Genomic Plasticity of the Human Fungal Pathogen *Candida albicans*\(^\text{▼}\)

Anna Selmecki,\(^1\) Anja Forche,\(^2\) and Judith Berman\(^3,4\)*

Department of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts;\(^1\) Department of Biology, Bowdoin College, Brunswick, Maine;\(^2\) Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, Minnesota;\(^3\) and Department of Microbiology, University of Minnesota, Minneapolis, Minnesota\(^4\)

The genomic plasticity of *Candida albicans*, a commensal and common opportunistic fungal pathogen, continues to reveal unexpected surprises. Once thought to be asexual, we now know that the organism can generate genetic diversity through several mechanisms, including mating between cells of the opposite or of the same mating type and by a parasexual reduction in chromosome number that can be accompanied by recombination events (2, 12, 14, 53, 77, 115). In addition, dramatic genome changes can appear quite rapidly in mitotic cells propagated in *vitro* as well as *in vivo*. The detection of aneuploidy in other fungal pathogens isolated directly from patients (145) and from environmental samples (71) suggests that variations in chromosome organization and copy number are a common mechanism used by pathogenic fungi to rapidly generate diversity in response to stressful growth conditions, including, but not limited to, antifungal drug exposure.

Since cancer cells often become polyploid and/or aneuploid, some of the lessons learned from studies of genome plasticity in *C. albicans* may provide important insights into how these processes occur in higher-eukaryotic cells exposed to stresses such as anticancer drugs.

The purpose of this review is to describe the tools used to detect genome changes, to highlight recent advances in our understanding of large-scale chromosome changes that arise in *C. albicans*, and to discuss the role of specific stresses in eliciting these genome changes. The types of genomic diversity that have been characterized suggest that *C. albicans* can undergo extreme genomic changes in order to survive stresses in the human host. We propose that *C. albicans* and other pathogens may have evolved mechanisms not only to tolerate but also to generate large-scale genetic variation as a means of adaptation.

*C. albicans* is a polymorphic yeast with a 16-Mb (haploid) genome organized in 8 diploid chromosomes (140, 154, 203). The *C. albicans* genome displays a very high degree of plasticity. This plasticity includes the types of genomic changes frequently observed with cancer cells, including gross chromosomal rearrangements, aneuploidy, and loss of heterozygosity (reviewed in references 100, 117, and 157). Similar to somatic cancer cells, *C. albicans* reproduces primarily through asexual clonal division (65, 84). Nonetheless, it has retained much of the machinery needed for mating and meiosis (189), yet meiosis has never been observed (13, 120).

*C. albicans* has two mating-type-like (*MTL*) alleles, *MTLα* and *MTLα* (76). The *MTL* locus is on the left arm of chromosome 5 (Chr5), approximately 80 kbp from the centromere. Most *C. albicans* isolates are heterozygous for the *MTL* locus, but approximately 3 to 10% of clinical isolates are naturally homozygous at *MTL* (104, 108). Mating can occur between strains carrying the opposite *MTL* locus, and most strains that were found to be naturally *MTL* homozygous are mating competent (104, 108). *MTL*-homozygous strains were also constructed from *MTL*-heterozygous strains by deletion of either the *MTLα* or *MTLα* locus (77) or by selection for Chr5 loss on sorbose (87, 115).

Mating between these diploid strains of opposite mating type can occur both *in vitro* (115) and *in vivo* (77, 97). The products are tetraploid and do not undergo a conventional meiotic reduction in ploidy (12, 120). Rather, they undergo random loss of multiple chromosomes, a process termed “concerted chromosome loss,” until they reach a near-diploid genome content (2, 12, 53, 85). A subset of these cells also undergoes multiple gene conversion events reminiscent of meiotic recombination, and most remain trisomic for one to several chromosomes (53). While mating and concerted chromosome loss have been induced in the laboratory, the role of the parasexual cycle during the host-pathogen interaction and in the response to stresses, such as exposure to antifungal drugs, remains unclear. The prevailing model is that adaptive mutations (such as those that occur with the acquisition of drug resistance) evolve through somatic events, including point mutations, recombination, gene conversion, loss of heterozygosity, and/or aneuploidy (13).

**TOOLS FOR DETECTING GENOME CHANGES**

The development and adaptation of tools that analyze genome structure have greatly enhanced our ability to detect changes in genome organization, chromosome number, allelic ratios, and genome size, to monitor the dynamics with which they arise, and to explore the molecular mechanisms that drive adaptive evolution. Here we discuss four major approaches to measure genome organization: copy number variation (CNV), loss of heterozygosity at single nucleotide polymorphisms (SNPs), whole-genome ploidy, and several other methods de-
veloped to monitor relationships between different *C. albicans* strains. While the focus here is on *C. albicans*, the approaches are generally applicable to other medically relevant eukaryotic organisms as well.

**Electrophoretic karyotyping.** The most common method used to analyze genome organization and to detect gross chromosomal rearrangements (GCRs) is to monitor chromosome size using electrophoretic karyotyping, the separation of whole chromosomes by pulsed-field gel electrophoresis (PFGE) (168). PFGE separates whole chromosomes or fragments of chromosomal DNA by changing the direction of the electrophoretic current in a regular pattern. DNA strands migrate relative to their size because they must recoil prior to aligning with the new direction of the current; shorter DNA fragments realign more rapidly than longer fragments. Chromosome-sized DNA is protected from shearing forces by extracting DNA from cells that are immobilized in agarose.

Multiple techniques are often generally termed PFGE, but they differ in their electric potential and the number of electrodes used per electric field. A PFGE apparatus has two alternating electric fields, one homogenous and one nonhomogenous, while the orthogonal field alternation gel electrophoresis (OFAGE) apparatus has two nonhomogenous fields (24, 163). Both PFGE and OFAGE were optimized for the separation of *C. albicans* chromosomes (46, 101, 113, 132, 160). A disadvantage of these techniques is that samples separate in a curved trajectory, making lane-to-lane comparisons difficult (151, 171). Later development of the contour-clamped homogenous electric field (CHEF) apparatus addressed this by using multiple electrodes around a closed contour to maintain a uniform, homogenous potential across the entire electric field (112, 191).

PFGE is used to separate all 8 *C. albicans* chromosomes and can be optimized to resolve either larger or smaller chromosomes (101, 112, 113, 160, 178, 191). While this enables the determination of a karyotype pattern (Fig. 1) and a “standard” karyotype has been codified (119), ~30 to 50% of isolates vary significantly from this standard karyotype (for examples, see Fig. 1B). The highly variable electrophoretic karyotypes observed with clinical isolates of *C. albicans* are an indication of the extensive GCRs that can occur during growth in the host (9, 47, 48, 142, 160). Even greater karyotypic variability is observed with clinical isolates of *Candida glabrata* (145) and *Candida dubliniensis* (116), suggesting that karyotype variability is common in at least several medically relevant fungal pathogens.

Karyotyping revealed that *C. albicans* chromosome size can be highly variable. Southern blot analysis of CHEF gels has further determined the identity of chromosome bands with altered mobility. Southern analyses showed that this plasticity is due largely to inter- and intrachromosomal recombination (described below). However, some changes are too small to be detected or can be masked by other chromosomes, requiring additional fine-mapping using restriction digests of chromosomal DNA within the agarose molds, followed by either CHEF or conventional gel electrophoresis. In particular, the restriction enzyme SfiI, which recognizes a rare 8-bp restriction site and is present primarily in the major repeat sequence (MRS; discussed below) (169), has been used to identify common chromosome fragments that become rearranged (35), while XhoI, which does not digest within the MRS, is used together with unique MRS-flanking probes to determine the length of the MRS on different chromosomes (105).

**Comparative genome hybridization.** The availability of the *C. albicans* genome sequence (20, 91, 190) facilitated the construction of microarrays for the analysis of gene expression at the RNA level and the gene copy number at the genomic DNA level. Initially, cDNA arrays were designed independently by consortia of laboratories in the United States, and in Europe by Eurogentec (Seraing, Belgium), based upon assembly 6 of the genome sequence. These early arrays were constructed by PCR amplification of open reading frame (ORF) probes from genomic DNA (gDNA) (from sequenced strain SC5314) that were spotted onto coated glass slides (15, 109, 155, 188). An Affymetrix array format was also constructed and used in a limited number of published experiments (98, 99).

Oligonucleotide tiling arrays, produced by Agilent or Nimblegen, in which the DNA sequence corresponding to both intra- and intergenic regions of the genome are represented in multiplex format, enabled the rapid detection of CNVs for multiple isolates on a single microscope slide. The next generation of these tiling arrays will include oligonucleotides for the detection of CNVs and single nucleotide polymorphisms in the same multiplex format (66).

Array-based comparative genome hybridization (aCGH) is the direct comparison of two genomes by competitive hybridization using a microarray platform (146). This method determines the gene copy number of essentially every ORF in the genome by comparing the DNA content of an experimental strain to the DNA content of a reference strain, each labeled with a different fluorophore. aCGH has been used to detect aneuploidy in sequenced organisms from *Escherichia coli* and *Saccharomyces cerevisiae* to mice and humans (78, 81, 153, 200). Whole-chromosome aneuploidies are visualized when array data are plotted as a function of chromosome location using programs such as Matlab (172), Treeview (161), or Excel (7). Because gene copy number often correlates with gene
and 87,000 and 89,000 SNPs were identified, suggesting that only a lower rate of variation is seen when SNPs are nearly homozygous. In contrast, in WO-1 both arms of Chr5 are present, on average, once every 390 bp (22, 91); in this strain SC5314 suggested that single nucleotide polymorphisms (SNPs) are by segmental deletions (Fig. 2). The genome sequence of DNA regions extending to the telomere), by chromosome loss, short DNA tracts), by break-induced replication (usually long-range LOH events can arise from single crossovers followed by mitotic segregation or from break-induced replication (BIR). (B) Events due to chromosome missegregation: chromosome nondisjunction yields aneuploid cells monosomic for that chromosome (b) as well as cells carrying an extra copy of the homolog (aab). Events due to reduplication of the monosomic chromosome via a subsequent nondisjunction event or by unlicensed recombination can generate a strain that is homozygous but disomic for that chromosome (bb).

expression levels (67, 75, 94, 173), expression array data profiles are a good proxy for gene copy number. Recently, two groups have retrospectively analyzed hundreds of C. albicans gene expression profiles for expression biases that indicate gene copy number changes (7, 17). Similarly, segmental aneuploidies are readily detected from aCGH data; the breakpoints, where euploid-to-aneuploid transitions occur, are often within moderately repetitive DNA (49, 75, 81, 172). aCGH also has been used to compare the genomes of related organisms, often before whole-genome sequences were available (135, 195). This approach has the potential to advance our understanding of genomic differences between a pathogenic species and a related, less pathogenic species, as was done with C. albicans and C. dubliniensis (134).

SNP. Single nucleotide polymorphism (SNP) analysis reveals allelic differences between strains due to maintenance or loss of heterozygosity (LOH). LOH can arise by recombination and subsequent allele segregation, by gene conversion (usually short DNA tracts), by break-induced replication (usually long DNA regions extending to the telomere), by chromosome loss, or by segmental deletions (Fig. 2). The genome sequence of strain SC5314 suggested that single nucleotide polymorphisms are present, on average, once every 390 bp (22, 91); in this strain both ends of ChrR, most of Chr3R, and Chr7L are nearly homozygous. In contrast, in WO-1 both arms of Chr5 and most of Chr2L, -6L, -3R, and -1R are nearly homozygous (22). Interestingly, a higher rate of variation is seen when SNPs of the two genomes are compared to each other; between 87,000 and 89,000 SNPs were identified, suggesting that only a very small number of SNPs are actually shared between the two genomes (22).

Based on the genome sequence of strain SC5314, SNP arrays were developed for high-throughput genotype analysis (55, 56). Short allele-specific oligonucleotides representing both alleles are spotted onto the array. When labeled probes are hybridized to the array, probes from a heterozygous strain will hybridize to both oligonucleotides, while probes from a homozygous strain will hybridize only to one of the oligonucleotides (54, 56). SNP genotypes are determined by calculating allelic fractions that are defined for each of the three possible genotypes (55).

One outcome of aneuploidy in C. albicans is allelic imbalance. In SNP array analysis, a heterozygous diploid should exhibit a 1:1 ratio of the two alleles, while a heterozygous trisomic strain should exhibit an allelic ratio of 2:1. When multiple SNP markers on a chromosome arm all exhibit ratios of ~2:1 (or 1:2), the region is very likely trisomic. In this way, quantification of allelic ratios can predict whole-chromosome or segmental trisomies. It is important to note, however, that tetraploid strains (e.g., those that arise following mating between diploids) will exhibit 1:1 ratios for all markers that were heterozygous in the parents; thus, SNP arrays are not reliable for detecting whole-genome polyploidy.

Haplotype mapping is used in sexual organisms to follow recombination events and to study genetic diversity by analyzing the segregation of SNPs in large numbers of meiotic progeny. Since C. albicans reproduces predominantly by clonal division, standard haplotype mapping approaches cannot be used. Legrand et al. (103) exploited naturally occurring aneuploidies and homozygositites in strain SC5314 and the clinical isolate T118 (41) to generate haplotype maps (HapMaps) for each strain (103). They compared HapMaps and discovered multiple recombination events that distinguished the two strains; most appeared to be the result of either gene conversion or single crossover and chromatid segregation events. These HapMaps were the first of their kind for a diploid, nonmeiotic organism. In addition, they showed that whole-chromosome trisomies arise upon use of the UAU1 method (51) for the disruption of essential genes. Thus, the UAU1 method can be exploited to generate heterozygous trisomic strains, which facilitates the determination of haplotypes for any clinical isolate (103).

The classical DNA-fingerprinting method (digest of genomic DNA followed by Southern hybridization) detects restriction fragment length polymorphisms (RFLP) among strains and has been used widely for strain typing in C. albicans (90, 122). Multilocus sequence typing (MLST) determines strain relatedness based on sequence variation and has been used successfully for strain typing and epidemiological studies (18, 19). However, MLST analyzes only sequence variation for a limited number of housekeeping genes (7 genes) that are distributed on only 5 of the 8 C. albicans chromosomes (ChrR, -1, -2, -4, and -6) (139). Recently, amplification of short genomic regions (SNP-RFLP markers) followed by an allele-specific restriction digest of PCR products (PCR-RFLP) has been used for genotype determination in population genetic studies (57, 208). Forche and coworkers have developed a set of 32 SNP-RFLP markers, two per C. albicans chromosome arm, which provides a simple and rapid method to not only analyze heterozygous
loci before and after strain transformation but also identify LOH events that may have occurred at unbiased loci across the genome (58).

**Flow cytometry.** Flow cytometry provides a rapid and accurate method for analyzing genome size changes (ploidy changes). Genomic DNA is stained with an intercalating fluorescent dye, and the amount of DNA per cell is quantified simultaneously for thousands of cells per second (68). Flow cytometry of DNA content was used to identify tetraploid mating progeny and to select for near-diploid progeny following concerted chromosome loss (CChrL) (12, 53). It generally does not have the sensitivity to detect the addition or loss of a single chromosome, but with *C. albicans*, flow cytometry can detect shifts in ploidy due to multiple aneuploidies in a strain (D. Abbey and J. Berman, unpublished data).

In summary, all four of these genomic tools (CHEF, aCGH, SNPs, and flow cytometry) provide important information about genome structure. CHEF reveals chromosome length changes and reciprocal translocations but is limited to detecting larger changes in chromosome size and structure. Southern hybridization to DNA separated by CHEF overcomes problems with overlapping chromosome bands and is indispensable for detecting translocations, yet it is difficult to detect whole-chromosome trisomies with this method alone. Furthermore, Southern hybridization is limited by the number of probes used. aCGH readily detects copy number changes for essentially every gene in the genome, but it cannot detect whole-genome ploidy shifts, reciprocal translocations, or LOH. Similarly, SNP arrays determine the ratio of different alleles and detect LOH and major recombination events such as gene conversion, but they cannot detect whole-genome ploidy changes or translocations. Furthermore, SNP array analysis is limited by the number of SNPs included on the array; this limitation will be much less of a concern when tiling arrays with limited by the number of SNPs included on the array; this limitation will be much less of a concern when tiling arrays with limited by the number of SNPs included on the array; this limitation will be much less of a concern when tiling arrays with limited by the number of SNPs included on the array; this limitation will be much less of a concern when tiling arrays with

In conclusion, all four of these genomic tools (CHEF, aCGH, SNPs, and flow cytometry) provide important information about genome structure. CHEF reveals chromosome length changes and reciprocal translocations but is limited to detecting larger changes in chromosome size and structure. Southern hybridization to DNA separated by CHEF overcomes problems with overlapping chromosome bands and is indispensable for detecting translocations, yet it is difficult to detect whole-chromosome trisomies with this method alone. Furthermore, Southern hybridization is limited by the number of probes used. aCGH readily detects copy number changes for essentially every gene in the genome, but it cannot detect whole-genome ploidy shifts, reciprocal translocations, or LOH. Similarly, SNP arrays determine the ratio of different alleles and detect LOH and major recombination events such as gene conversion, but they cannot detect whole-genome ploidy changes or translocations. Furthermore, SNP array analysis is limited by the number of SNPs included on the array; this limitation will be much less of a concern when tiling arrays with large numbers of probes are used. Finally, flow cytometry readily detects whole-genome ploidy shifts that are not evident by CHEF, aCGH, or SNP analyses; however, it cannot determine which specific chromosome(s) may be aneuploid, and it is not sufficiently sensitive to detect the presence of lower levels of aneuploidy (e.g., <20% of the normal DNA content). Thus, a combination of all four methods provides the most comprehensive view of genome changes to date. New array formats that permit SNP and copy number variation to be detected within the same custom designed array (e.g., see reference 66) will speed the analysis of genome changes.

**TYPES OF GENETIC DIVERSITY IN C. ALBICANS**

A striking feature of the *C. albicans* genome is the high degree of genetic diversity that exists among clinical isolates. This genetic diversity has been characterized from the whole-chromosome level down to the single-nucleotide level, using subsets of the methods described above.

**Translocations and gross chromosomal rearrangements.** Large changes in chromosome size, due to translocations, repeat length changes, chromosome truncations, telomere recombination events, and supernumerary chromosome formation, all have been detected in *C. albicans*. A number of these involved the major repeat sequence (MRS), a tract of repeats that varies in size from tens of kilobase pairs to approximately one hundred kilobase pairs (117, 118). There are 9 MRSs in the *C. albicans* genome, one on Chr R, -1, -2, -5, and -6 and two on Chr 4 and -7. Chr 3 contains only one portion (RB2) of the repeat (31, 33, 34, 82). The MRS is composed of a set of nested repeats contained in the RPS region (~2 kbp) flanked by longer regions termed HOK (~8 kbp) and RB2 (~6 kbp) (Fig. 1C) (33, 34). The RPS region is composed of tandem repeats of ~172 bp each, termed “alts,” that vary only slightly in sequence within and between chromosomes; this results in RPS size variability (31, 82). The alts contain ~6 to 8 copies of a 29-bp conserved sequence (termed COM29) that includes the 8-bp recognition site for the macrorestriction endonuclease Sfi I. Outside of the 9 RPS sequences, Sfi I cuts infrequently, such that the resulting large chromosome fragments can be analyzed on CHEF gels (35).

Because MRSs are the largest nontelomeric, homologous sequences present on each chromosome, they provide a potential hot spot for reciprocal recombination events between non-homologous chromosomes (30, 35, 86, 86, 105, 107). The classic example of chromosome translocation in *C. albicans* was described for strain WO-1, an *MTLa* homozygous strain for which white-opaque switching was first described (177). Chu and coworkers found that WO-1 had six new bands on whole-chromosome CHEF gels yet had very similar Sfi I digestion patterns (35), indicating that the strain had undergone 3 translocations, all of them at or near an MRS. Similar MRS translocations resulted in the formation of two new chromosomal bands in the clinical isolate 1001 (136). Thus, some strains appear to have undergone multiple reciprocal translocation events that involve the MRS.

Expansion and contraction of MRS repeats that are caused by unequal recombination between sister chromatid repeats also contribute to chromosome length polymorphisms. MRS length changes are detected by digestion with XhoI, which cuts outside the MRS, followed by Southern hybridization with an MRS-specific probe (32, 106). This intrachromosomal recombination is frequently observed as a separation of homologous chromosomes on CHEF gels, due to the MRS length change on only one homolog. MRS length changes are detected frequently for the smaller chromosomes (Chr 5 to -7), although this is biased by the increased resolution of the smaller chromosomes by CHEF (30). Deletion of an MRS has no phenotypic effect on the cell, but MRS length apparently affects chromosome segregation, such that homologs with a larger MRS have a higher rate of chromosome nondisjunction than the same chromosome homolog with a shorter MRS (106).

Length polymorphisms on Chr R also arise by changes in the number of ribosomal DNA (rDNA) repeats, all of which are present at a single ~660-kbp locus containing ~55 copies of the 12-kbp rDNA repeat that contains genes for 18s, 5s, 25s, and 5s rRNA (91). As in *S. cerevisiae* (186), expansions and contractions due to unequal recombinations of rDNA repeats occur frequently and are thought to be the basis for large variations in the size of Chr R in different *C. albicans* strains and upon serial passaging of the same strain (1, 84, 158). XhoI does not cut within the rDNA and is used in conjunction with CHEF to analyze rDNA repeat length changes as described above for the MRS (84). Recombination between rDNA repeats can also yield extrachromosomal rDNA circles and linear rDNA extrachromosomal plasmids (1, 74), but a correlation
between rDNA repeats and aging, as described for some \textit{S. cerevisiae} strains (62, 176), or genome integrity (80) has not been detected (61a). Similar changes for ChrR have been described for related \textit{Candida} species, including \textit{C. parapsilosis} (138).

Reciprocal translocation events, as well as nonreciprocal recombination events that alter MRS and rDNA copy length, are often detected on CHEF gels, yet they do not affect aCGH profiles, since the number of unique ORFs does not change. These are best distinguished using CHEF of XhoI digests to detect changes in rDNA length (84) and MRS length (106) and SfiI digestion to detect new chromosome fragments. Intriguingly, clinical isolates of \textit{C. glabrata} appear to undergo many translocation events (145), although the degree to which these translocations affect copy number remains unclear. In most cases Southern hybridization with chromosome-specific probes will assist in distinguishing these two types of changes in chromosome organization. Since \textit{C. glabrata} is a haploid, perhaps it will behave more like haploid \textit{S. cerevisiae}, by acquiring mostly segmental, rather than whole-chromosome, aneuploidies.

**Chromosome truncations.** Chromosome truncations and fragmentations have been reported for laboratory and clinical isolates. For example, laboratory strain BWP17 (206) and several of its parental progenitors (4) incurred a chromosome break during disruption of \textit{HIS1}, resulting in two separable homologs of Chr5 (54, 137, 172). \textit{HIS1} is 35 kbp from the right telomere of Chr5. A 9-bp sequence with similarity to the 23-bp \textit{C. albicans} telomere repeat is present nearby and apparently seeded the telomere addition following a chromosome break that accompanied \textit{HIS1} deletion. All 17 genes distal to \textit{HIS1} were lost, and thus, this strain is monosomic for the terminal portion of Chr5R (172). Importantly, there is at least one essential gene among them, such that loss of the full-length homolog results in cell death (P. T. Magee and J. Berman, personal communication).

Truncations also arise in clinical isolates. In isolate FH8 (125), from a bone marrow transplant patient treated with fluconazole, Chr5 was truncated ~200 kbp from the left telomere, resulting in a Chr5 homolog indistinguishable in size from Chr6 by CHEF analysis (174). The Chr5 breakpoint in this strain is near 3 genes with high sequence identity, which may provide a recombination hot spot or fragile site (8). Thus, highly similar gene sequences as well as sequences resembling the natural telomere repeat appear to facilitate unintended chromosome truncations, perhaps by a mechanism similar to that used to construct strains with specific telomere-mediated chromosome truncations in \textit{S. cerevisiae} (182, 192) and more recently in \textit{C. albicans} (6, 93, 174).

**SNCs.** Karyotype analyses have identified supernumerary chromosomes (SNCs) (74, 104, 136). A SNC is an additional chromosome, usually generated from extra copies of other chromosomes or a chromosome fragment(s). It contains extra, usually dispensable, information and thus exhibits a size (and thus CHEF gel mobility) different from that of the original chromosomes from which it was derived. Some SNCs include multiple copies of a single chromosome region, e.g., isochromosomes i(5R), i(5L), att-i(5L), and i(5L)+3R (173, 175). In some cases SNCs can be identified only by a combination of aCGH and Southern blot analysis, because they comigrate with a different chromosome (Fig. 3A to C). The most efficient way to identify the chromosomal constituents of a SNC is by aCGH of SNC DNA excised from a CHEF gel (67, 175). This should then be confirmed by Southern analysis of a CHEF blot using probes for the relevant chromosome fragments to determine if the SNC size is consistent with the presence of one or more copies of the fragments detected.

**Whole-chromosome aneuploidy.** Aneuploidy, an abnormality of chromosome number (presence of extra or the lack of chromosomes), is the result of errors in DNA replication or chromosome segregation and can involve whole chromosomes or chromosome segments. Whole-chromosome aneuploidies, including monosomy (2n − 1) and trisomy (2n + 1), were first detected in \textit{C. albicans} by studying the loss of nutritional markers and by densitometric analysis of ethidium bromide (EtBr)-stained CHEF gels and quantitative Southern hybridization with different chromosome probes (10, 28, 30, 70, 83, 114, 130, 142, 160, 179). Recently, RFLP analysis together with quantitative Southern hybridization was used to confirm a suspected whole-chromosome aneuploidy in several isolates of strain CAI4 (28), using probes to several loci on a given chromosome and quantitative comparison of allele frequencies. While Chr1 trisomy was detected in some isolates of laboratory strain CAI4, trisomy of Chr2, which is present in all of the CAI4 isolates (172), was not detected because no Chr2 probes were used. A more comprehensive analysis by Southern hybridization would require the use of multiple probes for each chromosome arm, a process that is quite labor intensive.

Genome changes in \textit{C. albicans} can be more complex. Translocated chromosomes can be lost or amplified (85, 136); for example, WO-2, a derivative of strain WO-1, lost one of the reciprocal recombination products observed with WO-1 (a Chr5,6 fusion) and thus became monosomic for each part of the original donor chromosome (Fig. 4A to C) (114). WO-2 also lost the only full-length Chr7 homolog of WO-1; the other copies of Chr7 are found on two translocated chromosomes, Chr7,4 and Chr4,7. Interestingly, aCGH analysis of WO-2 identified the two previously reported loss events and also revealed that a different translocated chromosome (the fusion of Chr7 and Chr4) was amplified, resulting in a segmental trisomy of Chr4 and a segmental disomy of Chr7 (Fig. 4D and E), and further highlights the comprehensive nature of aCGH. (Some of the aCGH and flow cytometry data in Fig. 3, 4, and 5 are primary data that have not been published previously. They are now available at the Genome Expression Omnibus, under accession number GSE21643.)

Several studies have identified whole-chromosome aneuploidies that correlate with phenotypic variation in \textit{C. albicans}. For example, some aneuploidies provide a selective advantage for growth on specific carbon sources (86, 87, 160). In response to the presence of l-sorbose or D-arabinose as the sole carbon source, colonies of strain 3153A underwent chromosome copy number alterations detectable by PFGE. aCGH analysis of strain 3153A did not detect any alterations in copy number (Fig. 5C) relative to the standard karyotype, suggesting that a translocation event involving Chr4R and Chr7 in this strain (87) was reciprocal. Densitometric comparison of EtBr-stained chromosome bands followed by quantitative Southern hybridization with different chromosome probes revealed that either loss of Chr6 or gain of Chr2 correlated with the assimilation of D-arabinose (159), while monosomy of Chr5 enabled strains to
grow better on \( L \)-sorbose (87). Importantly, strains monosomic for Chr5 could reduplicate Chr5, and in doing so, they lost the ability to grow rapidly on \( L \)-sorbose (87). Although the specific genes inhibiting sorbose utilization have not been identified, they have been mapped to a 209-kbp region on Chr5R (92), and the assumption is that this region must contain a negative regulator of sorbose utilization (87). We analyzed several of these strains using aCGH. Interestingly, one of these spontaneous \( Sou^+ \)/H11001 strains, Sor5 (92), includes a segmental trisomy of Chr4R in addition to a segmental monosomy of Chr5R (Fig. 5E). This is interesting because the trisomic region of Chr4R includes genes encoding Sou1 and Sou2, which facilitate sorbose utilization. \( L \)-Sorbose induction of Chr5 loss is now frequently used to generate strains homozygous for the mating type-like locus (\( MTL \)), which is located on Chr5L. Furthermore, strain SGY243 carries a Chr4R amplification similar to that seen with Sor5 (Fig. 5D), and both breakpoints occur at the MRS. It is important to note that while Chr5 loss certainly can occur following sorbose treatment, other chromosomes are often affected as well, not all \( Sou^+ \) strains are monosomic for Chr5 (6, 17, 53), and segmental trisomy of Chr4R has been observed with several \( Sou^+ \) derivatives (E. Rustchenko, personal communication).

In addition to carbon source utilization, a correlation between chromosome changes and resistance to fluconazole (142), the most commonly used antifungal drug, was reported. Fluconazole-resistant (FluR) isolates of a multiply aneuploid strain SGY-243 were generated \( in vitro \) using very high (1.5-mg/ml) concentrations of drug (142). After 7 days of exposure to 1.5 mg/ml fluconazole, resistant clones carried only two of the four parental copies of Chr4, while after 35 days of exposure resistant clones also gained an extra copy of Chr3. The chromosome changes were identified by CHEF and Southern blotting with 7 different probes. These studies suggested a role for Chr3 and Chr4 in fluconazole resistance. However, it remained unclear whether resistance was acquired because of the aneuploid nature of the parent and/or the very high levels of drug used. More recent work has found aneuploidy in \( ~50\% \) of cells that have become resistant to fluconazole, as described in more detail below (173).
Clearly, \textit{C. albicans} isolates have a remarkable tolerance for gross chromosomal rearrangements, including multiple aneuploidies. Furthermore, at least some of these chromosome changes can confer a selective advantage under different stresses, such as poor carbon source availability or antifungal drug exposure (9, 10, 87, 173). Others are correlated with phenotypes such as changes in morphogenesis (55, 56); for example, strains that had lost either homolog of ChrR exhibited colony morphologies different from those of the smooth parental phenotype. Indeed, given the ability to tolerate aneuploidies and major chromosomal rearrangements, it seems amazing that a “standard” karyotype is found in even 30 to 50% of clinical isolates (116).

Genetic diversity due to heterozygosity and LOH. Because it is a diploid, a major source of genetic variability in \textit{C. albicans} is provided by the maintenance of heterozygous alleles. Early investigations identified many different genes encoding nutritional markers, such as genes required for the biosynthesis of methionine and cysteine, that are heterozygous (147, 201, 203). Recent sequence analysis also indicates that the positions of SNP loci vary significantly between strains (22).

Despite the maintenance of heterozygosity, LOH is common in \textit{C. albicans}. The phenotypic advantages of certain alleles and the presence of specific haplotypes in drug-resistant isolates have been analyzed (39, 50, 141, 165). For example, LOH appears to play an important role in acquired azole resistance in \textit{C. albicans}: homozygosis of the hyperactive alleles of \textit{ERG11}, \textit{TAC1}, or \textit{MRR1} (which encode the target of azole drugs and transcription factors that upregulate drug efflux pumps) is associated with increased azole resistance (39, 50, 204). Homozygosis of the \textit{PAPI} allele within the \textit{MTL} locus also confers different levels of resistance to drugs (88, 121). Similarly, resistance to the antifungal 5-flucytosine (5-FC) was associated with homozygosis of a hyperactive \textit{FUR1} allele (45). Furthermore, different alleles have been found to have different levels of expression. Strains homozygous for a novel allele of \textit{HWP1}, which encodes a major \textit{C. albicans} hyphal cell wall protein, showed significantly lower levels of \textit{HWP1} expression during hyphal growth and biofilm formation (141). This suggests that the \textit{HWP1} locus has biofilm-specific differential expression of different \textit{HWP1} alleles. A natural heterozygosity in the promoter region of the gene \textit{CHS7} (a GC-rich sequence of 28 nucleotides) led to reduced levels of chitin, partial resistance to calcofluor, and a moderate degree of altered-morphology in hypha-inducing media; the allele lacking the GC-rich sequence did not appear to confer distinct phenotypic.
characteristics (165). Thus, in some cases homozygosis causes obvious phenotypes that can be attributed to allele-specific differences in genes of known function. In other cases, homozygosis confers no obvious phenotypic change.

The MTL locus is naturally heterozygous, yet ~3 to 7% (104, 108) of clinical isolates are homozygous for MTL (MTLhom). Interestingly, MTLhom was correlated withazole resistance in two of three different studies. In 96 clinical isolates, 11 of 12 MTLhom strains (92%) were FluR (156), with homozygosis of either MTL allele in the FluR strains detected at approximately equal frequencies; despite similar correlations, an independent study found that MTL homozygosity alone does not confer azole resistance (39). In a third study this correlation was not evident (104). Importantly, TAC1 is located ~14 kbp to the right of the MTL locus on Chr5L, and ERG11 is also located on Chr5L, ~150 kbp from the telomere. Therefore, the correlation of FluR to MTL LOH likely occurs only when in combination with LOH of the hyperactive TAC1 and/or ERG11 locus during Chr5L loss or i(5L) formation (37, 174).

**Changes in whole-genome ploidy.** An obvious mechanism for generation of ploidy changes (from diploid to tetraploid) is mating (described above) between cells of the opposite (76, 77, 115) or the same (2) mating type. Both types of mating require conversion to the opaque state, a process involving a positive
feedback loop that mediates a transcriptional switch (133, 210, 211), and same-sex mating was detected only in bar1ΔΔ cells (2). Nonetheless, both types of mating may occur outside the laboratory setting. Prior to the discovery of mating in C. albicans, tetraploids were generated through protoplast fusion between cells whose cell walls had been digested enzymatically (102, 166, 202).

Increased ploidy can also arise via mitotic defects, in which cells complete replication but nuclei fail to be properly separated and/or segregated to the two daughters. Such failures arise in a number of S. cerevisiae mutants (e.g., strains with defects in IPL1 and BEM2 [26] as well as in strains with defects in the RSC complex [23] and strains with defects in spindle pole body duplication [29]). Early work on C. albicans detected strains with mixed ploidy states using fluorescent microscopy (183–185). Additional work on these Sps strains, named for the inability to suppress ploidy shift changes, found that the polyploid yeast also generated increased chromosomal translocations relative to stable diploid controls (85). It is unclear whether the polyploid strains in this study represent cells that incurred defects in mitosis or replication, and the extent to which such mechanisms operate in C. albicans is not clear. Nonetheless, evidence for shifts in ploidy suggests yet another mechanism for generating diversity in this organism. For example, tetraploid S. cerevisiae cells display a 200- to 1,000-fold increase in chromosome loss rates, increased recombination between homologs, and gross chromosomal rearrangements relative to diploid cells (5, 129, 180).

Reduction in the ploidy of tetraploid strains generated by mating or protoplast fusion occurs by concerted chromosome loss (CChrL) in C. albicans. CChrL following mating is induced by stresses such as limited nitrogen availability or growth on sorbose (12, 53). Importantly, ~20% of cells undergoing CChrL also exhibit multiple short-range LOH events, and these recombination events require Spo11, a meiosis-specific protein that initiates double-strand breaks during meiosis in other organisms (53). However, 80% of the cells that underwent CChrL, including all of those that included multiple recombination events, are trisomic. Thus, mating progeny in C. albicans include cells that have become aneuploid (trisomic for one or more chromosomes) and have also undergone multiple LOH events. While CChrL is not equivalent to classic meiosis, it is reminiscent of the meiosis described for Candida lusitaniae (152), a closely related yeast with fewer conserved meiotic genes than C. albicans, in which high proportions of haploid mating progeny are aneuploid or remain diploid (152). This suggests that these “meioses” are highly aberrant and generate much genotypic diversity. It is not yet clear how much of this parasexual cycle is similar to, and different from, conventional meiosis and to what degree this benefits the organism. What is clear is that parasexual mating in C. albicans generates diversity by random segregation of different chromosomal homologs and generation of aneuploidy and can be accompanied by high levels of recombination events in a subset of the progeny.

The effects of ploidy and mating on virulence were tested using a murine model of disseminated candidiasis (79). In this study, tetraploid strains were less virulent than their diploid progenitors and during the infection they underwent ploidy shifts, resulting in <4N aneuploid cells. Thus, reduction in ploidy, presumably by CChrL, clearly can occur in vivo. In a different study (207), strains heterozygous for Chr5 and MTL exhibited increased virulence and a competitive advantage over MTL homozygous strains. An important question remaining is whether changes in ploidy per se or only specific combinations of aneuploid chromosomes provide a selective advantage during growth in the host.

**TYPES OF STRESSES THAT GENERATE GENOME CHANGES**

**Antifungal drugs.** Azole resistance is a serious problem in immunocompromised patients required to undergo prolonged courses of antifungal therapy. Numerous studies sought to identify genetic features that are common among drug-resistant isolates in hopes of elucidating the mechanisms of resistance. In many cases, changes were observed, but they could not be correlated conclusively with drug resistance (reviewed in references 40 and 162). We suggest that three phenomena contributed to this difficulty: first, resistance can arise through several mechanisms, only one of which is aneuploidy; second, multiple aneuploidies can be present, and only some of them may confer a selective advantage; and third, multiple yeast genotypes can be present in an individual patient. All three of these mechanisms confound the determination of which chromosomes are responsible for increased drug resistance.

**CHEF analysis of isolates that acquired fluconazole resistance in vitro (41) or in vivo (143, 151) frequently identified GCRs, but no specific karyotype change was correlated with increased resistance. Similarly, molecular typing methods, including DNA fingerprinting (209), RFLP analysis, multilocus genotyping (42), and linkage analysis, were not predictive of FluR, possibly because the markers or the primers used were not sufficient to identify a recurring FluR genotype. Finally, because there are multiple mechanisms of acquired FluR, it is not expected that all FluR strains will have the same electrophoretic karyotype or DNA fingerprint or that genetic linkage will identify all FluR strains.

Using aCGH to analyze a large number of drug-resistant C. albicans isolates revealed that changes in chromosome copy number arise with high frequency upon exposure to azole antifungals and other stresses. Analysis of the transcription profiles of drug-resistant isolates found a similar proportion of aneuploid strains (17). In a survey of over 90 clinical isolates, more than 50% of FluR isolates carried at least one aneuploidy (173; 175; unpublished data). Of these FluR aneuploid strains, 76% included at least one whole-chromosome aneuploidy, 76% included at least one segmental aneuploidy, and 59% included two or more of these aneuploidies. Trisomy was much more frequently observed than monosomy, and trisomy of the smaller chromosomes, Chr4 to -7, was detected most frequently. Importantly, the most frequent aneuploidy in all FluR strains was increased levels of DNA from Chr5, and about half of the aneuploid FluR strains carried a specific aneuploidy, isochromosome 5L (i(5L)), in which two copies of Chr5L are present in an inverted orientation together with a copy of the Chr5 centromere (Fig. 3F) (173).

Isochromosome 5L formation and fluconazole resistance also appeared twice independently, within 2 weeks of fluconazole exposure (174), in a series of 9 isolates from the same bone marrow transplant (BMT) patient (126). The resistance
in these strains was unstable or transient: it could be lost when cells were grown in the absence of fluconazole (125). Telomere truncation experiments and subsequent disruption experiments revealed that the fluconazole MIC is directly proportional to the number of copies of Chr5L and, specifically, to the total number of copies of ERG11 and TAC1 present in the strain (174). Deletion of an entire Chr5L arm had the same effect on the MIC as deletion of only the ERG11 and TAC1 loci, indicating that ERG11 and TAC1 are the major contributors to FluR on Chr5L. Consistent with this, i(5L) provides a fitness advantage in the presence ofazole drugs (175). A situation reminiscent of this was seen with C. glabrata, a pathogen more closely related to S. cerevisiae than to C. albicans: increased gene copies of CYPS1, the azole drug target, were correlated with azole resistance (123).

The appearance of aneuploidy can occur rapidly in vivo (124, 174) as well as in vitro (41, 175) following exposure to fluconazole. In vivo, two independent strains carrying i(5L) [one a diploid plus i(5L)] that is homozygous for Chr5 SNP markers and one that is almost tetraploid plus i(5L) and that is heterozygous for Chr5 SNPs arose within the bloodstream of a single patient (174) (Fig. 3D, green). In vitro experimental evolution detected i(5L) in three independent cultures following only 24 h of exposure to fluconazole and subsequent nonselective propagation (175). Strikingly, strains with ploidy intermediate between diploid and tetraploid were also found (Fig. 3D, red). Such a rapid appearance of multiple aneuploidies likely arises through an intermediate step, such as mating, to form tetraploids followed by partial C ChrL. Alternatively, cell cycle defects, such as a failure of mitosis or cytokinesis, can yield binucleate and/or mononucleate tetraploid cells (150, 181) that undergo rounds of chromosome nondisjunction due to a higher incidence of syntelic/monopolar kinetochore attachments to the spindle pole (180). Alternatively, diploid parental cells could undergo nondisjunction of multiple individual chromosomes, due to defects in chromosome segregation (Fig. 2B). Global defects of this type would likely arise through defects in centromere, cohesion, spindle, or mitotic checkpoint functions (reviewed in references 16, 36, 44, and 144).

It is evident that some aneuploidies confer a strong selective advantage under specific stress conditions. For example, S. cerevisiae cells lacking MYO1 die due to failure to form and contract the cytokinetic ring at the bud neck (150). However, some myo1Δ cells survive due to aneuploidies that increased the levels of RLM1 (a transcription factor that regulates cell wall remodeling genes), which induces the accumulation of aberrant cell wall structures near the bud necks and the bypass of myo1Δ lethality. Strong selection for specific segmental aneuploidies is also observed during in vitro evolution experiments with S. cerevisiae in media with limiting glucose (48) or limiting sulfur (64). In C. albicans, i(5L) clearly provides an advantage to cells growing in fluconazole via additional copies of both TAC1 and ERG11 (175). In some cases, this advantage in the presence of drug is accompanied by a high fitness cost in the absence of drug (e.g., in the case of a SNC composed of i(5L) plus Chr3R, such that the SNC is rapidly lost in cells grown in the absence of the drug). Similarly, the acquisition and loss of i(5L) can readily explain the “transient” FluR seen with sequential isolates from a bone marrow transplant patient (125, 174). Thus, the formation of aneuploids is a mechanism that is rapid and flexible: within one or two divisions it can confer a selective advantage under conditions of severe stress; once the stress condition has subsided, the extra chromosome is readily lost in a single step, often returning the cell to its original state. Less clear is how frequently aneuploidy is utilized as a stress response mechanism in C. albicans as well as in higher eukaryotes.

**Growth in the host.** An understudied area of Candida research is the role that genetic variation and its evolution play in host/pathogen interactions. While impaired host immune function clearly contributes to the severity of Candida infections (25), the fungus must possess characteristics that facilitate the transition from a harmless commensal to an aggressive pathogen (72). During the course of infection, C. albicans encounters many different host environments to which it must adapt so that it can grow and survive. Recent work has revealed the surprising result that a large amount of genetic and phenotypic variation appears over very short periods of time during passage in a mammalian host (55, 56; unpublished data). Isolates from an in vivo passaging experiment using a bloodstream infection (BSI) model compared to an in vitro control population found that rates of LOH at the GAL1 locus were increased ~28-fold compared to those of the in vitro environment, despite the slower in vivo growth rate. Furthermore, segmental and whole-chromosome events, including whole-chromosome LOH and whole-chromosome trisomy as well as several distinct colony morphologies, were evident only in the in vivo isolates. The association of genome rearrangements and colony phenotypes was statistically significant, implying that homozgyosity or copy number variations lead to phenotypic alterations (55).

Several groups have advanced our understanding of genome changes in vivo and ex vivo at the transcriptional level (3, 21, 59, 60, 73, 205). However, a recent study examined and compared transcription profiles of two C. albicans strains during in vivo infection and found only minimal overlap between their transcriptomes. This suggests that much of the gene regulation observed during infection may not be essential for virulence (193). This emphasizes the need to explore other mechanisms that play a role in virulence. Since exposure to the host results in a very high rate of genome-level changes, it is critical that we understand the different mechanisms by which C. albicans evolves genetic and phenotypic diversity while it is growing within the host. In addition, we need to determine which mechanisms of genome change (e.g., whole-chromosome aneuploidy versus recombination events) are specific to certain host environments, so that we can reach a deeper understanding of host-pathogen interactions from the pathogen’s perspective.

**Heat shock.** It is well documented that heat shock affects genome stability: chromosome loss was detectable following 1 or 2 min of exposure to 50°C (70). A more recent study detected both gain and loss of whole chromosomes using CHEF and aCGH analysis of genome alterations following 90 to 120 s at 50°C (17). This heat shock also reduced cell viability and gave rise to colony size changes; small colonies had twice as many karyotype changes as large colonies, and all small colonies analyzed by aCGH exhibited aneuploidies, such as whole-chromosome monosomies and/or trisomies. Therefore, limited
exposure to high temperatures most likely causes an increased number of chromosome nondisjunction events.

**Transformation with DNA.** Recently two groups have retrospectively analyzed transcription profiles for expression biases that predict gene copy number changes (7, 17). Arbour et al. (7) found that 35% of over 100 expression arrays analyzed predicted aneuploid chromosomes. In addition to urging researchers to plot expression array data and detect chromosomal or regional expression biases, this group also provided an inexpensive alternative to aCGH for detecting whole-chromosome aneuploidy by qPCR (7). Bouchonville et al. analyzed *C. albicans* expression array data from 411 experiments (7). This analysis identified entire chromosomes and chromosome regions with altered gene expression, predicting aneuploidy of those chromosomes. In several cases, aCGH confirmed several of these predicted aneuploidies. A major class of aneuploid strains were those that underwent DNA transformation, often to delete or modify specific genes. While this could be due to a low level of mutagenesis caused by 5-fluoroorotic acid (5-FOA) (1, 199), which was used for selection of most transformants, 5-FOA does not appear to cause major aneuploidies or LOH (A. Forche and J. Berman, unpublished data). Strains that were aneuploid prior to transformation were much more likely to undergo alterations in chromosome copy number as a consequence of undergoing a subsequent transformation. Transformation of strains carrying one or two aneuploid chromosomes confirmed this observation.

An important take-home message from these studies is that DNA transformation can cause changes in chromosome copy number and that transformants should be tested for possible aneuploidies. In addition, when expression profiles are generated, the data should always be plotted as a function of chromosomal location to ensure that chromosome copy number has not changed. A major concern is that several widely used *C. albicans* laboratory strains (e.g., BWP17 and CAI4) are aneuploid. We have determined the karyotype of the most commonly used *C. albicans* strains (Table 1). Importantly, the auxotrophic strains SN76, SN95, SN152, and SN148 that were generated from RM1000 #2 (137) are disomic, with no detectable aneuploidies (Fig. 5) (7). These strains are recommended for use in future strain construction and mutational analyses.

**MODELS FOR HOW ANEUPLOIDY AND LOH ARISE**

There are several models for how aneuploidy arises in a cell and how such events could simultaneously result in whole-chromosome and segmental LOH. First, polyploidy can be generated by mating, protoplast fusion, or defects in cell cycle events, such as mitosis or replication, as described above. Polyploid cells most likely lose extra chromosomes via nondisjunction events. What is not clear is whether such CChrL events are similar in tetraploids that arose from mating or from other defects. Furthermore, it is unclear whether CChrL is a regulated process, akin to aberrant meiosis, or whether it is simply a breakdown in centromere/spindle functions and/or in the checkpoints that monitor them (described below). In any case, this process results, primarily, in whole-chromosome aneuploidy. Second, mitotic nondisjunction, the failure of sister chromatids to separate during mitosis, will yield two aneuploid cells: one that is monosomic and thus necessarily homozygous for that chromosome and the other that is trisomic and heterozygous (with a 2:1 allelic ratio) for the same chromosome (Fig. 2B). Whole-chromosome homozygosis arises by reexpression of the monosomic chromosome or by loss of the heterozygous homolog in a trisomic strain. The mechanisms that allow these events to bypass cell cycle checkpoints, which normally inhibit such events, have not been explored.

Chromosome instability could be due to defects involving centromeres or the kinetochore proteins that associate with them. This may be especially relevant in *C. albicans*, as it has small regional centromeres. *C. albicans* centromeres were first identified using chromatin immunoprecipitation of the centromere-specific histone H3 isofrom CENP-A/CENH3/Cse4p (164). Because centromere sequences are not conserved between chromosomes and because it was not possible to establish centromere function from naked centromere DNA, Baum and colleagues proposed that the centromeres, like centromeres in most organisms, are epigenetic, lacking a specific DNA binding site (11). Indeed, deletion of a centromere leads to formation of neocentromeres at ectopic loci (96), a property of regional centromeres in humans, flies, and *Schizosaccharomyces pombe* (reviewed in reference 127). Most regional centromeres are very large (hundreds to thousands of megabase pairs in length), include long tracts of highly repetitive DNA, form pericentric heterochromatin, and attach to multiple microtubules per kinetochore. In contrast, *C. albicans* centromeres are small (~3 kb), are not flanked by highly repetitive DNA, and attach to a single microtubule per kinetochore (89). *S. cerevisiae* chromosomes, which are also tethered by a single microtubule, have point centromeres, which form kinetochores at distinct, conserved DNA sequence elements that confer centromere function. Thus, *C. albicans* lacks the DNA binding sequence found in *S. cerevisiae* point centromeres and the pericentric heterochromatin found at other regional centromeres. It is tempting to speculate that *C. albicans* centromeres may be more unstable than centromeres in other organisms and thus may be responsible for a significant proportion of the aneuploidy detected in *C. albicans* cells.

While genetic and genomic diversity is common in *C. albicans*, several questions remain to be answered: How is this diversity generated, and how does the fungus tolerate such diversity while maintaining a fully functional genome? And are most gross chromosomal rearrangements and LOH events beneficial, or do these changes confer a fitness cost that causes strains to be less fit in the host?

**CANCER RELEVANCE: TOLERANCE OF POLYPLOIDY AND ANEUPLOIDY**

Aneuploidy in human cancers is very prevalent yet is not well understood (reviewed in references 27 and 128). Some aneuploid genomes promote tumor development, while other aneuploidies suppress tumor development (197, 198). Aneuploidy and copy number variations (CNVs) are increasingly common during the acquisition of drug resistance in cancer cells, suggesting that aneuploidy increases the probability that a drug-resistant cell will arise within a tumor (148, 167, 194). Indeed, the degree of aneuploidy in nondiploid cancers often correlates with the severity of the disease (196).

While haploid *S. cerevisiae* has been used to systematically
### TABLE 1. Karyotypes, copy number variations, and genotypes of common *C. albicans* laboratory strains

| Strain          | CHEF (reference)                        | aCGH and sequence data (reference, source, and/or figure) | Strain, genotype, or description | Original strain reference |
|-----------------|-----------------------------------------|-----------------------------------------------------------|---------------------------------|---------------------------|
| SCS314          | Standard karyotype                      | Disomic (20, 91, 172, 190)                                | Wild type                       | 63                        |
| CAF2-1          | Standard karyotype                      | Disomic (172)                                             |                                 |                           |
| CAJ4            | Standard karyotype                      | Trisomic Chr1 and -2 or just Chr2 (172)                   |                                 |                           |
| RM1             | Standard karyotype                      | Disomic (172)                                             |                                 |                           |
| RM10            | Standard karyotype                      | Disomic (172)                                             |                                 |                           |
| RM100 #13       | Truncated Chr5b                         | Segmental monosomy Chr5 (HIS1→telomere) (172)             |                                 |                           |
| RM1000 #6       | Truncated Chr5b                         | Segmental monosomy Chr5 (HIS1→telomere) (172)             |                                 |                           |
| RM1000 #2       | Standard karyotype                      | Disomic, homozygous Chr5 (HIS1→telomere) (172)            |                                 | F. Navarro-Garcia         |
| SN87            | Standard karyotype                      | Disomic (Fig. 5) (7)                                      | leu2Δ/leu2Δ his1Δ/his1Δ URA3/ura3Δ::imm434 IRO1/iro1Δ::imm434 | 137                       |
| SN95            | Standard karyotype                      | Disomic (Fig. 5) (7)                                      | arg4Δ/arg4Δ his1Δ/his1Δ URA3/ura3Δ::imm434 IRO1/iro1Δ::imm434 | 137                       |
| SN152           | Standard karyotype                      | Disomic (Fig. 5) (7)                                      | arg4Δ/arg4Δ his1Δ/his1Δ URA3/ura3Δ::imm434 IRO1/iro1Δ::imm434 | 137                       |
| WO-1            | Six new fusion chromosomes due to three reciprocal translocations (35) | Disomic (Fig. 5) (22) or trisomic Chr1 (149)              | Prototrophic                    | 177                       |
| WO-2            | Loss of the full-length Chr7 homolog and the fusion Chr5,6 (114) | Segmental trisomy Chr4, segmental monosomies Chr5, -6, -7 (Fig. 5) | Prototrophic                    | 114                       |
| 3153A           | Altered ChrR, unstable Chr5 and -6 (reviewed in reference 157); translocation involving Chr4R and Chr7 (87) | Disomic (Fig. 5) (Magee Lab collection via Dexter Howard) | Prototrophic                    | 43                        |
| SGY-243         | Trisomic for CHS2 (ChrR) locus (64); trisomic Chr1 (28, 111); 4 copies of Chr 4 and -5, trisomic Chr6 and -7 (reviewed in reference 157) | Trisomic ChrR and -2, segmental trisomy Chr4R (Fig. 5) (Scherer Lab collection via Squibb and Myra Kurtz) | ade2/ade2 Δura3::ADE2/Δura3::ADE2 | 95                        |

* Strain 3153A (43) (and MRL 3153 of the Mycology Reference Laboratory, London, United Kingdom) is the same as the B311 strain of H. F. Hasenclever (see, e.g., reference 69) and is deposited in the American Type Culture Collection as ATCC 32354.
study aneuploidy, aneuploidy of single or multiple chromosomes strongly correlates with a high fitness cost in this organism (measured as slow growth relative to the euploid parental strain) (187). In contrast, C. albicans is more tolerant of aneuploidy: in some cases multiply aneuploid Flu6 strains do not incur a fitness cost in the absence of drug (175). Thus, the response of diploid C. albicans is more similar to the situation in cancer cells, which are often polyploid and/or highly aneuploid (181) yet continue to divide rapidly. We propose that C. albicans may be an excellent model yeast for the study of how aneuploidy arises and why it is tolerated in cancer cells.

A central question in the field is why C. albicans is maintained as a heterozygous diploid. In practice, different clinical isolates of C. albicans are homozygous for different chromosome arm regions (22), such that LOH certainly occurs during growth in the host. Since different chromosomes can be monosomic, diploidy does not appear to be maintained only to suppress lethal recessive mutations (53). Nonetheless, most strains that are homozygous for whole chromosomes do have growth phenotypes (55, 157). While some monosomies provide a selective advantage in the presence of a specific stress (87), other monosomic chromosomes cause filamentous growth or slow-growth phenotypes (10), and in the absence of selection, reduplication of monosomic chromosomes is associated with faster growth (87), suggesting that an imbalance of gene products is generally stressful to cells (187). Additionally, a fully diploid genome is not required for certain cellular processes, such as maintenance in the host, the yeast-to-hyphal transition, and white-opaque switching, as shown by the segmental monosomic, diploidy does not appear to be maintained only to suppress lethal recessive mutations (53). Nonetheless, most strains that are homozygous for whole chromosomes do have growth phenotypes (55, 157). While some monosomies provide a selective advantage in the presence of a specific stress (87), other monosomic chromosomes cause filamentous growth or slow-growth phenotypes (10), and in the absence of selection, reduplication of monosomic chromosomes is associated with faster growth (87), suggesting that an imbalance of gene products is generally stressful to cells (187). Additionally, a fully diploid genome is not required for certain cellular processes, such as maintenance in the host, the yeast-to-hyphal transition, and white-opaque switching, as shown by the segmental monosomies detected in clinical isolates as well as in the switching strain WO-2 (114). Thus, like cancer cells, somatic defects associated with aneuploidy. Consistent with this, in controlled evolution experiments, diploid S. cerevisiae cells often acquired large aneuploidies, while isogenic haploid strains underwent only shorter-range CNVs (67).

Importantly, aneuploidy and CNVs appear with a much higher frequency than point mutations in yeast cells (110). Additionally, it appears that many types of fungi, including S. cerevisiae, may resort to generating polyploidy or aneuploidy in response to severe stress (61, 71, 145, 150, 170). At least in some cases, this aneuploidy provides a selective advantage in the presence of the stress (67, 150, 173, 175). Furthermore, since polyploidy and aneuploidy can arise rapidly, through mating, unlicensed rerepllication, and/or mitotic nondisjunction, altering whole-chromosome or chromosome segment copy number may be a common mechanism for responding to stress. An important question that will certainly be answered with new deep-sequencing analysis of evolved isolates of different organisms is how frequently aneuploidy is employed as a general mechanism to rapidly and flexibly adapt to changing environmental assaults.

ACKNOWLEDGMENTS

This work was supported by NIH NIAID R01AI0624273 and NIAID R01AI075906 to J.B. and by a Microbial and Plant Genomics Institute of Minnesota Integrative Fellowship to A.S.

We are grateful to Elena Rusektenko, P. T. Magee, and B. B. Magee for providing unpublished data and comments on the manuscript.

REFERENCES

1. Ahmad, A., M. A. Kahir, A. Kravets, E. Andaluz, G. Larriba, and E. Rusektenko. 2008. Chromosome instability and unusual features of some widely used strains of Candida albicans. Yeast 25:433–448.
2. Alby, K., D. Schafer, and R. J. Bennett. 2009. Homothallic and heterothallic mating in the opportunistic pathogen Candida albicans. Nature 460: 893–893.
3. Almeida, R. S., Z. Storchova, C. Styles, T. Galitski, D. Pellman, and G. R. Fink. 2004. Defects arising from whole-genome duplications in Saccharomyces cerevisiae. Genetics 167:1109–1121.
4. Andaluz, E., J. Gomez-Raja, B. Hermosa, T. Ciudad, E. Rusektenko, R. Calderone, and G. Larriba. 2007. Loss and fragmentation of chromosome 5 in Candida albicans and C. glabrata. J. Gen. Microbiol. 153:2531–2538.
5. Barton, R. C., and K. Gull. 1992. Isolation, characterization, and genetic analysis of monosomic, aneuploid mutants of Candida albicans. Mol. Microbiol. 6:171–177.
6. Baum, M., K. Sanyal, P. K. Mishra, N. Thaler, and J. Carbon. 2006. Formation of functional centromeric chromatin is specified epigenetically in Candida albicans. Proc. Natl. Acad. Sci. U. S. A. 103:14877–14882.
7. Bennett, R. J., and A. D. Johnson. 2003. Completion of a parasexual cycle in Candida albicans by induced chromosome loss in tetraploid strains. EMBO J. 22:2505–2515.
8. Bennett, R. J., and A. D. Johnson. 2005. Matings in Candida albicans and the search for a sexual cycle. Annu. Rev. Microbiol. 59:233–255.
9. Bennett, R. J., M. G. Miller, P. R. Chua, M. E. Mazan, and A. D. Johnson. 2000. Nuclear fusion occurs during mating in Candida albicans strains that are homothallic and heterothallic. Proc. Natl. Acad. Sci. U. S. A. 105:11845–11850.
10. Asakura, K., S. Iwaguchi, M. Homma, T. Sukai, K. Higashide, and K. Tanaka. 1991. Electrophoretic karyotypes of clinically isolated yeasts of Candida albicans and C. glabrata. J. Gen. Microbiol. 137:2531–2538.
riman, J. Heitman, N. A. Gow, M. C. Lorenz, B. W. Birren, M. Kellis, and C. A. Cuomo. 2009. Evolution of pathogenicity and sexual reproduction in eight Candida genomes. Nature 459:657–662.

23. Campstain, C., A. M. Wijmants-Collin, and C. Logan. 2007. Reverse genetic analysis of the yeast RSC3 (RSC1) complex reveals a role for RSC3 and SNF5 homolog 1 in ploidy maintenance. PLoS Genet. 3:e92.

24. Carle, G. F., and M. V. Olson. 1984. Separation of chromosomal DNA molecules from yeast by orthogonal-field-alternation gel electrophoresis. Nucl. Acids Res. 12:5467–5464.

25. Casadavall, A., and L. A. Pirofski. 1999. Host-pathogen interactions: redefining the basic concepts of virulence and pathogenicity. Infect. Immun. 67:3703–3713.

26. Chen, C. S., and D. Botstein. 1993. Isolation and characterization of chromosome-gain and increase-in-ploidy mutants in yeast. Genetics 135:677–691.

27. Chanhok, N. S., and D. Pellman. 2009. A little CIN may cost a lot: resisting aneuploidy and cancer. Curr. Opin. Genet. Dev. 19:74–81.

28. Chen, X., B. B. Magee, D. Dawson, P. T. Magee, and C. A. Kumamoto. 2004. Chromosome 1 trisomy compromises the virulence of Candida albicans. Mol. Microbiol. 51:551–565.

29. Chial, H. J., and M. Winey. 1999. Mechanisms of genetic instability revealed by analysis of yeast spindle pole body duplication. Biol. Cell 91:439–450.

30. Chibana, H., J. L. Beckerman, and P. T. Magee. 2000. Fine-resolution physical mapping of genomic diversity in Candida albicans. Genome Res. 10:1655–1677.

31. Chibana, H., S. Iwaguchi, M. Homma, A. Chindamporn, Y. Nakagawa, and K. Tanaka. 1994. Diversity of tandemly repetitive sequences due to short periodic repetitions in the chromosomes of Candida albicans. J. Bacteriol. 176:3851–3858.

32. Chibana, H., B. B. Magee, S. Grindle, Y. Ran, S. Scherer, and P. T. Magee. 1998. A physical map of chromosome 7 of Candida albicans. Genetics 149:1739–1752.

33. Chindamporn, A., Y. Nakagawa, M. Homma, H. Chibana, M. Doi, and K. Tanaka. 1995. Analysis of the chromosomal localization of the repetitive sequences (RPSs) in Candida albicans. Microbiology 141:469–476.

34. Chindamporn, A., Y. Nakagawa, I. Mizuguchi, H. Chibana, M. Doi, and K. Tanaka. 1995. Repetitive sequences (RPSs) in the chromosomes of Candida albicans are sandwiched between two novel stretches, HOK and RBD2, common to each chromosome. Microbiology 144:849–857.

35. Chu, W. S., B. B. Magee, and P. T. Magee. 1993. Construction of an Sfl microrestriction map of the Candida albicans genome. J. Bacteriol. 175:6637–6651.

36. Clarke, D. J., and J. Bachtant. 2008. Kinetochoore structure and spindle assembly checkpoint signaling in the budding yeast, Saccharomyces cerevisiae. Front. Biosci. 13:6787–6819.

37. Coste, A., A. Selmecki, A. Forche, D. Diogo, M. E. Bougnoux, C. d’Enfert, J. Morschhauser, A. Forche, A. Selmecki, V. Turner, F. Ischer, and A. Forche. 2007. Genotypic evolution of azole resistance in Candida albicans isolates during a single passage through a mouse host. Genetics 182:799–811.

38. Forche, A., G. May, and P. T. Magee. 2005. Demonstration of loss of heterozygosity by single-nucleotide polymorphism microarray analysis and alterations in strain morphology in Candida albicans strains during infection. Eukaryot. Cell 4:156–165.

39. Forche, A., G. Schonian, Y. Graser, R. Vilgalys, and T. G. Mitchell. 1999. Genetic structure of typical and atypical populations of Candida albicans from Africa. Fungal Genet. Biol. 29:124–134.

40. Forche, A., M. Steinbach, and J. Berman. 2009. Efficient and rapid identification of Candida albicans allelic status using SNP-RFLP. FEMS Yeast Res. 9:1061–1069.

41. Fridkin, C., and B. Hube. 2003. Tissue infection and site-specific gene expression in Candida albicans, Adv. Appl. Microbiol. 53:271–290.

42. Fridkin, C., M. Kretschmar, T. NichteLeaner, C. Gaillardin, C. D’Enfert, and B. Hube. 2003. Stage-specific gene expression of Candida albicans in human blood. Mol. Microbiol. 48:1525–1543.

43. Frasier, J. A., J. C. Huang, R. Pukkila-Worley, J. A. Alspaugh, T. G. Mitchell, and J. Heitman. 2005. Chromosomal translocation and segmental duplication in Cryptococcus neoformans. Eukaryot. Cell 4:401–406.

44. Fu, X. H., F. L. Meng, Y. Hu, and J. Q. Zhou. 2008. Candida albicans, a distinctive fungal model for cellular aging study. Aging Cell 7:746–757.

45. Ganley, A. R., S. Ide, K. Saka, and T. Kobayashi. 2009. The effect of replication initiation on gene amplification in the DNA and its relationship to aging. Mol. Cell 36:683–693.

46. Gillies, A. M., E. Y. Tsang, and R. Kirsch. 1984. Isolation of the Candida albicans gene for orotidine-5-phosphate decarboxylase by complementation of S. cerevisiae ura3 and E. coli proP mutations. Mol. Gen. Genet. 198:179–182.

47. Gitlin, A. N., P. W. Robbins, J. W. Lester, A. J. Brown, W. A. Fonzi, T. Chapman, and O. S. Kinsman. 1994. A hyphal-specific chitin synthase gene (CHS2) is not essential for growth, dimorphism, or virulence of Candida albicans. Proc. Natl. Acad. Sci. U. S. A. 91:6216–6220.

48. Gräser, Y., M. Volovsek, J. Arrington, G. Schonian, W. Presber, T. G. Mitchell, and R. Vilgalys. 1996. Molecular markers reveal that population structure of the human pathogen Candida albicans exhibits both clonality and recombination. Proc. Natl. Acad. Sci. U. S. A. 93:12473–12477.

49. Gresham, D., B. Curry, A. Ward, D. B. Gordon, L. Brizuela, L. Kruglyak, and D. Botstein. 2010. Optimized detection of sequence variation in heterozygous genomes using DNA microarrays with isothermal-melting probes. Proc. Natl. Acad. Sci. U. S. A. 107:124–1248.

50. Gresham, D., M. M. Desai, C. M. Tucker, H. T. Jenq, D. A. Pai, A. Ward, C. Delavo, D. Botstein, and M. J. Dunham. 2008. The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. PLoS Genet. 4:e1000303.

51. Haase, S. B., and D. J. Lew. 1997. Flow cytometric analysis of DNA content in budding yeast. Methods Enzymol. 282:522–532.

52. Hasencluber, H. F. 1965. The antigens of Candida albicans. Am. Rev. Respir. Dis. 92:150–158.

53. Hilton, C., D. Markie, B. Corner, E. Rikkerink, and R. Poulter. 1985. Heat shock induces chromosome loss in the yeast Candida albicans. Mol. Gen. Genet. 200:162–168.

54. Hu, G., I. Liu, A. Sham, J. E. Stajich, F. S. Dietrich, and J. W. Kronstad. 2008. Comparative hybridization reveals extensive genome variation in the AIDS-associated pathogen Cryptococcus neoformans. Genome Biol. 9:R41.

55. Hu, B. 2000. Extracellular proteases of human pathogenic fungi. Curr. Microbiol. 51:126–137.

56. Hube, B. 2004. From commensal to pathogen: stage- and tissue-specific gene expression of Candida albicans. Curr. Opin. Microbiol. 7:336–341.
Anna Selmecki received her B.A. in chemistry from the College of Saint Benedict/Saint John’s University, Saint Joseph, MN, and her Ph.D. in molecular, cellular, and developmental biology and genetics from the University of Minnesota, Minneapolis. She is currently a postdoctoral fellow at the Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA. Her research focuses on the evolutionary consequences of aneuploidy and polyploidy and seeks a better understanding of the balance between beneficial and deleterious genome changes during adaptive evolution of yeast and its relation to the development of cancer.

Anja Forche received her B.S. in food technology from Humboldt University, Berlin, Germany, and her Ph.D. from the Technical University in Berlin, Germany. She did her postdoctoral work at Duke University and at the University of Minnesota. She is currently a research assistant professor in the Department of Biology at Bowdoin College. Her research focuses on host-pathogen interactions of Candida albicans and its human host and in particular how the host genome changes to facilitate adaptation of C. albicans to different host niches.
Judith Berman received her B.S. in plant pathology from Cornell University, Ithaca, NY, and her Ph.D. from the Weizmann Institute of Science, Rehovot, Israel. Her postdoctoral work was in genetics and biochemistry at Cornell University. She joined the University of Minnesota faculty in 1986, where she is now Distinguished McKnight University Professor in the Department of Genetics, Cell Biology and Development and the Department of Microbiology. She is a fellow of the American Academy of Microbiology and of the American Association for the Advancement of Science. Her research seeks to answer fundamental questions concerning how eukaryotic genomes are stably maintained and how they change in response to stresses such as antifungal drugs, using approaches of basic cell biology, genetics, and genomics. A major focus is Candida albicans, an opportunistic commensal that is the most frequent cause of nosocomial fungal infections and is of particular concern among those with weakened immune function.