A PCR-based strategy to generate integrative targeting alleles with large regions of homology

Robert C. Davidson, Jill R. Blankenship, Peter R. Kraus, Marisol de Jesus Berrios, Christina M. Hull, Cletus D’Souza, Ping Wang and Joseph Heitman

Departments of Molecular Genetics and Microbiology, Pharmacology and Cancer Biology, and Medicine, and the Howard Hughes Medical Institute, Duke University Medical Center, Research Drive, Durham, NC 27710, USA

Cryptococcus neoformans is an opportunistic fungal pathogen with a defined sexual cycle for which genetic and molecular techniques are well developed. The entire genome sequence of one C. neoformans strain is nearing completion. The efficient use of this sequence is dependent upon the development of methods to perform more rapid genetic analysis including gene-disruption techniques. A modified PCR overlap technique to generate targeting constructs for gene disruption that contain large regions of gene homology is described. This technique was used to disrupt or delete more than a dozen genes with efficiencies comparable to those previously reported using cloning technology to generate targeting constructs. Moreover, it is shown that disruptions can be made using this technique in a variety of strain backgrounds, including the pathogenic serotype A isolate H99 and recently characterized stable diploid strains. In combination with the availability of the complete genomic sequence, this gene-disruption technique should pave the way for higher throughput genetic analysis of this important pathogenic fungus.

Keywords: gene disruption, PCR overlap, Cryptococcus neoformans

INTRODUCTION

Cryptococcus neoformans is a basidiomycetous fungus that has been recognized as a human pathogen for more than a century (Casadevall & Perfect, 1998; Sanfelice, 1894). C. neoformans mainly infects immunocompromised individuals, and disease typically manifests as meningocerebralitis (Casadevall & Perfect, 1998). With the relative increase in recent years in the numbers of immunocompromised patients due to AIDS, cancer chemotherapy, transplant-rejection suppression therapy, and other factors, the incidence of opportunistic fungal infections has risen (Mitchell & Perfect, 1995). C. neoformans infection results from inhalation of spores or desiccated yeast cells, which then spread via the bloodstream to infect the central nervous system and cause life-threatening meningocerebralitis.

In addition to being an important cause of disease in immunocompromised patients, C. neoformans is an excellent model of fungal pathogenesis. Several well-defined virulence factors include the polysaccharide capsule (which promotes intracellular survival in macrophages), the pigment melanin (which prevents oxidation by macrophages), the production of the enzymes urease and phospholipase B, and the ability to grow at 37 °C (Alspaugh et al., 2000a; Casadevall & Perfect, 1998; Chen et al., 2000; Cox et al., 2000, 2001; Cruz et al., 2000; Fox et al., 2001; Kwon-Chung & Bennett, 1992; Odom et al., 1997). Several animal models have been developed, including a mouse tail-vein injection model, rat and mouse inhalation models, and a rabbit meningitis model (Goldman et al., 1994; Kwon-Chung et al., 1982; Perfect et al., 1980). Importantly, C. neoformans is genetically tractable, with a well-characterized sexual cycle involving mating between two haploid mating types, MATa and MATα (Alspaugh et al., 2000b; Kwon-Chung, 1975, 1976). Efficient transformation and gene-disruption protocols have been developed along with auxotrophic and dominant selectable markers and

Abbreviation: MAP, mitogen-activated protein.
overexpression plasmids (Davidson et al., 2000; Edman & Kwon-Chung, 1990; Hua et al., 2000; McDade & Cox, 2001; Sudarshan et al., 1999; Toffaletti et al., 1993). Furthermore, genetic analysis has been facilitated by the recent characterization of a stable diploid state, which allows identification and analysis of essential genes (Sta et al., 2000). Finally, the entire genome sequence is nearing completion for the serotype D strain JEC21, and a sequencing project has been initiated for the clinical serotype A isolate H99 (Stanford, TIGR, UBC, Duke, Whitehead; reviewed by Heitman et al., 1999). These advances will allow extensive comparative genomics experimentation between C. neoformans strains and other fungal genomes.

The efficient use of the complete genome sequence will require more rapid gene-function testing, a process that is limited by tedious construction of targeting alleles by cloning. PCR-generated targeting alleles have been employed for Saccharomyces cerevisiae and for the human pathogen Candida albicans, which has overcome the necessity for cloning and has allowed more high-throughput genetic analysis (Baudin et al., 1993; Eberhardt & Hohmann, 1995; Lorenz et al., 1995; Wach, 1996; Wach et al., 1994; Wilson et al., 1999). However, efficient homologous recombination in C. neoformans requires larger regions of homology, which has prevented the application of similar PCR-based strategies.

Here we present a modified application of a technique called PCR overlap to generate targeting alleles with larger regions of homology (Ho et al., 1989; Horton et al., 1989). This technique can be used in any system where homologous recombination requires larger regions than those that can be incorporated into synthetic oligonucleotides, and will effectively eliminate the time-consuming process of searching for convenient restriction sites and cloning targeting alleles.

**METHODS**

**Gene disruptions.** To create the ste11p mutant strain, first the 4.5 kb STE11 locus was amplified from MATα genomic DNA using primers JOHE5391 (GTCTGTTCTCCCTCCTGTAC) and JOHE5392 (CCTGCCAGCGCCGTGTAAT) (R. C. Davidson and others, unpublished). The ste11α::URA5 disruption allele was constructed using PCR overlap (outlined in Fig. 1). In the first round, the 5′ end of the STE11α gene was amplified with primers JOHE5391 (primer 1) (GTCTGTTCTCCCTCCTGTAC) and JOHE5306 (primer 3) (GTTGCCGACAACTTCCTACATTACGCGGTGG). The 3′ end of the STE11α gene with primers JOHE5307 (primer 4) (ACTCTCCCAAGGCAAGATATTTCATGCGTACGCGGTGG) and JOHE5392 (primer 6) (CAGTCGTGTTCGTCCTGCTGCT) and the URA5 gene with primers JOHE5305 (primer 2) (GTCTGTTCTCCCTCCTGTAC) and JOHE5308 (primer 5) (CGCCGCAACTCCGGCGGGCGGTGG) were then used to overlap the three products to yield the 4.2 kb ste11α::URA5 disruption allele. The PCR product was gel-purified, extracted and directly introduced into the serotype D ura5 strain JEC43 by biolistic transformation.

The tor1Δ::URA5/TOR1 mutant strain was generated by PCR overlap using six primers as outlined (Fig. 1). The 5′ end of the TOR1 gene was amplified with primers JOHE437 (no. 1) (CCTGAGTTGAGATTCTTAC) and JOHE5850 (no. 3) (CCTGGCTAGCTTTTCTTCGTGGCAAGCTGA) and the 3′ end of TOR1 with primers JOHE5851 (no. 4) (ACTGGCCGCTCGTATTCGAGAACTGCGG) and JOHE4413 (no. 6) (GTACAAAGTATGTGCTTGC) and the URA5 selectable marker with primers JOHE5849 (no. 2) (GCCTTTCTACGTTTGCGGAGGACAGCTATGACCATG) and JOHE5852 (no. 5) (GCTCTCCTCGAGCTCTATCGTAAACAGCAGCCCGAGT). The URA5 fragment in this disruption was amplified from plasmid pRCD69, which contains the C. neoformans serotype D URA5 sequence in the pCR2.1-TOPO vector, and the primers were designed to amplify from the vector sequence. The three amplified products were run on a 0.6% agarose gel and extracted together using the Qiapquick method (Qiagen) and used as templates for the overlap reaction. Primers JOHE6209 (no. 1) and JOHE4413 (no. 6) were used to overlap the three first-round products to yield the tor1Δ::URA5 allele. The PCR product was gel-purified, extracted and transformed directly into the C. neoformans serotype D diploid strain RAS008 by biolistic transformation.

The mpk1Δ::URA5 mutant strain was generated by PCR overlap using six primers as outlined (Fig. 1). The 5′ end of the MPK1 gene was amplified with primers JOHE7288 (no. 1) (ACTAGGCGTGGCATTGTTTAC) and JOHE7418 (no. 4) (GGTCGAGCAACTTTCGTCAGGATTGCTGCGGACAGTG), the 3′ end of MPK1 with primers JOHE7419 (no. 4) (CCACCTCTGGAGCAAGCAGATATTACGCGGTGG) and JOHE7289 (no. 6) (GGCGACTGGGCACCAGAGACAGCGACAGTG). The three amplified products were run on a 0.6% agarose gel and extracted together using the Qiapquick method (Qiagen) and were used as templates for the overlap reaction. Primers JOHE7288 (no. 1) and JOHE7289 (no. 6) were used to overlap the three first-round products to yield the mpk1Δ::URA5 disruption allele. This amplified product was gel-purified, extracted and directly transformed into the serotype D ura5 strain JEC43 by biolistic transformation.

**Overlap primer design.** Overlap oligos were approximately 35–40 bp in length. Primers were obtained from Integrated DNA Technology and were not PAGE-purified. In the STE11α and MPK1 deletions, the portion of the oligo corresponding to the selectable marker flanked the C. neoformans URA5 gene. In the TOR1 disruption, the portion of the oligo corresponding to the selectable marker was designed to amplify from a plasmid containing URA5. This was done so that any selectable marker cloned into the vector could be easily inserted into the disruption cassette.

The sequence of primer 3 was designed to be completely complementary to the sequence of primer 2, and the sequence of primer 4 was complementary to the sequence of primer 3 as outlined in Fig. 1. This strategy generates fragments with approximately 40 bp of overlap for the final PCR reaction. The 3′ ends of primer 3 and primer 4 were designed so that the
The first-round PCR of the reproducibly increases the yield of specific overlap products. Importantly, denaturation and annealing times were limited to 10–15 s during the final overlap amplification, which was 2 min at 95 °C and a final extension of 5 min at 72 °C. The PCR samples were visualized using standard DNA electrophoretic techniques (Sambrook et al., 1989).

**Southern blot analysis.** Genomic DNA was isolated from JEC21 and JEC20 wild-type, ste11α and mpk1 mutant strains by the method of Pitkin et al. (1996). Twenty micrograms genomic DNA was digested with the appropriate enzymes and electrophoresed on a 0.8% TBE agarose gel. Transfer, hybridization and autoradiography were performed as described by Sambrook et al. (1989). Fragments of the STE11α and MPK1 genomic ORFs were used as probes for Southern blot hybridization, using [α-32P]dCTP (Amersham) and the Prime-It II random primed labelling kit (Stratagene).

**RESULTS**

**Targeted disruption of the STE11α gene**

The STE11α gene encodes a mating type-specific MEK kinase homologue that is related to the Ste11 kinase of *S. cerevisiae*, a component of the pheromone-sensing mitogen-activated protein (MAP) kinase cascade. The *C. neoformans* STE11α gene was identified by its location in the MATα locus adjacent to the first pheromone gene discovered in *C. neoformans* by Moore & Edman (Clarke et al., 2001; Moore & Edman, 1993). To perform epistasis analysis with other members of this MAP kinase cascade, we made a ste11α mutant strain by homologous integration of a disruption allele. We used PCR to generate a ste11α::URA5 disruption allele. Six oligonucleotide primers (numbered 1–6) were directed against the STE11α ORF and the URA5 selectable marker (strategy outlined in Fig. 1). Separate PCR
The TOR1 gene encodes a homologue of the Tor kinases of S. cerevisiae. To obtain functional information about this gene, we took a gene-deletion approach. The gene was originally identified in serotype A (Cruz et al., 1999), and the serotype D sequence was then partially assembled from the Stanford and TIGR databases by comparison of serotype A and serotype D sequences. Six primers (numbered 1–6) were constructed to amplify the 5' and 3' portions of the TOR1 gene and the selectable marker, URA5 (Fig. 3a, b, lanes 1–3). These three products and primers 1 and 6 were used to amplify a single linear disruption allele that consisted of URA5 flanked by ~1 kb of TOR1 gene sequence on each side (Fig. 3b, lane 4). The linear overlap PCR product was gel-purified and used directly for transformation into a haploid recipient strain of C. neoformans. No homologous recombination events were obtained out of approximately 200 Ura+ isolates screened, indicating that TOR1 might be an essential gene in C. neoformans.

To establish whether TOR1 is essential or not, we sought to isolate tor1/TOR1 heterozygous mutants using the recently discovered stable, congeneric diploid strains of C. neoformans (Sia et al., 2000). A ura5/ura5 diploid strain of C. neoformans (RAS008) was transformed with the linear tor1::URA5 PCR-generated disruption allele, and Ura+ transformants were selected. Using PCR amplification and digestion with the Sall restriction enzyme, which specifically cleaves the mutant allele but not the wild-type at a pair of recognition sites present in the URA5 gene (Fig. 3c), one tor1::URA5/TOR1 strain was identified out of 56 Ura+ transformants screened (1.8%) (Table 1). These results indicate that PCR-generated disruption constructs can be used to target genes in the laboratory-constructed congeneric
**Table 1.** Genes disrupted using PCR-generated targeting alleles

<table>
<thead>
<tr>
<th>Gene</th>
<th>Strain (genotype)</th>
<th>Frequency* (%)</th>
<th>Homology†</th>
</tr>
</thead>
<tbody>
<tr>
<td>ste11Δ::URA5</td>
<td>JEC43 (Sero D MATα ura5)</td>
<td>2</td>
<td>1040…1180</td>
</tr>
<tr>
<td>tor1Δ::URA5</td>
<td>RA5008 (Sero D MATα/α ura5/ura5)</td>
<td>2</td>
<td>1222…1030</td>
</tr>
<tr>
<td>mpk1Δ::URA5</td>
<td>JEC43 (Sero D MATα ura5)</td>
<td>10</td>
<td>628…835</td>
</tr>
<tr>
<td>ste12Δ::URA5</td>
<td>JEC55 (Sero D ade2-27 ura5)</td>
<td>10</td>
<td>1476…1661</td>
</tr>
<tr>
<td>tlk1Δ::URA5</td>
<td>JEC34 (Sero D MATα/α ura5)</td>
<td>4</td>
<td>879…956</td>
</tr>
<tr>
<td>fbb1Δ::URA5</td>
<td>H99 5-FOA (Sero A MATα ura5)</td>
<td>22</td>
<td>682…1113</td>
</tr>
<tr>
<td>gno1Δ::NAT1</td>
<td>JEC43 (Sero D MATα ura5)</td>
<td>23‡</td>
<td>792…656</td>
</tr>
<tr>
<td>gno1Δ::NAT1</td>
<td>H99 5-FOA (Sero A MATα ura5)</td>
<td>77‡</td>
<td>549…656</td>
</tr>
<tr>
<td>pka2Δ::URA5</td>
<td>JEC43 (Sero D MATα ura5)</td>
<td>87</td>
<td>800…561</td>
</tr>
<tr>
<td>sxi1Δ::URA5</td>
<td>JEC43 (Sero D MATα ura5)</td>
<td>3</td>
<td>530…860</td>
</tr>
<tr>
<td>sxi1Δ::URA5</td>
<td>RA5009 (Sero D MATα/α ura5/ura5)</td>
<td>4</td>
<td>1102…1007</td>
</tr>
<tr>
<td>sxi1Δ::URA5</td>
<td>JEC43 (Sero D MATα ura5)</td>
<td>10</td>
<td>1102…1007</td>
</tr>
<tr>
<td>gno1Δ::NAT1</td>
<td>RA5009 (Sero D MATα/α ura5/ura5)</td>
<td>ND</td>
<td>1103…1014</td>
</tr>
<tr>
<td>sxi1Δ::NAT1</td>
<td>JEC43 (Sero D MATα ura5)</td>
<td>10</td>
<td>1103…1014</td>
</tr>
<tr>
<td>gpa1Δ::URA5</td>
<td>H99 5-FOA (Sero A MATα ura5)</td>
<td>12</td>
<td>650…540</td>
</tr>
<tr>
<td>crg1Δ::URA5</td>
<td>PPW28 (Sero A MATα ade2 sch9Δ::ADE2 ura5)</td>
<td>10</td>
<td>590…605</td>
</tr>
</tbody>
</table>

*Frequency refers to the homologous recombination frequency obtained. ND, Not determined.†Homology refers to the amount of the 5′ (left) and 3′ (right) regions of each gene incorporated into the disruption alleles in base pairs.‡In these cases, this is the frequency of homologous recombination in isolated stable transformants.

**Fig. 3.** Deletion of the **TOR1** gene. (a) The **TOR1** gene was deleted by replacing a portion of the **TOR1** ORF with the **URA5** selectable marker by PCR overlap. (b) The 5′ region of **TOR1** (lane 1), the **URA5** selectable marker (lane 2), and the 3′ region of **TOR1** (lane 3) were amplified as described (Methods). The three products were used as templates to amplify the tor1Δ::URA5 deletion allele (lane 4). (c) PCR analysis was performed to confirm the heterozygous gene deletion. Genomic DNA from wild-type **MATα** (WT) and the tor1/tor1 mutant strain was PCR-amplified with primers y and z and digested with **SalI** (S), which specifically digests the mutant allele but not the wild-type allele. Two of the fragments generated by the digestion of the mutant disruption allele were approximately the same size and thus appear as a single 21 kb band. The smaller **SalI–SalI** fragment inside the **URA5** sequence of approximately 650 bp is not shown in this figure.

diploid strains, and provide evidence that **TOR1** is an essential gene, since it could be disrupted in a diploid, but not a haploid, strain of **C. neoformans**. We do note that gene disruption in the congenic serotype D diploid strains typically results in a lower frequency of homologous recombination (1–5%) than that commonly observed in haploid recipient strains (2–25% on average).
Targeted disruption of the MPK1 gene

The gene encoding the MAP kinase homologue Mpk1 was identified by comparative genomics using the Stanford C. neoformans genome sequence database (P. R. Kraus and others, unpublished). A complete sequence of the putative gene was assembled using the Stanford sequence as well as that from the TIGR C. neoformans sequencing project. To assess the function of this MAP kinase homologue, we generated an allele for gene disruption using PCR overlap. As described above, primers were directed against the 5’ and 3’ flanking regions of the MPK1 gene and the URA5 selectable marker, and three products were amplified (Fig. 4a, b, lanes 1–3). These products were overlapped into a single product with the portions of the MPK1 gene flanking URA5 (Fig. 4b, lane 4). The mpk1Δ::URA5 PCR-generated disruption allele was transformed into a serotype D ura5 strain of C. neoformans (JEC43), and Ura+ colonies were selected. Two mpk1Δ::URA5 mutant strains were identified by PCR out of 20 screened (10%) (Table 1), which was confirmed by a Southern blot (Fig. 4c).

Summary of results

Together, these results indicate that PCR overlap can be used to generate targeting alleles for gene disruption in C. neoformans. Successful PCR overlap amplification products have been obtained in each case for which it has been attempted (Table 1). However, in some cases, obtaining sufficient product for transformation has required manipulation of PCR parameters, particularly when transforming diploid strains, which typically require more transforming DNA (3–5 µg) to obtain transformants. For example, reducing the denaturation and annealing times to 15 s or less has increased product yield dramatically in several cases, particularly when using the dominant selectable marker NAT1. Additionally, to maximize product, it is important to maintain relatively equimolar amounts of the templates in the overlap reaction. In all, we have obtained 17 gene disruptions using this PCR overlap technique to generate targeting alleles. Thus, we show that the PCR overlap technique is generally applicable and can be used with a variety of recipient strains as well as with both auxotrophic and dominant selectable markers (Table 1).

DISCUSSION

C. neoformans is an important human pathogenic fungus. As a genetically tractable organism with well-developed molecular biology and animal models, and a nearly completed genomic sequence that is publicly available, C. neoformans is an excellent model for the study of fungal pathogenesis (Casadevall & Perfect, 1998). However, efficient use of the sequence data
requires implementation of improved molecular techniques such as gene disruption. In many model systems, PCR has replaced cloning as a faster, more efficient method of generating disruption alleles. Generation of disruption constructs in S. cerevisiae is performed almost exclusively by the technique involving large synthetic oligonucleotides that contain 40 bp of homology to the gene of interest followed by 20 bp of homology to the selectable marker (Eberhardt & Hohmann, 1993; Lorenz et al., 1995; Wach, 1996; Wach et al., 1994). The selectable marker is then amplified and the PCR product used to directly transform and generate the mutant strain (Tong et al., 2001). This technique was also recently implemented for producing mutant strains in the human pathogen Candida albicans (Wilson et al., 1999). However, in Cryptococcus neoformans, a larger region of homology is required for integration and thus cannot be incorporated into synthetic oligonucleotides.

We describe a modified use of the PCR overlap technique originally described by Ho et al. (1989) and Horton et al. (1989) to generate disruption constructs without the need for cloning. This method can be used to generate both partial and complete deletions of the targeted gene’s ORF. Using this method, we have disrupted three genes, STE11x (encoding a MAP kinase kinase homologue), TORI (encoding a Tor kinase homologue) and MKP1 (encoding a MAP kinase homologue), at efficiencies consistent with previous studies (Davidson et al., 2000). We also report the disruption of 11 other genes using the same method and a variety of recipient strains, including the recently described serotype D diploids (Sia et al., 2000). Moreover, both the auxotrophic URA5 and dominant NAT1 selectable markers were used, which allows generation of double mutant strains or the use of prototrophic strains. In all cases, the efficiency of targeted integration was in the range 2–10% or higher, consistent with previous studies using biolistic transformation in C. neoformans (Alspaugh et al., 1997; Davidson et al., 2000; Toffaletti et al., 1993; Wang et al., 2000). These findings indicate that generation of constructs by two rounds of PCR does not seem to inhibit homologous recombination significantly.

Additional screening tests can make the PCR-disruption method even more efficient. Jennifer Lodge and co-workers recently found that a significant proportion of initial transformants are unstable, in accord with earlier results (Edman & Kwon-Chung, 1990), and by implementing an initial screening step for stable transformants they found that the efficiency of homologous targeting can be increased (Nelson et al., 2002). We have confirmed these observations in the background of our PCR-based approach to gene disruption and found that approximately 33% of transformants are stable. In the case of the gmo1::NAT1 disruption in strain JEC34, the frequency of homologous integration was increased from 10 to 23%, and, in the case of the gmo1::NAT1 disruption in strain H99, from 28 to 77%, by first screening for stable transformants. Thus, implementing this important advance of Lodge and coworkers, together with the PCR overlap approaches described here, should result in even more efficacious gene-disruption frequencies.

Implementation of this PCR-based gene-disruption technique will preclude the generation of mutant strains from being the rate-limiting step in performing genetic analyses. The recently reported RNAi and antisense methods for disrupting gene function and the ability to generate panels of strains containing randomly inserted signature tags should also aid in accelerating the analysis of gene function (Gorlach et al., 2002; Liu et al., 2002; Nelson et al., 2001). However, the PCR-based method proposed here has the potential to be applied to situations in which the ability to alter targeted genomic sequence is necessary. For instance, larger deletions that span multiple genes can be performed using this method. In addition to disruption mutations, this technique can also be applied to the efficient generation of other targeted insertions, including site-directed mutations and the insertion of regulatable promoters or epitope tags.

In conclusion, we show that this PCR-based gene-disruption approach is generally applicable for different genes using a variety of strains and genetic markers, and this should allow more efficient analysis of gene function in cases where the complete sequence is available.

ACKNOWLEDGEMENTS

We thank John Perfect, Jennifer Görlach, Andy Alspaugh, Cristina Cruz, Connie Nichols and Klaus Lengeler for valuable technical advice and scientific discussions, Debbie Fox for reviewing the manuscript, and Jennifer Lodge for sharing data with us prior to publication. These studies were supported, in part, by P01 grant AI44975 to the Duke University Mycology Research Unit, R01 grants AI39115, AI50438, AI50115 (J.H.) and a minority supplement, AI39115S1 (M. J.B.), all from the NIAID. C.M.H. is supported by a Damon Runyon Cancer Research Fund Fellowship (DRG-1694). J.H. is an associate investigator of the Howard Hughes Medical Institute and a Burroughs Wellcome Scholar in molecular pathogenic mycology. Some of the genomic sequences used in this paper were obtained from the C. neoformans Genome Project, Stanford Genome Technology Center, funded by the NIAID/NHGRI under cooperative agreement U01 AI47087, and The Institute for Genomic Research, funded by the NIAID/NIG under cooperative agreement U01 AI48594.

REFERENCES


Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F. &


PCR-based generation of large disruption alleles

deletion mutants, Yeast 12, 239–265.


Received 27 February 2002; revised 27 March 2002; accepted 18 April 2002.