Allele-specific gene targeting in *Candida albicans* results from heterology between alleles

Kyle Yesland and William A. Fonzi

Author for correspondence: William A. Fonzi. Tel: +1 202 687 1135. Fax: +1 202 687 1800.
e-mail: fonziw@gusun.georgetown.edu

Department of
Microbiology and
Immunology, Georgetown
University, 3900 Reservoir
Road NW, Washington,
DC 20007-2197, USA

The opportunistic fungal pathogen *Candida albicans* is asexual and diploid. Thus, introduction of recessive mutations requires targeted gene replacement of two alleles to effect expression of a recessive phenotype. This is often performed by recycling of a *URA3* marker gene that is flanked by direct repeats of *hisG*. After targeting to a locus, recombination between the repeats excises *URA3* leaving a single copy of *hisG* in the disrupted allele. The remaining functional allele is targeted in a second transformation with the same *URA3* marked construct. Replacement can be highly biased toward one allele. At the *PHR1* locus, there was an approximately 50-fold preference for replacement of the disrupted versus the functional allele in a heterozygous mutant. This preference was reduced six- to eightfold when the transforming DNA lacked the *hisG* repeats. Nonetheless, there remained a sixfold preference for targeting a particular allele of *PHR1* and this was evident even in transformations of the parental strain containing two wild-type alleles of *PHR1*. Both wild-type alleles were cloned and nucleotide sequence comparison revealed 24 heterologies over a 2 kb region. Using restriction site polymorphisms to distinguish alleles, it was observed that transformation with the cloned DNA of allele *PHR1-1* preferentially targeted allele 1 of the genome. Transformations with *PHR1-2* exhibited the reciprocal specificity. In both these instances, heterology was present in the flanking regions of the transforming DNA. When the transforming DNA was chosen from a region 100% identical in both alleles, alleles 1 and 2 were targeted with equal frequency. It is concluded that sequence heterology between alleles results in an inherent allele specificity in targeted recombination events.

Keywords: *Candida albicans*, recombination, gene targeting, heterozygosity

INTRODUCTION

Candida albicans is an opportunistic fungal pathogen of humans. Because of its clinical significance, the biology of this fungus has been the subject of numerous studies (Odds, 1988). Despite this extensive investigation, relatively little is known of its basic genetic behaviour. This is due, in part, to its genetic intractability, an asexual life cycle and a diploid genome (Odds, 1988). The use of parasexual techniques has provided some information

and suggests that *C. albicans* behaves much like other fungi with, perhaps, a few distinctions (Poulter, 1990).

Currently, genetic manipulation of *C. albicans* relies almost exclusively on reverse genetic methods. The demonstration of transformation (Kelly *et al.*, 1988, 1987) coupled with workable methods of gene replacement (Fonzi & Irwin, 1993; Gorman *et al.*, 1991) has fostered numerous molecular genetic studies of *C. albicans* biology. Although improved methods of gene replacement have been developed (Morschhauser *et al.*, 1999; Wilson *et al.*, 1999, 2000), most studies to date have relied on recycling of a *URA3* marker using positive and negative selection (Fonzi & Irwin, 1993). This method, first developed for *Saccharomyces cerevisiae*,

The GenBank accession numbers for the sequences reported in this paper are AF247189 and AF247190 for *PHR1-1* and *PHR1-2*, respectively.

employs a cassette consisting of direct repeats of *Salmonella typhimurium hisG* flanking a *URA3* marker gene (Alani *et al.*, 1987; Fonzi & Irwin, 1993). The cassette is inserted within the cloned gene of interest, which provides homologous flanking sequences to direct targeted gene replacement. Integration of transforming DNA in *C. albicans* occurs largely via homologous recombination (Kelly *et al.*, 1987). The integrated DNA restores uridine prototrophy to the *ura3* recipient cells. Spontaneous recombination between the *hisG* sequences results in loss of the *URA3* marker and retention of a single copy of *hisG* at the disrupted locus (Fonzi & Irwin, 1993). The *Uri*⁻ uridine auxotrophs resulting from this excision event are selected on 5-fluoroorotic-acid-containing medium (Boeke *et al.*, 1984).

Because *C. albicans* is diploid, expression of recessive mutations requires disruption of two alleles. The heterozygous *Uri*⁻ mutants are transformed with the same cassette construction, selecting again for uridine prototrophy. Statistically, the cassette is expected to integrate with equal probability into either the previously disrupted allele or the remaining wild-type allele. In practice this is not the case. The ratio of targeting to the wild-type versus the disrupted allele can vary from 1 in 3 to 1 in 35, depending on the locus under investigation (Fonzi & Irwin, 1993; Sentandreu *et al.*, 1998). The factors responsible for this allele specificity and locus-specific variation are unknown.

In this study we examined allele-specific integration at the *PHR1* locus (Saporito-Irwin *et al.*, 1995). In a heterozygous mutant with the genotype *phr1Δ::hisG/PHR1*, transforming DNA recombined with the *phr1Δ::hisG* allele approximately 50-times more frequently than with the wild-type allele. This selectivity was only partially due to the additional homology between the disrupted allele and the transforming DNA provided by the *hisG* sequences. Restriction site polymorphisms permitted identification and cloning of both wild-type alleles of *PHR1*, designated *PHR1-1* and *PHR1-2*. Transformation with the cloned alleles demonstrated that *PHR1-1* preferentially recombined with the resident copy of allele 1 and that *PHR1-2* preferentially recombined with the resident copy of allele 2. The nucleotide sequences of the two alleles were approximately 1% divergent overall and these heterologies were concentrated in the 5' and 3' flanking regions of the transforming DNAs. Transformation with *PHR1* sequences that were identical in both alleles eliminated the recombinational bias. The results suggest that in *C. albicans*, as in other organisms, homeologous recombination is less efficient than homologous recombination and that this imparts allele specificity to targeted integration events.

METHODS

Strains and growth conditions. *Candida albicans* strains CAF3-1 (*ura3Δ/ura3Δ*) and CAS-6 (*ura3Δ/ura3Δ phr1Δ::hisG/PHR1*) were described previously (Fonzi & Irwin, 1993; Saporito-Irwin *et al.*, 1995). Strains YZP1-3 and

YZP1-9 are independent mutants of the same genotype as CAS-6 and were constructed as described for CAS-6 (Saporito-Irwin *et al.*, 1995). The strains were routinely cultured on YPD medium (Sherman, 1991) supplemented with 25 µg uridine ml⁻¹. YNB medium, consisting of 2% glucose and 0.67% Difco yeast nitrogen base, was used in selecting *Uri*⁺ transformants. Medium 199 buffered at pH 8 was used to assess growth and morphological phenotypes as described previously (Saporito-Irwin *et al.*, 1995). Media were solidified with 1.5% agar. Cultures were incubated at 30 °C.

Gene cloning and plasmid constructions. Both alleles of *PHR1* were cloned by hybridization screening of a λGEM-12 genomic library derived from strain SC5314 (Gillum *et al.*, 1984). The library and screening conditions have been described by Birse *et al.* (1993). The 2.1 kb insert from plasmid pSMS-24 (Saporito-Irwin *et al.*, 1995) was used as the hybridization probe to detect *PHR1* sequences. The two alleles were distinguished by a *Bam*HI restriction site polymorphism, which resulted in hybridization bands of 8.4 and 7.4 kb on Southern blots of genomic DNA. λ clones containing the corresponding fragments were isolated and the inserts were subcloned into the *Bam*HI site of plasmid pBSK+ (Stratagene). The alleles contained on the 7.4 and 8.4 kb fragments were designated *PHR1-1* and *PHR1-2*, respectively.

Plasmids pΔPHR1-1::URA3 and pΔPHR1-2::URA3 contained analogous deletions in alleles 1 and 2, respectively. These were constructed by first subcloning a 2.5 kb *Eco*RI-*Eco*RV fragment containing *PHR1* into a modified pBSK+ vector (Stratagene). The vector lacked the *Cl*aI recognition site in the polylinker region, which was destroyed by Klenow polymerase fill-in of the restricted site followed by ligation. The subclones were digested with *Cl*aI to remove nt +397 to +1235 of the *PHR1* coding region and ligated with a 1.1 kb *Cl*aI fragment containing *URA3*. The *URA3* fragment was generated by PCR amplification using the forward primer USP-1 (5'-CCATCGATTGCTGTAGTGCCATTGAT-3') and the reverse primer RP-2 (5'-CCATCGATAAAGTGAA-AATCTCCCCCT-3'). The amplification product encompassed nt -283 to +840, relative to the start codon of *URA3* and incorporated flanking *Cl*aI restriction sites. *URA3* was cloned in the same transcriptional orientation as *PHR1* in both plasmids.

Plasmid pSMS-23 contains a deletion mutation of *PHR1* similar to that of pΔPHR1-1::URA3 and pΔPHR1-2::URA3. It differs in that nt +397 to +1235 of the *PHR1* coding region were substituted with *hisG-URA3-hisG* sequences (Fonzi & Irwin, 1993). Construction of this plasmid was described previously (Saporito-Irwin *et al.*, 1995).

Plasmid pSMS24-URA contained an insertional disruption of *PHR1-2*. This plasmid was constructed by blunt-end ligation of a 1.3 kb *Sca*I-*Xba*I fragment containing *URA3* into the *Bsp*EI site located at nt +889 of the *PHR1* coding region in plasmid pSMS-24 (Saporito-Irwin *et al.*, 1995). The *Bsp*EI and *Xba*I ends were made blunt by Klenow polymerase fill-in.

Transformation and screening. Strains were transformed as described previously (Mühlschlegel & Fonzi, 1997) and *Uri*⁺ transformants were selected on YNB plates. For phenotypic screening of strains CAS-6, YZP1-3 and YZP1-9, transformants were patched to Medium 199 plates buffered to pH 8. These were incubated at 37 °C for 48 h. The resulting colonies were examined macroscopically for the presence of hyphae and microscopically for cell morphology. Colonies lacking filamentation and containing enlarged, highly vacuolated cells were scored as *PHR1* null mutants (Saporito-Irwin *et al.*, 1995). Between 100 and 500 colonies were scored for

each transformation. The results are presented as the fraction of transformants exhibiting the null mutant phenotype.

Genotypic screening of CAF3-1 transformants was performed by Southern blot analysis. The methods for Southern blot preparation and hybridization were described previously (Mühlschlegel & Fonzi, 1997). The 2.1 kb *EcoRI* fragment of *PHR1* (Saporito-Irwin *et al.*, 1995) was used as the hybridization probe. Genomic DNA from cells transformed with p Δ PHR1-1::URA3 or p Δ PHR1-2::URA3 was digested sequentially with *EcoRI* and *MscI*. The *MscI* site is located at position +23 of allele 2, but is absent from allele 1. Replacement of *PHR1-1* was distinguished by the loss of the native 9 kb *EcoRI*–*EcoRI* hybridizing band and acquisition of a 1.21 kb *EcoRI*–*EcoRI* band due to the proximal *EcoRI* introduced by integration of URA3. The loss of the native 8.7 kb *MscI*–*EcoRI* band and acquisition of a 0.9 kb *MscI*–*EcoRI* band distinguished replacement of *PHR1-2*.

Genomic DNA from cells transformed with pSMS24-URA was digested with *MscI* and *PvuII*. Replacement of *PHR1-1* was distinguished by loss of a native 3.2 kb *MscI*–*PvuII* hybridization band and acquisition of a 2.8 kb *MscI*–*MscI* band due to insertion of a proximal *MscI* site present in URA3. Loss of a native 1.4 kb *MscI*–*PvuII* band and acquisition of a 0.99 kb *MscI*–*MscI* band distinguished replacement of *PHR1-2*.

RESULTS

Selective integration at the *PHR1* locus

In previous genetic studies the *PHR1* locus was mutated by gene replacement with a *HindIII*–*PvuII* fragment derived from plasmid pSMS-23 (Saporito-Irwin *et al.*, 1995). This fragment contained 660 bp of *PHR1* sequence 5' and 185 bp 3' of a *hisG*–URA3–*hisG* insertion. The *hisG*–URA3–*hisG* cassette replaced 839 bp of the *PHR1* ORF between nt +397 and +1236. Heterozygous mutants resulting from the replacement of one allele of *PHR1* were readily obtained (Saporito-Irwin *et al.*, 1995). Limited attempts to replace the remaining functional allele of the heterozygous mutant with this same *HindIII*–*PvuII* fragment were unsuccessful due to preferential recombination with the disrupted allele (unpublished data). The magnitude of this preference was examined by transforming heterozygous mutants and using a phenotypic assay to discriminate between recombination with the functional or the disrupted allele. If the transforming DNA recombined with the disrupted allele, the strain would retain a functional copy of *PHR1* and exhibit a wild-type phenotype. Recombination with the functional allele would result in a homozygous null mutant, which exhibits pH-dependent morphological defects. *PHR1* null mutants are normal when cultured at pH < 6.0, but fail to form hyphae and produce enlarged, rounded and highly vacuolated cells when cultured at alkaline pH (Saporito-Irwin *et al.*, 1995). Strain CAS-6 (*phr1 Δ ::hisG/PHR1*) was transformed with the *HindIII*–*PvuII* fragment from pSMS-23 and the Uri⁺ transformants were patched to Medium 199, pH 8.0. After growth at the restrictive pH, the phenotype of the cells was examined. Of the 147 transformants scored, only three exhibited the null mutant phenotype, a frequency of 2.0×10^{-2} . A very

similar allelic preference was observed in transformations of two independent heterozygous strains YZP1-3 and YZP1-9, which yielded mutant frequencies of 3.3×10^{-2} and 1.4×10^{-2} , respectively. These frequencies are minimal estimates since there is no phenotypic distinction between replacement of the disrupted allele and non-homologous integration at other loci, both of which would behave as wild-type. However, non-homologous integration of this DNA fragment is rare (Saporito-Irwin *et al.*, 1995), so failure to account for these events introduces only a minor error in the calculated frequencies. This highly biased targeting also cannot be attributed to selective recovery of wild-type cells, since the transformants were selected at acidic pH, which is permissive for growth of *PHR1* null mutants (Saporito-Irwin *et al.*, 1995). Thus we conclude that the exogenous transforming DNA preferentially recombines with the disrupted allele and only rarely replaces the wild-type allele.

Effect of *hisG* sequences on recombination

Preferential targeting of the disrupted allele might reflect the simultaneous presence of *hisG* sequences in the disrupted allele and the exogenous transforming DNA. As a consequence, the disrupted allele has homology with the exogenous DNA that the wild-type allele does not. By transforming with DNA lacking these sequences the effect of *hisG* on recombinational preference was assessed. Plasmid p Δ PHR1-2::URA3 was analogous to pSMS-23 used in the previous experiment in that both had deletions of the same *PHR1* sequences and, as discussed subsequently, both were derived from the same allele of *PHR1*. However, instead of a *hisG*–URA3–*hisG* insert, p Δ PHR1-2::URA3 contained an insert of URA3 alone. Digestion of p Δ PHR1-2::URA3 with *PstI* and *PvuII* released a DNA fragment in which URA3 was flanked by *PHR1* sequences essentially identical to those of the *HindIII*–*PvuII* digest of pSMS-23. The fragments differed in only a few nucleotides at the 5' end resulting from cleavage in the polylinker of the plasmid. Despite the absence of *hisG* sequences, preferential recombination with the disrupted allele was still evident (Table 1). The median frequency of mutant transformants was 15×10^{-2} , a roughly sixfold increase in comparison with the previous transformations using *hisG*-containing DNA. However, the disrupted allele was still targeted at a sixfold higher frequency than the wild-type allele. This indicates that the simultaneous presence of *hisG* sequences in the disrupted allele and on the transforming DNA does have a selective influence on recombination. However, it does not fully account for the allelic preference of recombination at the *PHR1* locus.

Inherent allele-specific recombination

Unlike the parental strain in which both alleles of *PHR1* are wild-type, the alleles in the heterozygous mutant are distinguished by the presence of the *hisG* insert within the disrupted allele. This creates an inherent structural asymmetry at the locus and, perhaps, in recombination

Fig. 1. Allele specific recombination of transforming DNA with *PHR1-2*. Genomic DNA from CAF3-1 (lane 1) and 11 independent transformants (lanes 2–12) was digested with *Bam*HI and examined by Southern blot hybridization with *PHR1*. The sizes of the wild-type bands are indicated on the left; the 7.4 and 8.4 kb bands correspond to alleles *PHR1-1* and *PHR1-2*, respectively.

events. To assess whether the insert per se was affecting recombination, allele targeting was examined in the parental strain. The alleles of *PHR1* can be distinguished on Southern blots by virtue of a *Bam*HI restriction site polymorphism (Fig. 1). Strain CAF3-1 was transformed with the *Hind*III–*Pvu*II fragment from pSMS-23 and the genotype of 11 *Uri*⁺ transformants was determined by Southern blot analysis. In all 11 isolates the 8.4 kb band, corresponding to allele 2, was absent and replaced with a fragment of approximately 12 kb (Fig. 1). Thus, the transforming DNA was preferentially recombining with *PHR1-2*. Furthermore, Southern blot analysis of the heterozygous strains CAS-6, YZP1-3 and YZP1-6 demonstrated that *PHR1-2* was the allele disrupted in these strains (data not shown). Therefore, the preferential targeting of the disrupted locus in the transformation of these strains was also due to preferential recombination with *PHR1-2*. Together these results indicated that *PHR1-2* recombined at a higher frequency than *PHR1-1* and that this was independent of the presence or absence of an insert within the allele.

The cause of this asymmetry could lie with either or both reactants in the recombination event, the resident

genomic DNA or the exogenous transforming DNA. *PHR1-1* and *PHR1-2* may have intrinsic differences in their ability to recombine, specificity may be imparted by some attribute of the transforming DNA or specificity may be an interdependent effect. Examination of restriction site polymorphisms demonstrated that the *PHR1* sequences in plasmid pSMS-23 were derived from allele 2 (data not shown). Thus, in the preceding experiments the transforming DNA derived from allele 2 was preferentially recombining with the resident allele 2. This suggested that specificity was interdependent. If so, then transforming DNA derived from *PHR1-1* should preferentially recombine with allele 1 in the genome.

By transforming with identical DNA fragments derived from either *PHR1-1* or *PHR1-2* the relationship between the allelic source of the transforming DNA and the allele specificity of recombination was tested. Plasmids pΔ*PHR1-1*::*URA3* and pΔ*PHR1-2*::*URA3* were identically constructed, except that the *PHR1* sequences were derived from the cloned allele 1 or 2, respectively. In the preceding experiment, transformation of the heterozygous disruptants with a *Pst*I–*Pvu*II digest of pΔ*PHR1-2*::*URA3* resulted in preferential recombination with the disrupted allele as indicated by the low frequency of null mutants (Table 1). The analogous transformation using pΔ*PHR1-1*::*URA3* DNA yielded the reciprocal results; the mutant frequency increased to 81×10^{-2} , indicating preferential replacement of the functional allele. Since the functional allele was *PHR1-1* in these strains, pΔ*PHR1-1*::*URA3* DNA selectively recombined with the corresponding allele in the genome. The magnitude of this bias was similar to that of pΔ*PHR1-2*::*URA3* DNA for allele 2. Several other restriction enzyme digests, which altered the length of the flanking sequences, were tested in the transformations and in each case the allele specificity of integration was maintained (Table 1).

Table 1. Effect of allelic source, fragment length and sequence divergence on recombinational preference

Transforming DNA	Allelic source*	Restriction fragment†	Mutant frequency‡	Length of flanking sequences§		No. of base pair substitutions		Percentage sequence divergence	
				5'	3'	5'	3'	5'	3'
pΔ <i>PHR1-1</i> :: <i>URA3</i>	1	<i>Pst</i> I/ <i>Pvu</i> II	0.81	660	185	16	5	2.4	2.7
pΔ <i>PHR1-2</i> :: <i>URA3</i>	2		0.15	659	185	16	5	2.4	2.7
pΔ <i>PHR1-1</i> :: <i>URA3</i>	1	<i>Sac</i> II/ <i>Pvu</i> II	0.75	232	185	3	5	1.3	2.7
pΔ <i>PHR1-2</i> :: <i>URA3</i>	2		0.19	232	185	3	5	1.3	2.7
pΔ <i>PHR1-1</i> :: <i>URA3</i>	1	<i>Bsa</i> BI/ <i>Ahd</i> I	0.72	86	78	0	2	0	2.6
pΔ <i>PHR1-2</i> :: <i>URA3</i>	2	<i>Bsa</i> BI/ <i>Bsa</i> BI	0.03	86	71	0	1	0	1.4

* Indicates which allele of *PHR1* was used in constructing the plasmid.

† The restriction endonucleases used to release the transforming fragment are indicated.

‡ The median frequency determined from one or more transformations of CAS-6, YZP1-3 and YZP1-9 is indicated, except for the *Sac*II/*Pvu*II fragments for which one transformation of CAS-6 is indicated.

§ The base pairs of the *PHR1* sequence either side of the *URA3* marker are indicated.

Fig. 2. Distribution of sequence heterologies at the *PHR1* locus. A restriction map of the region sequenced from both alleles is shown. The vertical bars below the map indicate the location of nucleotide differences between *PHR1-1* and *PHR1-2*. The rectangle indicates the coding region of *PHR1* and flanking regions are indicated by the solid line. The dashed line indicates vector sequence. The hatched region between the *ClaI* sites indicates the region replaced with either *hisG-URA3-hisG* or *URA3* in the transforming DNAs. The asterisk indicates the site of *URA3* insertion in pSMS24-*URA*. Polymorphic restriction sites are indicated by the particular allele containing the site.

To ensure that these results were not an artefact of the locus asymmetry in the heterozygous mutants, DNA from allele 1 or 2 was transformed into the parental strain CAF3-1. Plasmid pΔ*PHR1*-1::URA3 was digested with *BsaBI* and *AhdI*. The recognition sites for these enzymes are located at positions +313 and +1306, respectively. Plasmid pΔ*PHR1*-2::URA3 was digested with *BsaBI* alone since two recognition sites are present, one at +313 and the other +1300. The 3' ends differ due to a restriction site polymorphism. The transformants were screened by Southern blot hybridization using an *MscI* restriction site polymorphism to distinguish which allele had recombined. Of 24 isolates recovered from transformation with pΔ*PHR1*-1::URA3 DNA, 22 were disrupted in allele 1. Conversely, 15 of 16 transformants recovered from transformation with pΔ*PHR1*-2::URA3 DNA were disrupted in allele 2. Therefore, exogenous DNA derived from allele 1 preferentially recombines with allele 1 of the genome and the inverse preference is observed for allele 2, irrespective of the presence or absence of a non-homologous insert within one of the alleles.

Allele-specific recombination due to sequence heterology

One obvious explanation for the self-preference of allele targeting lies in the diverged nucleotide sequence of the two alleles as reflected in their multiple restriction site polymorphisms (Fig. 2). As a consequence, recombination between the exogenous DNA and its genomic counterpart is a homologous event, but recombination with the alternate allele would be a homeologous event. Homeologous recombination is less efficient than homologous recombination (Chen & Jinks-Robertson, 1998; Elliott *et al.*, 1998; Nassif & Engels, 1993; Stambuk & Radman, 1998) and this could account for the bias in gene replacement.

The nucleotide sequence of both alleles was determined for a region extending from –263 bp upstream to 119 bp downstream of the *PHR1* coding region. Comparison of alleles 1 and 2 over this 2028 bp region identified 24 nt differences, including 18 transitions, 5

Table 2. Sequence heterology of *PHR1* alleles

Location*	Nucleotide		Mutation class
	Allele 1	Allele 2	
–225	C	T	Transition
–151	C	A	Transversion
–11	G	C	Transversion
–9	A		Deletion
+26	T	C	Transition
+27	T	C	Transition
+72	T	G	Transversion
+78	T	C	Transition
+85	A	G	Transition
+90	G	A	Transition
+94	A	G	Transition
+123	C	T	Transition
+150	T	C	Transition
+183	T	C	Transition
+255	T	C	Transition
+286	G	A	Transition
+519	C	T	Transition
+540	C	T	Transition
+1245	A	G	Transition
+1308	C	T	Transition
+1323	T	G	Transversion
+1329	T	C	Transition
+1330	G	C	Transversion
+1501	A	G	Transition

* Nucleotide positions are numbered relative to the A of the translational start codon, which is designated +1.

transversions and a 1 bp deletion (Table 2). Seven of the 20 nt differences within the coding region were associated with amino acid substitutions (Table 3). Heterologies were unevenly distributed and concentrated in the 5' and 3' ends of the gene (Fig. 2). In relation to the transforming DNA fragments, the number of diverged nucleotides in the 5'*PHR1* sequences ranged from 16 in the *PstI*–*PvuII* digests to none in the *BsaBI* digests (Table

Table 3. Substitution mutations in *PHR1*

Nucleotide location	Codon no.	Codon		Amino acid	
		Allele 1	Allele 2	Allele 1	Allele 2
+26–27	9	GTT	GCC	V	A
+90	29	ATT	GTT	I	V
+94	32	ATT	GTT	I	V
+286	96	GCT	ACT	A	T
+1330	444	GAG	CAG	E	Q
+1501	501	ACT	TCT	T	A

1). The 3' flanking sequences contained from as many as 5 to as few as 1 diverged nucleotides (Table 1).

If this heterology was responsible for the observed allele-specific gene replacement, then transforming DNA derived from a region 100% identical between the two alleles should recombine with either allele at equal frequency. The previous plasmids could not be used to test this prediction since one of the heterologous sites was only 6 bp from the marker insertion site. Consequently, plasmid pSMS24-URA was constructed. Digestion of this plasmid with *BsaBI* and *NheI* released a fragment in which the 153 bp 5' and 202 bp 3' of the marker gene were identical in both alleles. Strain CAF3-1 was transformed with the pSMS24-URA DNA and the genotype of the Uri⁺ transformants were determined by Southern blot analysis. Of 26 transformants examined, *PHR1-1* was replaced in 14 and *PHR1-2* was replaced in the other 12. Therefore, allele-specific targeting is lost in the absence of sequence heterology.

DISCUSSION

Targeted gene replacements at the *PHR1* locus exhibited a highly biased allele preference. In transformations of heterozygous mutants, there was an approximately 50-fold preference for replacement of the disrupted allele over the wild-type allele. This specificity was a consequence of two factors, the parallel presence of *hisG* sequences in the disrupted allele and the exogenous DNA and the sequence divergence of *PHR1-1* and *PHR1-2*. Increasing the length of the complementary sequences has a positive affect on the frequency of targeted integration in a number of organisms (Hasty *et al.*, 1991; Papadopoulou & Dumas, 1997; Rothstein, 1991) and this may account for the selectivity imparted by *hisG*. However, the frequency of transformation in yeast is also proportional to the number of target sequences available for recombination (Wilson *et al.*, 1994). The presence of the *hisG* sequences in the disrupted allele, in a sense, provides an additional target in this allele since recombination could occur between the *hisG-URA3-hisG* cassette of the exogenous DNA even in the absence of flanking *PHR1* sequences. Furthermore, because the cassette has two copies of *hisG* this may also increase the opportunity for such recombination events.

The *hisG* insert in the disrupted allele could potentially affect recombination by other mechanisms as well. For instance, the insert increased by >500 bp, the distance between *PHR1* homologous sequences, and may alter the chromatin structure, nucleosome positioning or rate of transcription initiation, any of which might influence recombination frequency (Grigoriev & Hsieh, 1997; Kirkpatrick *et al.*, 1999; Wu & Lichten, 1994). However, the significance of these effects appear minimal since they would be predicted to have a similar effect on recombination with transforming DNA from either *PHR1-1* or *PHR1-2* and this was not the case. Exogenous *PHR1-1* DNA replaced the endogenous *PHR1-1* at a similar frequency in both parental and heterozygous strains.

The effect of allelic sequence divergence, which was approximately 1.2% overall for *PHR1*, presumably reflects the greater efficiency of homologous recombination in comparison with homeologous transformation. Recombination between like alleles was a homologous event while recombination between the diverged alleles was a homeologous event. It is well established that sequence heterology interferes with recombination in prokaryotic and eukaryotic species (Chen & Jinks-Robertson, 1998; Elliott *et al.*, 1998; Nassif & Engels, 1993; Stambuk & Radman, 1998) and this includes targeted gene replacements (Leung *et al.*, 1997; Negritto *et al.*, 1996). In *S. cerevisiae* a single nucleotide difference within a 350 bp region can reduce recombination fourfold (Datta *et al.*, 1997). Reduced recombination is due to interference by the mismatch repair system with heteroduplex formation between diverged sequences (Chen & Jinks-Robertson, 1998; Leung *et al.*, 1997).

The results of this study imply that the mismatch repair system is similarly effective in *Candida albicans*. As shown by the data in Table 1, a single nucleotide difference was sufficient to bias recombination frequencies. In all of these transformations, recombination with the homologous allele was favoured four- to fivefold over recombination with the heterologous allele. The only exception occurred in transformations using the *BsaBI* fragment of pΔ*PHR1-2::URA3*, which contained only a single nucleotide difference from *PHR1-1*. This fragment unexpectedly recombined 30-times more fre-

quently with the homologous allele. Preliminary experiments suggest that this atypical behaviour is a peculiarity of this particular DNA fragment (unpublished results). In these experiments, designed to test reciprocity, a heterozygous mutant was constructed in which *PHR1-1* was disrupted instead of *PHR1-2* as in the present experiments. Transformations of this heterozygote analogous to those in Table 1 yielded quantitatively inverted mutant frequencies as expected. The only exception was again with the *BsaBI* fragment of *pΔPHR1-2::URA3*. The frequency of mutants in these experiments was 14% rather than the expected 97%. However, it should be noted that the mutant frequency increased about fivefold.

Allele-specific gene targeting has been reported for other loci in *C. albicans* (Sentandreu *et al.*, 1998) and this is presumably related to the same factors influencing gene replacement of *PHR1*. Although few loci in *C. albicans* have been examined for heteroallelic differences at the nucleotide level, heterology has been evident at all loci examined. Both alleles of *SAP4*, *ALS2* and *ALS4* have been completely or partially sequenced (Hoyer *et al.*, 1998; Miyasaki *et al.*, 1994) and roughly 0.5% divergence was found for each locus. This compares to the 1.2% divergence of *PHR1* alleles. Variations in the degree and distribution of these heterologies could account for the observed locus to locus variations in allele-specific gene targeting. Of course this does not rule out additional factors. In *C. albicans* the frequency of recombination with exogenous DNA correlates with the transcriptional activity of the targeted locus (Srikantha *et al.*, 1995). Given that *C. albicans* is naturally heterozygous for many recessive mutations (Whelan & Magee, 1981; Whelan *et al.*, 1980; Whelan & Soll, 1982), it would not be surprising to find allelic variations in transcription or perhaps DNA topology, which could influence recombination (Paques & Haber, 1999).

There are both practical and theoretical implications to the results reported here. From a pragmatic perspective, the inability to inactivate both alleles of a locus cannot be taken as *prima facie* evidence that the gene is essential. Since a single cloned allele is typically employed in such experiments, sequence divergence between alleles may greatly reduce recombination and prevent production of the homozygous replacement. Of more biological significance is the implied effect of the mismatch repair system in generating genomic diversity in *C. albicans*. As discussed by Vulić *et al.* (1997), the mismatch repair system serves dual roles in the evolution of genetic diversity, it prevents the accumulation of mutations and inhibits genetic exchange between diverged sequences. In an asexual diploid like *C. albicans* this may have the interesting consequence of fostering allelic divergence. Although the recombination and repair system would limit the rate of accumulation of mutations, once a mutation did occur, mismatch repair would limit interchromosomal recombination and act to maintain heterozygosity. This would promote the accumulation of additional heterozygous mutations, which would

further inhibit interchromosomal recombination. The net effect would be divergence of alleles. Environmental and physiological influences on expression of the mismatch repair system would periodically relax the barrier on chromosomal exchange and allow improved traits to be brought to homozygosity and deleterious mutations to be lost (Vulić *et al.*, 1997). This periodic recombination between diverged alleles may serve a genetic role analogous to sexual recombination in other organisms.

ACKNOWLEDGEMENTS

We wish to thank Yonghong Zhang and Terrie Nghiem for their excellent technical assistance in conducting these studies. This work was supported by Public Health Services grant GM47727 from the National Institutes of Health and the Burroughs Wellcome Fund Scholar Award in Molecular Pathogenic Mycology.

REFERENCES

- Alani, E., Cao, L. & Kleckner, N. (1987). A method for gene disruption that allows repeated use of *URA3* selection in the construction of multiply disrupted yeast strains. *Genetics* **116**, 541–545.
- Birse, C. E., Irwin, M. Y., Fonzi, W. A. & Sypherd, P. S. (1993). Cloning and characterization of *ECE1*, a gene expressed in association with cell elongation of the dimorphic pathogen *Candida albicans*. *Infect Immun* **61**, 3648–3655.
- Boeke, J. D., LaCroute, F. & Fink, G. R. (1984). A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol Gen Genet* **197**, 345–346.
- Chen, W. & Jinks-Robertson, S. (1998). Mismatch repair proteins regulate heteroduplex formation during mitotic recombination in yeast. *Mol Cell Biol* **18**, 6225–6237.
- Datta, A., Hendrix, M., Lipsitch, M. & Jinks-Robertson, S. (1997). Dual roles for DNA sequence identity and the mismatch repair system in the regulation of mitotic crossing-over in yeast. *Proc Natl Acad Sci USA* **94**, 9757–9762.
- Elliott, B., Richardson, C., Winderbaum, J., Nickoloff, J. A. & Jasin, M. (1998). Gene conversion tracts from double-strand break repair in mammalian cells. *Mol Cell Biol* **18**, 93–101.
- Fonzi, W. A. & Irwin, M. Y. (1993). Isogenic strain construction and gene mapping in *Candida albicans*. *Genetics* **134**, 717–728.
- Gillum, A. M., Tsay, E. Y. H. & Kirsch, D. R. (1984). Isolation of the *Candida albicans* gene for orotidine-5'-phosphate decarboxylase by complementation of *S. cerevisiae ura3* and *E. coli pyrF* mutations. *Mol Gen Genet* **198**, 179–182.
- Gorman, J. A., Chan, W. & Gorman, J. W. (1991). Repeated use of *GAL1* for gene disruption in *Candida albicans*. *Genetics* **129**, 19–24.
- Grigoriev, M. & Hsieh, P. (1997). A histone octamer blocks branch migration of a Holliday junction. *Mol Cell Biol* **17**, 7139–7150.
- Hasty, P., Rivera-Perez, J. & Bradley, A. (1991). The length of homology required for gene targeting in embryonic stem cells. *Mol Cell Biol* **11**, 5586–5591.
- Hoyer, L. L., Payne, T. L. & Hecht, J. E. (1998). Identification of *Candida albicans* *ALS2* and *ALS4* and localization of Als proteins to the fungal cell surface. *J Bacteriol* **180**, 5334–5343.
- Kelly, R., Miller, S. M., Kurtz, M. B. & Kirsch, D. R. (1987). Directed

- mutagenesis in *Candida albicans*: One-step gene disruption to isolate *ura3* mutants. *Mol Cell Biol* **7**, 199–207.
- Kelly, R., Miller, S. M. & Kurtz, M. B. (1988). One-step gene disruption by cotransformation to isolate double auxotrophs in *Candida albicans*. *Mol Gen Genet* **214**, 24–31.
- Kirkpatrick, D. T., Wang, Y. H., Dominska, M., Griffith, J. D. & Petes, T. D. (1999). Control of meiotic recombination and gene expression in yeast by a simple repetitive DNA sequence that excludes nucleosomes. *Mol Cell Biol* **19**, 7661–7671.
- Leung, W.-Y., Malkova, A. & Haber, J. E. (1997). Gene targeting by linear duplex DNA frequently occurs by assimilation of a single strand that is subject to preferential mismatch correction. *Proc Natl Acad Sci USA* **94**, 6851–6856.
- Miyasaki, S. H., White, T. C. & Agabian, N. (1994). A fourth secreted aspartyl proteinase gene (SAP4) and a CARE2 repetitive element are located upstream of the SAP1 gene in *Candida albicans*. *J Bacteriol* **176**, 1702–1710.
- Morschhauser, J., Michel, S. & Staib, P. (1999). Sequential gene disruption in *Candida albicans* by FLP-mediated site-specific recombination. *Mol Microbiol* **32**, 547–556.
- Mühlschlegel, F. A. & Fonzi, W. A. (1997). *PHR2* of *Candida albicans* encodes a functional homolog of the pH-regulated gene *PHR1* with an inverted pattern of expression. *Mol Cell Biol* **17**, 5960–5967.
- Nassif, N. & Engels, W. (1993). DNA homology requirements for mitotic gap repair in *Drosophila*. *Proc Natl Acad Sci USA* **90**, 1262–1266.
- Negritto, M. T., Wu, X., Kuo, T., Chu, S. & Bailis, A. M. (1996). Influence of DNA sequence identity on efficiency of targeted gene replacement. *Mol Cell Biol* **17**, 278–286.
- Odds, F. C. (1988). *Candida and Candidosis. A Review and Bibliography*, 2nd edn. London: Baillière Tindall.
- Papadopoulos, B. & Dumas, C. (1997). Parameters controlling the rate of gene targeting frequency in the protozoan parasite *Leishmania*. *Nucleic Acids Res* **25**, 4278–4286.
- Paques, F. & Haber, J. E. (1999). Multiple pathways of recombination induced by double-stranded breaks in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* **63**, 349–404.
- Poulter, R. (1990). Classical methods for the genetic analysis of *Candida albicans*. In *The Genetics of Candida*, pp. 75–123. Edited by D. R. Kirsch, R. Kelly & M. B. Kurtz. Boca Raton, FL: CRC Press.
- Rothstein, R. (1991). Targeting, disruption, replacement, and allele rescue: Integrative transformation in yeast. *Methods Enzymol* **194**, 281–301.
- Saporito-Irwin, S. M., Birse, C. E., Sypherd, P. S. & Fonzi, W. A. (1995). *PHR1*, a pH-regulated gene of *Candida albicans*, is required for morphogenesis. *Mol Cell Biol* **15**, 601–613.
- Sentandreu, M., Elorza, M. V., Sentandreu, R. & Fonzi, W. A. (1998). Cloning and characterization of *PRA1*, a gene encoding a novel pH-regulated antigen of *Candida albicans*. *J Bacteriol* **180**, 282–289.
- Sherman, F. (1991). Getting started with yeast. *Methods Enzymol* **194**, 3–21.
- Srikantha, T., Morrow, B., Schröppel, K. & Soll, D. R. (1995). The frequency of integrative transformation at phase-specific genes of *Candida albicans* correlates with their transcriptional state. *Mol Gen Genet* **246**, 342–352.
- Stambuk, S. & Radman, M. (1998). Mechanism and control of interspecies recombination in *Escherichia coli*. I. Mismatch repair, methylation, recombination and replication functions. *Genetics* **150**, 533–542.
- Vulić, M., Dionisio, F., Taddei, F. & Radman, M. (1997). Molecular keys to speciation: DNA polymorphisms and the control of genetic exchange in enterobacteria. *Proc Natl Acad Sci USA* **94**, 9763–9767.
- Whelan, W. L. & Magee, P. T. (1981). Natural heterozygosity in *Candida albicans*. *J Bacteriol* **145**, 896–903.
- Whelan, W. L. & Soll, D. R. (1982). Mitotic recombination in *Candida albicans*: recessive lethal alleles linked to a gene required for methionine biosynthesis. *Mol Gen Genet* **187**, 477–485.
- Whelan, W. L., Partridge, R. M. & Magee, P. T. (1980). Heterozygosity and segregation in *Candida albicans*. *Mol Gen Genet* **180**, 107–113.
- Wilson, J. H., Leung, W., Bosco, G., Dieu, D. & Haber, J. E. (1994). The frequency of gene targeting in yeast depends on the number of target copies. *Proc Natl Acad Sci USA* **91**, 177–181.
- Wilson, R. B., Davis, D. & Mitchell, A. P. (1999). Rapid hypothesis testing with *Candida albicans* through gene disruption with short homology regions. *J Bacteriol* **181**, 1868–1874.
- Wilson, R. B., Davis, D., Enloe, B. M. & Mitchell, A. P. (2000). A recyclable *Candida albicans* *URA3* cassette for PCR product-directed gene disruptions. *Yeast* **16**, 65–70.
- Wu, T. C. & Lichten, M. (1994). Meiosis-induced double-strand break sites determined by yeast chromatin structure. *Science* **263**, 515–518.

Received 30 March 2000; revised 30 June 2000; accepted 4 July 2000.