

Review

To Repeat or Not to Repeat: Repetitive Sequences Regulate Genome Stability in *Candida albicans*

Matthew J. Dunn¹ and Matthew Z. Anderson^{1,2,*} ¹ Department of Microbiology, The Ohio State University, Columbus, OH 43210, USA; dunn.658@osu.edu² Department of Microbial Infection and Immunity, The Ohio State University, Columbus, OH 43210, USA

* Correspondence: anderson.3196@osu.edu; Tel.: +614-247-0058

Received: 30 September 2019; Accepted: 23 October 2019; Published: 30 October 2019



Abstract: Genome instability often leads to cell death but can also give rise to innovative genotypic and phenotypic variation through mutation and structural rearrangements. Repetitive sequences and chromatin architecture in particular are critical modulators of recombination and mutability. In *Candida albicans*, four major classes of repeats exist in the genome: telomeres, subtelomeres, the major repeat sequence (MRS), and the ribosomal DNA (rDNA) locus. Characterization of these loci has revealed how their structure contributes to recombination and either promotes or restricts sequence evolution. The mechanisms of recombination that give rise to genome instability are known for some of these regions, whereas others are generally unexplored. More recent work has revealed additional repetitive elements, including expanded gene families and centromeric repeats that facilitate recombination and genetic innovation. Together, the repeats facilitate *C. albicans* evolution through construction of novel genotypes that underlie *C. albicans* adaptive potential and promote persistence across its human host.

Keywords: genome stability; telomere; subtelomere; gene family expansion; LTR; MRS; *Candida albicans*

1. Introduction

Considerable variability in genome organization exists across the tree of life, ranging from single circular chromosomes in bacterial species to thousands of linear chromosomes in some ferns [1]. These changes in chromosome complement arise through underlying processes of genome instability, giving way to structural reorganization and locus shuffling. Emergence of novel genetic and karyotypic arrangements can increase phenotypic diversity within a population and provide adaptive solutions to emerging selective pressures. Yet, extreme forms of genome plasticity can lead to DNA fragmentation, chromosome loss, and cell death. Thus, a balance between prohibitive and unrestricted genome instability is required to maintain DNA integrity while allowing evolution to produce novel and potentially advantageous genotypes.

Repetitive regions of the genome often serve as hotspots of genome rearrangements and evolutionary innovation. These elements can exist as single or multi-copy loci encoding tandemly-repeated identical or near-identical sequences of variable length ranging from one to thousands of nucleotides. By their very nature, repetitive loci are prone to recombination, producing insertion/deletions (indels) and translocations between both adjacent and distal repeats. Recombination within repeats alters copy number variants in addition to promoting de novo mutations due to errors in DNA exchange and repair [2]. In contrast, recombination between distal, non-allelic elements can result in large-scale chromosomal aberrations such as fusion events, truncations, and translocations. Genetic exchange between repeats does not require perfect sequence identity and commonly occurs between imperfect repetitive sequences [3]. Repetitive loci are also particularly prone to replication fork collapse and environmental damaging agents, including oxidative stress and ultraviolet irradiation, that promote further mutation and recombination [4–6]. Condensed heterochromatin across repetitive regions of the genome is thought to help shield these regions from

DNA damage, but targeted investigation of specific loci has revealed that open chromatin often borders these repetitive loci, suggesting they are susceptible to these same mutagenic forces [7]. Recent work has further complicated the relationship between chromatin architecture and DNA damage susceptibility as the damaging agent and chromatin status of the DNA both contribute to altered susceptibility to double strand breaks, mutation, and recombination [8,9]. Thus, a more complex interplay between the locus and genetic insult determine the likelihood for mutations to accrue and recombination to take place.

Fungi are well established model systems for the investigation of cellular, molecular, and genetic processes [10]. Even prior to the emergence of genome sequencing, fungi were pilot organisms for the analysis of repetitive DNA, including the investigation of cornerstone sequences such as centromeres and telomeres [11–15]. Progression into the genomics era led to rapid advances in the catalogue of available fungal genomes for mining and deciphering common features of functional repetitive sequences [16–18]. In particular, budding yeast genomes provided much of the initial understanding of eukaryotic genome architecture and intra-species sequence diversity because of their relative smaller sizes and well-defined gene features. The prevalent fungal commensal and opportunistic human pathogen *Candida albicans* possesses many of these same hallmark attributes. The diploid *C. albicans* genome is comprised of eight chromosomes numbered 1 through 7 by size with the exception of chromosome R (ChrR), which encodes the rDNA locus [19–21]. *C. albicans* isolate SC5314, the genome reference strain, possesses approximately 6400 genes over ~14 megabases with one single nucleotide polymorphism (SNP) approximately every 300 basepairs (bp) between chromosome homologs [22,23]. Although *C. albicans* is most commonly isolated as a heterozygous diploid, strains are capable of propagating with altered ploidy, copy number variants, and patterns of heterozygosity [24–26]. Some of these genotypic changes confer selective advantages to specific environmental conditions in vitro and in vivo, including the presence of antifungal drugs and the oral cavity [25,27–32]. Thus, genomic rearrangements may contribute to the success of *C. albicans* colonization, persistence, and adaptability across host niches.

Four major repetitive regions are present in the *C. albicans* genome: telomeres, subtelomeres, the Major Repeat Sequence (MRS), and the rDNA locus (Figure 1). Additional repetitive sequences are interspersed throughout the genome where they contribute to the collective process of chromosome instability and genome evolution (summarized in Table 1). Here, we describe these repetitive sequences and highlight the role they play in genome dynamics in this important model organism and human fungal pathogen.

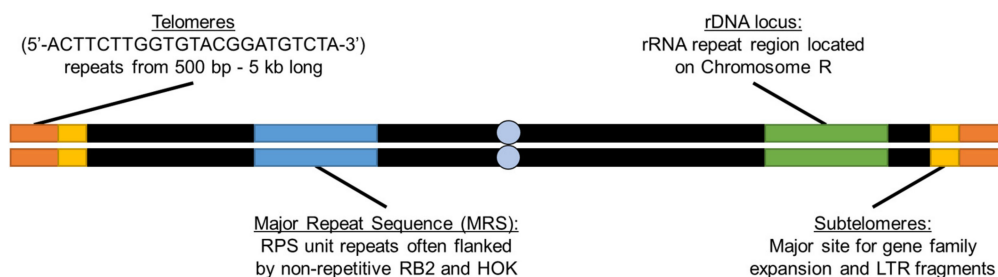


Figure 1. Major classes of genetic repeats in *Candida albicans*. *C. albicans* contains four major categories of repeat sequences: the telomeres that contain multiple copies of a 23 bp repeat; the Major Repeat Sequence (MRS) composed of repetitive sequence (RPS) repeats; the rDNA locus, which encodes the polycistronic rRNA transcripts; and the subtelomeres, a telomere-proximal region containing transposable elements and a gene family expansion. Centromeres are indicated by grey circles.

Table 1. Repeat composition summary in *C. albicans*.

Repeat Element	Repeat Composition	Size	Repeat Copy Number (Haploid Genome)	Mechanisms of Instability
Telomeres	5'-ACTTCTTGGTGTACGGATGTCTA-3'	500 bp–5 kb	>20 repeats	Recombination, t-circles
Subtelomeres	Expanded gene families, LTRs, non-LTR retrotransposons	15 kb proximal to the telomeric repeats	14 <i>TLO</i> genes, 13 LTRs, 1 non-LTR retrotransposon	Recombination, LTR insertions, LOH, copy number variations
Centromeres	Unique regions without common sequence motifs, inverted repeats	Core 3 kb regions	Eight loci	Interchromosomal recombination, isochromosome formation
Major Repeat Sequence	RPS repeat units often flanked by non-repetitive RB2 and HOK elements	50 kb on average	Nine complete MRS loci, 14 RB2 elements, and two HOK elements	Chromosome translocations, intrachromosomal recombination
rDNA locus	18 s, 5.8s, 25 s, and 5 s rRNAs organized as tandem repeating units	11.6 kb - 12.5 kb per <i>RDN1</i> locus	21 to 176 copies of repeating <i>RDN1</i> locus	Intrachromosomal recombination
ALS adhesin family	Tandemly repeated Ser/Thr-rich domain	~3–6 kb	Eight genes	Intragenic recombination, intergenic recombination
Long repeat sequences	20 bp or longer repeats present more than once in the genome	65–6499 bp	1974 long repeats (2.87% of the haploid reference genome)	Recombination, LOH, chromosomal inversions, copy number variations

2. Telomeres and Recombination

C. albicans maintains the ends of its linear chromosomes through the use of telomeres that are composed of repeating subunits produced by the enzyme telomerase. The telomeric repeats of most eukaryotes are built from a repeating 6 bp GT motif, but many fungal species including the closely-related Saccharomycotina yeasts use repeats of diverse lengths and complexity [33]. In *C. albicans*, the telomeric repeat is 23 bp (5'-ACTTCTTGGTGTACGGATGTCTA-3'), and has diverged substantially from other *Candida* species by increasing in both length and complexity [34]. Other *Candida* CUG paraphyletic species, such as *Candida guilliermondii* and *Debaryomyces hansenii*, decode the "CUG" codon as serine instead of leucine and use shorter telomeric repeats of 5'-ACTGGTGT-3' and 5'-ATGTTGAGGTGTAGGG-3', respectively [35,36]. Telomerase reverse transcriptase (TERT) and the telomerase RNA component (TERC) work together to extend telomeres by adding single 23 bp repeats to nascent chromosome ends to form the full telomere, which can range in size from 500 bp to 5 kilobases (kb) in *C. albicans* laboratory isolates [37,38].

2.1. Telomere Structure and Maintenance

The telomerase complex, which includes TERT and TERC, forms the basic functional unit required for telomere repeat production and maintenance [39–41]. Both TERC and TERT, encoded by *TER1* and *EST2*, respectively, perform repeated rounds of reverse transcription to build G overhangs as part of the end-protective T-loops that prevent telomere attrition [42]. More specifically, Est2 uses the *TER1* RNA as a template to add telomeric repeat subunits to the chromosome end in order to maintain replicative capacity and avoid cell senescence. The dual roles of end protection and telomere repeat addition by Est2 are separable. A catalytically inactive Est2 retains the ability to suppress excessive G-strand accumulation despite being unable to add telomeric repeats [42]. Two additional telomerase protein subunits, Est1 and Est3, also contribute to appropriate telomere structural maintenance by suppressing excessive recombination and preventing telomere loss although their precise molecular mechanisms remain obscure [43].

Maintenance of telomeric repeats in *C. albicans* requires the heterotrimeric CST complex comprised of Cdc13, Stn1, and Ten1. Loss of any of these factors results in dysregulation of telomere length, resulting in either overextended or shrunken telomeres [37]. This contrasts with inactivation of the telomerase complex that only leads to iterative telomere shortening due to the inability to add telomeric repeats. The canonical Ku70/Ku80 complex also controls telomere length and loss of either gene produces heterogeneous telomere lengths indicative of dysfunctional telomerase activity [44,45]. Rap1 (repressor/activator protein 1), an essential protein in telomere maintenance in other budding yeasts, is surprisingly not essential in *C. albicans* despite also functioning in telomere length regulation [46,47].

2.2. Telomere Replication and Recombination

Telomere maintenance through recombination serves a critical function in chromosome stability but can also contribute directly to genome instability during telomere length shortening. As such, the actions of telomere length control and protection from overhang accumulation must be tightly controlled. Telomere length can be regulated through addition of telomeric repeats following their loss during DNA replication or via recombination within single telomeres or between chromosomes [48]. Alternative lengthening of telomeres (ALT), a process that is telomerase-independent but recombination-dependent, adds telomeric repeats to chromosome ends by using pre-formed telomeres as a template [49]. The 3' overhang of the extending telomere invades other telomeric DNA that can be either: within the same telomere through telomere looping, another chromosome's telomere, or an extrachromosomal telomere circle. Telomere circles (t-circles), autonomous plasmids composed entirely of telomere repeat subunits, can be produced as a byproduct of telomere length regulation via intra-chromosomal recombination events [50]. In *C. albicans*, t-circle frequency increased following loss of Ku70, suggesting these proteins suppress intra-telomeric recombination [44]. However, no direct studies of ALT have been performed

in *C. albicans* despite the presence of homologous genes that appear capable of conducting these functions (e.g., Rad52 [45,49,51]).

Conserved telomeric proteins are required for telomere length regulation. For example, Rap1 likely plays a role in this process, as its deletion produced aberrant telomere repeat-containing DNA structures and t-circles [47]. Yet, all observed nuclear telomere lengthening events in *C. albicans* have occurred through concatenation of linear elements by *TER1* [52], leaving the function of Rap1 in t-circle production and ALT unclear. Our current understanding of the recombination machinery (e.g., Rad52, etc.) and precise mechanisms for telomere maintenance are severely deficient in *C. albicans* as little direct experimental work has been performed in recent decades. Investigations into *C. albicans* telomere biology, in particular, stands out for its potential scientific gains given the paucity of existing data combined with the common involvement of telomeres in contributing to genome plasticity. Newly developed linear plasmids containing in *C. albicans* share similar mechanisms of linear telomere formation and maintenance for their retention and propagation and may present as a simplified model to study telomere dynamics in *C. albicans* [53].

3. *Candida albicans* Subtelomeres Are Hotspots of Genomic Rearrangements

Subtelomeres are defined as the genomic regions adjacent to telomeric repeats that are enriched for repetitive genetic elements [54]. Repetitive sequences within the subtelomeres include intact transposable elements and expanded gene family members as well as remnants of these functional units following gene disruption or inactivation. Formation of heterochromatin conducive to epigenetic silencing is also a common feature of subtelomeres, although it may be interspersed with euchromatic regions surrounding intact genes. In *C. albicans*, the histone deacetylase Sir2 maintains the heterochromatin within subtelomeres [55,56]. Surprisingly, Sir2 appears to perform these roles without other Sir complex proteins typically required for subtelomeric silencing through “telomere position effect”, as they are absent from the *C. albicans* genome [57]. The *SIR2* paralog, *HST1*, also regulates subtelomeric gene expression, although these effects are less uniform across *C. albicans* subtelomeres with only some loci being affected in a $\Delta/\Delta hst1$ background [58]. These combined effects of heterochromatin formation and gene silencing in *C. albicans* were recently shown to be strongest within the proximal 10–15 kb of the telomere repeats [59]. As such, we choose to define the *C. albicans* subtelomeric region as the most telomere-proximal 15 kb of each chromosome arm.

3.1. Subtelomeric Gene Functions

C. albicans subtelomeres are relatively gene poor compared to chromosome-internal regions of the genome. Approximately 50 protein-coding genes are annotated within *C. albicans* subtelomeres in the genome reference strain SC5314 based on the above definition from the *Candida* Genome Database [60], but most genes remain largely uncharacterized (Table 2). Most of the subtelomeric genes with functional assignments are either associated with biofilm formation (23 of 55), growth in Spider or other hyphae-inducing media (22 of 55), or have predicted roles in metabolism (21 of 55; Table 2). Coincidentally, many of these genes are also assigned roles in virulence either by directly promoting pathogenicity or through adaptive responses promoting persistence across host niches. Genes that may be used in niche adaptation include multiple transporters, cell wall proteins, and metabolite utilization genes.

Table 2. *C. albicans* characterized subtelomeric genes.

Gene Name	Description ¹	Genomic Location
<i>TLO</i> α 1	Member of a family of telomere-proximal genes of unknown function; hypha-induced expression; rat catheter biofilm repressed	Ca22chrRA_C_albicans_SC5314:9111 to 9863
INO1	Inositol-1-phosphate synthase; antigenic in human; repressed by farnesol in biofilm or by caspofungin; upstream inositol/choline regulatory element; glycosylation predicted; rat catheter, flow model induced; Spider biofilm repressed	Ca22chrRA_C_albicans_SC5314:2157044 to 2155482
<i>TLO</i> β 2	Putative transcription factor/activator; Med2 mediator complex domain; transcript is upregulated in an RHE model of oral candidiasis; member of a family of telomere-proximal genes; Efg1, Hap43-repressed	Ca22chrRA_C_albicans_SC5314:2286198 to 2285377
ATP16	Subunit of the mitochondrial F1F0 ATP synthase; sumoylation target; protein newly produced during adaptation to the serum; Spider biofilm repressed	Ca22chrRA_C_albicans_SC5314:2285228 to 2284743
XYL2	D-xylulose reductase; immunogenic in mice; soluble protein in hyphae; induced by caspofungin, fluconazole, Hog1 and during cell wall regeneration; Mnl1-induced in weak acid stress; stationary phase enriched; flow model biofilm induced	Ca22chrRA_C_albicans_SC5314:2284454 to 2283372
MAL2	Alpha-glucosidase; hydrolyzes sucrose for sucrose utilization; transcript regulated by Suc1, induced by maltose, repressed by glucose; Tn mutation affects filamentous growth; upregulated in RHE model; rat catheter and Spider biofilm induced	Ca22chrRA_C_albicans_SC5314:2276745 to 2278457
IAH1	Protein similar to <i>S. cerevisiae</i> Iah1p, which is involved in acetate metabolism; mutation confers hypersensitivity to tunicamycin; transposon mutation affects filamentous growth	Ca22chrRA_C_albicans_SC5314:2276572 to 2275766
CSM1	Putative component of the monopolin complex with role in rDNA silencing, homologous chromosome segregation, protein localization to nucleolar rDNA repeats	Ca22chrRA_C_albicans_SC5314:2272097 to 2272774
<i>TLO</i> α 3	Putative transcription factor; Med2 mediator domain; activates transcription in 1-hybrid assay in <i>S. cerevisiae</i> ; repressed by Efg1; member of a family of telomere-proximal genes; Tbf1-induced	Ca22chr1A_C_albicans_SC5314:10718 to 11485
TUP1	Transcriptional corepressor; represses filamentous growth; regulates switching; role in germ tube induction, farnesol response; in repression pathways with Nrg1, Rfg1; farnesol upregulated in biofilm; rat catheter, Spider biofilm repressed	Ca22chr1A_C_albicans_SC5314:12163 to 13701
MVD	Mevalonate diphosphate decarboxylase; functional homolog of <i>S. cerevisiae</i> Erg19; possible drug target; regulated by carbon source, yeast-hypha switch, growth phase, antifungals; gene has intron; rat catheter, Spider biofilm repressed	Ca22chr1A_C_albicans_SC5314:13778 to 14917
<i>TLO</i> γ 4	Member of a family of telomere-proximal genes of unknown function; transcript induced in an RHE model of oral candidiasis; Hap43-repressed	Ca22chr1A_C_albicans_SC5314:3186663 to 3186463

Table 2. Cont.

Gene Name	Description ¹	Genomic Location
FGR24	Protein encoded in retrotransposon Zorro2 with similarity to retroviral endonuclease-reverse transcriptase proteins; lacks an ortholog in <i>S. cerevisiae</i> ; transposon mutation affects filamentous growth	Ca22chr1A_C_albicans_SC5314:3185097 to 3184381
SAM35	Predicted component of the sorting and assembly machinery (SAM complex) of the mitochondrial outer membrane, involved in protein import into mitochondria	Ca22chr1A_C_albicans_SC5314:3177333 to 3176578
GDI1	Putative Rab GDP-dissociation inhibitor; GlcNAc-induced protein; Spider biofilm repressed	Ca22chr1A_C_albicans_SC5314:3172988 to 3171639
<i>TLO</i> γ 5	Member of a family of telomere-proximal genes of unknown function; may be spliced in vivo	Ca22chr2A_C_albicans_SC5314:4248 to 4778
RRN3	Protein with a predicted role in recruitment of RNA polymerase I to rDNA; caspofungin induced; flucytosine repressed; repressed in core stress response; repressed by prostaglandins	Ca22chr2A_C_albicans_SC5314:5665 to 7335
MPP10	Putative SSU processome and 90S preribosome component; repressed in core stress response; repressed by prostaglandins	Ca22chr2A_C_albicans_SC5314:11063 to 9360
FAV3	Putative alpha-1,6-mannanase; induced by mating factor in MTLa/MTLa opaque cells	Ca22chr2A_C_albicans_SC5314:12443 to 11100
PGA52	GPI-anchored cell surface protein of unknown function; Hap43p-repressed gene; fluconazole-induced; possibly an essential gene, disruptants not obtained by UAU1 method	Ca22chr2A_C_albicans_SC5314:14415 to 13261
RDH54	Putative DNA-dependent ATPase with a predicted role in DNA recombination and repair; transcriptionally induced by interaction with macrophages	Ca22chr2A_C_albicans_SC5314:2221521 to 2223911
<i>TLO</i> γ 7	Member of a family of telomere-proximal genes of unknown function; may be spliced in vivo; rat catheter biofilm repressed	Ca22chr3A_C_albicans_SC5314:13756 to 14265
<i>TLO</i> γ 8	Member of a family of telomere-proximal genes of unknown function; may be spliced in vivo	Ca22chr3A_C_albicans_SC5314:1788114 to 1787605
SRB1	Essential GDP-mannose pyrophosphorylase; makes GDP-mannose for protein glycosylation; functional in <i>S. cerevisiae</i> <i>psa1</i> ; on yeast-form, not hyphal cell surface; alkaline induced; induced on adherence to polystyrene; Spider biofilm repressed	Ca22chr3A_C_albicans_SC5314:1786177 to 1785089
<i>TLO</i> α 9	Member of a family of telomere-proximal genes of unknown function; Hap43p-repressed gene	Ca22chr4A_C_albicans_SC5314:983 to 1660
CNH1	Na ⁺ /H ⁺ antiporter; required for wild-type growth, cell morphology, and virulence in a mouse model of systemic infection; not transcriptionally regulated by NaCl; fungal-specific (no human or murine homolog)	Ca22chr4A_C_albicans_SC5314:5033 to 7435
PHR3	Putative beta-1,3-glucanosyltransferase with similarity to the <i>A. fumigatus</i> GEL family; fungal-specific (no human or murine homolog); possibly an essential gene, disruptants not obtained by UAU1 method	Ca22chr4A_C_albicans_SC5314:12793 to 14309

Table 2. Cont.

Gene Name	Description ¹	Genomic Location
<i>TLOα10</i>	Member of a family of telomere-proximal genes of unknown function	Ca22chr4A_C_albicans_SC5314:1597812 to 1597159
VID21	Subunit of the NuA4 histone acetyltransferase complex; soluble protein in hyphae; Spider biofilm repressed	Ca22chr4A_C_albicans_SC5314:1596377 to 1594317
<i>TLOγ11</i>	Member of a family of telomere-proximal genes of unknown function; may be spliced in vivo	Ca22chr5A_C_albicans_SC5314:1918 to 2427
CDC11	Septin; cell and hyphal morphology, agar-invasive growth, full virulence and kidney tissue invasion in mouse, but not kidney colonization, immunogenicity; hyphal and cell-cycle-regulated phosphorylation; rat catheter biofilm repressed	Ca22chr5A_C_albicans_SC5314:8946 to 10154
THS1	Putative threonyl-tRNA synthetase; transcript regulated by Mig1 and Tup1; repressed upon phagocytosis by murine macrophages; stationary phase enriched protein; Spider biofilm repressed	Ca22chr5A_C_albicans_SC5314:14674 to 12554
<i>TLOα12</i>	Putative transcription factor; positive regulator of gene expression; Efg1-repressed; member of a family of telomere-proximal genes; transcript upregulated in RHE model of oral candidiasis	Ca22chr5A_C_albicans_SC5314:1182868 to 1182110
HTS1	Putative tRNA-His synthetase; downregulated upon phagocytosis by murine macrophage; stationary phase enriched protein; Spider biofilm repressed	Ca22chr5A_C_albicans_SC5314:1180902 to 1179397
DES1	Putative delta-4 sphingolipid desaturase; planktonic growth-induced gene	Ca22chr5A_C_albicans_SC5314:1178288 to 1179400
EST1	Telomerase subunit; allosteric activator of catalytic activity, but not required for catalytic activity; has TPR domain	Ca22chr5A_C_albicans_SC5314:1176356 to 1178194
NRG2	Transcription factor; transposon mutation affects filamentous growth	Ca22chr6A_C_albicans_SC5314:5 to 346
<i>TLOγ13</i>	Member of a family of telomere-proximal genes of unknown function; may be spliced in vivo; overlaps orf19.6337.1, which is a region annotated as blocked reading frame	Ca22chr6A_C_albicans_SC5314:5545 to 6069
PGA25	Putative GPI-anchored adhesin-like protein; fluconazole-downregulated; induced in oralpharyngeal candidiasis; Spider biofilm induced	Ca22chr6A_C_albicans_SC5314:9894 to 7276
HET1	Putative sphingolipid transfer protein; involved in localization of glucosylceramide which is important for virulence; Spider biofilm repressed	Ca22chr6A_C_albicans_SC5314:12543 to 11950
NAG4	Putative transporter; fungal-specific; similar to Nag3p and to <i>S. cerevisiae</i> Ypr156Cp and Ygr138Cp; required for wild-type mouse virulence and wild-type cycloheximide resistance; gene cluster encodes enzymes of GlcNAc catabolism	Ca22chr6A_C_albicans_SC5314:1026875 to 1025130
NAG3	Putative MFS transporter; similar to Nag4; required for wild-type mouse virulence and cycloheximide resistance; in gene cluster that includes genes encoding enzymes of GlcNAc catabolism; Spider biofilm repressed	Ca22chr6A_C_albicans_SC5314:1024922 to 1023237

Table 2. Cont.

Gene Name	Description ¹	Genomic Location
DAC1	N-acetylglucosamine-6-phosphate (GlcNAcP) deacetylase; N-acetylglucosamine utilization; required for wild-type hyphal growth and virulence in mouse systemic infection; gene and protein are GlcNAc-induced; Spider biofilm induced	Ca22chr6A_C_albicans_SC5314:1021987 to 1023228
NAG1	Glucosamine-6-phosphate deaminase; required for normal hyphal growth and mouse virulence; converts glucosamine 6-P to fructose 6-P; reversible reaction in vitro; gene and protein is GlcNAc-induced; Spider biofilm induced	Ca22chr6A_C_albicans_SC5314:1021746 to 1021000
HXK1	N-acetylglucosamine (GlcNAc) kinase; involved in GlcNAc utilization; required for wild-type hyphal growth and mouse virulence; GlcNAc-induced transcript; induced by alpha pheromone in SpiderM medium	Ca22chr6A_C_albicans_SC5314:1019276 to 1020757
RVS161	Protein required for endocytosis; contains a BAR domain, which is found in proteins involved in membrane curvature; null mutant exhibits defects in hyphal growth, virulence, cell wall integrity, and actin patch localization	Ca22chr7A_C_albicans_SC5314:2040 to 1246
RAD3	Ortholog of <i>S. cerevisiae</i> Rad3; 5' to 3' DNA helicase, nucleotide excision repair and transcription, subunit of RNA polIII initiation factor TFIIH and Nucleotide Excision Repair Factor 3 (NEF3)	Ca22chr7A_C_albicans_SC5314:5761 to 3464
SAC7	Putative GTPase activating protein (GAP) for Rho1; repressed upon adherence to polystyrene; macrophage/pseudohyphal-repressed; transcript is upregulated in RHE model of oral candidiasis and in clinical oral candidiasis	Ca22chr7A_C_albicans_SC5314:9430 to 7583
CSA1	Surface antigen on elongating hyphae and buds; strain variation in repeat number; ciclopirox, filament induced, alkaline induced by Rim101; Efg1-, Cph1, Hap43-regulated; required for WT RPMI biofilm formation; Bcr1-induced in a/a biofilms	Ca22chr7A_C_albicans_SC5314:13080 to 10024
FRP2	Putative ferric reductase; alkaline induced by Rim101; fluconazole-downregulated; upregulated in the presence of human neutrophils; possibly adherence-induced; regulated by Sef1, Sfu1, and Hap43	Ca22chr7A_C_albicans_SC5314:14047 to 15825
<i>TLOγ16</i>	Member of a family of telomere-proximal genes of unknown function; may be spliced in vivo	Ca22chr7A_C_albicans_SC5314:943352 to 942825
SYS1	Putative Golgi integral membrane protein; transcript regulated by Mig1	Ca22chr7A_C_albicans_SC5314:941499 to 940993
SPT6	Putative transcription elongation factor; transposon mutation affects filamentous growth; transcript induced in an RHE model of oral candidiasis and in clinical isolates from oral candidiasis	Ca22chr7A_C_albicans_SC5314:934878 to 939083

¹ As assigned in the Candida Genome Database [49].

Approximately one quarter of *C. albicans* subtelomeric genes (13 of 55) belong to the telomere-associated (*TLO*) gene family (Figure 2). *TLO* genes underwent a recent lineage-specific expansion from a single gene in most *Candida* species to 14 paralogs in *C. albicans* and are the only gene family with a significant presence within the subtelomeres [61–63]. Each *TLO* contains a MED2 domain, indicating their function as homologs of the Med2 subunit in Mediator, the major eukaryotic transcriptional regulatory complex [63]. *TLO*s participate in regulating a variety of virulence traits, including growth, resistance to stressors, and biofilm formation [64,65]. Strikingly, individual *TLO* genes regulate distinct virulence properties despite sharing >98% nucleotide identity between individual paralogs.

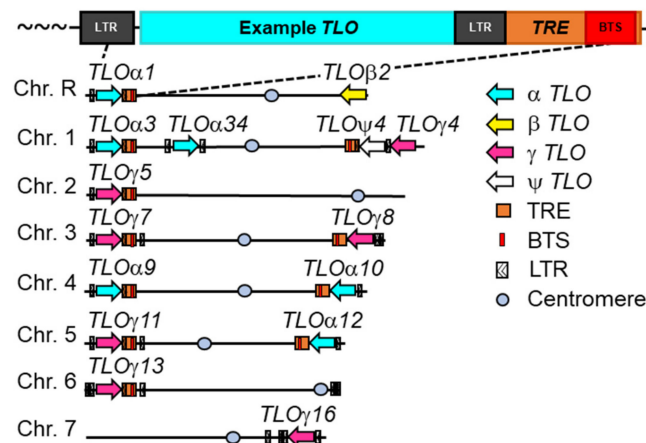


Figure 2. Organization of subtelomeric repetitive sequences in *C. albicans*. The telomere-associated (*TLO*) genes and other subtelomeric repetitive sequences are marked on each *C. albicans* chromosome arm for the genome reference strain SC5314. The *TLO* open reading frame (ORF; cyan) is commonly flanked by two long terminal repeats (LTR) elements, indicated in black. The *TLO* recombination element (TRE; orange) overlaps with the *TLO* 3' untranslated region (UTR) and extends towards the centromere to encompass the Bermuda Triangle Sequence (BTS; red). All repetitive sequences are oriented similarly on all chromosome arms. Both the TRE and BTS contribute to subtelomeric recombination. Telomeric repeats are denoted by “~~~”. Figure adapted from [66].

3.2. Repetitive DNA Elements in the Subtelomere

Repetitive elements commonly cluster in subtelomeric regions where they can buffer gene-rich chromosomal interiors from the detrimental effects of telomere length variation [67]. In particular, retrotransposons are often enriched within subtelomeres as new insertions are expected to incur less of a fitness cost in these gene-poor regions and can provide selective advantages by rescuing telomerase-deficient cells [68,69]. Over 350 unique retrotransposon insertions by 34 distinct families of transposable elements have been identified in the *C. albicans* SC5314 genome [70], which can be distinguished by the unique sequences of flanking long terminal repeats (LTRs). Of these 34 families, seven LTR families are found within the *C. albicans* subtelomeres, including an intact copy of Zorro2, the only complete non-LTR retrotransposon in the subtelomeres (Table 3) [70]. Thirteen individual LTRs are annotated in the subtelomeres, ~2.5x higher than the genome average. Importantly, LTR sequences incorporated as part of functional genes (e.g., *TLO*s) are not included among annotated repetitive sequences, suggesting these frequencies are likely an underestimate. Most evidence for transposon integration events in the subtelomeres comes from abandoned LTR repeats that mark previous insertions which subsequently reactivated the intervening transposase to excise itself and reintegrate elsewhere in the genome. In addition to inducing genotypic variation via continual transposon movement, highly abundant non-LTR retrotransposons in *C. albicans* can generate genetic diversity through recombination between dispersed LTR sequences [71].

Table 3. *C. albicans* subtelomeric LTRs and retrotransposons.

ORF Name	Description ¹	Genomic Location
gamma-1c	Solo copy of the long terminal repeat (LTR) associated with the transposon Tca2; about 280 bp long, 5-10 copies per genome	Ca22chr1A_C_albicans_SC5314:3187028 to 3187306
gamma-1b	Solo copy of the long terminal repeat (LTR) associated with the transposon Tca2; about 280 bp long, 5-10 copies per genome	Ca22chr1A_C_albicans_SC5314:3187383 to 3187662
Zorro2-1	Non-LTR retrotransposon, encodes a potential DNA-binding zinc-finger protein and a polyprotein similar to pol with conserved endonuclease and reverse transcriptase domains; member of L1 clade of transposons	Ca22chr1A_C_albicans_SC5314:3185887 to 3181663
san-1a	Solo copy of the long terminal repeat (LTR) associated with the transposon Tca4; about 381 bp long, 1-4 copies per genome	Ca22chr1A_C_albicans_SC5314:3178575 to 3178194
rho-2a	Long terminal repeat (LTR); about 275 bp long, 13 copies per genome	Ca22chr2A_C_albicans_SC5314:4611 to 4885
rho-3a	Long terminal repeat (LTR); about 275 bp long, 13 copies per genome	Ca22chr3A_C_albicans_SC5314:14098 to 14372
tui-3a	Long terminal repeat (LTR); about 199 bp long, 19 copies per genome	Ca22chr3A_C_albicans_SC5314:14751 to 14553
rho-3b	Long terminal repeat (LTR); about 275 bp long, 13 copies per genome	Ca22chr3A_C_albicans_SC5314:1787772 to 1787498
rho-5a	Long terminal repeat (LTR); about 275 bp long, 13 copies per genome	Ca22chr5A_C_albicans_SC5314:2260 to 2534
rho-6a	Long terminal repeat (LTR); about 275 bp long, 13 copies per genome	Ca22chr6A_C_albicans_SC5314:5902 to 6176
chi-6a	Long terminal repeat (LTR) associated with the transposon Tca10; about 192 bp long, 11 copies per genome	Ca22chr6A_C_albicans_SC5314:1032389 to 1032200
kappa-6a	Long terminal repeat (LTR) associated with the transposon Tca6; about 280 bp long, 10-15 copies per genome	Ca22chr6A_C_albicans_SC5314:1028505 to 1028227
weka-6a	Long terminal repeat (LTR); about 167 bp long, 9 copies per genome	Ca22chr6A_C_albicans_SC5314:1027905 to 1028066
san-7a	Solo copy of the long terminal repeat (LTR) associated with the transposon Tca4; about 381 bp long, 1-4 copies per genome	Ca22chr7A_C_albicans_SC5314:942855 to 942475

¹ As assigned in the Candida Genome Database [49].

Additional sequence elements, detected through molecular investigations, exist in *C. albicans* subtelomeres. These elements, the Bermuda Triangle sequence (BTS) and the *TLO* recombination element (TRE), lie immediately centromeric to *TLO* genes (Figure 2). The BTS is defined by a 50 bp sequences which share ~88% sequence similarity across the 11 BTS-containing subtelomeres [72]. Each BTS is encompassed within the longer *TLO* recombination element (TRE). The TRE stretches ~300 bp, beginning immediately following the *TLO* coding sequence and extending towards the centromere [59]. While no clear phylogenetic hierarchy is observed among the BTS sequences, TREs can be organized into three groups based on near complete sequence identity within any single cluster. Both the BTS and TRE overlap with a member of the *tui*-class family of LTR sequences which are annotated in some subtelomeres but missing in others (Figure 2). Given the high sequence homology across the BTS and TRE, it is likely this LTR element is present next to all *TLO* genes with the exception of ChrRR and Chr7R, which both lack the BTS and TRE [56].

3.3. Subtelomeres Are Prone to Mutation

C. albicans subtelomeres experience high rates of recombination, including frequent loss of heterozygosity (LOH). LOH events result from non-reciprocal crossing over or chromosome loss and reduplication of the lost region from the remaining chromosome homolog. Rates of LOH increase along all chromosomes towards the telomeres of *C. albicans* isolates, indicating a general relationship between the distance from the centromere and allelic homozygosity mediated by recombination [73] (Figure 3). Yet, subtelomeric LOH rates increase an order of magnitude compared to centromere-proximal regions interior to the TRE and loss of the TRE in subtelomeres results in a significant decrease in LOH [56]. Integration of the TRE into an exogenous locus increased LOH rates of an adjacent selectable marker [56]. Thus, the TRE sequence is crucial to the elevated recombination rates in *C. albicans* subtelomeres where it suppresses mitotic recombination additively with Sir2 silencing. However, some stressors, including fluconazole, promote such a strong genome instability phenotype that the protective effects of Sir2 can be masked in their presence [56].

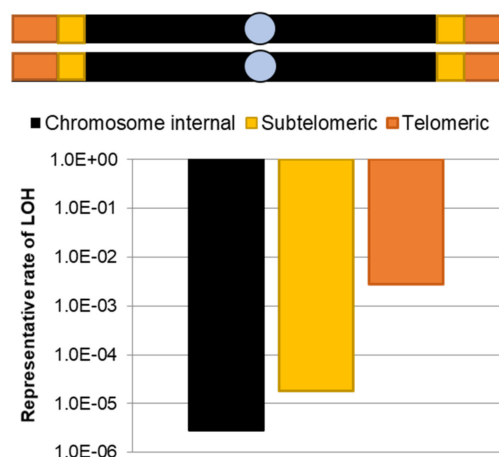


Figure 3. Rates of loss of heterozygosity (LOH) increase towards chromosome ends. An average LOH rate for each genomic region is depicted for chromosome internal sequences (black), the subtelomeres (yellow), and telomeres (orange) based on published (chromosome internal and subtelomeric) [28,56] and unpublished results (telomeric). All data is derived from LOH assays in which a *URA3* marker was inserted within different genomic regions and its location determined by either sequencing or contour clamped homogenous electric field (CHEF) gel analysis and Southern blotting. Centromeres are indicated by grey circles.

Phosphorylation of serine 129 on histone H2A, commonly known as γ H2A, is a common marker for heterochromatin in *S. cerevisiae* [74]. Consistent with localizing to heterochromatic regions, γ H2A is enriched at subtelomeres, but is also found at origins of replication and convergent genes where it

labels double-strand DNA breaks. More specifically, γ H2A abundance is statistically enriched on 13 of the 16 *C. albicans* subtelomeres (all except ChrRR, 1R, and 7R). Its conspicuous absence in select subtelomeres is thought to be due to incomplete assembly in the current genome assembly and not the absence of γ H2A [75]. These γ -sites mark putative recombination-prone genomic loci that are intrinsically more fragile and susceptible to DNA damage [75].

Subtelomeric recombination directly affects their genic repertoire. Continuous passaging of SC5314-derived strains for 30 months led to the gain and loss of specific *TLO* genes and all telomere-proximal sequences through non-reciprocal recombination events among different chromosome arms [72]. These non-reciprocal exchanges occurred once every 5000 generations on average, with some chromosome arms being favored as either DNA donors or recipients. Bias in amplification and loss of specific subtelomeric sequences due to non-reciprocal DNA exchange suggests that selection favored expansion or loss of certain genomic regions during passaging. Approximately half of the detected recombination events occurred within the BTS element with two additional events initiating in the TRE. Two recombination events took place within *TLO* genes and produced novel *TLO* sequences, highlighting the potential for significant sequence diversity to arise via subtelomeric recombination that can contribute to phenotypic plasticity [72].

Genes within the subtelomeres are subject to not only increased rates of recombination, but also increased formation of indels [73]. *TLO*s can be separated into three clades (α , β , and γ) based on sequence diversity that primarily clusters in the 3' end of the coding sequence. The *TLO* β -clade is unique in containing two major indel events in this 3' region that distinguish it from the other *TLO* clades. The 3' end of *TLO* γ -clade genes uniformly contains a *rho* LTR sequence that likely disrupted an ancestral *TLO* α -clade homolog which then expanded across the subtelomeres through non-reciprocal exchange of this newly-formed *TLO* γ -clade gene [62]. Further demonstrating the evolutionary potential of continual retrotransposon activity, a subsequent *psi* LTR-containing retrotransposon insertion disrupted *TLO* γ 4 in the SC5314 background, producing a truncated *TLO* γ 4 gene, which is actively transcribed, and a *TLO* ψ 4 pseudogene with no detectable transcript [62,72].

4. The *C. albicans* Major Repeat Sequence

Seven of the eight *C. albicans* chromosomes, the exception being Chr3, contain large tracts of repetitive DNA known as the MRS. These regions are comprised of large, variable arrays of the 2 kb repetitive sequence (RPS) unit that are often flanked by the non-repetitive RB2 (6 kb) and HOK (8 kb) elements [21,76]. RPS-1, the main unit of the MRS, is itself comprised of smaller repeat segments of about 80–170 bp, which are then assembled into two families of ordered segments, REP1 and REP3, which both contain the 29 bp COM29 common sequence (5'-CAAAAAGGCCGTTTTGGCCATAGTTAAG-3') [77]. Altogether, *C. albicans* possesses nine complete MRS loci, 14 RB2 elements, and two HOK elements (Figure 4). MRS loci contain rare 8-bp SfiI restriction sites that have been used extensively to separate chromosome arms for mapping genetic loci prior to construction of the genome reference sequence and for detection of chromosomal rearrangements by contour clamped homogenous electric field (CHEF) electrophoresis [78].

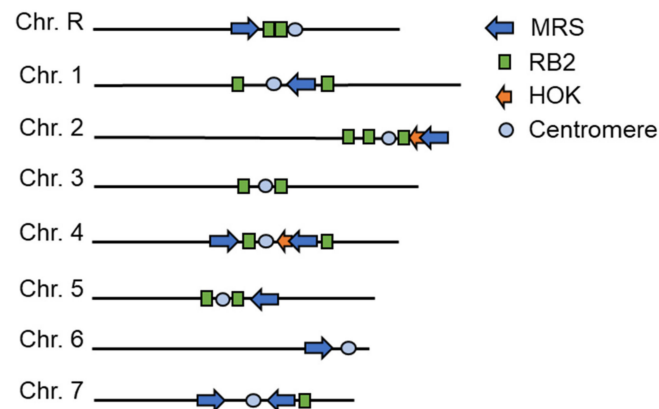


Figure 4. MRS elements in the *C. albicans* genome. The nine complete MRS loci (blue), 14 RB2 elements (green) and two HOK elements (orange) have been placed on their relative chromosome positions in *C. albicans* strain SC5314.

Impacts of the MRS on Chromosome Loss and Recombination

Despite the highly nested repeat structure of the MRS sites, recombination frequencies in the MRS are equivalent to the genome average rates of recombination. Thus, the MRS is unlikely to function as a recombination hotspot, but instead often contributes to translocations between heterologous chromosomes [79]. Insertion of a *URA3* selectable marker into the RB2 element of various MRS sites allowed translocation events to be tracked that involved multiple different chromosomes as well as truncations of Chr7 in a few instances [80]. Broad karyotypic surveys across clinical isolates have repeatedly identified chromosomal translocations involving the MRS of different chromosomes, suggesting that these chromosomal rearrangements are rare but stably maintained over longer evolutionary time frames by *C. albicans* isolates [81–83]. In particular, the WO-1 strain, in which the white-opaque cell state switch was first described, contains three chromosomal translocations that are coincident with the MRS of different chromosomes [84,85].

Intrachromosomal recombination also occurs within the MRS. The *XhoI* restriction enzyme cuts chromosomal DNA on either side of the MRS that can be used to define MRS size by Southern blotting with a RPS-specific probe [79]. Changes in MRS size can also be observed by CHEF electrophoresis of chromosome homologs that resolve from each other due to repeat expansion or contraction. Detection of these events is more readily observable for smaller chromosomes (e.g., Chr7) that have greater length resolution [83]. Loss of the MRS through targeted deletion suggests that these repeats are not required for gross viability in *C. albicans* [86]. Yet, chromosome loss due to nondisjunction was directly related to MRS size based on selection for loss of Chr5 by growth in sorbose [86]. Chr5 homologs encoding the smaller MRS were preferentially retained compared to homologs with longer MRS regions [79,86].

5. The rDNA locus

C. albicans rDNA repeats are encoded as an array on the right arm of ChrR at the *RDN1* locus. The rDNA locus encompasses the 18s, 5.8s, 25s, and 5s rRNAs that are organized as tandem repeating units. This locus ranges in size from 11.6 kb to 12.5 kb and shifts between 21 and 176 copies of repeating rDNA units across strains and growth conditions [87]. Consequently, the full *RDN1* locus can vary between 244 and 2200 kb or approximately 10% of the total size of the *C. albicans* genome. The presence of the rDNA locus and such massive shifts in chromosome size underlie the distinct name of the encoding chromosome, ChrR [88]. Naming of this chromosome breaks from the numbering system based on length used for all other nuclear chromosomes, as changes in rDNA copy number can greatly alter its size.

Recombination between rDNA Repeats

Unequal intrachromosomal recombination events within the rDNA repeats are frequent and produce the large-scale shifts in *RDN1* size. Changes in rDNA repeats contribute to approximately 92% of *C. albicans* chromosomal variation and instability in colony morphology mutants [87]. The frequency of these recombination events lacks precise quantification but can be observed within five cell divisions by next-generation sequencing (unpublished data). Extrachromosomal rDNA circles and linear rDNA plasmids can also be released during recombination [89,90]. The monopolin complex, shown to be important for regulating rDNA and telomeric repeats in *S. cerevisiae* [91], also contributes to *C. albicans* rDNA maintenance. Strains lacking monopolin encoded shorter and more variably-sized rDNA, implicating monopolin in specifically maintaining rDNA length [92].

6. Additional Repetitive Sequences Shape *Candida albicans* Genomic Stability

Although telomeres, subtelomeres, the MRS, and the rDNA locus are all clearly defined genomic regions in *C. albicans* that play important roles in genome stability, multiple additional repetitive sequences throughout the genome promote recombination and shape genome evolution.

6.1. The Agglutinin-Like Sequence (ALS) Family of Adhesins

Critical to *C. albicans*' success within the human host is its ability to adhere to a wide range of substrates. Colonization and adherence are mediated, in part, by the ALS gene family of adhesins. ALS genes family members are encoded by eight distinct loci that each produce a glycoprotein capable of forming amyloid-like aggregates [93,94]. Within each ALS locus, a conserved Ser/Thr-rich domain provides a platform for intergenic recombination between ALS homologs and production of new genetic variants [93–96]. Indeed, ALS51 is a product of recombination between the ALS1 and the ALS5 loci that is present in multiple clinical isolates of *C. albicans*. Vast allelic differences can also arise through intragenic recombination. Hypermutability of the ALS7 ORF has produced 60 ALS7 alleles identified across a collection of 66 strains [97]. Variant alleles arise via frequent recombination within tandem repeats and conserved 3' repetitive domains. Mutation and LOH rates within this gene family are also remarkably high, comparable to that within the MRS, subtelomeric regions, and the TLO gene family [24]. Thus, recombination within and among these cell surface antigens allows for rapid gene turnover and adaptation via cellular adherence to environmental substrates that can promote colonization and persistence across body sites [98].

6.2. Genome Evolution through Centromere-Associated Repeats

Centromeres in *C. albicans* are unique, gene devoid stretches of DNA typically flanked by inverted repeats [60,99]. Recombination between these flanking repeats can lead to the formation of isochromosomes, which are chromosomes comprised of two identical chromosome arms joined by a functional centromere. Formation of an isochromosome of Chr5L (i5L) commonly follows *C. albicans* exposure to azole-class antifungal drugs [29]. Amplification of the left arm of Chr5 confers azole resistance through the acquisition of extra copies of *ERG11* and *TAC1*, encoding the canonical azole drug target and the transcriptional activator of efflux pumps, respectively [100]. Recombination between the centromeric repeats flanking either side of the Chr5 centromere is responsible for producing i5L [29]. Similarly, clinical isolate P78042, which was isolated with a Chr4 trisomy [25], produced an isochromosome of the right arm of Chr4 during passage in the presence of FLC for 100 generations [101]. These examples highlight the importance of recombination and isochromosome formation through centromere-associated repeats for adaptation of *C. albicans* to antifungal drugs [101].

6.3. Recombination Facilitated by Cryptic Long Repeats

A recent scan of the *C. albicans* SC5314 genome identified 1974 long repeat sequences, excluding the rDNA, the MRS, subtelomeric repeats, and previously characterized complex tandem repeats [101].

Called repeats were required to have sequence matches of 20 bp or longer, which occur more than once in the genome and accounted for 2.87% of the haploid reference genome altogether. Repeat location did not correlate with GC content or ORF density and many repeats encompassed complete and actively transcribed genes [101]. Recombination at these long repeats produced strains harboring copy number variations, LOH, and chromosomal inversions. Consequently, repetitive sequences scattered throughout the genome must be included as additional arbiters of genome instability [101].

7. Conclusions

Thus, repetitive sequences in the *C. albicans* genome promote genome instability through various mechanisms that can range from minor changes in sequence length to massive chromosomal rearrangements. Previous work has focused on defining the characteristics of four major repetitive regions (telomeres, subtelomeres, the MRS, and the rDNA locus), but their contributions to genetic and phenotypic diversity still remain generally unresolved. Furthermore, the importance of less well-understood repetitive sequences to genome evolution and karyotypic innovation are just beginning to emerge. What is clear is that repetitive elements within the *C. albicans* genome produce novel genotypes through changes in DNA sequences or karyotypes that often prove to be beneficial. Some of these changes, such as isochromosome formation during antifungal drug exposure and changes to gene family copy number, are more intuitive, whereas the function of others involving changes in MRS repeat length and rDNA copy number are less easy to grasp. It could be that repeat length modulates the probability of recombination occurring or that the presence of repeats is itself sufficient to promote genome instability. Association of additional repetitive sequences at major regions within the genome (e.g., telomeres, the MRS) suggest that these sites are likely under relaxed selection for new insertions that increase the likelihood of additional recombination events. In contrast, repeats scattered throughout the genome beyond these defined regions may function as back-up sites for genome instability. These cryptic repetitive sequences throughout the *C. albicans* genome may not only promote recombination but reduce the associated fitness costs of novel genotypes by increasing the resolution of affected sequences to only the locus (loci) that is under selection. Regardless, these repetitive sequences collectively construct a dynamic environment of common and infrequent sites of genomic instability to generate genotypic and phenotypic variants that allow *C. albicans* adaptation to its diverse niches across the human host and ultimately promote its success as a commensal and pathogen.

Author Contributions: Writing—original draft preparation, M.J.D.; writing—review and editing, M.Z.A. and M.J.D.; visualization, M.J.D. and M.Z.A.; supervision, M.Z.A.

Funding: This research received no external funding

Acknowledgments: We would like to thank members of the Anderson lab for helpful discussion in organizing and preparing this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Khandelwal, S. Chromosome evolution in the genus *Ophioglossum* L. *Bot. J. Linn. Soc.* **1990**, *102*, 205–217. [[CrossRef](#)]
2. Carvalho, C.M.B.; Lupski, J.R. Mechanisms underlying structural variant formation in genomic disorders. *Nat. Rev. Genet.* **2016**, *17*, 224–238. [[CrossRef](#)] [[PubMed](#)]
3. Mézard, C.; Pompon, D.; Nicolas, A. Recombination between similar but not identical DNA sequences during yeast transformation occurs within short stretches of identity. *Cell* **1992**, *70*, 659–670. [[CrossRef](#)]
4. Oikawa, S.; Tada-Oikawa, S.; Kawanishi, S. Site-specific DNA damage at the GGG sequence by UVA involves acceleration of telomere shortening. *Biochemistry* **2001**, *40*, 4763–4768. [[CrossRef](#)]
5. von Zglinicki, T. Oxidative stress shortens telomeres. *Trends Biochem. Sci.* **2002**, *27*, 339–344. [[CrossRef](#)]
6. Madireddy, A.; Gerhardt, J. *Replication Through Repetitive DNA Elements and Their Role in Human Diseases*; Springer: Singapore, 2017; pp. 549–581. ISBN 9789811069543.

7. Riggi, N.; Knoechel, B.; Gillespie, S.M.; Rheinbay, E.; Boulay, G.; Suvà, M.L.; Rossetti, N.E.; Boonseng, W.E.; Oksuz, O.; Cook, E.B.; et al. EWS-FLI1 Utilizes Divergent Chromatin Remodeling Mechanisms to Directly Activate or Repress Enhancer Elements in Ewing Sarcoma. *Cancer Cell* **2014**, *26*, 668–681. [[CrossRef](#)]
8. García-Nieto, P.E.; Schwartz, E.K.; King, D.A.; Paulsen, J.; Collas, P.; Herrera, R.E.; Morrison, A.J. Carcinogen susceptibility is regulated by genome architecture and predicts cancer mutagenesis. *EMBO J.* **2017**, *36*, 2829–2843. [[CrossRef](#)]
9. Smith, K.S.; Liu, L.L.; Ganesan, S.; Michor, F.; De, S. Nuclear topology modulates the mutational landscapes of cancer genomes. *Nat. Struct. Mol. Biol.* **2017**, *24*, 1000–1006. [[CrossRef](#)]
10. Deacon, J.W. *Fungal Biology*, 4th ed.; Wiley-Blackwell: Hoboken, NJ, USA, 29 July 2005; ISBN 978-1405130660.
11. Dutta, S.K. Repeated DNA sequences in fungi. *Nucleic Acids Res.* **1974**, *1*, 1411–1420. [[CrossRef](#)]
12. Toth, G. Microsatellites in Different Eukaryotic Genomes: Survey and Analysis. *Genome Res.* **2000**, *10*, 967–981. [[CrossRef](#)]
13. Smith, K.M.; Galazka, J.M.; Phatale, P.A.; Connolly, L.R.; Freitag, M. Centromeres of filamentous fungi. *Chromosom. Res.* **2012**, *20*, 635–656. [[CrossRef](#)] [[PubMed](#)]
14. Kupiec, M. Biology of telomeres: lessons from budding yeast. *FEMS Microbiol. Rev.* **2014**, *38*, 144–171. [[CrossRef](#)] [[PubMed](#)]
15. Daboussi, M.J. Fungal transposable elements and genome evolution. *Genetica* **1997**, *100*, 253–260. [[CrossRef](#)] [[PubMed](#)]
16. Galagan, J.E.; Henn, M.R.; Ma, L.J.; Cuomo, C.A.; Birren, B. Genomics of the fungal kingdom: Insights into eukaryotic biology. *Genome Res.* **2005**, *15*, 1620–1631. [[CrossRef](#)]
17. Karaoglu, H.; Lee, C.M.Y.; Meyer, W. Survey of simple sequence repeats in completed fungal genomes. *Mol. Biol. Evol.* **2005**, *22*, 639–649. [[CrossRef](#)]
18. Lim, S.; Notley-McRobb, L.; Lim, M.; Carter, D.A. A comparison of the nature and abundance of microsatellites in 14 fungal genomes. *Fungal Genet. Biol.* **2004**, *41*, 1025–1036. [[CrossRef](#)]
19. Bartelli, T.F.; Bruno, D.D.C.F.; Briones, M.R.S. Whole-Genome Sequences and Annotation of the Opportunistic Pathogen *Candida albicans* Strain SC5314 Grown under Two Different Environmental Conditions. *Genome Announc.* **2018**, *6*, 4–5. [[CrossRef](#)]
20. Muzzey, D.; Schwartz, K.; Weissman, J.S.; Sherlock, G. Assembly of a phased diploid *Candida albicans* genome facilitates allele-specific measurements and provides a simple model for repeat and indel structure. *Genome Biol.* **2013**, *14*, R97. [[CrossRef](#)]
21. Van het Hoog, M.; Rast, T.J.; Martchenko, M.; Grindle, S.; Dignard, D.; Hogues, H.; Cuomo, C.; Berriman, M.; Scherer, S.; Magee, B.B.; et al. Assembly of the *Candida albicans* genome into sixteen supercontigs aligned on the eight chromosomes. *Genome Biol.* **2007**, *8*. [[CrossRef](#)]
22. Butler, G.; Jackson, A.P.; Gamble, J.A.; Yeomans, T.; Moran, G.P.; Saunders, D.; Harris, D.; Aslett, M.; Barrell, J.F.; Citiulo, F.; et al. Comparative genomics of the fungal pathogens *Candida dubliniensis* and *Candida albicans*. *Genome Res.* **2009**, *19*, 2231–2244.
23. Jones, T.; Federspiel, N.A.; Chibana, H.; Dungan, J.; Kalman, S.; Magee, B.B.; Newport, G.; Thorstenson, Y.R.; Agabian, N.; Magee, P.T.; et al. The diploid genome sequence of *Candida albicans*. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7329–7334. [[CrossRef](#)] [[PubMed](#)]
24. Ene, I.V.; Farrer, R.A.; Hirakawa, M.P.; Agwamba, K.; Cuomo, C.A.; Bennett, R.J. Global analysis of mutations driving microevolution of a heterozygous diploid fungal pathogen. *Proc. Natl. Acad. Sci.* **2018**, *115*, E8688–E8697. [[CrossRef](#)] [[PubMed](#)]
25. Hirakawa, M.P.; Martinez, D.A.; Sakthikumar, S.; Anderson, M.Z.; Berlin, A.; Gujja, S.; Zeng, Q.; Zisson, E.; Wang, J.M.; Greenberg, J.M.; et al. Genetic and phenotypic intra-species variation in *Candida albicans*. *Genome Res.* **2015**, *25*, 413–425. [[CrossRef](#)] [[PubMed](#)]
26. Ropars, J.; Maufrais, C.; Diogo, D.; Marcet-Houben, M.; Perin, A.; Sertour, N.; Mosca, K.; Permal, E.; Laval, G.; Bouchier, C.; et al. Gene flow contributes to diversification of the major fungal pathogen *Candida albicans*. *Nat. Commun.* **2018**, *9*, 2253. [[CrossRef](#)] [[PubMed](#)]
27. Forche, A.; Magee, P.T.; Selmecki, A.; Berman, J.; May, G. Evolution in *Candida albicans* populations during a single passage through a mouse host. *Genetics* **2009**, *182*, 799–811. [[CrossRef](#)]
28. Forche, A.; Abbey, D.; Pisithkul, T.; Weinzierl, M.A.; Ringstrom, T.; Bruck, D.; Petersen, K.; Berman, J. Stress alters rates and types of loss of heterozygosity in *Candida albicans*. *MBio* **2011**, *2*, 1–9. [[CrossRef](#)]

29. Selmecki, A. Aneuploidy and Isochromosome Formation in Drug-Resistant *Candida albicans*. *Science* (80-). **2006**, *313*, 367–370. [[CrossRef](#)] [[PubMed](#)]
30. Forche, A.; Solis, N.V.; Swidergall, M.; Thomas, R.; Guyer, A.; Beach, A.; Cromie, G.A.; Le, G.T.; Lowell, E.; Pavelka, N.; et al. Selection of *Candida albicans* trisomy during oropharyngeal infection results in a commensal-like phenotype. *PLOS Genet.* **2019**, *15*, e1008137. [[CrossRef](#)]
31. Liang, S.-H.; Anderson, M.Z.; Hirakawa, M.P.; Wang, J.M.; Frazer, C.; Alaalm, L.M.; Thomson, G.J.; Ene, I.V.; Bennett, R.J. Hemizygoty Enables a Mutational Transition Governing Fungal Virulence and Commensalism. *Cell Host Microbe* **2019**, *25*, 418–431. [[CrossRef](#)] [[PubMed](#)]
32. Forche, A.; Cromie, G.; Gerstein, A.C.; Solis, N.V.; Pisithkul, T.; Srifa, W.; Jeffery, E.; Abbey, D.; Filler, S.G.; Dudley, A.M.; et al. Rapid Phenotypic and Genotypic Diversification After Exposure to the Oral Host Niche in *Candida albicans*. *Genetics* **2018**, *209*, genetics.301019.
33. Yu, E.Y. Telomeres and telomerase in *Candida albicans*. *Mycoses* **2012**, *55*, e48–e59. [[CrossRef](#)] [[PubMed](#)]
34. McEachern, M.J.; Hicks, J.B. Unusually large telomeric repeats in the yeast *Candida albicans*. *Mol. Cell. Biol.* **1993**, *13*, 551–560. [[CrossRef](#)] [[PubMed](#)]
35. Lépingle, A.; Casaregola, S.; Neuvéglise, C.; Bon, E.; Nguyen, H.-V.; Artiguenave, F.; Wincker, P.; Gaillardin, C. Genomic Exploration of the Hemiascomycetous Yeasts: 14. *Debaryomyces hansenii* var. *hansenii*. *FEBS Lett.* **2000**, *487*, 82–86. [[CrossRef](#)]
36. McEachern, M.J.; Blackburn, E.H. A conserved sequence motif within the exceptionally diverse telomeric sequences of budding yeasts. *Proc. Natl. Acad. Sci.* **1994**, *91*, 3453–3457. [[CrossRef](#)] [[PubMed](#)]
37. Sun, J.; Yang, Y.; Confer, L.A.; Wan, K.; Lei, M.; Sun, S.H.; Yu, E.Y.; Lue, N.F. Stn1-Ten1 is an Rpa2-Rpa3-like complex at telomeres. *Genes Dev.* **2009**, *23*, 2900–2914. [[CrossRef](#)] [[PubMed](#)]
38. Sadhu, C.; McEachern, M.J.; Rustchenko-bulgac, E.P.; Schmid, J.A.N.; Soll, D.R.; Hicks, J.B. Telomeric and Dispersed Repeat Sequences in *Candida* Yeasts and Their Use in Strain Identification. *J. Bacteriol.* **1991**, *173*, 842–850. [[CrossRef](#)]
39. Counter, C.M. The roles of telomeres and telomerase in cell life span. *Mutat. Res. Genet. Toxicol.* **1996**, *366*, 45–63. [[CrossRef](#)]
40. Nugent, C.I.; Lundblad, V. The telomerase reverse transcriptase: Components and regulation. *Genes Dev.* **1998**, *12*, 1073–1085. [[CrossRef](#)]
41. Zhou, J.; Ding, D.; Wang, M.; Cong, Y.S. Telomerase reverse transcriptase in the regulation of gene expression. *BMB Rep.* **2014**, *47*, 8–14. [[CrossRef](#)]
42. Hsu, M.; McEachern, M.J.; Dandjinou, A.T.; Tzfati, Y.; Orr, E.; Blackburn, E.H.; Lue, N.F. Telomerase core components protect *Candida* telomeres from aberrant overhang accumulation. *Proc. Natl. Acad. Sci.* **2007**, *104*, 11682–11687. [[CrossRef](#)]
43. Steinberg-Neifach, O.; Lue, N.F. Modulation of telomere terminal structure by telomerase components in *Candida albicans*. *Nucleic Acids Res.* **2006**, *34*, 2710–2722. [[CrossRef](#)] [[PubMed](#)]
44. Chico, L.; Ciudad, T.; Hsu, M.; Lue, N.F.; Larriba, G. The *Candida albicans* Ku70 modulates telomere length and structure by regulating both telomerase and recombination. *PLoS ONE* **2011**, *6*. [[CrossRef](#)] [[PubMed](#)]
45. Legrand, M.; Chan, C.L.; Jauert, P.A.; Kirkpatrick, D.T. Role of DNA Mismatch Repair and Double-Strand Break Repair in Genome Stability and Antifungal Drug Resistance in *Candida albicans*. *Eukaryot. Cell* **2007**, *6*, 2194–2205. [[CrossRef](#)] [[PubMed](#)]
46. Biswas, K.; Rieger, K.J.; Morschhäuser, J. Functional analysis of CaRAP1, encoding the repressor/activator protein 1 of *Candida albicans*. *Gene* **2003**, *307*, 151–158. [[CrossRef](#)]
47. Yu, E.Y.; Yen, W.-F.; Steinberg-Neifach, O.; Lue, N.F. Rap1 in *Candida albicans*: an Unusual Structural Organization and a Critical Function in Suppressing Telomere Recombination. *Mol. Cell. Biol.* **2010**, *30*, 1254–1268. [[CrossRef](#)]
48. Lundblad, V.; Blackburn, E.H. An alternative pathway for yeast telomere maintenance rescues est1- senescence. *Cell* **1993**, *73*, 347–360. [[CrossRef](#)]
49. Zhang, J.M.; Yadav, T.; Ouyang, J.; Lan, L.; Zou, L. Alternative Lengthening of Telomeres through Two Distinct Break-Induced Replication Pathways. *Cell Rep.* **2019**, *26*, 955–968. [[CrossRef](#)]
50. Tomaska, L. Extragenomic double-stranded DNA circles in yeast with linear mitochondrial genomes: potential involvement in telomere maintenance. *Nucleic Acids Res.* **2000**, *28*, 4479–4487. [[CrossRef](#)]

51. Ciudad, T.; Andaluz, E.; Steinberg-Neifach, O.; Lue, N.F.; Gow, N.A.R.; Calderone, R.A.; Larriba, G. Homologous recombination in *Candida albicans*: Role of CaRad52p in DNA repair, integration of linear DNA fragments and telomere length. *Mol. Microbiol.* **2004**, *53*, 1177–1194. [[CrossRef](#)]
52. Gerhold, J.M.; Aun, A.; Sedman, T.; Jøers, P.; Sedman, J. Strand invasion structures in the inverted repeat of *Candida albicans* mitochondrial DNA reveal a role for homologous recombination in replication. *Mol. Cell* **2010**, *39*, 851–861. [[CrossRef](#)]
53. Bijlani, S.; Thevandavakkam, M.A.; Tsai, H.-J.; Berman, J. Autonomously Replicating Linear Plasmids That Facilitate the Analysis of Replication Origin Function in *Candida albicans*. *mSphere* **2019**, *4*, 1–19. [[CrossRef](#)] [[PubMed](#)]
54. Pryde, F.E.; Gorham, H.C.; Louis, E.J. Chromosome ends: all the same under their caps. *Curr. Opin. Genet. Dev.* **1997**, *7*, 822–828. [[CrossRef](#)]
55. Froyd, C.A.; Kapoor, S.; Dietrich, F.; Rusche, L.N. The Deacetylase Sir2 from the Yeast *Clavispora lusitaniae* Lacks the Evolutionarily Conserved Capacity to Generate Subtelomeric Heterochromatin. *PLoS Genetics* **2013**, *9*, e1003935. [[CrossRef](#)] [[PubMed](#)]
56. Freire-Benítez, V.; Gourlay, S.; Berman, J.; Buscaino, A. Sir2 regulates stability of repetitive domains differentially in the human fungal pathogen *Candida albicans*. *Nucleic Acids Res.* **2016**, *44*, gkw594. [[CrossRef](#)]
57. Gottschling, D.E.; Aparicio, O.M.; Billington, B.L.; Zakian, V.A. Position effect at *S. cerevisiae* telomeres: Reversible repression of Pol II transcription. *Cell* **1990**, *63*, 751–762. [[CrossRef](#)]
58. Anderson, M.Z.; Gerstein, A.C.; Wigen, L.; Baller, J.A.; Berman, J. Silencing Is Noisy: Population and Cell Level Noise in Telomere-Adjacent Genes Is Dependent on Telomere Position and Sir2. *PLoS Genet.* **2014**, *10*. [[CrossRef](#)]
59. Freire-Benítez, V.; Price, R.J.; Tarrant, D.; Berman, J.; Buscaino, A. *Candida albicans* repetitive elements display epigenetic diversity and plasticity. *Sci. Rep.* **2016**, *6*. [[CrossRef](#)]
60. Skrzypek, M.S.; Binkley, J.; Binkley, G.; Miyasato, S.R.; Simison, M.; Sherlock, G. The *Candida* Genome Database (CGD): Incorporation of Assembly 22, systematic identifiers and visualization of high throughput sequencing data. *Nucleic Acids Res.* **2017**, *45*, D592–D596. [[CrossRef](#)]
61. Zhang, A.; Petrov, K.O.; Hyun, E.R.; Liu, Z.; Gerber, S.A.; Myers, L.C. The Tlo proteins are stoichiometric components of *Candida albicans* Mediator anchored via the Med3 subunit. *Eukaryot. Cell* **2012**, *11*, 874–884. [[CrossRef](#)]
62. Anderson, M.Z.; Baller, J.A.; Dulmage, K.; Wigen, L.; Berman, J. The three clades of the telomere-associated Tlo gene family of *Candida albicans* have different splicing, localization, and expression features. *Eukaryot. Cell* **2012**, *11*, 1268–1275. [[CrossRef](#)]
63. Sullivan, D.J.; Berman, J.; Myers, L.C.; Moran, G.P. Telomeric ORFS in *Candida albicans*: Does Mediator Tail Wag the Yeast? *PLOS Pathog.* **2015**, *11*, e1004614. [[CrossRef](#)] [[PubMed](#)]
64. Dunn, M.J.; Kinney, G.M.; Washington, P.M.; Berman, J.; Anderson, Z. Functional diversification accompanies gene family expansion of MED2 homologs in *Candida albicans*. *PLOS Genet.* **2018**, 1–29. [[CrossRef](#)] [[PubMed](#)]
65. Flanagan, P.R.; Fletcher, J.; Boyle, H.; Sulea, R.; Moran, G.P.; Sullivan, D.J. Expansion of the TLO gene family enhances the virulence of *Candida* species. *PLoS ONE* **2018**, *13*, e0200852. [[CrossRef](#)] [[PubMed](#)]
66. Moran, G.P.; Anderson, M.Z.; Myers, L.C.; Sullivan, D.J. Role of Mediator in virulence and antifungal drug resistance in pathogenic fungi. *Curr. Genet.* **2019**, *65*, 621–630. [[CrossRef](#)] [[PubMed](#)]
67. Tashiro, S.; Nishihara, Y.; Kugou, K.; Ohta, K.; Kanoh, J. Subtelomeres constitute a safeguard for gene expression and chromosome homeostasis. *Nucleic Acids Res.* **2017**, *45*, 10333–10349. [[CrossRef](#)] [[PubMed](#)]
68. Kim, J.M.; Vanguri, S.; Boeke, J.D.; Gabriel, A.; Voytas, D.F. Transposable elements and genome organization: A comprehensive survey of retrotransposons revealed by the complete *Saccharomyces cerevisiae* genome sequence. *Genome Res.* **1998**, *8*, 464–478. [[CrossRef](#)]
69. Scholes, D.T.; Kenny, A.E.; Gamache, E.R.; Mou, Z.; Curcio, M.J. Activation of a LTR-retrotransposon by telomere erosion. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 15736–15741. [[CrossRef](#)]
70. Goodwin, T.J.D. Multiple LTR-Retrotransposon Families in the Asexual Yeast *Candida albicans*. *Genome Res.* **2000**, *10*, 174–191. [[CrossRef](#)]
71. Goodwin, T.J.D.; Ormandy, J.E.; Poulter, R.T.M. L1-like non-LTR retrotransposons in the yeast *Candida albicans*. *Curr. Genet.* **2001**, *39*, 83–91. [[CrossRef](#)]

72. Anderson, M.Z.; Wigen, L.J.; Burrack, L.S.; Berman, J. Real-time evolution of a subtelomeric gene family in *Candida albicans*. *Genetics* **2015**, *200*, 907–919. [[CrossRef](#)]
73. Wang, J.M.; Bennett, R.J.; Anderson, M.Z. The Genome of the Human Pathogen *Candida albicans* Is Shaped by Mutation and Cryptic Sexual Recombination. *MBio* **2018**, *9*, 1–16. [[CrossRef](#)] [[PubMed](#)]
74. Szilard, R.K.; Jacques, P.-É.; Laramée, L.; Cheng, B.; Galicia, S.; Bataille, A.R.; Yeung, M.; Mendez, M.; Bergeron, M.; Robert, F.; et al. Systematic identification of fragile sites via genome-wide location analysis of γ -H2AX. *Nat. Struct. Mol. Biol.* **2010**, *17*, 299–305. [[CrossRef](#)] [[PubMed](#)]
75. Price, R.J.; Weindling, E.; Berman, J.; Buscaino, A. Chromatin Profiling of the Repetitive and Nonrepetitive Genomes of the Human Fungal Pathogen *Candida albicans*. *MBio* **2019**, *10*, 475996. [[CrossRef](#)] [[PubMed](#)]
76. Chibana, H.; Iwaguchi, S.; Homma, M.; Chindamporn, A.; Nakagawa, Y.; Tanaka, K. Diversity of tandemly repetitive sequences due to short periodic repetitions in the chromosomes of *Candida albicans*. *J. Bacteriol.* **1994**, *176*, 3851–3858. [[CrossRef](#)]
77. Iwaguchi, S.I.; Homma, M.; Chibana, H.; Tanaka, K. Isolation and characterization of a repeated sequence (RPS1) of *Candida albicans*. *J. Gen. Microbiol.* **1992**, *138*, 1893–1900. [[CrossRef](#)]
78. Scherer, S.; Stevens, D.A. A *Candida albicans* dispersed, repeated gene family and its epidemiologic applications. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 1452–1456. [[CrossRef](#)]
79. Lephart, P.R.; Magee, P.T. Effect of the Major Repeat Sequence on Mitotic Recombination in *Candida albicans*. *Genetics* **2006**, *174*, 1737–1744. [[CrossRef](#)]
80. Iwaguchi, S.I.; Suzuki, M.; Sakai, N.; Nakagawa, Y.; Magee, P.T.; Suzuki, T. Chromosome translocation induced by the insertion of the URA blaster into the major repeat sequence (MRS) in *Candida albicans*. *Yeast* **2004**, *21*, 619–634. [[CrossRef](#)]
81. Lockhart, S.R.; Fritch, J.J.; Meier, A.S.; Schroppel, K.; Srikantha, T.; Galask, R.; Soll, D.R. Colonizing populations of *Candida albicans* are clonal in origin but undergo microevolution through C1 fragment reorganization as demonstrated by DNA fingerprinting and C1 sequencing. *J. Clin. Microbiol.* **1995**, *33*, 1501–1509.
82. Iwaguchi, S.I.; Sato, M.; Magee, B.B.; Magee, P.T.; Makimura, K.; Suzuki, T. Extensive chromosome translocation in a clinical isolate showing the distinctive carbohydrate assimilation profile from a candidiasis patient. *Yeast* **2001**, *18*, 1035–1046. [[CrossRef](#)]
83. Chibana, H.; Beckerman, J.L.; Magee, P.T. Fine-resolution physical mapping of genomic diversity in *Candida albicans*. *Genome Res.* **2000**, *10*, 1865–1877. [[CrossRef](#)] [[PubMed](#)]
84. Slutsky, B.; Staebell, M.; Anderson, J.; Risen, L.; Pfaller, M.; Soll, D.R. “White-opaque transition”: A second high-frequency switching system in *Candida albicans*. *J. Bacteriol.* **1987**, *169*, 189–197. [[CrossRef](#)] [[PubMed](#)]
85. Chu, W.S.; Magee, B.B.; Magee, P.T. Construction of an SfiI macrorestriction map of the *Candida albicans* genome. *J. Bacteriol.* **1993**, *175*, 6637–6651. [[CrossRef](#)] [[PubMed](#)]
86. Lephart, P.R.; Chibana, H.; Magee, P.T. Effect of the Major Repeat Sequence on Chromosome Loss in *Candida albicans*. *Eukaryot. Cell* **2005**, *4*, 733–741. [[CrossRef](#)]
87. Rustchenko, E.P.; Curran, T.M.; Sherman, F. Variations in the number of ribosomal DNA units in morphological mutants and normal strains of *Candida albicans* and in normal strains of *Saccharomyces cerevisiae*. *J. Bacteriol.* **1993**, *175*, 7189–7199. [[CrossRef](#)]
88. Wickes, B.; Staudinger, J.; Magee, B.B.; Kwon-Chung, K.J.; Magee, P.T.; Scherer, S. Physical and genetic mapping of *Candida albicans*: several genes previously assigned to chromosome 1 map to chromosome R, the rDNA-containing linkage group. *Infect. Immun.* **1991**, *59*, 2480–2484.
89. Ahmad, A.; Kabir, M.A.; Kravets, A.; Andaluz, E.; Larriba, G.; Rustchenko, E. Chromosome instability and unusual features of some widely used strains of *Candida albicans*. *Yeast* **2008**, *25*, 433–448. [[CrossRef](#)]
90. Huber, D.H.; Rustchenko, E. Large circular and linear rDNA plasmids in *Candida albicans*. *Yeast* **2001**, *18*, 261–272. [[CrossRef](#)]
91. Johzuka, K.; Horiuchi, T. The cis Element and Factors Required for Condensin Recruitment to Chromosomes. *Mol. Cell* **2009**, *34*, 26–35. [[CrossRef](#)]
92. Burrack, L.S.; Appen Clancey, S.E.; Chacón, J.M.; Gardner, M.K.; Berman, J. Monopolin recruits condensin to organize centromere DNA and repetitive DNA sequences. *Mol. Biol. Cell* **2013**, *24*, 2807–2819. [[CrossRef](#)]
93. Otoo, H.N.; Lee, K.G.; Qiu, W.; Lipke, P.N. *Candida albicans* Als Adhesins Have Conserved Amyloid-Forming Sequences. *Eukaryot. Cell* **2008**, *7*, 776–782. [[CrossRef](#)] [[PubMed](#)]

94. Hoyer, L.L.; Green, C.B.; Oh, S.H.; Zhao, X. Discovering the secrets of the *Candida albicans* agglutinin-like sequence (ALS) gene family - A sticky pursuit. *Med. Mycol.* **2008**, *46*, 1–15. [[CrossRef](#)] [[PubMed](#)]
95. Sheppard, D.C.; Yeaman, M.R.; Welch, W.H.; Phan, Q.T.; Fu, Y.; Ibrahim, A.S.; Filler, S.G.; Zhang, M.; Waring, A.J.; Edwards, J.E. Functional and Structural Diversity in the Als Protein Family of *Candida albicans*. *J. Biol. Chem.* **2004**, *279*, 30480–30489. [[CrossRef](#)] [[PubMed](#)]
96. Klotz, S.A.; Gaur, N.K.; De Armond, R.; Sheppard, D.; Khardori, N.; Edwards, J.E.; Lipke, P.N.; El-Azizi, M. *Candida albicans* Als proteins mediate aggregation with bacteria and yeasts. *Med. Mycol.* **2007**, *45*, 363–370. [[CrossRef](#)]
97. Zhang, N.; Harrex, A.L.; Holland, B.R.; Fenton, L.E.; Cannon, R.D.; Schmid, J. Sixty alleles of the ALS7 open reading frame in *Candida albicans*: ALS7 is a hypermutable contingency locus. *Genome Res.* **2003**, *13*, 2005–2017. [[CrossRef](#)]
98. Zhao, X.; Oh, S.; Coleman, D.A.; Hoyer, L.L. ALS51, a newly discovered gene in the *Candida albicans* ALS family, created by intergenic recombination: analysis of the gene and protein, and implications for evolution of microbial gene families. *FEMS Immunol. Med. Microbiol.* **2011**, *61*, 245–257. [[CrossRef](#)]
99. Sanyal, K.; Baum, M.; Carbon, J. Centromeric DNA sequences in the pathogenic yeast *Candida albicans* are all different and unique. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 11374–11379. [[CrossRef](#)]
100. Selmecki, A.; Gerami-Nejad, M.; Paulson, C.; Forche, A.; Berman, J. An isochromosome confers drug resistance in vivo by amplification of two genes, ERG11 and TAC1. *Mol. Microbiol.* **2008**, *68*, 624–641. [[CrossRef](#)]
101. Todd, R.T.; Wikoff, T.D.; Forche, A.; Selmecki, A. Genome plasticity in *Candida albicans* is driven by long repeat sequences. *Elife* **2019**, *8*, 153–165. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).