

Efficient and rapid identification of *Candida albicans* allelic status using SNP-RFLP

Anja Forche, Musetta Steinbach & Judith Berman

Department of Genetics, Cell Biology and Development, University of Minnesota, Minneapolis, MN, USA

Correspondence: Present address: Anja Forche, Department of Biology, Bowdoin College, 234 Druckenmiller Hall, 6500 College Station, Brunswick, ME 04011, USA. Tel.: +1 207 725 3365; fax: +1 207 725 3405; e-mail: aforche@bowdoin.edu

Received 26 March 2009; revised 22 May 2009; accepted 5 June 2009.
Final version published online 20 July 2009.

DOI:10.1111/j.1567-1364.2009.00542.x

Editor: Richard Calderone

Keywords

Candida albicans; SNP-RFLP; LOH; aneuploidy; genome rearrangement.

Abstract

Candida albicans is the most prevalent opportunistic fungal pathogen in the clinical setting, causing a wide spectrum of diseases ranging from superficial mucosal lesions to life-threatening deep-tissue infections. Recent studies provide strong evidence that *C. albicans* possesses an arsenal of genetic mechanisms promoting genome plasticity and that it uses these mechanisms under conditions of nutritional or antifungal drug stress. Two microarray-based methods, single nucleotide polymorphism (SNP) and comparative genome hybridization arrays, have been developed to study genome changes in *C. albicans*. However, array technologies can be relatively expensive and are not available to every laboratory. In addition, they often generate more data than needed to analyze specific genomic loci or regions. Here, we have developed a set of SNP-restriction fragment length polymorphism (RFLP) (or PCR-RFLP) markers, two per chromosome arm, for *C. albicans*. These markers can be used to rapidly and accurately detect large-scale changes in the *C. albicans* genome including loss of heterozygosity (LOH) at single loci, across chromosome arms or across whole chromosomes. Furthermore, skewed SNP-RFLP allelic ratios are indicative of trisomy at heterozygous loci. While less comprehensive than array-based approaches, we propose SNP-RFLP as an inexpensive, rapid, and reliable method to screen strains of interest for possible genome changes.

Introduction

Candida albicans is the most prevalent opportunistic fungal pathogen in the clinical setting, causing a wide spectrum of diseases ranging from superficial mucosal lesions to life-threatening deep-tissue infections. Although host immunity is a major contributor to the development and progression of disease (Ashman *et al.*, 2004; Tuite *et al.*, 2004; Ashman, 2008), genome plasticity of the eukaryotic pathogen is an important factor in the adaptation of this fungus to its many niches in the warm-blooded host (Fradin *et al.*, 2003; Hube, 2004; Forche *et al.*, 2009). *Candida albicans* is a diploid yeast and, although a parasexual cycle has been defined (Bennett & Johnson, 2003, 2005), no meiosis has been observed and the range of diversity observed in clinical strains is consistent with a predominantly clonal and asexual life cycle (Pujol *et al.*, 1993; Gräser *et al.*, 1996; Forche *et al.*, 1999). Thus, classical segregation analyses (such as those applied to *Saccharomyces cerevisiae*) are not used to study and detect

genetic and genomic changes. The diploid assembly of the *C. albicans* SC5314 genome sequence (Jones *et al.*, 2004; Van het Hoog *et al.*, 2007) revealed a high level of natural heterozygosity, with *c.* 55 000 single nucleotide polymorphisms (SNPs) (*c.* 3% of the 16-Mb genome).

In many *C. albicans* isolates, segmental or whole-chromosome aneuploidies arise in response to environmental stresses (Janbon *et al.*, 1998; Perepnikhatka *et al.*, 1999; Selmecki *et al.*, 2006; Rustchenko, 2007). For example, gain and loss of nutrient assimilation is associated with chromosomal alterations (Rustchenko *et al.*, 1994), and the growth on specific carbon sources can result in the loss of a specific chromosome (Janbon *et al.*, 1998). A large proportion of *C. albicans* strains that acquire resistance to fluconazole, the most commonly used antifungal drug, form a segmental aneuploidy, isochromosome 5L [i(5L)], in which the left arm of Chr5 [i(5L)] is duplicated. *Candida albicans* that acquire i(5L) become highly resistant to the antifungal drug fluconazole due to extra copies of two genes on Chr5L: *ERG11*,

which encodes the drug target and *TAC1*, a transcription factor that activates *CDR1* and *CDR2*, as well as ABC transporters that lower the intracellular drug concentration (Selmecki *et al.*, 2008). *In vitro* studies have detected loss of *i(5L)* (and a reduction in Flu^R) when the drug pressure is relieved (Selmecki *et al.*, 2006, 2008). Increased resistance to antifungal drugs can also arise by point mutations in *ERG11* (White, 1997a, b), *TAC1* (Coste *et al.*, 2004, 2006), or *MRR1* (a transcriptional activator of *MDR1*, a major facilitator efflux pump) (Morschhäuser *et al.*, 2007; Dunkel *et al.*, 2008). Once such mutations arise, loss of allelic variation via recombination causes increased drug resistance by duplicating the hyperactive allele (Coste *et al.*, 2006; Holmes *et al.*, 2006; Dunkel *et al.*, 2008). These studies provide a strong evidence that *C. albicans* possesses an arsenal of genetic mechanisms promoting genome plasticity and that it uses these mechanisms under conditions of nutritional or antifungal drug stress.

Recently two microarray-based technologies have been developed to study genome changes in *C. albicans*. SNP microarrays detect loss of heterozygosity (LOH) for many loci across all eight chromosomes (Forche *et al.*, 2004, 2005, 2008). Comparative genome hybridization (CGH) utilizes microarrays containing virtually all ORFs in the genome, to detect segmental and whole-chromosome aneuploidies (Selmecki *et al.*, 2005). These methods provide comprehensive large-scale analyses of the genome. However, CGH and SNP microarray technology can be relatively expensive and is not available to every laboratory. Furthermore, these approaches generate more data than needed to analyze specific loci or defined regions in the genome, or for checking strains after laboratory manipulations, which can give rise to genome-scale changes (Selmecki *et al.*, 2005; Bouchonville *et al.*, in preparation).

Here we have developed a set of SNP-RFLP [or PCR-restriction fragment length polymorphism (RFLP)] markers for *C. albicans* and show that they can be used to detect LOH and aneuploidy for whole chromosomes and individual chromosome arms across the genome. SNP-RFLP analysis is based on restriction enzyme (RE) analysis of genomic regions that are heterozygous for RE recognition sites (McEwen *et al.*, 2000). While a limited number of SNP-RFLP markers have already been used successfully to analyze the genome status of strains after transformation (Legrand *et al.*, 2008), for population genetics studies (Gräser *et al.*, 1996; Forche *et al.*, 1999; Xu *et al.*, 1999), to analyze progeny after parasexual mating (Forche *et al.*, 2008), or to analyze events on Chr5 (Wu *et al.*, 2005; Coste *et al.*, 2006), here we define a diagnostic set of 32 SNP-RFLP markers, with four markers per chromosome (usually two per chromosome arm). Each SNP includes a single polymorphic RE site that yields digestion products that are readily distinguishable by size. We show that these markers can be used to rapidly and

accurately detect large-scale changes in the *C. albicans* genome, including LOH at single loci, as well as across chromosome arms (concerted LOH of two markers on a single arm) or whole chromosomes (concerted LOH of all four markers for a given chromosome). Furthermore, skews in SNP-RFLP allelic ratios provide an initial indicator of trisomy at heterozygous loci. While less comprehensive than array-based approaches, we propose SNP-RFLP as an inexpensive, rapid, reliable, and interpretable method to screen for chromosome changes in strains of interest.

Materials and methods

Strain maintenance and growth

The strains used in this study are listed in Supporting Information, Table S1. Strains were stored in 50% glycerol at -80°C , and grown and maintained on YPAD medium (1% yeast extract, 1% peptone, 2% glucose, and 1.5% agar). Genomic DNA was prepared as described previously (Selmecki *et al.*, 2005).

Selection of a diagnostic minimal set of SNP-RFLP markers

The two original heterozygous Contig 19 sequences from each SNP marker used for SNP arrays (Forche *et al.*, 2005) were analyzed using NEB webcutter software (<http://tools.neb.com/NEBcutter2/index.php>) to identify REs that would digest at heterozygous loci and yield restriction fragments of different sizes that were easily separable on an agarose gel. For each chromosome, four SNP markers were selected, with the goal of two markers per chromosome arm, so that events involving long tracts of LOH on a single arm could be inferred. Additional criteria for SNP marker choice were that each RE should cut only once within the SNP marker and that a minimal number of REs should be used with the exception of marker 1765/2519 on Chr3, in which there are two RE sites within the marker sequence (Table 1). The current set of markers requires the use of 13 different REs.

Amplification of SNP markers

The diagnostic set of 32 markers was tested using strain SC5314. PCRs were performed in a final volume of 25 μL with 5 μL of e2TAK buffer (Mg^{2+} ; Takara Bio Inc., Japan), 2.5 mM each of dATP, dCTP, dGTP, and dTTP, 10 μM of each primer (Table 1), 0.25 μL of e2TAK polymerase, and 1.0 μL of genomic DNA (30 ng μL^{-1}). PCR conditions were initial denaturation at 98°C for 3 min, 30 cycles each of denaturation at 98°C for 10 s, annealing at 55°C for 10 s, and extension at 72°C for 1 min, and a final extension at 72°C for 5 min. Amplification was verified by analysis of 3 μL of PCR product run in a 1% agarose gel (Seakem LE

Table 1. Diagnostic set of SNP-RFLP markers

Unique ID	SNP-RFLP marker	Chromosome	Contig19 number	Size of the marker (bp)	Primer sequences	RE	Enzyme recognition sequence*	Size of digestion products (bp)	Allele [†]
12	1642/2179	RL	19-10040	231	F: TGATTACCCTTGGGGCATT R: ATGGGTGGCCAAATTTGACG	ApoI	RAAT <u>T</u> Y	None 64 167	a b
1	B13B19	RL	19-10161	956	F: TGCCCAAATGCTTCCGAT R: GAGGTAAGGGTTCAAGTCCA	AseI	ATTA <u>A</u> T	93 863 None	a b
20	2388/2421	RL	19-10238	139	F: GACGTCTCTGTAGATGGGTT R: AAGTCGGTGCCCAAGTAATG	DdeI	CT <u>N</u> AG	None 6277	a b
16	1158/2429	RR	19-10247	187	F: GGAACCTCCTGCGGTCAATTA R: CGGTGCATCGGTAAACGATT	DdeI	CT <u>N</u> AG	71 116 None	a b
23	1799/2450	1L	19-10237	195	F: AGCCAACCATATTCAGGATTGAC R: GTGCCAACTAGTAATGGTTGTCAT	AluI	<u>A</u> CGT	78 115 None	b a
41	2347/2406	1L	19-10241	221	F: CAATATTAGGGACTTTGAGCCCTTC R: TGGGAGTAGTGTAGTCGCATAAGTA	AseI	ATTA <u>A</u> T	101 120 None	a b
32	F12n4	1R	19-10218	938	F: AGCATTGGGTCATCCAATAACGAC R: GTGATGAAGGTCTTACTGAAGTGC	HpaII	CC <u>G</u> GG	363 575 None	? ?
42	2106/2441	1R	19-10216	252	F: GTACTTCTATACACGCACATCTTCA R: GAAATCCACCGCATAAGAAATGGTT	PstI	CTGC <u>A</u> G	114 138 None	b a
60	2051/2483	2L	19-10119	193	F: GCTGTTGATAGGACTGATTGG R: AGGTGGAACACATGAACCAG	TaqI	<u>T</u> CGA	None 83 109	a b
53	1414/2481	2L	19-10135	162	F: TTAGTCGCTGAGCTGAAGCG R: CACAAGTGCTAATGGCAACT	TaqI	<u>T</u> CGA	None 58 104	b a
66	2046/2319	2R	19-10113	256	F: AGAAGGGCAGCCAGAAAT R: GGGCTCTGAATTGGTTTC	DdeI	CT <u>N</u> AG	None 120 136	a b
68	1314/2038	2R	19-10125	268	F: TCGTCGCGATCTCATTCA R: CGCTTAATGCCGGTTGTA	AluI	<u>A</u> GCT	None 63 205	a b
44	2091/2447	3L	19-10150	264	F: CAGCCAAAGAAACGGCTAGGTAT R: AACGGTGAATGAATGCATCGCC	DdeI	CT <u>N</u> AG	52 210 None	a b
76	2195/2207	3L	19-10123	219	F: TCCGATGGGGCATTATTGCT R: GCAGAGGCCAATGAAATTG	TaqI	<u>T</u> CGA	89 130 None	a b
79	1838/2419	3L	19-10046	754	F: CTCTCTTTGCTCTTTGTC R: TGTTCTGGATTTGGTATG	EcoRI	<u>G</u> AATTC	None 373 381	b a
71	1765/2519	3R	19-10236	147	F: ACCATCAGATGAGGTAGGAC R: GGACTCCCTTATATTGTCTAG	Hpy188I	TC <u>N</u> GA	7140 742 98	a b
156	ERG8-1	4L	19-10212	845	F: GTTCAGTCACGCATAAATCC R: TGGGTTTGGACATCAGGTTCAA	TaqI	TC <u>G</u> A	None 148 697	b a
85	2159/2205	4L	19-10109	244	F: CTCGGGAGAATAAGCTTACCATCTG R: TTAATTGTTGGGAAATCTGAACAGC	TaqI	TC <u>G</u> A	95 147 None	b a
96	2074/2337	4R	19-10126	215	F: GGAGAGTGTATCAACAATGAGGTG R: GCTACTACATTTAACACCGCGCT	AluI	<u>A</u> GCT	None 91 122	a b
95	1654/2166	4R	19-10204	156	F: CAACTGGAGGAACTGACAATAGTG R: TCACAAATTCCTTAGGGTGGTTAG	AluI	<u>A</u> GCT	29 125 None	a b
167	10080A	5L	19-10080	175	F: GCGTCATCCCAACCTGCTA R: TCTGAGTTCACAGGTTGCAG	TaqI	TC <u>G</u> A	None 66 109	b a
107	SNF1-4	5L	19-10137	298	F: CACCTGAGAGAGAGGTAAC R: TAGCACTTTAGCCACAGG	Bfal	CT <u>A</u> G	None 53 245	b a
116	2093/2390	5R	19-10194	265	F: TCTCTCTTTGGAGTGAGC R: CAATACGCAACTTCCAG	Bfal	CT <u>A</u> G	84 184 None	b a
177	2222A	5R	19-10194	263	F: AATCCACCAGCTGCTAGACA R: GTGGTAAATACCACTCCA	TaqI	TC <u>G</u> A	Uncut 100 163	a b
122	1578/2118	6L	19-10035	227	F: GTGTGCTGGGTAATACAGCT R: GGCAAGACAATACCACAGT	ApoI	RA <u>A</u> TTY	91 136 None	a b
123	1534/2336	6L	19-10181	268	F: GAGGAGATAGTTGTTGCTGTC R: GTAGGCATTGACTACTGTGC	Bfal	CT <u>A</u> G	122 144 None	a b
127	1870/2315	6L	19-10230	271	F: AGGGAACCATAGAGTAGC R: CTGGCGACTCTCCACGATAA	Hpy188I	TC <u>N</u> GA	None 128 143	b a
132	2114/2404	6R	19-10140	256	F: TCTGTGCGCTGAGCTGTTTA	Sau96I	<u>G</u> GNCC	47 209	a

Table 1. Continued.

Unique ID	SNP-RFLP marker	Chromosome	Contig number	Size of the marker (bp)	Primer sequences	RE	Enzyme recognition sequence*	Size of digestion products (bp)	Allele [†]
136	2397/2496	7L	19–10248	773	R: CCAAACGACCCAAAATACCC F: CAGTCATTCCTCATCTAC	PciI	AC <u>AT</u> GT	None	b
135	2236/2498	7L	19–10248	442	R: CTATTGTGTCGCAGTGGG F: CAACCACGTCAGGACCA	DdeI	CTN <u>AG</u>	250 524	a
139	1491/2436–1	7R	19–10219	319	R: GTGGACTTGGCTGTATAG F: CATAGTCCGCTGACATA	TaqI	TC <u>G</u> A	34 408	b
144	1530/2473–2	7R	19–10253	282	R: TAGAAGACCCGCTTGATG F: ACCACCACAATTGGCTTC	TaqI	TC <u>G</u> A	43 276	b
					R: GATCAGCCCATTGTTGAT			98 184	a

For marker F12n4 (Chr1) no SNP-RFLP haplotype could be assigned because the original trace sequence was homozygous (resequencing revealed polymorphisms for this marker) and cannot be used for distinguishing the two alleles (unless the fragment is cloned and sequenced).

*The polymorphic nucleotide is underlined.

[†]allele (homolog) assignment based on Legrand *et al.* (2008).

RE, restriction enzyme; L, left of centromere; R, right of centromere.

agarose, Lonza, Rockland, ME; in 1 × TBE). Gels were stained with ethidium bromide and photographed.

RE digest of PCR products

Each PCR product was digested overnight with the relevant RE (see Tables 1 and S2–S9) in a total volume of 15 µL with 1 µL RE, 10 × restriction buffer, distilled water, and *c.* 5 µL of the digested PCR product was run on a 3% agarose gel (Seakem LE agarose, Lonza; 1.0 × TBE buffer) along with an undigested control PCR sample. Gels were stained with ethidium bromide and photographed. Genotypes were assigned based on banding patterns for each SNP marker.

Results and discussion

The *C. albicans* genome is quite plastic, exhibiting a broad range of mitotic recombination events (Lephart *et al.*, 2005; Lephart & Magee, 2006; Forche *et al.*, 2008; Legrand *et al.*, 2008) in addition to aneuploidies due to chromosome loss (Janbon *et al.*, 1998) or acquisition of whole chromosomes or chromosome segments (Selmecki *et al.*, 2005). To distinguish these different mechanisms, the minimal number of markers required is four per chromosome, with two markers per chromosome arm. A total of 152 SNP markers used previously on our SNP microarray (Forche *et al.*, 2005, 2008; Legrand *et al.*, 2008; Fig. 1) were screened for the presence of RE sites that digest SNPs within the SNP marker sequence. These SNP markers are on average 200 bp in length and often include several SNPs (Forche *et al.*, 1999, 2004, 2005). Of the 152 SNP markers analyzed, 134 (88%) included one or more SNPs that were recognized by a RE (Tables 2 and S2–S8). Of these, 112 (74%) included unique RE sites (those digested only once within the marker sequence) (Table 2).

We selected a set of 32 markers (eight chromosomes, four per chromosome) from the 112 SNP-RFLP markers with unique RE recognition sites (Fig. 1). PCR amplification of these markers, followed by digestion with the relevant RE, results in three distinguishable fragments: one fragment from the uncut allele and two smaller fragments from the allele that carries the polymorphism recognized by the RE. Strain SC5314, which is heterozygous for all SNP markers, was tested for the correct restriction pattern (data not shown). Thirty-one SNP markers yielded the expected three restriction fragments (one fragment for uncut allele and two fragments for the cut allele), and the restriction digest of marker 1765/2519 (Chr3) yielded the expected four restriction fragments (Fig. 2).

In a previous study, haplotypes for each of the eight pairs of chromosomes for strain SC5314 were assigned based on SNP microarray analysis of strains with existing and induced aneuploidies (Legrand *et al.*, 2008). For nine out of 32 markers, the SNPs analyzed by microarray and by SNP-RFLP were identical (Table 1), and thus the assignment of haplotype identity was straightforward. For marker F12n4 on Chr1, no SNP-RFLP haplotype could be assigned (Table 1). Based on the assumption that original trace sequences for strain SC5314 were *c.* 400 bp in length, and based on the fact that array SNPs and RFLP SNPs within a specific marker were on an average no further apart than *c.* 80 bp, each SNP (nucleotide) analyzed by SNP-RFLP was matched up with the appropriate homolog (a or b) containing the SNP (nucleotide) present on the SNP array (Legrand *et al.*, 2008). Based on the restriction digest results (cut or uncut) haplotypes were assigned for the remaining 21 SNP-RFLP markers (Table 1).

By screening strains of interest, we asked whether the SNP-RFLP markers would detect known genotype changes

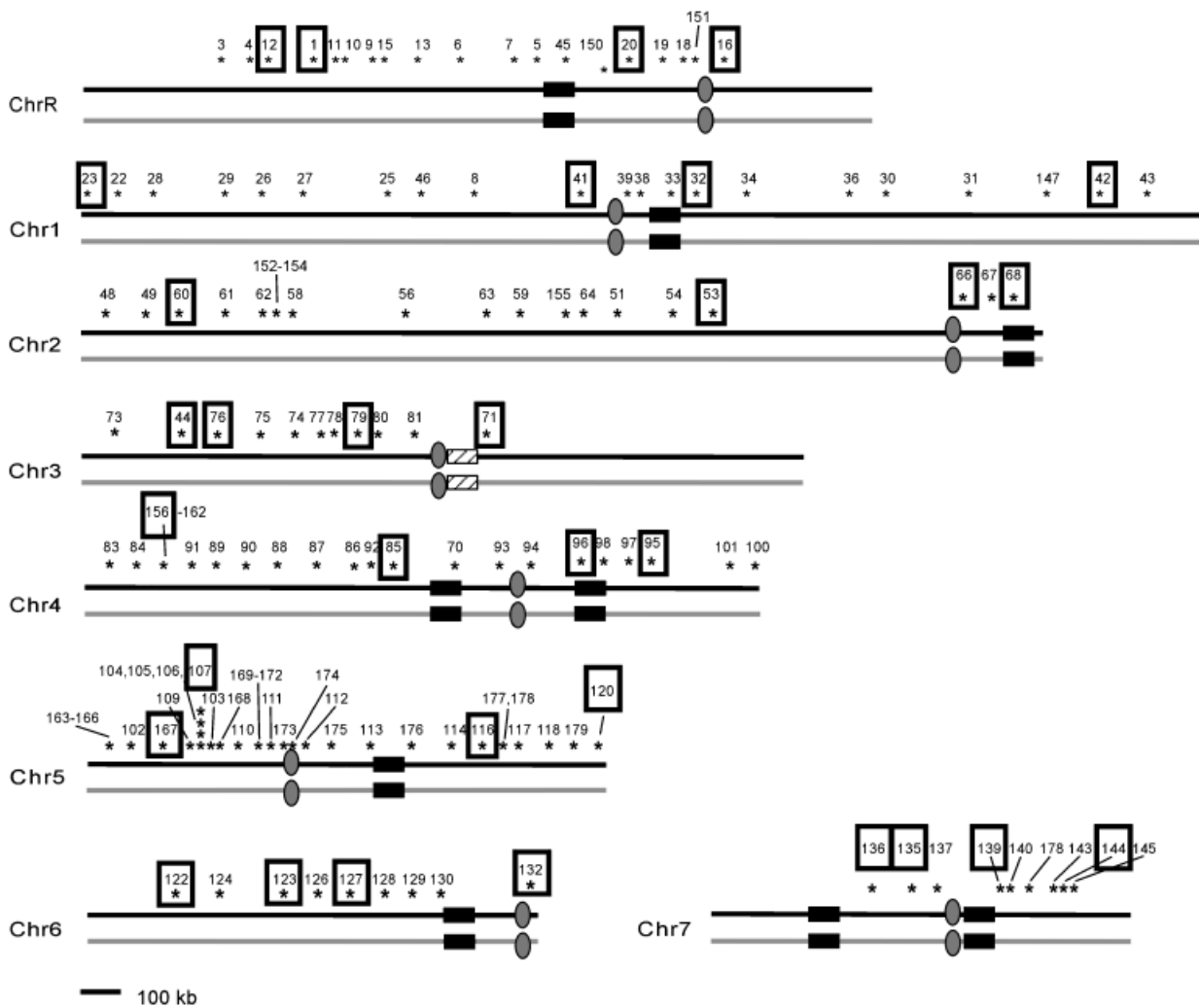


Fig. 1. Whole genome map of SNP-RFLP markers. Homolog 'a' is colored in black, homolog 'b' is colored in gray, and major repeat sequences (MRS) are indicated as black-filled boxes except for Chr3 (RB2 only; hashed box). The centromere is indicated as a gray oval. SNP-RFLP markers from the diagnostic set are boxed in black. Stars indicate the location of individual markers. Numbers above stars indicate the unique identifier number of each marker. Centromere and MRS are not to scale.

Table 2. Summary of SNP-RFLP markers sorted by chromosome

Chromosome	Number of SNP markers	Number of SNP-RFLP markers	Number of SNP-RFLP markers with single-cut enzymes
R	19	18	15
1	26	24	21
2	21	16	15
3	11	11	8
4	22	20	17
5	31	26	21
6	11	9	8
7	11	9	7
Total	152	134 (87.5%)	112 (74%)

including LOH (the appearance of only a single uncut band or the two cut bands) at a single locus, LOH of a chromosome arm, or LOH at all loci across an entire chromosome.

In addition, we asked whether trisomy of a chromosome arm or of an entire chromosome could be detected as a change in the relative ratio of the uncut band to the two cut bands (Legrand *et al.*, 2008; Selmecki *et al.*, 2008).

As a proof of principle, we performed SNP-RFLP analysis of strains known to have genotypic alterations when compared with their parental genotype profiles (Table S1, Fig. 2). For example, strain YJB10019 (derived from SC5314) was known from a previous study (Forche *et al.*, 2008) to have undergone changes on multiple chromosomes including a single LOH event on Chr1, a chromosome arm LOH event on Chr2, and a whole Chr LOH on ChrR. Along with strain SC5314, YJB10019 was subjected to SNP-RFLP analysis of markers from chromosomes R, 1, and 2. Figure 3 shows the results of the restriction digests. Strain SC5314 was heterozygous for all four markers on ChrR, Chr1, and Chr2,

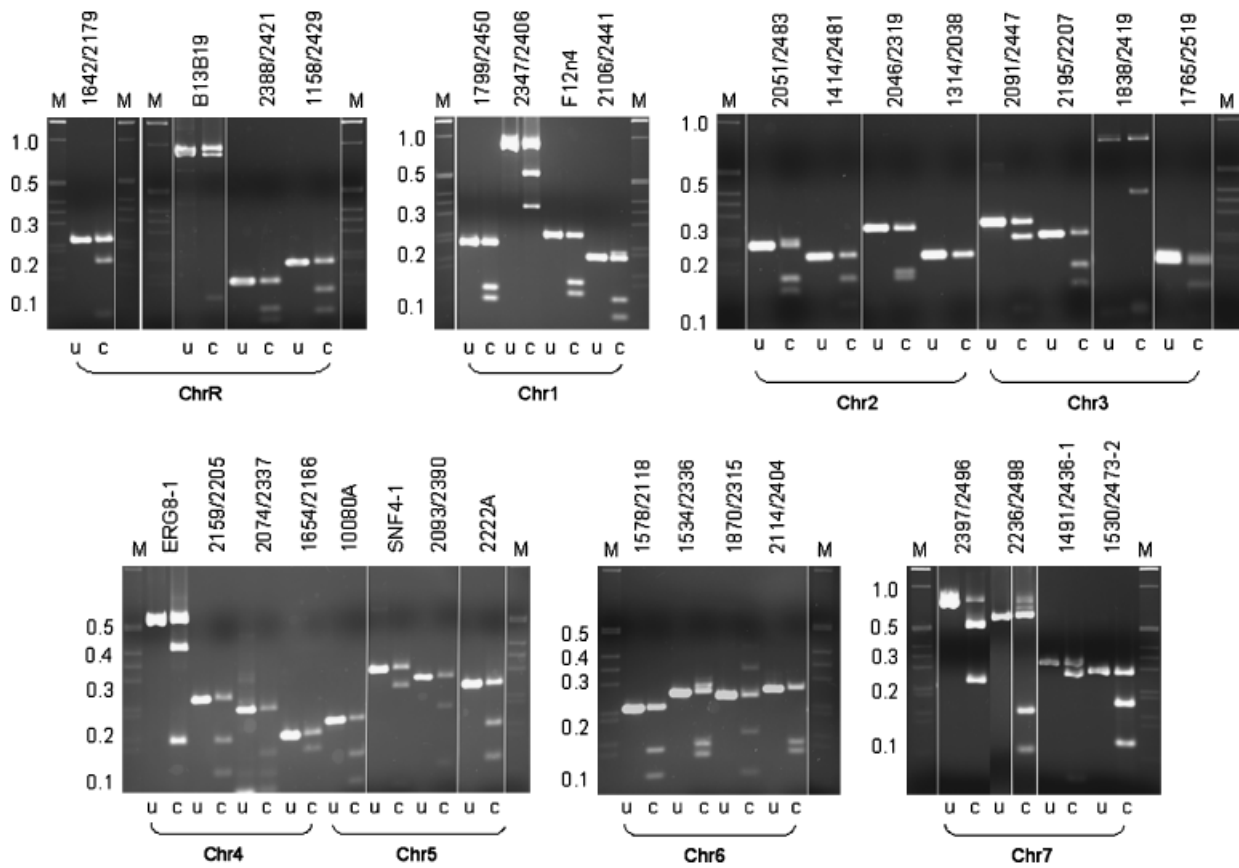


Fig. 2. Use of the diagnostic SNP-RFLP marker set confirms the heterozygosity of all 32 markers in strain SC5314. The chromosome is indicated below each set of gel panels. For each digest, the uncut (u) and cut (c) PCR products are shown. Order of the SNP markers corresponds to their position on the chromosome from left to right. M, molecular size marker; fragment sizes on the left are shown in kb. Precise SNP marker location and RE fragment sizes are detailed in Table 1.

respectively (Fig. 3a, b, c, left images). In strain YJB10019, a single LOH event on Chr1 was detected at the telomere-proximal 1799/2450 marker (Fig. 3a, right image). As expected, the Chr2 arm event spanned both markers on Chr2L (2051/2483 and 1414/2481) (Fig. 3b, right image). The most extensive LOH event was observed for ChrR, where all four SNP markers were homozygous. This indicates that LOH occurred across most parts or the entire chromosome (Fig. 3c, right image). Whole-chromosome LOH most likely arises via one or more chromosome nondisjunction events.

Aneuploidy, especially trisomy, is frequently observed in drug-resistant *C. albicans* isolates (Selmecki *et al.*, 2006, 2008). Array CGH (aCGH) (Selmecki *et al.*, 2005) is the most comprehensive method for detecting aneuploidies. However, aCGH can be expensive and requires microarray technology that may not be readily available to all labs. Because trisomy results in skewed allelic ratios (Legrand *et al.*, 2008), we asked whether it is possible to detect a skewed allele copy number by SNP-RFLP. For this experiment, we chose two sets of

strains that had included chromosomes with segmental or whole-chromosome aneuploidies.

We used a well-documented set of strains derived, by transformation, from CAF-2 (diploid). These strains had become trisomic for Chr2 (CAI4-F2 and CAI4-F3) (Selmecki *et al.*, 2005). The CAI4-F3 strain also became trisomic for Chr1 (Chen *et al.*, 2004; Selmecki *et al.*, 2005). SNP-RFLP analysis was carried out for four markers of Chr1 and Chr2 for the parental strain CAF-2 and for the two versions of CAI4. Strain CAF-2 was heterozygous for all four markers on Chr1 and Chr2 with relative band intensities appropriate for 1:1 amounts of the two different alleles (Fig. 3d and e, left images). In contrast, in strain CAI4-F3, allele ratio bias was evident for four markers on Chr2 (Fig. 3d, right image) and for one marker on Chr1 (Fig. 3e, right image). Furthermore, as expected, markers on Chr2, but not on Chr1, exhibited skewed allele ratios for strain CAI4-F2 (Fig. 3d and e, middle images). This shows that SNP-RFLP analysis generates data consistent with the idea that these chromosomes are trisomic in the strain. When

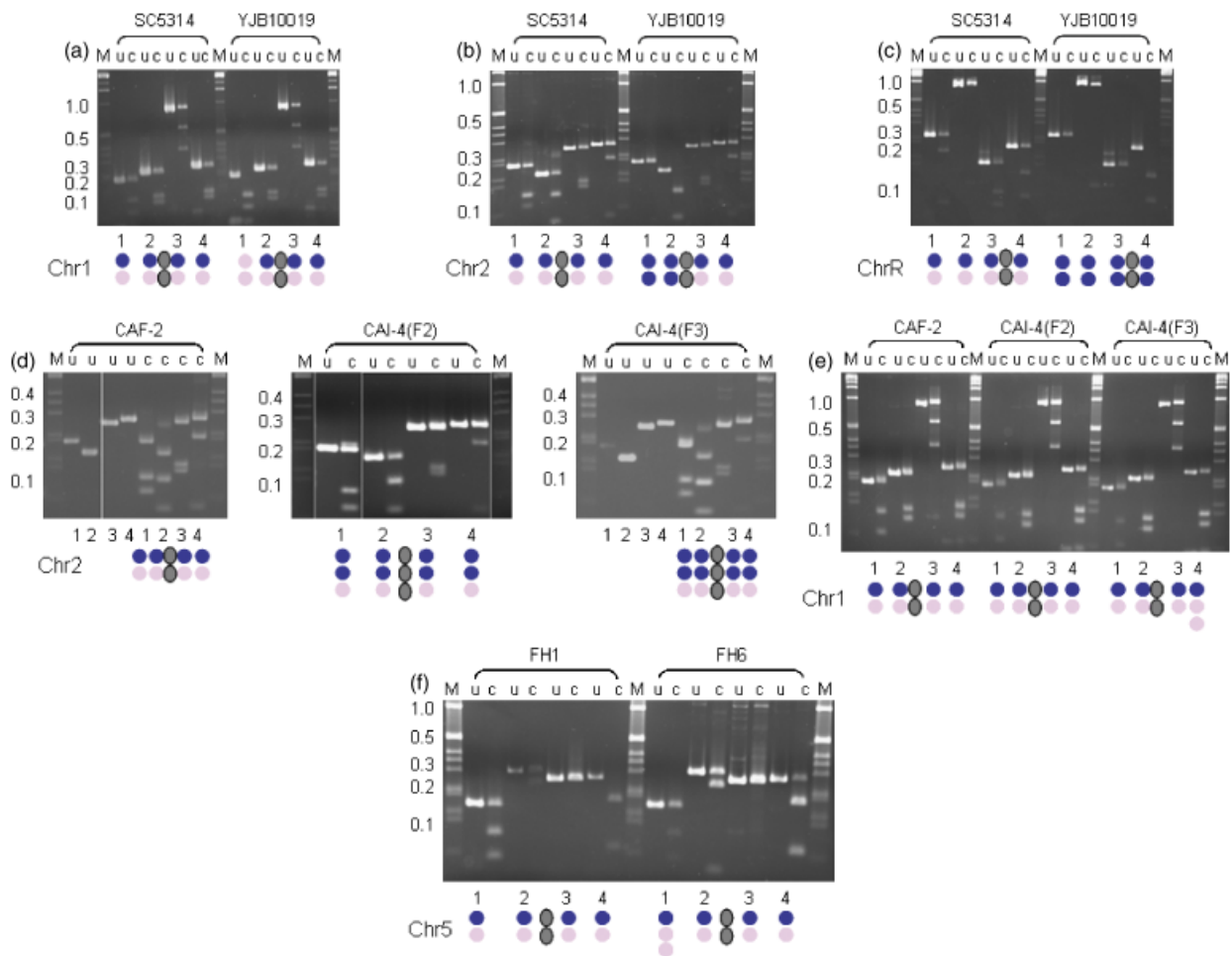


Fig. 3. Detection of LOH and trisomy by SNP-RFLP analysis. Each gel is labeled as in Fig. 2, with SNP markers aligned from left to right along the relevant chromosome. The uncut control (u) lane is shown for each cut (c) PCR product. Colored circle cartoons illustrate the LOH events for each strain. Blue indicates allele 'a' and pink indicates allele 'b.' Two vertical circles represent one locus. Gray ovals represent centromeres. (a) An LOH occurred at the left most SNP marker on Chr1L in strain YJB10019. (b) LOH of both markers on Chr2L in strain YJB10019. (c) Whole chromosome LOH is detected as LOH of all four SNP markers on ChrR in strain YJB10019. (d) Skewed ratios of the two alleles of Chr2 suggest trisomy of Chr2 in CAI4-F2 and CAI4-F3 strains. Allele 'a' is present in two copies and allele 'b' is present in one copy (see also Table 1 for restriction patterns and allele assignments). (e) Skewed ratios are seen for only one of the four markers on Chr1 in strain CAI4-F3. (f) Of the two heterozygous SNPs present in both clinical strains FH1 and FH6 (markers 1 and 2) marker 2 has a skewed allelic ratio, consistent with the idea that Chr5L is trisomic in strain FH6, but not in parental strain FH1. M, DNA size marker; u, undigested; c, cut; 1–4, SNP markers from left to right on the chromosome.

multiple markers on a single chromosome exhibit a skewed allelic ratio, it provides more confidence in the idea that the chromosome may be aneuploid. Of course, aCGH would be necessary to confirm this hypothesis.

We next compared strains FH1 and FH6 (White, 1997a, b; Selmecki *et al.*, 2008). From aCGH, it was known that FH6 carries i(5L), a segmental aneuploidy in which there are two extra copies of the left arm of chromosome 5, and SNP array analysis had found that markers on Chr5L were present in a 2:1 ratio (Selmecki *et al.*, 2008). It is important to note that FH1 and FH6 are clinical isolates and likely possess a different SNP distribution across their genome. Thus, although we analyzed all four diagnostic SNP-RFLP markers from Chr5,

only two of the Chr5 markers were heterozygous for both FH1 and FH6 (Fig. 3f). Similarly, when markers are not heterozygous, it is difficult to infer changes in ploidy with high confidence. For example, while a skewed allelic ratio (2:1) for the first marker (10080A) on Chr5L can be detected for strain FH6 (Fig. 3f), and this is consistent with trisomy of Chr5L detected in this strain (Selmecki *et al.*, 2008), no skewed allelic ratio (2:1) can be detected for strain FH6 for the second marker (SNF1-4) on Chr5L. Marker 2222A on Chr5R was homozygous in both strains and is thus not informative with regard to alterations in the genome (Fig. 3f).

In conclusion, here we describe a practical, inexpensive, and simple approach to determine whether chromosomes

underwent LOH using a diagnostic set of SNP-RFLP markers to detect LOH events in the *C. albicans* genome. This method provides the ability to distinguish short-range events that generate LOH of a single SNP, LOH events that involve multiple SNPs on a chromosome arm as well as LOH of SNPs spanning entire chromosomes. Combined with our HapMap (Legrand *et al.*, 2008), haplotype exchanges can also be detected. Furthermore, this SNP-RFLP approach can be used as a preliminary test for trisomy of chromosomes that remain heterozygous. Trisomy is more difficult to detect and thus the absence of a skewed ratio of RE fragment alleles is not definitive. However, the presence of a skewed ratio at multiple markers on a chromosome provides more confidence in the interpretation of a potential trisomic chromosome.

A major advantage is that this diagnostic set of SNP-RFLP markers provides a rapid and accurate method to detect genomic changes on all eight chromosomes after strain manipulations such as transformation, which can lead to increased levels of aneuploidy and/or LOH (Bouchonville *et al.*, in preparation). The use of two SNP markers per chromosome arm reveals distinct mechanisms of LOH. Importantly, this method relies on simple techniques available in all molecular biology labs, enabling the generation of data that is easily compared between research groups. In this regard, it should be considered a simple, rapid, and accessible alternative to SNP microarrays, and MLST for testing strain integrity.

Acknowledgement

This work was supported by grant RO1-AI062427 to J.B. and A.F. from the National Institutes of Health/National Institute of Allergy and Infectious Diseases.

References

- Ashman RB (2008) Protective and pathogenic immune responses against *Candida albicans* infection. *Front Biosci* **13**: 3334–3351.
- Ashman RB, Farah CS, Wanasaengsakul S, Hu Y, Pang G & Clancy RL (2004) Innate versus adaptive immunity in *Candida albicans* infection. *Immunol Cell Biol* **82**: 196–204.
- Bennett RJ & Johnson AD (2003) Completion of a parasexual cycle in *Candida albicans* by induced chromosome loss in tetraploid strains. *EMBO J* **22**: 2505–2515.
- Bennett RJ & Johnson AD (2005) Mating in *Candida albicans* and the search for a sexual cycle. *Annu Rev Microbiol* **59**: 233–255.
- Chen X, Magee BB, Dawson D, Magee PT & Kumamoto CA (2004) Chromosome 1 trisomy compromises the virulence of *Candida albicans*. *Mol Microbiol* **51**: 551–565.
- Coste AT, Karababa M, Ischer F, Bille J & Sanglard D (2004) *TAC1*, transcriptional activator of CDR genes, is a new transcription factor involved in the regulation of *Candida albicans* ABC transporters *CDR1* and *CDR2*. *Eukaryot Cell* **3**: 1639–1659.
- Coste AT, Turner V, Ischer F, Morschhauser J, Forche A, Selmecki A, Berman J, Bille J & Sanglard D (2006) A mutation in *Tac1p*, a transcription factor regulating *CDR1* and *CDR2*, is coupled with loss of heterozygosity at chromosome 5 to mediate antifungal resistance in *Candida albicans*. *Genetics* **172**: 2139–2156.
- Dunkel N, Blass J, Rogers PD & Morschhäuser J (2008) Mutations in the multi-drug resistance regulator *MRR1*, followed by loss of heterozygosity, are the main cause of *MDR1* overexpression in fluconazole-resistant *Candida albicans* strains. *Mol Microbiol* **69**: 827–840.
- Forche A, Schönian G, Gräser Y, Vilgalys R & Mitchell TG (1999) Genetic structure of typical and atypical populations of *Candida albicans* from Africa. *Fungal Genet Biol* **28**: 107–125.
- Forche A, Magee PT, Magee BB & May G (2004) Genome-wide single-nucleotide polymorphism map for *Candida albicans*. *Eukaryot Cell* **3**: 705–714.
- Forche A, May G & Magee PT (2005) Demonstration of loss of heterozygosity by single-nucleotide polymorphism microarray analysis and alterations in strain morphology in *Candida albicans* during infection. *Eukaryot Cell* **4**: 156–165.
- Forche A, Alby K, Schaefer D, Johnson AD, Berman J & Bennett RJ (2008) The parasexual cycle in *Candida albicans* provides an alternative pathway to meiosis for the formation of recombinant strains. *PLoS Biol* **6**: e110.
- Forche A, Magee PT, Selmecki A, Berman J & May G (2009) Evolution in *Candida albicans* populations during single passage through a mouse host. *Genetics*, DOI: 10.1534/genetics.109.103325.
- Fradin C, Kretschmar M, Nichterlein T, Gaillardin C, d'Enfert C & Hube B (2003) Stage-specific gene expression of *Candida albicans* in human blood. *Mol Microbiol* **47**: 1523–1543.
- Gräser Y, Volovsek M, Arrington J, Schönian G, Presber W, Mitchell TG & Vilgalys R (1996) Molecular markers reveal that population structure of the human pathogen *Candida albicans* exhibits both clonality and recombination. *P Natl Acad Sci USA* **93**: 12473–12477.
- Holmes AR, Tsao S, Ong S-W *et al.* (2006) Heterozygosity and functional allelic variation in the *Candida albicans* efflux pump genes *CDR1* and *CDR2*. *Mol Microbiol* **62**: 170–186.
- Hube B (2004) From commensal to pathogen: stage- and tissue-specific gene expression of *Candida albicans*. *Curr Opin Microbiol* **7**: 336–341.
- Janbon G, Sherman F & Rustchenko E (1998) Monosomy of a specific chromosome determines L-sorbose utilization: a novel regulatory mechanism in *Candida albicans*. *P Natl Acad Sci USA* **95**: 5150–5155.
- Jones T, Federspiel N, Chibana H *et al.* (2004) The diploid genome of *Candida albicans*. *P Natl Acad Sci USA* **101**: 7329–7334.
- Legrand M, Forche A, Selmecki A, Chan C, Kirkpatrick DT & Berman J (2008) Haplotype mapping of a diploid non-meiotic

- organism using existing and induced aneuploides. *PLoS Genet* **4**: e1.
- Lephart PR & Magee PT (2006) Effect of the major repeat sequence on mitotic recombination in *Candida albicans*. *Genetics* **174**: 1737–1744.
- Lephart PR, Chibana H & Magee PT (2005) Effect of the major repeat sequence on chromosome loss in *Candida albicans*. *Eukaryot Cell* **4**: 733–741.
- McEwen JG, Taylor JW, Carter D *et al.* (2000) Molecular typing of pathogenic fungi. *Med Mycol* **38**: 189–197.
- Morschhäuser J, Barker KS, Liu TT, Blass-Warmuth J, Homayouni R & Rogers PD (2007) The transcription factor Mrr1p controls expression of the *MDR1* efflux pump and mediates multidrug resistance in *Candida albicans*. *PLoS Pathog* **3**: e164.
- Perepnikhatka V, Fischer FJ, Niimi M, Baker RA, Cannon RD, Wang YK, Sherman F & Rustchenko E (1999) Specific chromosome alterations in fluconazole-resistant mutants of *Candida albicans*. *J Bacteriol* **181**: 4041–4049.
- Pujol C, Reynes J, Renaud F, Raymond M, Tibayrenc M, Ayala FJ, Janbon F, Mallie M & Bastide JM (1993) The yeast *Candida albicans* has a clonal mode of reproduction in a population of infected human immunodeficiency virus-positive patients. *P Natl Acad Sci USA* **90**: 9456–9459.
- Rustchenko E (2007) Chromosome instability in *Candida albicans*. *FEMS Yeast Res* **7**: 2–11.
- Rustchenko EP, Howard DH & Sherman F (1994) Chromosomal alterations of *Candida albicans* are associated with the gain and loss of assimilating functions. *J Bacteriol* **176**: 3231–3241.
- Selmecki A, Bergmann S & Berman J (2005) Comparative genome hybridization reveals widespread aneuploidy in *Candida albicans* laboratory strains. *Mol Microbiol* **55**: 1553–1565.
- Selmecki A, Forche A & Berman J (2006) Aneuploidy and isochromosome formation in drug-resistant *Candida albicans*. *Science* **313**: 367–370.
- Selmecki A, Gerami-Nejad M, Paulson C, Forche A & Berman J (2008) An isochromosome confers drug resistance *in vivo* by amplification of two genes, *ERG11* and *TAC1*. *Mol Microbiol* **68**: 624–641.
- Tuite A, Mullick A & Gros P (2004) Genetic analysis of innate immunity in resistance to *Candida albicans*. *Genes Immun* **5**: 576–587.
- Van het Hoog M, Rast TJ, Martchenko M *et al.* (2007) Assembly of the *Candida albicans* genome into sixteen supercontigs aligned on the eight chromosomes. *Genome Biol* **8**: R52.
- White TC (1997a) Increased mRNA levels of *ERG16*, *CDR*, and *MDR1* correlate with increases in azole resistance in *Candida albicans* isolates from a patient infected with human immunodeficiency virus. *Antimicrob Agents Ch* **41**: 1482–1487.
- White TC (1997b) The presence of an R467 K amino acid substitution and loss of allelic variation correlate with an azole-resistant lanosterol 14 α demethylase in *Candida albicans*. *Antimicrob Agents Ch* **41**: 1488–1494.
- Wu W, Pujol C, Lockhart DJ & Soll DR (2005) Chromosome loss followed by duplication is the major mechanism of spontaneous mating-type locus homozygosis in *Candida albicans*. *Genetics* **169**: 1311–1327.
- Xu J, Vilgalys R & Mitchell TG (1999) Lack of genetic differentiation between two geographically diverse samples of *Candida albicans* isolated from patients infected with human immunodeficiency virus. *J Bacteriol* **181**: 1369–1373.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Strains used in this study.

Table S2. Additional SNP-RFLP markers for Chromosome R.

Table S3. Additional SNP-RFLP markers for Chromosome 1.

Table S4. Additional SNP-RFLP markers for Chromosome 2.

Table S5. Additional SNP-RFLP markers for Chromosome 3.

Table S6. Additional SNP-RFLP markers for Chromosome 4.

Table S7. Additional SNP-RFLP markers for Chromosome 5.

Table S8. Additional SNP-RFLP markers for Chromosome 6.

Table S9. Additional SNP-RFLP markers for Chromosome 7.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.