quality score lower than 20 (fastq_quality_trimmer). An in-house algorithm was used to intersect the read-pairs after quality trimming. The reads were aligned to the *C. albicans* WO-1 genome (downloaded from Genbank as assembly GCA_000149445.2) using BWA (version 0.6.1-r104). The depth of reads at each position on the genome for identification of deleted genes was calculated using genomeCoverageBed from BEDTools (version 2.16.2), as previously described[46]. The depth for each gene was normalized to the median depth across all genes and

genes having a depth of zero in all samples were removed. The 15 genes showing abnormal depth across samples were isolated using R[47]. The depth for the region of the gene ±1 kilobase was extracted for each of the 15 samples, normalized to the median genome depth, and evaluated manually. The sequencing data were deposited to the NCBI Short Read Archive under Bioproject PRJNA356375.

Mouse model of systemic infection assay. Female BALB/c mice of 8–10 weeks were used. *C. albicans* yeast cells were grown in YPD at 30 °C to log phase, harvested by centrifugation and washed twice with and resuspended in PBS at a concentration of 1×10^6 cells ml^{-1}. Each mouse ($n = 8$ for each *C. albicans* strain) was injected via the tail vein with 1×10^6 cells. Forty-eight hours after the injection, two mice from each group were euthanized and both kidneys were removed, homogenized and spread, after tenfold serial dilution, onto YPD plates for the enumeration of colony-forming units. The rest of the mice were monitored daily for survival over 14 days. Sample size determination was based on previous experiments, and from standard practice in the field. The assignment of animals into different infection groups was determined at random, and the study was conducted in a blinded manner, such that the experimenter was not aware of which strains were being infected into each mouse. All procedures of the animal experiment were approved by the Institutional Animal Care and Use Committee (IACUC no. 151010).

Code availability. All custom code from this work is available from https://github.com/cbmporter/CandidaKOanalysis

Data availability. The data sets generated and analysed during the current study are available either in this published article, or from the following sources: the full sequence of the plasmids is shown in Supplementary Text File 1, and we have made the plasmids available via Addgene (reference numbers 89576, 89577 and 89578); the whole-genome sequencing data were deposited to the NCBI Short Read Archive under Bioproject PRJNA356375.

Received: 14 June 2017; Accepted: 15 September 2017;
Published online: 23 October 2017

References

1. Shapiro, R. S., Robbins, N. & Cowen, L. E. Regulatory circuitry governing fungal development, drug resistance, and disease. *Microbiol. Mol. Biol. Rev.* **75**, 213–267 (2011).
2. Nobile, C. J. & Johnson, A. D. *Candida albicans* biofilms and human disease. *Annu. Rev. Microbiol.* **69**, 71–92 (2015).
3. Ramage, G., Mowat, E., Jones, B., Williams, C. & Lopez-Ribot, J. Our current understanding of fungal biofilms. *Crit. Rev. Microbiol.* **35**, 340–355 (2009).
4. Tong, A. H. Y. et al. Global mapping of the yeast genetic interaction network. *Science* **303**, 808–813 (2004).
5. Enkler, L., Richer, D., Marchand, A. L., Ferrandon, D. & Jossinet, F. Genome engineering in the yeast pathogen *Candida glabrata* using the CRISPR-Cas9 system. *Sci. Rep.* **6**, 35766 (2016).
6. Fuller, K. K., Chen, S., Loros, J. J. & Dunlap, J. C. Development of the CRISPR/Cas9 system for targeted gene disruption in *Aspergillus fumigatus*. *Eukaryot. Cell* **14**, 1073–1080 (2015).
7. Vyas, V. K., Barrasa, M. I. & Fink, G. R. A *Candida albicans* CRISPR system permits genetic engineering of essential genes and gene families. *Sci. Adv.* **1**, e1500248 (2015).
8. Min, K., Ichikawa, Y., Woolford, C. A. & Mitchell, A. P. *Candida albicans* gene deletion with a transient CRISPR-Cas9 system. *mSphere* **1**, e00130-16 (2016).
9. Nguyen, N., Quail, M. M. F. & Hernday, A. D. An efficient, rapid, and recyclable system for CRISPR-mediated genome editing in *Candida albicans*. *mSphere* **2**, e00149-17 (2017).
10. DiCarlo, J. E., Chavez, A., Dietz, S. L., Esvelt, K. M. & Church, G. M. Safeguarding CRISPR-Cas9 gene drives in yeast. *Nat. Biotechnol.* **33**, 1250–1255 (2015).
11. Hickman, M. A. et al. The 'obligate diploid' *Candida albicans* forms mating-competent haploids. *Nature* **494**, 55–59 (2013).
12. Miller, M. G. & Johnson, A. D. White-opaque switching in *Candida albicans* is controlled by mating-type locus homeodomain proteins and allows efficient mating. *Cell* **110**, 293–302 (2002).
13. Cannon, R. D. et al. Efflux-mediated antifungal drug resistance. *Clin. Microbiol. Rev.* **22**, 291–321 (2009).
14. Sundstrom, P. Adhesins in *Candida albicans*. *Curr. Opin. Microbiol.* **2**, 353–357 (1999).
15. Baryshnikova, A., Costanzo, M., Myers, C. L., Andrews, B. & Boone, C. Genetic interaction networks: toward an understanding of heritability. *Annu. Rev. Genomics Hum. Genet.* **14**, 111–133 (2013).
16. Boone, C., Bussey, H. & Andrews, B. J. Exploring genetic interactions and networks with yeast. *Nat. Rev. Genet.* **8**, 437–449 (2007).
17. Mani, R., St Onge, R. P., Hartman, J. L. 4th, Giaever, G. & Roth, F. P. Defining genetic interaction. *Proc. Natl Acad. Sci. USA* **105**, 3461–3466 (2008).
18. Morschhäuser, J. The genetic basis of fluconazole resistance development in *Candida albicans*. *Biochim. Biophys. Acta* **1587**, 240–248 (2002).
19. Hawser, S. P. & Douglas, L. J. Biofilm formation by *Candida* species on the surface of catheter materials in vitro. *Infect. Immun.* **62**, 915–921 (1994).
20. Lawrence, E. L. & Turner, I. G. Materials for urinary catheters: a review of their history and development in the UK. *Med. Eng. Phys.* **27**, 443–453 (2005).
21. Fuller, KevinK. J. C. R. Protein kinase A and fungal virulence: a sinister side to a conserved nutrient sensing pathway. *Virulence* **3**, 109 (2012).
22. Sibley, L. D., Howlett, B. J. Heitman, J. (eds) *Evolution of Virulence in Eukaryotic Microbes* (Wiley, USA, 2012).
23. Shapiro, R. S., Ryan, O., Boone, C. & Cowen, L. E. Regulatory circuitry governing morphogenesis in *Saccharomyces cerevisiae* and *Candida albicans*. *Cell Cycle* **11**, 4294–4295 (2012).
24. Cui, Z., Hirata, D., Tsuchiya, E., Osada, H. & Miyakawa, T. The multidrug resistance-associated protein (MRP) subfamily (Yrs1/Yor1) of *Saccharomyces cerevisiae* is important for the tolerance to a broad range of organic anions. *J. Biol. Chem.* **271**, 14712–14716 (1996).
25. Tomitori, H., Kashiwagi, K., Sakata, K., Kakinuma, Y. & Igarashi, K. Identification of a gene for a polyamine transport protein in yeast. *J. Biol. Chem.* **274**, 3265–3267 (1999).
26. Costanzo, M. et al. The genetic landscape of a cell. *Science* **327**, 425–431 (2010).
27. Byrne, K. P. & Wolfe, K. H. The Yeast Gene Order Browser: combining curated homology and syntenic context reveals gene fate in polyploid species. *Genome Res.* **15**, 1456–1461 (2005).
28. Vallabhaneni, S. et al. Investigation of the first seven reported cases of *Candida auris*, a globally emerging invasive, multidrug-resistant fungus-United States, May 2013-August 2016. *Am. J. Transplant* **17**, 296–299 (2017).
29. Chatterjee, S. et al. Draft genome of a commonly misdiagnosed multidrug resistant pathogen *Candida auris*. *BMC Genomics* **16**, 686 (2015).
30. Nobile, C. J. et al. Complementary adhesin function in *C. albicans* biofilm formation. *Curr. Biol.* **18**, 1017–1024 (2008).
31. Liu, Y. & Filler, S. G. *Candida albicans* Als3, a multifunctional adhesin and invasin. *Eukaryot. Cell* **10**, 168–173 (2011).
32. Shapiro, R. S., Zaas, A. K., Betancourt-Quiroz, M., Perfect, J. R. & Cowen, L. E. The Hsp90 co-chaperone Sgt1 governs *Candida albicans* morphogenesis and drug resistance. *PLoS ONE* **7**, e44734 (2012).
33. Shen, J., Guo, W. & Köhler, J. R. CaNAT1, a heterologous dominant selectable marker for transformation of *Candida albicans* and other pathogenic *Candida* species. *Infect. Immun.* **73**, 1239–1242 (2005).
34. Shapiro, R. S. et al. Pho85, Pcl1, and Hms1 signaling governs *Candida albicans* morphogenesis induced by high temperature or Hsp90 compromise. *Curr. Biol.* **22**, 461–470 (2012).
35. Ryan, O. et al. Global gene deletion analysis exploring yeast filamentous growth. *Science* **337**, 1353–1356 (2012).
36. Ramage, G., Vande Walle, K., Wickes, B. L. & López-Ribot, J. L. Standardized method for in vitro antifungal susceptibility testing of *Candida albicans* biofilms. *Antimicrob. Agents Chemother.* **45**, 2475–2479 (2001).
37. Robbins, N. et al. Hsp90 governs dispersion and drug resistance of fungal biofilms. *PLoS Pathog.* **7**, e1002257 (2011).
38. St Onge, R. P. et al. Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. *Nat. Genet.* **39**, 199–206 (2007).
39. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
40. Warnes, G. R. et al. gplots. R Package v.3.0.1 (CRAN, 2016); https://CRAN.R-project.org/package=gplots
41. *RStudio* (accessed 6 June 2017); https://www.rstudio.com
42. *R: The R Project for Statistical Computing* (accessed 6 June 2017); https://www.R-project.org
43. Baym, M. et al. Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS ONE* **10**, e0128036 (2015).
44. Cohen, N. R. et al. A role for the bacterial GATC methylome in antibiotic stress survival. *Nat. Genet.* **48**, 581–586 (2016).
45. Rohland, N. & Reich, D. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* **22**, 939–946 (2012).
46. Kaas, C. S., Kristensen, C., Betenbaugh, M. J. & Andersen, M. R. Sequencing the CHO DXB11 genome reveals regional variations in genomic stability and haploidy. *BMC Genomics* **16**, 160 (2015).
47. Ihaka, R. & Gentleman, R. R: A language for data analysis and graphics. *J. Comput. Graph. Stat.* **5**, 299 (1996).

Acknowledgements

We thank G. Fink, J. Berman, M. Hickman, V. Vyas and A. Baryshnikova for helpful discussions. We also thank V. Vyas, J. Köhler and L. Cowen for strains. This work was supported by the Paul G. Allen Frontiers Group, a Banting postdoctoral fellowship from the Canadian Institutes of Health Research, National Cancer Institute grant no. 5T32CA009216-34, US National Institutes of Health National Human Genome Research Institute grant no. RM1 HG008525 and the Wyss Institute for Biologically Inspired Engineering.

Author contributions

R.S.S., A.C., J.E.D., G.M.C. and J.J.C. conceptualized the project; R.S.S., A.C., M.H., A.V.R. and X.X. performed the experiments; C.B.M.P., C.S.K. and R.S.S. performed analysis and visualization of experimental results; G.Z. and Y.W. generated and provided strains; R.S.S., A.C. and C.B.M.P wrote and edited the manuscript; Y.W., N.V.K., G.M.C. and J.J.C. supervised the project; J.J.C and G.M.C. acquired funding.

Competing interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41564-017-0043-0.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to J.J.C.

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

nature research

Corresponding author(s): James J Collins

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

Each deletion mutant in our libraries was created in duplicate copies. For mouse experiments, 8 mice were included in each experimental condition.

2. Data exclusions

Describe any data exclusions.

No data was excluded

3. Replication

Describe whether the experimental findings were reliably reproduced.

Experimental findings were reliably reproduced. Each deletion library was created in duplicate, and screened in duplicate under each relevant condition.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Randomization is not relevant to the library screening component of this study, as all strains were tested under all conditions. For mouse experiments The assignment of animals into different infection groups was determine at random

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

The analysis of the data from library screens was performed by an independent researcher from the one who screened the libraries. The samples were blinded using an accession code, and therefore the person doing the analysis did not know which mutant strains were associated with which data. For mouse experiments, the study was conducted in a blinded manner, such that the experimenter was not aware of which strains were being infected into each mouse.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	FlowJo Software for FACS analysis, Cytoscape v. 3.4.0 for genetic interaction network visualization, RStudio version 1.0.136 and R version 3.2.3, BEDTools version 2.16.2.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	All materials are available from the authors.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	No antibodies were used.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	Various Candida albicans fungal cell lines were used. No mammalian cell lines were used.
b. Describe the method of cell line authentication used.	*Describe the authentication procedures for each cell line used OR declare that none of the cell lines used have been authenticated OR state that no eukaryotic cell lines were used.*
c. Report whether the cell lines were tested for mycoplasma contamination.	*Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination OR state that no eukaryotic cell lines were used.*
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	*Provide a rationale for the use of commonly misidentified cell lines OR state that no commonly misidentified cell lines were used.*

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	Female BALB/c mice of 8-10 weeks were used in this study.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	No human participants were used.