nature genetics

# Systematic screens of a *Candida albicans* homozygous deletion library decouple morphogenetic switching and pathogenicity

Suzanne M Noble[1,2], Sarah French[2], Lisa A Kohn[1], Victoria Chen[1] & Alexander D Johnson[1]



***Candida albicans* is the most common cause of serious fungal disease in humans. Creation of isogenic null mutants of this diploid organism, which requires sequential gene targeting, allows dissection of virulence mechanisms. Published analyses of such mutants show a near-perfect correlation between *C. albicans* pathogenicity and the ability to undergo a yeast-to-hypha morphological switch *in vitro*. However, most studies have used mutants constructed with a marker that is itself a virulence determinant and therefore complicates their interpretation. Using alternative markers, we created ~3,000 homozygous deletion strains affecting 674 genes, or roughly 11% of the *C. albicans* genome. Screening for infectivity in a mouse model and for morphological switching and cell proliferation *in vitro*, we identified 115 infectivity-attenuated mutants, of which nearly half demonstrated normal morphological switching and proliferation. Analysis of such mutants revealed that virulence requires the glycolipid glucosylceramide. To our knowledge, this is the first *C. albicans* small molecule that has been found to be required specifically for virulence.**

*C. albicans* is the major cause of serious fungal infections in the United States, and *Candida* species are the fourth most commonly cultured microbe from blood[1]. Whereas some human fungal pathogens exist primarily as budding yeast cells (for example, *Cryptococcus neoformans*) or filamentous hyphal structures (for example, *Aspergillus* spp.), *C. albicans* belongs to a group that freely alternates between these and other morphologies in response to specific environmental cues. Unlike other major fungal pathogens, *C. albicans* composes part of the normal human microbiome, inhabiting the skin and gastrointestinal tract as a commensal organism. However, it also produces serious disease in people who have immune deficits, who have undergone surgical instrumentation or who have been treated with long courses of antibiotics. The attributable mortality from bloodstream *C. albicans* infections in adults is at least 15% (ref. 2). Despite the frequency and seriousness of these infections, there are few tests and antifungal drugs to diagnose and treat them.

*C. albicans* has an obligate diploid genome and lacks a complete sexual cycle, two properties that have historically hindered genetic approaches to studying this organism. Nonetheless, numerous genes important for virulence have been identified through 'reverse genetics', in which both alleles of a selected gene are inactivated and the resulting mutant evaluated for ability to cause disease. These studies have found a strong correlation between virulence and the ability to alternate among three morphological forms—yeasts, hyphae and pseudohyphae (**Fig. 1a** and **Supplementary Table 1**; ref. 3). Because the morphological switch is itself a complex transition involving changes in cellular shape, mechanical properties, interactions with other cells and differential expression of myriad genes[4–7], its close association with virulence has made it difficult to dissect out specific virulence mechanisms and effectors. Indeed, it remains uncertain whether the observed virulence defects result from defects in morphology *per se* or from genes whose expression is correlated with a particular morphological state[8].

A second factor complicating the understanding of *C. albicans* virulence is the use of *URA3* as a selectable marker in the majority of published mutants. Several studies have established that adequate expression of *URA3* is essential for normal virulence as well as for efficient transitions between yeast and filamentous forms[9–13]. Because *URA3* has been used to replace the target gene of most published mutants and because its expression can vary substantially as a function of chromosomal position[9–11], rigorous interpretation of published mutants has often proven difficult.

To investigate *C. albicans* virulence afresh, we created a large number of homozygous gene disruption mutants using auxotrophic markers that are neutral for virulence[14] (**Fig. 1b**). Because many *C. albicans* laboratory strains are aneuploid[15,16], we verified that the parental strain SN152 had a diploid chromosome number by contour clamped homogeneous electric field gel analysis[14] and comparative genome hybridization[17]. Gene targets were selected to represent a wide variety of cellular processes, and the majority were previously unstudied. We barcoded each mutant by embedding one of 48 unique oligonucleotide sequences in the gene disruption cassette. Rigorous quality control was implemented to verify deletion of both target alleles, as well as the absence of a third copy of the ORF. To control for unwanted genetic rearrangements that sometimes arise during

[1]Department of Microbiology and Immunology, University of California San Francisco, San Francisco, California, USA. [2]Department of Medicine, Division of Infectious Diseases, University of California San Francisco, San Francisco, California, USA. Correspondence should be addressed to S.M.N. (suzanne.noble@ucsf.edu).

Received 6 January; accepted 12 May; published online 13 June 2010; doi:10.1038/ng.605

mutant construction in *C. albicans*[17,18], two or more independent isolates were derived for each targeted gene. In total, we constructed ~3,000 strains that correspond to deletions of 674 ORFs, or ~11% of annotated coding genes in the *C. albicans* genome.

Three phenotypes that have been associated with virulence in animal models—infectivity in a mouse model of disseminated candidiasis, the yeast-to-hypha transition *in vitro* and proliferation rates *in vitro*—were monitored across the entire deletion collection. Contrary to expectation, these traits were not strictly correlated with each other. In particular, many infectivity-defective mutants had normal morphological transitions and proliferation *in vitro*. We hypothesized that, unlike mutants with associated defects in morphogenesis and proliferation, those with isolated infectivity defects could be used to define discrete effectors of pathogenesis. We therefore pursued four such mutants with directed biochemical and virulence experiments, resulting in the identification of a biosynthetic pathway for glucosylceramide as a novel virulence factor that is important in the mouse, but dispensable for morphogenesis and proliferation *in vitro*. To our knowledge, glucosylceramide is the first small molecule synthesized by *C. albicans* to be found to have these characteristics.

## RESULTS

### Constructing a high-quality *C. albicans* gene deletion library

We designed our knockout library to include genes involved in a broad spectrum of cellular processes and to avoid a strong bias toward factors already known to affect *C. albicans* virulence. When this work was begun, the genome of clinical strain SC5314 had been sequenced and partially assembled by the Stanford Genome Technology Center[19] We created a database that assembled the putative ORFs with results of pairwise BLAST comparisons to multiple sequenced genomes and predicted functional motifs (using the SMART algorithm). We next compiled the following nonexclusive sets of target genes: (i) genes without clear homologs in two nonpathogenic yeasts, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*; (ii) genes apparently unique to *C. albicans* (not conserved in other sequenced genomes); (iii) genes with a functional motif potentially related to virulence (for example, all genes with signatures of cell surface proteins such as signal sequences and glycosylphosphatidylinositol anchors); and (iv) additional ORFs of interest (for example, genes whose transcripts are associated with the She3-dependent RNA transport system[20]). After excluding candidates that appeared unlikely to encode proteins, those in transposons and those that were homologous to essential genes in *S. cerevisiae*, we identified 1,152 targets for disruption.

Mutants were constructed in *C. albicans* strain SN152, a derivative of SC5314 with auxotrophies for histidine, leucine and arginine[14] (**Fig. 1b**). These nutritional markers are not required for virulence in the murine tail vein injection model that is most commonly used for virulence analysis[14]. Briefly, a fusion PCR technique was used to synthesize gene disruption cassettes containing either *Candida dubliniensis HIS1* or *Candida maltosa LEU2* flanked by ~350 nucleotides matching sequences upstream and downstream of the target gene. One of 48 20-nucleotide barcodes was included adjacent to the selectable markers. Heterozygous gene disruption strains were constructed by transformation of SN152 with a *HIS1*-marked gene disruption cassette; $His^+$ transformants were screened by colony PCR for the presence of expected 5′ and 3′ junctions of the integrated DNA. Homozygous gene disruption strains were constructed by transformation of the heterozygous knockout strain with a *LEU2*-marked gene disruption cassette; $His^+$ $Leu^+$ transformants were screened for expected 5′ and 3′ junctions of the second disrupted allele and for absence of the original target ORF. At least two homozygous knockout isolates were obtained for each target gene.

We were successful in obtaining homozygous knockouts of 64% of target genes, based on the original ORF designations (**Fig. 1b**). Some of the failures undoubtedly corresponded to essential genes. In the current *C. albicans* genome annotation, the disruption mutants—comprising more than 3,000 strains—correspond to 674 unique genes. Below, we describe the characterization of these mutants by means of three genetic screens: an *in vivo* screen for infectivity, and *in vitro* screens for morphogenesis and rate of proliferation (**Fig. 1c**).

### Infectivity screen identifies 115 candidate virulence factors

Classical virulence analysis of *C. albicans* mutants typically involves infecting groups of mice with an individual mutant or a wild-type strain and comparing the course of illness; virulence-defective mutants produce less disease or a delay in disease onset. To minimize the number of animals required to characterize our library of mutants, we adapted the signature-tagged mutagenesis technique, which has been highly successful in identifying novel virulence factors of bacterial

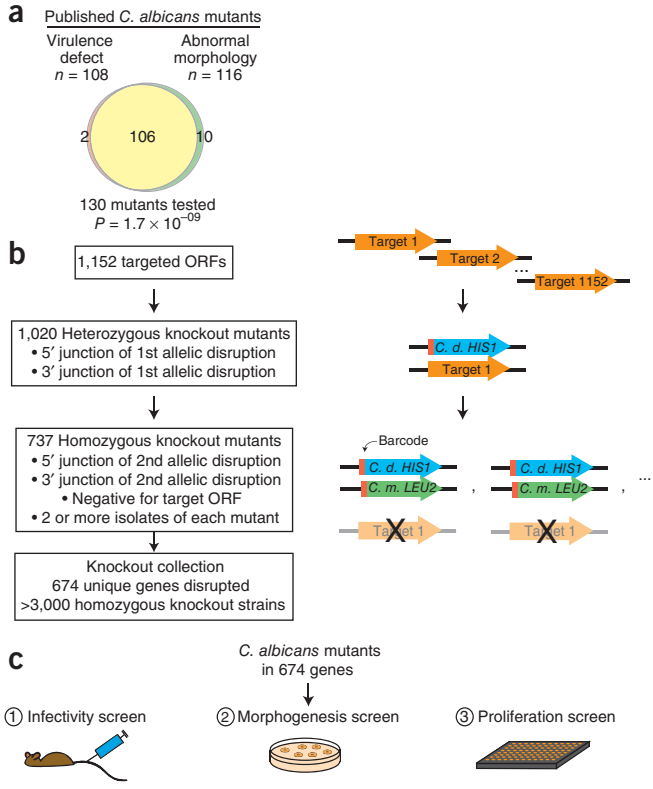


**Figure 1** *C. albicans* mutants and screens. (**a**) Venn diagram, based on the published literature, depicting virtually complete overlap between *C. albicans* mutants with virulence defects and mutants with abnormal morphology (see Online Methods). Significance was calculated using the hypergeometric distribution. (**b**) Construction of the homozygous gene disruption library by homologous recombination. One allele of each target gene was replaced by *C. dubliniensis HIS1*, and the second allele was replaced by *C. maltosa LEU2*. Red rectangles indicate the oligonucleotide barcodes. Successful knockout strains were defined as those with PCR-verified junctions at the 5′ and 3′ ends of the selectable markers, as well as absence of the target ORF. (**c**) *In vivo* and *in vitro* screens of the *C. albicans* mutant library. As described in Online Methods, the 674 mutants were screened for infectivity in the mouse, for colony morphology on solid Spider medium at 30 °C and for proliferation in liquid synthetic complete medium at 37 °C.

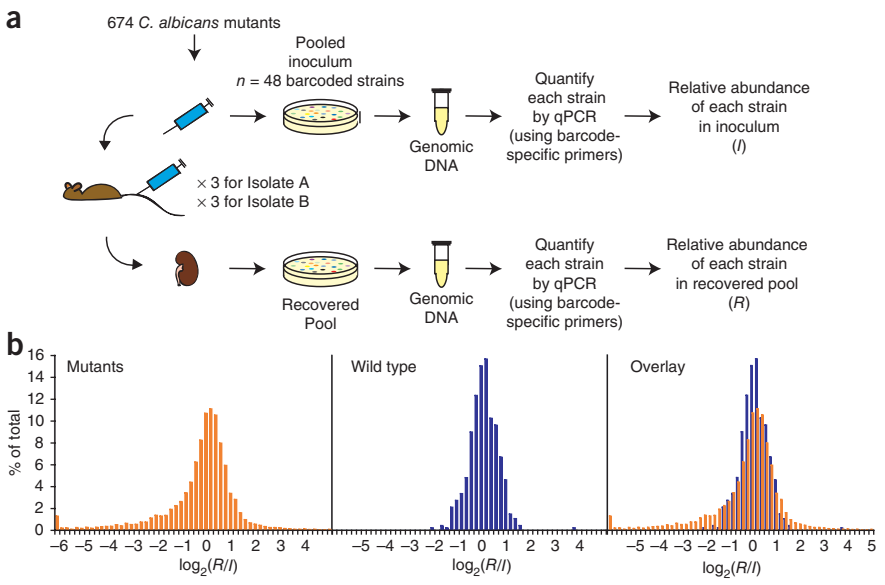


**Figure 2** Infectivity screen. (**a**) Schematic of the infectivity screen. Pools of up to 48 mutants and the wild-type strain were used to infect BALB/c mice. Samples of *C. albicans* from the infecting inoculum and recovered from mouse kidneys were plated on Sabouraud agar. Genomic DNA was recovered, and the relative abundance of each strain in the inoculum and the recovered pool was determined by real-time PCR (qPCR) using primers specific to the DNA barcodes. (**b**) Histograms depicting $\log_2(R/I)$ values for *C. albicans* mutants (orange), the wild-type comparator (blue) and the overlay between the two groups.

pathogens[21–23]. The method involves simultaneous infection of an animal model with numerous mutants of interest, each marked by an oligonucleotide 'signature tag' that allows for individual quantification. Theoretically and in practice, mutants with virulence defects have reduced competitive fitness and are selectively depleted in the host.

In the mouse model of disseminated candidiasis, kidneys are the organs of greatest fungal proliferation[24], and kidney tissue burden has been highly correlated with virulence among published mutants[3] (see **Supplementary Table 2**). We used pools of 48 barcoded *C. albicans* mutants and a matched wild-type strain to infect BALB/c mice by tail vein injection (**Fig. 2a**). After disease had progressed, the mice were euthanized, and the abundance of *C. albicans* strains recovered from host kidneys ($R$) and the infecting inoculum ($I$) was determined by real-time PCR (using primers to the signature tags). Each mutant was evaluated in at least six mice, and at least two independent isolates of each mutant were assessed.

We screened a total of 118 pools, and the results—representing 6,917 real-time PCR measurements of *C. albicans* knockout mutants and 472 measurements of wild-type yeast—are presented in **Figure 2b**. The $x$-axis represents the $\log_2$ ratio of the abundance of each strain in the recovered pool relative to the inoculum. A strain whose representation neither increased nor decreased in the infected mouse would have a ratio of 1, or a $\log_2$ ratio of 0. A strain that became underrepresented over the course of infection would be shifted to the left, and one that accumulated would be shifted to the right. Inspection of the histogram of wild-type *C. albicans*, derived from data acquired from the majority of mice throughout the screen (blue bars), revealed a bell-shaped distribution consistent with simple variation around a mean. In contrast, the distribution of the mutants (orange bars) was asymmetric, skewed to the left. We assumed that the prominence of the left-sided tail resulted from mutants with fitness defects in infecting or proliferating in mouse kidneys; we hypothesized that many of these mutants would also be defective in causing disease. Mutants shifted to the right may have had a fitness advantage resulting in hyperproliferation or a defect, for example, in exiting the kidney.

Although many mutants showed obvious defects in infectivity (with $\log_2(R/I)$ values shifted far to the left in **Fig. 2b**), we used two systematic, statistics-based tests to also capture mutants with moderate defects. This analysis, along with estimates of false discovery rates and our results with previously published virulence mutants, are described in detail in the Online Methods. Using these tests, we identified 115 mutants with substantial defects in infectivity (**Supplementary Table 3**). These mutations disrupt genes involved in a broad range of cellular processes; as might be expected, they also affect a sizeable group of genes whose functions are completely unknown.

## Morphogenesis screen reveals determinants of dimorphism

As described above, the yeast-to-hypha transition is closely linked to virulence in the *C. albicans* literature. Usually referred to as a

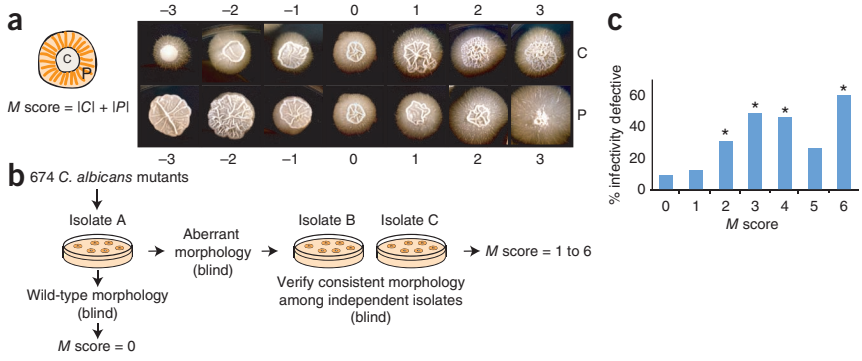


**Figure 3** Colony morphology screen. (**a**) *C. albicans* colony morphology. Left, schematic showing that *C. albicans* colonies are bipartite: a central region (C) comprises yeast, pseudohyphae and true hyphae, and a peripheral region (P) comprises mainly pseudohyphae and hyphae. Right, examples of *C. albicans* colonies that illustrate the spectrum of morphologies observed in the screen. Wild-type morphology is scored as 0, and aberrant morphology is scored from 1 to 3 (increased relative to wild type) and –1 to –3 (decreased relative to wild type) for both $C$ and $P$ scores. The morphology score ($M$ score) of each mutant is calculated as the sum of the absolute values of $C$ and $P$. (**b**) Schematic of the colony morphology screen. One isolate of each of the 674 mutants and the wild-type strain was initially plated for single colonies on Spider medium. Mutants showing morphology different from the wild type underwent testing of one to three additional isolates (depending on availability) to confirm consistency of the phenotype. (**c**) Histogram of infectivity defects among mutants with different $M$ scores. *Significant association with infectivity defect (using the hypergeometric distribution); $P = 0.0077$ for $M$ score of 2, $P = 4.8 \times 10^{-6}$ for $M$ score of 3, $P = 4.8 \times 10^{-6}$ for $M$ score of 4, $P = 0.10$ for $M$ score of 5, $P = 0.027$ for $M$ score of 6.

**Table 1 Results of the morphogenesis and proliferation screens**

| Morphology | | Proliferation | | |
|---|---|---|---|---|
| *M* score | Number of mutants | Doubling time (min) | No. s.d. above $DT_{WT}$ | Number of mutants |
| 0 | 504 (77%) | <92.2 | 0 | 396 (59%) |
| 1 | 16 (2%) | 92.3–97.4 | 0.5 | 107 (16%) |
| 2 | 39 (6%) | 97.5–102.8 | 1 | 68 (10%) |
| 3 | 33 (5%) | 102.9–108.1 | 1.5 | 35 (5%) |
| 4 | 37 (6%) | 108.2–113.4 | 2 | 21 (3%) |
| 5 | 19 (3%) | 113.5–118.7 | 2.5 | 5 (<1%) |
| 6 | 5 (<1%) | >118.8 | 3 or more | 42 (6%) |

Morphogenesis results are presented as the number and percentage of mutants with a given *M* score. Results were omitted for four mutants with inconsistent results and 17 mutants with poor growth on Spider medium. Proliferation results are presented as the number and percentage of mutants with a given doubling time. Doubling time is also expressed as the number of s.d. (No. s.d.) greater than the doubling time of wild-type yeast ($DT_{WT}$).

'dimorphic' transition, this process actually involves interconversions among three types of cells: budding oval yeast cells that separate after cell division, and two types of elongated cells that remain attached as filaments, called hyphae and pseudohyphae[4,25]. Interconversions among the three cell types are induced *in vitro* by modifying the pH, phosphate concentration, temperature or serum concentration in the growth medium, or a combination of these variables, which are thought to mimic environmental cues found within different niches of the host.

The macroscopic appearance of a *C. albicans* colony derives from the types of the cells that make it up, and changes in colony morphology have been used as a sensitive indicator of mutants with defects in the dimorphic transition[26]. On a variety of media, wild-type *C. albicans* generates colonies with two distinct regions (**Fig. 3a**): a central region (C) consisting of yeast, pseudohyphae and true hyphae, and a peripheral region (P) consisting primarily of filamentous forms (hyphae and pseudohyphae). Mutants shifted toward filamentous forms produce colonies with increased central wrinkling or peripheral filamentation, or both; mutants shifted toward the yeast morphology produce smooth colonies with decreased peripheral filamentation.

We examined colony morphology across our mutant library. Five to ten cells of each mutant were plated on solid Spider medium and incubated at 30 °C (**Fig. 3b**). Colonies were photographed at 14 days, and central and peripheral regions were each assigned a score (*C* and *P*, respectively) on a scale of −3 (decreased morphology) to +3 (increased morphology), with 0 indicating wild-type morphology. The overall morphology score (*M* score) was the sum of the absolute values of *C* and *P* and ranged from 0 (wild-type) to 6 (highly aberrant). Mutants with *M* scores of other than 0 underwent testing of additional isolates to confirm the consistency of the phenotype. Ultimately, 504 mutants were indistinguishable from the wild type (*M* score of 0), 16 had mildly altered morphology (*M* score of 1) and 133 had substantial aberrations (*M* score of 2 or greater; **Table 1**). Four mutants had inconsistent defects (different abnormal morphologies), most likely resulting from unlinked mutations in at least one of the isolates, and 17 mutants grew slowly on Spider medium; these were excluded from subsequent analysis. Mutants with substantial morphology defects (*M* score of 2 or more) are presented in **Supplementary Table 4**, and data for all of the mutants are in **Supplementary Table 5**.

We next examined the correlation between defects in infectivity and defects in the yeast-to-hypha transition by determining the frequency of infectivity defects among mutants with a given *M* score. If the ability to undergo the dimorphic transition were the major determinant of virulence, then the proportion of infectivity-defective mutants would be expected to increase as a function of the *M* score. This prediction was only partially borne out by our data (**Fig. 3c**). Mutants with an *M* score >0 were more likely to have infectivity defects than those with wild-type morphology (37% versus 10%; $P = 6.5 \times 10^{-14}$, calculated using the hypergeometric distribution), in keeping with the prevailing view that morphogenesis and infectivity are associated. However, among 24 mutants with the highest *M* scores of 5 and 6, two-thirds showed normal infectivity. Likewise, among 103 infectivity-defective mutants that were scored for morphology, 48, or 47%, had normal colony morphology on Spider medium. These results demonstrate that the relationship between infectivity and morphogenesis cannot be one of simple cause and effect.

Although colony morphology is a sensitive indicator of mutants with defects in the dimorphic switch, we also employed a specific test of the ability of yeast cells to form hyphae. We tested the group of 115 infectivity-defective mutants for a hyphal response to serum, which is the 'germ tube test' used by clinical laboratories to distinguish *C. albicans* from other yeasts. Only 15 mutants did not elaborate true hyphae (**Supplementary Table 6**), all of which were also captured by the colony morphology assay ($n = 12$) or grew too slowly on this medium to be scored ($n = 3$).

## Proliferation screen identifies mutants with slow growth

It is a reasonable assumption, and one often made in the literature[27], that a *C. albicans* mutant showing slow proliferation *in vitro* will have a corresponding virulence defect *in vivo*. We addressed this assumption systematically by measuring proliferation rates across the mutant library and comparing them with the infectivity data. Wild-type *C. albicans* and two isolates of each mutant were grown in chemically defined liquid medium at 37 °C, and doubling times were measured (**Supplementary Table 5**). The median doubling time of wild-type yeast under these conditions was 87 min.

Sixty-eight mutants had significant defects in proliferation, defined as a doubling time more than two s.d. greater than that of the wild type (**Table 1**). Contrary to our expectations, the majority (42, or 62%) of these proliferation-defective mutants were competent for infectivity (**Fig. 4a**). Likewise, among the 115 infectivity-defective mutants, the vast majority (89 or 77%) showed a normal rate of

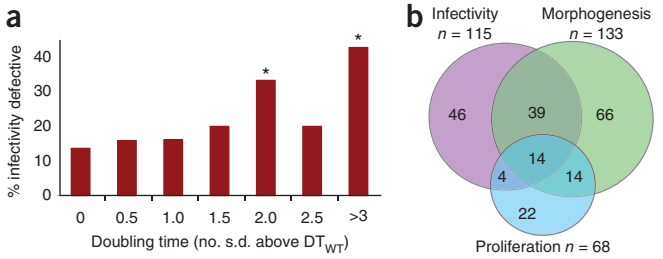


**Figure 4** Proliferation screen and Venn diagram. (**a**) Histogram of infectivity defects among mutants with different doubling times. *Significant association with infectivity defect (using the hypergeometric distribution); $P = 0.11$ for 0.5 s.d. above $DT_{WT}$, $P = 0.13$ for 1 s.d. above $DT_{WT}$, $P = 0.15$ for 1.5 s.d. above $DT_{WT}$, $P = 0.034$ for 2 s.d. above $DT_{WT}$, $P = 0.40$ for 2.6 s.d. above $DT_{WT}$ and $P = 3.3 \times 10^{-5}$ for >3 s.d. above $DT_{WT}$. (**b**) Venn diagram illustrating the overlap of mutants from the three screens. Mutants with slow growth or inconsistent growth on Spider medium could not be scored for morphogenesis and are not represented in the figure; these included four mutants with infectivity-specific defects, eight with combined infectivity and proliferation defects, and one with a proliferation-specific defect.

**Table 2 Mutants with defects in infectivity but not in morphogenesis or proliferation**

| Disrupted gene | Function of protein product |
|---|---|
| *orf19.3384* | Cell wall component: protein |
| *orf19.529* | Cell wall component: protein |
| *PGA32* (*orf19.6784*) | Cell wall component: protein |
| *CHS4* (*orf19.7349*) | Cell wall composition: carbohydrate |
| *CHS7* (*orf19.2444*) | Cell wall composition: carbohydrate |
| *CHT2* (*orf19.3895*) | Cell wall composition: carbohydrate |
| *orf19.753* | Cell wall composition: carbohydrate |
| *DUN1* (*orf19.4002*) | DNA damage response |
| *ESC4* (*orf19.1445*) | DNA damage response |
| *CRZ1* (*orf19.7359*) | Gene expression: transcription factor |
| *MTLA1* (*orf19.3201*) | Gene expression: transcription factor |
| *orf19.2315* | Gene expression: transcription factor |
| *orf19.7397* | Gene expression: transcription factor |
| *SEF1* (*orf19.3753*) | Gene expression: transcription factor |
| *orf19.2961* | Gene expression: transcription factor |
| *CYB1* (*orf19.7049*) | Lipid biosynthesis and regulation |
| *HET1* (*orf19.6327*) | Lipid biosynthesis and regulation |
| *orf19.3283* | Lipid biosynthesis and regulation |
| *orf19.4831* | Lipid biosynthesis and regulation |
| *orf19.6411* | Lipid biosynthesis and regulation |
| *MID1* (*orf19.3212*) | Metal ion homeostasis: calcium |
| *CFL2* (*orf19.1264*) | Metal ion homeostasis: iron |
| *HAK1* (*orf19.6249*) | Metal ion homeostasis: potassium |
| *orf19.3290* | Miscellaneous: aerobic respiration |
| *CDC10* (*orf19.548*) | Miscellaneous: cell cycle |
| *NMD5* (*orf19.4188*) | Miscellaneous: nuclear import of proteins |
| *orf19.1392* | Miscellaneous: protein folding |
| *CDC19* (*orf19.3575*) | Nutrient acquisition and metabolism |
| *HGT12* (*orf19.7094*) | Nutrient acquisition and metabolism |
| *HGT8* (*orf19.2021*) | Nutrient acquisition and metabolism |
| *orf19.449* | Nutrient acquisition and metabolism |
| *orf19.4905* | Nutrient acquisition and metabolism |
| *orf19.5499* | Nutrient acquisition and metabolism |
| *orf19.7288* | Nutrient acquisition and metabolism |
| *TRP2* (*orf19.10080*) | Nutrient acquisition and metabolism |
| *RGA2* (*orf19.4593*) | Signaling: cell polarity |
| *KAR2* (*orf19.2013*) | Signaling: unfolded protein response |
| *orf19.1276* | Unknown |
| *orf19.2653* | Unknown |
| *orf19.2726* | Unknown |
| *orf19.3335* | Unknown |
| *orf19.5509* | Unknown |
| *orf19.557* | Unknown |
| *orf19.5782* | Unknown |
| *PDR6* (*orf19.2094*) | Unknown |
| *RBT4* (*orf19.6202*) | Unknown |

Gene names and proven or predicted protein functions are presented for mutants that showed defects in infectivity (as defined in text) but not in colony morphology (*M* scores were 1 or less) or proliferation ($DT_{mutant} < DT_{WT} + 2$ s.d.).

proliferation *in vitro* (**Supplementary Table 5**). The simplest interpretation of these results is that the requirements for proliferation *in vitro* are likely to differ substantially from those required for growth in the host. For example, nutrient limitation in the host plus direct attack by the immune system may result in very long doubling times *in vivo*, rendering irrelevant many of the differences that we observed under optimal conditions *in vitro*. At a minimum, these results invalidate the common assumption that poor growth *in vitro* necessarily correlates with decreased virulence.

A Venn diagram summarizing the results of the three systematic screens is presented in **Figure 4b**. Overall, approximately one-third of the mutants in the library showed a defect in at least one screen. The majority had defects in a single phenotype.

### Analysis of mutants predicted to affect glycolipid biosynthesis

Notwithstanding the previous arguments, we focused on infectivity-defective mutants with normal colony morphology and normal proliferation *in vitro* (**Table 2**), hypothesizing that mutants with these characteristics would be most likely to reveal pathways used exclusively for virulence. Among the mutants meeting these criteria, two caught our attention because their mutations were predicted to affect proteins involved in the synthesis of sphingolipids—a class of molecules not previously associated with *C. albicans* virulence. Orf19.4831 was homologous to known sphingolipid methyltransferases/cyclopropane synthases, and Het1 was homologous to known sphingolipid transfer proteins. Inspection of mutants with less consistent infectivity defects revealed two additional mutations affecting proteins that could be ordered with Orf19.4831 and Het1 into a hypothetical biosynthetic pathway for glucosylceramide (**Fig. 5a**). Hsx11 was previously shown to possess sphingolipid glucosyltransferase activity[28], and Orf19.260 was homologous to known sphingolipid desaturases.

To experimentally determine whether *hsx11*, *orf19.260* (also called *sld1*), *orf19.4831* and *het1* truly defined a glucosylceramide biosynthetic pathway in *C. albicans*, we compared membrane lipids from the four mutants and the wild-type strain. Crude lipid fractions prepared from equal masses of cells were resolved using thin layer chromatography (TLC), and glucosylceramide was visualized with anthrone reagent (**Fig. 5b**); soybean glucosylceramide served as a mobility marker (lanes 1 and 2). The wild-type strain synthesized a product that comigrated with bona fide glucosylceramide (lane 3). By comparison, *hsx11* and *orf19.260* produced substantially less comigrating product (lanes 4 and 5), *orf19.4831* had only a mild decrease (lane 6), and *het1* produced as much comigrating product as the wild type (lane 7).

TLC cannot resolve mature glucosylceramide from the preceding two intermediates (**Fig. 5a**). Therefore, to determine the identities of the comigrating lipids, we used mass spectrometry. After a further purification step, we resolved lipids from wild-type, *hsx11*, *orf19.260*, *orf19.4831* and *hsx11* strains by TLC, recovered species migrating at the position of mature glucosylceramide and subjected the lipids to matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF) mass spectrometry. Soybean glucosylceramide was analyzed as a positive control and showed its expected molecular weight (736 for the sodium salt, **Fig. 5c**). Wild-type *C. albicans* produced a product consistent with mature fungal glucosylceramide (778 molecular weight for sodium salt; **Fig. 5d**). The *hsx11* mutant produced no product (**Fig. 5e**), which was expected, as an intermediate lacking the glucose moiety would not comigrate by TLC. *orf19.260* and *orf19.4831* accumulated products of molecular weights 766 and 764, respectively (**Fig. 5f,g**); these are the weights of the sodium salts of the final two intermediates in **Figure 5a** and indicate that, consistent with our model, Orf19.260 and Orf19.4831 mediate sphingolipid C8 desaturation and C9 methylation reactions, respectively. *het1* accumulated a product consistent with mature glucosylceramide (**Fig. 5h**); this mutation affects a predicted sphingolipid transfer protein expected to alter the localization but not synthesis of glucosylceramide. Together, these results confirm both the biosynthetic pathway of **Figure 5a** and our assignments of the genes responsible for each step.

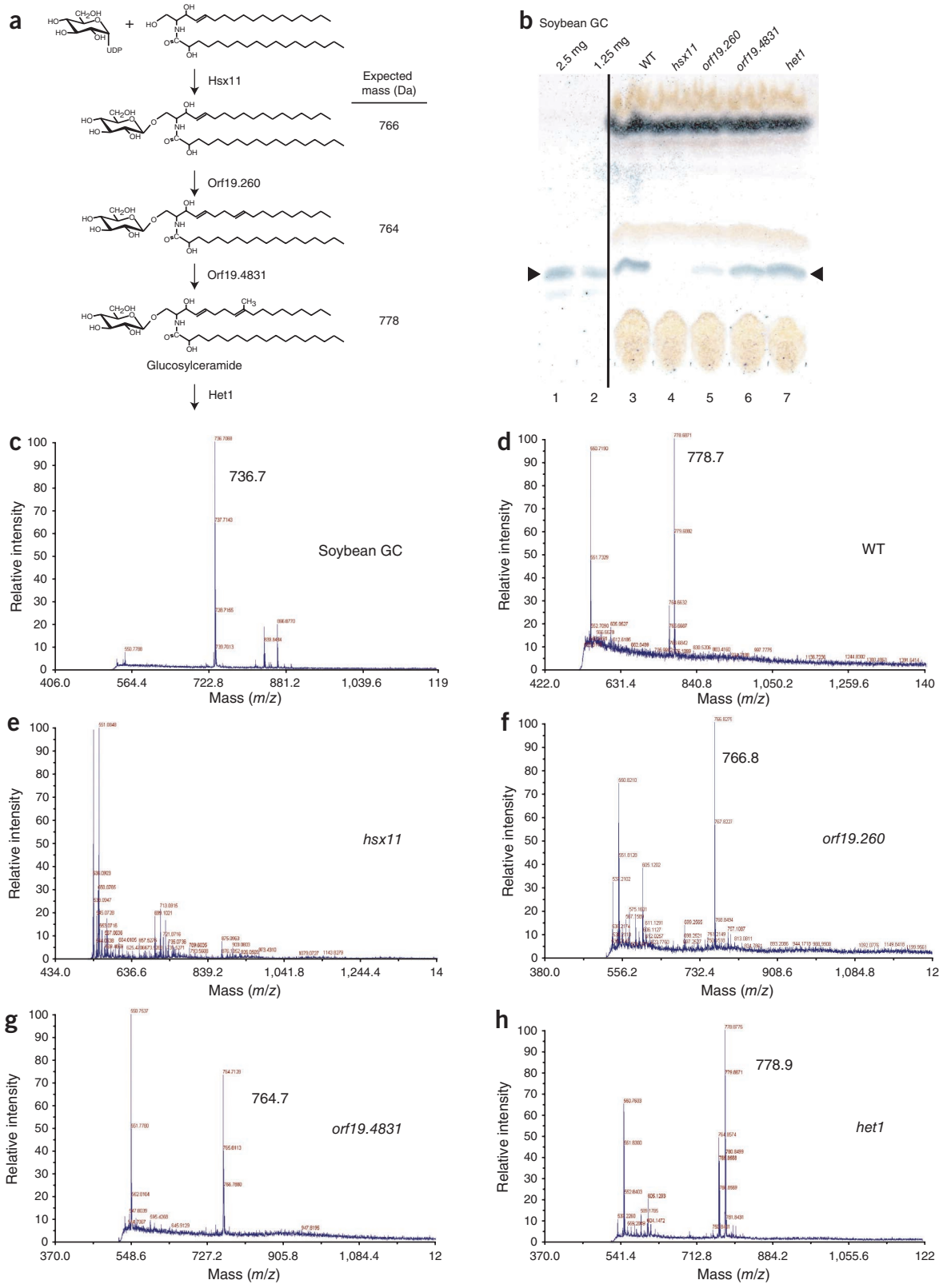


**Figure 5** Characterization of mutants affecting glucosylceramide biosynthesis. (**a**) Schematic of the predicted glucosylceramide synthesis pathway and expected molecular masses. The position of the glucosyltransferase reaction mediated by Hsx1 has not been determined and may occur as depicted or elsewhere in the pathway. Note that the masses given are for the sodium salts of each molecule. (**b**) TLC plate with glucosylceramide (GC) from soybean (lanes 1 and 2) and lipids extracted from equal masses of wild-type (lane 3), *hsx11* (lane 4), *orf19.260* (lane 5), *orf19.4831* (lane 6) and *het1* (lane 7) strains. The markers indicate the position of glucosylceramide, visualized with anthrone reagent. (**c**–**h**) MALDI-TOF spectra for a soybean glucosylceramide standard and lipid extracted from wild-type, *hsx11*, *orf19.260*, *orf19.4831* and *het1* strains.

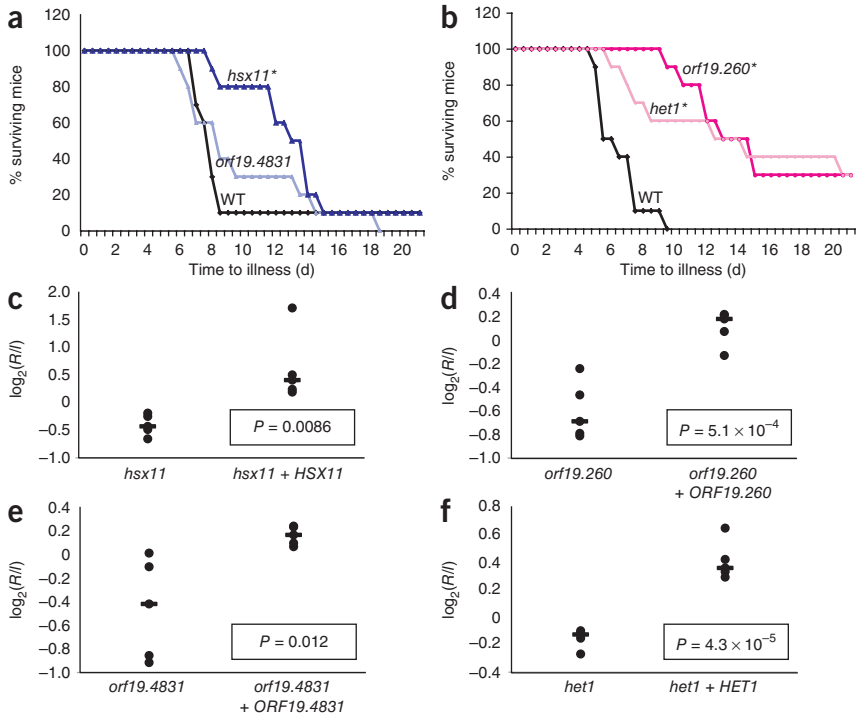


**Figure 6** Virulence analysis of mutants affecting glucosylceramide biosynthesis. (**a**,**b**) The glucosylceramide pathway is required for virulence. As described in the text, BALB/c mice were infected with wild-type *C. albicans* or one of the glucosylceramide pathway mutants, and time to illness was monitored. *Significant difference from wild type by the logrank test; $P = 0.0020$ for *hsx11*, $P < 0.0001$ for *orf19.260*, $P > 0.05$ for *orf19.4831* and $P = 0.032$ for *het1*. (**c**–**f**) Wild-type genes restore competitive fitness to the glucosylceramide pathway mutants. Two-strain infectivity assays were performed comparing each mutant to a gene add-back strain in five BALB/c mice. Shown are the *R*/*I* values of each strain in individual mice (black dots), the median *R*/*I* value (black line) and the *P* value of the observed differences.

## Glucosylceramide is required for virulence

We next tested the virulence of mutants affecting the glucosylceramide pathway in single (as opposed to pooled) infections of BALB/c mice. Ten mice were infected with each mutant or wild-type *C. albicans*, and time to illness was monitored. *hsx11*, *orf19.260* and *het1* had significant virulence defects (**Fig. 6a**,**b**), whereas *orf19.4831* showed a trend toward reduced virulence that fell outside our significance threshold of 0.05 (**Fig. 6a**). To establish whether deletion of the target genes was responsible for these phenotypes, we constructed 'add-back' strains into which one copy of the wild-type gene was restored. When mice were co-infected with each mutant and its cognate gene add-back strain, the add-back strain had superior fitness in each case (**Fig. 6c**–**f**). The observation that *ORF19.4831* complemented the infectivity defect of the *orf19.4831* mutant supports our conclusion above that this mutant is mildly attenuated for virulence. These results establish glucosylceramide biosynthesis as a novel and important determinant of *C. albicans* virulence.

The *C. neoformans* homolog of *hsx11* is also required for virulence in a murine model, but this mutant has an *in vitro* growth defect at neutral or alkaline pH[29]. We assessed the growth of the *C. albicans* glucosylceramide pathway mutants on solid laboratory media buffered to pH 7, 7.5, 8, 9 or 10, and found no differences from the wild-type strain (**Supplementary Fig. 1**). Therefore, in contrast to *C. neoformans*, *C. albicans* glucosylceramide probably mediates virulence through a mechanism other than facilitating growth at mammalian pH.

## DISCUSSION

*Candida albicans* is the most common fungal pathogen of humans, yet progress in understanding its virulence program has been difficult. Its diploid genome, chromosomal instability and incomplete sexual cycle have made genetic approaches especially challenging. In this study, we describe the construction and analysis of homozygous gene deletion mutants that encompass 674 genes. The library contains at least two independent isolates affecting each gene of interest and was created without the use of *URA3*—an auxotrophic marker that has clouded the interpretation of most of the previously described mutants in this field. The knockouts in our collection of ~3,000 strains cover ~11% of the *C. albicans* genome, including genes predicted to affect a wide range of cellular processes, as well as ones whose functions are completely unknown. We detected no aneuploidy in a survey of 12 mutants using a quantitative PCR assay (six with infectivity defects and six without; **Supplementary Table 7**); however, because aneuploidy can occur during strain construction[17], we recommend the use of complementation tests for mutants of special interest. Below, we summarize insights gained from systematic screens of the library for (i) host (kidney) colonization, (ii) the yeast-to-hypha transition and (iii) rate of proliferation *in vitro*.

Prior to this work, fewer than 200 *C. albicans* mutants had been tested for virulence in the mouse model of disseminated disease[3], and there was a near-perfect correlation between mutants with virulence defects and those affecting the yeast-to-hypha transition (**Fig. 1a**). This correlation was based on myriad studies of individual gene disruption mutants that had been created in different laboratories, using different *C. albicans* strain backgrounds, and tested in different, nonstandardized assays for virulence and morphogenesis. An exception is an infectivity study of a large number of *C. albicans* mutants created by disrupting one copy of each target gene and placing the other allele under control of a repressible promoter[30]. In this case, however, the identities of the disrupted genes were not revealed, the collection was not made freely available, and effects of the mutations on morphogenesis were not reported.

Using our library, we were able to directly compare the results of a standardized screen for infectivity to those for morphogenesis and proliferation *in vitro*, resulting in several general conclusions. First, across the mutant library, defects in the three properties were relatively common. Approximately 17% of mutants showed reproducible defects in infectivity, 21% of mutants had abnormal morphology, and 10% of mutants were defective for proliferation (**Supplementary Table 5**). The proportion of mutants with infectivity defects was somewhat higher than that reported in signature-tagged mutagenesis screens of bacterial pathogens[21–23] and far lower than the >80% rate of virulence-defective mutants reported in the *C. albicans* literature[3]. Second, infectivity-defective mutants had a higher-than-average rate of morphology abnormalities (43% had *M* scores of 2 or more), and morphogenesis-defective mutants were similarly enriched for infectivity defects (40%

by the Boolian test), supporting the well-described crucial role of the yeast-to-hypha transition in virulence. Finally, our unexpected discovery that a large group of mutants had substantial defects in infectivity without associated defects in morphogenesis or proliferation *in vitro* suggests that the process of virulence can be teased apart from the complex morphogenetic program, thereby providing a strategy for identifying discrete effectors of virulence.

We tested the ability of our data set to predict novel virulence effectors by focusing on four mutants with specific defects in infectivity but not in morphogenesis or proliferation. Using biochemical methods and mass spectrometry, we demonstrated that Hsx11, Orf19.260 and Orf19.4831 mediate sequential steps in the biosynthesis of fungal-specific glucosylceramide (**Fig. 5**). Given its homology to lipid transfer proteins, Het1 probably localizes the final product. Hsx11 is already known to possess glucosyltransferase activity[28], and recently Orf19.260 (also known as Sld1) was independently shown to have sphingolipid desaturase activity[31], consistent with our findings. In two-strain infectivity experiments, all four mutants showed defects in colonizing host kidneys that were complemented by restoration of the wild-type genes; *hsx11*, *orf19.260* and *het1* were also considerably less virulent in classical monotypic infections (**Fig. 6**). Thus, the glucosylceramide biosynthetic pathway is crucial for causing disease as well as for proliferation in the kidney. Not all fungi retain this pathway; for example, *S. cerevisiae* and *S. pombe* lack all four genes involved in glucosylceramide synthesis and (proposed) localization. Although several small molecules have previously been associated with virulence of *C. albicans*[10,32–42], glucosylceramide is the first such virulence effector found to act independently of the yeast-to-hypha transition.

*C. albicans* is unique in being both a normal human fungal commensal and the most common cause of serious fungal disease. As exemplified by the glucosylceramide pathway, our identification of 46 genes whose disruption leads to reproducible defects in infectivity in the host but to no observable changes in morphogenesis or proliferation *in vitro* (**Table 2**) offers opportunities for understanding the relationship between *C. albicans* and the mammalian host. The knockout collection should also facilitate investigations of other clinically relevant aspects of *C. albicans* biology. For example, little is known regarding how *C. albicans* (or any other organism) functions as a commensal or how the determinants of superficial epithelial infections differ from those required for invasive, disseminated infections. The development of experimental models of these broader aspects of *Candida* biology, coupled with the genetic resource and strategies described here, promise new insights into the biology of this important human pathogen.

*Note added in proof: Another group has independently demonstrated[43] that Orf19.4831 has C9-methyltransferase activity; they name the gene MTS1.*

## METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

**Accession codes.** Homozygous disruption mutants will be deposited with the Fungal Genetics Stock Center.

*Note: Supplementary information is available on the Nature Genetics website.*


### ACKNOWLEDGMENTS

We are extremely grateful to the Stanford Genome Technology Center for releasing the *C. albicans* genomic sequence before publication, and to the curators of the *Candida* Genome Database, which has been an invaluable tool in our analysis of the published literature. We are also grateful to many colleagues at the University of California, San Francisco (UCSF), and elsewhere who were generous with advice, protocols and reagents. S. O'Brien constructed our *C. albicans* bioinformatics database. B. Hromatka assisted with construction of *C. albicans* mutants. J. Zhe and J. De Risi provided a database of putative unique *C. albicans* ORFs. M. Fischbach suggested the link between our sphingolipid-associated mutants and the glucosylceramide pathway. E. Chow and J. Cox provided expertise and equipment for lipid extraction. P.E. Bailey and the Agard, Morgan, Shokat and Walter labs at UCSF provided training and access to their mass spectrometer. B. al-Sady and G. Narlikar shared their mass spectrometry plate. H. Madhani wrote software for data analysis. S. Desta prepared hundreds of liters of laboratory media. H. Madhani, J. Cox, L. Connolly and R. Locksley provided helpful comments on the manuscript. This work was supported by US National Institutes of Health grants RO1 A149187 and K08 AI062800, and the UCSF Program for Breakthrough Biomedical Research. S.M.N. received a Howard Hughes Medical Institute Postdoctoral Research Fellowship for Physicians and a Burroughs Wellcome Fund Career Award in the Biomedical Sciences during the course of these studies.


### AUTHOR CONTRIBUTIONS

S.M.N. constructed mutants, conceived and carried out the majority of experiments, and wrote the manuscript. S.F. assisted with validation of mutants, contributed to the morphogenesis screen and performed the germ tube screen. L.A.K. assisted with mutant construction and validation. V.C. assisted with the morphogenesis screen. A.D.J. provided guidance and participated in writing the manuscript.

### COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.


1. Edmond, M.B. *et al.* Nosocomial bloodstream infections in United States hospitals: a three-year analysis. *Clin. Infect. Dis.* **29**, 239–244 (1999).
2. Zaoutis, T.E. *et al.* The epidemiology and attributable outcomes of candidemia in adults and children hospitalized in the United States: a propensity analysis. *Clin. Infect. Dis.* **41**, 1232–1239 (2005).
3. Skrzypek, M.S. *et al.* New tools at the *Candida* Genome Database: biochemical pathways and full-text literature search. *Nucleic Acids Res.* **38**, D428–D432 (2010).
4. Sudbery, P., Gow, N. & Berman, J. The distinct morphogenic states of *Candida albicans*. *Trends Microbiol.* **12**, 317–324 (2004).
5. Enjalbert, B. & Whiteway, M. Release from quorum-sensing molecules triggers hyphal formation during *Candida albicans* resumption of growth. *Eukaryot. Cell* **4**, 1203–1210 (2005).
6. Kadosh, D. & Johnson, A.D. Induction of the *Candida albicans* filamentous growth program by relief of transcriptional repression: a genome-wide analysis. *Mol. Biol. Cell* **16**, 2903–2912 (2005).
7. Whiteway, M. & Bachewich, C. Morphogenesis in *Candida albicans*. *Annu. Rev. Microbiol.* **61**, 529–553 (2007).
8. Saville, S.P. *et al.* Inhibition of filamentation can be used to treat disseminated candidiasis. *Antimicrob. Agents Chemother.* **50**, 3312–3316 (2006).
9. Brand, A., MacCallum, D.M., Brown, A.J., Gow, N.A. & Odds, F.C. Ectopic expression of URA3 can influence the virulence phenotypes and proteome of *Candida albicans* but can be overcome by targeted reintegration of URA3 at the RPS10 locus. *Eukaryot. Cell* **3**, 900–909 (2004).
10. Cheng, S. *et al.* Evaluation of the roles of four *Candida albicans* genes in virulence by using gene disruption strains that express URA3 from the native locus. *Infect. Immun.* **71**, 6101–6103 (2003).
11. Lay, J. *et al.* Altered expression of selectable marker URA3 in gene-disrupted *Candida albicans* strains complicates interpretation of virulence studies. *Infect. Immun.* **66**, 5301–5306 (1998).
12. Sharkey, L.L., Liao, W.L., Ghosh, A.K. & Fonzi, W.A. Flanking direct repeats of hisG alter URA3 marker expression at the HWP1 locus of *Candida albicans*. *Microbiology* **151**, 1061–1071 (2005).
13. Sundstrom, P., Cutler, J.E. & Staab, J.F. Reevaluation of the role of HWP1 in systemic candidiasis by use of *Candida albicans* strains with selectable marker URA3 targeted to the ENO1 locus. *Infect. Immun.* **70**, 3281–3283 (2002).
14. Noble, S.M. & Johnson, A.D. Strains and strategies for large-scale gene deletion studies of the diploid human fungal pathogen *Candida albicans*. *Eukaryot. Cell* **4**, 298–309 (2005).
15. Chen, X., Magee, B.B., Dawson, D., Magee, P.T. & Kumamoto, C.A. Chromosome 1 trisomy compromises the virulence of *Candida albicans*. *Mol. Microbiol.* **51**, 551–565 (2004).
16. Selmecki, A., Bergmann, S. & Berman, J. Comparative genome hybridization reveals widespread aneuploidy in *Candida albicans* laboratory strains. *Mol. Microbiol.* **55**, 1553–1565 (2005).
17. Arbour, M. *et al.* Widespread occurrence of chromosomal aneuploidy following the routine production of *Candida albicans* mutants. *FEM. Yeast Res.* **9**, 1070–1077 (2009).
18. Bouchonville, K., Forche, A., Tang, K.E., Selmecki, A. & Berman, J. Aneuploid chromosomes are highly unstable during DNA transformation of *Candida albicans*. *Eukaryot Cell* **8**, 1554–1566 (2009).
19. Jones, T. *et al.* The diploid genome sequence of *Candida albicans*. *Proc. Natl. Acad. Sci. USA* **101**, 7329–7334 (2004).

20. Elson, S.L., Noble, S.M., Solis, N., Filler, S.G. & Johnson, A.D. An RNA transport system in *Candida albicans* regulates hyphal morphology and invasive growth. *PLoS Genet.* **5**, e1000664 (2009).
21. Mecsas, J. Use of signature-tagged mutagenesis in pathogenesis studies. *Curr. Opin. Microbiol.* **5**, 33–37 (2002).
22. Saenz, H.L. & Dehio, C. Signature-tagged mutagenesis: technical advances in a negative selection method for virulence gene identification. *Curr. Opin. Microbiol.* **8**, 612–619 (2005).
23. Shea, J.E., Santangelo, J.D. & Feldman, R.G. Signature-tagged mutagenesis in the identification of virulence genes in pathogens. *Curr. Opin. Microbiol.* **3**, 451–458 (2000).
24. MacCallum, D.M. & Odds, F.C. Temporal events in the intravenous challenge model for experimental *Candida albicans* infections in female mice. *Mycoses* **48**, 151–161 (2005).
25. Bastidas, R.J. & Heitman, J. Trimorphic stepping stones pave the way to fungal virulence. *Proc. Natl. Acad. Sci. USA* **106**, 351–352 (2009).
26. Uhl, M.A., Biery, M., Craig, N. & Johnson, A.D. Haploinsufficiency-based large-scale forward genetic analysis of filamentous growth in the diploid human fungal pathogen *C. albicans*. *EMBO J.* **22**, 2668–2678 (2003).
27. Rieg, G. *et al.* Unanticipated heterogeneity in growth rate and virulence among *Candida albicans* AAF1 null mutants. *Infect. Immun.* **67**, 3193–3198 (1999).
28. Leipelt, M. *et al.* Glucosylceramide synthases, a gene family responsible for the biosynthesis of glucosphingolipids in animals, plants, and fungi. *J. Biol. Chem.* **276**, 33621–33629 (2001).
29. Rittershaus, P.C. *et al.* Glucosylceramide synthase is an essential regulator of pathogenicity of *Cryptococcus neoformans*. *J. Clin. Invest.* **116**, 1651–1659 (2006).
30. Roemer, T. *et al.* Large-scale essential gene identification in *Candida albicans* and applications to antifungal drug discovery. *Mol. Microbiol.* **50**, 167–181 (2003).
31. Oura, T. & Kajiwara, S. Disruption of the sphingolipid $\Delta^8$-desaturase gene causes a delay in morphological changes in *Candida albicans*. *Microbiology* **154**, 3795–3803 (2008).
32. Arguelles, J.C., Rodriguez, T. & Alvarez-Peral, F.J. Trehalose hydrolysis is not required for human serum-induced dimorphic transition in *Candida albicans*: evidence from a tps1/tps1 mutant deficient in trehalose synthesis. *Res. Microbiol.* **150**, 521–529 (1999).
33. Cao, F. *et al.* The Flo8 transcription factor is essential for hyphal development and virulence in *Candida albicans*. *Mol. Biol. Cell* **17**, 295–307 (2006).
34. Donovan, M. *et al.* Virulence of a phosphoribosylaminoimidazole carboxylase-deficient *Candida albicans* strain in an immunosuppressed murine model of systemic candidiasis. *Infect. Immun.* **69**, 2542–2548 (2001).
35. Hornby, J.M. *et al.* Quorum sensing in the dimorphic fungus *Candida albicans* is mediated by farnesol. *Appl. Environ. Microbiol.* **67**, 2982–2992 (2001).
36. Huh, W.K., Kim, S.T., Kim, H., Jeong, G. & Kang, S.O. Deficiency of D-erythroascorbic acid attenuates hyphal growth and virulence of *Candida albicans*. *Infect. Immun.* **69**, 3939–3946 (2001).
37. Jia, N. *et al. Candida albicans* sterol C-14 reductase, encoded by the ERG24 gene, as a potential antifungal target site. *Antimicrob. Agents Chemother.* **46**, 947–957 (2002).
38. Kirsch, D.R. & Whitney, R.R. Pathogenicity of *Candida albicans* auxotrophic mutants in experimental infections. *Infect. Immun.* **59**, 3297–3300 (1991).
39. Mio, T., Kokado, M., Arisawa, M. & Yamada-Okabe, H. Reduced virulence of *Candida albicans* mutants lacking the GNA1 gene encoding glucosamine-6-phosphate acetyltransferase. *Microbiology* **146**, 1753–1758 (2000).
40. Pedreno, Y. *et al.* Disruption of the *Candida albicans* ATC1 gene encoding a cell-linked acid trehalase decreases hypha formation and infectivity without affecting resistance to oxidative stress. *Microbiology* **153**, 1372–1381 (2007).
41. Rocha, C.R. *et al.* Signaling through adenylyl cyclase is essential for hyphal growth and virulence in the pathogenic fungus *Candida albicans*. *Mol. Biol. Cell* **12**, 3631–3643 (2001).
42. Zaragoza, O., Blazquez, M.A. & Gancedo, C. Disruption of the *Candida albicans* TPS1 gene encoding trehalose-6-phosphate synthase impairs formation of hyphae and decreases infectivity. *J. Bacteriol.* **180**, 3809–3815 (1998).
43. Oura, T. & Kajiwara, S. *Candida albicans* sphingolipid C9-methyltransferase is involved in hyphal elongation.. *Microbiology* **156**, 1234–1243 (2010).

## ONLINE METHODS

**Analysis of published *C. albicans* mutants.** Published *C. albicans* 'features' and 'phenotypes' were downloaded from the *Candida* Genome Database website[3] on 6 July 2009. Mutants affecting virulence, morphogenesis and organ colonization were identified through searches for, respectively, the terms 'virulence'; 'filamentation' and 'hyphal'; and 'kidney', 'tissue' and 'organ'. Phenotypes were associated with homozygous knockout or transcriptional depletion mutants. Contradictory reports were omitted unless one report addressed a deficiency in the other. Because phenotypes can differ between model systems, virulence measurements were limited to time to illness in the mouse tail vein model of disseminated disease. Morphology phenotypes describe colony morphology on Spider medium; reports of defective germ tube formation were also included, as this phenotype reliably predicts defects on Spider medium (for example, see **Supplementary Table 6**).

***C. albicans* annotation database and target selection.** Predicted ORFs in *C. albicans* assembly 19 (Stanford Genome Technology Center)[19] were compared by BLAST to the genomes of *S. cerevisiae*, *S. pombe*, *Aspergillus nidulans*, *Magnaporthe grisea*, *Neurospora crassa* and the Nonredundant Protein Database. They were also analyzed with the SMART motif finding algorithm (European Molecular Biology Laboratory).

Targets were identified by multiple, partially overlapping criteria, as follows. (i) ORFs without homologs in *S. cerevisiae* and *S. pombe*, $n = 352$. These were similar to an ORF in the Nonredundant Protein Database with a chance probability $<10^{-6}$ and similar to ORFs in *S. cerevisiae* or *S. pombe* with $P > 10^{-2}$. (ii) ORFs without known homologs, $n = 1{,}374$. The *C. albicans* genome was compared by BLAST to genomes of numerous bacteria as well as *Arabidopsis thaliana, Caenorhabditis elegans, Drosophila melanogaster, Rattus norvegicus, Plasmodium falciparum, Homo sapiens, Saccharomyces bayanus, S. cerevisiae, Saccharomyces paradoxus, Saccharomyces mikatae, Saccharomyces castellii, S. pombe, Saccharomyces kluyveri,* and *Saccharomyces mikatae*. *C. albicans* ORFs that appeared unique after cluster analysis were selected. (iii) ORFs with motifs potentially related to virulence, $n = 1215$. These included ORFs with SMART motifs associated with immunity (for example, 'LRR'), pathogenesis (for example, 'pathogen', 'toxin'), transcription (for example, 'ZnF') signaling (for example, 'G alpha'), and cell surface and secreted proteins (for example, 'signal sequence'). (iv) ORFs encoding She3-associated mRNAs, $n = 51$ (ref. 20). (The She3-dependent RNA transport system localizes mRNAs to the tips of hyphal cells, and we were interested to know whether any of these is required for infectivity.) ORFs were eliminated if they appeared to be false ORFs (for example, a short ORF within a larger predicted ORF), were associated with transposons or were recognized homologs of essential genes in *S. cerevisiae*.

**Construction of the mutant library.** Primers used for mutant construction are listed in **Supplementary Table 8**. Fusion PCR was used to produce gene disruption fragments as described[14]. Two alleles of each target gene were sequentially disrupted in parental strain SN152 (**Supplementary Table 9**), with disruption fragments marked by *C. dubliniensis HIS1* and *C. maltosa LEU2*, respectively. $His^+$ $Leu^+$ transformants were screened by colony PCR for the four expected 5′ and 3′ gene disruption junctions and absence of the target ORF.

**Infectivity screen.** Procedures involving animals were approved by the UCSF Institutional Animal Care and Use Committee. Overnight cultures of 48 *C. albicans* mutants and the isogenic wild-type reference strain QMY23 (ref. 44; **Supplementary Table 9**) were diluted 1:20 into fresh YEPD medium and incubated for ~4 h at 30 °C. The log-phase cultures were pooled and washed twice with sterile normal saline, and ~$1.1 \times 10^6$ colony-forming units (CFUs) were injected by tail vein into 18- to 20-g female BALB/c mice. The inoculum was also plated onto Sabouraud medium (with ampicillin 50 μg/ml, gentamicin 15 μg/ml) and incubated at 30 °C. Mice were monitored closely and euthanized when morbid (defined as weight loss >15%, hunched posture, reduced activity). *C. albicans* was recovered by plating homogenates of mouse kidneys onto Sabouraud with ampicillin and gentamicin (antibiotics were included to suppress contaminating bacteria) and incubating them at 30 °C. Genomic DNA was prepared from yeast scraped off the plates[45].

Abundance of individual mutants in each pool was determined using real-time PCR and primers specific to each signature tag; the second, common primer hybridized to *C. dubliniensis HIS1*. Reactions included 200 ng of genomic DNA, 2 μM of each primer, 2 mM $MgCl_2$, 50 mM KCl, 10 mM Tris 8.3, 0.075× SYBR green (Sigma) and 2 units Taq polymerase. Wild-type *C. albicans* was quantified using the *C. dubliniensis HIS1* primer and primer SNO183 (which hybridizes upstream of *LEU2*, where *C. dubliniensis HIS1* is integrated in this strain). Cycling was performed in an Opticon device (Bio-Rad) using the following parameters: 92 °C for 6 min; followed by 40 cycles of 92 °C for 30 s, 52 °C for 45 s, 72 °C for 1 min; and finally 72 °C for 10 min.

Threshold cycle ($C_T$) values were converted to a linear scale by the following equation: linear value = $2^{-C_T}$. Experiments comparing 48 strains resulted in 48 values for the inoculum ($I$) and another 48 for the recovered pool ($R_{raw}$). $R_{raw}$ values were multiplied by median($I$)/median($R_{raw}$) to generate normalized $R$ values. Finally, $R/I$ was calculated for each mutant and the wild-type strain; these are presented in **Supplementary Tables 10** and **11**. Please see "Statistical analysis", below, for an explanation of statistical methods. Plots and calculations were based on median-centered data.

We used two methods to determine which mutants had significant infectivity defects. First, the $t$-test was used to compare the $\log_2(R/I)$ values of mutants to those of the wild-type data set, using a significance threshold of $P < 7.4 \times 10^{-5}$ (that is, $P < 0.05$, corrected for 674 comparisons) to identify mutants that were significantly different from the wild type. The $t$-test captured all of the infectivity mutants that were obviously depleted in **Supplementary Figure 2a** (for example, mutants 7, 9, 23, 29, 39 and 47) and had a very low false discovery rate, estimated at <0.04%. However, close inspection of the data revealed that this test—based on the average $\log_2(R/I)$ of each mutant—missed a class of mutants that had a mix of aberrant phenotypes. Certain mutants (for example, mutants 6, 17, 30, 31 and 41 in **Supplementary Fig. 2a**) were alternately increased or decreased in kidneys of different hosts, such that the mean $\log_2(R/I)$ ratios were normal even though individual $\log_2(R/I)$ ratios were well outside the mean. This phenomenon is reflected by mutants with high s.d. around the mean in **Supplementary Figure 2b** and may reflect the presence of multiple fitness deficits that are manifested stochastically in different infected mice. To include candidates with this phenotype, we devised a Boolean test for aberrant infectivity, defined as having a preponderance (>70%) of $R/I$ measurements that are abnormally high, low or both (<0.7, >2). When applied to $10^6$ random sets of six values from the wild-type data set, the false discovery rate of this test was 1.3%.

We compared the ability of the $t$-test with that of the Boolean test to discriminate among 57 previously published *C. albicans* mutants—all previously tested in a mouse tail vein model of disseminated disease—that were included in our library. We note that the published mutants are not a bona fide gold standard for assessing virulence because most were constructed using *URA3*, and ~30% of reported virulence defects are thought to result from deficient *URA3* expression rather than deletion of the target genes[9]. Despite these caveats, the Boolean test showed superior sensitivity and specificity, properly identifying 21 of 47 mutants previously reported as virulence-defective (**Supplementary Fig. 2a** and **Supplementary Table 12**) and only one of ten mutants previously reported as competent for virulence (in *GPA2* (ref. 46); **Supplementary Table 12**). The Boolean method also captured all four virulence-defective mutants that had been constructed without the use of *URA3* (refs. 47–50). Results of both statistical tests are presented in **Supplementary Table 5**.

**Morphogenesis screen.** Mutants and wild-type reference strain SN250 (**Supplementary Table 9**) were plated onto Spider agar at a density of 5–10 CFUs per plate. Colonies were photographed after 7 and 14 d of incubation at 30 °C. Morphology scores were based on colony appearance at 14 d, with consideration of appearance at 7 d in questionable cases.

Initial assignments were based on the first isolate of every mutant in the library. The center and periphery were scored separately on a scale from −3 (fewer central wrinkles or fewer peripheral filaments) to +3 (increased central wrinkles or peripheral filaments); wild-type morphology was defined as 0. Scoring was blind in that the mutants were numerically coded. One to three additional isolates (depending on availability) of mutants with abnormal phenotypes were analyzed by a different investigator. If an additional isolate had wild-type morphology, the mutant was scored as wild-type (with the presumption that the isolate with abnormal morphology had acquired an

additional mutation). Three mutants showed inconsistent abnormal morphologies and were scored as 'IC'; 18 mutants had severe growth defects on Spider and were not scored.

**Proliferation screen.** Two isolates of every mutant and the wild-type reference strain QMY23 (**Supplementary Table 9**) were diluted 1:30 into liquid synthetic complete medium in 96-well optically clear plates (Nunc) and incubated with shaking at 37 °C. Absorbance at 600-nm wavelength was monitored every 60–90 min for ~10 h with a plate reader (Molecular Devices). Doubling times were calculated from the peak rate of growth between any two time points, and the average of the doubling times of the two isolates was reported.

**Virulence analysis.** Cultures of SN285 (*hsx11*, also called *orf19.4592*), SN255 (*orf19.260*), SN290 (*orf19.4831*), SN297 (*het1*, also called *orf19.6327*) and the isogenic reference strain SN250 (wild type) were prepared in liquid synthetic complete medium and grown for 2 d at room temperature. Mid- to late log-phase cells were washed twice with sterile normal saline and counted using a hemacytometer. We used ~$4.5 \times 10^5$ CFUs of each strain to infect ten female 18- to 20-g BALB/c mice by tail vein injection. Mice were monitored twice daily and euthanized when morbid, as judged by criteria described above.

**Complementation analysis.** Plasmids for complementation were constructed using PCR and homologous recombination in *S. cerevisiae*[51] The plasmids (pSN103, *HSX11*; pSN104, *orf19.260*; pSN101, *orf19.4831*; and pSN102, *HET1*) were engineered to sandwich a wild-type copy of each test gene and *C. dubliniensis ARG4* between sequences upstream and downstream of the *C. albicans LEU2* gene. Inserts were sequenced to verify the absence of mutations. Disruption fragments were liberated by digestion with PmeI and introduced into the appropriate mutants (SN285, *hsx11*; SN255, *orf19.260*; SN290, *orf19.4831*; and SN297, *het1*; **Supplementary Table 9**). An additional plasmid (pSN105) was created with just the *C. dubliniensis ARG4* gene and *C. albicans LEU2* flanking sequences; this was used to create auxotrophic marker-matched versions of the homozygous deletion mutants. Strain designations appear in **Supplementary Table 9**.

We performed two-strain infectivity experiments by infecting female 18- to 20-g BALB/c mice by tail vein with ~$8 \times 10^5$ CFUs of a 1:1 pool of each mutant and its gene add-back comparitor. Five mice were infected with each pool, and the abundance of strains in the inoculum compared with the infected kidneys was determined by real-time PCR. Standard curves were created using dilutions of genomic DNA from individual strains.

**Lipid extraction and mass spectrometry.** We incubated 100-ml cultures (in YEPD) of SN425 (wild type), SN442 (*hsx11*), SN444 (*orf19.260*), SN446 (*orf19.4831*) and SN448 (*het1*) with shaking at 30 °C overnight. Cell pellets were frozen in liquid nitrogen and lyophilized overnight. We extracted 10 mg of each lyophilized cell pellet for 48 h with chloroform/methanol (2:1, vol/vol). The extract was mixed with 0.2 volumes of 0.34% (wt/vol) $MgCl_2$. The aqueous layer was discarded and the organic phase mixed with 5 ml of 2 N KCl/methanol (4:1, vol/vol). The aqueous layer was discarded, and the organic phase was washed six times with 5 ml artificial upper phase (chloroform/methanol/water, 3:48:47, vol/vol/vol). The organic phase was dried under $N_2$ vapor and then dissolved in 500 µl CHCl3/methanol (2:1, vol/vol). We then applied 30 µl of each preparation to a silica gel 60 high-performance TLC plate (Whatman); soybean glucosylceramide (Polar Lipids) was included as a mobility marker, and chloroform/methanol (90:25, vol/vol) was the mobile phase. Glucosylceramide was visualized with anthrone reagent (in concentrated $H_2SO_4$) and baking.

For mass spectrometry, 450 µl of each lipid preparation was resolved on silica gel 60 plates with divided lanes (J.T. Baker). For each test sample, the region of TLC resin that aligned with soybean glucosylceramide was scraped into 5 ml $CHCl_3$/methanol (2:1). Samples were dried under $N_2$ vapor, solubilized in 5 ml of 0.1 N NaOH/methanol, then incubated at 55 °C for 1 h. Mixtures were acidified with 1 ml of 1 N HCl/methanol, mixed with five drops of water and 5 ml hexane, and centrifuged. The lower phase was removed and evaporated under $N_2$ vapor, and the lipid residue dissolved in 2 ml theoretical lower phase (chloroform/methanol/water 86:14:1). We added 8 ml of theoretical upper phase (chloroform/methanol/water 3:48:47) to create a Folch partition. After centrifugation, the lower phase was recovered, evaporated to dryness and solubilized in 200 µl chloroform/methanol (2:1, vol/vol). Samples were mixed 1:1 with 10 mg/ml 2,5-dihydroxybenzoic acid (in chloroform/methanol 2:1, vol/vol) and assessed with a Voyager-DE MALDI-TOF mass spectrometer (Applied Biosystems) in reflector mode.

**Statistical analysis.** The hypergeometric distribution was used to evaluate the significance of associations of virulence with morphogenesis among published mutants, of virulence with kidney burden among published mutants, of previously published virulence defects with infectivity defects in our reconstructed mutants, and of all combinations of defects in infectivity, morphogenesis and proliferation in our screens. $P < 0.05$ was considered significant.

The *t*-test (two-tailed, comparison of unpaired samples) was used to evaluate the significance of the following differences: (i) $\log_2(R/I)$ values of mutants in the infectivity screen versus the wild-type reference strain, with $P < 0.05/674$ considered significant (Bonferroni correction for multiple comparisons); and (ii) $\log_2(R/I)$ values of the glucosylceramide mutants and their cognate gene add-back strains in the paired infectivity experiments, with $P < 0.05$ considered significant.

The logrank test was used to compare the time-to-illness curves of mice infected with wild-type *C. albicans* versus mutants in the glucosylceramide pathway; $P < 0.05$ was considered significant.

False discovery rates for the *t*-test and the Boolean methods of identifying infectivity mutants were determined by applying these tests to 1,000 random sets of six values from the wild-type data set 1,000 times and calculating the median percentage identified as positive.

**URL.** Candida Genome Database, http://www.candidagenome.org/.


44. Mitrovich, Q.M., Tuch, B.B., Guthrie, C. & Johnson, A.D. Computational and experimental approaches double the number of known introns in the pathogenic yeast *Candida albicans*. *Genome Res.* **17**, 492–502 (2007).
45. Ausubel, F.M. *et al.* (eds.) *Current Protocols in Molecular Biology* (John Wiley & Sons, New York, 1990).
46. Miwa, T. *et al.* Gpr1, a putative G-protein-coupled receptor, regulates morphogenesis and hypha formation in the pathogenic fungus *Candida albicans*. *Eukaryot. Cell* **3**, 919–931 (2004).
47. Badrane, H. *et al.* The *Candida albicans* phosphatase Inp51p interacts with the EH domain protein Irs4p, regulates phosphatidylinositol-4,5-bisphosphate levels and influences hyphal formation, the cell integrity pathway and virulence. *Microbiology* **154**, 3296–3308 (2008).
48. Kelly, M.T. *et al.* The *Candida albicans* CaACE2 gene affects morphogenesis, adherence and virulence. *Mol. Microbiol.* **53**, 969–983 (2004).
49. Warenda, A.J., Kauffman, S., Sherrill, T.P., Becker, J.M. & Konopka, J.B. *Candida albicans* septin mutants are defective for invasive growth and virulence. *Infect. Immun.* **71**, 4045–4051 (2003).
50. Wu, W., Lockhart, S.R., Pujol, C., Srikantha, T. & Soll, D.R. Heterozygosity of genes on the sex chromosome regulates *Candida albicans* virulence. *Mol. Microbiol.* **64**, 1587–1604 (2007).
51. Oldenburg, K.R., Vo, K.T., Michaelis, S. & Paddon, C. Recombination-mediated PCR-directed plasmid construction in vivo in yeast. *Nucleic Acids Res.* **25**, 451–452 (1997).





doi:10.1038/ng.605