Published in final edited form as:

Fungal Genet Biol. 2020 November ; 144: 103464. doi:10.1016/j.fgb.2020.103464.

An intergenic “safe haven” region in *Cryptococcus neoformans* serotype D genomes

Yumeng Fan¹, Xiaorong Lin¹

¹Department of Microbiology, University of Georgia, Athens, Georgia 30602

Abstract

Cryptococcus neoformans is an opportunistic human fungal pathogen and serves as a model organism for studies of eukaryotic microbiology and microbial pathogenesis. *C. neoformans* species complex is classified into serotype A, serotype D, and AD hybrids, which are currently considered different subspecies. Different serotype strains display varied phenotypes, virulence, and gene regulation. Genetic investigation of important pathways is often performed in both serotype A and D reference strains in order to identify diversification or conservation of the interrogated signaling network. Many genetic tools have been developed for *C. neoformans* serotype A reference strain H99, including the gene free “safe haven” (*SH*) regions for DNA integration identified based on genomic features. However, no such a genomic safe haven region has been identified in serotype D strains. Here, capitalizing on the available genomic, transcriptomic, and chromatin data, we identified an intergenic region named as *SH3* for the serotype D reference strains JEC21 and XL280. We also designed a sgRNA and a vector facilitating any alien gene integration into *SH3* through a CRISPR-Cas9 system. We found that gene inserted in this region complemented the corresponding gene deletion mutant. Fluorescent reporter gene inserted in *SH3* can also express efficiently. Insertion in *SH3* itself did not alter the expression of adjacent genes and did not affect the growth or mating of *C. neoformans*. Thus, *SH3* provides a resource for genetic manipulations in serotype D strains and will facilitate comparative analyses of gene functions in this species complex. In addition, the incorporation of the multi-omic data in our selection of the safe haven region could help similar studies in other organisms.

Keywords

Cryptococcus neoformans; genome editing; safe haven; ectopic integration; complementation; CRISPR-Cas9

*correspondence: xiaorong.lin@uga.edu.

Y.F. and X.L. conceived and designed the experiments. Y.F. performed the experiments. Y.F. and X.L. analyzed the data. X.L. contributed reagents, materials, and analysis tools. Y.F. wrote the paper. X.L. edited the paper.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Cryptococcus neoformans is an opportunistic human fungal pathogen responsible for hundreds of thousands of deaths annually (Perfect and Bicanic, 2015; Rajasingham et al., 2017). This organism was first discovered in 1890s (Otto, 1894; Sanfelice, 1895), but the biology of this fungus was poorly understood at the molecular level until targeted mutagenesis was developed for *C. neoformans* about three decades ago (Edman and Kwon-Chung, 1990; Toffaletti et al., 1993). Gene deletion (loss of function) is the classic and the most common approach to study gene function. To verify the gene function inferred from the knockout mutants, gene complementation is usually necessary (Falkow, 1988). Gene overexpression (often gain-of-function) is another common approach to investigate gene function. Gene complementation, gene overexpression, and construction of a fluorescent reporter stain all require the introduction of a DNA fragment into the genome. This is sometimes achieved through knock-in at the native locus or more often through ectopic integration. Native knock-in is not necessarily risk-free, as introducing a gene copy (often with a selection marker) back to its native locus may change the local chromatin structure and affect expression of neighboring genes. In addition, some cryptic cis-regulatory elements may be disrupted, leading to unexpected phenotypes. For this reason and for technical simplicity, random ectopic insertion is more commonly used.

However, ectopic integration of gene of interest has its own drawbacks. The expression level of randomly integrated gene varies because of the position effect. Ectopic integration could also be mutagenic due to unintended gene disruption, and the risk is particularly high in organisms with highly compacted genomes like *C. neoformans* (Arras et al., 2015; Tenney et al., 2004). Thus, integration in a designated gene free region in the genome is preferred. Several gene free regions were named as “Safe Harbor” sites for alien gene integration in human cellular studies (Pellenz et al., 2018; Sadelain et al., 2011). Similarly, “safe haven” sites have recently been identified in the fungus *Aspergillus fumigatus* reference strains CEA10/Af293 and *C. neoformans* serotype A reference strain H99 (Arras et al., 2015; Pham et al., 2020; Upadhyay et al., 2017).

As a species complex, *Cryptococcus neoformans* encompasses serotype A, D, and their hybrid based on their immunologic properties of capsular polysaccharides (Evans, 1950; Levitz, 1991; Wilson et al., 1968), which correspond to the molecular type VNI, VNIV, and VNIII (Desjardins et al., 2017; Meyer et al., 2003). The two serotypes differ genetically and phenotypically. The serotype A reference strain H99 is one of the most aggressive strains in causing lethal infection in animal models (Barchiesi et al., 2005; Lin et al., 2008; Nielsen et al., 2005; Toffaletti et al., 2004; Zhao et al., 2019). The serotype D reference strains JEC21 and XL280 are lower in virulence, with XL280 displaying intermediate level of virulence between H99 and JEC21 (Barchiesi et al., 2005; Zhai et al., 2013). These strains also differ in their ability to undergo mating and morphological changes, with JEC21/XL280 superior in mating and filamentation. In addition, there is evidence of diversification in genetic regulation between serotype A and serotype D. For example, mitogen-activated protein kinase Hog1 in JEC21 is phosphorylated rapidly after exposure to 1M NaCl. By contrast, the Hog1 in H99 strain is dephosphorylated under the same condition (Bahn et al., 2005). Two protein kinase A (PKA) catalytic subunits Pka1 and Pka2 also played diverged function in

the serotype A and D reference strains. In H99, Pka1 plays its roles in melanin and capsule production and mating while Pka2 does not. By contrast, Pka2 control these three process in JEC21 (Hicks and Heitman, 2007). Consequently, all reference strains are commonly used for cryptococcal research and many important pathways have been examined in both serotype A and D reference strains to identify functional diversification or conservation (Bahn et al., 2005; Chacko et al., 2015; Xu et al., 2017). Many genetic tools have been developed for these reference strains, including congenic pairs and annotated genome sequences (Janbon et al., 2014; Loftus et al., 2005; Nielsen et al., 2003; Zhai et al., 2013).

In H99, two safe haven regions, *SH1* and *SH2*, have been identified based on genomic organization of intergenic regions (Arras et al., 2015; Upadhyay et al., 2017). They are being widely adopted by researchers in their genetic investigations in H99 (Arras et al., 2016; Lin et al., 2020). Safe haven gains more popularity with the advent of CRISPR-Cas9 genome editing systems in *Cryptococcus* (Arras et al., 2016; Fan and Lin, 2018; Lin et al., 2019; Wang et al., 2016). Cas9 and sgRNA create a double stranded DNA break at the specific safe haven region, which facilitates integrating the introduced DNA to the designated site (Fan and Lin, 2018).

The goal of this study is to identify and characterize a safe haven site in the serotype D reference strains XL280 and JEC21. Here we took advantage of not only the genomic data, but also recent transcriptome and chromatin data generated for these strains to identify gene free and unsuppressed regions as potential safe haven sites. We identified one gene-free region *SH3* that met our criteria. We demonstrated that genes integrated into *SH3* are consistently and reproducibly expressed, and the integration of foreign DNA into *SH3* region itself does not affect the expression of the neighboring genes. More importantly, DNA integration at *SH* does not alter cryptococcal phenotypes examined, including growth, thermotolerance, capsulation, or mating. Thus, the safe haven *SH3* provides a new tool for *Cryptococcus* serotype D strains and will facilitate comparative analyses of gene functions in this species complex.

2. Materials and methods

2.1 Strains and growth conditions

Cryptococcus neoformans strains generated and used in this study are listed in Supplemental Table 1. Cells were stocked in 15% glycerol at -80°C . Cells were freshly cultured in YPD (1% yeast extract, 2% BactoPeptone and 2% dextrose) at 30°C unless stated otherwise. For selection, G418 (200 $\mu\text{g/ml}$) or HYG (hygromycin, 250 $\mu\text{g/ml}$) was supplemented in YPD medium.

2.2 Construction of plasmids

Plasmids generated and used in this study are listed in Supplemental Table 1 and all primers used in this study are listed in Supplemental Table 2. To construct the mCherry expression cassette without homologous arms, the coding sequence of mCherry was first amplified from plasmid pYZ2 (Zhao et al., 2018) using primers carrying *FseI* and *PacI* restriction sites. Purified PCR product was then digested with *FseI* and *PacI*. The plasmid pXL1-PHD11-

mcherry (Meng et al., 2018), which served as the backbone vector, was also digested with *FseI* and *PacI* and then purified. After T4 ligation of these two digested products, the plasmid pYF10 carrying P_{GPD}-mCherry was generated.

To construct vector pYF49, which carries the *SH3* homologous arms flanking the mCherry expression construct, we amplified the mCherry construct from pYF10, the HYG marker from pPZP-HYG2 (Walton et al., 2005), and two homologous arms of *SH3* from XL280 genomic DNA separately. These four PCR purified fragments, together with *Bam*HI digested pUC19 plasmid as backbone, were then assembled (Figure 2F) using NEBuilder® HiFi DNA Assembly Master Mix according the manufacture's instruction. Primers used in the assembly were designed via NEBuilder Assembly tool (<https://nebuilder.neb.com/#/>) and are listed in Supplemental Table 2.

2.3 Generation of mCherry reporter strains

To generate XL280 strains carrying the mCherry gene in the *SH3* region, we PCR amplified the mCherry expression construct along with a NEO drug marker using M13F and M13R primers from plasmid pYF10. The amplified mCherry-NEO construct was then used as the donor DNA for cryptococcal transformation (Figure 2A). Here we used the Transient CRISPR-Cas9 Coupled with Electroporation (TRACE) System that we developed recently (Fan and Lin, 2018). In this system, three independent DNA fragments donor DNA (with drug selection), sgRNA expression cassette, and Cas9 expression cassette are introduced to recipient strain through electroporation. Here, 2 µg of the purified donor DNA, 100 ng sgRNA construct, and 150 ng of Cas9 expression cassette were mixed and introduced into XL280 via TRACE transformation as we described previously (Fan and Lin, 2018). Transformants were selected on YPD +NEO plates. Stable transformants verified by passages on non-selective media (Lin et al., 2020) were then further confirmed by diagnostic PCR using primers shown in Figure 2C. Transformants with correct integration into *SH3* in the forward orientation were expected to yield a band of ~1.8kb with primers Linlab4950/Linlab4483 and these positive transformants were named F-mCherry isolates. Transformants with correct integration into *SH3* in the reverse orientation were expected to yield a band of ~1.5kb with primers Linlab4951/Linlab4483 and these transformants were named R-mCherry isolates (Supplemental Figure 2A).

Similarly, we generated H-mCherry strains by introducing a donor DNA where the mCherry expression cassette was flanked by the homologous arms of *SH3*. This donor DNA was amplified from plasmid pYF49 via PCR using primers Linlab4961/Linlab4966. The proper integration of the donor DNA in the *SH3* site in the stable transformants was confirmed by diagnostic PCR using primers Linlab4903 and Linlab4983. The H-mCherry strains yielded the expected band of ~3.4 kb (Supplemental Figure 2A).

2.4 Fluorescence microscopy

XL280α and SH3-mCherry reporter strains were cultured in YPD liquid medium at 30°C overnight. Cells were collected by centrifugation, washed twice with ddH₂O, and resuspended in ddH₂O. Then they were observed under Zeiss Imager M2 with filter 63 HR

Red Fluores and DIC (Carl Zeiss). Images were captured with an AxioCam 506 camera and analyzed with Zeiss ZEN 3.0 software.

To calculate the relative fluorescence intensity, arithmetic mean intensities of mCherry of each cell and its background were collected for 50 randomly selected cells of each genotype (25 cells per strain) through the Zeiss ZEN 3.0 software. Using the fluorescent intensity of wild-type XL280 as the reference, the relative fluorescence fold change was plotted and the mean with standard deviation (SD) were calculated with GraphPad Prism 8. One-Way ANOVA was used for multiple comparison and *P* values .05 were considered significant.

2.5 Growth and phenotypic assays

In vitro phenotypic assays were performed as we reported previously (Lin et al., 2007; Lin et al., 2009). Briefly, cryptococcal strains were cultured in YPD liquid medium at 30°C with shaking overnight. Cells were then washed twice with sterile water and adjusted to the same cell optical density ($OD_{600}=3$). Then 3 μ l of 10X serial dilutions of each cell suspension was spotted onto various agar media for phenotypical analyses as described previously (Lin et al., 2008). To test melanization, yeast cells were spotted onto L-3,4-dihydroxyphenylalanine (L-Dopa) medium and incubated at 30°C or 37°C in the dark for 4 days. To examine the production of capsule, cells were spotted onto RPMI medium and incubated at 37 °C for 3 days. Capsule was visualized microscopically as a halo surrounding the yeast cell by India ink exclusion. For the thermotolerance test, yeast cells were spotted onto YPD medium and incubated at 22°C, 30°C, or 37°C for 2 days. For self-filamentation (unisex) and bisexual mating, cells were cultured on V8 juice agar (pH=7.0) at 22°C for 96 hours before observation.

2.6 Quantitative real-time PCR

Cells of the wild type strain XL280 and the indicated strains were cultured in YPD liquid medium at 30°C for 16 hr. Three biologically replicates were used. Total RNA samples were extracted using the PureLink® RNA Mini Kit (Life Technology) and their integrity was confirmed with RNA gel electrophoresis. Each RNA sample (8 μ g) was treated with DNase (TURBO DNA-free™) according to the manufacture's instruction. DNase-treated RNA samples were used as templates in the synthesis of the first strand cDNA using GoScript™ Reverse Transcription System (Promega) per the instruction from the manufacture. The resulting cDNA products were diluted to 8 ng/ μ l and 2 μ l was used as template for quantitative real-time PCR (qRT-PCR). qRT-PCR was performed using the SYBR FAST qPCR master mix (KAPA Biosystems, Wilmington, MA) on a realplex² instrument (Eppendorf). The house-keeping gene *TEF1* was used as the internal control for each sample to normalize the gene transcript level as we described previously (Chacko et al., 2015). The relative levels of transcripts were quantified using the C_t method as we described previously (Wang et al., 2014; Wang et al., 2012). We used One-Way ANOVA for multiple comparison and *P* values .05 were considered statistically significant. Primers for qRT-PCR are included in the Supplemental Table 2.

2.7 *Galleria mellonella* (great wax moth) infection

Galleria mellonella in their final instar larval stage were purchased from Snackworms (Ohio). *C. neoformans* strains were grown overnight in YPD liquid medium. Cells were collected by centrifugation, washed three times with PBS buffer, and resuspended in PBS buffer to the final concentration of 2×10^8 cells/ml. Cell concentration was confirmed by counting the cells using a hemocytometer. Five μ l of the cell suspension (1×10^6 cells) were injected into larval hemocoel through the second left proleg via a 10 μ l Hamilton syringe. The control group was injected with 5 μ l of PBS buffer. After injection, the larvae were incubated in dark at 37°C and monitored daily. Two independent repeats were performed. The Kaplan-Meier survival curve was plotted using GraphPad Prism 8. The Log-rank (Mantel-Cox) test was used for statistical analysis.

3. Results

3.1 Identification of a putative safe haven region *SH3* in the serotype D strains JEC21 and XL280

Serotype D reference strain JEC21 was generated in the 1990s (Kwon-Chung et al., 1992) and is the first sequenced and annotated *C. neoformans* strain (Loftus et al., 2005). Another serotype D reference strain XL280, which has superior ability to undergo yeast-hypha transition and much enhanced virulence compared to JEC21 (Zhai et al., 2013), was generated in 2006 and sequenced in 2013 (Lin et al., 2006; Ni et al., 2013). Both strains have congenic pair strains for classic genetic analyses (Kwon-Chung et al., 1992; Zhai et al., 2013) and are widely used as references in cryptococcal research (Bahn et al., 2005; Lin et al., 2019; Liu et al., 2018; Meng et al., 2018; Wang et al., 2016; Zhao et al., 2018). In this study, we sought to identify putative safe haven region(s) conserved in XL280 and JEC21 to provide a research tool for these *C. neoformans* serotype D isolates.

The two safe haven regions, *SH1* and *SH2*, characterized in the serotype A reference strain H99 are both located at large intergenic regions to avoid disruption of neighboring genes (Arras et al., 2015; Upadhyaya et al., 2017). To identify a safe haven region in the serotype D reference strains, we first examined the homologous regions of H99 *SH1* and *SH2* in JEC21 or XL280. *SH1* locates between genes *CNAG_00777* and *CNAG_00778* (Arras et al., 2015) in H99 with the intergenic region being 1544 bp. In JEC21, the homologous genes of *SH1* flanking genes, *CNA07550* and *CNA07560*, are 1912bp apart. In XL280, however, the distance between these two homologous genes is only 20 bp (*CNXL_007860* and *CNXL_007870*). *SH2* in H99 is surrounded by genes *CNAG_01561* and *CNAG_01562* (Upadhyaya et al., 2017). However, this synteny of the homologs *CNXL_021490* and *CNXL_015620* does not exist in either JEC21 or XL280 genomes and these two genes are 160 kb apart in these serotype D genomes instead. Thus, there is no good equivalent of *SH1* or *SH2* conserved in the serotype D reference strains. This is disappointing but not unexpected given that serotype A and D diverged from a common ancestor about 20 million years ago (Xu et al., 2000).

Therefore, we decided to screen the serotype D reference genome to identify potential safe haven regions *de novo* using the following criteria (Figure 1A). We focused only on

intergenic regions of non-overlapping genes from the XL280 genome (Ni et al., 2013; Zhai et al., 2013), a strategy also used in searching for safe haven regions in H99 (Arras et al., 2015). Based on the orientation of the two neighboring genes, the intergenic regions between non-overlapping genes were classified into convergent (tail-tail), tandem (head-tail), or divergent (head to head) (Figure 1A).

Although the promoter and terminator regions are not defined in *Cryptococcus* genome, a region of 1 to 1.5 kb prior to an open reading frame (ORF) is considered enough to include a typical promoter and a region of 0.2 to 0.5 kb after a stop codon is usually sufficient to include a terminator. For the divergent group (head to head), we only considered intergenic regions that are greater than 2.5 kb to avoid potential disruption of promoters of the two neighboring genes. Three divergent intergenic regions meet this criterium. For the tandem intergenic group (head-tail), we decided to only consider regions larger than 2 kb as they should harbor one promoter and one terminator of the neighboring genes. Ten candidate loci meet this criterium. Likewise, for the intergenic regions of convergent genes (tail-tail), we only considered those greater than 1 kb and fifteen candidates meet this bar (Figure 1A).

In addition to the distance from potential important elements like promoters and terminators, a safe haven site should also be accessible to the transcription machineries to facilitate the expression of an inserted gene. Accessibility of the local chromatin can greatly influence gene expression as genes inserted in a heterochromatin region (e.g. the centromere region) will be suppressed for expression. Such position effects are often reflected in the wide variation in the expression level of an ectopically integrated gene driven by the same promoter among different transformants. Hence, we decided to exclude candidates bordered by either extremely highly expressed or lowly expressed genes. To this end, we examined the transcript levels of the flanking genes of the 28 candidate intergenic regions based on our previously published RNA-seq data of XL280 cultured in YPD liquid medium that favors yeast growth and on V8 juice agar medium that favors filamentous growth (Chacko et al., 2015; Gyawali et al., 2017). The medium Fragments Per Kilobase Million (FPKM) values of all genes encoded in the genome in both YPD medium and V8 medium are around 36. The constitutively highly expressed house-keeping gene *TEF1* has FPKM values ranging from 1000 to 2200 under these conditions. Thus, we decided to exclude regions where the FPKM values of either the upstream or the downstream gene are consistently lower than 30 or higher than 2200. 12 candidate intergenic regions (4 convergent, 6 tandem, and 2 divergent) meet this criterion (Figure 1A).

To ensure that the putative safe haven regions identified in XL280 can also work for the other serotype D reference strain JEC21, we then examined if the 12 candidate regions are also conserved in the JEC21 genome. This is important as gene and chromosome arrangements in JEC21 and XL280 are not identical (Zhai et al., 2013). Indeed, 6 candidate intergenic regions either have no annotated homologous sequences or the neighboring genes have different chromosomal arrangements in JEC21. Two other candidates are found to reside between a protein-coding gene and a non-coding RNA gene in JEC21 and thus are also excluded from further analyses. Eventually, four candidate safe haven regions meet all of our criteria (Supplemental Table 3). We decided to prioritize on the two intergenic regions with convergent (tail-tail) flanking genes to eliminate the risk of disrupting any promoter of

neighbor genes with insertions. The intergenic region between *CNXL_025530* and *CNXL_025540* (*CND03070* & *CND03080* in JEC21) on chromosome IV is larger compared to the region between *CNXL_053410* & *CNXL_053420*. After excluding the 500 bp near the 3' ends of both neighboring genes, we decided to focus on the 291 bp long region between *CNXL_025530* and *CNXL_025540* (Figure 1B) and examine its potential as a safe haven site. We named it *SH3* (Figure 1B).

According to our RNA-Seq data of XL280 cultures collected from YPD (16h) and V8 media (16h and 72h) (Chacko et al., 2015), there are no transcripts made from the *SH3* region (Figure 1C). This supports the idea that the *SH3* region is a *bona fide* gene free region. We also decided to check the accessibility of the chromatin in this region using our recently published data of Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) for XL280 (Lin et al., 2019). ATAC-seq is a novel tool for detecting chromatin accessibility (Buenrostro et al., 2015), which often reveal promoter regions of actively expressed genes. For example, the translation elongation factor gene *TEF1* (*CNXL_065470*) is highly expressed in YPD liquid with a FPKM over 15000. As expected, the ATAC-seq data showed enriched reads in the putative promoter region of *TEF1* (Supplemental Figure 1). In comparison, the read depth in *SH3* was slightly lower than the ORF regions of its neighbor genes under both conditions (Figure 1D), suggesting that there is indeed no unannotated promoter region or transcription in *SH3*.

3.2 *ADE2* inserted into the *SH3* region can complement the *ade2* phenotype

To test if a gene inserted in *SH3* can be efficiently expressed, we decided to introduce a wild-type copy of the *ADE2* gene into *SH3* of an *ade2* mutant recipient (YF195). Deleting phosphoribosylaminoimidazole carboxylase gene *ADE2* results in the accumulation of red pigments, which changes colony color from white to red. Thus, *ade2* mutants yield red colonies. If the *ADE2* gene inserted into *SH3* of the *ade2* mutant can be expressed properly, we expect that it will change the colony color back to white. For this purpose, we designed a sgRNA targeting the *SH3* region (Figure 1B) using our previously reported method (Fan and Lin, 2018; Lin et al., 2020). We also constructed an *ADE2* expression cassette carrying the NEO (G418) drug marker as the donor DNA (~5.4 kb). As an alien DNA fragment can be inserted into a designated target site without any homologous arms in TRACE transformation (Fan and Lin, 2018), the *ADE2* expression cassette was designed without *SH3* homologous arms (Figure 2A). We then transformed the *ade2* mutant with the donor DNA *ADE2-NEO* mixed with the sgRNA and Cas9 using TRACE. The majority of the transformants turned the originally small red *ade2* colonies back to the larger wild-type white colonies (Figure 2B). This result indicated that the *ADE2* gene integrated at *SH3* was expressed and functional.

3.3 The mCherry expression construct integrated at *SH3* yielded consistent fluorescent signal irrespective of the orientation of integration

As the integration and the expression of *ADE2* could provide growth advantage to the *ade2* mutant, it is possible that most of our transformants had *ADE2* expressed at *SH3* due to this selective pressure. To determine if a gene expression construct that offers no growth benefit to the organism can be expressed when inserted at *SH3*, we decided to transform

Cryptococcus with a fluorescent protein marker. Fluorescent reporter strains are routinely used to investigate biological and pathogenic process of *C. neoformans*. Previously, a reporter strain with red fluorescent mCherry gene integrated into *SH2* was used to characterize the impact of DNA integration at that designated site in the serotype A strain H99 (Upadhyaya et al., 2017). Likewise, we decided to integrate an expression construct carrying the mCherry gene and a HYG selective marker into *SH3* of XL280 and examine the effect of DNA integration at *SH3*.

We first constructed a plasmid pYF10, carrying a *GPD1* promoter driven mCherry and a HYG selective marker. This cassette was then amplified by PCR and introduced into XL280 through TRACE with the sgRNA targeting *SH3*. Lacking homologous arms to *SH3*, this 5.2 kb construct could be inserted into the *SH3* site either in the forward (F-mCherry) or in the reverse direction (R-mCherry) (Figure 2C). Indeed, we found transformants with the construct integrated in either direction by diagnostic PCR using one primer inside of the expression construct and one primer outside of *SH3* (Figure 2C, Supplemental Figure 2A). We then examine the expression of mCherry gene by observing the red fluorescence signal microscopically. As expected, randomly selected and confirmed transformants (Supplemental Figure 2A) showed red fluorescent signal in their cytosol (Figure 2D). To evaluate the effect of the orientation of insertion into *SH3*, we then compared the fluorescent intensity of F-mCherry and R-mCherry isolates. As shown in Figure 2E, both the F-mCherry and the R-Cherry isolates showed strong fluorescent signal compared to the wild type control and there was no significant difference between them. This result indicates the expression level of mCherry integrated at *SH3* is similar regardless of the orientation of integration.

Sometimes it is desirable to control the direction of integration. To this end, we used NEBuilder Assembly and generated a vector pYF49 in which the mCherry together with the HYG marker was surrounded by the 5' and 3' homologous arms of *SH3* (Figure 2F). The whole construct including the *SH3* arms (~6.3kb) were amplified via PCR and transformed into XL280 using TRACE with the *SH3* gRNA as described earlier. To make the plasmid pYF49 versatile for integrating different constructs into *SH3*, we included *FseI*/*PacI* cloning sites so that an expression cassette of other genes of interests can be easily made through restriction enzyme digestion and ligation (Figure 2F). A *NoI* cutting site which locates before the *GPD1* promoter can be used to swap the promoter easily. Thus, this vector serves as a good backbone for gene integration in *SH3*.

3.4 DNA insertion in *SH3* does not affect the expression of its neighboring genes

Although our safe haven selection criteria help minimize the chance of disrupting the neighboring genes (Figure 1A), the effect of inserting a large DNA fragment into *SH3* on the neighboring genes needs to be tested experimentally. To that end, we examined the transcript level of *CNXL_025530* and *CNXL_025540* (*CND03070* and *CND03080* in JEC21) in the presence or absence of DNA insertion in *SH3*. We included wild type XL280 and the F-mCherry and R-mCherry isolates that have 5.2 kb foreign DNA inserted in *SH3*. We extracted RNA samples from these strains cultured in liquid YPD for 16 hours and measured the relative transcript level of these two flanking genes via qRTP-CR. As shown in Figure

3A, all five mCherry reporter strains tested showed transcript levels of *CNXL_025530* and *CNXL_025540* comparable to those in the wild-type XL280. Therefore, DNA insertion into *SH3* itself, regardless of the direction of integration, does not alter the transcript level of the neighboring genes under the condition tested.

3.5 Insertion in *SH3* itself has no significant impact on growth or other phenotypes examined

For an ideal safe haven site, we expect that DNA insertion in this region itself should not alter cryptococcal phenotype. To test the impact of DNA insertion in *SH3* on *Cryptococcus*, we performed *in vitro* phenotypic assays. As an environmental opportunistic human pathogen, *C. neoformans* possesses several classic virulence traits that enable it to survive in the host, including thermo-tolerance, melanization, and capsulation. First, we tested the temperature sensitivity of the wild type strain and the mCherry reporter strains (F-mCherry and R-mCherry). All the strains cultured on rich YPD medium grew similarly well at 22°C, 30°C, and 37°C (Figure 3B), indicating that DNA insertion into *SH3* does not affect cryptococcal thermotolerance or growth in the rich media. The growth of these strains was also comparable on the minimal YNB medium (Figure 3B), indicating that DNA insertion at *SH3* does not cause auxotrophy. Likewise, no apparent difference in melanization or capsulation was observed when cells were cultured on L-Dopa or RPMI medium (Figure 3C). Similarly, H-mCherry strains also showed no difference in phenotype under all the tested conditions comparing to the wild-type XL280 (Supplemental Figure 2B).

Since XL280 strain is commonly used in studying mating and yeast-to-hyphal switch, we also tested these strains for their ability to undergo self-filamentation and α - α bisexual mating. For self-filamentation, these α strains were dropped onto solid V8 juice medium (pH=7) individually. For α - α bisexual mating, these α cells were mixed with the wild-type congenic strain XL280 α and co-cultured on V8 medium. As expected, wild-type XL280 produced robust filamentation, which reflects successful self-filamentation (unisexual) or bisexual mating. Longer and denser filamentation was observed during bisexual mating than the self-filamentation. No difference in self-filamentation or bisexual mating filamentation was observed among these strains (Figure 3D). Collectively, the results indicate that insertion into *SH3* itself does not affect cryptococcal growth or virulence traits analyzed *in vitro*.

4. Discussion

In this study, we screened the genomes of serotype D reference strains XL280 and JEC21 and identified one gene free region *SH3* that can serve as a safe haven site for DNA integration. In addition to the information gathered from genome sequence and protein-coding gene annotations, RNA-seq and ATAC-seq data are included in our selection criteria to minimize the risk of selecting heterochromatin regions or regions harboring non-coding but transcribed genes. Although the safe haven region we found is specific to *Cryptococcus* serotype D strains, our screening criteria should be useful for the identification of similar genetic loci in other species.

Our evidence presented here demonstrates that genes inserted in *SH3* site can be expressed efficiently and consistently. We also showed that insertion of DNA into *SH3* did not alter the expression of native genes flanking the *SH3* site. Based on the mCherry reporter strains, DNA insertion into *SH3* does not impact on cryptococcal growth or its classical virulence traits. DNA insertion into *SH3* itself does not alter any of the *C. neoformans* phenotypes examined. We also tested if DNA insertion into *SH3* would affect cryptococcal virulence in the wax moth model (Mylonakis et al., 2005; Perdoni et al., 2014; Slater et al., 2011). Surprisingly, XL280 was no more harmful to the larvae in terms of survival than the control buffer PBS, consistent with the generally low virulence level of serotype D strains. Similar survival rates were observed in the mCherry reporter isolates, suggesting that insertion in *SH3* at least did not significantly enhance cryptococcal virulence in this model (Supplemental Figure 3). That said, our data suggest that the wax moth model is not suitable to assess the virulence of less aggressive strains.

Taking together, this newly identified *SH3* provides a good target for genetic manipulation in *C. neoformans* serotype D strains. Since CRISPR-Cas9 based genome manipulation has been increasingly adopted in *C. neoformans* study, *SH3* can easily serve as the target site for the integration of introduced DNA. In addition, the design of sgRNA targeting *SH3* and the plasmid pYF49 carrying *SH3* homologous arms provide ready-to-use tools for gene complementation and overexpression in these reference serotype D strains. Such tools will facilitate future investigation of gene functions in this species complex.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank all Lin lab members for their helpful suggestions.

Funding

This work was supported by National Institutes of Health (<http://www.niaid.nih.gov/Pages/default.aspx>) (R01AI140719 to XL). The funder had no role in study design, data collection, and interpretation, or the decision to submit the work for publication.

References

- Arras SD, et al., 2015 A genomic safe haven for mutant complementation in *Cryptococcus neoformans*. PLoS One. 10, e0122916. [PubMed: 25856300]
- Arras SD, et al., 2016 Targeted Genome Editing via CRISPR in the Pathogen *Cryptococcus neoformans*. PLoS One. 11, e0164322. [PubMed: 27711143]
- Bahn Y-S, et al., 2005 Specialization of the HOG pathway and its impact on differentiation and virulence of *Cryptococcus neoformans*. Molecular biology of the cell. 16, 2285–2300. [PubMed: 15728721]
- Barchiesi F, et al., 2005 Comparative analysis of pathogenicity of *Cryptococcus neoformans* serotypes A, D and AD in murine cryptococcosis. J Infect. 51, 10–6. [PubMed: 15979484]
- Buenrostro JD, et al., 2015 ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Current Protocols in Molecular Biology. 109, 21.29.1–21.29.9.
- Chacko N, et al., 2015 The lncRNA RZE1 Controls Cryptococcal Morphological Transition. PLoS Genet. 11, e1005692. [PubMed: 26588844]

- Desjardins CA, et al., 2017 Population genomics and the evolution of virulence in the fungal pathogen *Cryptococcus neoformans*. *Genome Res.* 27, 1207–1219. [PubMed: 28611159]
- Edman JC, Kwon-Chung KJ, 1990 Isolation of the *URA5* gene from *Cryptococcus neoformans* var. *neoformans* and its use as a selective marker for transformation. *Mol Cell Biol.* 10, 4538–44. [PubMed: 2201894]
- Evans EE, 1950 The antigenic composition of *Cryptococcus neoformans*. I. a serologic classification by means of the capsular and agglutination reactions. *J Immunol.* 64, 423–30. [PubMed: 15415610]
- Falkow S, 1988 Molecular Koch's postulates applied to microbial pathogenicity. *Rev Infect Dis.* 10 Suppl 2, S274–6. [PubMed: 3055197]
- Fan Y, Lin X, 2018 Multiple Applications of a Transient CRISPR-Cas9 Coupled with Electroporation (TRACE) System in the *Cryptococcus neoformans* Species Complex. *Genetics.* 208, 1357–1372. [PubMed: 29444806]
- Gyawali R, et al., 2017 Pheromone independent unisexual development in *Cryptococcus neoformans*. *PLOS Genetics.* 13, e1006772. [PubMed: 28467481]
- Hicks JK, Heitman J, 2007 Divergence of Protein Kinase A Catalytic Subunits in *Cryptococcus neoformans* and *Cryptococcus gattii* Illustrates Evolutionary Reconfiguration of a Signaling Cascade. *Eukaryotic Cell.* 6, 413. [PubMed: 17189488]
- Janbon G, et al., 2014 Analysis of the Genome and Transcriptome of *Cryptococcus neoformans* var. *grubii* Reveals Complex RNA Expression and Microevolution Leading to Virulence Attenuation. *PLOS Genetics.* 10, e1004261. [PubMed: 24743168]
- Kwon-Chung KJ, et al., 1992 Genetic association of mating types and virulence in *Cryptococcus neoformans*. *Infection and immunity.* 60, 602–605. [PubMed: 1730495]
- Levitz SM, 1991 The ecology of *Cryptococcus neoformans* and the epidemiology of cryptococcosis. *Rev Infect Dis.* 13, 1163–69. [PubMed: 1775849]
- Lin J, et al., 2020 Transformation of *Cryptococcus neoformans* by electroporation using a transient CRISPR-Cas9 expression (TRACE) system. *Fungal Genetics and Biology.* 138, 103364. [PubMed: 32142753]
- Lin J, et al., 2019 Transcription factor Znf2 coordinates with the chromatin remodeling SWI/SNF complex to regulate cryptococcal cellular differentiation. *Communications Biology.* 2, 412. [PubMed: 31754642]
- Lin X, et al., 2006 Virulence attributes and hyphal growth of *C. neoformans* are quantitative traits and the *MAT* α allele enhances filamentation. *PLoS Genet.* 2, e187. [PubMed: 17112316]
- Lin X, et al., 2007 α AD α hybrids of *Cryptococcus neoformans*: evidence of same-sex mating in nature and hybrid fitness. *PLoS Genet.* 3, 1975–90. [PubMed: 17953489]
- Lin X, et al., 2008 Impact of Mating Type, Serotype, and Ploidy on the Virulence of *Cryptococcus neoformans*. *Infection and Immunity.* 76, 2923–2938. [PubMed: 18426889]
- Lin X, et al., 2009 Diploids in the *Cryptococcus neoformans* serotype A population homozygous for the α mating type originate via unisexual mating. *PLoS Pathog.* 5, e1000283 (1–18).
- Liu LX, et al., 2018 Genetic basis for coordination of meiosis and sexual structure maturation in *Cryptococcus neoformans*. *eLife.* 7, e38683. [PubMed: 30281018]
- Loftus BJ, et al., 2005 The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. *Science (New York, N.Y.)* 307, 1321–1324.
- Meng Y, et al., 2018 Plant Homeodomain Genes Play Important Roles in Cryptococcal Yeast-Hypha Transition. *Appl Environ Microbiol.* 84.
- Meyer W, et al., 2003 Molecular typing of IberoAmerican *Cryptococcus neoformans* isolates. *Emerg Infect Dis.* 9, 189–195. [PubMed: 12603989]
- Mylonakis E, et al., 2005 *Galleria mellonella* as a model system to study *Cryptococcus neoformans* pathogenesis. *Infection and immunity.* 73, 3842–3850. [PubMed: 15972469]
- Ni M, et al., 2013 Unisexual and Heterosexual Meiotic Reproduction Generate Aneuploidy and Phenotypic Diversity De Novo in the Yeast *Cryptococcus neoformans*. *PLOS Biology.* 11, e1001653. [PubMed: 24058295]
- Nielsen K, et al., 2005 *Cryptococcus neoformans* α strains preferentially disseminate to the central nervous system during coinfection. *Infect Immun.* 73, 4922–33. [PubMed: 16041006]

- Nielsen K, et al., 2003 Sexual cycle of *Cryptococcus neoformans* var. *grubii* and virulence of congenic α and α isolates. *Infect Immun*. 71, 4831–41. [PubMed: 12933823]
- Otto B, 1894 Ueber parasitäre Zelleinschlüsse und ihre Züchtung. *Centralblatt für Bakteriologie und Parasitenkunde*. 16, 175–180.
- Pellenz S, et al., 2018 New human chromosomal safe harbor sites for genome engineering with CRISPR/Cas9, TAL effector and homing endonucleases. *bioRxiv*. 396390.
- Perdoni F, et al., 2014 A histological procedure to study fungal infection in the wax moth *Galleria mellonella*. *European journal of histochemistry : EJH*. 58, 2428–2428. [PubMed: 25308852]
- Perfect JR, Bicanic T, 2015 Cryptococcosis diagnosis and treatment: What do we know now. *Fungal Genetics and Biology*. 78, 49–54. [PubMed: 25312862]
- Pham T, et al., 2020 An intergenic “safe haven” region in *Aspergillus fumigatus*. *Medical Mycology*.
- Rajasingham R, et al., 2017 Global burden of disease of HIV-associated cryptococcal meningitis: an updated analysis. *Lancet Infect Dis*. 17, 873–881. [PubMed: 28483415]
- Sadelain M, et al., 2011 Safe harbours for the integration of new DNA in the human genome. *Nature Reviews Cancer*. 12, 51. [PubMed: 22129804]
- Sanfelice F, 1895 Sull’azione patogena dei blastomiceti. *Annali d’Igiene Sperimentale*. 5, 239–62.
- Slater JL, et al., 2011 Pathogenicity of *Aspergillus fumigatus* mutants assessed in *Galleria mellonella* matches that in mice. *Medical Mycology*. 49, S107–S113. [PubMed: 20950221]
- Tenney AE, et al., 2004 Gene prediction and verification in a compact genome with numerous small introns. *Genome research*. 14, 2330–2335. [PubMed: 15479946]
- Toffaletti DL, et al., 2004 *Cryptococcus neoformans* mitochondrial genomes from serotype A and D strains do not influence virulence. *Curr Genet*. 46, 193–204. [PubMed: 15309505]
- Toffaletti DL, et al., 1993 Gene transfer in *Cryptococcus neoformans* by use of biolistic delivery of DNA. *J Bacteriol*. 175, 1405–11. [PubMed: 8444802]
- Upadhy R, et al., 2017 A fluorogenic *C. neoformans* reporter strain with a robust expression of m-cherry expressed from a safe haven site in the genome. *Fungal Genetics and Biology*. 108, 13–25. [PubMed: 28870457]
- Walton FJ, et al., 2005 Novel gene functions required for melanization of the human pathogen *Cryptococcus neoformans*. *Mol Microbiol*. 57, 1381–96. [PubMed: 16102007]
- Wang L, et al., 2014 Morphotype transition and sexual reproduction are genetically associated in a ubiquitous environmental pathogen. *PLoS Pathog*. 10, e1004185. [PubMed: 24901238]
- Wang L, et al., 2012 The link between morphotype transition and virulence in *Cryptococcus neoformans*. *PLoS Pathog*. 8, e1002765. [PubMed: 22737071]
- Wang Y, et al., 2016 A ‘suicide’ CRISPR-Cas9 system to promote gene deletion and restoration by electroporation in *Cryptococcus neoformans*. *Scientific Reports*. 6, 31145. [PubMed: 27503169]
- Wilson DE, et al., 1968 Serologic grouping of *Cryptococcus neoformans*. *P Soc Exp Biol Med*. 127, 820–3.
- Xu JP, et al., 2000 Multiple gene genealogies reveal recent dispersion and hybridization in the human pathogenic fungus *Cryptococcus neoformans*. *Mol Ecol*. 9, 1471–81. [PubMed: 11050543]
- Xu X, et al., 2017 Glucosamine stimulates pheromone-independent dimorphic transition in *Cryptococcus neoformans* by promoting Crz1 nuclear translocation. *PLoS Genet*. 13, e1006982. [PubMed: 28898238]
- Zhai B, et al., 2013 Congenic strains of the filamentous form of *Cryptococcus neoformans* for studies of fungal morphogenesis and virulence. *Infect Immun*. 81, 2626–37. [PubMed: 23670559]
- Zhao Y, et al., 2019 Life Cycle of *Cryptococcus neoformans*. *Annual Review of Microbiology*. 73, 17–42.
- Zhao Y, et al., 2018 PAS Domain Protein Pas3 Interacts with the Chromatin Modifier Bre1 in Regulating Cryptococcal Morphogenesis. *mBio*. 9, e02135–18. [PubMed: 30425151]

Highlights

- A gene free but unsuppressed “safe haven” (*SH*) region is a preferred locus for alien DNA integration.
- A safe haven region *SH3* in *Cryptococcus neoformans* serotype D strains is identified through a screen incorporating genomic, transcriptomic, and chromatin accessibility data.
- Genes inserted into *SH3* showed reliable expression.
- DNA insertion into *SH3* itself does not affect the expression of adjacent genes and did not alter cryptococcal growth or mating.

Figure 1. Identification of potential safe haven sites in *Cryptococcus neoformans* serotype D strains. (A) Flow chart and criteria for screening potential safe haven candidates. (B) Genome location and feature of the selected safe haven *SH3*. Brown line represents the sgRNA target sequence. (C) RNA-seq profile of the *SH3* region on chromosome IV in XL280 when cells were cultured under YPD, V8 16h, and V8 72h. The FPKM range (y-axis) is set to 0–1000. Our RNA-seq data suggest that there is probably a mis-annotation for CNXL_025530, which is based on the annotation of *CND03070* in JEC21. (D) ATAC-seq profile of the *SH3* region on chromosome IV in XL280 under YPD and V8 16h condition. Reads range is set to 0–100.

Figure 2. Genes inserted in SH3 can be expressed efficiently.

(A) Diagram of $ADE2$ complementation via TRACE. (B) A plate with $ADE2$ complementation transformants is shown. Colonies of $ade2$ show red color while complemented colonies show white color. (C) Diagram of a mCherry reporter in $SH3$ with two possible directions of integration resulting in F-mCherry and R-mCherry. Arrows indicate primers for screening. (D) Two isolates of F-mCherry and R-mCherry along with the wild-type XL280 strain were examined microscopically under DIC and the mCherry filter. Scale Bar: 10 μ m. (E) Quantification of mCherry intensity in selected F-mCherry isolates and R-mCherry isolates. The relative strength of the mCherry signal in the WT strain was used as reference. P values $\leq .05$ were considered significant (One-Way ANOVA). n.s.: non-significant, ****: $P \leq .0001$ (F) Diagram of pYF49 with homologous arms of $SH3$.

Figure 3. DNA insertion in SH3 does not impact the transcript level of the neighboring genes or the examined phenotypes.

(A) The transcript levels of the *SH3* flanking genes in three biological replicates of XL280 and five mCherry reporter strains were tested via qRT-PCR. The transcript level of the house-keeping gene *TEF1* was used as the internal control for each sample and the transcript level of each gene in the wild-type XL280 was normalized to 1. P values $\leq .05$ were considered significant (One-Way ANOVA), n.s.: non-significant. (B) Wild-type XL280, F-mCherry, and R-mCherry isolates were serially diluted and dropped onto YPD or YNB medium to test their growth and temperature sensitivity. Cells were cultured at 22°C, 30°C or 37°C for two days to test thermotolerance. (C) Cells were cultured 4 days on L-Dopa medium to test melanization or RPMI medium to test capsule formation. (D) Wild type and three SH3-mCherry isolates were diluted to equal concentration and cultured on V8 agar

medium alone (unisex) or with the mating partner XL280a cells in same concentration (bisexual mating) at 22°C for 96 h before the images of the edges of colonies were taken.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript